IDENTIFICATION OF ENZYMES POTENTIALLY INVOLVED IN THE FORMATION OF THE OOCYST WALL OF COCCIDIAN PARASITES

Thesis submitted by

Philippa Anne Sharman

BSc. Biomedical Science (Hons)

October 2013

for the degree of

Doctor of Philosophy

in the School of Public Health, Tropical Medicine

and Rehabilitation Sciences,

James Cook University

CERTIFICATE OF AUTHORSHIP

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that this thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Philippa Sharman

ACKNOWLEGEMENTS

First and foremost, I would like to thank my principal supervisor Professor Nick Smith. Thank you for all your guidance, support, encouragement and particularly your patience throughout this challenging project that travelled all over Australia and the world. Without you, this would not have been possible.

I would also like to thank my co-supervisors Dr Kate Miller for all your help, particularly in the later stages of my PhD where your support and guidance helped me finally get to the finish line. I would also like to thank Professor Alex Loukas for his support in my new lab in Cairns.

My PhD has spanned over many labs and universities, all of which have contributed immensely to my work and this project would not have been possible without them all. I would like to thank all the students and staff in each lab that made this project possible and bearable. I would like to especially thank Dr Marilyn Katrib, not only for your continued support, encouragement and friendship, but also my chicken house memories, those days would not be the same without you. I would like to thank my fellow researchers Dr Robert Walker, Dr Alana Wylie, Dr Amanda Hudson, Dr Michael Johnson, Michelle Robinson and Ivana Ferreira for their assistance with countless experiments and advice. Your support and friendship helped me get through this long journey. I would like to thank Fabien Brossier and his laboratory at the National Institute of Agronomic Research (INRA) in France for their guidance and assistance with transgenics and cell culture, and their patience with my broken French.

This project would not have been possible without the funding provided by the Australian Research Council. Furthermore, I would like to acknowledge my PhD scholarships, an Australian Postgraduate Award, as well as funding from the Institute for the Biotechnology of Infectious Diseases (IBID) and James Cook University. My research at INRA was funded by a travel grant from the Australian Society of Parasitology (ASP) as well as an ARC Linkage Grant. The funding for my national and international conferences kindly provided by ASP.

Finally, I would not be here without the loving support of my husband Leon Stepan. His patience, guidance and love was always there throughout this whole experience and never waivered. I would have fallen short of the line if it weren't for you. And to my family, thank you for all your love and support. Your excitement at seeing me in a floppy hat has kept me going to the end.

ABSTRACT

The Coccidia are a group of single-celled, protozoan parasites within the phylum Apicomplexa. Adaptable and widely distributed, they are able to reproduce asexually and sexually within their host, causing serious diseases in poultry, livestock, pets, wildlife and people all over the world. One of the defining features of the group is the oocyst, which is the end-point of sexual reproduction and is pivotal for transmission of these parasites. Oocysts are surrounded by a wall, which has a chemical structure that enables them to persist for long periods in the environment until ingested by new hosts. The resilience of the oocyst wall has been attributed to the formation of dityrosine cross-links between tyrosine-rich proteins that originate in the sexual stages (gametocytes) of the parasite. These tyrosine-rich proteins are processed into smaller proteins, then undergo oxidation to form the dityrosine bonds as they are incorporated into the oocyt wall matrix.

The hypotheses investigated in this study were that: processing of tyrosine-rich proteins is via proteases; and catalysis of dityrosine crosslinking is via oxioreductases. The aim of the thesis was to identify proteases and oxioreductases that are expressed specifically in gametocytes as they transform into oocysts. The poultry coccidian, *Eimeria tenella*, was chosen as the model for this study because of the availability of a genome database and reliable methods to purify gametocytes and oocysts of this parasite. Bioinformatic analyses identified 45 potential proteases in the *E. tenella* genome. However, expression of only three subtilisin-like serine proteases was up-regulated in gametocytes, as revealed by qualitative and quantitative RT-PCR. Expression of one, *Et*SUB2, was specific to the sexual phase of development and localised within the cytoplasm of macrogametocytes, adjacent to the wall forming bodies that store tyrosine-rich proteins prior to their incorporation into the oocyst wall. Data mining of the *E. tenella* database also identified five genes coding for oxioreductases. Quantitative RT-PCR revealed that the relative transcript abundance of three of these, amine oxidase 2, glutathione peroxidase and an oxioreductase, was increased dramatically in gametocytes. All three proteins could be detected specifically in macrogametocytes within infected caeca of chickens, displaying a punctate distribution in the cytoplasm of the parasite. Thus, this thesis identifies candidates for two putative enzymatic steps in the generation of chemical bonds that are crucial for the formation of oocyst walls and, therefore, the completion of the lifecycle of coccidian parasites. These enzymes are potential targets for novel parasite control strategies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

JOURNAL PUBLICATIONS

Katrib M, Ikin RJ, Brossier F, Robinson M, Slapetova I, Sharman PA, Walker RA, Belli SI, Tomley FM and Smith NC (2012) Stage-specific expression of protease genes in the apicomplexan parasite, *Eimeria tenella. BMC Genomics,* 12: 685

Sharman PA, Smith NC, Wallach MG, Katrib M (2010) Chasing the Golden Egg: vaccination against poultry coccidiosis. *Parasite Immunology,* 32 (8): 590-8. **This paper constitutes a large proportion (60%) of Chapter 1 of this thesis.**

Mai K, Sharman PA, Walker RA, Katrib M, De Souza D, McConville MJ, Wallach MG, Belli SI, Ferguson DJP and Smith NC (2009) Oocyst wall formation and composition in coccidian parasites.Mem. Inst. Oswaldo Cruz, 104 (2)

PRESENTATIONS

Sharman PA, Walker RA, Katrib M, Miller K, Smith NC (2012) Identification of oxioreductases in the coccidian parasite *Eimeria tenella*and their potential involvement in oocyst wall formation. Australian Society of Parasitology Annual Conference, Launceston, Australia

Sharman PA, Katrib, M, Ikin, RJ, Walker RA, and Smith NC (2010) Characterisation of Gametocyte-specific Subtilase-like Proteases of the Coccidian Parasite *Eimeria tenella*, 10th International Coccidiosis Conference, Guangzhou, China

Sharman PA, Katrib, M, Slapetova I, Walker RA, Belli SI, Wallach MG and Smith NC (2010) Subtilase-like proteases and oocyst wall formation of the coccidian parasite *Eimeria tenella*, 12th International Congress of Parasitology (ICOPA) Conference, Melbourne, Australia

Sharman PA, Katrib, M, Ikin RJ, Slapetova I, Walker RA, Belli SI, Wallach MG and Smith NC (2009) Investigation into the role of subtilase-like proteases in the formation of the oocyst wall of coccidian parasites, Australian Society of Parasitology Annual Conference, Sydney, Australia

Sharman PA, Katrib M, Ikin RJ, Slapetova I, Walker RA, Belli SI, Wallach MG and Smith NC (2008) The Identification of Subtilsin-like Serine Proteases in *Eimeria tenella*, Australian Society of Parasitology Annual Conference, Glenelg, Australia

ABBREVIATIONS

Abbreviation	Full name
ATP	adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyphosphate
BLAST	basic local alignment search tool
Bp	base pair
BSA	bovine serum albumin
°C	degrees
cDNA	complementary DNA
ddH_2O	double-distilled water
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphates
DTT	dithiolthreitol
E	efficiency
EDTA	ethylene diamine tetra acetic acid
ELISA	enzyme-linked immunosorbent assay
EtOH	ethanol
FA	formaldehyde agarose
FAD	flavin adenine dinucleotide
hrs	hours
g	relative centrifugal force (gravity)
G	gauge
IPTG	isopropyl- β-thiogalactopyranoside
Kb	kilobases
kDa	kilodaltons
m	metres
M	molar
min	minutes
mRNA	messenger RNA

NADH	beta-nicotinamide adenine dinucleotide
NADPH	beta-nicotinamide adenine dinucleotide phosphate reduced form
ng	nanograms
nm	nanometer
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pI	isoelectric point
p.i.	post infection
PMSF	phenyl methane sulphonyl fluoride
qRT-PCR	quantitative reverse transcriptase polymerase chain reaction
QS	quackenbush
RACE PCR	random amplification cDNA polymerase chain reaction
RNA	ribonucleic acid
RNAi	RNA interference
ROS	reactive oxygen species
Rpm	revolutions per minute
RT	room temperature
RT-PCR	reverse transcriptase polymerase chain reaction
secs	seconds
SDS	sodium dodecyl sulphate
SOD	superoxide dismutase
TBE	tris borate buffer with EDTA
T_m	melting temperature
Trx	Thioredoxin
UV	ultraviolet
V	volts

CHAPTER 1

INTRODUCTION

A large proportion (60%) of this chapter has been previously published in *Parasite Immunology*

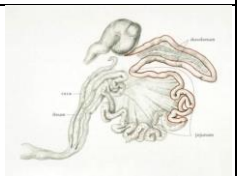
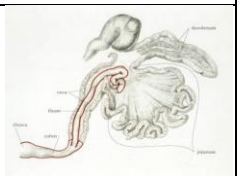
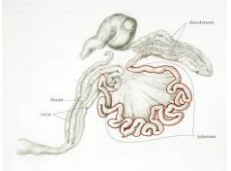
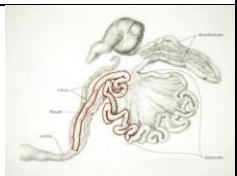
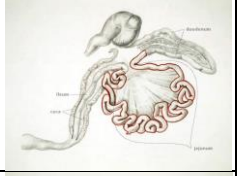
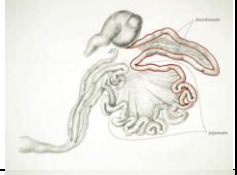
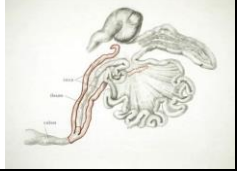
1.1 Introduction

Coccidiosis, still one of the most widely reported diseases within the poultry industry (Biggs, 1982, Xie *et al.*, 2001) is caused by one or more of at least seven species of the apicomplexan genus, *Eimeria* - *E. tenella, E. maxima, E. acervulina, E. brunetti, E. necatrix, E. praecox* and *E. mitis.* They characteristically infect different regions of the intestine (Table 1.1) causing the disease, coccidiosis, symptoms of which include listlessness, asocial behaviour, anorexia, weight loss and haemorrhagic diarrhoea. However, different species result in variant pathogenicity (Table 1.1). For example, whilst infection with *E. tenella* may cause considerable haemorrhagic diarrhoea and high mortality, infection with *E. praecox* results in a much milder disease (Long and Millard, 1968, Tyzzer, 1929). With about 40 billion chickens raised annually worldwide, the disease is estimated to cost upwards of GBP1.5 billion per annum (Williams, 1998), with the majority of the expense associated with poor flock performance (weight gain, food conversion rate) and administration of coccidiosis control systems such as prophylactics.

Coccidiosis occurs when the chicken host ingests environmentally-resistant oocysts, which are commonly found in the floor litter of a typical poultry house, or in the natural environment, such as is the case for free range poultry. Upon ingestion, a total of eight sporozoites are released from the four sporocysts contained within each oocyst (Figure 1.1). These rapidly attach to and invade the host intestinal epithelium, beginning the first of a limited number of asexual cycles that result in rapid amplification of merozoites. Eventually, the merozoites differentiate into sexual stages, the male microgametes fertilising the female macrogametes to produce oocysts that are shed in the faeces. For every oocyst ingested, several hundred thousand may be produced, which then contaminate the floor of poultry houses. Continual recycling through a flock leads to a high number of oocysts in the litter within 3-4 weeks (Williams, 1994). This situation is exacerbated by the high intensity rearing conditions within the industry (Shirley and Bedrnik, 1997). Good husbandry techniques have helped control the disease, however, the use of additional control measures, including anticoccidial drugs, are still essential.

Over the past 70 years, heavy reliance on chemoprophylaxis has led to the emergence of resistant parasite strains, rendering the use of anticoccidial drugs less effective (Chapman, 1997, Chapman and Hacker, 1994, Stephen *et al.*, 1997). Furthermore, with increasing health awareness, there is also a greater public concern regarding drug residues in poultry products, and growing pressure from government and consumer groups to ban such drugs from animal feeds (Williams, 2002). In Australia alone, the growth rate in

Table 1.1: *Eimeria* and different patterns of infection

Species	Site of Infection	Site of Infection	Pathogenicity
E. acervulina	Anterior small intestine		Watery, viscid and mucoid droppings. Severe growth depression, impaired feed efficiency and cessation in egg production (Williams *et al.*, 2009). Also associated with increased mortality.
E. brunetti	Mid and lower small intestine, rectum, caeca and cloaca		Severe weight loss, sloughing of the mucosa and hemorrhagic cases (Taylor *et al.*, 2007)
E. maxima	Mid-small Intestine		Lesions of pink exudate, thickened intestinal wall and haemorrhage (Schnitzler and Shirley, 1999). Haemorrhagic enteritis results in mild to severe diarrhoea and dehydration, decreased body weight (Long, 1929, McDonald *et al.*, 1986)
E. mitis	Lower small intestine, caeca and rectum		Reduced feed conversion, growth suppression and cessation in egg production (Fitz-Coy and Edgar, 1992, Williams *et al.*, 1999)
E. necatrix	Anterior and middle portion of the small intestine		Watery, mucoid and bloody diarrhoea leading to dehydration (Hein, 1971, Tyzzer, 1932)
E. praecox	Upper small intestine		Severe growth depression, impaired feed efficiency (Williams *et al.*, 2009)
E. tenella	Caeca only		Bloody and mucoid diarrhoea. Growth depression, severe bloody diarrhoea, high mortality (Allen and Fetterer, 2002, Jeurissen *et al.*, 1996)

* Images by Dr. Gheorghe Constantinescu for The Merck Veterinary Manual: Overview of Coccidiosis in Poultry

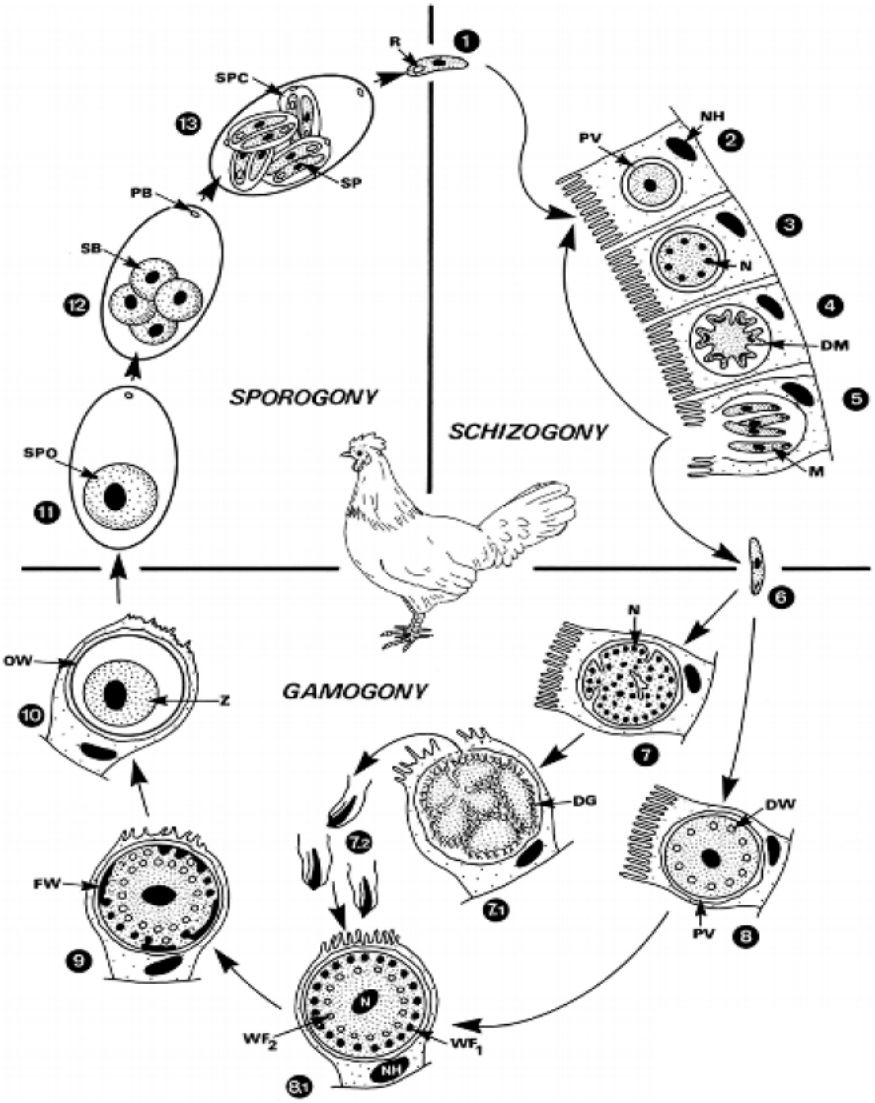


Figure 1.1: Lifecycle of *Eimeria.*

Chickens ingest sporulated oocysts in contaminated food and/or water. (1) Excystation of a sporulated oocyst in the chicken gizzard releases eight sporozoites; (2-5) sporozoites invade host epithelial cells of the gut and undergo asexual reproduction (schizogony), to produce merozoites; (6) merozoites are released from schizonts and re-infect host epithelium, or differentiate into micro – (7) or macrogametocytes (8) (gamogony); microgametes fertilise macrogametes and a zygote develops (9); the zygote develops into an unsporulated oocyst (10) which ruptures from host cells into the gut and is excreted in faeces (11), where environmental conditions cause sporulation (12-13). Modified from Melhorn (1988).

Abbreviations - DG, developing microgametes; DM, developing merozoites; DW, developing wall forming bodies; FW, fusion of WF1; M, merozoite, N, nucleus; NH, nucleus of host cell; OW, oocyst wall; PB, polar body (granule); PV, parasitophorous vacuole; R, refractile body; SB, sporoblast; SP, sporozoite; SPC, sporocyst; SPO, sporont; WF1, wall forming bodies I; WF2, wall forming bodies II; Z, zygote.

the demand for organic produce is expected to continue to increase by 10-30% per annum, including organic poultry free from antibiotics, chemotherapeutics and growth enhancers (Chang and Zepeda, 2005). Consequently, the use of vaccines has become more desirable.

1.2 Live Unattenuated Vaccines

Observations of *Eimeria* infections and subsequent immunity in several early studies indicated that development of an anticoccidial vaccine was feasible (Johnson, 1932, Tyzzer, 1929, Tyzzer *et al.*, 1932). It has since been established that any infection with *Eimeria* causes a strong, species-specific protective immunity that has also been found to be strain specific, at least with regard to *E. maxima* (Fitz-Coy, 1992, Martin *et al.*, 1997); therefore, any vaccine administered should include the common pathogenic species and strains that affect poultry. Immunity to *Eimeria* is stimulated by the initial developing parasite stages, particularly the schizonts, and subsequently boosted and maintained by multiple re-exposure to oocysts in the litter. Thus, the recycling of infection following administration of live oocysts is critical to the development of protective immunity (Chapman and Cherry, 1997).

The first commercial anticoccidial vaccine, CocciVac®, was introduced to the US market in 1952 and was a live vaccine comprising several wild-type strains of *E. tenella* oocysts. Criticism of the early vaccine was based on the observation that inclusion of only one species of *Eimeria* would not protect flocks from other species (Hinshaw, 1952). In the subsequent 50 years, the vaccine has been through a number of reformulations and variants of the original product - Coccivac®-B, Coccivac®-D and Immucox® - are still in use today and are registered in over 40 countries. However, the use of live unattenuated vaccines is limited somewhat by the pathogenicity of the parasites used. Thus, up until the late 1990's, vaccination with live vaccines was accompanied by chemotherapy to control pathology often induced by the live parasites (Martin *et al.*, 1997), though this is usually not required today due to improved means of administration of oocysts (Danforth, 1998, Danforth *et al.*, 1997). Hence, although virulent strains are still widely used, especially in North America, strains that have attenuated lifecycles are now, arguably, the preferred products.

1.3 Live Attenuated Vaccines

The effectiveness of attenuated vaccines also relies on administration of low doses of oocysts that are recycled through the litter, with protective immunity induced after two to three consecutive infections (Joyner and Norton, 1973, Long *et al.*, 1986). Attenuated vaccines are produced by selecting parasites that have fewer and smaller schizogonous cycles and enter the sexual phase faster than wild-type parasites but are equally

immunogenic (reviewed by McDonald and Shirley, 2009). Hence, recycling of oocysts with an attenuated vaccine results in a lower risk of disease occurring, as there is a reduction in proliferation of the parasites and less damage to the intestinal lining after passage through the gut (Williams, 1994, Allen and Fetterer, 2002).

Early attempts to attenuate *Eimeria* parasites included heat treatment (Jankievicz and Schofield, 1934) and x-irradiation (Albanese and Smetana, 1937), both of which were unsuccessful. The first successful attempt to develop attenuated parasites of *Eimeria* began when Long (Long, 1972) showed that *E. tenella* was able to complete its lifecycle in the chorio-allantoic membrane of the chicken embryo, and that serial passage in eggs resulted in significant attenuation of the parasite. The loss of pathogenicity of the parasites was attributed to a reduction in the size and invasiveness of the second generation schizonts (Long, 1974). Based on this, an embryo-adapted line of *E. tenella,* derived after more than 100 passages, is included in the commercially available Livacox® vaccine along with precocious lines of *E. acervulina, E. brunetti* and *E. maxima* (Shirley and Bedrnik, 1997, Williams, 2002). Although embryo-adapted, attenuated lines of *E. necatrix* have since been described (Kogut *et al.*, 1983, Shirley, 1980) there has been a failure to produce the equivalent in *E. acervulina, E. maxima and E. praecox* (Shirley and Bedrnik, 1997). This is thought to be mainly due to failure of the sporozoites to develop in the embryo, or any oocysts produced not sporulating properly (Shirley *et al.*, 1981). Therefore, a different means of attenuation was required for vaccine development. Today, the second of the two commonly used methods of attenuation of *Eimeria* species for inclusion in vaccination formulations, precociousness, is the most widely used method.

Precociousness is a selectable trait of the parasite and refers to a population of parasites that complete their lifecycle from sporozoite to oocyst up to 30 hours faster than parasites from the same parent strain (Shirley and Bedrnik, 1997, Vermeulen, 2005). The selection of such parasites was first described by Jeffers (Jeffers, 1975) who showed that serial passage of oocysts through a chicken and collection at earlier and earlier time points post-infection resulted in parasites of attenuated virulence. Importantly, infection of chickens with these parasites induced a high level of immunity against a challenge with the parent line (Johnson *et al.*, 1979). Precocious lines have been described for all seven species of *Eimeria* (Williams, 2002). Characteristically, precocious parasites of *Eimeria* have a marked reduction in oocyst reproduction and pathogenicity, and yet are still highly immunogenic. Studies also demonstrated the genetic stability of precocious lines, where precociousness was retained through serial passage without selection for early maturation of oocysts (McDonald and Shirley, 2009); thus, lines do not revert back to virulence. With this inherent improvement in

safety, and parasites being more predictable and reliable than embryo-adapted lines, precocious lines of *Eimeria* became the basis for the development of the first attenuated anticoccidial vaccine, Paracox®.

Paracox® was launched in 1989 to protect laying and breeding hens and it contained precocious lines of all seven species of *Eimeria,* including two lines of *E. maxima* due to antigenic variation seen in this species (Shirley, 1989, Shirley and Millard, 1986, Williams, 1994). Since its introduction, several other formulations and attenuated vaccines have become commercially available for use in different poultry flocks. Generally, *E. maxima, E. tenella* and *E. acervulina* are the only species included in vaccines for broiler birds as younger flocks rarely encounter the pathogenic species *E. brunetti* or *E. necatrix* (Chapman, 2000, Williams, 1998). In 2003, EIMERIAVAX 4M, was the first live coccidiosis vaccine registered for use in Australian poultry. It is comprised of drug-sensitive, precocious lines of *E. acervulina, E. maxima, E. tenella and E. necatrix,* each isolated from backyard flocks of Australian chickens (Jorgensen *et al.*, 2006a, Jorgensen *et al.*, 2006b). Field trials showed that the vaccine could protect broiler breeders, broilers, free range and barn flocks of egg-laying hens by eye-drop or coarse aerosol application (Richards and Woods, 2001). Efforts continue to be directed towards the derivation of further vaccines based on precociousness and it is probably fair to say that reliance on these type of vaccines will, if anything, increase in years to come (McDonald and Shirley, 2009).

1.4 Subunit Vaccine Development

An anticoccidial vaccine composed of protective antigens, either native or recombinant, has been pursued as an alternative to live vaccines and the problems and costs associated with them. The identification of protective antigens is vital in development of any future vaccine and many of the *Eimeria* antigens investigated thus far are primarily associated with asexual stages of the parasite such as sporozoites and merozoites. In particular, proteins associated with invasion and the apical complex, characteristic of this phylum of parasites, have been investigated as potential subunit vaccine components. These include antigens associated with micronemes (Tomley *et al.*, 1996, Tomley *et al.*, 1991), rhoptries (Tomley, 1994) and refractile bodies (Vermeulen *et al.*, 1993). Ultimately, these studies have revealed that use of these asexual stage antigens to immunise chickens only provides a moderate and, often, inconsistent protection against challenge with *Eimeria* infections (Jenkins, 2001, Vermeulen, 1998). Studies have also highlighted that there is distinct antigenic variability between the endogenous developmental stages of the parasite, and that antigenic modification during successive asexual generations may aid the parasite in evading immune responses (McDonald *et al.*, 1988, Tomley, 1994). The various antigens and strategies used in attempts

to develop subunit vaccines against the asexual stages of *Eimeria* have been reviewed thoroughly (Jenkins, 1998, McDonald and Shirley, 2009, Shirley *et al.*, 2007, Shirley *et al.*, 2005, Vermeulen, 1998) and, so, will not be reiterated here.

Work has also been conducted to investigate antigens of *Eimeria* sexual stages as vaccine candidates with the aim of developing a transmission blocking vaccine. The goal of transmission-blocking vaccines is to reduce oocyst output, resulting in a low level of exposure to allow natural immunity to asexual stages to also develop. The outcome of this research, described in more detail below, has led to the successful development and marketing of the first subunit vaccine against any protozoan parasite as an alternative means to control coccidiosis – CoxAbic®.

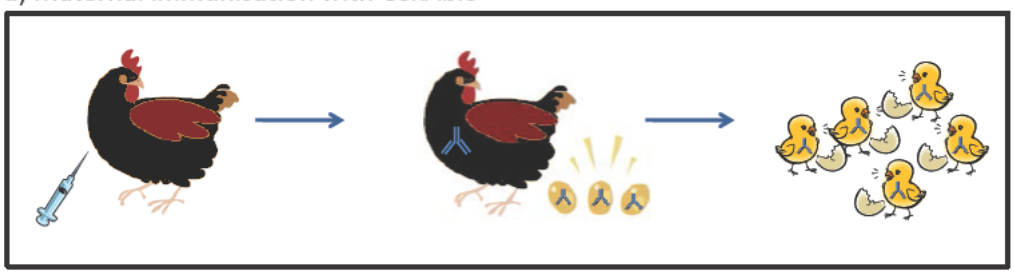
1.4.1 A Novel Sexual Stage Subunit Vaccine

The subunit vaccine, CoxAbic®, is comprised of affinity purified gametocyte antigens (APGA) from *E. maxima* in a proprietary oil in water adjuvant. The vaccine is cost effective on a commercial scale through a novel strategy of maternal immunisation, where vaccination of laying hens can lead to protection of broiler offspring (Figure 1.2a). More specifically, injection of gametocyte antigens into the breast muscle of breeder hens stimulates the production of large amounts of specific IgG (also referred to as IgY) maternal antibodies that are transferred to their offspring, via the egg yolk, to provide protective immunity (Smith *et al.*, 1994b, Wallach *et al.*, 1992, Wallach *et al.*, 1995a). Immunisation occurs prior to hatching, thus eliminating stress imposed by vaccination of the hatchlings, which are protected against coccidiosis from one day of age. The vaccine functions as a transmission blocker by inhibiting development of macrogametes into oocysts (Figure 1.2b), thereby reducing levels of oocysts shed in the litter. Thus, broiler chicks, once exposed to the parasite in the field, are able to develop active immunity against re-infection without suffering the economically damaging effects of the disease.

1.4.2 Identification of the antigenic potential of gametocytes

Initial studies in the development of CoxAbic® aimed to identify major antigenic proteins of *Eimeria* gametocytes. It was postulated that by immunologically inhibiting the growth, development and/or fertilisation of sexual stages of *Eimeria,* the lifecycle would be interrupted and, thus, reduce parasite transmission. Of the seven species that affect the poultry industry*, E. maxima* is considered the most immunogenic (Rose, 1974). Early studies,

a) Maternal immunisation with CoxAbic®



b) Molecular basis of CoxAbic®

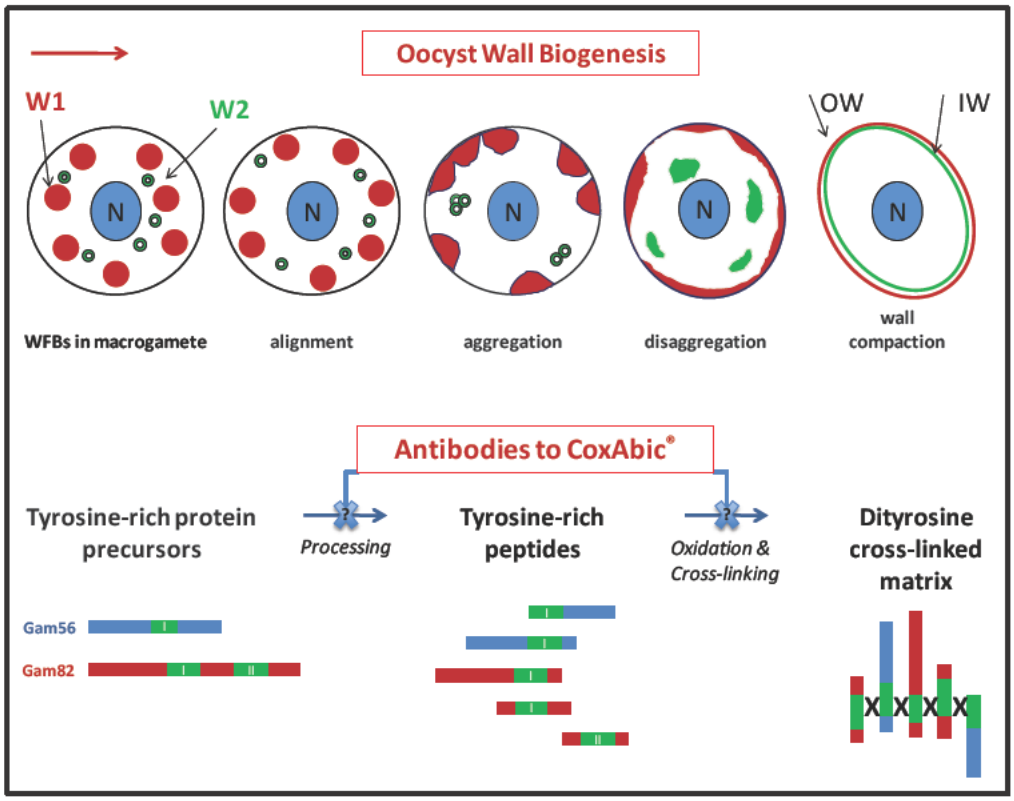


Figure 1.2: A schematic representation of the CoxAbic® vaccine strategy

The vaccine is administered by immunisation of broiler breeder hens. Maternal IgG (or IgY) antibodies develop and are then transferred to broiler offspring via the eggyolk, conferring protection against coccidiosis (a). The proposed mechanism of action of CoxAbic® is by inhibition of oocyst wall formation by maternal antibodies, resulting in reduced oocyst shedding. CoxAbic® is composed of tyrosine-rich precursor proteins, GAM56 and GAM82, which are located in the wall forming bodies of the macrogametoctye. Upon fertilisation, the wall forming bodies of type I (red) and type II (green) begin to form the outer (OW) and inner layers (IW) of the oocyst wall, respectively. In the process, GAM56 and GAM82 are released, processed, and cross-linked via their tyrosine residues to form the hardened, dehydrated and environmentally-resistant oocyst wall. Antibodies induced by vaccination with CoxAbic® could potentially interfere with oocyst wall formation either by "protecting" GAM56 and GAM82 from processing into smaller tyrosine-rich peptides, or by inhibiting their cross-linking (b).

therefore, concentrated on the *E. maxima* model and much work was focused on better understanding the basis of immunity and the lifecycle stages that are predominantly involved in immune responses (Rose, 1971, Rose, 1974, Rose and Hesketh, 1976). In addition, *E. maxima* was selected as the model species for initial work on development of a subunit vaccine as its gametocytes are very large in size and relatively easily visualised and purified (Wallach, 1997).

The induction of immunity using the *E. maxima* model was first demonstrated by Rose (1974), who showed that a single low dose of *E. maxima* oocysts could protect chickens against a challenge with high doses of oocysts of the homologous strain, and that one cycle of infection was enough to stimulate this protective immunity. It was further demonstrated that sera taken 14 days post infection with *E. maxima* can give passive protection to naive chickens against a challenge (Rose, 1971, Rose and Hesketh, 1976). However, it was first thought that the asexual stages of the lifecycle of *Eimeria* were predominantly responsible for the immune response invoked in chickens. Analysis of convalescent sera taken 14-20 days after an *E. maxima* infection, which was shown to passively protect naive birds, demonstrated that the early stages of development had a strong immunogenic effect, and the later sexual stages were poorly immunogenic in *E. maxima, E. necatrix and E. tenella* (Rose, 1967, Rose and Hesketh, 1976). Kouwenhoven and Kuil (1976) also reported that sera taken from chickens infected with *E. tenella* showed no reaction with sexual stages or first generation schizonts, indicating that gametocytes had poor immunising capabilities in chickens. Later studies, however, contradicted these earlier findings (Pugatsch *et al.*, 1989, Wallach *et al.*, 1989).

The antigenicity of sexual stages of *Eimeria* was first demonstrated when a monoclonal antibody to an *E. tenella* gametocyte antigen was shown to inhibit fertilisation *in vitro* (Laxer *et al.*, 1987). In 1989, Pugatsch *et al*. developed a method to isolate and purify gametocytes by enzymatic digestion of the infected mucosa with hyaluronidase, followed by size separation. They showed that whole gametocytes were highly antigenic both in the course of an infection and when injected into rabbits/mice. That same year, convalescent sera from *E. maxima* immune chickens was found to recognise two immunodominant macrogametocyte antigens of 56 kDa and 82 kDa in size (Wallach *et al.*, 1989). When these two proteins were administered to a variety of hosts in the form of a crude extract, they were found to be highly immunogenic (Wallach *et al.*, 1989). Following this discovery, it was hypothesised that these macrogametocyte antigens may play a role in conferring protective immunity to the host (Wallach *et al.*, 1989).

1.4.3 Passive Immunity to Affinity-Purified Gametocyte Antigens (APGA)

Earlier work had demonstrated, as aforementioned, that convalescent sera collected from chickens had the capacity to provide up to 97% inhibition in oocyst output in passively immunised and challenged birds (Rose, 1971). Work by Wallach *et al.* (Wallach *et al.*, 1990) investigated antibodies to the previously identified immunodominant gametocyte antigens and their potential to transfer immunity passively.

Sera from mice immunised with enriched gametocyte extracts were found to contain antibodies to the predominant 56kDa and 82kDa macrogametocyte proteins. A monoclonal antibody, 1E11-11, which recognised the 56kDa antigen, was bound to a Sepharose column and used to purify the 56kDa macrogametocyte protein. Surprisingly, the 82kDa macrogametocyte protein co-eluted, sometimes with a third 230-250 kDa gametocyte protein (Wallach *et al.*, 1990). Thus, affinity purification could successfully extract the macrogametocyte antigens. These affinity purified macrogametocyte antigens were then used to produce highly specific chicken anti-gametocyte sera, which were pooled and used in passive immunisation studies. Naïve, two week old chicks were immunised passively with sera containing the anti-56kDa and anti-82kDa protein IgG antibodies, resulting in a reduction in oocyst output by 40-50% in chickens. Based on this result, it was determined that these antibodies provided partial protective immunity against *E. maxima* (Wallach *et al.*, 1990). Although the exact mechanism of inhibition remained unknown, it was obvious that the antibodies were affecting parasite development. Studies showed that mouse antibody raised to the 56kDa and 82kDa antigens bound predominantly to macrogametocytes (Pugatsch *et al.*, 1989). As such, it was hypothesised that these antibodies were either inhibiting the growth, development or fertilisation of the macrogametes and, thus, inhibiting oocyst formation (Figure 1.2b), reducing the total number of oocysts produced (Wallach *et al.*, 1990). As work progressed, the ability of the macrogametocyte antigens to induce protective immunity was investigated.

1.4.4 Maternal immunisation studies

Previously, maternal transfer of IgG antibodies via the egg yolk had been shown to effectively prevent infection with *Eimeria* in chickens (Rose, 1971, Rose, 1972). This mechanism of maternal antibody transfer was investigated as a means of immunising hens with *E. maxima* APGA (Wallach *et al.*, 1990, Wallach *et al.*, 1989). Work showed that APGA, when used as a vaccine to immunise laying hens, could provide a good level of immunity to hatched chicks through passive transfer of protective maternal anti-gametocyte antibodies (Figure 1.2a). This level of immunity resulted in up to an 83% reduction in oocyst shedding,

when chicks were challenged with *E. maxima* oocysts, which was similar to that observed in chicks from hens vaccinated with a live infection (Wallach *et al.*, 1992).

These results led to further maternal immunisation studies (Smith *et al.*, 1994a, Smith *et al.*, 1994b, Smith *et al.*, 1994c, Wallach *et al.*, 1995b). Maternal transfer of protective antibodies to chicks from hens given a high dose of *E. maxima* oocysts was also observed, where passive immunity in the chicks correlated to the amount of IgG transferred via the egg yolk, and was detected in the sera of chicks for up to 3 weeks post-hatching (Smith *et al.*, 1994b). It also appeared that maternal immunisation with *E. maxima* APGA could induce cross-protection to heterologous species, *E. tenella* and *E. acervulina,* as well as *E. maxima* infections (Smith *et al.*, 1994c, Wallach *et al.*, 1995b). At this point, it was believed that this cross-reactivity was most probably due to conserved epitopes of the major gametocyte antigens in the different species and, thus, led to the hypothesis that a vaccine of *E. maxima* gametocyte antigens could possibly be used to control coccidiosis caused by, at least, the three predominant *Eimeria* species.

In floor pen maternal immunisation trials, the resistance of chicks from APGA-immunised breeder hens was compared with that of hatchlings from control parent flocks by introducing oocysts to the pens via infection of 'seeder' birds infected with 50 oocysts of *E. maxima, E. tenella* and *E. acervulina.* Analogous to the laboratory trials, a reduction of 60-80% in oocyst shedding into the litter was observed; comparable to the reduction observed using coccidiostats (Wallach, 1997). These trials were repeated three times under the same conditions, showing that there was an average reduction of 60-70% oocyst output in vaccinated groups up to 4-6 weeks in age (Wallach, 1997, Wallach, 2003). These results were encouraging, as they supported the idea that the highly conserved *E. maxima* antigens could provide good levels of protective immunity against at least three major species that cause coccidiosis in broilers. Despite this, questions still remained about whether this vaccine could provide maternal protection against all *Eimeria* species (and strains) encountered in the field and if maternal immunity was applicable in controlling coccidiosis within industry management schemes and a variety of climatic conditions.

1.4.5 Field Trials with Affinity-Purified Gametocyte Antigens

To further assess the efficacy of maternal immunisation with APGA, testing was undertaken in a multi-centred, multinational field trial involving five countries from four different continents: Israel; Brazil; Argentina; South Africa; and Thailand (Michael, 2003, Wallach *et al.*, 2008). The safety and immunogenicity of the vaccine in breeding hens was assessed on a large scale, with birds vaccinated twice prior to the start of their laying period (15 and 20

weeks respectively). Immunisations were found to have no deleterious effect on the hens (Wallach *et al.*, 2008): no adverse reactions; no damage at the site of injection; and no effect on hen mortality or on the number of eggs produced (*eg*, in the Argentine trials, 119 eggs were produced per immunised hen versus 116 per control hen). In all four countries, IgG antibody titers remained at a presumptive protective level throughout the life of the laying hens. The maintenance of highly specific IgG antibody levels in vaccinated flocks is thought to be due to the boosting effect that is naturally acquired from exposure to infection with oocysts in the environment (Wallach *et al.*, 2008).It is even conceivable that maternal antibodies may increase due to this natural exposure. However, in the absence of immunisation, these titres are variable and, therefore, do not necessarily provide protective levels of maternal immunity (Wallach *et al.*, 2008). Thus, maternal immunisation with APGA, which lasts for the life of the laying hens, is important.

Extensive field trials also assessed the protection provided to chicks from vaccinated breeder hens. Hatchlings were challenged with *E. tenella* oocysts to assess oocyst output; it was found that there was a significant reduction of 67.9%, similar to results found in laboratory and pen trials performed earlier (Wallach, 1997, Wallach *et al.*, 2008). An important outcome of these studies was the active immunity seen in maternally immunised birds up to 8 weeks old. Broiler chickens are bred to live for 5-7 weeks, before being slaughtered for poultry meat production; therefore, maternal immunisation with gametocyte antigens has the capacity to protect broiler flocks for the entirety of their lifetime. It has also been observed that resistance to infection from vaccinated progeny can outlast the life of maternal antibodies (Wallach *et al.*, 2008). This is because maternal immunity does not interfere with exposure to asexual development within vaccinated birds. Thus, passively transferred protective antibodies reduce, rather than completely stop, transmission of oocysts between birds, thereby allowing birds to develop their own active anti-asexual stage immunity in addition to the already induced maternal immunity. Immunity based on the asexual stages of *Eimeria* has previously been demonstrated to be strong and effective (McDonald *et al.*, 1986, Rose, 1987, Shirley *et al.*, 1995). Hence, the protective immunity of CoxAbic® is two-fold – on one hand, reducing exposure of hatchlings to oocysts, yet, at the same time, allowing them to acquire natural immunity by exposure to asexual stages, thus, providing effective and long-lasting control of coccidiosis.

The same study by Wallach *et al.* (2008) also revealed that hatchlings from vaccinated hens performed at least as well as positive control groups treated with anticoccidial drugs or live vaccines. In the poultry industry, the main performance parameter of any coccidiosis vaccine is its effect on weight-gain, especially in regard to broiler flocks. As poultry farmers would

not leave any of their flock unprotected, the performance of maternal immunisation was assessed in comparison to a 'gold standard', either anticoccidial drug administered in feed or a live vaccine. At least 1 million CoxAbic® vaccinated breeder hens and positive control chickens were assessed, resulting in a total of over 60 million progeny from immunised hens and 112 million positive control progeny (Wallach *et al.*, 2008). To assess the economic feasibility of the vaccine, lesion scores were graded and overall performance assessed including parameters such as mortality, daily weight gain (DWG) and food conversion ratio (FCR). When compared with flocks vaccinated with a live coccidiosis vaccine, in field trials in Argentina, no significant difference was observed. In Brazil, broiler flocks were vaccinated with gametocyte antigens and performance measured against broiler flocks treated with an ionophore anticoccidial in their feed. In this case, the vaccinated birds showed, if anything, an advantage over the drug treated group of 9.6% total income. Therefore, it was concluded that the use of CoxAbic® as a method of vaccination offers at least the same level of protection and economic advantage as those commonly accepted and used in the poultry market. Further evidence of the effectiveness of the maternal immunisation approach in the field was obtained in Thailand and South Africa. In a challenge trial in Thailand, three groups of vaccinated birds – CoxAbic®, a commercial live vaccine and salinomycin-treated – were challenged with 60,000 virulent *E. tenella* oocysts orally. Lesion scores between the three flock groups revealed that the CoxAbic® vaccinated groups had the lowest lesion score (<0.5) at 24, 30 and 35 days of age. By contrast, live vaccine treated flocks had a lesion score >2 during the same period, whilst salinomycin-treated flocks peaked at 30 days of age with a score >2.5, but recovered to ~1.0 at day 35 (Wallach *et al.*, 2008), again confirming the effectiveness of vaccination with CoxAbic®.

These results demonstrated that maternal immunisation with gametocyte antigens provides the potential for controlling coccidiosis under different rearing conditions in various climates and environmental surroundings. The basis of control, rather than eradication, means that both sexual and asexual stage protective immunity develops in the birds. Importantly, several recent studies demonstrated the conserved and functional importance of the two gametocyte antigens, GAM56 and GAM82, and explained why their inclusion in the vaccine formula confers protection against a range of *Eimeria* species (Belli *et al.*, 2009).

1.4.6 Characterisation of GAM56 and GAM82

Concurrent to the development of CoxAbic®, studies were conducted to characterise the GAM56 and GAM82 antigens that are the main components of the vaccine. Initial studies showed that GAM56 and GAM82 are glycoproteins (Belli *et al.*, 2002b) and further immunofluorescence studies localised these antigens to the wall-forming bodies of the

macrogametocyte and in the oocyst wall (Ferguson *et al.*, 2003a). These two antigens were identified as key players in the formation of the oocyst wall (Belli *et al.*, 2003d, Belli *et al.*, 2002c, Wallach *et al.*, 1992, Wallach *et al.*, 1995b). The oocyst wall, which facilitates the transmission of *Eimeria* by protecting the parasite when it is in the outside world, originates from the fusion of specialised organelles – wall forming bodies (WFB's) – found in the macrogametocytes of *Eimeria* (Ferguson *et al.*, 2003a). During maturation of the macrogametocyte, the WFB's align beneath the cell surface before degranulating and releasing GAM56 and GAM82 (Figure 1.2b). The proteins, and/or truncated versions thereof, are believed to then cross-link via dityrosine bonds to form the resilient wall structure (Belli *et al.*, 2003b). The inclusion of these proteins in CoxAbic® means that the stimulated antibodies probably interfere with the formation of cross-links between the proteins (Figure 1.2b), and therefore, prevent effective transmission by interrupting oocyst wall formation (Belli *et al.*, 2006b, Wallach *et al.*, 2008). As these antigens are involved in a process as fundamental as oocyst wall formation, it is not surprising that homologues of these proteins occur in other species of *Eimeria.* An important study by Belli *et al.* (Belli *et al.*, 2009) demonstrated that antibodies to *E. maxima* GAM56 and GAM82 recognised proteins in the WFB's of macrogametocytes and oocysts of *E. tenella* and *E. acervulina*. Homologous genes encoding for GAM56 and GAM82 were also identified in these species and, when the three sequences were aligned, they were found to be highly homologous around the epitopes of these proteins, hence explaining the cross-species protection afforded by the vaccine. It is not yet understood, however, how the protective IgG antibodies gain access to what is essentially an intracellular parasite.

1.4.7 Future development of second generation maternally-delivered vaccines

Despite the obvious success of CoxAbic® as the first and, currently, only subunit vaccine for the control of coccidiosis within the poultry industry now registered in many countries worldwide, a drawback is the expense associated with production. This is because production of the vaccine relies on affinity purification of native gametocyte antigens from parasites. As it is not possible to reliably culture sexual stages of *Eimeria* in an *in vitro* culture system, parasites are passaged and isolated from the intestines of chickens raised under strict specific pathogen free (SPF) conditions, which is not only expensive, but also time consuming and laborious (Belli *et al.*, 2004b). Therefore, recent work has been aimed at determining whether recombinant forms of the gametocyte proteins in APGA could maintain antigenic and immunogenic properties analogous to the native antigens and, therefore, perhaps replace them. A study by Belli *et al.* (Belli *et al.*, 2004b) examined bacterially-expressed recombinants of GAM56 and GAM82 to investigate if they could maintain the

antigenic determinants recognised by protective antibodies to their native protein counterparts. Antibodies to the native proteins appeared to recognise the same epitopes of the recombinant GAM56 and GAM82, suggesting the epitope sites had been retained (Belli *et al.*, 2004b). Moreover, immunisation of chickens with these recombinant proteins induced a strong antibody response, and sera from these birds recognised the native proteins (Belli *et al.*, 2004b), further indicating that these recombinant proteins mimicked the antibody response elicited by immunisation with the native antigens.

1.5 A Common Theme in Control of Coccidiosis – Transmission Blocking

Many of the symptoms of coccidiosis are directly related to the number of parasites invading, replicating within, and egressing from host cells and damaging the epithelium and sub-mucosa of the chicken intestine as a result (Reid, 1990). Even other effects, such as the shortening in height of the intestinal villi and destruction of epithelia cell function leading to malabsorption (Pout, 1967), which are at least in part a result of the host's reaction to the parasites, are correlated with parasite burden (Lillehoj, 2005, Reid, 1990). Not surprisingly, therefore, the successful management of coccidiosis – whether through prophylactic chemotherapy, vaccination with virulent or attenuated parasites or maternal immunisation with CoxAbic® – ultimately depends on reducing transmission of parasites within flocks of birds. Understanding transmission, and its molecular basis, is likely to be the key to the development of future strategies and products to control coccidiosis. Crucial to the successful transmission of coccidian parasites is the oocyst wall. Renowned for its robustness, it protects the parasite from desiccation and mechanical damage, facilitates its survival in the environment and, thus, its transmission.

1.6 The Coccidian Oocyst

The hard-walled oocyst is a defining feature of the Coccidia, a group that includes *Isospora* and *Cyclospora,* species similar to *Eimeria,* that have a monoxenous life cycle, where asexual and sexual development occurs within a single host and transmission is exclusively via the faecal-oral route (Belli *et al.*, 2006b). The Coccidia also include species of *Toxoplasma, Neospora* and *Sarcocystis,* parasites with a heteroxenous life cycle. These parasites can replicate within an intermediate host, though their sexual development is limited to the definitive host. These parasites can form cysts in tissue, and as such can be transmitted through the ingestion of tissue cysts as well as oocysts within a contaminated environment.

Similar to that which was described for *Eimeria,* the oocysts of the majority of coccidian parasites are shed, unsporulated and non-infectious, in faeces into the environment. It is only with moisture, warmth and oxygenation that sporulation occurs by meiotic division

(Canning and Anwar, 1968, del Cacho *et al.*, 2005). The oocyst is able to remain for extended lengths of time in even harsh external environments and remain infective (Korich *et al.*, 1990, Kuticic and Wikerhauser, 1996, Langkjaer and Roepstorff, 2008, Williams, 1995).

1.6.1 The oocyst wall of the coccidia

The basic structure of the oocyst is consistent across the Coccidia (Ferguson *et al.*, 1977, Ferguson *et al.*, 1975, Scholtyseck *et al.*, 1971, Speer and Duszynki, 1975). Its development and formation have been described at an ultrastructural level in a number of Coccidia (Ferguson *et al.*, 1977, Ferguson *et al.*, 1975, Lindsay *et al.*, 1999a, Pittilo and Ball, 1980, Scholtyseck *et al.*, 1971, Speer and Duszynki, 1975), and its morphology, in particular the oocyst wall has been studied by electron microscopy especially in *T. gondii* and *Eimeria.* In contrast, only a limited number of papers have been published on other Coccidia, such as *Besnoitia, Hammondia*, *Neospora* and *Sarcocystis* (Lindsay *et al.*, 2004, Lindsay *et al.*, 1999b). Despite difficulties studying the oocyst wall, studies undertaken thus far have shown general similarities amongst the Coccidia.

In *Eimeria,* early oocysts are enclosed by a loose membrane – the outer veil (Ferguson *et al.*, 2003a, Ferguson *et al.*, 2000) – followed by the oocyst wall itself, which is comprised of two layers. The outer veil is not found in mature oocysts purified from faeces. It is most likely shed during excretion and, therefore, plays no protective role for the parasite in the external environment. The oocyst wall of *Toxoplasma gondii* has been reported to have four or five layers, however, it has since been shown that the two or three outermost layers correspond to the loose outer veil of *Eimeria* that is not present in excreted oocysts (Ferguson *et al.*, 1977, Ferguson *et al.*, 1975, Speer *et al.*, 1998). Multiple membranes similar to the outer veil of *Eimeria* that surround the oocyst wall have also been reported in other coccidian parasites (Ferguson *et al.*, 1977, Ferguson *et al.*, 1975, Scholtyseck *et al.*, 1971, Speer and Duszynki, 1975). Characterization, at the molecular level, of the oocyst wall of Coccidia has to this time focused on a limited number of reports from *Eimeria* species (Belli *et al.*, 2003b, Belli *et al.*, 2003d, Eschenbacher *et al.*, 1996, Fried *et al.*, 1992, Monné and G, 1954, Stotish *et al.*, 1978a), due largely to limitations in isolating large numbers of oocyst wall samples, and the inherent nature of the wall (Belli *et al.*, 2006b) though recent studies have also shed some light on *T. gondii* (Fritz *et al.*, 2012).

1.6.2 Morphology and composition of the oocyst wall of *Eimeria*

The structure, composition and formation of the oocyst wall of *Eimeria* has been reviewed recently (Belli *et al.*, 2006b, Chapman *et al.*, 2013). It is a bilayered structure composed of an electron-dense outer layer, ~200μm in width, with a roughened outer surface and a uniform

electron-lucent inner layer, ~40μm in width. The oocyst remains viable after treatment with 10% sodium hypochlorite, storage in the strong oxidising agent, potassium dichromate (1-2%) or sulphuric acid (2%) (Belli *et al.*, 2006a). The wall is also resistant to enzyme proteolysis and impermeable to water-soluble substances, including many detergents and disinfectants (Monné and G, 1954, Ryley, 1969).

Mai *et al* (2009) showed by gas chromatography and mass spectrometry that the *Eimeria* oocyst wall is composed mainly of protein (>90%) with small amounts of lipid (1.4-7.6%) and carbohydrate (0.3-2.0%). There is little difference between the unsporulated and sporulated oocyst walls of *E. tenella* and *E. maxima* (Mai *et al.*, 2008). Only a small number of oocyst wall proteins have been identified, *eg*: a single 10kDa protein band detected with SDS-PAGE gels of *E. tenella* oocyst walls (Stotish *et al.*, 1978b); a single 12kDa protein band of the oocyst wall of *E. tenella* (Karim *et al.*, 1996); and a 14kDa oocyst wall protein in *E. tenella* and *E. acervulina* by Eschenbacher *et al.* (1996). Additionally, Moufafo (2002) reported that a monoclonal antibody raised against the macrogametocytes of *E. tenella,* reacted with three proteins of 23, 25 and 30kDa in the oocyst wall and Krücken *et al.* (2008) (Krucken *et al.*, 2008) described a 22kDa protein that localises to the inner oocyst wall of *E. tenella.*

It is probable that the 10, 12 and 14kDa proteins described by Stotish *et al.,* (1978), Mouafa *et al.,* (2002) and Eschenbacher *et al.,* (1994), are essentially the same as the 8, 10 and 12kDa proteins uncovered by Belli *et al.* (2003a) in *E. maxima* using methods that enabled the separation of an apparent single band on an SDS-PAGE gel into three bands. Belli *et al* (2003a) also described a 31kDa protein in the oocyst wall. All these proteins are unusually rich in the amino acid tyrosine. Furthermore, Belli *et al.* (2003a) demonstrated, by N-terminal sequencing, that the 8, 10, 12 and 31kDa oocyst wall proteins all originate from either GAM56 or GAM82 (Figure 1.3). Thus, it was found that the N-terminus sequence of wp12 was identical to the tyrosine rich internal domain I of GAM56 and that the N-terminus of wp31 was identical to the N-terminus of GAM56 (Figure 1.3). Similarly, the N-terminus of both wp10 and wp8 were identical to two tyrosine-rich domains, I and II respectively, in GAM82 (Belli *et al.*, 2002a, Belli *et al.*, 2003c, Belli *et al.*, 2006a).

Using antibodies to GAM56 and GAM82, as well as antibodies to APGA, Ferguson *et al.* (2003) demonstrated that the outer and inner layers of the wall are derived from specialised organelles, wall forming bodies I (WFB1) and II (WFB2), found exclusively in the macrogametocytes (Figure 1.4). There is some controversy, and a distinct lack of hard data, about what initiates wall formation (Ferguson *et al.*, 2003b, Ferguson, 2002) but perhaps the

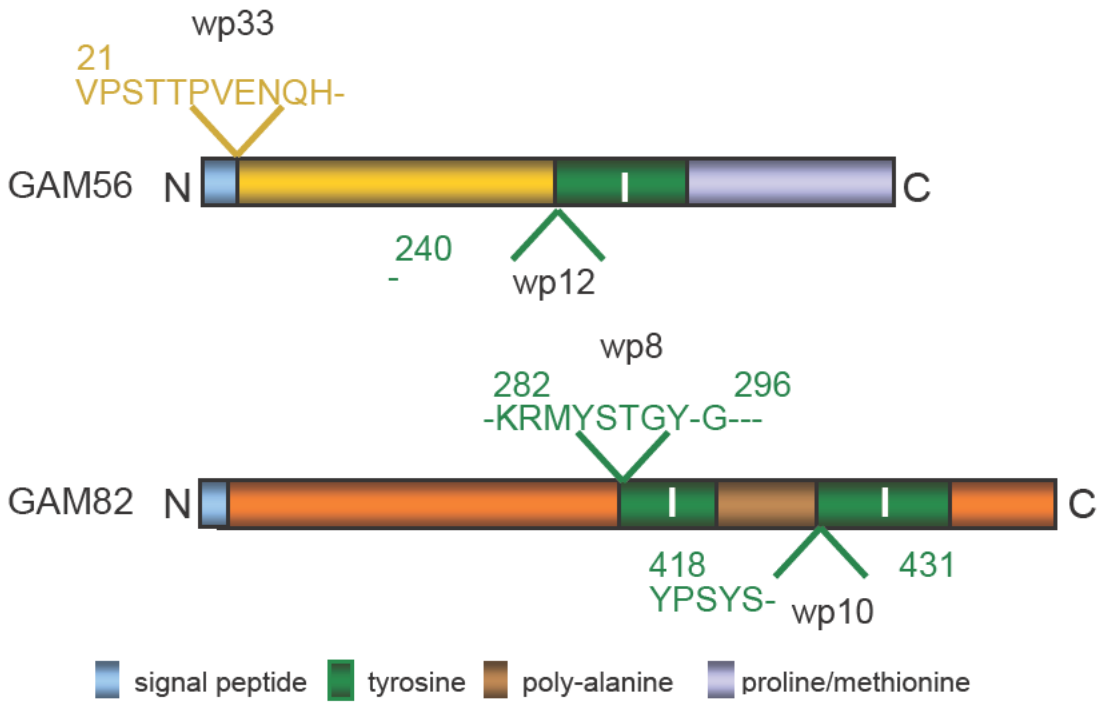


Figure 1.3: Oocyst wall protein sequences found in proteins GAM56 and GAM82.
The N-terminal amino acid sequences of oocyst wall proteins wp33, wp8, wp10 and wp12 map to the gametocyte protein GAM56 and GAM82 (Belli *et al.*, 2003b).

most likely scenario is that, after fertilization of macrogametocytes, the parasite receives a signal that stimulates the WFBs to aggregate around the periphery of the developing zygote and secrete their contents (Belli *et al.,* 2006). The WFBI move towards the surface, degranulate, and appear to fuse together and form the outer layer. Subsequently, WFBII also disaggregate and fuse beneath the outer layer to form the inner layer (Ferguson *et al.*, 2003b, Mouafo *et al.*, 2002a).

1.6.3 Molecular basis of oocyst wall formation in *Eimeria*

The prevalence of relatively large numbers of tyrosine residues in the amino acid sequences of oocyst wall proteins provides a clue to the assembly of the oocyst wall at a molecular level (Belli *et al.*, 2006b, Belli *et al.*, 2003b). Under oxidising conditions, particularly in reactions catalysed by, for example, peroxidase or catechol oxidase, tyrosine residues can be cross-linked to form dityrosine bonds or dihydroxyphenylalanine (DOPA), with resulting DOPA-quinone cross-linking of proteins (Waite, 1990a). These chemical reactions, sometimes referred to as quinone tanning or sclerotization, facilitate dehydration, close packing of polymers, and gradual hardening, and are widely used in nature to produce stable, protective extracellular matrices (Belli *et al.*, 2006b, Waite, 1990b). Thus, dityrosine bonding

and/or DOPA have been implicated in, for example, the formation of cuticles of insects (Hopkins, 1992) and nematodes (Parise and Bazzicalupo, 1997b), the egg shells of trematode parasites (Chen *et al.*, 1992), the fertilisation membrane of sea urchins (Foerder and Shapiro, 1977a, Weidman *et al.*, 1985), the outer ascophore layer of yeast (Briza *et al.*, 1990, Smail *et al.*, 1995) and cocoon formation in the silk moth, *Bombyx mori* (Zhou *et al.*, 2001).

Monne and Honnig (1954) were the first to propose that the oocyst wall contained quinone-tanned protein, an idea reiterated by Pittilo and Ball (1980) and resurrected by Eschenbacher *et al.* (1996), when they discovered tyrosine-rich 14 kDa proteins in the oocyst walls of *E. tenella* and *E. acervulina.* However, it was not until recently that chemical evidence of this phenomenon was obtained. A characteristic of protein-tyrosine crosslinks and quinone-tanned protein is its natural blue UV autofluorescence and, significantly, Belli *et al.* (2003a) demonstrated that, under an excitation wavelength of 330-385ηm, the oocysts of *E. maxima* autofluoresced blue but macrogametocytes did not (Figure 1.5). The oocyst walls of other coccidians (*eg, Toxoplasma, Neospora, Isospora* and *Cyclspora*) have also been shown to autofluoresce blue in a similar fashion to *Eimeria,* suggesting they too share a common mechanism that involves protein-tyrosine crosslinks (Belli *et al.,* 2006). Importantly, High Performance Liquid Chromatography (HPLC) confirmed the presence of dityrosine bonds and DOPA in the oocyst walls of *E. maxima*. The levels found were higher than levels detected in other organisms or in human arthrosclerosis samples (Davies *et al.*, 1999).

These high levels of dityrosine and DOPA imply a deliberate synthesis of these bonds rather than an incidental formation of these oxidation products, an idea lent credibility by the detection of peroxidase activity specifically in the wall forming bodies and immature oocyst wall of *E. maxima* (Belli *et al.,* 2003a; 2006). On the strength of these observations, Belli *et al.* (2006) put forward a model for the molecular basis of oocyst wall formation (Figure 1.6). Briefly, they proposed that large tyrosine-rich proteins (GAM56 and GAM82), which they referred to as “scaffolding proteins”, are synthesised and stockpiled in the wall forming bodies of the macrogametocyte until a (undefined) signal is received, initiating their processing into smaller, tyrosine-rich proteins. The tyrosine residues are cross-linked via an oxidative reaction, and the now cross-linked proteins form a stable matrix that dehydrates and hardens to form the oocyst wall, with its inherent resistance to chemical and environmental challenges. Homologues to GAM56 and GAM82 have yet to be identified in

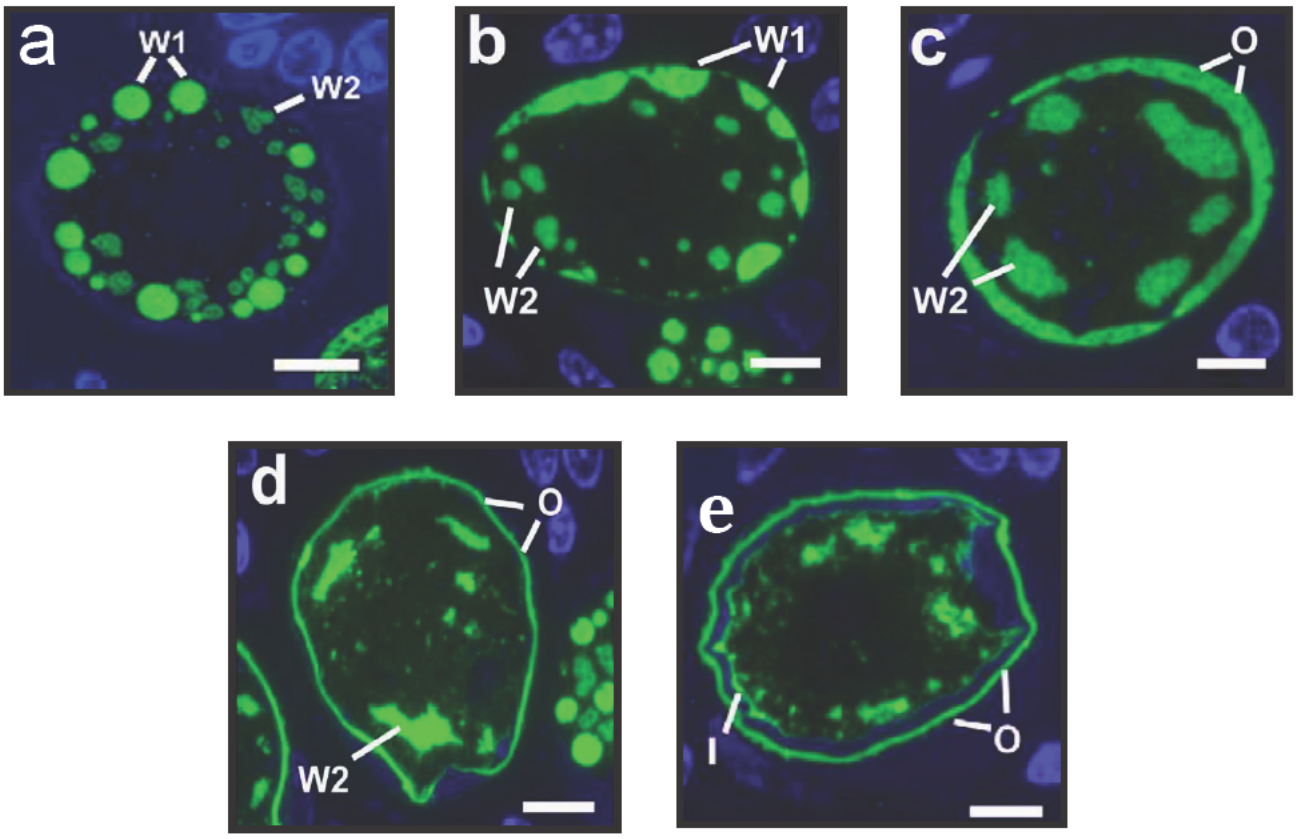
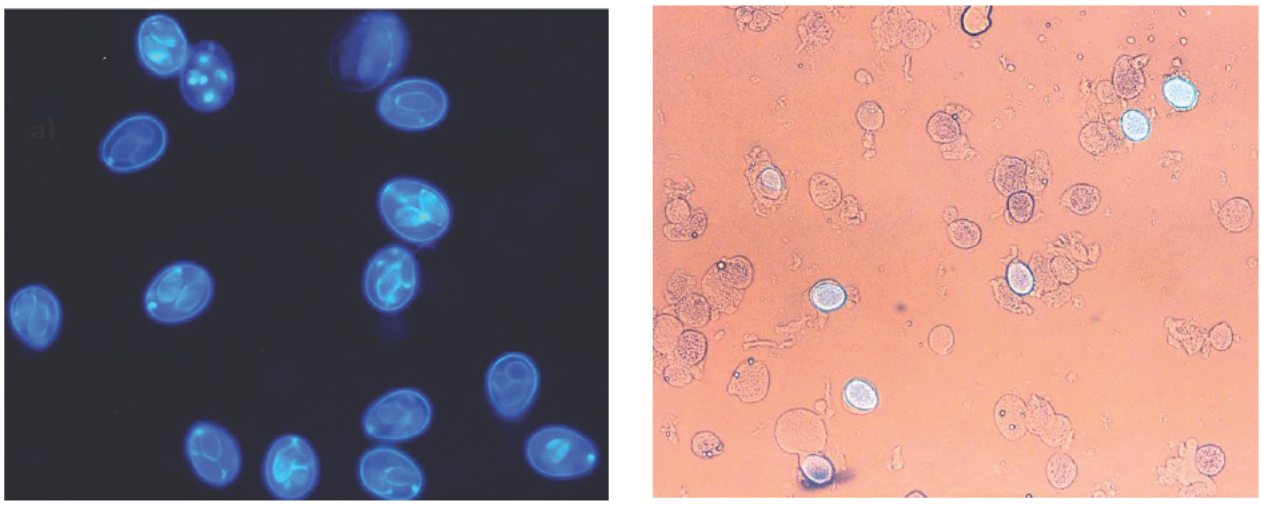


Figure 1.4: Oocyst wall biogenesis stages in *Eimeria maxima.*

Oocyst wall biogenesis stages – macrogametocyte, zygote and developing unsporulated oocyst immune-labelled with anti-APGA and visualised with fluorescein isothiocyanate (FITC). (a) Wall forming bodies 1 and 2 (WFB1 and WFB2) present in a macrogametocyte. (b-d) Contents of wall forming bodies 1 begin forming a layer around the periphery of the developing macrogametocyte to form the outer layer of the oocyst wall, whilst content from WFB2 disaggregate and move to periphery of immature oocyst (d,e) and begin to degranulate to form the inner oocyst wall (Ferguson *et al.*, 2003b).

Abbreviations: WI, wall forming body type 1; W2, wall forming body type 2; I, inner oocyst wall; and O, outer oocyst wall.



a) b)

Figure 1.5: Figure 1.5: UV autofluorescence of *E. maxima* oocysts.

(a) Purified oocysts visualized under 330- to 385-nm; b) Mixed culture of purified gametocytes (and developing oocysts) visualized under 330- to 385-n. Only the oocysts autofluoresce (Belli *et al.*, 2003a).

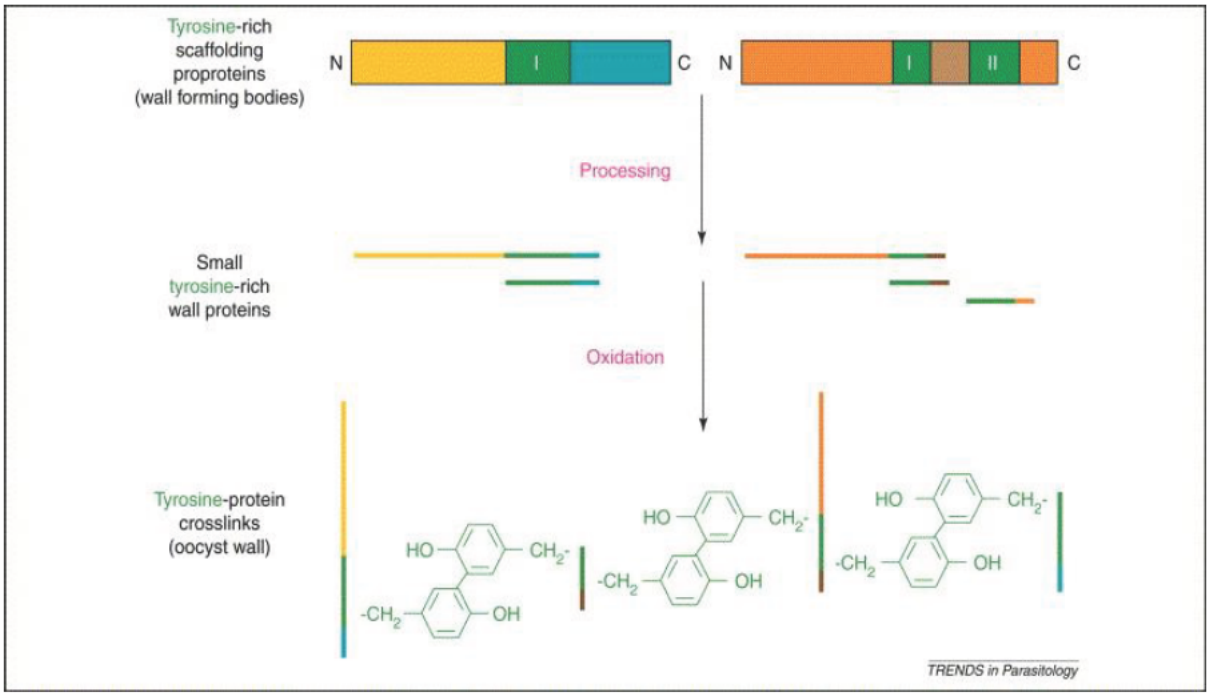


Figure 1.6: Molecular basis of oocyst wall formation.

Tyrosine-rich precursor proteins (GAM56 and GAM82), derived from WFB2 in macrogametocytes, are processed into smaller tyrosine-rich proteins found in the oocyst wall. These are cross-linked via dityrosine bonds to form a hard, resistant oocyst wall (Belli *et al.*, 2006a)

the genomes of other Coccidia. However, the recent identification of six tyrosine-rich proteins in the oocyst proteome of *T. gondii* (Fritz *et al.*, 2012) further supports the idea that these proteins may be cross-linked through their tyrosine residues, suggesting that it shares a common mechanism of oocyst wall assembly with *Eimeria.*

The evidence presented suggests that the formation of the oocyst wall of *Eimeria* is an enzyme-mediated reaction (Belli *et al.*, 2006b). However, several constituents required for the formation of these bonds are, as yet, to be identified in this parasite. The method by which GAM56 and GAM82 are processed to smaller wall peptides is not yet known. Available evidence suggests that subtilase serine proteases may proteolytically cleave the larger proteins into smaller peptides that are then incorporated into the wall as they undergo oxidation (Katrib *et al.*, 2012). Similarly, the enzyme(s) involved in the oxidative reaction of oocyst wall formation, specifically the formation of the dityrosine crosslinks, is yet to be discerned although peroxidase activity has been shown to localise exquisitely to wall forming bodies and developing oocyst walls (Belli *et al.,* 2006). Whilst there is only limited evidence available supporting either process during oocyst wall formation, there is considerable evidence in the best-studied analogous system to that of the oocyst wall, namely, the formation of the cuticle of nematode worms, especially the model organism, *Caenorhabditis elegans.*

1.7 Mechanism of Cuticle Formation in Nematodes – Analogies to Oocyst Wall Formation

Caenorhabditis elegans is a free-living nematode that is covered with a cuticle that consists predominantly of small collagen-like proteins. The cuticle functions as: (i) a barrier between the animal and its surrounding environment; (ii) a scaffold to maintain body form and; (iii) an anchor to facilitate the movement of attached muscles (Johnstone, 2000a). The collagens that form the cuticle are, like the oocyst wall proteins, initially synthesised as pro-proteins. These pro-proteins possess a conserved subtilisin-like serine protease cleavage pattern RX(K/R)R, recognition motif for proteolytic cleavage (Yang and Kramer, 1994). In the first step leading to cuticle formation, these pro-proteins are modified, trimerised and cleaved at the N-terminus by a subtilisin-like protease, blisterase (Johnstone, 2000a, Thacker *et al.*, 1995, Thacker and Rose, 2000, Yang and Kramer, 1994), and at the C-terminus by an astacin-like zinc metalloproteinase, to form truncated collagen proteins (Novelli *et al.*, 2004, Thacker and Rose, 2000). Blisterase was first described by Thacker *et al* (1995), who characterised the gene *Bli-4*, which is alternatively spliced at its 3` end to produce four serine proteases belonging to the Kex-2/Subtilisin-like Proprotein Convertase family of endoproteases.

Mutations at this allele, *bli-4 e937,* resulted in the blistering of the cuticle, therefore directly indicating its functional role in cuticle formation. The cuticle collagens form di- and tri-tyrosine cross-links, formation of which are oxidatively catalysed by the enzyme, dual oxidase (Page and Winter, 2003a).

Similar observations have recently been made concerning the formation of the cuticle of a parasitic nematode, the filarial parasite, *Onchocerca volvulus,* a major cause of blindness in Africa (Berger and Nnadozie, 1993). Poole *et al* (2003) were the first to biochemically characterise a subtilisin-like proprotein convertase in *O. volvulus.* Cloned and expressed, the enzyme was found to exhibit the conserved catalytic domain of subtilisin-like serine proteases and was 84% homologous to the *C. elegans* enzyme. Studies showed it has maximal activity in 1mM [Ca^{2+}] and is inhibited by its own propeptide (Poole *et al.*, 2003), a characteristic of subtilases. It is, therefore, thought to play a role in cuticle formation in the parasitic nematode, homologous to the *C. elegans* model. The analogies between this model of cuticle formation and oocyst wall formation of *Eimeria* make it tempting to speculate involvement of subtilases in the processing of GAM56 and GAM82 into the smaller tyrosine-rich proteins ultimately incorporated into the oocyst wall.

1.8 Subtilisin-like Serine Proteases

The superfamily of subtilisin-like serine proteases is a group of over 40 serine endo/exo-peptidases that catalyse the hydrolysis of covalent peptide bonds (Rawlings and Barrett, 1994). At present, subtilases (S8) are grouped into two subfamilies, S8A-subtilisins and S8B-kexins (http://merops.sanger.ac.uk/). Subtilases were first identified by Siezen *et al.* (1997) (Siezen and Leunissen, 1997a) and were characterised by the order in which their catalytic triad residues, aspartic acid, histidine and serine, are arranged in their primary sequence (Siezen and Leunissen, 1997b). The arrangement of these catalytic residues is shared with the prototypical subtilisin from *Bacillus licheniformis*. This core region gives the enzyme its proteolytic function; serine acts as a nucleophile towards peptide bonds, histidine is the general base that accepts the proton from the nucleophilic OH- group, and the aspartic acid residue stabilises and orientates the general base in the correct position (Siezen *et al.*, 2007b). An asparagine residue on a side chain to the catalytic domain contributes to the oxyanion binding site (Siezen *et al.*, 2007a). Further investigation into this highly conserved domain by Siezen *et al.* (2007) has shown that, whilst this catalytic domain is highly conserved throughout all subtilases, there are other amino acids that could occupy the domain as long as the integrity of the site is maintained. Thus, it always has a fully conserved nucleophile serine residue, whilst the nature and position of the general base (histidine) and acid (aspartic acid) residues can be found in different combinations (Siezen *et al.*, 2007b). In

addition, four calcium binding sites have been described in the known crystal structures of subtilases and [Ca^{2+}] has been shown to be essential for stability and activity (Siezen and Leunissen, 1997b).

Subtilases are characteristically synthesised as zymogens (precursor proteins) that consist of a signal peptide, a propeptide and a protease domain (Siezen and Leunissen, 1997b). The propeptide acts as an intracellular chaperone, essential for correct folding of the catalytic domain and enzyme maturation (Baker *et al.*, 1993, Ikemura and Inouye, 1988). Folding precedes auto-proteolytic cleavage of the propeptide domain at the N-terminal, activating the enzyme (Siezen and Leunissen, 1997b). However, the cleaved propeptide often remains transiently associated with the catalytic domain and, therefore, enzyme activation can require further proteolytic cleavage and release of the propeptide (Li *et al.*, 1995, Siezen and Leunissen, 1997b, Yabuta *et al.*, 2001).

1.9 Subtilisin-like serine proteases in Apicomplexa

Several subtilisin-like serine proteases have been identified in apicomplexan parasites (Table 1.2). These enzymes possess all the functional characteristics of subtilisin-like serine proteases, distinguishing them from other groups of serine proteases. Thus, like all subtilases they have the conserved catalytic triad of Asp-His-Ser (Siezen *et al.*, 1991, Siezen and Leunissen, 1997b) and function optimally in 1mM [Ca^{2+}](Miller *et al.*, 2001a, Montero *et al.*, 2006a, Withers-Martinez *et al.*, 2002). They also all undergo auto-post-translational processing themselves by a serine protease, in a two-step process that begins with cleavage of the synthesised zymogen and results in an active enzyme stored predominantly in the secretory organelles of the parasite; for example, *Plasmodium falciparum* subtilisin-1 is stored in the dense granules of merozoites, prior to red blood cell invasion (Blackman *et al.*, 1998b).

Studies on subtilases in the Apicomplexa have focused on *Plasmodium* and *Toxoplasma,* due to the severe resulting clinical disease in humans and the relative ease of *in vitro* cultivation of the asexual stages of these parasites. It has been suggested that targeting of proteases involved in invasion may have therapeutic value (Blackman, 2000b). The inhibitors 3,4-dicholoroisocoumarin (3,4-DCI) and 4-(2-aminoethyl)-benzenesulfonyl fluoride (AEBSF), both general serine protease inhibitors, have been shown to inhibit host cell invasion by *T. gondii* (Carruthers and Sibley, 1999, Conseil *et al.*, 1999). The inhibitors tosyl phenylalanyl chloromethyl ketone (TPCK), an irreversible inhibitor of chymotrypsin and some cysteine proteases, and tosyl-L-lysine chloromethyl ketone (TLCK), a specific inhibitor of the serine protease trypsin as well as thrombin-like cerastocytin, have similar effects on *Babesia*

Subtilisin-like Serine Protease	Storage Organelle	Proteolytic Function
Toxoplasma gondii TgSUB1 (Miller *et al.*, 2001b)	Micronemes	Secreted from micronemes concomitantly to invasion, and is therefore suggested to play a role in processing parasite adhesion and invasion of proteins; (Miller *et al.*, 2001b, Opitz *et al.*, 2002b, Zhou *et al.*, 2004b) is homologous to PfAM1 (Hehl *et al.*, 2000).
Toxoplasma gondii TgSUB2 (Miller *et al.*, 2003b)	Rhoptries	Current evidence suggests it has a role in Rhoptry protein maturation, with ROP1 (Rhoptry Protein-1) proposed as the best candidate (Miller *et al.*, 2003b, Kim, 2004).
Toxoplasma gondii Rhomboid-like (MPP1) (Zhou *et al.*, 2004a, Carruthers *et al.*, 2000b, Freeman, 2004, Carruthers *et al.*, 2000a, Zhou *et al.*, 2004b)	Micronemes	MPPI cleaves TgMIC2 , part of the MgCI2/M2AP complex in , (which it itself is cleaved by anonymous proteases MPP2 and MPP3)(Rabenau *et al.*, 2001b) (Rabenau *et al.*, 2001a) from the trans-membrane domain on the tachytzoite surface however it remains transiently attached. (Carruthers *et al.*, 2000b, Opitz *et al.*, 2002a, Zhou *et al.*, 2004a). A second unknown protease further cleaves TgMIC2, releasing it from the surface, culminating in shedding of the adhesion microneme proteins concomitantly to invasion.
Plasmodium falciparum PfSUB1 (Blackman *et al.*, 1998b)	Newly discovered parasite organelles, the exonemes (Yeoh *et al.*, 2007)	Released into the parasitophorous vacoule (PV) in the latter stages of intracellular merozoite maturation. Shown to cleave SERA4, SERA5 and SERA6 (serine repeat antigens) which triggers the pathway that leads to parasite egress from erythrocyte (Arastu-Kapur *et al.*, 2008, Yeoh *et al.*, 2007)
Plasmodium falciparum PfSub2 (Barale *et al.*, 1999a, Hackett *et al.*, 1999a, Barale *et al.*, 1999b, Hackett *et al.*, 1999b)	Micronemes	Secreted onto the parasite surface following schizont rupture and cleaves surface antigen protein MSP1 (Merozoite Surface Protein-1) (Blackman, 2000a, Goel *et al.*, 2003a, Li *et al.*, 2004, Blackman, 2000b, Goel *et al.*, 2003b) and apical invasion protein AMA1 (Apical Membrane Antigen-1) (Howell *et al.*, 2001a, Howell *et al.*, 2001b) by a single proteolytic cleave at the juxtamembrane site of attachment (Harris *et al.*, 2005). This results in shedding of the proteins, which occurs concomitantly with invasion of RBCs.
Plasmodium falciparum PfSub3 (Alam *et al.*, 2012)	Unknown	Expression in late stages of erythrocytic schizogony indicate it may have a role in merozoite egress and/or invasion (Alam *et al.,* 2012)
Plasmodium berghei PbSub2 (Han and Barillas-Mury, 2002a, Han and Barillas-Mury, 2002b)	Ookinetes	Secreted concomitantly as motile ookinetes penetrate the peritrophic matrix (PM) and invade midgut epithelium (Han and Barillas-Mury, 2002a) Also secreted concomitantly in merozoites stage during RBC invasion (Han and Barillas-Mury, 2002a, Uzureau *et al.*, 2004b, Han and Barillas-Mury, 2002b, Uzureau *et al.*, 2004a)
Neospora caninum NcSUB1 (Louie and Conrad, 1999)		Unknown, proposed substrate NcAMA1 (Miller *et al.*, 2001a, Zhang *et al.*, 2007).
Babesia divergens BdSUB1 (Montero *et al.*, 2006a, Montero *et al.*, 2006b)	Dense Granules	Unknown

Table 1.2. Subtilisin-like serine proteases identified in apicomplexan parasites

divergens, suggesting intracellular invasion by parasites is accompanied by proteolytic modification of proteins associated with host cell entry.

Invasion of *Plasmodium* into host erythrocytes is potently inhibited by leupeptin, a calpain protease inhibitor that inhibits both serine and cysteine proteases, and chymostatin, a chymotrypsin protease inhibitor (Blackman, 2000a, Dluzewski *et al.*, 1986). Yeoh *et al.* (2007) demonstrated that the inhibitor MRT12113 abrogated host cell rupture and the few *Plasmodium* merozoites released were less able to re-invade (Yeoh *et al.*, 2007). Further work with the serine protease inhibitor, biotin-tagged chloroisocoumarin JCP104, specifically inhibited parasite egress (Arastu-Kapur *et al.*, 2008), implicating proteolytic activity and maturation of invasion pathway proteins by subtilisin-like serine proteases, not only in parasite invasion, but also in egress from the host cell (Arastu-Kapur *et al.*, 2008).

1.10 Subtilisin-like serine proteases in *Eimeria*

Similar to the above studies of *Toxoplasma* and *Plasmodium,* a number of proteases have been implicated in the complex processes involved in *Eimeria* spp. survival including sporulation, intracellular invasion, evasion of host immune responses, reproduction and nutrition (Fetterer *et al.*, 2007). Research on serine proteases in *E. tenella* has also focused on their role in intracellular invasion. A 20kDa trypsin-like protease was partially purified from *E. tenella* sporulated oocysts (Michalski *et al.*, 1994, Fuller and McDougald, 1990). Fuller and McDougald (1990) showed that serine protease inhibitors leupeptin, TPCK and phenyl methylsufonyl fluoride (PMSF) reduced merozoite and sporozoite invasion of *E. tenella* into host epithelial cells *in vitro,* suggesting that trypsin or serine proteases may be involved. However, bioinformatic analysis showed that serine proteases in the *E. tenella* genome lack homology to trypsin (Fetterer *et* al., 2007) and, instead, have homology to subtilisin-like serine proteases. Recent characterisation of genes differentially expressed during sporulation showed transcripts of several proteases present in *E. tenella* (Miska *et al.*, 2004), including a subtilisin-like serine protease (Miska *et al.*, 2004), a 24kDa rhomboid protease (Li *et al.*, 2006), prolyl-endopeptidase, a metalloprotease, and aminopeptidase-N (Fetterer and Barfield, 2003, Fetterer *et al.*, 2007, Miska *et al.*, 2004). The timing of expression of these proteases again suggests a role for serine proteases in sporozoite invasion of cells. To date, exploration of subtilisin-like proteases in apicomplexan parasites has been limited largely to the asexual stages of the lifecycle and has focused on their potential role in cell invasion and egress; recently, however, some interest has been shown in the role of proteases in sexual stages of *E. tenella.*

In similar studies to those described above, it has recently been shown that the serine protease inhibitors aprotinin, leupetin, and the metal chelating agent ethylenediaminetetraacetic acid (EDTA), inhibit the processing of *E. tenella* GAM56 protein (Katrib *et al.* 2012). In contrast, a range of protease inhibitors had little or no effect on the prevention of degradation of GAM56. These included aspartic and cathepsin, metallo-protease and cysteine protease inhibitors. Interestingly, AEBSF was found to also have little effect on the degradation of GAM56. However, in a study by Sajid *et al* (Sajid *et al.*, 2000b) it was found that *Pf*Sub1 (see Table 1.2), a merozoite subtilase involved in RBC invasion, is not inhibited by either AEBSF or PMSF, both standard serine protease inhibitors. It has also been shown that, whilst EDTA is not commonly a subtilase inhibitor, it has been known to inhibit the activity of subtilisin-like serine proteases (Poole *et al.*, 2003), as also reflected in the GAM56 assay (Katrib *et al.,* 2012), indicating a requirement for divalent cations ($Ca^{2+)}$ for enzymatic activity. This suggests that a serine protease is active in the processing of GAM56 protein. As stated above, GAM56 processing occurs during oocyst wall formation; therefore, serine proteases may be fundamental in oocyst wall formation.

In the study by Katrib *et al* (2012) over 40 proteases were identified in the genomic database of *E. tenella.* These proteases were classified as aspartic, cysteine, metallo and serine proteases. Interestingly, expression of thirteen of these protease genes was specific or up-regulated in gametocytes of the parasite. Of the twelve serine proteases putatively identified in *E. tenella,* six were found to be subtilisin-like serine proteases, based on BLAST homology and signature protein motif searches (Katrib *et al.,* 2012), and expression of four of these was specific or up-regulated in gametocytes. The identification of these subtilases, similar to the *Blisterase* subtilisin-like protease identified in *C. elegans,* warrants further investigation into whether any of these genes play a role in oocyst wall formation of *E. tenella,* specifically, the processing of oocyst wall proteins GAM56 and GAM82.

1.11 Oxioreductase Catalysation of Dityrosine Formation

1.11.1 Oxioreductases

Oxioreductases are the broad range of enzymes that catalyse the transfer of electrons from one molecule (the oxidant, or donor) to another (the reductant, or acceptor) according to the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB). Oxioreductases can be classified into 22 subclasses, depending on the reaction they catalyse. These include oxidases, dehydrogenases, peroxidases, hydroxylases, oxygenases, and reductases. These enzymes play an important role in both aerobic and anaerobic mechanisms, found in many pathways including glycolysis (Kim and Dang, 2005,

Lowry and Passonneau, 1964), the TCA cycle (Fernie *et al.*, 2004, Velot *et al.*, 1997), oxidative phosphorylation (Chance and Williams, 1955, Hatefi, 1985) and amino acid metabolism (Felig, 1975). They are also important in defence against the host's immune response (Quinn and Gauss, 2004) and in the formation of structural biomaterials (Lambeth, 2004).

1.11.2 Oxioreductases in Apicomplexa

Apicomplexan parasites are critically dependent on host-cell nutrients for energy and carbon sources, particularly during host-cell invasion and intracellular growth and replication. As such, study of their respiration and metabolism has focused heavily on this area and the oxioreductases identified reflect this. Both *Plasmodium* and *T. gondii* possess the complete set of genes coding for the TCA cycle and glycolytic enzymes (Polonais and Soldati-Favre, 2010). There is also evidence of a functional respiratory chain and oxidative phosphorylation in both *T. gondii* (Lin *et al.*, 2009, Vercesi *et al.*, 1998) and *P. yoelii* (Uyemura *et al.*, 2004) including the identification of succinyl-CoA synthase in *T. gondii* (Fleige *et al.*, 2008). Two respiratory dehydrogenases have also been characterised in active mitochondria of isolated *P. yoelii* infected red blood cells (Kawahara *et al.*, 2009). Similarly, both *T. gondii* and *Plasmodium* have highly effective defense mechanisms to cope with oxidative stress imposed by the host during their intracellular stages and a number of oxioreductases identified within these parasites are involved in these pathways. During the erythrocytic stage of the life cycle of *P. falciparum* the parasite is exposed to reactive oxygen species (ROS) as a result of oxidative metabolism of hemoglobin and also toxic ferriprotoporphyrin IX (Muller, 2004, Nickel *et al.*, 2006). As the parasite lacks the antioxidant enzymes, catalase and glutathione peroxidase, it relies instead on the antioxidant capacity of its host cells and own peroxidases. Central to the response to this oxidative stress is the oxioreductase, superoxide dismutase (Muller, 2004), as well as five described peroxidases – two 2-Cys peroxiredoxins (Prx), a 1-Cys Prx, an antioxidant protein and a glutathione peroxidase-like thioredoxin peroxidase(Nickel *et al.*, 2006). A thioredoxin reductase is also essential for survival of intraerythrocytic *P. falciparum* (Muller 2004). *Plasmodium falciparum* also contains a fully functional glutathione redox system (Muller, 2004).

A number of oxioreductases have also been identified in *T. gondii.* These form an antioxidant network to protect against endogenously produced ROS intermediates from cellular respiration and metabolism, and oxidative stress imposed by the host. In contrast to *P. falciparum,* a cystolic catalase is present in *T. gondii* and it is thought to protect the parasite from H_2O_2 exposure (Ding *et al.*, 2005). A superoxide dismutase and a thioredoxin-like dependent peroxidase have also been identified in *T. gondii* (Pino *et al.*, 2007). Interestingly, these two antioxidant enzymes target both the apicoplast and mitochondrion of *T. gondii*

(Pino *et al.*, 2007), and are hypothesised to act as antioxidant protectants during carbon metabolism in both organelles.

1.11.3 Oxioreductases in *Eimeria*

The intestinal pathology of poultry coccidiosis is not just the result of the mechanics of invasion by *Eimeria*, but also the ROS-induced damage that can occur as a result of the immune response to infection (Fridovich, 1978, Georgieva and Gadjeva, 2005, Georgieva *et al.*, 2006, Halliwell, 1974). As an intracellular parasite, *Eimeria* is faced with these chicken host cellular immune responses, including the generation of highly reactive species like superoxide and nitric oxide (Allen and Teasdale, 1994, Allen, 1997, Ovington and Smith, 1992). As discussed previously, the coccidian parasite, *T. gondii,* employs a number of oxioreductase enzymes in its antioxidant defence against host immune responses, including superoxide dismutase, *Tg*SOD, catalase and thioredoxin-like peroxidase *Tg*TPX1/2 (Pino *et al.*, 2007). Similarly, superoxide dismutase has been described in *E. tenella,* with high levels detected in unsporulated oocysts, and low levels in sporozoites (Michalski and Prowse, 1991). Additionally, a number of oxioreductase enzymes were identified by Williams (Williams, 1999) including a hydroxybutyrate dehydrogenase, an alanine aminotransferase and a γ-glutamyltransferase in both *E. tenella* and *E. maxima.* This same study utilised a number of oxioreductase group enzymes identified previously in *Eimeria* including a creatine kinase (Andrews *et al.*, 1990), glutamate dehydrogenase (Wang *et al.*, 1979) and lactate dehydrogenase (Shirley, 1989, Shirley and Millard, 1986) to try and identify a inhibitory compound to use as an anticoccidial. These enzymes are predominantly associated with the metabolism and respiration of the parasite. Two oxidases have also been identified in *E. tenella,* a succinate oxidase and a NADH oxidase, both with a role in mitochondrial respiration in *Eimeria* (Wang, 1975, Fry *et al.*, 1984, Fry and Williams, 1984).

Previous work has identified peroxidase activity in the wall-forming bodies of macrogametocytes of *E. maxima*, but not in microgametocytes (Belli *et al.*, 2003b). An *in vitro* study also demonstrated that exogenous peroxidases from horseradish and *Arthromyces* were highly effective at catalyzing the rapid formation of a recombinant truncated oocyst wall protein into dimer, trimer and polymers of the protein (Mai *et al.*, 2011). High levels of dityrosine were detected in the polymer samples (Mai *et al.*, 2011), consistent with the dityrosine crosslinks that were observed in oocyst walls. Both *Arthromyces* peroxidase and horseradish peroxidase have been used previously to catalyze dityrosine cross-links. Peroxidase from *Arthromyces* can catalyse dityrosine cross-linking of calmodulin (Aeschbach *et al.*, 1976). Similarly, horseradish peroxidases have been shown to catalyse the formation of dityrosine dimers between peptides *in vitro* (Malencik and

Anderson, 1996), as well as dityrosine cross-linking for silk fibroin (Anderson 1966), and CUT-2, the protein component of the *C. elegans* cuticle (Parise and Bazzicalupo, 1997a). Indeed, a number of different extracellular matrices in physiological systems involve a peroxidase, or an oxidase-type enzyme in the catalysis of tyrosine crosslink formation, resulting in di- and tri-tyrosine formation, most especially nematode cuticles and the fertilisation membrane of sea urchins.

1.11.4 Peroxidase activity in cuticle formation of *Caenorhabditis elegans*

In the model of nematode cuticle biogenesis described above, collagens that form the cuticle are, like the oocyst wall proteins, initially synthesised as pro-proteins. Following processing of these, cross-linking of collagen and other cuticle proteins in *C. elegans* occurs through di- and trityrosine linkages, which bridge and stabilise the proteinaceous structure (Fetterer and Hill, 1993, Fetterer and Rhoads, 1990). A dual oxidase (Duox) Ce-Duox1 is thought to catalyse this reaction (Edens *et al.*, 2001). A high molecular weight homologue of gp91*phox*, the superoxide-generating subunit of phagocyte nicotinamine adenine dinucleotide phosphate (NADPH)-oxidase, Ce-Duox-1, is expressed in the hypodermal cells, underlying the cuticle of the larval worm (Edens *et al.*, 2001). RNA interference (RNAi) was used to investigate the role of this protein in developing nematodes. In animals where expression was absent, gross abnormalities were observed in the cuticle, *eg:* superficial blisters; 'dumpy' (or shortened) worms; worms with retained eggs or larvae; and worms that also appeared translucent (whereas wild-type worms are dark in appearance) (Edens *et al.*, 2001). These observations have previously been associated with mutations in the collagen biosynthetic pathway (Gupta *et al.*, 1997, Johnstone, 2000b, Levy *et al.*, 1993), consistent with a function of Ce-Duox1 in cuticle biogenesis (Edens *et al.*, 2001). Further investigation revealed that interference with the expression of Ce-Duox1 in *C. elegans* eliminated the formation of di- and trityrosine crosslinks that form between the collagen and other proteins during cuticle formation. Expression of the peroxidase homologous domain of Ce-Duox1 resulted in peroxidase activity towards 3,3`,5,5`-tetramethylbenzidine (TMB), and this activity was inhibited by a peroxidase inhibitor, aminobenzohydrazide. It was further demonstrated that this expressed peroxidase domain catalysed the cross-linking of tyrosine ethyl esters, consistent with the idea that Ce-Duox1 catalyses the cross-linking of tyrosine residues involved in the stabilization of the cuticular extracellular matrix (Donko *et al.*, 2005, Edens *et al.*, 2001). However, it is not yet known if this is a coupled reaction, whereby the NADPH oxidase portion of Ce-Duox1 generates the H_2O_2 for the crosslinking reaction, which is then catalysed by the peroxidase-like portion of the molecule, or if there is a distinct peroxidase to catalyse

the cross-linking reaction and the Duox is solely responsible for H_2O_2 production (Edens *et al.*, 2001)

1.11.5 Peroxidase activity in the formation of fertilization membrane of the sea urchin

Early evidence of involvement of a peroxidase in extracellular matrix structure was observed in the fertilisation reaction in sea urchin oocytes (Foerder and Shapiro, 1977b). Formation of the fertilisation membrane in echinoderm eggs is initiated by an increase in Ca^{2+} that induces the exocytosis of secretory granules. Content from the cortical granules then associates with the vitelline membrane, forming a soft fertilisation membrane (Heinecke and Shapiro, 1992). Amongst the secreted proteins is ovoperoxidase, a peroxidase that, using the oxidative substrate H_2O_2, catalyses cross-links in the envelope proteins through dityrosine formation, leading to the formation of a hard, protective envelope (Deits *et al.*, 1984, Heinecke and Shapiro, 1992). The production of H_2O_2 in this reaction is a deliberate mechanism, produced by a second oxioreductase, sea urchin dual oxidase (Udx1) (Wong *et al.*, 2004). Similar to the mammalian homologues, the peroxidase-like domain of Udx1 lacks conserved histidines, necessary for heme binding, typical of a 'classical' peroxidase. Therefore, this enzyme's sole function is the production of H_2O_2, suggesting that the primary outcome of its activation is the catalysis of a cross-linking reaction (Donko *et al.*, 2005).

1.12 Hypothesis and Aims of this Study

The overarching hypothesis for this research project is that a subtilisin-like serine protease proteolytically processes glycoproteins GAM56 and GAM82 into smaller proteins that are, subsequently, enzymatically oxidised by an oxioreductase to form dityrosine cross-links within the oocyst wall of *E. tenella.* The specific aims of this thesis are to:

[1] identify which, if any, of the subtilisin-like serine protease genes identified previously by Katrib *et al* (2012) have specific or up-regulated expression in gametocytes of *E. tenella;*

[2] confirm which, if any, of the subtilisin-like serine proteases identified to have gametocyte up-regulated expression are gametocyte-specific at the protein level and determine their subcellular localisation*;*

[3] identify putative oxioreductase genes in the *E. tenella* genome by data mining both the *T. gondii* and *E. tenella* genome databases;

[4] identify which, if any, oxioreductases in the *E. tenella* genome have gametocyte-specific or highly up-regulated expression in gametocytes.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

For convenience, the chemical reagents, biological reagents and miscellaneous materials listed below are listed alphabetically.

2.1.1 Chemical Reagents

Unless otherwise stated, all chemicals used throughout this study were of molecular grade.

- Absolute ethanol (Sigma, USA), Acetic acid (Amresco®, USA), Advantage 2 DNA polymerase (50x) (Clontech), 10x Advantage Buffer (Clontech), (Amresco®, USA), Agarose (Amresco®, USA), Antarctic phosphatase (New England® Biolabs Inc, USA)
- Bacteriological agar (Difco Laboratories, USA), Baxter water (Baxter), β-mercaptoethanol (Sigma Aldrich), Bovine Serum Albumin (Cohn Fraction V)(Sigma, USA), Bromophenol Blue (Sigma, USA)
- Calcium chloride ($CaCl_2$) (Sigma Aldrich®), Chloroform (Sigma, USA), Coomassie Brilliant Blue R250 (Sigma, USA), CPS-1 Chemiluminescent Peroxidase Substrate-1 (Sigma, USA)
- 4', 6-Diamidino-2-phenylindole dihydrochloride (DAPI) (Sigma, USA), Diethylaminoethyl cellulose (DE-52)- pre-swollen microgranular anion exchange (Whatman International Ltd. England), Dipotassium phosphate (K_2HPO_4) (Sigma Aldrich®), Dithiothreitol (DTT) (Sigma, USA), 10mM dNTP's (Clontech), 2mM dNTP's (Fisher Biotech)
- Ethanol (Merck Pty Ltd), Ethylenediaminetetraacetic acid (EDTA) (Astral Scientific), ethylene glycol tetra-acetic acid (EGTA) (Life Technologies™ (Invitrogen™), USA)
- FAST™ BCIP/NBT Buffered Substrate tablet (Sigma, USA), Foetal calf serum (Sigma, USA), Formaldehyde solution (Sigma Aldrich®), Formamide (Sigma Aldrich®), Freund's incomplete adjuvant (Sigma, USA), Freund's complete adjuvant (Sigma, USA)
- GelRed® Nucleic acid Gel Stain (10,000x in water) (Biotium, Inc. USA), Glacial acetic acid (Amresco®, USA), Glass wool (low in lead)(Ajax Chemicals, Australia), Glucose (BDH, Australia), L-glutamine (Sigma, USA), Glutathione (Sigma, USA), Glycerol (Sigma Aldrich), Glycine (Amresco®, USA)
- HEPES (Sigma, USA), Hyaluronidase (Sigma Aldrich), Hydrochloric acid (Chem Supply Pty Ltd)
- Iodoacetamide (Sigma Aldrich®), Imidazole (Sigma Aldrich), Isoamyl alcohol (Sigma, USA), Isopropyl-β-thiogalactopyranoside (IPTG) (Astral Scientific)

- Luria broth base (Millers) (Oxoid, England), Lysozyme from chicken egg white (Sigma Aldrich®)
- Methanol, (Merck Pty Ltd), 25mM $MgCl_2$ (Fisher Biotech), Monopotassium phosphate (KH_2PO_4) (Sigma Aldrich®)
- Penicillin (Sigma, USA), Phenol (Sigma, USA), phenol: chloroform: isoamyl alcohol mixture (Sigma, USA), phenyl methylsulfonyl fluoride (PMSF) Sigma, USA), 104® Phosphatase Substrate (p-nitrophenyl phosphate disodium) (Sigma, USA), polyoxyethyline sorbitan monolaurate (TWEEN-20®) (Sigma, USA), Potassium chloride (KCl) (Sigma Aldrich®), potassium dichromate (Sigma, USA), ProLong® Gold Antifade Reagent, (Life Technologies™ (Invitrogen™), USA), Protease Cocktail Inhibitor (Sigma Aldrich®)
- RNase free water (Gibco®, Australia)
- Sodium chloride (NaCl) (Sigma, USA), Sodium citrate (Ajax Chemicals, Australia), sodium dihydrogen orthophosphate monohydrate (NaH_2PO_4) (Sigma, USA), sodium dodecyl sulphate (SDS) (Sigma, USA), sodium hypochlorite (12.5% v/w) solution (Ajax Chemicals, Australia), sodium phosphate monohydrate disodium (Na_2HPO_4) (Sigma, USA)
- *Taq* polymerase (Fisher Biotech), Taurocholic acid (Sigma, USA), 10 x transcription buffer (Fisher Biotech), Triton x-100 (Amresco, USA), Tris base (Amresco, USA), Tris Hydrochloride (Sigma Aldrich®), Trizol® Reagent (Life Technologies™ (Invitrogen™), USA), Tween®-20 (Sigma Aldrich®)
- Whatman® No. 1 Filter Paper (Whatman International Ltd. England)
- Xylene cyanol (Sigma, USA), Xylol/xylenes (histological grade) (Sigma, USA)

2.1.2 Biological Reagents

- Agarose I (Amresco®, USA)
- SOB media (Amresco®, USA)
- Diploma Skim Milk Powder (SMP)(Bonlac Foods, Australia)
- Luria-Bertani base (Difco Laboratories)
- Nutrient Mixture F-12 (HAMS F-12) (Sigma Aldrich®, USA)
- Trypsin (Difco Laboratories), Tryptone (Difco Laboratories)
- Yeast extract (Difco Laboratories)

2.1.3 Miscellaneous equipment used

- Bioline (UK): Incubator Shaker 4500
- Bio-Rad (USA): Profinia Protein Purification Instrument, Gene Pulser cuvette (4mm gap), VersaDoc™ Imaging System
- Black and Decker®: 900W, 30L Microwave Oven
- Dynavac: Freeze Drier
- Carl Zeiss (Australia): Zeiss Axio Imager M1
- Endecotts (UK): stainless steel laboratory sieves – 1mm and 250µm
- FLUOstar OMEGA® multi-detection microplate reader (BMG Technologies, USA)
- Greiner Bio-One (Germany): 96 Well ELISA Microplates, Test tubes 15ml and 50ml
- Life Technologies™ (Invitrogen™) (USA): DynaMag™ Magnet, NuPAGE® 4-12% Bis-Tris Gel (10 and 12 x 1.0mm wells), NuPAGE® MES SDS Running Buffer (20x), XCell *SureLock*™ Novex Mini-Cell (SDS-PAGE Unit), XCell II™ Blot Module (Western blot unit)
- Hettich Instruments (Germany): Hettich Rotina 420R
- Millipore Corporation (USA): Immobilon™ Transfer Membranes (PVDF)
- MJ Research, Inc. (USA): PTC-200, Peltier Thermal Cycler
- Nanodrop Technologies, Inc. (USA): Nanodrop® ND-1000 Spectrophotometer
- Qiagen Rotor-Q Real Time PCR System (Qiagen, USA)
- Qsonica, LLC (USA): Misonix Sonicator Ultrasonic Processor S-4000
- Olympus Optical Co., Ltd. (Japan): Research BX51 Microscope (fluorescent)
- Penn-Plax, Inc. (USA): Air-Tech 2KOUK™ Aquarium Air Pump (air sparger), plant airline tubing
- Promega, USA: MagneSphere® Technology Magnetic Separation Stand (two-position)
- Sefar Filtration Oceania, Sydney: Sefar Nitex polymon mesh - 10µm and 17µm
- TissueLyser LT (Qiagen, USA)
- Ultralum, Inc. (USA): ultraviolet transilluminator

2.1.4 General solutions and growth media

- Alkylating buffer: reducing buffer supplemented with 200mM iodoacetamide
- Blocking buffer 1: 10% (w/v) skim milk powder (SMP) in PBS
- Blocking buffer 2: 1% (w/v) BSA in PBS
- Citrate buffer: 10mM sodium citrate, pH6.0
- Coomassie Blue solution: 0.2% Coomassie Brilliant Blue R250, 50% methanol, 10% glacial acetic acid
- Cytomix (Electroporation) Buffer: 10mM K_2HPO_4/KH_2PO_4 pH 7.6, 120mM KCl, 0.15mM $CaCl_2$, 25mM HEPES, 2mM EGTA, 5mM $MgCl_2$
- Destaining solution: 10% ethanol, 10% glacial acetic acid
- DNA loading buffer (6x): 18% glycerol, 0.15% Orange G, 0.15% bromophenol blue, 0.15% xylene cyanol
- ELISA Buffer 1: 0.1M carbonate buffer pH 9.5, 0.02% Sodium azide
- ELISA wash buffer: 0.05% (v/v) Tween 20 in PBS
- Excystation media: 10mM $MgCl_2$, 0.5% taurocholic acid and 0.25% trypsin in Hanks Balanced Salts solution
- Formaldehyde agarose gel buffer (10x): 200mM MOPS, 50mM sodium acetate, 10mM EDTA, pH to 7.0 with NaOH
- Formaldehyde agarose gel: 1.2% agarose in 1x formaldehyde agarose gel buffer
- Formaldehyde agarose gel running buffer (1x): 100ml 10x FA gel buffer, 20ml 40% formaldehyde, 880ml Baxter water
- His-tag wash buffer: 0.05% (v/v) Tween 20, 0.01% (v/v) Triton (Tx-100) in TBS
- 4 x Laemmli buffer: 250mM Tris-HCl (pH 6.8), 8% SDS, 40% Glycerol, 10% β-mercaptoethanol, 0.02% Bromophenol Blue
- Luria Broth media (LB media): 2.5% Luria broth base
- Luria Broth Plates (LB Plates): 2.5% Luria broth base, 0.25% agar
- Phosphate buffered saline (PBS) pH 7.4 (1x): 0.14 M NaCl, 2.68 mM KCl, 10.14 mM sodium phosphate dibasic, 1.76 M potassium dihydrogen orthophosphate
- PBS/Tween: 0.05% Tween-20 in PBS
- PBS/Tween/Triton: 0.05% Tween-20, 0.01% Triton-100 in PBS 1
- RNA loading buffer (5x): 16µl saturated aqueous bromophenol blue solution, 80µl 500mM EDTA, pH 8.0, 720µl 40% formaldehyde, 2ml glycerol, 3ml formamide, 4ml 10x FA gel buffer, water to 10ml

- SAC medium: 170mM NaCl, 10mM Tris-HCl, pH 7.0, 10mM glucose, 5mM $CaCl_2$, 1mg/ml BSA
- Sodium phosphate buffer pH 6.8 (0.1 M): 463mM M Na_2HPO_4, 537 mM NaH_2PO_4
- TBE electrophoresis buffer (1x): 89mM Tris, 89mM borate, 2mM EDTA
- TBS buffer: 10 mM Tris-HCl pH 7.5, 150 mM NaCl
- TE buffer: 10 mM Tris-HCl pH8.0, 1 mM EDTA
- Transfer buffer: 25 mM Tris, 192mM glycine, 20% (v/v) methanol
- Tris-HCl buffer, pH 10.0 (IFM): 100mM Tris, pH10.0
- Triton solution: 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 0.05% Tween-20, 0.2% Triton (Tx-100)
- Laemmli loading buffer (1x): 62 mM Tris-HCl pH6.8, 10% glycerol, 2.3% SDS, 5% β-mercaptoethanol, 0.5% bromophenol blue

2.1.5 Commercial kits and columns

- Advantage 2 Polymerase Kit (Clontech, USA)
- Bio-Scale Mini Profinity IMAC Cartridges (1ml) (Bio-Rad, USA)
- Bio-Scale Mini Bio-Gel P-6 Desalting Cartridges (10ml) (Bio-Rad, USA)
- Dynabeads® mRNA DIRECT™ Kit (Invitrogen™, Australia)
- High Pure Agarose Gel Purification Kit (Roche)
- High Pure Plasmid Purification Kit (Roche)
- Marathon® cDNA Amplification Kit (Clontech Laboratories Inc, USA)
- Profinia Native IMAC Buffer Kit (Bio-Rad, USA)
- Purelink™ Quick Gel Extraction Kit (Invitrogen ™, Australia)
- Purelink™ Quick Plasmid Miniprep Kit (Invitrogen ™, Australia)

2.1.6 Molecular Weight Markers

- Invitrogen ™ (USA): BenchMark™ protein ladder, SeeBlue® Plus 2 pre-stained protein ladder, 100bp DNA ladder,
- Promega (USA): 1kb DNA ladder, RNA Marker
- Thermo Fisher Scientific Inc (Fermentas)(Australia): PageRuler ™ Prestained Protein Ladder

2.1.7 Antibodies

2.1.7.1 Primary Antibodies

- Mouse anti-penta His antibodies (QIAGEN, USA)
- Mouse anti-recombinant 56 kDa antibodies (Belli *et al.,* 2003)

2.1.7.2 Secondary Antibodies

- Goat anti-mouse IgG (whole molecule), alkaline phosphatase conjugates (Sigma, Australia)
- Goat anti-rabbit IgG (whole molecule), alkaline phosphatase conjugates (Sigma, Australia)
- Goat anti-mouse IgG (H+L), horseradish peroxidase conjugates (Life technologies™, USA)
- Goat anti-rabbit IgG (H+L), horseradish peroxidase conjugates (Life technologies™, USA)
- Alexa Fluor® 488 Goat Anti-Mouse IgG (H+L), highly cross-absorbed (Life technologies™, USA)
- Alexa Fluor® 594 Goat Anti-Mouse IgG (H+L) (Life technologies™, USA)

2.1.8 DNA Cloning and Sequencing Vectors

- pGEM® -T Easy Vector System 1 Kit (Promega, USA)
- Champion™ pET Directional TOPO® Expression Kits (Invitrogen™, Australia)

2.1.9 Bacterial Strains

- *E. coli* strain TOP10 Genotype: F^+, *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80 *lac*ZΔM15 Δ*lac*X74 *deo*R *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G; provided with TOPO TA Coning Kit (Invitrogen™, Australia)
- *E. coli* strain BL21 Star™ (DE3) Genotype: F^+ *omp*T *hsdS*$_B$ (r_B-m_B-) *gal dcm rne131* (DE3); provided with Champion™ pET Directional TOPO® Expression Kits (Invitrogen™, Australia)

2.1.10 Animal Strains

2.1.10.1 Chickens

The Australorp strain of domestic chicken, *Gallus gallus domesticus,* was used throughout this study in accordance with UTS ACEC protocol 2008-096. They were supplied as day-old cockerels by Barter & Sons Pty Ltd, Huntingwood, NSW (Australia).

2.1.10.2 Mice

Quackenbush (QS) female mice were used for antibody production. Animals were housed in accordance with UTS ACEC protocol 2008-188 at 21°C with a 12 hr light/dark cycle and free access to food and water.

2.1.11 Parasite Strains

The Houghton strain of *E. tenella* was used throughout this study. Sporulated oocysts of *E. tenella* were kindly provided by Prof. Martin Shirley, from the Institute for Animal Health, Compton Laboratory (United Kingdom) as a cryopreserved sample in 1998, and passaged regularly at UTS.

2.1.12 Computer Software

- Graphpad Prism® Version 5 (GraphPad Software, Inc., USA)
- ImageJ Version 1.47b available here http://rsbweb.nih.gov/ij/index.html.
- Lasergene™10, Software Suite for Sequence Analysis (DNASTAR® Inc., USA)
- Quantity One® 1-D Analysis Software (BioRad)
- ZEN lite Zeiss (Carl Zeiss, Australia)

2.1.12.1 Online Computer Software

- BLAST: http://www.ncbi.nlm.gov/blast/Blast.cgi
- ClustalW (EMBL-EBI): (http://www.ebi.ac.uk/Tools/msa/clustalw2/)
- EMBOSS (version 6.3.1): (http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::antigenic)
- EpiC: http://epic.embl.de/
- InterProScan, Release 38.0 (EMBL-EBI): http://www.ebi.ac.uk/Tools/pfa/iprscan/
- Jalview (v2)

- ProtScale (ExPASY): http://web.expasy.org/protscale/
- ScanProsite tool (release 20.83) (ExPASY): http://prosite.expasy.org/
- TMpred (EMBnet): http://www.ch.embnet.org/software/TMPRED_form.html)

2.2 Methods

2.2.1 Parasite purification

2.2.1.1 Chickens

All animal research was performed following the UTS ACEC protocols 2008-096 and 2008-188. Australorp cockerals (4-5 weeks old) were infected with *E. tenella* via oral inoculation with the appropriate dose of sporulated oocysts of *E. tenella* in a final volume of 200µl PBS, pH 7.4. The birds were housed in the Ernst Facility at UTS, at 21°C, on a 12 hour light/dark cycle with free access to food and water. Parasite samples were collected at different time points post infection (p.i.) to coincide with specific lifecycle stages of the parasite. Immediately before parasite isolation, the birds were euthanased by CO_2 inhalation and cervical dislocation.

2.2.1.2 Merozoites

2.2.1.2.1 DE-52 Column preparation

DE-52 (Whatman) on cellulose matrix in 50ml PBS, was incubated for 1h at room temperature, allowing it to settle as described previously (Shirley, 1995). The supernatant was gently removed and the wash repeated a further three times. The DE-52 was resuspended in PBS and adjusted to pH8.0 using 5% (w/v) solution of ortho-phosphoric acid. The resin was allowed to settle overnight at 4°C. The column was set-up on the day of merozoite purification on a retort stand as follows: glass wool was pressed into the nozzle of 50ml syringe and polyester wool added to a height of 2cm; the wool was then wet with PBS; and equilibrated DE-52 was added to a height of 4cm. Buffer was then allowed to drain, and extra DE-52 added to maintain the level. The column was washed with PBS supplemented with 1% glucose (PBS/1% glucose) and kept at 4°C until parasite purification.

2.2.1.2.2 Merozoite purification

Merozoites were purified from *E. tenella*- infected chickens as described previously (Shirley, 1995).

Chickens aged 4.5 weeks were infected orally with 2 x 10^5 sporulated *E. tenella* oocysts per bird. The birds were euthanased at 112hrs p.i. to purify mature merozoites from the chicken caeca. The caecum was removed immediately, slit open longitudinally and washed twice with ice-cold PBS. Caecal tissue was cut into 1cm^2 pieces and incubated at 41°C in 100ml of excystation media. The release of merozoites was monitored by microscopic examination and ceased within 30min

incubation. The sample was filtered through two layers of cheesecloth, centrifuged at 800 x g for 10min at 4°C to pellet parasites and the excystation media replaced with PBS/1% glucose. This wash was repeated a further two times. The merozoite pellet was then resuspended in 50ml of PBS/1% glucose and kept on ice. The suspension was then applied to the top of a pre-equilibrated DE-52 column (prepared as described in section 2.2.1.2.1), allowing the merozoites to migrate through the column by gravity flow (Shirley, 1995). A further 10 ml of PBS/1% glucose was applied to the column to ensure all the parasites washed through.

Six separate 10ml fractions were collected (50ml merozoite resuspension plus the 10ml PBS/1% glucose wash). Merozoite yield was determined for each fraction under bright light microscopy and the level of host debris contamination was assessed. Fractions that appeared to have little contamination and high yield were pooled together and washed in PBS three times by centrifugation at 800 x g for 10min at 4°C. The final yield of merozoites was then calculated, and the pellet stored at -80°C.

2.2.1.3 Gametocyte purification

The method developed for *E. tenella* gametocyte purification was adapted from that described by Pugatsch *et al.* (Pugatsch *et al.*, 1989). Chickens aged 4-5 weeks were infected orally with 1 x 10^4 sporulated oocysts per bird. The birds were euthanased at 134hrs p.i. or 144hrs p.i., and the caeca immediately removed. For future immunolocalisation studies, some caecum samples were dissected into smaller pieces and stored in 2% paraformaldehyde in 0.1M phosphate buffer, pH7.4 at room temperature (at least overnight) until sectioning. Each remaining caecum was slit open longitudinally and the lining and mucosa scraped into a beaker with 100ml of SAC medium (pre-warmed to 34°C) containing 1mM PMSF and 0.5g/ml hyaluronidase (both added just prior to use), and incubated at 34°C for 20min. The treated lining and mucosal scrapings were then added to a food processor and homogenised by performing 3 x 10 sec pulses, each separated by a 5 sec pause. The tissue homogenate was then filtered through a fine-pore metal colander, two layers of cheese cloth and a wide-pore metal colander into a 2L beaker. This process was repeated if a high fat content was observed. The filtrate was then passed through a 17µm polymon mesh into another 2L beaker, using SAC medium to wash the filtrate through and maximise the flow-through of gametocytes. The filtrate was then filtered through a 10µm polymon mesh, retaining the gametocytes on the mesh. The volume of liquid on the mesh was monitored and, when it had decreased to approximately 30ml, it was transferred to a fresh 50ml tube and centrifuged at 5,000 x g. The gametocyte pellet was then

resuspended in a minimal volume of PBS (5ml). Bright field microscopy was used to enumerate the number of gametocytes in the sample. The samples were stored in aliquots of 10^6 at -80°C.

2.2.1.4 Oocyst purification

The methods used for *E. tenella* oocyst purification were adapted from those described by Shirley (1995). Chickens aged 4-5 weeks were infected orally with 2.5 x 10^3 sporulated oocysts each. At 168h p.i. the chickens were euthanased and the caecum from each chicken immediately removed. Each caecum was slit open longitudinally and the lining and mucosa scraped into a beaker with 100ml of cold PBS. The sample was homogenised in a food processor for 30 sec at full power and then transferred to a flask. Trypsin was added to 1.5% (w/v) to degrade connective tissue and the homogenate was incubated at 41°C in a water bath for 30min with constant mixing. The homogenate was then filtered through two layers of cheesecloth into a beaker to remove large particles of host contaminants such as fat and cellular debris, followed by centrifugation at 12,000 x g for 12min at 4°C. The supernatant was discarded and the pellet resuspended in H_2O and centrifuged at 12,000 x g for 12min at 4°C. This step was repeated a further three times. Unsporulated oocysts were then stored at -80°C. If oocysts were to be sporulated, oocysts were resuspended in 2% potassium dichromate ($K_2Cr_2O_7$) at a final concentration of 2.5 x 10^5/ml and aerated using a piece of plastic tubing attached to an air sparger placed directly into the sample. The oocysts were left aerating at room temperature (25°C) for 3 days. The level of sporulation was assessed daily using a haemocytometer under bright field microscopy. Aeration was stopped when sporulation was greater than 90% (or had ceased) and the oocysts were stored at 4°C in 2% $K_2Cr_2O_7$.

Oocyst samples were bleached with sodium hypochlorite to eliminate contaminants from faeces or caeca, as described by Shirley (1995). Oocysts were transferred into 50ml tubes and centrifuged at 13,000 x g for 10min at room temperature to wash away the $K_2Cr_2O_7$. The pellets were then resuspended in H_2O, with this wash being repeated a further two times until the washed sample was colourless. Each pellet was resuspended in 45ml of H_2O, 5ml of 25% sodium hypochlorite was added (for a final concentration of 2.5%) and the sample left on ice for 10min. The oocyst samples were then centrifuged at 13,000 x g for 10min at 4°C, the pellet was resuspended in saturated NaCl, and then centrifuged at 5,000 x g for 10min at 4°C to float the oocysts out of the sample. The supernatant containing the oocysts was then decanted through a 10µm polymon mesh, retaining the oocysts on the mesh. The oocysts were washed with 2L of H_2O to remove any excess sodium hypochlorite and NaCl. The washed oocysts were collected from the mesh and centrifuged in 10ml tubes at 12,000 x g for 10min at 4°C. The pellets were resuspended in a small volume of H_2O and the

concentration of oocysts determined under bright light microscopy, using a haemocytometer. Sporulated and unsporulated oocysts required for RNA and protein extraction were centrifuged at 3,000 x g for 10min at 4°C and the pellets stored at -80°C in aliquots of 1 x 10^6 oocysts. Bleached sporulated oocysts required for re-infection were stored at 4°C at a concentration of 2.5 x 10^5 oocysts/ml in 2% potassium dichromate.

2.2.1.5 Sporozoite purification

Sporulated oocysts were pelleted by centrifugation at 540 x g for 8min and resuspended in PBS (pH7.6) at approximately 1 x 10^6/ml. One volume of glass beads was then added and the sample vortexed for 1min. This was followed by 1min incubation on ice. This process was repeated for five 1min pulses and the disruption of the oocyst wall was confirmed by bright field microscopy.

The sporocyst/oocyst/debris suspension was centrifuged at 540 x g for 8 min and the pellet resuspended in 20ml PBS (pH7.6). This wash was repeated to prevent any adverse effects arising from increased acidity produced by the cracked oocysts and to remove some oocyst wall debris. The sporocyst suspension was then centrifuged at 540 x g for 8 min and the pellet resuspended in a suitable volume of excystation solution to produce a final concentration within the range 2 x 10^5 to 2 x 10^6/ml. Sporocysts were then incubated for approximately 1 hour at 41°C, with intermittent mixing and checked for excystation progress every 10-15 min microscopically.

A purification column was prepared with 0.2g nylon wool placed in a 50ml syringe barrel to a depth of 15-20ml. The nylon wool was washed through with sufficient PBS to clean it and remove any bubbles. Following excystation, the sporozoite suspension was centrifuged at 420 x g for 8 min and the excystation solution removed. The discarded excystation solution was checked microscopically for the presence of sporozoites. The pellet was then resuspended in 15ml PBS and passed through the syringe and nylon wool. Sporocysts and wall debris were trapped in the wool whilst sporozoites travelled through the wool and were purified. A further 50ml of PBS was used to wash the sporozoites through the nylon wool column. Following this, the sporozoite solution was passed through a 5µm filter to remove any remaining sporocysts. Following sporozoite purification the solution was centrifuged at 540 x g for 5 min and the pellet resuspended in 1ml PBS. The numbers of sporozoites purified were then enumerated and immediately used for transfections as described in Section 2.2.7.

2.2.1.6 Uninfected chicken caeca

Uninfected chicken caeca were collected as negative controls in RT-PCR, qRT-PCR and immunoblots. Chickens were euthanased and the caeca removed. Each caecum was slit open longitudinally and the caecal lining scraped into a sterile tube. One hundred milligrams was snap frozen in liquid nitrogen and then homogenised using a mortar and pestle. Once the sample was ground to a powdered state, it was suspended in 1ml Trizol® Reagent and kept on ice until RNA extraction.

2.2.2 RNA

2.2.2.1 Total RNA extraction

Purified merozoite and gametocyte pellets were resuspended in 1ml Trizol® Reagent in preparation for RNA extraction (1 x 10^7 merozoites or 1 x 10^6 gametocytes) and homogenised by pipetting the sample until all visible clumps were resuspended. Oocyst samples were washed in PBS and centrifuged at 1500 x g for 10min at 4°C, a process that was repeated another three times to remove residual $K_2Cr_2O_7$. Oocysts were resuspended at 5 x 10^5 sporulated oocysts and 2 x 10^5 for unsporulated oocysts per 1ml Trizol® Reagent. One volume of glass beads was added and the sample vortexed for 1min. This was followed by 1min incubation on ice. This process was repeated for a total of eight 1min pulses for sporulated oocysts and five 1min pulses for unsporulated oocysts. The disruption of the oocyst wall was confirmed by bright field microscopy. The samples were then kept on ice until the RNA was extracted.

Following homogenisation, samples were incubated at room temperature for 10min. Chloroform (200µl: 1/5 volume of total homogenate) was added and the samples vortexed for 15s. The samples were incubated for another 3min at room temperature, then centrifuged at 10,000 x g for 10min at 4°C. The aqueous phase containing the RNA was transferred to a clean 1.5ml tube, an equal volume (approximately 50µl) of ice-cold isopropanol added to precipitate RNA and the sample gently mixed by inverting the tube several times. The samples were incubated at room temperature for 10min and the RNA pelleted by centrifugation at 10,000 x g for 10min at 4°C. The supernatant was removed and the RNA pellet washed with 1ml of 75% ethanol, ensuring not to disturb the pellet. The pellet was centrifuged again at 10,000 x g for 5min at 4°C, the supernatant removed a second time, and centrifuged at 10,000 x g for 5min at 4°C. The ethanol was removed and the pellet left to air dry in the fume hood for 10min. Nuclease-free water (20µl) was used to resuspend the pellet and then it

was placed in a 55°C water bath for 10min. The samples were then stored at - 80°C until cDNA synthesis.

2.2.2.2 Dynabead™ mRNA Extraction

A purified gametocyte pellet (1 x 10^6) was resuspended in 250µl of sterile PBS for mRNA extraction and homogenised by pipetting the sample until all visible clumps were resuspended. Approximately 250µl of silica glass beads were added and the sample vortexed for 3 x 1min with 30s intervals on ice to ensure the gametocytes were not over-heated. The lysate (approximately 250µl) was then removed, placed on ice and the beads washed with a further 250µl of PBS. The 250µl wash was added to the sample on ice to give a total volume of 500µl. An equal volume of lysis/binding buffer (Invitrogen™) was added to the lysate and the sample passaged through a 26-gauge needle to shear any contaminating DNA. Poly A+ RNA was purified from the gametocyte lysate sample using the Dynabead mRNA DIRECT™ kit (Invitrogen™), according to the manufacturer's instructions detailed below.

2.2.2.2.1 Preparation of Dynabead® oligo (dT) magnetic beads

Prior to mRNA extraction 50µl Dynabead® magnetic beads were transferred to an RNase-free 1.5ml microcentrifuge tube and placed on a DynaMag™ magnet. After 30 sec (or when the suspension was clear), the supernatant was removed and the beads washed in 50µl fresh lysis/binding buffer.

2.2.2.2.2 Direct mRNA Isolation

The lysis/binding buffer from pre-washed Dynabead® magnetic beads (Section 2.2.2.2.1) was removed by placing the tube on the DynaMag™ magnet for 30 sec or until the suspension was clear. The microcentrifuge tube was removed from the magnet, the lysis/binding buffer removed and the gametocyte sample lysate added to the beads. The beads were resuspended completely in the gametocyte lysate by pipetting. The sample was then incubated with continued mixing for 3-5min at room temperature to allow poly A+ mRNA to bind to the oligo (dT) on the beads. If the solution was still viscous after 5min, the incubation time was increased. The microcentrifuge tube with the sample was then placed on the magnet for 2 min and the supernatant removed once the solution had cleared. If the solution was noticeably viscous this time was increased to 10min. The

beads/mRNA complex was then washed twice with 600µl Washing buffer A (Invitrogen™) at room temperature, using the magnet to separate the beads from the solution between washes. The beads/mRNA complex was then washed once with Washing buffer B (Invitrogen™) at room temperature, again using the magnet to separate the beads from the solution. The mRNA was then eluted from the beads by adding 10µl of 10mM Tris-HCl and incubating at 65-80°C for 2min. The tube was placed immediately onto the magnet, the supernatant containing the purified mRNA was transferred to a sterile RNase free microcentrifuge tube and placed on ice for immediate use, or stored at -80°C until Marathon® cDNA amplification.

2.2.2.3 Quantification of the concentration of nucleic acids

Quantification of nucleic acid samples, including DNA and RNA, was carried out using the Nanodrop® ND-1000 spectrophotometer (Nanodrop Technologies, Montchanin, DE, USA). A volume of 2µL per sample was analysed against the diluent blank in each experiment. Default settings for either double-stranded DNA (dsDNA) or RNA was used so that when the optical density at 260nm (A_{260}) equalled 1.0, the nucleic acid concentration was 50ng/µl (dsDNA) or 40ng/µl (RNA). In addition to this the $A_{260/280}$ ratio was calculated for an indication of nucleic purity. An $A_{260/280}$ ration of ~1.8 was considered 'pure' for DNA, and ~2.0 for RNA. Ratios with lower values than these were indicative of contaminants in the sample such as protein and phenol.

2.2.2.4 Determination of RNA quality

Gel electrophoresis was used to determine the quality of RNA, i.e. whether it was degraded or not.

Both total and mRNA samples were added to 5x RNA loading buffer and separated on a 1.2% formaldehyde (FA) gel alongside RNA ladder. GelRed™ at an appropriate working concentration was added to the gel prior to casting for visualisation of samples with UV light following electrophoresis. Electrophoresis was performed using the BioRad Wide Mini-Sub® Cell GT electrophoresis unit in 1x FA gel running buffer and gels were run at 80V for one hour. The Gel Red™ stained nucleic acids were visualised using the VersaDoc™ Imaging System (BioRad) and images analysed using the Quantity One® 1-D Analysis Software (BioRad). The level of RNA degradation was assessed by observing the large subunit rRNA (26S) and small subunit rRNA (18S). Due to differences in size of the large subunit ribosomal RNA between *Eimeria* (26S) and *Gallus gallus* (28S) (Schaap *et al.*, 2005)

it was also possible to quantitatively discriminate between parasite and host total RNA levels and determine relative levels of host contamination.

2.2.3 Synthesis of cDNA

2.2.3.1 DNase treatment

Total RNA (1µg) was incubated at 37°C for 30 min with 1U of DNase I, 1 µL 10x DNase I reaction buffer and ddH_2O to 10 µL. The DNase I was inactivated by the addition of 1 µL RQ1 DNase stop solution (Promega) and the reaction heated at 65°C for 10 min.

2.2.3.2 Reverse transcriptase PCR (RT-PCR) using random hexamer primers

Reverse transcriptase of RNA was performed with the Superscript™ III First-Strand Synthesis System (Invitrogen ™ (USA). DNase treated RNA (up to 1µg of total RNA) was mixed with 1µl of random hexamer primers (150ng/µl), 1µl of 10mM dNTP mix (10mM each dATP, dGTP, dCTP and dTTP) and DEPC H_2O (up to 13µl) and the solution incubated at 65°C for 5min followed by incubation on ice for 1min. A master mix of 4µl 5x First-Strand buffer, 1µl 0.1M DTT, 1µl RNaseOUT™ (40U/µl) and 1µl Superscript ™ RT (200U/µl) per sample was prepared and 7µl added to each reaction for a total of 20µl reaction volume. Reactions were incubated at 25°C for 5min followed by 50°C for 60min. The reverse transcriptase enzyme was inactivated by incubating at 70°C for 15min. To remove complementary RNA, 1µl (2U/µl) of RNase H was added to the reaction and incubated at 37°C for 20min. Samples were stored at -30°C until analysis.

2.2.3.3 Marathon cDNA amplification

2.2.3.3.1 First strand synthesis

Marathon cDNA amplification was performed using the Marathon® cDNA amplification kit (Clontech Laboratories Inc). *Eimeria tenella* gametocyte mRNA (1µg in 3µl) was added to 10µM cDNA synthesis primer (CDS) and sterile water in a final volume of 5µl, and the reaction incubated at 70°C for 2min, then cooled on ice for a further 2min. A master mix of 2µl of 5x First-Strand buffer, 1µl 10µM dNTP mix (10mM each dATP, dGTP, dCTP and dTTP at neutral pH) and 1µl AMV reverse transcriptase

(20U/μl) was added and the reaction incubated at 42°C for 1h in an air incubator. Placing the sample on ice then terminated the first strand synthesis reaction.

2.2.3.3.2 Second strand synthesis

A master mix containing 4μl of 20x second strand enzyme cocktail, 1.6μl dNTP mix (10μM), 16μl 5x Second-Strand buffer, 48.4μl sterile H_2O was prepared and 70μl added to 10μl first strand reaction (Section 2.2.3.3.1) and the tube incubated at 16°C for 90min. Two microlitres (3U/μl) of T4 DNA polymerase was then added to the reaction and the tube incubated for a further 45min. To terminate the second strand synthesis reaction, 4μl of EDTA/glycogen mix was added and the reaction transferred to a 1.5ml eppendorf tube. Phenol:chloroform:isoamyl alcohol (100μl of 25:24:1) was added and the sample vortexed thoroughly. The sample was then spun in an Eppendorf 5415D microcentrifuge at 14,000 rpm for 10min to separate the phases. The top aqueous phase was removed and placed in a clean 0.5ml tube, discarding the interface and lower phase. Chloroform:isoamyl alcohol (100μl of 24:1) was added to the aqueous phase and the sample vortexed. The tube was centrifuged again at 14,000 rpm for 10min. Following this, the top aqueous layer was again removed and placed in a clean 0.5ml tube. To this, one-half volume of 4M ammonium acetate was added, followed by 2.5 volume of room temperature 95% ethanol. The sample was vortexed thoroughly and the tube centrifuged immediately at 14,000 rpm at room temperature for 20min. The supernatant was then gently removed, 80% ethanol added and the sample re-centrifuged at 14,000 rpm for 10min. Following this, the supernatant was removed and the pellet air dried to remove residual ethanol. The pellet was then dissolved in 10μl sterile H_2O and stored at -20°C.

2.2.3.3.3 Adaptor ligation

Double stranded cDNA (5μl; Sections 2.2.3.3.1 and 2.2.3.3.2) was added to 2μl Marathon cDNA Adaptor (1μM), 2μl 5x DNA ligation buffer and 1μl T4 DNA ligase (400U/μl), and incubated overnight at 16°C. T4 DNA ligase was inactivated the following day by incubation at 70°C for 5min. The adaptor-ligated ds cDNA was then diluted 1:50 and used in RACE PCR.

2.2.4 Polymerase Chain Reaction

2.2.4.1 Oligonucleotide design

Sequence data available in the genome database was used to design primers that would amplify a unique part of the gene sequence. Primer pairs were designed to be approximately 18-24 mers with an optimal Tm of 55°C-72°C. The optimal product size was 300 – 800bp, with a GC content of 45-55%. Poly G, C, A, T, and poly T/C or A/G were avoided, Gs or Cs were included at the 3` end to anchor the primer. In addition primer sequences were designed to contain no inter- or intra-molecular homology with the primer. BLASTn nucleotide-nucleotide on NCBI was used to assess the complementation of each primer to other *E. tenella* genes, and *G. gallus* genes to ensure there was no redundancy. Each oligonucleotide primer Tm was calculated using the 'Oligocalculator' tool on the Sigma Genosys (Australia) website (http://www.sigma-aldrich.com/sigmagenosys-australia).

2.2.4.2 Semi-quantitative PCR

For the amplification of gene-specific products, each reaction contained 1µl DNA template (50ng), 1µl of 2mM dNTPs, 1µl of 25mM $MgCl_2$, 1µl of 10 x Taq DNA polymerase buffer, 0.2µl each of 10µM sense and antisense primers, 0.04µl of Taq DNA polymerase (5U/µl) and sterile H_2O to a final volume of 10µl per reaction. Reactions were cycled in a MJ Research PTC 200 thermal cycler with a heated lid using the following conditions: initial denaturation at 95°C for 5min; denaturation at 95°C for 30s; primer annealing of specific primer for 1min; extension at 72°C for 1min per kb of PCR product; 30-35 repetitions of the entire cycle from the second denaturation step; and an additional extension at 72°C for 5min. PCR products were visualised by agarose gel electrophoresis (Section 2.2.4.4).

2.2.4.3 Real-time PCR

2.2.4.3.1 Real-time PCR reaction conditions

PCR master mixes were optimised for 10µl reaction volumes using the QuantiFast SYBR Green PCR Kit (Qiagen, USA). A master mix containing 5µl of 2x SYBR Green RT-PCR master mix (HotstarTaq® *Plus* DNA polymerase, Quantifast SYBR green PCR buffer, dNTP mix and ROX passive reference dye), 0.5µl each of 10µM sense and antisense primer and 3.9µl of ddH_2O was added to 0.1µl of cDNA. Reactions were assayed in the Qiagen Rotor-Q Real Time PCR system (Qiagen, USA) using the

following hot-start PCR conditions: initial activation of reaction 95°C for 5min; denaturing at 95°C for 10sec; combined annealing and extension at 60°C for 30sec for 40 cycles. Data was collected at the 60°C extension step and a dissociation step was performed at the end of the PCR using default parameters (ramp from 55°C to 95°C) for melt curve analysis of the products. Threshold cycles (C_T) were calculated from baseline-corrected normalised fluorescence data. The threshold fluorescence level was set to an arbitrary value (0.03) for all assays so that the amplification curves were in log-linear phase at that level.

2.2.4.3.2 Primer efficiency analysis

Four-fold serial dilutions were prepared from *E. tenella* cDNA so there were six-point standards of 1/1, 1/4, 1/16, 1/64, 1/256, and 1/1024 and a no template negative control. Each six-point standard curve and no template negative control was run as described in Section 2.2.4.3.1. For comparison, each cycle run also contained a six-point standard curve of the *E. tenella* reference gene, 18S ribosomal RNA. The individual reaction efficiencies (E) were calculated as described by Pfaffl (Pfaffl, 2001). Mean C_T values were regressed against the log of the template dilutions and the amplification efficiency calculated from the slope of the standard curve using the formula:

$$\mathbf{E} = 10^{-1/\text{slope}}$$

where an ideal efficiency is 2, and the optimal slope of the standard curve is -3.32.

Amplification efficiency was then converted into a percentage, following the formula:

$$\textbf{\% Efficiency} = (E-1) \times 100\%$$

To determine the relative transcript abundance of target genes in the different life cycle stages of *E. tenella* the level of transcript expression of the target gene, measured as C_T, was normalised to the transcript expression of the reference *E. tenella* ribosomal 18s RNA gene C_T. This calculation was made as follows:

$$\text{Relative quantification } \mathbf{R} = 2^{-(\text{CT of target gene} - \text{CT of reference gene})}$$

as described in Pfaff (Pfaffl, 2001) and Livak and Schmittgen (Livak and Schmittgen, 2001) and where the reference gene is *E. tenella* ribosomal 18S RNA. This ratio was then compared between the different life cycle stages to determine levels of transcript expression. A no template negative

control was run alongside each qRT-PCR experiment to ensure no contaminants were present. Each target gene was investigated three times in each life cycle stage in at least two different experiments to calculate mean and standard deviation from n= 6. Two-way analysis of variance (ANOVA) and a Bonferroni post-test calculated statistical significance between transcript levels in life cycle stages of *E. tenella.*

2.2.4.4 Random amplification of cDNA ends (RACE) PCR

For the amplification of RACE PCR products, a master mix containing 5µl of 10x Advantage 2 PCR Mix buffer, 1µl of 50x Advantage® *Taq*, 5µl of 2mM dNTPs, 1µl each of 10µM gene specific primer and Marathon Adaptor primer 1 and ddH_2O to 50 µL was added to 1 µl cDNA. Reactions were cycled in a MJ Research PTC 200 thermal cycler with a heated lid using the following conditions for touchdown PCR: 94°C for 1min; 94°C for 30s, 65°C for 30s, 72°C for 3min for 5 cycles; 94°C for 30s, 60°C for 30s, 72°C for 3min for 5 cycles; and 94°C for 30s, 55°C for 30s, 72°C for 3min for 35 cycles. PCR products were visualised by agarose gel electrophoresis (Section 2.2.4.4).

2.2.4.5 Nucleic acid electrophoresis

DNA samples were analysed by running samples on agarose gels made with TBE electrophoresis buffer in percentages ranging from 0.8% to 1.5%. Gels were pre-stained with Gel Red™ (10,000x) at an appropriate working concentration. Nucleic acid samples and molecular markers were diluted in 6x loading buffer and electrophoresed using a Bio-Rad Mini-Sub® Cell GT electrophoresis unit in TBE buffer at 80V-100V for 60min. The Gel Red™ stained nucleic acids were visualised using the VersaDoc™ Imaging System (BioRad) and images analysed using the Quantity One® 1-D Analysis Software (BioRad).

2.2.5 Gene cloning

2.2.5.1 Restriction digests

Single and double digestion reactions were made for both insert and vector preparations. Digestions were performed using ten units of restriction enzyme for every microgram of DNA. Standard reaction mixtures contained 1µg of DNA, 10 units of restriction enzyme, 1µl-2µl enzyme-specific

buffer (10x), 1µl BSA (10x) and ddH_2O to a final volume of 10µl for single restriction enzyme digestion reactions or to a final volume of 20µl for double restriction enzyme digestion reactions. Samples were incubated in a 37°C water bath for 2h and then heated at 65°C for 10min to deactivate enzymes. Digested DNA products were visualised by agarose gel electrophoresis and extracted from the gel as described below (Section 2.2.4.4).

2.2.5.2 DNA gel extraction

DNA samples electrophoresed on agarose gels were excised and purified prior to cloning into a vector. DNA bands were visualised using UV light and bands of interest were cut from the gel with a sterile scalpel and purified using the High Pure Agarose Gel Purification Kit (Roche), according to the manufacturer's instructions. The concentration of the purified DNA was determined using the NanoDrop® ND-1000 Spectrophotometer (Nanodrop Technologies, Inc) and the final elution was stored at –20°C.

2.2.5.3 Dephosphorylation treatment of pActmCherry

Antarctic alkaline phosphatase was used to catalyse the removal of 5` phosphate groups from the ends of the pActmCherry vector (provided by Tsai Lab). Standard reaction mixtures contained 1µg of pActmCherry vector, 1µl of 5U/µl Antarctic alkaline phosphatase, 1µl of 10x Antarctic phosphatase reaction buffer and sterile H_2O to a total volume of 10µl. The reaction was then incubated at 37°C for 1h, and heat inactivated at 65°C for 5min. The dephosphorylated pActmCherry vector was stored at -20°C, without ethanol precipitation, for later use in ligations.

2.2.5.4 Ligation reactions

Purified PCR product (Section 2.2.5.2) was cloned into the sequencing vector pGEM® -T easy, using the pGEM®-T easy Vector System 1 Kit (Promega) according to the manufacturer's instructions. Briefly, 5µl of purified DNA (20-45ng) was added to 1µl pGEM® -T Easy vector (50ng), 5µl 2x Rapid Ligation Buffer T4 DNA ligase and 1µl T4 DNA Ligase (3 Weiss units/µl) to a total volume of 10µl. Reactions were incubated at room temperature for 1h or, alternatively, overnight at 4°C.

For the ligation of the *Et*Sub5 insert with the pActmCherry vector (provided initially by the Tsien Lab, as described in Shaner *et al.* (Shaner *et al.*, 2004), Mlu I/Sal I-digested *Et*Sub5 insert was added to

Mlu I/Sal I-digested and dephosphorylated pActmCherry vector to obtain a 3:1 vector to insert molar ratio:

$$\text{Amount of insert}_{[ng]} = 3(\text{Vector}_{[ng]} \times \text{Insert length}_{[Kb]}/\text{Vector length}_{[Kb]})$$

T4 DNA ligase (0.5µL) and 1µl of 2X T4 buffer from the pGEM® -T easy Vector System 1 Kit were added with ddH_2O to a final volume of 10µL. The reaction was incubated at 4°C overnight.

2.2.5.5 Transformation of competent cells

Following ligation, transformation of competent cells was performed by adding 10 µL DNA ligation mixture (Section 2.2.5.4) to 90 µL of competent OneShot® TOP10 *E. coli* cells (Invitrogen™, USA), mixed by gentle flicking and incubated on ice for 20 min. Solutions were then heat shocked for 1min at 42°C and returned to ice for 1 min. The cells were allowed to recover by adding 900µl LB media to the transformation mixture and incubating at 37°C in a shaking incubator at 200 rpm for 1h. Appropriate aliquots were spread on LB agar plates (supplemented with 100 µg/ml ampicillin, 0.5mM IPTG and 80µL X-Gal), inverted and incubated overnight at 37°C.

2.2.5.6 Colony PCR

Colony PCR with vector-specific primers was performed to select bacterial clones that were successfully transformed with the recombinant plasmid of interest. Single white colony transformants (representing colonies containing recombinant pGEM T easy plasmids) were picked with a sterile toothpick from LB/ampicillin/IPTG/X-Gal agar plates and streaked onto a new LB/ampicillin agar plate with reference numbers and the plate incubated overnight at 37°C. The following day, the numbered colonies were picked and used to inoculate a 10µl PCR reaction containing 1µl of 10x reaction buffer, 1 µl of 25mM $MgCl_2$, 1 µl of 2mM dNTPs, 0.04µl of Platinum® *Taq* DNA polymerase (Invitrogen™), sense and antisense primers (0.2µL of 10µM each) and ddH_2O to 10µL. Reactions were cycled in a MJ thermal cycler with a heated lid using the following conditions: 95°C 3min for one cycle; then 95°C for 30s, 55°C for 1min, 72°C 3min for 30 cycles. PCR products were visualised by agarose gel electrophoresis (Section 2.2.4.4).

2.2.5.7 Plasmid preparations

Clones containing the recombinant plasmid of interest were picked from the reference LB/ampicillin agar plate (Section 2.2.5.6) and used to inoculate 5ml LB media with ampicillin (100 μg/ml). The cultures were incubated overnight at 37°C in a shaking incubator (200 rpm). The following day, a 20% glycerol stock of each clone was prepared by adding 200μl filter-sterilised glycerol to 800μl of overnight culture. The glycerol stocks were stored at -80°C for long-term storage. The remaining 4.2ml of overnight culture was kept on ice for immediate plasmid purification.

2.2.5.8 Plasmid purification

Plasmid purifications were performed using the High Pure Plasmid Purification Kit (ROCHE) according to the manufacturer's instructions. Briefly, 4.2 ml of overnight culture was centrifuged at 9000rpm for 1min to pellet bacterial cells. The remaining culture was removed and the pellet allowed to dry for 5-10min. The bacterial cell pellet was then resuspended in 250μl Suspension buffer/RNase. The bacterial cells were lysed by adding 250μl of Lysis buffer and incubated at room temperature for 5min. The sample was then placed on ice and 350μl cold Binding buffer added, mixed gently, incubated on ice for 5 min then centrifuged in a bench-top Eppendorf 5415D microcentrifuge (Eppendorf) at 13,000 rpm for 10min. The released plasmid DNA was bound to the High Pure spin column by spinning in a bench-top microcentrifuge at 13,000 rpm for 1min. The column was washed twice with Wash buffer then centrifuged at 13,000 rpm for 1min to remove any excess ethanol. Plasmid DNA was then eluted in 50μl nuclease-free H_2O. The concentration of purified DNA was determined using the NanoDrop® ND-1000 Spectrophotometer and the sample stored at -20°C.

2.2.5.9 DNA sequencing

DNA sequencing was carried out at Macrogen Inc (Seoul, Korea) or at the Australian Genome Research Facility (AGRF, Brisbane). Approximately 1μg of double-stranded plasmid DNA was sequenced with 3.2pmol of appropriate sense and antisense primer.

2.2.6 Bioinformatics

2.2.6.1 DNA sequencing analysis

The analysis of sequence data of recombinant clones was performed with Seqman program (Lasergene™ 10 software suite). Contig assemblies were generated automatically from multiple sequences of the same recombinant plasmid. Individual sequences were manually trimmed of vector sequence and poor trace data. Consensus sequence was exported into Seqbuilder for further analyses.

2.2.6.2 Sequence annotation

DNA or protein sequences were exported into the Seqbuilder program (Lasergene™ 10 software suite) for annotation and management. Open-reading frames and protein translations of non-annotated DNA sequences were determined using this program. This program was also used to design primer locations for PCR and to manage PCR results for comparison with cDNA and genomic DNA. Furthermore, this program was used to export DNA or protein sequences to other programs, such as Protean (Section 2.2.6.4), for further analyses.

2.2.6.3 Sequence alignments

DNA and protein sequence alignments were performed using ClustalW (Thompson *et al.*, 1994) through the publically available European Bioinformatics Institute (EMBL-EBI) website (http://www.ebi.ac.uk/Tools/msa/clustalw2/) according to default parameters (Chenna *et al.*, 2003). The alignments produced could also be manually altered using the multiple alignments editor, Jalview (v2) (Waterhouse *et al.*, 2009), also available on this website. A report of sequence distances was used to help identify protein homologues between different organisms.

2.2.6.4 Protein sequence analysis

The Protean program (Lasergene™ 10 software suite) was used to perform a number of different analyses of protein sequences, including the calculation of molecular weight and analysis of amino acid composition. Protein sequences were scanned for known patterns or motifs using two publically available online tools: InterProScan, Release 38.0, publically available on the EMBL-EBI website

http://www.ebi.ac.uk/Tools/pfa/iprscan/ according to default settings (Quevillon *et al.*, 2005); and ScanProsite tool (release 20.83)(Sigrist *et al.*, 2002) accessed via ExPASY at http://prosite.expasy.org/ using default parameters. The online tool ProtScale (Gasteiger *et al.*, 2005), also accessed via Expasy at http://web.expasy.org/protscale/, was used to analyse the hydrophobicity/hydrophilicity of protein sequences using the default settings. EMBOSS (version 6.3.1) antigenic tool, available on the European Molecular Biology Open SoftwareSuite (http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::antigenic), was used to analyse antigenicity of protein sequences (Rice *et al.*, 2000). To compare antigenicity and hydrophobicity and construct antigenic/hydrophobic plots of protein sequences, the online program EpiC (also known as Epic Antigenic), available at http://epic.embl.de/, was used (Haslam and Gibson, 2010). TMpred, a tool available online at EMBnet (http://www.ch.embnet.org/software/TMPRED_form.html), was used to predict transmembrane domains in protein sequences.

2.2.6.5 BLAST

Protein and DNA similarity searches of public domain databases were performed using BLAST (Altschul *et al.*, 1997) and were accessed through the National Centre for Biotechnology Information (NCBI) http://www.ncbi.nlm.gov/blast/Blast.cgi, including the *G. gallus* genome. BLAST analyses within organism-specific genomes were performed for a number of Apicomplexa using their respective online BLAST tools as shown below:

Eimeria tenella	http://www.genedb.org/Homepage/Etenella
Eimeria maxima	http://www.genomemalaysia.gov.my/emaxdb/
Toxoplasma gondii	http://toxodb.org/toxo/
Cryptosporidium parvum	http://cryptodb.org/cryptodb/
Plasmodium falciparum	http://plasmodb.org/plasmo/

The *Eimeria* Transcript database (Eimeria TDB; http://www.coccidia.icb.usp.br/eimeriatdb/) was also accessed for BLAST analysis.

Homologues of protein sequences were identified in protein databases using BLASTp (protein vs. protein) and tBLASTn (protein vs. translated DNA). Homologues of nucleic acid sequences were identified in DNA/EST databases using BLASTn (DNA vs. DNA) or in protein databases using BLASTx (translated DNA vs. protein).

2.2.6.6 *Toxoplasma gondii* database mining

The *Toxoplasma gondii* genome database, ToxoDB, Toxoplasma Genomics Resource (Version 7.3) http://toxodb.org/toxo/ was mined to identify potential homologues of oxioreductase genes in *E. tenella.* An integrated search was performed using the following parameters on the 'New Search' toolbar: (1) 'Search for Genes'; (2) 'Protein Expression'; (3) 'Mass Spectrometry Evidence'. The 'Experiment/Sample' selected was 'Oocyst proteome' (M4 Type II) (Fritz *et al.*, 2012)which is comprised of proteomic data described as 'Oocyst Walls', 'Sporocysts/Sporozoites Cytosol' and 'Sporocysts/Sporozoites Membrane'; the minimum number of unique peptides was set as one; the minimum number of spectra was set at one and weight given to the search was set at ten (Figure 2.1).

For further analysis of genes present in *T. gondii* at certain time points in the parasite life cycle, a second integrated search was performed. This integrated search was performed using the same parameters on the 'New Search' toolbar as for the first search: (1) 'Search for Genes'; (2) 'Protein Expression'; (3) 'Mass Spectrometry Evidence'. However, the 'Experiment/Sample' selected was all experiments except for 'Oocyst proteome' (M4 Type II) (Fritz *et al.*, 2012). This selected all proteomic data from the tachyzoite proteome of *T. gondii* (Figure 2.2). The proteomic data used for this search was generated from experiments with the calcium dependent tachyzoite phosphoproteome (Nebl *et al.*, 2011); tachyzoite rhoptry proteome (Bradley *et al.*, 2005); tachyzoite conoid proteome (Hu *et al.*, 2007); tachyzoite membrane and cytosolic fractions (Dybas *et al.*, 2008); tachyzoite secretome (Zhou *et al.*, 2005); tachyzoite phosphoproteome from purified parasites or infected host cells (Treeck *et al.*, 2011); subcellular fractions (Moreno et al., unpublished, available online http://toxodb.org/toxo/) and total proteome (Xia *et al.*, 2008) of *T. gondii* tachyzoites. The minimum number of unique peptides was set at one; the minimum number of spectra was set at one and weight given to the search was set at ten (Figure 2.2). The resulting proteins from this search were then subtracted from the first search and a list of proteins of interest compiled.

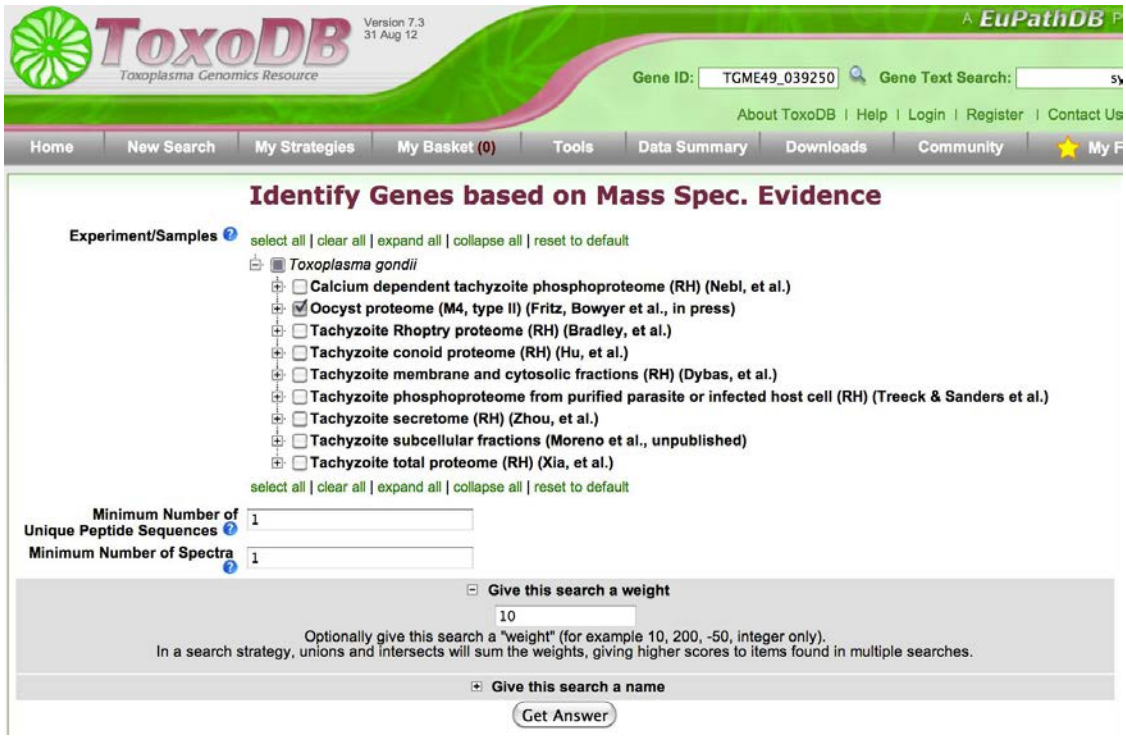
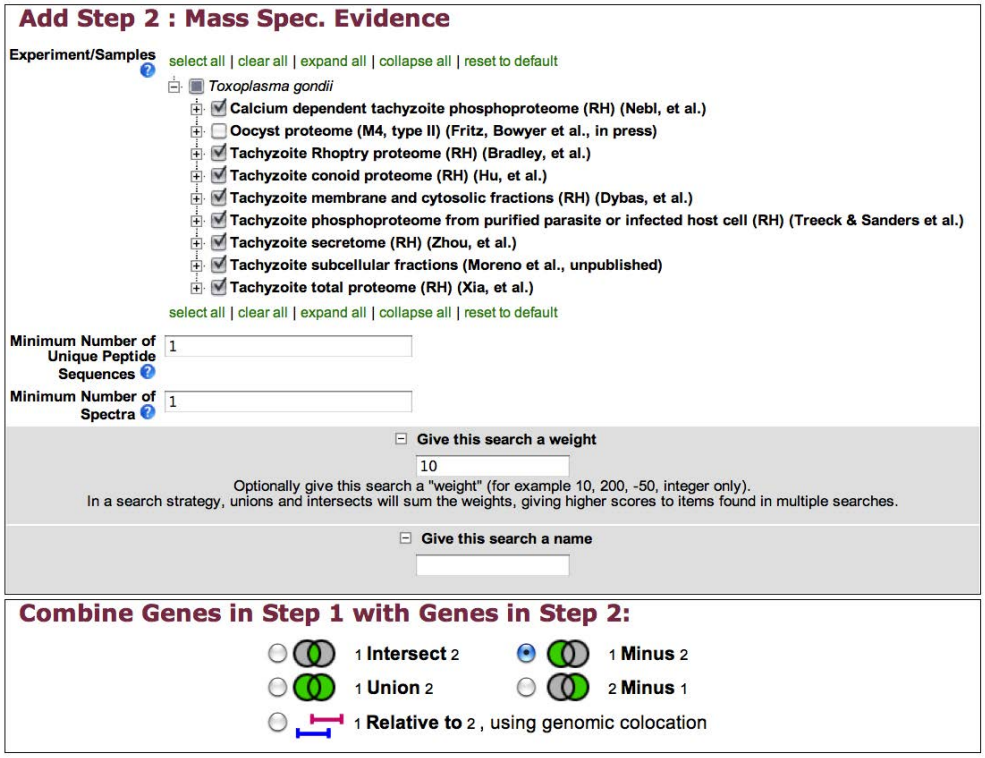


Figure 2.1 Identification of oocyst/sporozoite-specific genes in *T. gondii* based on mass spectrometry results.

Genes coding for proteins detected in oocysts by mass spectrometry were identified in *T. gondii.* Various parameters were set to only identify proteins detected in the oocyst, bleached and untreated, sporozoites and sporocysts in *T. gondii.*

(a)



(b)

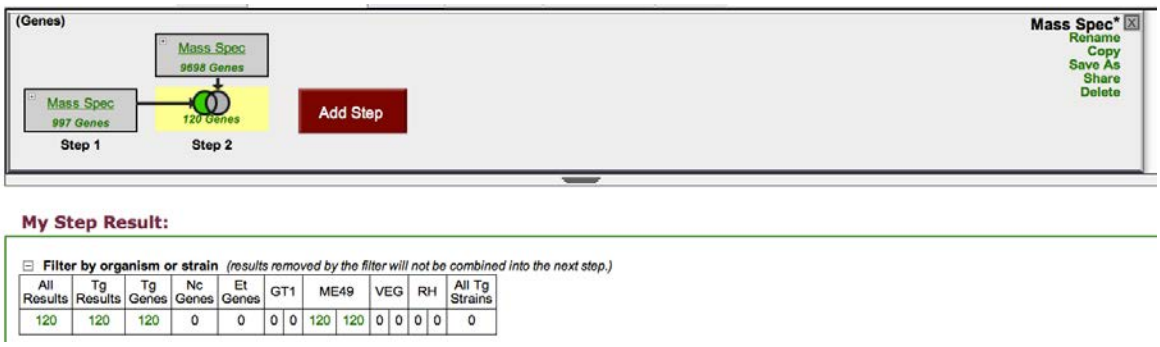
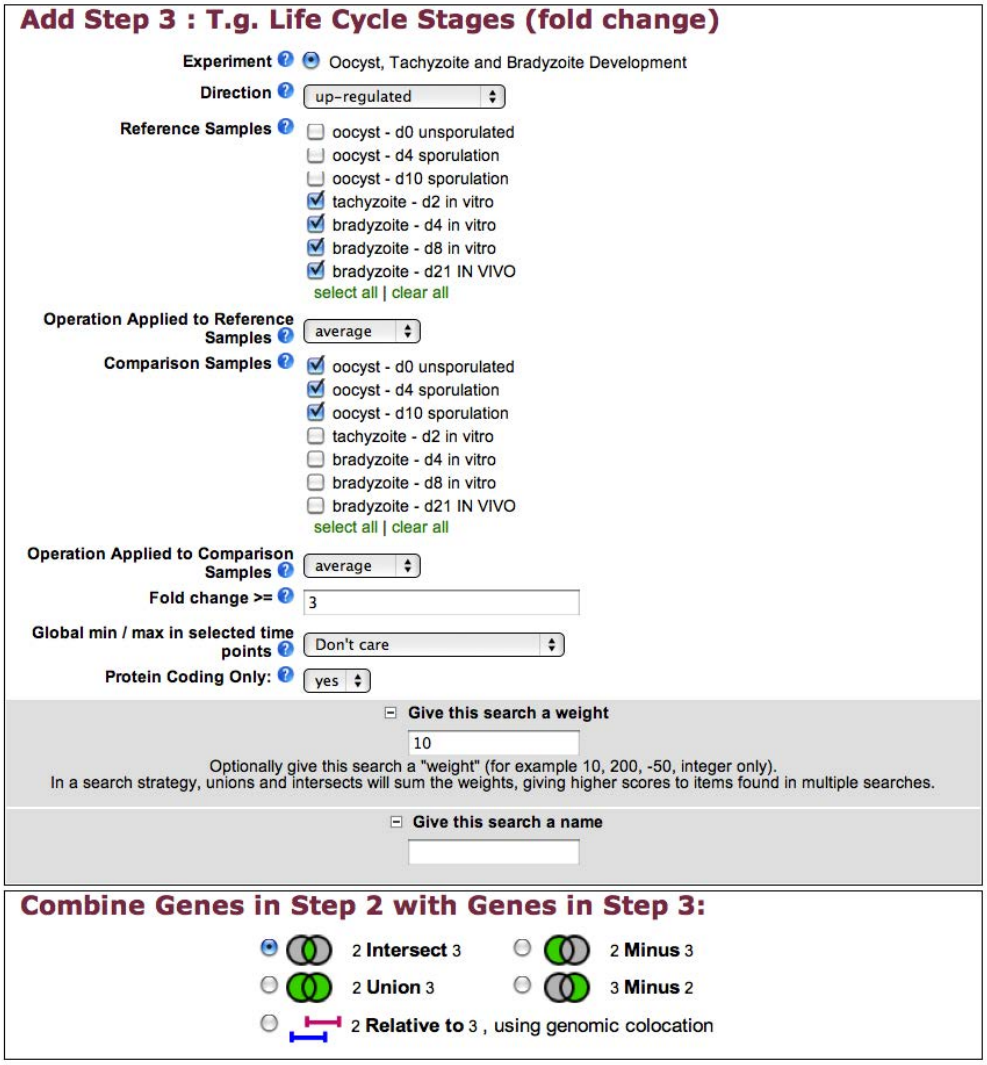


Figure 2.2: Second strategy employed in identification of genes in *T. gondii* based on mass spectrometry results.

Genes coding for proteins detected in tachyzoite extracts were identified in *T. gondii.* Various parameters were set to eliminate any genes identified that were also identified in genes coding for proteins in oocysts bleached and untreated, sporozoites and sporocysts in *T. gondii*.

A new integrated search was then performed to identify further experimental evidence available on the *T. gondii* genome database to 'intersect' with previous results. The new data mined was a list of genes that showed up-regulated transcript levels in *T. gondii* oocysts by microarray. To run this analysis, a 'New Search' was run using the following parameters: (1) 'Genes'; 'Transcript Expression'; 'Mircoarray Evidence'. The search chosen was entitled, 'Expression profiling of *T. gondii* oocyst/tachyzoite/bradyzoite stages (str M4) (Boothroyd/Conrad)' – 'Tg cell cycle (fold change)'. The parameters for the search were then entered as shown in Figure 2.3. Two gene samples were selected for this search – unsporulated fresh (day 0) oocysts, sporulated 4 day old oocysts and sporulated 10 day old oocysts (comparison sample) were compared to tachyzoite (d2 *in vitro*), bradyzoite (d4 *in vitro*), bradyzoite (d8 *in vitro*) and bradyzoite (d21 *in vivo*) (reference sample). The gene direction selected was 'Up-regulated' so that the resulting comparison would select for genes up-regulated in *T. gondii* oocysts with a fold-change in gene transcript levels set at >3.

This resulted in a list of proteins that were then scanned for oxioreductase patterns and motifs with ProSite and InterProScan. BLAST was used to identify homologues of these proteins of interest in *E. tenella.*



(b)

Figure 2.3: Third strategy employed in identification of genes in *T. gondii* based on mircoarray results.

A third bioinformatics step was employed to intersect oocyst/sporozoite-specific proteins/genes found in previous steps (Figure 2.1) with microarray data in *T. gondii.* Parameters were set to search for genes up-regulated in sporulated oocysts of *T. gondii* in comparison with tachyzoites and bradyzoites (a). These results were then cross-referenced with previous results, and genes of interest identified (b).

2.2.8 Protein methods

2.2.8.1 Antibody production

Polyclonal sera and monoclonal antibodies were produced for the specific detection of a number of different *E. tenella* proteins. These antibodies were raised against: a) synthetic peptides representative of the target protein; and b) purified truncated protein of the target proteins. The design and synthesis of these peptides/proteins is described below.

2.2.8.1.1 Synthetic Peptide Design

Synthetic peptides were designed to be 15 amino acids in length from the reference protein sequence of gametocyte-specific subtilases, *Et*SUB1, *Et*SUB2 and *Et*SUB2. An antigenic profile, generated with the Antigenicity Index of the protein sequence using the Jameson-Wolf plot, through the Protean program (Lasergene™ v8.0), was used to predict the antigenic regions of the protein sequence as described in Section 2.2.6.4. Antigenic regions were examined for potential peptide sequences. A ClustalW alignment was also performed so that the peptides were designed in regions that were unique when compared with other proteins within *E. tenella,* to minimise the chance of developing antisera that recognises not only the reference protein but also a potential paralogue in *E. tenella.* BLAST searches were also performed to ensure the peptide designed showed no obvious homology to chicken predicted proteins, thereby reducing the potential for cross-reactivity to the host tissue. Finally, where possible, the peptide contained no more than one cysteine amino acid to aid in the efficient conjugation to Keyhole Limpet Hemocyanin (KLH) (Auspep Pty. Ltd.).

2.2.8.1.2 Synthetic peptide synthesis

Peptides were synthesised and conjugated by Auspep Pty. Ltd. (Parkville, Victoria). Synthesis of un-conjugated peptide was carried out to a purity of >70% and a yield of 10-20mg each. Synthesised peptide (10mg) was then conjugated to the carrier protein KLH, via a cysteine residue, either already within the peptide sequence, or added to the N-terminal of the peptide sequence if it was absent. The efficiency of conjugation was calculated by Auspep and used to determine the yield of conjugated peptide, excluding the mass associated with the KLH conjugate. For example, if 10mg of peptide was conjugated to KLH at a calculated efficiency of 50%, then the total yield of conjugated

peptide was defined as 5mg. This conjugated peptide was then used for immunisation and antibody production by Genscript Corporation (Piscataway, USA). All synthetic peptides, (upon delivery) were made to a 10mg/ml solution in ddH_2O and stored at -80°C.

2.2.8.1.3 Recombinant protein production

2.2.8.1.3.1 Directional Cloning

Protein expression vector-specific primers were designed for the pET101/D-TOPO® fusion vector. Primers were designed so that ligation of one of the gametocyte-specific subtilase genes into the vector would include the C-terminal V5 epitope and 6 x His tag. The reverse primer was designed to preserve the reading frame through the C-terminal end of the sequence. Plasmid of each gene for protein expression was used as a template in PCR with pet101/D-TOPO® specific primers. The reaction contained 1µl of plasmid template (10ng), 0.2µl of 10µM pet101/D-TOPO® specific forward primer, 0.2µl of 10µM pet101/D-TOPO® specific reverse primer, 1µl of 2mM dNTP's, 1µl of 25mM $MgCl_2$, 1µl of 10 x Reaction Buffer, and 0.04µl Taq DNA polymerase (5U/µl) and sterile H_2O to a final reaction volume of 10µl per reaction. The samples were then loaded into the MJ thermal cycler and the reaction was carried out as follows: initial denaturation at 95°C for 3min; denaturation at 95°C for 30s; annealing at 65°C for 1min; extension at 68°C for 1min; and the entire cycle from the second denaturation step was repeated 30-35 times, with a final extension at 68°C for 3min.

The PCR product was then gel purified and cloned into the pET101/D-TOPO® vector. Following colony PCR (Section 2.2.5.6) with vector-specific primers T7 and T7 reverse and plasmid purification (see Section 2.2.5.7), the construct was sequenced (Section 2.2.5.9) with vector specific T7 forward and reverse primers to confirm that the sequence was in frame with the his-tag at the C-terminal end and that there were no mutations generated in the sequence. The construct was then transformed into BL21 Star™ cells for expression.

2.2.8.1.3.2 Protein expression

Escherichia coli BL21 Star™ cells containing the recombinant protein construct pET101/D-TOPO® with the correct DNA sequence were grown in 20ml of LB + 10% amp medium at 37°C with agitation at 200 rpm O/N. The following day, 1L LB + 10% amp medium was inoculated with 10ml (1%) of the O/N culture and incubated at 37°C with shaking at 200 rpm until the cell density OD_{600nm} was 0.5-0.6. Expression of the recombinant protein was then chemically induced by the addition of 1ml of 1M

IPTG to a final concentration of 1mM. The culture was grown for 4h after induction at 37°C with shaking at 200 rpm. The cell pellet was harvested by centrifugation at 4000 rpm for 30min at 4°C, and the pellet stored at -80°C until purification.

Recombinant proteins of oxioreductases, *Et*AO1, *Et*AO2, *Et*Oxio1, *Et*GPX1 and *Et*QO1, were produced by Genscript Corporation (Piscataway, NJ, USA). Recombinant proteins were directionally cloned into E3, E6 or C5 (GenScript USA Inc.), which fused a his-tag at the N-terminal of the protein; a dcTag + his-tag + rTev expression vector; or his-tag + SUMO tag vector expression vectors to fuse a his-tag at the N-terminal of the protein, respectively. Recombinant plasmids were then transformed into *E. coli* BL21 Star™ (DE3) cells or *E. coli* Artic Express™ (DE3) cells and a single colony used to inoculate LB medium containing ampicillin or kanamycin. The culture was incubated at 37°C at 200 rpm until cell density was OD= 0.6-0.8 at 600nm. IPTG was then added for induction as described above.

2.2.8.1.3.3 Protein purification

Recombinant subtilase proteins were purified from induced cultures by affinity purification using the His-tag on the expressed protein. The *E. coli* cell pellet (from 1L culture) was resuspended in 2x Native IMAC lysis buffer (Bio-Rad) containing 20µl protease inhibitor cocktail and 125 µl of 4mg/µl lysozyme was added to a final concentration of 50mg/ml to lyse cells. The suspension was left on ice for 30min followed by sonication on ice using a Misonix Sonicator Ultrasonic Processor S-4000 (Qsonica, LLC, USA) with the following settings: 50W power; 2 x 20s pulses of 10% output; 2 x 20s pulses of 20% output; 2 x 20s pulses of 40% output; with a 30s rest on ice between each pulse. The lysate was centrifuged with the Hettich Rotina 420R (Hettich Instruments, Germany) at 10,000 x g for 30min at 4°C. The supernatant was removed and filtered through a 0.2µm filter membrane. Recombinant protein in the resulting filtrate was purified using the Profinia™ protein purification system and Native IMAC Purification Kit (Bio-Scale Mini Profinity IMAC nickel-chelating column and desalting column) according to manufacturer's instructions. Briefly, the recombinant protein lysate was loaded onto the Profinia™ and washed through the Bio-Scale™ Mini Cartridges (Bio-Rad) prepacked with Profinity™ IMAC (Bio-Rad) to allow the His-tagged recombinant protein to bind to the column. The protein was washed with several Profinia™ purification buffers (Bio-Rad). Automatic UV peak detection is then used to direct the target protein from the affinity cartridge to desalting cartridge prepacked with Bio-Gel® P-6 desalting media. Following purification the purified protein was stored at -20°C.

2.2.8.1.4 Antiserum production

Purified recombinant protein sample *Et*SUB1 was diluted in protein stabiliser (final concentration of 3% lactose/ 1% MSG) in PBS. Samples were then lyophilised overnight using the DYNAVAC freeze drier. Lyophilised protein was then sent to Genscript Corporation (Piscataway, NJ, USA) for immunisations in mice.

Polyclonal antibodies were raised in mice by Genscript Corporation (Piscataway, NJ, USA) using: a) synthetic peptides; b) recombinant protein *Et*SUB1; or c) recombinant proteins of *E. tenella* oxioreductases provided by Genscript Corporation (section 2.2.8.1.3.2). Antibodies were raised in five mice per protein, according to the following immunisation schedule: each mouse was injected i.p., with a dose of antigen at day 0 (1st), 14 (2nd), 35 (3rd) and 56 (4th). Each dose amount was equal to 500μg of purified protein in Genscript T-Max™ adjuvant. A pre-bleed (0.1ml) was taken prior to immunisation at day 0, a test bleed (0.1ml) was taken 10 days after the 2nd immunisation booster shot (on day 24), and the final bleed (0.5ml) was taken 10 days following the 4th immunisation (on day 66). Upon delivery all polyclonal antisera were stored at -20°C.

Recombinant proteins *Et*SUB2 and *Et*SUB5 were excised from a destained Coomassie gel and homogenised in 100μl PBS to a final concentration of 100ng/μl. This was repeated several times to obtain a final concentration of 100ng/μl in a total volume of 1ml. An equal volume of Freund's Complete Adjuvant was added and the mixture emulsified by repeated pipetting through a 1ml pipette. Six QS mice per protein, were immunised by subcutaneous injection with 5μg of protein (100μl). Booster injections were given on days 14, 35 and 56 post immunisation using antigen emulsified in Freund's Incomplete Adjuvant. A pre-bleed (0.05ml) was taken prior to immunisation at day 0, a test bleed (0.05ml) was taken 10 days after the 2nd immunisation booster shot (on day 24), and the final bleed (1ml) was taken 10 days following the 4th immunisation (on day 66). Mice were euthanased as per UTS-ACEC protocol number 2003-11-R-GPC. All polyclonal antisera were stored at -20°C.

2.2.8.1.5 Monoclonal antibody production

Recombinant versions of *Et*Sub1, *Et*Sub2 and *Et*Sub5 proteins were produced as described in Section 2.2.8.1.3.2 and sent to Promab Biotechnologies, Inc (Richmond, California), to be used in the production of monoclonal antibodies.

2.2.8.1.5.1 Monoclonal production

Five Balb/c mice were immunised with purified protein produced by Promab following the standard immunisation protocol described in Section 2.2.8.1.4. Antiserum from each mouse was evaluated by ELISA to determine the level of response to the immunising antigen. Once an acceptable titre was obtained, splenocytes were purified from mice with the highest titres and fused with SP2/0 myeloma cells. Fused cells were then plated onto a 96-well culture plate and screened by ELISA to detect positive clones. Positive clones were sub-cloned by limiting dilution until monoclonal and stable hybridomas. Cells of each subclone were then cryopreserved and stable hybridomas were stored at -20°C. Stable hybridomas were then provided for experimental work.

2.2.8.2 ELISA

Recombinant protein (antigen) was diluted in ELISA buffer 1 to a final concentration of 5ng/µl, and 100µl aliquots (500ng per well) added to a 96-well plate. Plates were incubated overnight at 4°C to allow the antigen to bind. The following day, plates were washed twice with PBS/0.05% TWEEN20® and once with PBS to remove unbound antigen. Plates were then blocked with 200µl per well of PBS/5% SMP for 1h at room temperature with gentle agitation. Plates were then washed as before, and 100µl aliquots of primary antibody diluted in PBS/2.5% SMP added to each well and incubated for 1h at room temperature with gentle agitation. The plates were washed again as above and 100 µL aliquots of secondary antibody (goat-derived, anti-rabbit or anti-mouse IgG, H&L chain-alkaline phosphatase conjugate) diluted at 1:2000 in PBS/2.5% SMP were added to each well. Plates were incubated for a further hour at room temperature, washed again and 200µL of 1 mg/ml p-nitrophenyl phosphate added to each well. The reaction was allowed to develop for 20mins at 37°C, with a maximum of 60min, and absorbance was read at 405nm using the FLUOstar Omega® multi-detection microplate reader (BMG Labtech). All samples were tested in quadruplicate.

2.2.8.3 Protein lysates

Protein lysates were prepared from a number of different developmental stages of *E. tenella.* Merozoites and gametocytes were resuspended directly in 2x SDS-reducing buffer. Unsporulated and sporulated oocysts were mechanically lysed by repeated vortexing with an equal volume of glass beads. Unsporulated oocysts were subjected to 5 x 1min pulses, whilst sporulated oocysts were subjected to 8 x 1min pulses, both with 1min intervals on ice between pulses. The disruption of oocyst and sporocyst walls was confirmed by bright field microscopy. Following disruption, samples

were resuspended in 2x SDS-reducing buffer. Chicken caeca tissue sections (approximately 1cm x 1cm) were lysed using the TissueLyser LT (Qiagen) system. Frozen sections were suspended in PBS and homogenised with 5mm stainless steel balls for 15min. Lysed sample was then resuspended in 2x SDS-reducing buffer.

All samples were boiled at 95°C for 10min, returned briefly to ice, then centrifuged at 10,000 x g for 10min to remove insoluble material. Soluble protein was then transferred to a new tube and protein concentration quantified using the NanoDrop® ND-1000 Spectrophotometer (Section 2.2.8.4).

2.2.8.3 Reduction and Alkylation

Protein lysates were prepared in SDS-reducing buffer as described in Section 2.2.8.3. Dithiothreitol (DTT) was added to each sample to a final concentration of 10mM and incubated at 60°C for 2hrs. Iodoacetamide (200mM final) was then added and each sample incubated for 30min at room temperature in the dark. Additional DTT (4M) was added to each sample to a final concentration of 250mM. Insoluble material was removed by centrifugation at 10,000 x g for 10min, soluble protein was transferred to a new tube and the protein concentration quantified using the NanoDrop® ND-1000 Spectrophotometer (Section 2.2.8.4).

2.2.8.4 Protein quantification

Quantification of protein samples was performed using the NanoDrop™ ND-1000 Spectrophotometer. The A_{280nm} of each sample (2µl) was measured and the concentration of protein calculated using the extension coefficient of BSA (default setting of NanoDrop™ ND-1000 Spectrophotometer software).

2.2.8.5 SDS-PAGE

Electrophoresis of protein samples was performed using NuPAGE Novex 4-12% Bis-Tris gels in an XCell *SureLock* Novex Mini-Cell SDS-PAGE electrophoresis unit, with 1x MES-SDS Running buffer at 150V. Following electrophoresis, gels used for total protein visualisation were stained with Coomassie Brilliant Blue for 1h. Gels were then placed in destaining solution until the background was clear.

2.2.8.6 Western blotting and immunodetection

Following SDS-PAGE, proteins were transferred to PVDF membrane using the XCell II™ Blot Module with the XCell *SureLock*™ unit in transfer buffer at 100mA for 1h. Prestained molecular weight protein markers were included in these gels as both a size marker and indicator of successful protein transfer. Membranes were then blocked in PBS/5% SMP or PBS/3% BSA at room temperature for 1h or overnight at 4°C, with gentle agitation. After blocking, the membrane was incubated in a heat sealed bag with primary antibody diluted in PBS/5% SMP or PBS/3% BSA at room temperature for 1h on a rotating wheel.

Membranes were then washed twice for 10min with PBS/TWEEN20® and once for 10min with PBS and again incubated in a heat sealed bag with secondary antibody diluted 1:2000 in PBS/5% SMP or PBS/3% BSA at room temperature for 1h on a rotating wheel. After probing, the membrane was washed as before and target protein detected using CPS-1 Chemiluminescent Peroxidase Substrate-1 (Sigma) according to the manufacturer's instructions. The membrane was visualised using the VersaDoc™ Imaging System (BioRad) and images analysed using the Quantity One® 1-D Analysis Software (BioRad).

2.2.8.7 Immunofluorescent microscopy (IFM)

Paraffin sections were prepared of *E. tenella*-infected chicken caeca at 112h p.i., 134h p.i. and 144h p.i.. Tissue sections were prepared at the Veterinary Pathology Diagnostic Services (Histopathology, University of Sydney) and were cut to a thickness of 4μm. Sections were stored at room temperature and used for IFM within 3 days of cutting and mounting.

Rehydration of tissue sections was performed by soaking the slides in xylol (xylene) for 3x 10min, 100% ethanol for 2x 3min, 95% ethanol for 3min, 90% ethanol for 3min, 70% ethanol for 3min and citrate buffer (pH6) thrice for 10min. Samples were then boiled for 15min in citrate buffer (pH6) and for 10min in Tris-HCl (pH10) for antigen retrieval. Slides were then cooled to room temperature in Tris-HCl (pH10).

Sections were blocked with 5% goat serum in 0.1M phosphate buffer (pH7.4) for 1h at room temperature. Primary antibodies, from either rabbit or mouse, were then diluted to appropriate concentrations in 5% goat serum in 0.1M phosphate buffer (pH7.4) and 50μl added to each tissue section. Sections were incubated overnight at 4°C in a humidified chamber. The following day, slides

were washed three times for 10min each wash in 0.1M phosphate buffer (pH7.4). Secondary antibodies, either anti-mouse or anti-rabbit conjugated to Alexa Fluor® 488 or 594, were diluted 1:2000 in 0.1M phosphate buffer (pH7.4) and 100µl aliquots added to each section. Sections were then incubated in the dark at room temperature (25°C) for 1h. The slides were washed twice for 10min in 0.1M phosphate buffer (pH7.4) and then counterstained with DAPI (1µg/ml) in 0.1M phosphate buffer (pH7.4) for 5min. The sections were then washed as before and sections mounted with 10 µl of ProLong® Gold Antifade Reagent (Life Technologies™) under a glass cover slip. The cover slip was sealed with nail polish and viewed under a fluorescent microscope.

All fluorescently labelled slides were examined using the Zeiss Axio Imager M1 (Carl Zeiss, Australia). The localisation of proteins indirectly labelled with either Alexa Fluor® 488 or Alexa Fluor® 594 conjugated antibodies was visualised using either a 470-490nm or 520-550nm excitation filter, respectively, while cells stained with DAPI were visualised with a 330-385nm excitation filter. DAPI staining was also used to discriminate between microgametocytes (DAPI-positive) and macrogametocytes (DAPI-negative). The acquisition of images and the overlay of images visualised under different filters was performed using ZEN lite (Carl Zeiss). Images were also analysed with ImageJ Version 1.47b available at http://rsbweb.nih.gov/ij/index.html.

2.2.9 Statistical analysis

One way and two-way analysis of variance (ANOVA) and Bonferroni's multiple comparison post-tests were carried out on ELISA and Real-Time PCR results. All tests were conducted in quadruplicate samples (ELISA) or triplicate, repeated twice (Real-Time PCR) using GraphPad Prism® Version 5 program (GraphPad Software Inc., USA).

CHAPTER 3

IDENTIFICATION OF GAMETOCYTE UP-REGULATED SUBTILISIN-LIKE PROTEASES IN *EIMERIA TENELLA*

3.1 Introduction

Renowned for its robustness, the oocyst wall of the coccidia, protects the parasite from desiccation and mechanical damage, facilitates its survival in the environment and, thus, its transmission. The resilience of the wall is attributed to the presence of di-tyrosine bonds and DOPA (Belli *et al.*, 2003b), where tyrosine residues are cross-linked, forming a stable protein matrix that dehydrates and hardens to form the oocyst wall, with its inherent resistance to chemical and environmental challenges (Belli *et al.*, 2006b). Two proteins, GAM56 and GAM82, are central to this process. They are stored in wall-forming bodies type 2 (WFBII), organelles specific to sexual stage macrogametocytes of *Eimeria,* and are processed into smaller, tyrosine rich proteins prior to cross-linking and incorporation into the oocyst wall (Belli *et al.*, 2004a, Belli *et al.*, 2003a, Belli *et al.*, 2003c). There is some evidence to show that dityrosine crosslinking may be initiated by a peroxidase (Belli *et al.*, 2003b), however, it has yet to be elucidated by what mechanism GAM56 and GAM82 are processed into the smaller tyrosine-rich proteins targeted for cross-linking.

Studies on an analogous system in *C. elegans* encourage speculation that subtilisin-like serine proteases, or subtilases, may play a role. Subtilase serine proteases have been shown to be involved in the processing of proteins in the formation of the cuticle of nematodes (Page and Winter, 2003b). Pro-proteins are proteolytically processed by a subtilase, blisterase, to collagens that are incorporated into the cuticle of the nematode. There they undergo oxidation to form di- and tri-tyrosine crosslinks and result in the protective nematode cuticle (Page and Winter, 2003b). A similar series of events can be readily imagined to occur in *Eimeria* oocyst wall formation and, indeed, serine protease inhibitors, specifically aprotonin, leupeptin and EDTA, inhibited the degradation of native GAM56 *in vitro* in a manner that suggests participation of subtilase-like enzymes in the process (Katrib *et al.,* 2012).

The aim of this chapter was to characterise the six subtilases of *E. tenella* that were identified previously in the *E. tenella* genome database (Katrib *et al*., 2012). Gene expression of these proteases was investigated to determine their stage-specific expression profile in *E. tenella*. Subtilase genes with gametocyte-upregulated or gametocyte-specific expression were sequenced and further characterised. Scanprosite was used to scan these subtilases for characteristic patterns and protein motifs, and BLASTp used to search for homologous sequences in other apicomplexan organisms. The sequences were classified as belonging to the subtilase family of serine proteases based on the presence of a catalytic triad of aspartic acid, histidine and serine active sites typically found in subtilase proteases (Siezen and Leunissen, 1997b, Siezen and Leunissen, 1997a, Siezen *et al.*, 2007a), as well as the side chain asparagine residue which contributes to the oxyanion binding site

3.2 Stage-specific gene expression of subtilisins

A two-fold approach was taken to analyse the gene expression of subtilases in *E. tenella.* Semi-quantitative PCR was used initially to determine in which stages of the life cycle the six subtilases were expressed. These results were then followed up with qRT-PCR to determine transcript abundance of each gene in the life cycle stages.

3.2.1 Semi-quantitative analysis of gene expression

3.2.1.1 Oligonucleotide design

Sequence data available in the *E. tenella* genome database was used to design primers to amplify a unique part of the gene sequence from each of the six subtilases. Each primer pair was designed as described in Section 2.2.4.1. Each primer pair and the expected amplicon size are shown in Table 3.1

Table 3.1: Primer for semi-quantitative real-time PCR for *E. tenella*

Gene	Primer	Sequence 5′→3′	Tm (°C)	Amplicon (Bp)	Optimal Annealing Temperatutre (°C)
etsub1	Sense	TAATTACCTCCATCCCGAACTG	64.32	260	53
	Antisense	CCAGAATCTTCAGCGCCATCAC	68.91		
etsub2	Sense	GGGTCTTCAACATTTGCTGCTGC	70.69	690	65
	Antisense	ATAAGTGCTGCTGCCAACCACC	69.17		
etsub3	Sense	AGGCTGAAGACCTTGAGCATAG	63.58	1361	65
	Antisense	CTACATCTTTGTGGCCCTTTTC	63.44		
etsub4	Sense	CCTTTGTGGCGTGTTCGTGAG	69.94	566	55
	Antisense	CCAGCAGAAGCAGTACCGTGGCC	74.49		
etsub5	Sense	TTGAAGCCGACAGGACGTGG	70.74	360	70
	Antisense	CCAGCCAAAGCAGTCGTCAATA	68.12		
etsub6	Sense	AGCGGCTGCGACTTGAACC	69.89	297	na
	Antisense	CCGTAGCCGCCGTAGGAGTT	69.23		

3.2.1.2 Semi-quantitative gene expression of subtilase genes in *E. tenella*

cDNA synthesised from total RNA of merozoites (112hrs p.i.), gametocytes (134hrs p.i.), unsporulated oocysts (164hrs p.i.) and sporulated oocysts was used for PCR amplification. Genomic DNA was included for each parasite stage in each set of reactions as a control. PCR annealing temperatures were optimised for each primer pair by starting at 5°C below the Tm of the sense and antisense primer pairs and increasing or decreasing temperature until one distinct band was observed.

β-actin specific primers were used to show that equal quantities of parasite cDNA were used in PCRs. Gene-specific primers to the *E. tenella gam56* gene were used as a positive control in the PCR experiments to validate expression or otherwise of the surveyed genes. Products were produced only in gametocyte preparations of *E. tenella,* confirming the presence of gametocyte cDNA, with no contamination from host or merozoites, as well as confirming that lack of product equated to lack of expression of that gene.

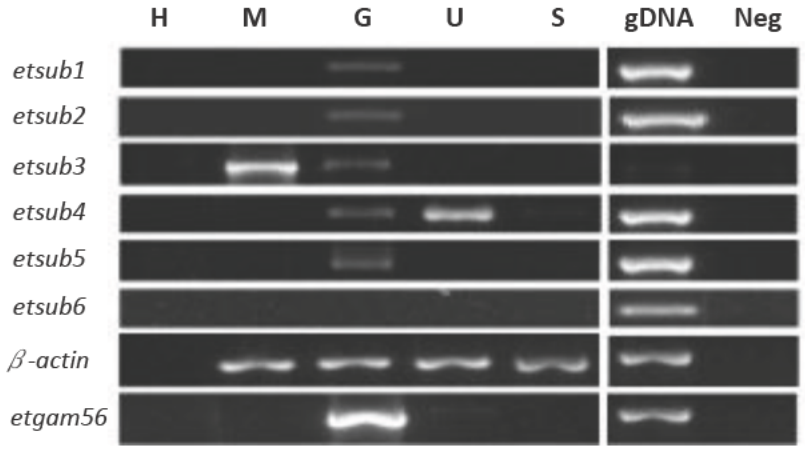
In each semi-quantitative PCR run, there was no visible contamination present in negative, no cDNA controls and, thus, products observed in reactions with *E. tenella* cDNA were unlikely to be due to non-specific binding. No product was observed in reactions with uninfected chicken caeca, eliminating the possibility of amplification of host gene cDNA.

The subtilase gene, *etsub3* produced an expected product size of 1361bp expressed in merozoites and gametocytes (Figure 3.1), with a stronger expression profile in merozoites. A genomic product of 1680bp was also observed (Figure 3.1).

The subtilase gene, *etsub4,* produced a PCR product of 566bp, as predicted, in gametocytes, unsporulated and sporulated oocysts of *E. tenella*. There appeared to be more product amplified from unsporulated oocysts compared with gametocytes and sporulated oocysts. A genomic product of 1456bp was also observed (Figure 3.1).

Despite numerous attempts, using a range of annealing temperatures, *etsub6* was unable to be amplified from any of the life cycle stages (Figure 3.1). However, a product of the expected size of 408bp was observed in the genomic DNA control (Figure 3.1).

Subtilase genes, *etsub1*, *etsub2* and *etsub5* were all found to have gametocyte stage-specific expression in cDNA of *E. tenella.* The gene, *etsub1,* produced a PCR product the expected size of 290bp, while *etsub2* produced a single PCR product of 690bp, as expected (Figure 3.1). The gene, *etsub5* produced a product of 400bp, slightly larger than the predicted 360bp (Figure 3.1).



H: Uninfected host chicken caeca cDNA
M: Merozoite cDNA, 112hrs p.i.
G: Gametocytes cDNA, 134hrs p.i.
U: Unsporulated oocyst cDNA, 164hrs p.i.
S: Sporulated oocyst cDNA
gDNA: *E. tenella* genomic DNA positive control
Neg: No DNA negative control

Figure 3.1: Subtilase gene expression in different developmental stages of *E. tenella*

cDNA from merozoites, gametocytes, unsporulated and sporulated oocysts as well as uninfected chicken caeca was amplified with gene-specific primers under optimised annealing temperatures for each primer pair. Subtilase genes *etsub1*, *etsub2* and *etsub5,* all showed gametocyte-specific expression, whilst *etsub3* was expressed in merozoites as well as gametocytes. The gene, *etsub4,* was expressed in gametocytes, unsporulated oocysts and sporulated oocysts, whilst expression of *etsub6* could not be detected in any developmental stages. A no DNA negative control was included to ensure there was no contamination, and inclusion of host cDNA ensured any product observed was not due to contaminating chicken tissue. Parasite cDNA samples were standardised using *β-actin*-specific primers. The gene, *etgam56,* was amplified from gametocyte cDNA using *etgam56* specific primers as a positive control. PCR products were electrophoresed in a 1% agarose gel and visualised using Gel Red® and a VersaDoc® Imaging system as described in Section 2.2.4.5.

3.2.2 Quantitative RT-PCR analysis

The levels of gene expression appeared to vary when analysed with semi-quantitative RT-PCR, but were unable to be quantified to allow comparison of levels of expression between different life cycle stages using this technique. The next step was to perform quantitative RT-PCR to provide more detail on the levels of gene transcript in each life cycle stage.

The five-subtilase genes shown to be expressed in *E. tenella, etsub1, etsub2, etsub3, etsub4* and *etsub5* were selected for further analysis by qRT-PCR. The gene, *etsub6*, which could not be shown to be expressed in *E. tenella,* was not analysed further.

Relative quantitation of gene abundance has been shown to be suitable for investigating gene expression levels (Pfaffl, 2001). To determine the relative transcript abundance of subtilase genes in the different life cycle stages of *E. tenella,* the level of transcript expression of the target gene was normalised to the transcript expression of a reference gene, *etr18s,* the *E. tenella* ribosomal 18S gene. This ratio was then compared between the different life cycle stages to determine levels of transcript expression. Two different time points of gametocytes were used in qRT-PCR analysis for a more detailed examination of transcript expression.

3.2.2.1 Oligonucleotide design

New primers for each gene as well as the reference gene *et18s* were designed for use with the QuantiFast SYBR Green PCR Kit (Qiagen, USA) to allow amplification of single amplicons with minimal secondary structure or self-dimerisation. Table 3.2 shows the qRT-PCR primers designed from the predicted mRNA sequences for the *E. tenella* genes as well as the expected size of the amplicon. Each pair was designed using Primer3 software (Rozen and Skaletsky, 2000) to ensure efficiency, sensitivity, self-complimentarity, primer length and a melting temperature of 60°C. BLASTn nucleotide-nucleotide on NCBI was used to assess the complementation of each primer to other *E. tenella* genes, and *Gallus gallus* genes to ensure there was no redundancy.

Each primer pair (Table 3.2) was used to amplify a gene product from *E. tenella* cDNA as described in Section 2.2.4.3.1. Resultant products were cloned into pGEM T-Easy vector (Section 2.2.5.4) and sequenced to verify identity (Section 2.2.5.9).

Table 3.2: Primers for quantitative real-time PCR for *E. tenella*

Gene	Primer	Sequence 5′→3′	Tm (°C)	Amplicon (Bp)
etsub1	Sense	GCGAAGAAGGAGTTGACGAC	60.0	227
	Antisense	ATGCCTTCTCCGTTGTTTTG	59.95	
etsub2	Sense	GTCTGAGGGTTGGAGACCAG	59.88	137
	Antisense	AGCCTCCCGGTATACTTCGT	59.96	
etsub3	Sense	ATGGAGGGGTGTACAAGCTG	59.91	187
	Antisense	GCAGAGCTAAAGGAACGGTG	59.94	
etsub4	Sense	GAAGAGACACGCACGATCAA	59.99	129
	Antisense	CAAACTTCCTTGCAGAAGGC	59.86	
etsub5	Sense	GGGCTTCAGAGCTTCCTTTT	60.0	142
	Antisense	GCCGTCGGAGTAAGTGTAGC	59.92	
etr18s	Sense	CTGATGCATGCAACGAGTTT	60.05	234
(Supercontig 190rc)	Antisense	GACCAGCCCCACAAATAAG	60.0	
etgam56	Sense	GACACCATGCCCGAGTTC	60.05	120
(ETH_00007320)	Antisense	GCGGTAGTTGGTGGTGTTTT	59.90	
etmic4	Sense	AGGAGGCTGAGCAAGTTGAG	59.75	109
(ETH_00024085)	Antisense	GCCCAGTAGTCGTCTTCGTC	59.87	

**Eimeria tenella* database (http://www.genedb.org/Homepage/Etenella) accession numbers for each sequence are indicated in brackets; Tm: melting temperature; Bp: base pairs.

3.2.2.2 Validation of the qRT-PCR assay: primer efficiencies

Relative quantification describes the ratio between the gene of interest, and an endogenous reference gene present in all samples. To accommodate the calculation of this ratio, it is necessary that primer pairs designed for each individual gene display sufficient amplification efficiency. In order to calculate the amplification efficiency of each primer pair, serial dilutions made of *E. tenella* gametocyte cDNA (Section 2.2.4.3.2) were amplified with QuantiFast SYBR Green technology and different gene primer sets. Cycle threshold (C_T) values were then plotted against the log dilution of cDNA and a standard curve produced (Figure 3.2) to produce a linear standard curve with $R^2 > 0.98$. A line of regression was then calculated for each standard curve of each primer pair and the slope of the line determined. The slope of the line was then used to calculate the amplification efficiency of each primer pair (E) as described in Section 2.2.4.3.2, and was presented as a percentage. If a primer pair demonstrated an amplification efficiency of >85%, this was deemed satisfactory to determine relative transcript abundance (Table 3.3) (Pfaffl, 2001, Tichopad *et al.*, 2004). Melt curve analysis was determined, with one dominant peak observed for each primer pair confirming a single amplification product in the reaction.

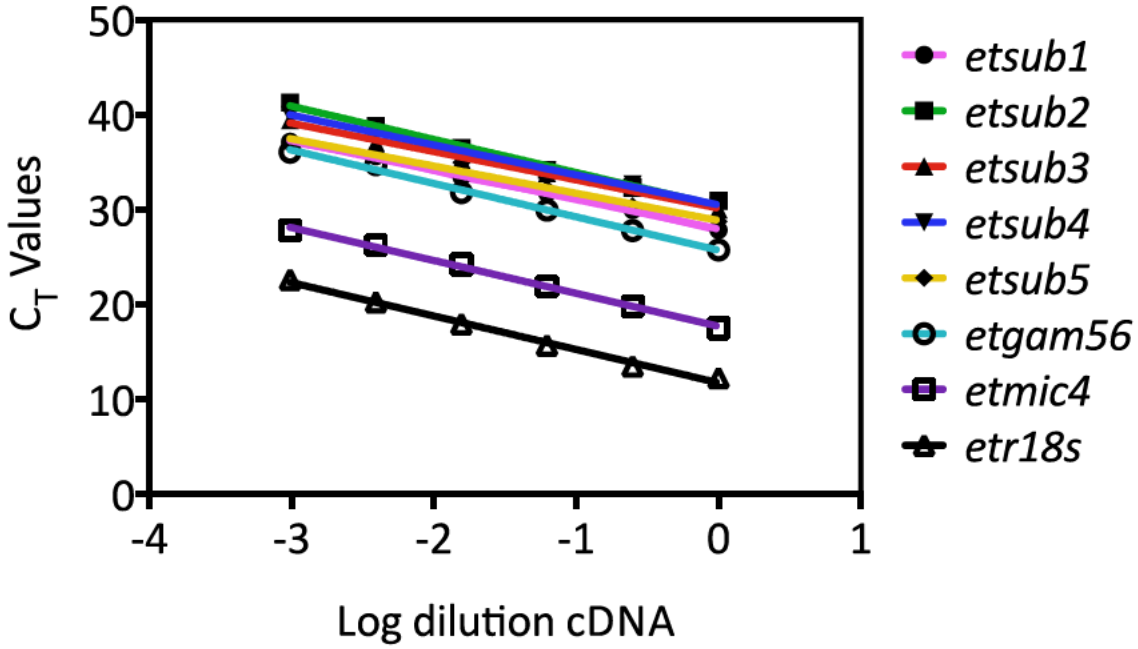


Figure 3.2: Linear regression of C_T versus template dilution of *E. tenella* in qRT-PCR reactions

The C_T values of the qRT-PCR reactions were plotted versus the log dilution factor of cDNA for *etsub1*, *etsub2*, *etsub3*, *etsub4*, *etsub5*, *etgam56*, *etmic4* and *etr18s* to determine primer efficiency of each primer pair. A line of regression was then calculated for each primer pair reaction to produce a linear standard curve to determine the R^2 value.

Table 3.3: Primer efficiency percentage for subtilase qRT-PCR primers

Gene	R^2	Slope of C_T vs. log dilution cDNA	Primer efficiency (E)	Amplification efficiency (%)
etr18s	0.992	-3.52	1.92	92.4
etsub1	0.996	-3.28	2.02	101.8
etsub2	0.991	-3.49	1.93	93.4
etsub3	0.993	-3.2	2.05	105.3
etsub4	0.998	-3.15	2.07	107.7
etsub5	0.980	-3.44	1.95	95.3
etgam56	0.995	-3.52	1.92	92.4
etmic4	0.996	-3.46	1.95	94.5

3.2.2.3 Relative transcript abundance of subtilase genes in *E. tenella*

Total RNA was extracted and the relative transcript abundance of subtilase genes in the different life cycle stages of *E. tenella* determined. Each sample was run in triplicate. No-template negative controls were included. Threshold cycles (Ct) were calculated from baseline-corrected normalised fluorescence data. The threshold fluorescence level was set to an arbitrary value (0.03) for all assays so that the amplification curves were in log-linear phase at that level. Each target gene was normalised to the transcript expression of *etr18s*. This ratio was then compared between the different life cycle stages to determine levels of transcript expression. Statistical analysis was performed using Two-Way ANOVA with a post-Bonferroni test to compare the level of expression between all of the life cycle stages. Note, the term 'transcript abundance' is used here rather than 'transcription' due to known mechanisms of mRNA-storage and repressed translation in *Plasmodium* gametocytes.

Genes with known gene expression profiles were also investigated to validate the relative quantities observed. Semi-quantitative RT-PCR analysis of *etgam56* has shown previously that this gene is expressed in a gametocyte-specific manner (Walker *et al.*, 2010), so it was used here as a positive control for gametocyte expression. Figure 3.3a shows the gametocyte-specific expression of *etgam56* with no transcript detected in merozoites, unsporulated or sporulated oocysts. A significant up-regulation of *etgam56* relative gene abundance was observed in late stage gametocytes ($p<0.001$) with a small level of transcript in early stage gametocytes ($p<0.05$).

The microneme encoding gene, *etmic4*, has been shown to be an asexual stage-specific protein in *E. tenella* (Tomley *et al.*, 2001) and, as such, was used as a positive control for merozoite-specific gene expression. Levels of transcript of *etmic4* were significantly upregulated in merozoites in comparison with early stage gametocytes, unsporulated and sporulated oocysts ($p<0.001$), as well as late stage gametocytes ($p<0.005$). Levels of transcript abundance detected in gametocytes, particularly late stage gametocytes, were also significantly higher ($p<0.05$) when compared with unsporulated oocysts where no *etmic4* transcripts were detected (Figure 3.3b). This can be explained by the fact that the development of the parasite is not perfectly synchronous and gametocyte preparations are sometimes contaminated with merozoites.

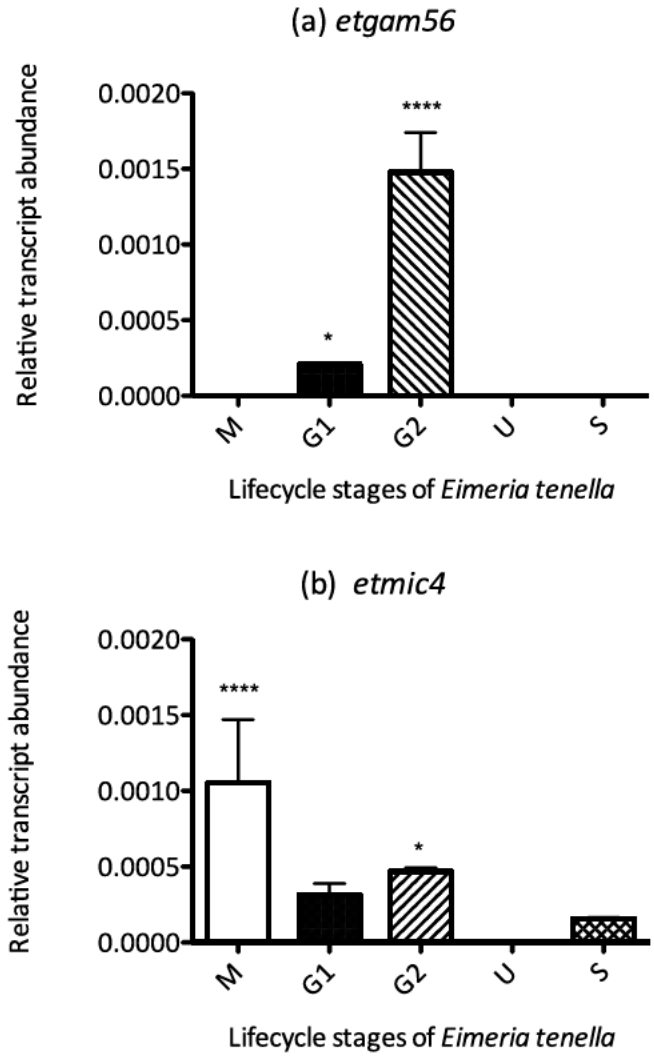


Figure 3.3: Relative transcript abundance of genes *etgam56* (a) and *etmic4* (b) in different life cycle stages of *E. tenella*

The relative transcript abundance of genes *etgam56* (a) and *etmic4* (b) were examined in different life cycle stages of *E. tenella.* Total RNA was reverse transcribed with random hexamer primers and the genes were amplified from cDNA using gene-specific primers. Relative transcript abundance of each gene was determined by calculating the difference between the C_T value of the target gene and the C_T of the reference gene, *etr18s*. This ratio was then compared between the different life cycle stages. Columns show the mean and the error bars show the standard deviation of three replicates per life cycle stage from at least two different experiments. A statistical comparison was made between each of the five life cycle stages, represented by '****' (P<0.001); '***' (P<0.005); '**' (P<0.05);'*' (P<0.01) and 'ns' (not significant). Merozoites (M), 134hr p.i. gametocytes (G1), 144hr p.i. gametocytes (G2), unsporulated oocysts (U) and sporulated oocysts (S) are indicated.

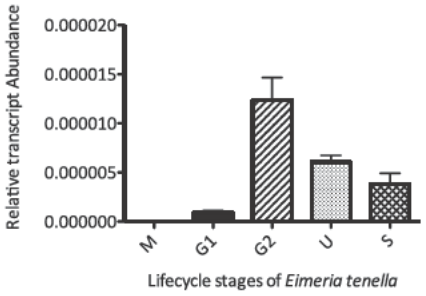
The relative transcript abundance of *etsub1* in the different life cycle stages of *E. tenella* is shown in Figure 3.4a where transcript was detected at low levels in early stage gametocytes (G1), and then at significantly higher levels in late stage gametocytes (G2) and unsporulated oocysts ($p<0.001$). A significant level of transcript was also observed in sporulated oocysts ($p<0.01$) when compared with the no DNA control.

Relative transcript abundance of *etsub2* was found to be gametocyte-specific with no transcript expression observed in any other life cycle stages (Figure 3.4b). Although 10-fold less abundant than other subtilase genes measured, the transcript abundance was significantly greater ($p<0.001$) in early and late stage gametocytes compared with the other stages.

Figure 3.4c shows the relative transcript abundance of *etsub3* in the life cycle stages of *E. tenella*. Although transcripts were detected in merozoites and early and late stage gametocytes, levels were significantly lower compared with expression observed in the sporulated oocyst stage of the parasite ($p<0.001$). No transcript of *etsub3* was detected in unsporulated oocysts.

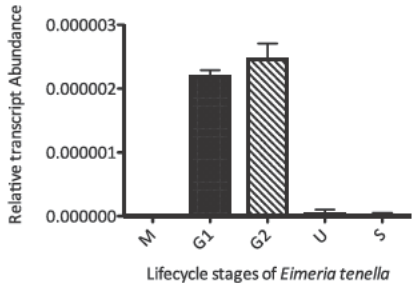
Transcripts of *etsub4* were detected in all life cycle stages of *E. tenella* except merozoites (Figure 3.4d). There was a significant increase in transcript abundance observed in early stage gametocytes ($p<0.05$), late stage gametocytes ($p<0.005$) and sporulated oocysts ($p<0.01$) compared with expression in merozoites. The highest amount of transcript detected, however, was observed in unsporulated oocysts where it was significantly greater than all the other life cycle stages ($p<0.001$).

Expression of the gene, *etsub5,* was detected in all five life cycle stages as shown in Figure 3.4e. A low level of transcript was detected in merozoites and unsporulated oocysts with significantly higher levels of transcript detected in late stage gametocytes and sporulated oocysts in comparison with those two life cycle stages ($p<0.01$). However, there was a prominent up-regulation observed in early stage gametocytes, with transcript levels being significantly greater in this stage in comparison with all the other life cycle stages ($p<0.001$).



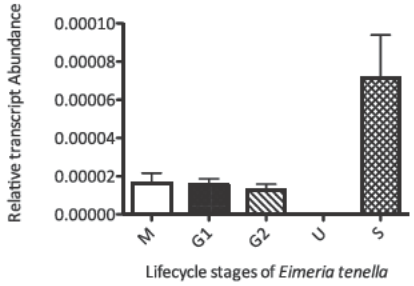
a)

	M	G1	G2	U	S
M		ns	****	****	***
G1			****	****	**
G2				****	****
U					*



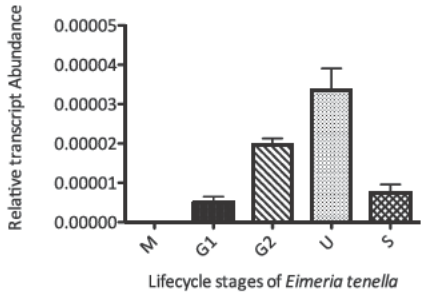
b)

	M	G1	G2	U	S
M		****	****	ns	ns
G1			*	****	****
G2				****	****
U					ns



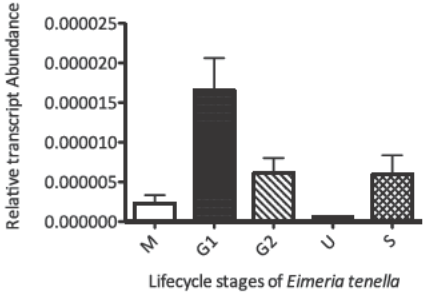
c)

	M	G1	G2	U	S
M		ns	ns	ns	****
G1			ns	ns	****
G2				ns	****
U					****



d)

	M	G1	G2	U	S
M		*	****	****	**
G1			****	****	ns
G2				****	****
U					****



e)

	M	G1	G2	U	S
M		****	ns	ns	ns
G1			****	****	****
G2				**	ns
U					++

Figure 3.4: Relative transcript abundance of subtilase gene, *etsub1, etsub2, etsub3, etsub4* and *etsub5* in different life cycle stages of *E. tenella.*

The relative transcript abundances of subtilase genes *etsub1* (a), *etsub2* (b), *etsub3* (c), *etsub4* (d) and *etsub5* (e) were examined in different life cycle stages of *E. tenella.* Total RNA was reverse transcribed with random hexamer primers and the gene was amplified from cDNA using gene specific primers. Relative transcript abundance of each gene was calculated for each life cycle stage as described in Section 2.2.4.3.2; the level of transcript expression of the target gene, measured as C_T, was normalized to the transcript expression of *etr18s,* the reference *E. tenella* ribosomal 18s RNA gene. Columns show the mean and the error bars show the standard deviation of three replicates per life cycle stage from at least two different experiments. A statistical comparison was made between each of the five life cycle stages, represented by '****' (P<0.001); '***' (P<0.005); '**' (P<0.05);'*' (P<0.01) and 'ns' (not significant). Merozoites (M), 134hr p.i. gametocytes (G1), 144hr p.i gametocytes (G2), unsporulated oocysts (U) an sporulated oocysts (S) are indicated.

3.2.2.4 Analysis of subtilases with gametocyte relative abundance

Following qRT-PCR, the results were compared with those obtained in the initial semi-quantitative PCR to identify which subtilases were most suitable for further investigation. Proteins were chosen on the basis of their gametocyte-specific expression in semi-quantitative analysis, and/or high level of transcript abundance in early and/or late gametocytes in comparison with other lifecycle stages as identified in qRT-PCR.

*S*emi-quantitative RT-PCR analysis showed that *etsub1* has a gametocyte-specific expression profile, with a product amplified from early stage gametocytes. A low level of transcript abundance was detected in early stage gametocytes by qRT-PCR, however, there was also relatively high transcript abundance in late stage gametocytes and in unsporulated oocysts. Therefore, it may be too late for transcripts of *etsub1* that are expressed in late gametocyte and unsporulated oocyst stages to have a role in the proteolytic processing of the GAM56 and GAM82 proteins. Given however, that the semi-quantitative RT-PCR results showed an early gametocyte - specific expression profile, *etsub1* was selected for further investigation.

Expression of *etsub2* was gametocyte-specific, as shown by semi-quantitative RT-PCR. Moreover, high levels of gametocyte-specific transcripts were detected in qRT-PCR analysis. Thus, in light of this strong indication of a gametocyte-specific role, *etsub2* was selected for further analysis.

Both *etsub3* and *etsub4* were not analysed further in this study since results from both semi-quantitative RT-PCR and qRT-PCR results indicated transcript abundance was highest in sporozoites for *etsub3*, and unsporulated oocysts for *etsub4*, making it unlikely that either of them play a significant role in processing of oocyst wall proteins.

The gene, *etsub5* showed a significant up-regulation of expression in early gametocytes of *E. tenella*, indicated by both semi-quantitative RT-PCR and qRT-PCR compared with other lifecycle stages. Thus, *etsub5* was also selected for further analysis.

PCR products of *etsub1, etsub2* and *etsub5* from qRT-PCR and/or semi-quantitative PCR in *E. tenella* were cloned into pGEM® T-Easy cloning vector and sequenced with vector specific M13 primers to confirm identity (Section 2.2.5.4) before further analyses were done.

3.2.3 Sequence analysis of subtilases

Gene sequences obtained for the three subtilases with gametocyte stage-specific expression – *etsub1*, *etsub2* and *etsub5* – were analysed to determine the correct open reading frame and translated amino acid sequence (Section 2.2.6.3 and 2.2.6.4).

ClustalW alignment was performed to compare the similarities and differences between the amino acid sequences of *Et*SUB1, *Et*SUB2 and *Et*SUB5. A high level of homology was observed between the three sequences with identical ('*'), highly similar (':') and similar ('.') amino acid residues shown in Figure 3.5. The histidine residue (H) of the active site was completely conserved between *etsub1, etsub2* and *etsub5* (Figure 3.5). Furthermore, half of the residues in the histidine active site were conserved (*) between the three proteins, whilst most of the non-conserved residues consisted of highly similar (':') residues; hence, the functional structure of the active site is conserved across the three proteins. Scanprosite identified the histidine active site, characteristic of subtilases, in *Et*SUB1 and *Et*SUB5, though, despite the conservation, not in *Et*SUB2. The aspartic acid active site (highlighted in blue), including the functional aspartic acid residue, was conserved between sequences *Et*SUB2 and *Et*SUB5 but was mostly absent in *Et*SUB1 (Figure 3.5).

```
EtSUB1          -------------------------------------------------------NYLHP 5
EtSUB2          AAANVEDPYFPLQWHLQAAEGLYNLGADKAWQQMQSAQQQQPQQQQVVVALLDTGCGPHE 60
EtSUB5          -------------------------EADRTWRKLEATENTKP----VVLAILDTGCYLHQ 31
                                                                         *

EtSUB1          ELSLS-------MWVNRKELHGEEGVDDDGNGFVDDLFGWNFIQDNNNPMDDN-GHGSHV 57
EtSUB2          DINN-------ILWINSKE-NCSDGVDHDNNGYVNDCKGWDFVDDT-NDITDTYGHGTAV 111
EtSUB5          DFIDDFDIAKSIFWDNAGETDCSDGIDNDGNGYIDDCFGWNFVEDNGHPFTDDSGHGTSV 91
                ::          :* *  * . .:*:*.*.**:::*  **:*::*. : : *  ***: *

EtSUB1          AGIVAALQNNGEGISGISERAKVMALKIL------------------------------- 86
EtSUB2          AALMAA-SFDKKGGRGVSPTGRVMCLRVGDQGGVQLSRQLMAMDYAIQQGAQVSVHPFSL 170
EtSUB5          TSVAAARAHDGKGGRGVLPNPTVMCLRVGSERGVWTSATIPALDYA-------------- 137
                :.: **   : :*  *:     **.*::

EtSUB1          ------------------------------------------------------------
EtSUB2          ATENEVYREAFEKLLSSNHLAVVAAGDNGCDLDEDTCKSYPASFKIPEPSGLLVVGSSTY 230
EtSUB5          ------------------------------------------------------------
```

Figure 3.5: ClustalW alignment of subtilases up-regulated in gametocytes of *E. tenella*

ClustalW was performed to compare amino acid sequences of subtilases *Et*SUB1, *Et*SUB2 and *Et*SUB5 identified to be up-regulated in gametocytes of *E. tenella* by semi-quantitative PCR and qRT-PCR. Identical amino acid residues between the three subtilase proteases are denoted as '*', whilst highly similar amino acids are denoted as ':' and similar amino acids are denoted as '.'. Aspartic active sites of the subtilase family (blue) are found in the putative subtilase serine proteases, *Et*SUB2 and *Et*SUB5, with the aspartic acid residue '**D**' highlighted in bold. Histidine active sites (green) are found in subtilase serine proteases, *Et*SUB1 and *Et*SUB5, with the histidine residue '**H**', highlighted, observed in all three sequences.

3.3 Characterisation of gametocyte subtilases in *E. tenella*

3.3.1 Full-length amplification of cDNA

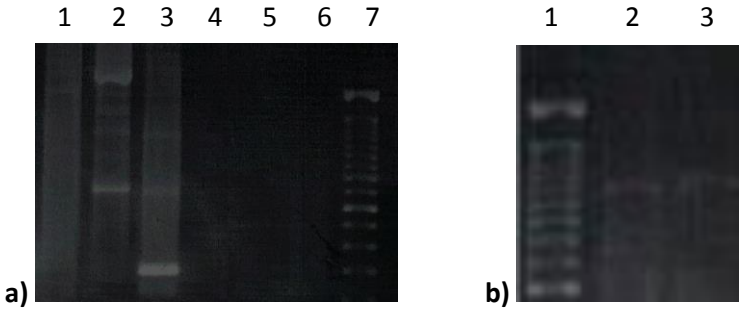
Bioinformatic analysis of *Et*SUB1, *Et*SUB2 and *Et*SUB5 revealed the absence or partial absence of conserved residues that form the catalytic triad characteristic of subtilase active sites suggesting that sequence was missing at the 3' end of each gene. ClustalW alignment of each gene with the relevant *T. gondii* homologue also indicated that sequence was missing at the 5' end. Therefore, rapid amplification of cDNA ends (RACE) PCR was employed to amplify the full-length coding sequence of *etsub1*, *etsub2* and *etsub5* using Marathon® cDNA (Clontech©) as described in section 2.2.3.3.

3.3.1.1 Isolation of the full-length gene sequence encoding subtilases *Et*SUB1, *Et*SUB2 and *Et*SUB5

The Marathon® ds cDNA amplification kit was used to produce ds cDNA for use in RACE PCR. Marathon® ligated-adaptor poly A^+ RNA (mRNA) was purified from *E. tenella* gametocytes using Dynabead® magnetic beads (isolated at 134hrs p.i. as described in section 2.2.2.2.2) and used to synthesise Marathon® ds cDNA using the amplification kit (in Section 2.2.3.3). Following synthesis, the adaptor-ligated ds cDNA from *E. tenella,* as well as control ds cDNA (supplied in the kit) was included in a positive control RACE PCR to ensure the synthesis reaction was successful. 5`-RACE TFR primer and 3`-RACE TFR primers were cycled with the Marathon® API primer (adaptor specific-primer) using the control ds cDNA, and, as expected, a 2.6kb band was observed in the 5` control reaction (Figure 3.6a, lane 3), and a 2.9kb band observed in the 3` reaction (Figure 3.6a, lane 3).

3.3.1.2 RACE PCR reaction controls

A control RACE PCR was performed with Marathon® API primer (adaptor specific-primer) and *E. tenella* β-actin primers to confirm that ds cDNA had been produced from *E. tenella* gametocytes with the Marathon® adaptor ligated to ds cDNA. Lanes 2 and 3 in Figure 3.6b show that products of the expected sizes of 1023bp and 1107bp were amplified with the Marathon® API primer and *E. tenella* β-actin sense and antisense primers, respectively. Both bands observed in the control RACE PCR with *E. tenella* β-actin primers were gel purified and then sequenced to confirm amplification of the *E. tenella* β-actin gene. The *E. tenella* Marathon® ds cDNA synthesised was used for subsequent RACE amplification of the gametocyte-specific subtilase genes, *etsub1, etsub2* and *etsub5*.



a)
Lane 1: Marathon® API primer and 5` TFR, control ds cDNA
Lane 2: Marathon® API primer and 3` TFR, control ds cDNA
Lane 3: 5` TFR and 3` TFR internal control, control ds cDNA
Lane 4: Negative control, β-actin sense and Marathon® API primer
Lane 5: Negative control, β-actin antisense and Marathon® API primer
Lane 6: Negative control, 5` TFR and 3` TFR
Lane 7: 100bp ladder

b)
Lane 1: 100bp ladder
Lane 2: β-actin sense primer and Marathon® API primer, *E. tenella* ds cDNA
Lane 3: β-actin antisense primer and Marathon® API primer, *E. tenella* ds cDNA

Figure 3.6: *E. tenella* gametocyte Marathon® ds cDNA synthesis control RACE PCR

Marathon® ds cDNA was synthesised from gametocyte mRNA (isolated 134 hours p.i.) and amplified with (a) 5`-RACE TFR primer and 3`-RACE TFR primers, and Marathon® API primer; and (b) *E. tenella* β-actin gene specific primers and Marathon® API primer. Negative controls were included for all primer reactions. PCR products were electrophoresed in 1% agarose gel and visualised using Gel Red® and VersaDoc® Imaging system (Section 2.2.4.5).

3.3.2 Characterisation of *etsub1* sequence

3.3.2.1 RACE PCR of *etsub1*

Marathon® ds cDNA synthesised from gametocyte mRNA (isolated 134 hrs p.i) was used in RACE PCR to characterise the unknown 5` and 3` sequence of *etsub1*. A 5` gene specific reverse primer (GSRP), CAGTTCGGGATGGAGGTAATTA, and a 3` gene specific forward primer (GSFP), GAAGGCATAAGCGGAATTAGTG, were designed using the previously confirmed *etsub1* sequence (Figure 3.5). Touchdown PCR thermal cycling conditions were then employed to amplify the unknown 5` and 3` sequence of *etsub1*. A small 5` product, approximately 100bp in size, was observed in 5` RACE PCR (data not shown). This band was excised from the gel, purified and cloned into sequencing vector pGEM® – T Easy (Section 2.2.5.4). The sequence of the 5` region of *etsub1* was then determined using vector-specific M13 primers.

Despite various attempts to optimise the annealing temperature and PCR conditions, a 3` product was unable to be amplified with the 3` GSP and Marathon® API primer. Therefore, a different strategy was employed using bioinformatic analysis of databases to try to obtain the unknown 3` region of *etsub1*.

3.3.2.2 5` sequence analysis of *etsub1*

Following cloning into pGEM® – T Easy and sequencing, analysis was performed using Seqman program (Lasergene™ 10), revealing that only an extra 65bp sequence was obtained using RACE PCR upstream from the known sequence of *etsub1*. Translation of this sequence revealed the full aspartic acid active site completing the catalytic triad motif, which is conserved between the three gametocyte-specific subtilases as showed by ClustalW analysis (Figure 3.7).

However, no start codon for the gene was included in the sequence obtained. A new 5` reverse gene specific primer was designed for *etsub1*, CACAACAGCAACGGGGACAG, and RACE PCR repeated, however, despite numerous attempts to optimise annealing temperature and PCR cycling conditions, no further sequence was obtained. Therefore, bioinformatic analysis was also performed to try and determine the remaining 5` sequence.

```
EtSUB1          ---------------------------------------------SVPVAVVDTGNYLHP 5
EtSUB2          AAANVEDPYFPLQWHLQAAEGLYNLGADKAWQQMQSAQQQQPQQQQVVVALLDTGCGPHE 60
EtSUB5          -------------------------EADRTWRKLEATENTKP----VVLAILDTGCYLHQ 31
                                                                         *

EtSUB1          ELSLS-------MWVNRKELHGEEGVDDDGNGFVDDLFGWNFIQDNNNPMDDN-GHGSHV 57
EtSUB2          DINN-------ILWINSKE-NCSDGVDHDNNGYVNDCKGWDFVDDT-NDITDTYGHGTAV 111
EtSUB5          DFIDDFDIAKSIFWDNAGETDCSDGIDNDGNGYIDDCFGWNFVEDNGHPFTDDSGHGTSV 91
                ::          :* *  * . .:*:*.*.**:::*  **:*::*. : : *  ***: *

EtSUB1          AGIVAALQNNGEGISGISERAKVMALKIL------------------------------- 86
EtSUB2          AALMAA-SFDKKGGRGVSPTGRVMCLRVGDQGGVQLSRQLMAMDYAIQQGAQVSVHPFSL 170
EtSUB5          TSVAAARAHDGKGGRGVLPNPTVMCLRVGSERGVWTSATIPALDYA-------------- 137
                :.: **   : :*  *:     **.*::

EtSUB1          ------------------------------------------------------------
EtSUB2          ATENEVYREAFEKLLSSNHLAVVAAGDNGCDLDEDTCKSYPASFKIPEPSGLLVVGSSTY 230
EtSUB5          ------------------------------------------------------------
```

Figure 3.7: ClustalW alignment of subtilases up-regulated in gametocytes of *E. tenella*

ClustalW was performed to compare amino acid sequences of subtilases *Et*SUB1, *Et*SUB2 and *Et*SUB5 identified to be up regulated in gametocytes of *E. tenella* following RACE PCR of 5` region of *Et*SUB1. Identical amino acid residues between the three-subtilase proteases are denoted as '*', whilst highly similar amino acids are denoted as ':' and similar amino acids are denoted as '.'. The aspartic acid active site, including the conserved aspartic acid residue '**D**' was identified in the *Et*SUB1 sequence, similar to the aspartic active site previously identified in *Et*SUB2 and *Et*SUB5. Histidine active sites (green) are found in subtilase serine proteases *Et*SUB1 and *Et*SUB5, with the histidine residue '**H**', highlighted, observed in all three sequences.

BLASTp was performed with the translated sequence of *etsub1* previously confirmed by PCR against the *T. gondii* database. A homologue, TGME49_035950 (E=4.8e^{-49}), was identified as a subtilase family serine protease (denoted as *tgsub8*). ClustalW was then performed to determine the level of homology between the two genes. Figure 3.8 shows that there was a high level of conservation between the known sequence of *Et*SUB1 and *Tg*SUB8; therefore, it was likely that this is the orthologue of the *E. tenella* subtilase. Approximately 1000 more amino acids were identified further upstream in the 5` region of the *tgsub8* gene compared with the *E. tenella* sequence. Given that the current *E. tenella* database prediction of *etsub1* is that only a further 18 amino acids are missing in this region, this annotation could be incorrect and there is actually a lot more 5` region still to be identified for *etsub1*.

When the 3' regions of *etsub1*and *tgsub8* were aligned, it appeared that approximately 250 amino acids were missing from the 3' region of *etsub1* rather than just the single amino acid predicted by the database. The absence of a predicted serine active site in the 3` sequence of *etsub1* further indicated that there was more sequence in the 3` region of *etsub1* than was predicted by the database. Following analysis of both the 5` and 3` region of the two homologous genes, it was concluded that the introns and exons of *etsub1* had been incorrectly predicted by the *E. tenella* genome database and the genomic sequence initially annotated as an intron contains the coding sequence of *etsub1*. To confirm this, the genomic sequence upstream and downstream of ETH_00009790 was analysed for the possible coding sequence of the 5` and 3` region of *etsub*1 using ToxoDB GBrowser v2.48 genome browser. The gene was found to be located on supercontig_570, which is 6.319kb in length. The predicted gene is located from bases 117-5072 on the supercontig and is 4.955kb in length including exons and introns.

With only 117bp of genomic sequence upstream from the mis-annotated start codon of the *etsub1* gene, there was insufficient sequence to allow determination of the 5` region of *etsub1* on this supercontig. It was also unlikely that the promoter sequence for the gene was included on the supercontig. Given that the predicted start codon is believed to be false, it seems that the 5` sequence for *etsub1* is contained on a different supercontig and cannot be identified at present.

Analysis of the 3` region of the annotated gene ETH_00009790 also revealed that there was very little sequence downstream from the gene on supercontig_570 (Figure 3.9), with a further 1.3kb of genomic sequence following the predicted gene until the contig ends. Looking at the gene in ToxoDB GBrowser v2.48 genome browser, the coding sequence, or exon, was separated by large introns of genomic sequence, often larger than 2kb. Therefore, there may not be enough sequence downstream from the gene before the contig ends for the entire coding sequence to be present. To investigate this, the 1.2kb genomic sequence was downloaded from the *E. tenella*

database and translated (Lasergene™ 10) to reveal the different open reading frames available. Manual analysis of each of the open reading frames revealed that none of them contained the third active site, the serine active site motif, of subtilase proteases. Therefore, it can be assumed that the sequence containing this motif, as well as the remaining 3` region of *etsub1* is located on another contig. No further information, therefore, could be obtained about the coding sequence of *Et*Sub1 and so the sequence obtained previously from PCR was used for further analysis of the *etsub1* protein (Section 3.2.3).

```
etsub1          ------------------------------------------------------------------
T.gondii        MNLRSDDSDIDPAASNRISGRRGEARSLGLPTREVNYGRTGGQSSSTLGGSLRVQGVSALPLSAFEV 67

etsub1          ------------------------------------------------------------------
T.gondii        HAIESKDDWQPRSSQEGPSHQDVSLDASSPMQKLHRLVGDASELQDAASRPASDDTLNNEKRNSVPL 133

etsub1          ------------------------------------------------------------------
T.gondii        LPGKRAGDANFLSEPRGVATNNSPSSQPDTSRVNSATANQPKRKTTKARNSVPDSGQSSPRGQMVVV 200

etsub1          ------------------------------------------------------------------
T.gondii        NVEALGSPIDKDTGRPLFYNDRILVGWRCSDHELLALKALEMEEEEEEDDVPQGSSLSAPASMDDPN 268

etsub1          ------------------------------------------------------------------
T.gondii        ASITDEKPQVGIPCAQSGFDPLTLGASQSEIRLCGYVPPDLGIDIRSLRDSRQTTSSASASSNPAVP 334

etsub1          ------------------------------------------------------------------
T.gondii        PLDLSQVSALWPNTRRPRNNNENSAHFGNFNRHGNMSPADSPSAVFASSDNSKRRTSTPGQPGLSDG 401

etsub1          ------------------------------------------------------------------
T.gondii        MVGAPPRRLQVLDRMLKSYASPEVESVKTLFRRLPQKRSLRKKVPHGETWKRPESQVDNQETLSASD 468

etsub1          ------------------------------------------------------------------
T.gondii        AQKNNSNSGLGISSKDMAEERQSLGLSSGNHATSKGATRPSKGNLDRSGGTVYSEHTRSPGTGDNLE 535

etsub1          ------------------------------------------------------------------
T.gondii        GGRAGKNIRDSVTEDSKTEEFVKKRLLDDDELCGMDLIKLSLKHYRARVEGTRYDGYPENNSENVEA 602

etsub1          ------------------------------------------------------------------
T.gondii        LKAFMKEFASHDYVGDILFLAPDAPIHAHATSPELNGEASPTDGEECTSASCTSKHPTEIGNKEEQA 669

etsub1          ------------------------------------------------------------------
T.gondii        RSAFRVNPPQPPALAPDDRDSVQRSNSNVNKTVKISEYPNGSADKQQSIFLKRTLNSDNSKNLDFWA 736

etsub1          ------------------------------------------------------------------
T.gondii        VNGFGPYPTSAPGDKRAEKVSQDEDGVAEEPFRVGKPSVPEQTIHAVSSNRSLQETLQVSESGKEES 803

etsub1          ------------------------------------------------------------------
T.gondii        EESGRTVFPNDPLFSLQWSLGSPPQDVQMSAEFCFLRAREELQGRSADADERGAGGSFQQTPNKSQL 867

etsub11         ------------------------------------------------------------------
T.gondii        LAKVRQLCEAAGHKDKYATGGKLRATTLAAEVEDGQGDTDQHKKDGITPDSASTLQAEANEKQVQDL 933

etsub1          ------------------------------------------------------------------
T.gondii        QDLEGDMEERQTTPSLQESAELHVLQGSHAGRGGGTHIGENAADANSIDMVKAWRLSLSPQDEDNKK 996

etsub1          ----PVAVVDTGINYLHPELSLSMWVNRKELHGEEGVDDDGNGFVDDLFGWNFIQDNNNPMDDNGHG 83
T.gondii        KQEVLVAVIDTGVNYIHSDLANSIWVNEQELNGIPGFDDDQNGYIDDVYGWNFLHGNNNPMDDNGHG 1064
                     ***:***:**:*.:*: *:***.:**:*  *.*** **::**::****::.***********

etsub1          HGSHVAGIVAALQNNGEGISGISERAKVMALKILDQKGEGDVSHAIPAIQYAVDNGAKVLTNSWG-- 126
T.gondii        HGSHVAGIIGAQRNNNQGVSGISSHARIIALKILNKKGEGDVSHAIPAIRYALDNGAKVITNSWGGI 1128
                ******:.* :**.:* : ****.:*:::*****::**************:**:******:*****

etsub1          ------------------------------------------------------------------
T.gondii        SGTQILGVLLKHLTAPGENITSTWMGSGYRLSSGTSMAAPMVSGVAAEILAYNPMLQPQQVVDVLVQ 1196

etsub1          ------------------------------------------------------------------
T.gondii        SAITDKRHANVSLTGARLNAYRAIILSQLQFVSLSPAELHVGGTVDPVGEVEMIFRSLMLPPGVYE  1322

etsub1          ---------------------------------------------
T.gondii        GTIELVYSRRLASTRGLGMLSGVLKRRIVQLVQVTLPVKLTIIQ 1364
```

Figure 3.8: ClustalW alignment of *Et*SUB1 and *Tg*SUB8 amino acid sequences

ClustalW was performed to compare the amoino acid sequence of *Et*SUB1 and a subtilase serine protease orthologue in *T. gondii.* The aspartic acid active site motif is highlighted in blue, the histidine active site motif in green, and the serine active site motif is highlighted in yellow. There was a high level of conservation between the sequences at the active sites in each subtilase with conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids identified.

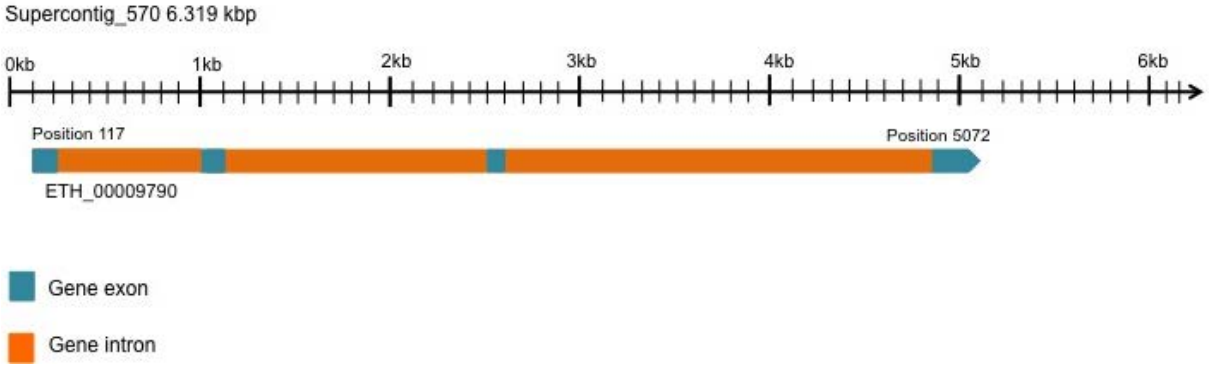


Figure 3.9: Gene ETH_00009790

The annotated *E. tenella* subtilase gene ETH_00009790 position in supercontig_570 in the *E. tenella* genome was analysed in ToxoDB GBrowser v2.48 genome browser. The supercontig is 6.319kb in length, and the ETH_00009790 gene is located at 117-5072bp. Introns are shown as connecting lines (orange) between the exons, or coding sequence (blue).

3.3.3 Characterisation of *etsub2*

3.3.3.1 RACE PCR of *etsub2*

Marathon® ds cDNA synthesised from gametocyte mRNA (134hrs p.i.) was used in RACE PCR to characterise the unknown 5` and 3` sequence of *etsub2*. A 5` gene specific reverse primer (GSRP), GGGTCTTCAACATTTGCTGCTGC, and a 3` gene specific forward primer (GSFP), GGTGGTTGGCAGCAGCACTTAT, were designed using the previously confirmed *etsub2* sequence (Figure 3.5) Touchdown PCR thermal cycling conditions were then employed to amplify the unknown 5` and 3` sequence of *etsub2*.

Despite attempts to optimise the annealing temperature and PCR conditions, neither a 5` nor 3` product amplified with the GSP and Marathon® API primer in the *etsub2* RACE PCR. Therefore, a different strategy was employed using bioinformatic analysis to try and obtain the unknown 5` and 3'regions of *etsub2*.

3.3.3.2 Bioinformatic analysis of *etsub2*

The *E. tenella* genomic database prediction of *etsub2* (ETH_00025145) is of a 439 amino acid protein. Alignment of known sequence with ETH_00025145 indicated that there were a further 311 amino acids predicted in the 5` region of *etsub2*, and no further sequence at the 3` region of the gene. Given that the serine active site motif, (Siezen and Leunissen, 1997a) was missing from the sequence, this prediction is likely to be inaccurate. The genome browser ToxoDB GBrowser v2.48 located ETH_00025145 on supercontig_1463 of the *E. tenella* genome, a contig that is 2.977kbp in size (Figure 3.10). The predicted gene of *etsub2* was positioned on the supercontig from bases 389-2961 and was 2.572kbp, in total, including exons and introns (Figure 3.10). With the aspartic acid and histidine active sites present in the gene, it is believed to be a subtilase; however, the currently predicted sequence does not include the serine active site in the 3` region, raising the possibility that significant coding sequence has been incorrectly annotated as an intron. The supercontig_1463 for the gene terminates a mere 16 bases after the incorrect 'stop' codon of the gene making it impossible to determine the missing 3` region sequence of *etsub2*.

The predicted database sequence of the gene (ETH_00025145) was again analysed to investigate the unknown 5` region of *etsub2*. As mentioned above, there were 311 amino acids, or 933bp, predicted in the 5` region of *etsub2* located upstream from the PCR-confirmed product. Given the problems highlighted with the annotations of the subtilase gene in the *E. tenella* genomic database, this part of the sequence was analysed to determine if it did contain *etsub2* sequence. BLASTp was performed to identify the orthologue of the *etsub2* confirmed sequence in *T. gondii.*

A serine protease/ subtilase, TGME49_319970, from this point named *Tg*SUB10*,* was identified, and aligned with *etsub2* to analyse the missing 5` region (Figure 3.11). The *T. gondii* sequence contained approximately 400 amino acids further upstream than were present in the 5' region of *etsub2*. The database sequence, ETH_00025145, predicted 311 amino acids in the 5` region, (Figure 3.11) and, therefore, given there was not a great discrepancy between the two homologues, the 5` region of ETH_00025145 may be an accurate prediction of the *etsub2* gene.

A tBLASTn was then performed with the *etsub2* sequence against the *E. maxima* genome database searching the 454 + Sanger database. The contig_10436 was identified to have homology (E = $6e^{-17}$) with the subtilase in *E. tenella.* This search was repeated with this contig back to the ETH_00025145 predicted sequence and gave an E score of $1e^{-62}$. Analysis of the two sequences identified homologous regions between the ETH_00025145 5` region and translated 5` region of contig_10436 (Figure 3.12). Given the homology between these two regions it was possible that the 5` region of ETH_00025145, as predicted in the *E. tenella* genome database, was correct. To test this, a 5` gene specific primer was designed to cover the start sequence of the 5` region of ETH_00025145, ATGCGACAAATAGCAGCCG, to try and amplify the missing 5` region of *etsub2* with the previously designed *etsub2* 5` reverse GSP RACE primer.

Several PCRs were performed to amplify the 5` region of *etsub2* using a variety of annealing temperatures and cycling parameters; however, all failed to produce usable sequence. This was thought to be due to the highly repetitive 5` sequence of ETH_00025145. There are long regions of repetitive sequence, including chains of ten alanine's or glutamine's in a row, with these sequences occurring more than twice in this region (Figure 3.12). It was possible that the 5` region predicted by ETH_00025145 was inaccurate, however, it was conserved in an homologous contig in the *E. maxima* database. It was also possible that the primer designed to amplify the coding sequence may not be part of the exon of the gene, and therefore could not be amplified. Therefore, only the previously confirmed sequence of *etsub2* was used for further investigation.

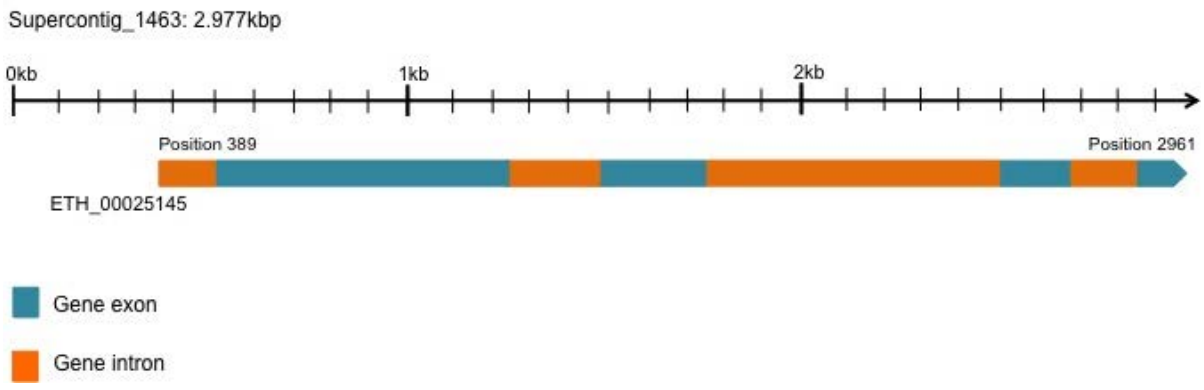


Figure 3.10: ETH_00025145 located on Supercontig_1463

The predicted coding sequence of *etsub*2, ETH_00025145 is located on the Supercontig_1463 in the *E. tenella* genome database as shown in ToxoDB GBrowser v2.48 genome browser. It is positioned from bases 389-2961 and is 2.572kbp in total including exons (blue) and introns (orange). Supercontig_1463 is 2.977kbp in length.

```
EtSUB2          ------------------------------------------------------------
TgSUB10         MASPTVPRRESVSRQGETGSRFPEDIFTKYPARNRHRLHGFSLCVVFACLAVFSSLSICP 60

EtSUB2          ------------------------------------------------------------
TgSUB10         WVQSALSVSPDLDFSESVKTVRSFAQPESEHEEKAASSEDSHVENSDSSNSGSATKRRHA 120

EtSUB2          ------------------------------------------------------------
TgSUB10         GILGFSFPTRTPPGDRAGEATWEGKADWTDEMTPPGDAESSVLASMVKKTSEKARVEMLS 180

EtSUB2          ------------------------------------------------------------
TgSUB10         SAVTKNCPYEQPLRLLFADMHKKVEMKRGRDRSRRLDPTENIILRWDTECLGSLAGKLSA 240

EtSUB2          ------------------------------------------------------------
TgSUB10         SPDVLDPEERSPDEIHLFRSNRDSSFSARAFPQRPLDPSLTQPAVVFRKGVFYGDRVTNK 300

EtSUB2          ------------------------------------------------------------
TgSUB10         KQIPEKSETSQAQVLAGRTLALSILAGALGVRVSTAGLRECSRDCSQAATETNIKESAEK 360

EtSUB2          ----------------------------------------AAANVEDPYFPL------- 12
TgSUB10         GKGKGERRGKRTGEEVGRESNDEYKCTKSCERELERRFWSATVESLHATFYPLFHIDVLQ 420
                                                         :. .:. .::**

EtSUB2          -----------------------QWHLQAAEGLYNLGADKAWQQMQSAQQQQPQQQQVVV 49
TgSUB10         ITNLSGVLAQTRTDLERNSEYKKQWYHLSQVSDFSLGSDKAWAEMN----YQSEENVPLV 476
                                       **:  :  . :.**:**** :*:     *.:::  :*

EtSUB2          ALLDTGCGPHEDINN------ILWINSKENCSDGVDHDNNGYVNDCKGWDFVDDTNDIT- 102
TgSUB10         ALLDTGCSKHEDYWDDSEGNARLWRNSQEDCSNNLDDDGNGYVDDCWGWNFAENNNDIFR 536
                *******. ***  :       ** **:*:**:.:*.*.****:** **:*.::.***

EtSUB2          DTYGHGTAVAALMAASFD--KKGG----RGKP---------------------------- 128
TgSUB10         DDSAHGTTMASLLVAKHTDSKRGVGIMKTGKIMCLKTGANNRVYASAVIAALQYAIINGA 596
                *  .***::*:*:.*..   *:*      **

EtSUB2          ------------------------------------------------------------
TgSUB10         QISICAFTFSKLYPAAFRVPSLLVVGSSKLTGGPCCSSNWGKNTVHVFAPGNRLWTGTNV 656

EtSUB2          ------------------------------------------------------------
TgSUB10         DHDAQTIKCTTCAAFAYGTSAAAALAGGVAAMVWAHLQTKQPIGWTSSKDKESVRVKKAL 716

EtSUB2          ------------------------------------------------------------
TgSUB10         VYGSRPSYRLAGACEANGIVNMHSAMHYYDTVPLQFEPVYHDYPSIFQLNIPPLLGASTF 776

EtSUB2          -----------------------------
TgSUB10         STPLSRTLVLGSVGASTLSLSLFGVPPLV 805
```

Figure 3.11: ClustalW alignment of *Et*SUB2 and *Tg*SUB10

ClustalW was performed with *etsub2* PCR confirmed amino acid sequence and its orthologue in *T.gondii*, a serine protease/subtilase protein, *Tg*SUB10. The aspartic acid active site motif is highlighted in blue, the histidine active site motif in green, and the serine active site motif is highlighted in yellow. There was a high level of conservation between the sequences at and around the active sites, with conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids identified.

a)

```
>contig10436  length=2673   numreads=110
          Length = 2673

 Score =  232 bits (591), Expect(2) = 1e-62
 Identities = 145/298 (48%), Positives = 190/298 (63%), Gaps = 50/298 (16%)
 Frame = -3

Query: 6    AAATAAATLAAAADPAAAAAAASAPQGAL-----RKLKNSPDSITLLLKWDFSCLAAQRQ 60
            AA  AA  L +AA  AAAAAAA+A   AL      K+K+ PD++TLLLKWDFSCLAA+RQ
Sbjct: 2551 AAVAAAVRLLSAAAAAAAAAAAAACAAALY*EVLNKIKSGPDAVTLLLKWDFSCLAAKRQ 2372

Query: 61   QREQQ----------QQQQQQQPVEVYLQLDRDVTR----------AAAAAE-------- 92
            Q++QQ          Q+QQQQQ  E+Y+ LD D+TR          AAAAA+
Sbjct: 2371 QQQQQQRQQQEQQQDQEQQQQQQFEMYIDLDEDLTRTGAAGTAAGAAAAAADPAAAAAAD 2192

Query: 93   ---AGKGQVLFTRGVFLRDKA----QKQQQQQQQQ----------QQKQQQQKQKQKQQQ 135
               A +GQ+LFT+GVFLRD      Q QQQQQQQQ          +QKQQ+++Q+Q QQQ
Sbjct: 2191 APAAKRGQLLFTKGVFLRDSGPSLEQLQQQQQQQQDSKQEEGSEEEQKQQEEEQQQLQQQ 2012

Query: 136  QQQQQQQFDDRVFGLQFLSQLLGHNIDLEAEAAELQQELQQELQQEQQQEQQQELLPQRQ 195
             QQQQQQ +DR FGL+ LS LL HN+DL A AAE Q+E  ++LQQ+  +E++
Sbjct: 2011 LQQQQQQQEDRRFGLEALSHLLDHNLDLHAAAAEQQEE--EQLQQDIDEEEK-------- 1862

Query: 196  QADGAAEQQKSPCVSSPSLCQELLLQKLLQQQEGLTVSKYLSSFDVDVVTFSRLKEFM 253
                  +++ SPCVSSP LC++LLLQ L QQQE + +S YLS  D+DV+TF++++ FM
Sbjct: 1861 ----GTKKKVSPCVSSPLLCKDLLLQLLQQQQEAILLSYYLSILDIDVLTFTKMRTFM 1700
```

b)

```
Score = 86.7 bits (213), Expect = 4e-17
Identities = 51/112 (45%), Positives = 61/112 (54%)
Frame = -2

Query: 292 VPAAAAAATAARGSSPRAAAAAANVEDPYFPLQWHLQAAEGLYNLGADKAWQQMQSAQQQ 351
           V   AAAA AA  ++  AAAAA +                    G  KA      A
Sbjct: 665 VAVVAAAAAAAAEAAATAAAAAGDAS-------------------GTKKA-----DALAA 558

Query: 352 QPQQQQVVVALLDTGCGPHEDINNILWINSKENCSDGVDHDNNGYVNDCKGW 403
             QQQ VVVA+LDTGCGPHEDI+++LWIN KE+  D  D+DNNGY+ND  GW
Sbjct: 557 ARQQQGVVVAVLDTGCGPHEDIDSVLWINPKEHKDDKKDNDNNGYINDYNGW 402
```

Figure 3.12: tBLASTn of *E. maxima* genome database

The predicted *E. tenella* coding sequence of *etsub2* ETH_00025145 was used to perform a tBLASTn against the *E. maxima* 454 + Sanger genomic database. A translated gene was identified in the *E. maxima* database (subject) with a high level of conservation to ETH_00025145 (query). The conserved protein residues between the two sequences are shown, with residues of similar function denoted by '+'. The first region of homology between the two sequences spanned between amino acids 6-235 of ETH_00025145 (a); whilst the second region of homology spanned between amino acids 292-403 (b).

3.3.4 Characterisation of *etsub5*

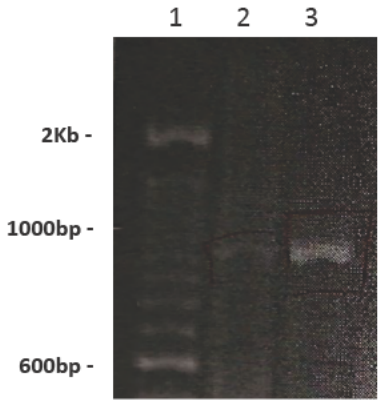
3.3.4.1 RACE-PCR of *etsub5*

Marathon® ds cDNA, synthesised from gametocyte mRNA (isolated 134 hrs p.i.), was used in RACE PCR to characterise the 5` and 3` sequence of *etsub5*. A 5` gene-specific reverse primer, TTGGCAGAAAGTCAAGAGACGG and a 3` gene specific forward primer ATCCCTGCTCTTGACTACGCGG were designed using the confirmed *etsub5* sequence (Section 2.2.4.5) for amplification with Marathon® API primer in RACE PCR. Touchdown PCR thermal cycling conditions were then employed to amplify the unknown 5` and 3` sequences of *etsub5*. A band was observed in the 3` reaction, amplified with 3` gene specific primer (GSP) and Marathon® API primer, although it was larger than the expected 815bp predicted by the *E. tenella* database sequence (Figure 3.13, lane 3). This band was excised from the gel, purified and cloned into sequencing vector pGEM® – T Easy (Section 2.2.5.4). The 3` region of *etsub5* was then sequenced using vector-specific M13 primers.

Despite various attempts to optimise the annealing temperature and PCR conditions, no product was amplified with the 5' GSP and Marathon® API primer. Therefore, another strategy was employed using further bioinformatics analysis and different gene-specific primers for amplification of the 5` sequence of etsub5 (Section 3.3.4.2).

3.3.4.2 Sequence analysis of 3` RACE PCR product of *etsub5*

Following cloning into pGEM® – T Easy and sequencing, analysis of the 3` RACE PCR product of *etsub5* was performed using the Seqman program (Lasergene™ 10) where the DNA sequence was translated into an amino acid sequence. Analysis of the coding sequence revealed the translated protein sequence at the 3` region of the protein, as well as the stop codon, the un-translated sequence (UTR) and the poly-A tail at the end of the gene. The residues that make up the serine active site, the third active site needed to complete the triad catalytic site (Siezen and Leunissen, 1997a) were identified within the 3` region of the protein (Figure 3.14). ClustalW analysis revealed that the sequence was highly homologous to the predicted protein sequence on the *E. tenella* genome database for the gene (ETH_00011340), however, there were significant differences in amino acids at the 3` end of the protein. Figure 3.14 shows the putative amino acid sequence of *Et*Sub5, including the correct 3` region, derived from the cDNA sequence of *etsub5,* compared with the sequence predicted by the *E. tenella* genome database. The serine active site was revealed by RACE PCR, however, it was not present in the genome database sequence prediction of the gene ETH_00011340. The stop codon for the *etsub5* gene was also located further downstream than originally predicted in the genome database (Figure 3.14). In order to



Lane 1: 100bp ladder
Lane 2: *etsub5* 3` GSP and Marathon® API primer, touchdown 65°C, 60°C, 55°C
Lane 3: *etsub5* 3` GSP and Marathon® API primer, touchdown 60°C, 55°C, 50°C

Figure 3.13: RACE PCR of 3` region of *etsub5*

Marathon® ds cDNA was synthesised from *E. tenella* gametocyte mRNA (isolated 134 hours p.i.) amplified with *etsub5* 3` GSP and Marathon® API primer. Using touchdown PCR (Section 2.2.4.4) a 3` product was amplified. PCR products were electrophoresed in a 1% agarose gel and visualised using Gel Red® and VersaDoc® Imaging system (Section 2.2.4.5).

```
Database        MRYLEGPSMKAAYYGTLSTDILRISTSTSDQPSLDFLPKIREVAKSLSSGTDEVAGALGR 60
EtSUB5          ------------------------------PSLDFLPKIREVAKSLSSGTDEVAGALGR 29
                                              ****************************

Database        VITAAVACVDKISLDRKVDVQRLGSLATNSQGNMSLQEPIKGTVNDGATSPQLLRAPSGL 120
EtSUB5          VITAAVACVDKISLDRKVDVQRLGSLATNSQGNMSLQEPIKGTVNDGATSPQLLRAPSGL 89
                ************************************************************

Database        LGDPFQMRQWYLHSFFGNFTVEADRTWRKLEATENTKPVVLAILDTGCYLHQDFIDDFDI 180
EtSUB5          LGDPFQMRQWYLHSFFGNFTVEADRTWRKLEATENTKPVVLAILDTGCYLHQDFIDDFDI 149
                ************************************************************

Database        AKSIFWDNAGETDCSDGIDNDGNGYIDDCFGWNFVEDNGHPFTDDSGHGTSVTSVAAARA 240
EtSUB5          AKSIFWDNAGETDCSDGIDNDGNGYIDDCFGWNFVEDNGHPFTDDSGHGTSVTSVAAARA 209
                ************************************************************

Database        HDGKGGRGVLPNPTVMCLRVGSERGVWTSATIPALDYAVKMKARVSNHSYGGPG------ 294
EtSUB5          HDGKGGRGVLPNPTVMCLRVGSERGVWTSATIPALDYAVKMKARVSNHSYGGPGFVSAEY 269
                ****************************************************

Database        ------------------------------------------------------------
EtSUB5          EAFLRALRHDHLIVTAAGNAGCNIDENERCYFTPGAFRLRGLVNVMASDISGFRASFSNY 329

Database        ------------------------------------------------------------
EtSUB5         GSRSVDVAAPGTLIYAGTNYRANSGNRRCTSCYTYSDGTSFAAPIVAGMAAALWGYFERT 389

Database        ------------------------------------------------------------
EtSUB5          NPIGWQESTEPASRKVEQAIMYSVTGSKALAGAVGTNGVVNLWRALSYYDTVPGPFTPEY 449

Database        ----------------------------------------
EtSUB5          PTPESGYNPTYTGGRPQCTQYMVLLLMLWLTGTATTTAYIL 490
```

Figure 3.14: ClustalW analysis of the amino acid sequence of *Et*SUB5, predicted from 3` RACE PCR, versus the amino acid sequence predicted by the *E. tenella* genome database

A ClustalW analysis was performed on the amino acid sequence predicted for the 3` region of *etsub5* sequenced with RACE PCR against the predicted amino acid sequence for ETH_00011340 in the *E. tenella* genome database. The residues of the serine active site are highlighted in yellow. Conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids are shown. The aspartic active site of the subtilase family with the aspartic acid residue '**D**' is highlighted (blue); the histidine active site is highlighted (green) with the histidine residue '**H**'; and the serine active site is highlighted (yellow), showing the serine residue '**S**'.

confirm that the *etsub5* sequence obtained was correct, new primers were designed based on the new sequence. Sequencing of the PCR product matched the 3` region obtained from RACE PCR (data not shown).

3.3.4.3 Sequence analysis of 5` region of *etsub5*

3.3.4.3.1 Bioinformatics analysis of 5` sequence of *etsub5*

Analysis of the *E. tenella* genome database predicted that 31 amino acids (93 bp) were missing from the 5' end of *etsub5*. Following the failure of RACE PCR of the 5' end to yield any products, a gene-specific primer, ATGAGGTATTTAGAAGGTCCATCGATG, was designed from the putative start codon for ETH_00011340 to amplify the 5` region with the previously designed 5` GSP reverse primer (Section 3.3.4.1). However, despite numerous attempts to optimise annealing temperatures and PCR cycling conditions, as well as the design of a second gene-specific primer, GGCAGCGTATTATGGAACTCTT, based on the database-predicted sequence, there was still no amplification of a product. This suggested that there were mis-predictions in the coding sequence (and consequent predicted amino acid sequence) of the gene from the original genomic sequence and another strategy was needed to obtain the 5' sequence of *etsub5*.

Therefore, 2kb of genomic sequence upstream from the previously confirmed *etsub5* sequence, containing the, as yet undetermined, 5` region of the gene was downloaded from the *E. tenella* genome database (Section 2.2.6.5) and used to BLAST the *E. maxima* genome database (Section 2.2.6.5). This was performed to determine if a homologue was present in *E. maxima* that could potentially provide more information on the 5` sequence of *etsub5*. An nBLASTx was performed on the 454 + Sanger contigs database of the *E. maxima* genome and results revealed a region in the genomic sequence of *etsub5* that was highly homologous to a predicted sequence in the *E. maxima* gene Contig_01404 with a homology score of $3e^{-82}$ (Figure 3.15). Interestingly, it was not part of the 5` sequence of ETH_00011430 that was predicted by the *E. tenella* genome database. Therefore, the predicted intron and exon predictions of *etsub5* may be incorrect, and the genomic sequence that is homologous to a similar gene in *E. maxima* may, indeed, code for the sequence of the subtilase *etsub5*, which is likely to be conserved across the two species.

To further analyse the possible 'missing' sequence in the 5` region of *etsub5* BLASTp was performed with the known *etsub5* protein sequence with both the *T. gondii* and *N. caninum* database sequences to identify the subtilase orthologues in both apicomplexan parasites. Putative subtilase family serine protease homologues were identified in both *T. gondii* (TGME49_048470) and *N. caninum* (NCLIV_064430). ClustalW was then performed to analyse

the 5` region of each of these genes against the putative 5` region of *Et*SUB5 taken from genomic sequence that was identified as highly homologous to the previously identified *E. maxima* sequence. Figure 3.16 shows the 5` sequence of both *T. gondii* and *N. caninum* homologues compared with the putative 5` region of *etsub5*. A high level of homology can be seen between the subtilases from *T. gondii* and *N. caninum* and the putative 5` region of *Et*Sub5 and it appears that the 5` sequence is conserved between different coccidian parasites. Analysis also showed that there were 129 extra amino acids in the 5` region of the *T. gondii* subtilase, and 61 extra amino acids in *N. caninum* compared with *E. tenella etsub5* putative 5` sequence. It was also noted that the two homologous genes in *T. gondii* and *N. caninum* showed a high level of homology further upstream (Figure 3.16) suggesting that there was more sequence in the 5` region of *etsub5* than had previously been revealed.

Given the evidence of high homology in the 5` region amongst various coccidians, it is likely that the genomic sequence of *etsub5*, previously thought to be the 5` sequence and genomic promoter of *etsub5,* is part of the 5` coding sequence. Since it was unclear where the exons that form the coding sequence of *etsub5* were located, Lasergene™ 10 was used to analyse the three different open reading frames available from the genomic sequence. The protein sequences from the two homologous subtilases from *T. gondii* and *N. caninum* were then used to manually identify any conserved residues in the three translated sequences of the *etsub5* putative 5` region. Several motifs were conserved and, therefore, thought to be part of the coding sequence of *etsub5*. Sequence upstream of these motifs and the putative 5` region of *etsub5*, was examined and several possible methionine start residues were identified (Figure 3.17). Gene specific primers (as described in Section 2.2.4.1), *etsub5* genomic P1 CCATGGCAAGGAAGATGATACGGC, *etsub5* genomic P1 AAGGAGAATGTACGCTCTGAGACTG and *etsub5* genomic P1 CATCACATTTCGCTGTATGCAGA, were then designed using the genomic sequence of *etsub5* that included the bases coding for each methionine, to try and identify the start methionine of *etsub5*.

```
                                                                  Score    E
Sequences producing significant alignments:                       (bits) Value

contig01404  length=22504   numreads=993                              258   3e-82


>contig01404  length=22504   numreads=993
          Length = 22504

 Score =  258 bits (557), Expect(2) = 3e-82
 Identities = 104/139 (74%), Positives = 123/139 (88%)
 Frame = +3 / +3


Query: 1659  KKQSRLIVAWDPKCLQGLPKKVPPSLHELQSKETFPLQVFGARDSSLQGRVVLLRGGVYM 1838
             KK SRLI+AW+P CL+ LP+KVP  L     +ET PLQVFG +DS+LQGR VLLRGGVYM
Sbjct: 19341 KKASRLILAWNPGCLKTLPQKVPAVLQAHSVRETVPLQVFGEKDSALQGRKVLLRGGVYM 19520


Query: 1839  QDRVGANELQEYSTREVVNGRHLALWVFARALGLHAKLGLSLTGCLANCRDRLRMRYLEG 2018
             QDRVGAN+LQE STR +VNGRHLALW+FARALGLHA+LGL+LTGC+ANCRDRLR+R+L+G
Sbjct: 19521 QDRVGANDLQEESTRNIVNGRHLALWLFARALGLHARLGLTLTGCMANCRDRLRLRFLQG 19700


Query: 2019  PSMKAAYYGTLSTDILRIS 2075
             PSM+AAYYG L+TD+LRIS
Sbjct: 19701 PSMRAAYYGALNTDVLRIS 19757




 Score =  192 bits (414), Expect(2) = 2e-52
 Identities = 86/140 (61%), Positives = 99/140 (70%)
 Frame = -1 / -1


Query: 2077  VLMRRISVLRVP*YAAFIDGPSKYLILSRSRQFAKQPVRLKPSLACNPRARAKTQRAR*R 1898
             +LMR  SV   P*YAA +DGP K L LSRSRQ A QPV +KPSLAC PRARAK+ RAR R
Sbjct: 19759 LLMRNTSVFSAP*YAARMDGPCKNLSLSRSRQLAIQPVSVKPSLACRPRARAKSHRARWR 19580


Query: 1897  PFTTSRVEYSCSSFAPTRSCIYTPPRRSTTRPWRELSLAPNT*SGKVSLL*SSCKDGGTF 1718
             P T   V+ SC SFAPTRSC+YTPPR +T RP + LS +P T* G VS    +CK  GTF
Sbjct: 19579 PLTMFLVDSSCKSFAPTRSCMYTPPRSNTLRPCKALSFSPKT*RGTVSRTL*ACKTAGTF 19400


Query: 1717  FGRPCKHLGSHATMSRDCFF 1658
             +GR  +H G HA+MS D FF
Sbjct: 19399 WGRVFRHPGFHASMSLDAFF 19340
```

Figure 3.15: BLAST of *E. maxima* genomic database with *etsub5*

2kb genomic DNA sequence upstream from the annotated *etsub5* gene was used to BLAST the *E. maxima* genome database. A tBLASTx was performed of 454 + Sanger contig database of the *E. maxima* genome to determine if there was homologous sequence present to compare with the 5` region of *etsub5* genomic sequence. A sequence was identified in *E. maxima* on the contig01404 (Subject) that was homologous to the genomic sequence of *etsub5* (Query). Conserved residues are shown in the centre.

```
T.gondii        MYPWHSLVLLLFFIYPYVDLVVGRAAGSDKKYNGSQATQEISGFAEAPKSAEREAWISTE 60
N.caninum       ----------MFFTW------VG-ARKSD------------------------------ 12
EtSUB5          ------------------------------------------------------------

T.gondii        HLKRALCKKYYTVVLGMHPSFSVDKLDCEHRPDPEE-RIIESS-ELRFMDPAQVRQDRIE 118
N.caninum       ----------------------VDKLDCERRPGPEEERIIESAGELKPMDPAQARQDRIE 50
EtSUB5          ------------------------------------------------------------

T.gondii        ALKRLQKEVKTSPTDRVIVEWDTSCLKDLPSRVVESTPSQYDGSVPVFDEYELHDSSER- 177
N.caninum       ALKRLQKEVEASPTDRVIVEWDTTCLKDLPSRVAESTAAQDDGSVPVFDEYELHDSSER- 109
EtSUB5         -----------KKQSRLIVAWDPKCLQGLPKKVPPSLHELQSKETFPLQVFGARDSSLQG 49
                            .  .*:** **..**:.**.:*  *     . ..  :: :  :*** :

T.gondii        --IVLKRGVYLRDRIGLKDKTDPEVRQMLKGREYALGIIASALGLQVGGPGHGVPCEGQC 235
N.caninum       --IVLKRGVYLRDRIGVKDKTDPEVRQTLRGREYALGIIASALGLQVWGPASGAPCEGQC 167
EtSUB5          RVVLLRGGVYMQDRVGANELQEYSTREVVNGRHLALWVFARALGLHAKLGLSLTGCLANC 109
                  ::*: ***::**:* ::  : ..*: :.**. ** ::* ****:.      . * .:*

T.gondii        RKELQNVMLAKDVPTLEAAHYSRLGVDVLHLKNLKKHKEIWGITKASEIPAVVRRLAEKR 295
N.caninum       RKALQSVMLAKDVPTLEAAHYSRLGVDVLHLKNLRKQKE--------------------- 206
EtSUB5          RDRLR--MRYLEGPSMKAAYYGTLSTDILRIS----------------------------- 139
                *. *:  *   : *:::**:*. *..*:*::.
T.gondii        ETTSTANGERNMAASAREATSCILKMSDDQKREVQVPDIAKTPMAATGGNAIADDFVVAV 355
N.caninum       ------------------------------------------------------------
EtSUB5          ------------------------------------------------------------

T.gondii        LDTGCFIHQDFIDGFDIENSRFWINKGETDCENRIDDDGNGYVDDCFGWDFVGDQAYPFV 415
N.caninum       ------------------------------------------------DFVGDQAYPFV 217
EtSUB5         ------------------------------------------------------------

T.gondii        DDLGHGTSVTSILAARHGNNAGGRGIMSTGKVMCLRVGGRDGIWLSGTIPALDYALRMGA 475
N.caninum       DDLGHGTSVTSVLAARHGNNAGGRGIMSTGKVMCLRVGGRDGIWLSDTIPALDYALRMGA 277
EtSUB5          ------------------------------------------------------------

T.gondii        KVSNHSYGGKGFDQAEFTAFRRASEADHLAVTAAGNSGCDVDSDNTCLFTPGAFDLEGIL 535
N.caninum       KVSNHSYGGKGFDQAEFTAFRRASEADHLAVTAAGNSGCDVDSDDTCLFTPGAFDLEGIL 337
EtSUB5          ------------------------------------------------------------

T.gondii        NVGANDVTGHRATFSNYGKNAVDVFAPGQSIFVGTNKMPNQVEKTCKYCYEFIDGTSFAA 595
N.caninum       NVGANDVTGHRATFSNYGKNAVDVFAPGQSIFVGTNKMPNQVEKTCKYCYEFIDGTSFAA 397
EtSUB5          ------------------------------------------------------------
```

Figure 3.16: ClustalW alignment of *Et*SUB5 and homologues in *T. gondii and N. caninum*

ClustalW was performed to align the putative 5` region of *Et*SUB5, previously identified as homologous to *E. maxima*, with subtilase orthologues in *T. gondii* and *N. caninum.* A high level of conservation was observed between the translated genomic sequence of *etsub5,* recognised as the putative 5` region of the *etsub5* gene; and the two subtilase sequences from *T. gondii* and *N. caninum*. Conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids are shown, with the predicted sequence encoded by the 5` sequence of *etsub5* highlighted (green).

```
CCAAACTGTAGGTCAGAACATTCTGTGGTAGTCTTGCCAGGACTTGAGCATCTTAGTTTTCCCTTATCTTCCCATCCCAATGCCGCTTCCCATAAGCGCCATCACATGCA
  Q  T  V  G  Q  N  I  L  W  .  S  C  Q  D  L  S  I  L  V  F  P  Y  L  P  I  P  M  P  L  P  I  S  A  I  T  C
T  K  L  .  V  R  T  F  C  G  S  L  A  R  T  .  A  S  .  F  S  L  I  F  P  S  Q  C  R  F  P  .  A  P  S  H  A
 P  N  C  R  S  E  H  S  V  V  V  L  P  G  L  E  H  L  S  F  P  L  S  S  H  P  N  A  A  S  H  K  R  H  H  M  H

CAATCACTAGGCAGAACTGAGTACAGCGGGGCATCACATTTCGCTGTATGCAGAAAAAGAGTAAATATTTCAACAGATTTCTTGGAAAGGAGAATGTACGCTCTGAGACT
                                  etsub5 genomic P3 Tm67                                  etsub5 genomic P2 Tm65
T  I  T  R  Q  N  .  V  Q  R  G  I  T  F  R  C  M  Q  K  K  S  K  Y  F  N  R  F  L  G  K  E  N  V  R  S  E  T
 Q  S  L  G  R  T  E  Y  S  G  A  S  H  F  A  V  C  R  K  R  V  N  I  S  T  D  F  L  E  R  R  M  Y  A  L  R  L
  N  H  .  A  E  L  S  T  A  G  H  H  I  S  L  Y  A  E  K  E  .  I  F  Q  Q  I  S  W  K  G  E  C  T  L  .  D

GTGCGCATTGAAGGCAATCGGGGTGCTGCTTCTTGCCAAGGCGTCCTCTTGTGGCTCCAACGAGGTCACTGAGCAAACGCCGCTTCCTAGTAGTGAAACCTTACGGCGAG
 V  R  I  E  G  N  R  G  A  A  S  C  Q  G  V  L  L  W  L  Q  R  G  H  .  A  N  A  A  S  .  .  .  N  L  T  A  R
  C  A  L  K  A  I  G  V  L  L  L  A  K  A  S  S  C  G  S  N  E  V  T  E  Q  T  P  L  P  S  S  E  T  L  R  R
C  A  H  .  R  Q  S  G  C  C  F  L  P  R  R  P  L  V  A  P  T  R  S  L  S  K  R  R  F  L  V  V  K  P  Y  G  E

AGCATCTGCGCACATCCCTCTGCGCCAGCTTTTACCTTCAGGAGCGTAGGCAACGGCTTCCTCAGGACCCTGAACAGCGCAGGAAAAGATGCGGTTCGGGAGCATACAAG
  A  S  A  H  I  P  L  R  Q  L  L  P  S  G  A  .  A  T  A  S  S  G  P  .  T  A  Q  E  K  M  R  F  G  S  I  Q
E  H  L  R  T  S  L  C  A  S  F  Y  L  Q  E  R  R  Q  R  L  P  Q  D  P  E  Q  R  R  K  R  C  G  S  G  A  Y  K
 S  I  C  A  H  P  S  A  P  A  F  T  F  R  S  V  G  N  G  F  L  R  T  L  N  S  A  G  K  D  A  V  R  E  H  T  S

CACCATGGCAAGGAAGATGATACGGCTACTTACGGCCCCCTCAAGACCAGAAGTAATTTTCGAGAGGCAAGCATCTGGCCTCCTTTCAGCTGGCAACATTTATCAGGAAG
  etsub5 genomic P1 Tm 67
A  P  W  Q  G  R  .  Y  G  Y  L  R  P  P  Q  D  Q  K  .  F  S  R  G  K  H  L  A  S  F  Q  L  A  T  F  I  R  K
 H  H  G  K  E  D  D  T  A  T  Y  G  P  L  K  T  R  S  N  F  R  E  A  S  I  W  P  P  F  S  W  Q  H  L  S  G  S
  T  M  A  R  K  M  I  R  L  L  T  A  P  S  R  P  E  V  I  F  E  R  Q  A  S  G  L  L  S  A  G  N  I  Y  Q  E

TATTCTGCGCAGCCAAAATGGCAAAAGCTGTGTTTTCCTCTTTTTCGTGTTTCTCGTTGCAAATCGCAACAAACAGACCCACGAAGACCTCGTCATTTTCTCCATTTAAC
 Y  S  A  Q  P  K  W  Q  K  L  C  F  P  L  F  R  V  S  R  C  K  S  Q  Q  T  D  P  R  R  P  R  H  F  L  H  L  T
  I  L  R  S  Q  N  G  K  S  C  V  F  L  F  F  V  F  L  V  A  N  R  N  K  Q  T  H  E  D  L  V  I  F  S  I  .
V  F  C  A  A  K  M  A  K  A  V  F  S  S  F  S  C  F  S  L  Q  I  A  T  N  R  P  T  K  T  S  S  F  S  P  F  N

AATTTTGGTGCAGGACCTCGCAGATGTGTGGGGGAAGCAAGAACATGGAAAAAAGCAGTCCCGACTCATAGTTGCGTGGGATCCTAAGTGCTTGCAAGGCCTGCCAAAGA
                                                                                  5' sequence from MAXIMA
  I  L  V  Q  D  L  A  D  V  W  G  K  Q  E  H  G  K  K  Q  S  R  L  I  V  A  W  D  P  K  C  L  Q  G  L  P  K
Q  F  W  C  R  T  S  Q  M  C  G  G  S  K  N  M  E  K  S  S  P  D  S  .  L  R  G  I  L  S  A  C  K  A  C  Q  R
 N  F  G  A  G  P  R  R  C  V  G  E  A  R  T  W  K  K  A  V  P  T  H  S  C  V  G  S  .  V  L  A  R  P  A  K  E

AAGTGCCGCCGTCCTTGCATGAGCTTCAAAGTAAGGAAACCTTTCCGCTTCAGGTGTTTGGTGCAAGAGACAGCTCCCTCCAGGGCCGCGTTGTGCTTCTTCGTGGTGGA
                                                                                  5' sequence from MAXIMA
K  V  P  P  S  L  H  E  L  Q  S  K  E  T  F  P  L  Q  V  F  G  A  R  D  S  S  L  Q  G  R  V  V  L  L  R  G  G
 K  C  R  R  P  C  M  S  F  K  V  R  K  P  F  R  F  R  C  L  V  Q  E  T  A  P  S  R  A  A  L  C  F  F  V  V  E
  S  A  A  V  L  A  .  A  S  K  .  G  N  L  S  A  S  G  V  W  C  K  R  Q  L  P  P  G  P  R  C  A  S  S  W  W

GTATATATGCAAGACCGCGTGGGAGCAAACGAGCTGCAGGAATATTCGACCCGTGAGGTCGTAAACGGGCGTCACCTTGCCCTCTGGGTTTTTGCACGAGCTCTAGGGTT
                                                                                  5' sequence from MAXIMA
 V  Y  M  Q  D  R  V  G  A  N  E  L  Q  E  Y  S  T  R  E  V  V  N  G  R  H  L  A  L  W  V  F  A  R  A  L  G  L
  Y  I  C  K  T  A  W  E  Q  T  S  C  R  N  I  R  P  V  R  S  .  T  G  V  T  L  P  S  G  F  L  H  E  L  .  G
S  I  Y  A  R  P  R  G  S  K  R  A  A  G  I  F  D  P  .  G  R  K  R  A  S  P  C  P  L  G  F  C  T  S  S  R  V

GCATGCCAAGCTAGGCTTGAGCCTAACAGGTTGTTTAGCGAACTGCCGGGACCGGCTAAGAATGAGGTATTTAGAAGGTCCATCGATGAAGGCAGCGTATTATGGAACTC
                                                                                  5' sequence from MAXIMA
  H  A  K  L  G  L  S  L  T  G  C  L  A  N  C  R  D  R  L  R  M  R  Y  L  E  G  P  S  M  K  A  A  Y  Y  G  T
C  M  P  S  .  A  .  A  .  Q  V  V  .  R  T  A  G  T  G  .  E  .  G  I  .  K  V  H  R  .  R  Q  R  I  M  E  L
 A  C  Q  A  R  L  E  P  N  R  L  F  S  E  L  P  G  P  A  K  N  E  V  F  R  R  S  I  D  E  G  S  V  L  W  N  S

TTAGCACAGATATTCTTCGCATCAGCACCAGCACTGTGAGCATATATCTTGA
  5' sequence from MAXIMA
L  S  T  D  I  L  R  I  S  T  S  T  V  S  I  Y  L  D
 L  A  Q  I  F  F  A  S  A  P  A  L  .  A  Y  I  L
  .  H  R  Y  S  S  H  Q  H  Q  H  C  E  H  I  S  .
```

Figure 3.17: Genomic sequence of *etsub5* with three alternative translations and primers designed to amplify each option

A 2kb genomic DNA sequence of the 5` region upstream of *etsub5* was analysed with bioinformatics including Lasergene™10 and ExPASy. Three alternative protein translations were identified which included three putative start methionines in each of the protein translation possibilities. The putative sequence homologous to *E. maxima*, initially used to identify the genomic sequence of *etsub5* as possible coding sequence, is also shown. Three primers were designed upstream from this putative sequence over three likely start methionines of the *etsub5* gene – *etsub5* genomic P1, *etsub5* genomic P2 and *etsub5* genomic P3.

3.3.4.3.2 Amplification of 5` region of *etsub5*

The previously designed 5` gene specific reverse primers (Section 3.3.4.1) were combined with each of the primers *etsub5* genomic P1, *etsub5* genomic P2 and *etsub5* genomic P3 and used to amplify cDNA at an optimal annealing temperature of 60°C. Results are shown in Figure 3.18. A band approximately 900bp in size was observed in the amplification reaction with primers *etsub5* genomic P3 and 5` GSPR (Figure 3.18, lane 4). A smaller band approximately 650-700bp in size was also observed in the reaction with primers *etsub5* genomic P1 and 5` GSPR (Figure 3.18, lane 3). No band was observed with the *etsub5* genomic P2 reactions. Both bands were excised from the gel, purified and cloned into pGEM® T-Easy. Plasmids were then sequenced with vector specific M13 primers and sequencing analysis was performed using Seqman program (Lasergene™ 10).

3.3.4.3.2 Analysis of 5` sequence of *etsub5*

Sequencing confirmed that the 5` region of *etsub5*, including the start methionine (Figure 3.19), was amplified by using both the *etsub5* P3 primer and 5'GSPR and the *etsub5* genomic P1 primer and 5' GSPR although more sequence was revealed by the longer product obtained with the *etsub5* genomic P3 and 5'GSPR pair (Figure 3.19).

The gene sequence of the 5' region of *etsub5* was analysed in Lasergene ™ 10 and combined with earlier sequencing results to manually construct the full-length gene sequence. The full-length *etsub5* gene sequence was found to be 2244bp in length, much longer than the 885bp predicted by the *E. tenella* genomic database. The gene sequence was then translated into the corresponding amino acid sequence, resulting in a protein that was 748 amino acids in length. InterproScan (version 4.8) (Section 2.2.6.4) online tool was used to scan the *etsub5* protein sequence for a range of protein signatures to identify the different protein families as well as any functional or structural domains. Analysis revealed the protein was classified as a Peptidase S8/S53 subtilisin-like protein as it contained the aspartic acid, histidine and serine active sites of the catalytic triad of subtilases, and the conserved oxyanion hole asparagine residue (Siezen and Leunissen, 1997a). The aspartic acid active site was between amino acids 388-398 with the conserved aspartic acid residue at position 392; the histidine active site was between amino acids 455-465 with the conserved histidine residue at position 455; the oxyanion hole asparagine motif was between amino acids 543-546 with the conserved asparagine residue at position 546; and the serine active site was between 625-635, with the conserved serine residue at position 627 (Figure 3.19).

After identification as a subtilase serine protease, *etsub5* was also analysed with SignalP 4.0 Server to predict the presence and location of a signal peptide in the amino acid sequences. Signal peptides are present in all Apicomplexa subtilases that have been characterised to date (Barale *et al.*, 1999b, Blackman *et al.*, 1998a, Cevallos *et al.*, 2000, Hackett *et al.*, 1999b, Louie and Conrad, 1999, Miller *et al.*, 2001b, Miller *et al.*, 2003a, Montero *et al.*, 2006b). A signal peptide sequence was identified in the *Et*SUB5 amino acid sequence with the cleavage site between amino acids 22 and 23, confirming the start methionine as the true start codon (Figure 3.19). This was further confirmed in analysis with homologous subtilase signal peptides, where the numbers of amino acids in the protein sequence of signal peptides in homologous Apicomplexa were identical to that of *etsub5*.

Subtilase enzymes also often possess a propeptide sequence in their amino acid protein sequence. Post-translational processing at the N-terminal of the sequence occurs to produce a biologically active subtilase enzyme (Siezen and Leunissen, 1997a). ProP 1.0b server (integrated with the SignalP online program) (Duckert *et al.*, 2004) was used to determine if there was a propeptide sequence present in the *etsub5* protein sequence. A peptide cleavage site was predicted at position 72 of the protein, with the sequence PEQRRKR/CG recognised as the cleavage motif for this subtilase (Figure 3.20).

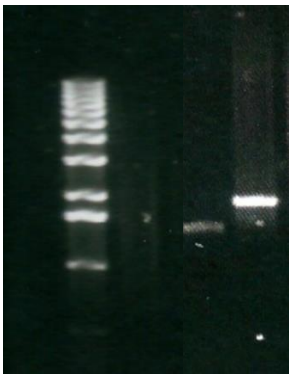


Figure 3.18: PCR of 5` sequence with *etsub5* P3 GSP and *etsub5* P3 GSP

E. tenella cDNA was PCR amplified with gene specific primers, *etsub5* genomic P3 and *etsub5* 5` RACE, and *etsub5* genomic P1 and *etsub5* 5` RACE primer to amplify the 5` region of *etsub5*. The PCR was cycled as described in Section 2.2.4.2, with the Tm annealing at 60°C. Products of 900bp and 700bp, respectively, were produced. PCR products were electrophoresed in a 1% agarose gel and visualised using Gel Red® and VersaDoc® Imaging system (Section 2.2.4.5). PCR was repeated with the *etsub5* genomic P3 primer and 5' GSPR pair and the product cloned and sequenced again to confirm the correct 5` sequence of *etsub5*.

```
MYALRLCALKAIGVLLLAKASSCGSNEVTEQTPLPSSETLRQEHLRTSLCASFYLQERRQRLPQDPEQRRKRCGSGAYKH
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  80
```

Signal Peptide

Propeptide

```
HGKEDDTATYGPLKTRSNFREDLADVWGKQEHGKKQSRLIVAWDPKCLQGLPKKVPPSLHELQSKETFPLQVFGARDSSL
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  160

QGRVVLLRGGVYMQDRVGANELQEYSTREVVNGRHLALWVFARALGLHAKLGLSLTGCLANCRDRLRMRYLEGPSMKAAY
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  240

YGTLSTDILRISTSTSDQPSLDFLPKIREVAKSLSSGTDEVAGALGRVITAAVACVDKISLDRKVDVQRLGSLATNSQGN
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  320

MSLQEPIKGTVNDGATSPQLLRAPSGLLGDPFQMRQWYLHSFFGNFTVEADRTWRKLEATENTKPVVLAILDTGCYLHQD
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  400
```

Aspartic Acid

```
FIDDFDIAKSIFWDNAGETDCSDGIDNDGNGYIDDCFGWNFVEDNGHPFTDDSGHGTSVTSVAAARAHDGKGGRGVLPNP
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  480
```

Histidine

```
TVMCLRVGSERGVWTSATIPALDYAVKMKARVSNHSYGGPGFVSAEYEAFLRALRHDHLIVTAAGNAGCNIDENERCYFT
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  560
```

OHA

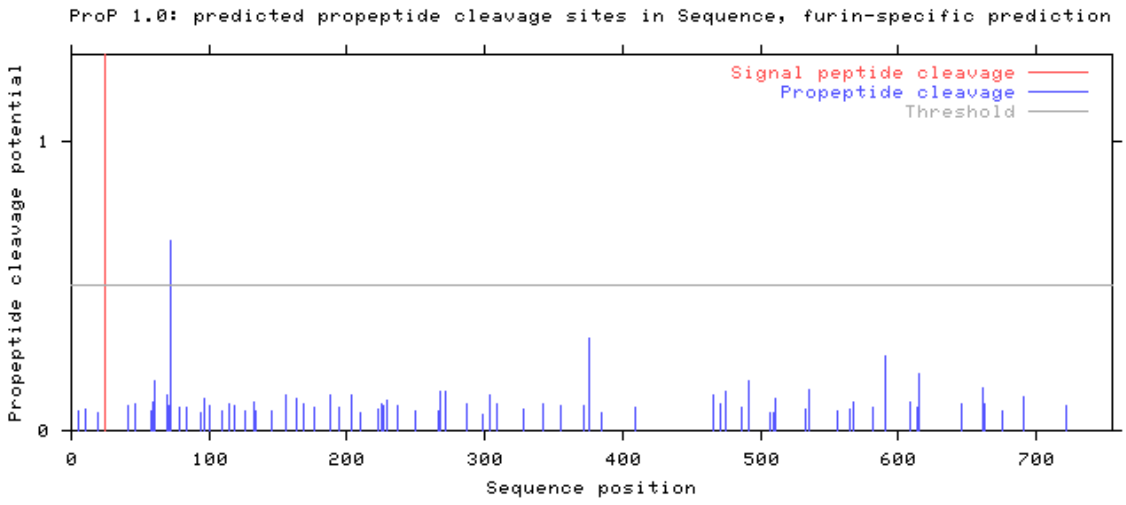
```
PGAFRLRGLVNVMASDISGFRASFSNYGSRSVDVAAPGTLIYAGTNYRANSGNRRCTSCYTYSDGTSFAAPIVAGMAAAL
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  640
```

Serine

```
WGYFERTNPIGWQESTEPASRKVEQAIMYSVTGSKALAGAVGTNGVVNLWRALSYYDTVPGPFTPEYPTPESGYNPTYTG
++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|++++|  720

GRPQCTQYMVLLLMLWLTGTATTTAYIL.
++++|++++|++++|++++|++++|++++
```

Figure 3.19: Full length protein sequence of subtilase *Et*SUB5

The full-length gene sequence of *Et*SUB5 was obtained through PCR and translated using Lasergene (Version 10) into the 748 amino acid protein sequence. The start methionine, signal peptide (1-22), pro-peptide region (1-72), aspartic acid active site motif (388-398), histidine active site motif (455-465), oxyanion hole asparagine motif (OHA) (543-546), serine active site motif (625-635), and stop codon of the protein are shown.




```
Measure   Position   Value     Cutoff    signal peptide?

 max. C     26        0.242

 max. Y     23        0.446

 max. S     11        0.919

 mean S      1-22     0.833

   D      1-22     0.655    0.450    YES

Propeptide cleavage sites predicted:    Arg (R)/Lys (K): 1

Sequence     PEQRRKR/CG     Score: 0.656 *ProP*

Name=Sequence SP='YES' Cleavage site between pos. 22 and 23: ASS-CG
D=0.655 D-cutoff=0.450 Networks=SignalP-noTM
```

Figure 3.20: SignalP and ProP prediction of signal peptide and propeptide cleavage sites in subtilase *Et*SUB5

SignalP 4.0 was used to analyse the translated amino acid sequence of *Et*SUB5. A signal peptide was predicted for subtilase *Et*SUB5 between amino acids 1-22, with a cleavage score of 0.655, greater than the 0.450 cutoff, predicting cleavage between amino acids 22 and 23 of the protein sequence. ProP 1.0b was then used to analyse the sequence for a propeptide. A propeptide cleavage site was predicted at position 72 in the protein sequence with the cleavage site PEQRRKR/CG recognised.

3.3.4.3.3 Homology analysis of *Et*SUB*5* to other apicomplexan subtilases

ClustalW was performed to compare the *Et*SUB5 protein with known subtilase sequences in other apicomplexan parasites, including *T. gondii, N. caninum, C. parvum* and *P. falciparum.* There was a high level of conservation of the subtilases around the three active sites of the catalytic triad (Siezen and Leunissen, 1997a), with the aspartic acid, histidine and serine active site residues conserved in all the Apicomplexa sequences (Figure 3.21). Also conserved between the sequences was the oxyanion hole asparagine, which is structurally required for correct folding and activity of subtilase proteins (Siezen and Leunissen, 1997a). There was a low level of homology between the sequences at the 3` region of the subtilases downstream from the serine active site. There also appears to be a lot more sequence found in *Cp*Sub1 than in other apicomplexan subtilases in this region. This is represented diagrammatically in Figure 3.22, which is a scale diagram showing that the catalytic triad appears to be not only highly conserved, but also structurally similar, with the same number of amino acids separating the active sites and the oxyanion hole asparagine. This conserves the 2D and 3D structure of the active site between the sequences from different Apicomplexa. A high level of homology also exists between the signal peptide present in all the subtilase protein sequences, with all subtilases having the same size signal peptide (22 amino acids) except for *Pf*SUB1 (Figure 3.22).

Figure 3.21: ClustalW alignment of *E. tenella etsub5* with previously identified apicomplexan subtilases

ClustalW was used to perform a multiple sequence alignment of apicomplexan subtilases that have been sequenced and characterised previously in T*oxoplasma gondii, Plasmodium falciparum, Neospora caninum* and *Cryptosporidium parvum*. Sequences were then analysed in Jalview to observe similarities/differences between the sequences. Jalview highlighted the protein features that were: aliphatic/hydrophobic residues (pink); aromatic (orange); positive (red); negative (green); hydrophilic (blue); conformational special proline/glycine (magenta); or cysteines (yellow).

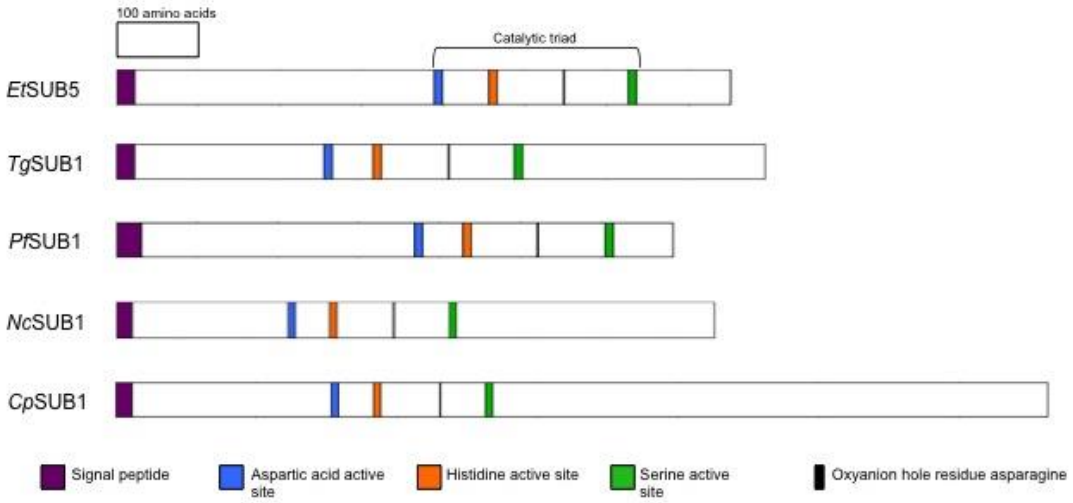


Figure 3.22: Conservation of *EtSub5* and other apicomplexan subtilases

Locations of the active site motifs, as well as the signal peptide and oxyanion hole residue asparagine are illustrated in a schematic diagram. Locations of active sites motifs – aspartic acid (blue), histidine (orange) and serine (green) – as well as the signal peptide (purple) and oxyanion hole asparagine motif (black) are shown. The diagram is drawn to scale.

3.3.5 Amino acid sequence analysis of subtilases

Amino acid sequences of *Et*SUB1, *Et*SUB2 and *Et*SUB5 were analysed with Clustal W to compare the similarities and differences. A high level of homology was observed between the three sequences with identical ('*'), highly similar (':') and similar ('.') functions of the amino acid residues shown in Figure 3.23. The histidine residue (H) of the active site was conserved between *Et*SUB1, *Et*SUB2 and *Et*SUB5 (Figure 3.23). Further, half of the residues in the histidine active site were conserved (*) between the three proteins, whilst most of the non-conserved residues consisted of highly similar (':') residues so the functional structure of the active site remained. Scanprosite identified the histidine active site of a subtilase in *Et*SUB1 and *Et*SUB5 although, despite the conservation, not *Et*SUB2. The aspartic acid active site (highlighted in blue), including the functional aspartic acid residue, was conserved between all three sequences (Figure 3.23). The serine residue (S) of the active site was identified in *Et*SUB5, however, was not identified in *Et*SUB1 or *Et*SUB2 as the 3` regions of these two sequences could not be sequenced.

```
EtSUB2          ------------------------------------------------------------
EtSUB5          MYALRLCALKAIGVLLLAKASSCGSNEVTEQTPLPSSETLRQEHLRTSLCASFYLQERRQ 60
EtSUB1          ------------------------------------------------------------

EtSUB2          ------------------------------------------------------------
EtSUB5          RLPQDPEQRRKRCGSGAYKHHGKEDDTATYGPLKTRSNFREDLADVWGKQEHGKKQSRLI 120
EtSUB1          ------------------------------------------------------------

EtSUB2          ------------------------------------------------------------
EtSUB5          VAWDPKCLQGLPKKVPPSLHELQSKETFPLQVFGARDSSLQGRVVLLRGGVYMQDRVGAN 180
EtSUB1          ------------------------------------------------------------

EtSUB2          ------------------------------------------------------------
EtSUB5          ELQEYSTREVVNGRHLALWVFARALGLHAKLGLSLTGCLANCRDRLRMRYLEGPSMKAAY 240
EtSUB1          ------------------------------------------------------------

EtSUB2          ------------------------------------------------------------
EtSUB5          YGTLSTDILRISTSTSDQPSLDFLPKIREVAKSLSSGTDEVAGALGRVITAAVACVDKIS 300
EtSUB1          ------------------------------------------------------------

EtSUB2          -------------------------------------------AAANVEDPYFPLQWHLQ 17
EtSUB5          LDRKVDVQRLGSLATNSQGNMSLQEPIKGTVNDGATSPQLLRAPSGLLGDPFQMRQWYLH 360
EtSUB1          ------------------------------------------------------------

EtSUB2          AAEGLYNLGADKAWQQMQSAQQQQPQQQQVVVALLDTGCGPHEDINN-------ILWINS 70
EtSUB5          SFFGNFTVEADRTWRKLEATENTKP----VVLAILDTGCYLHQDFIDDFDIAKSIFWDNA 416
EtSUB1          ----------------------------SVPVAVVDTGINYLHPELS------LSMWVNR 26
                                             * :*::***    .   .        :* *

EtSUB2          KE-NCSDGVDHDNNGYVNDCKGWDFVDDT-NDITDTYGHGTAVAALMAA-SFDKKGGRGV 127
EtSUB5          GETDCSDGIDNDGNGYIDDCFGWNFVEDNGHPFTDDSGHGTSVTSVAAARAHDGKGGRGV 476
EtSUB1          KELHGEEGVDDDGNGFVDDLFGWNFIQDNNNPMDDN-GHGSHVAGIVAALQNNGEGISGI 85
                 * . .:*:*.*.**:::*  **:*::*. : : *  ***: *:.: **   : :*  *:

EtSUB2          SPTGRVMCLRVGDQGG-VQLSRQLMAMDYAIQQGAQVSVHPFSLATENEVYREAFEKLLS 186
EtSUB5          LPNPTVMCLRVGSERG-VWTSATIPALDYAVKMKARVSNHSYGGPGFVSAEYEAFLRALR 535
EtSUB1          SERAKVMALKILDQKGEGDVSHAIPAIQYAVDNGAKVLTNSWG----------------- 128
                     **.*:: .: *    *  : *::**:.  *:*  :.:.

EtSUB2          SNHLAVVAAGDNGCDLDEDT-CKSYPASFKIP---------------------------E 218
EtSUB5          HDHLIVTAAGNAGCNIDENERCYFTPGAFRLRGLVNVMASDISGFRASFSNYGSRSVDVA 595
EtSUB1          ------------------------------------------------------------

EtSUB2          PSGLLVVGSSTY------------------------------------------------ 230
EtSUB5          APGTLIYAGTNYRANSGNRRCTSCYTYSDGTSFAAPIVAGMAAALWGYFERTNPIGWQES 655
EtSUB1          ------------------------------------------------------------

EtSUB2          ------------------------------------------------------------
EtSUB5          TEPASRKVEQAIMYSVTGSKALAGAVGTNGVVNLWRALSYYDTVPGPFTPEYPTPESGYN 715
EtSUB1          ------------------------------------------------------------

EtSUB2          ---------------------------------
EtSUB5          PTYTGGRPQCTQYMVLLLMLWLTGTATTTAYIL 748
EtSUB1          ---------------------------------
```

Figure 3.23: ClustalW alignement of *E. tenella* subtilases *Et*SUB1, *Et*SUB2 and *Et*SUB5 amino acid sequences

ClustalW was performed for analysis of *Et*SUB1 with *Et*SUB2 and *Et*SUB5 to compare the sequences of all three following completion of RACE PCR and bioinformatics analysis. The aspartic active site of the subtilase family with the aspartic acid residue '**D**' is highlighted (blue); the histidine active site is highlighted (green) with the histidine residue '**H**'; and the third serine active site is highlighted (yellow), showing the serine residue '**S**'. There is a high level of conservation between the sequence at the active sites in each subtilase with conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids identified.

3.4 Discussion

Six putative subtilase serine proteases were identified previously in the *E. tenella* genome using bioinformatic analysis (Katrib et al. 2012). This chapter describes the determination of stage-specific expression of these putative subtilases, the underlying assumption being that significant up-regulation of expression in gametocytes is a pre-requisite for them to have any potential role in the proteolytic cleavage of GAM56 and/or GAM82 as part of the process of oocyst wall formation. Those serine proteases with up-regulated expression in gametocytes were further characterised to confirm their classification as subtilases.

The time line of events of oocyst wall formation is important when analysing stage-specific expression while investigating the oocyst wall of *Eimeria,* and its development. In this study, gametocyte-specific expression was investigated using two PCR methods. Initially, stage-specific expression in the six putative subtilases was analysed by semi-quantitative analysis with *etsub1*, *etsub2* and *etsub5* showing gametocyte-specific expression (Figure 3.1). Expression of *etsub3* and *etsub4* was also detected in merozoites and unsporulated oocysts, respectively, while *etsub6* failed to amplify at all (Figure 3.1) and, so, was not investigated further. Semi-quantitative RT-PCR was followed by qRT-PCR to confirm the results and to further investigate the relative transcript abundance of each of these genes in *E. tenella.* Up-regulation or specificity of expression in gametocytes was confirmed for *etsub*1 (Figure 3.4(a)), *etsub2* (Figure 3.4(b)), and *etsub5* (Figure 3.4(e)), but, whilst gametocyte expression was also observed for *etsub3* (Figure 3.4(c)), and *etsub4* (Figure 3.4(d)), the highest relative transcript levels for these genes were observed in unsporulated and sporulated oocysts, respectively. Both of these genes were classified as subtilases, given the identification of active sites characteristic of subtilase-like serine protease genes (the aspartic acid and histidine active sites) but they were not investigated further as a significant role for them in processing of GAM56 or GAM82 in the formation of the oocyst wall appeared unlikely given their gene expression profiles.

The expression profile of *etsub2* was quite definitive, with gametocyte-specific expression observed in both semi-quantitative RT-PCR analysis and qRT-PCR results where transcript was detected only in early and late gametocytes. This made it a good candidate for a role in oocyst wall formation; the tight timing of expression of *etsub2* suggested that translation of the protein would coincide with the development of the oocyst wall of *E. tenella* as the macrogametocyte matures and the wall begins to form and that it would be present at an appropriate time to proteolytically cleave the oocyst wall proteins, including GAM56 and GAM82. Interestingly, the transcript levels detected for *etsub2* were 10-fold less than those levels measured for the other five subtilase genes tested. This may be indicative of the specific nature of this subtilase where it

may be speculated to act on specific single-site cleavage of GAM56 and/or GAM82 and, therefore, is not required in high abundance.

Expression of *etsub1* was observed only in gametocytes purified at 134hr p.i. by semi-quantitative RT-PCR analysis (Figure 3.1). qRT-PCR analysis confirmed that relative transcript of *etsub1* was up-regulated at this time point, however, the more detailed method highlighted that gene transcript was also significantly up-regulated in late-stage gametocytes. Transcript was also detected in unsporulated oocysts using qRT-PCR, however levels observed in this life cycle stage were much lower than in the gametocyte stages. Given the increased abundance of transcript in late stage gametocytes, the role of this subtilase may be at a later time point. Thus, the relatively late stage transcript abundance of *etsub1,* whilst not ruling out participation in processing of GAM56 or GAM82 as a part of oocyst wall formation, may suggest that this gene plays additional roles in *E. tenella,* perhaps even in sporocyst wall formation. Alternatively, it may be that the transcript is stored under a regulatory control in the parasite as is observed in other Apicomplexa. Studies in the homologous apicomplexan, *Plasmodium,* have suggested that there are regulatory events at the mRNA level of stability and/or translation whereby the transcriptome and proteome in life cycle stages of the parasite do not correlate (Le Roch *et al.*, 2004). Within this parasite, six ubiquitin hydrolases show delayed translation where transcript is detected in early gametocytes, but stored during the later gamete stage (Le Roch *et al.*, 2004). In this context, the fact that antibodies to GAM56 recognise the sporocyst wall (Belli *et al.*, 2002c) may be significant.

Similar to *etsub1* and *etsub2*, the gene expression of subtilase *etsub5* was found to be gametocyte-specific in semi-quantitative RT-PCR analysis of gametocytes purified at 134hrs p.i. qRT-PCR analysis confirmed that relative transcript of *etsub5* was significantly up-regulated at this time point, with the transcript also present in late stage gametocytes. There was also transcript present in merozoites and sporulated oocysts but the levels observed in these two life cycle stages were much lower than gametocyte stages. Given the increased abundance of the transcript in early gametocytes in both semi-quantitative and qRT-PCR analysis, *etsub5* is a legitimate candidate for a role in oocyst wall formation.

Putative sequences of the three subtilases found to be predominantly expressed in gametocytes, *etsub1, etsub2* and *etsub5,* were shown to have at least two of the three characteristic active site motifs. The aspartic acid sites first identified in the putative database sequences were confirmed in all three sequences, as was the histidine active site in *etsub1* and *etsub5* (Figure 3.23). *etsub2* was manually found to possess the highly conserved histidine residue, although Scanprosite failed to recognise it as such, despite also displaying high conservation in that region of the protein sequence (GHGT_VA). In 2007 Siezen *et al.* updated the parameters surrounding the conservation

of the catalytic triad, and found that, whilst the active site remains predominantly D-H-S, the nature of the region and position in the protein sequence of the general base (H) and acid residue (D) could be found in different combinations. Another reason may be as described by Siezen *et al.* (2007), who found that, because of the large sequence diversity among subtilases, misidentification of the catalytic triad motif by Scanprosite often leads to unidentified subtilases in many databases. Therefore, whilst the sequence motifs may not be recognised as the active site, the observation that the residues are there means that they are most probably subtilases. For each of the three genes investigated, the annotation of each gene in the *E. tenella* genomic database was identified as incorrect, previously identified as a problem in *de novo* gene prediction, as was used for the *E. tenella* genome database (Brent, 2008). With attempts to perform RACE PCR and obtain the full-length sequence of these three genes proving quite difficult, a number of bioinformatics techniques were employed to obtain full-length sequences of *etsub1, etsub2* and *etsub5.* The full-length sequence of *etsub5* was identified using RACE PCR, semi-quantitative PCR, BLAST analysis of the *E. tenella, E. maxima* genomic database and identification of homologous sequences in *T. gondii* and *N. caninum.* The presence of the three active sites (aspartic acid, histidine and serine) of the catalytic triad characterised this gene as a subtilase. The conserved oxyanion hole asparagine, which is structurally required for correct folding and activity of subtilase proteins (Siezen and Leunissen, 1997a) was also identified, as well as a signal peptide which displayed a high level of homology between subtilase protein sequences from other Apicomplexa including *T. gondii* and *N. caninum* (Figure 3.22). Bioinformatics were also used to further elucidate the full-length sequences for *etsub 1* and *etsub2,* however, limitations with the *E. tenella* genomic database and the mis-annotation of, not only these genes, but also the contig within the genome, meant that no further sequence was identified. Despite this, with confirmation of two of the three active sites of the catalytic triad of subtilases in the gene sequences of *etsub1, etsub2* it seems likely that all three genes code for legitimate *E. tenella* subtilases similar to *etsub5* and all three have some potential to play a role(s) in processing of GAM56 and/or GAM82 as a prelude to oocyst wall formation.

CHAPTER 4

LOCALISATION OF GAMETOCYTE UP-REGULATED SUBTILSIN-LIKE PROTEASES IN *EIMERIA TENELLA*

4.1 Introduction

The coccidian parasite, *Eimeria tenella,* possesses a characteristic oocyst wall that protects the parasite from the external environment. Key components of the oocyst wall include a series of tyrosine-rich peptides that are derived from larger proteins designated GAM56 and GAM82, which are found in wall-forming bodies type II in macrogametocytes. It has been proposed (Belli *et al.*, 2006b) that subtilase serine proteases proteolytically cleave GAM56 and GAM82 proteins into smaller peptides prior to their incorporation into the wall.

Subtilases identified in the genome of *E. tenella* have been investigated (Chapter 3) to determine their gene expression profile in different life cycle stages of the parasite, including merozoites, early and late stage gametocytes, unsporulated and sporulated oocysts and, thus, three gametocyte-specific subtilisin-like proteases were identified in *E. tenella*. The aims of this chapter were to (1) confirm the gametocyte-specificity of these three proteins by Western blotting and (2) determine their subcellular localisation by immunofluorescence microscopy. To achieve this, it was first necessary to generate a panel of antibodies, to maximise the likelihood of generating any that would be effective reagents for immunolocalisation studies. This, in turn, required the production of a number of different recombinant and synthetic versions of *Et*SUB1, *Et*SUB2 and *Et*SUB5. Thus: (i) a series of monoclonal antibodies to ~14-18kDa recombinant versions of each subtilase was produced in an attempt to find reagents that would be highly specific (*ie,* to a single epitope) for each enzyme; (ii) polyclonal antisera was raised to short (15 amino acid), synthetic peptides of putative high antigenicity but low cross-homology for each of the subtilases in an effort to produce reagents that were balanced for specificity and reactivity; and (iii) polyclonal sera was raised against recombinant proteins representing the longest PCR-confirmed sequence of each subtilase to produce reagents with as high a reactivity as possible for immunofluorescence microscopy, nothwithstanding that this approach may compromise specificity.

4.2 Production of Antibodies

4.2.1 Production of Antibodies. 1. Monoclonal Antibodies to *Et*SUB1, *Et*SUB2 and *Et*SUB5

The gene sequences of *etsub1, etsub2* and *etsub5* (determined in the previous chapter) were analysed for the correct open reading frame and translated using Seqbuilder (Lasergene™ version 10) to determine the predicted protein sequence for each; a region where homologies between the three truncated proteins ranged from 37.5 – 45% was chosen (Figure 4.1). The peptides produced were 14.6kDa (*Et*SUB1), 17.6kDa (*Et*SUB2) and 17.6kDa (*Et*SUB5) as described in Section 2.2.8.1.3.2. Each peptide was used to immunise five BALB/c mice (Section 2.2.8.1.4). Spleen cells from the immunised mice were then fused with myeloma sp2/0 cells and hybridomas of each were produced following the Promab standard protocol (Section 2.2.8.1.5.1). Ten clones were produced from each parent cell. These clones were then analysed by immunoblot and ELISA to confirm specificity.

The level of reactivity and the specificity of each monoclonal supernatant to the immunising peptide were analysed by ELISA (Section 2.2.8.2). Each ELISA microplate was coated with one of the antigens (500ng per well) and probed with monoclonal antibody supernatant from all 30 clones raised against each antigen as well as a negative control. Each sample was assayed in triplicate, and the average and standard error of the mean (SEM) were calculated (Figure 4.2). Two-way ANOVA was performed to determine statistical significance.

Anti-*Et*SUB1 clones 1-10 reacted specifically to *Et*SUB1 antigen (Figure 4.2a) with no cross-reactivity observed with the other subtilases (Figure 4.2b, c). Clone 1 was selected for further analysis since it recognised the immunising antigen and had the highest level of reactivity by ELISA (Figure 4.2a).

High levels of reactivity of anti-*Et*SUB2 reactivity were observed only in anti-*Et*SUB2 clones 1, 2 and 5 ($p<0.001$)(Figure 4.2b). These clones did not cross react with *Et*SUB5 antigen (Figure 4.2c); however, they did react significantly ($p<0.001$) with *Et*SUB1, as did anti-*Et*SUB2 clone 4. Clone 5 had the highest reactivity to *Et*SUB2, whilst no significant reactivity to either *Et*SUB1 or *Et*SUB5 clones and, therefore, was selected for further analysis (Figure 4.2b).

Significant levels of reactivity ($p<0.001$) against the *Et*SUB5 protein were detected in anti-*Et*SUB5 clones 1, 2, 3, 4, 5, 8 and 10 (Figure 4.2c). Clones 4 and 6 showed significant cross reactivity ($p<0.001$) against *Et*SUB2 but none of the clones showed significant cross-reactivity with *Et*SUB1. Clone 10 was selected for further analysis as it had high reactivity with *Et*SUB5 with no cross-reactivity to either *Et*SUB1 or *Et*SUB2.

```
EtSub1 ------------------------------------------------------------
EtSub2 -------------------------------------------------------AAAN-- 4
EtSub5 PSLDFLPKIREVAKSLSSGTDEVAGALGRVITAAVACVDKISLDRKVDVQRLGSLATNSQ 60

EtSub1 ------------------------------------------------------------
EtSub2 ----------------------------VEDPYFPLQWHLQAAEGLYNLGADKAWQQMQ 35
EtSub5 GNMSLQEPIKGTVNDGATSPQLLRAPSGLLGDPFQMRQWYLHSFFGNFTVEADRTWRKLE 120

EtSub1 ----------SVPVAVVDTGINYLHPELS------LSMWVNRKELHGEEGVDDDGNGFVD 44
EtSub2 SAQQQQPQQQQVVVALLDTGCGPHEDINN-------ILWINSKE-NCSDGVDHDNNGYVN 87
EtSub5 ATENTKP----VVLAILDTGCYLHQDFIDDFDIAKSIFWDNAGETDCSDGIDNDGNGYID 176
* :*::*** . . :* * * . .:*:*.*.**:::

EtSub1 DCKGWDFVDDT-NDITDTYGHGTAVAALMA------------------------------ 116
EtSub2 DLFGWNFIQDNNNPMDDN-GHGSHVAGIVAALQNNGEGISGISERAKVMALKILDQKGEG 103
EtSub5 DCFGWNFVEDNGHPFTDDSGHGTSVTSVAAARAHDGKGGRGVLPNPTVMCLRVGSERGVW 236
* **:*::*. : : * ***: *:.: *

EtSub1 ------------------------------------------------------------
EtSub2 DVSHAIPAIQYAVDNGAKVLTNSWG----------------------------------- 128
EtSub5 -TSATIPALDYAVKMKARVSNHSYGGPGFVSAEYEAFLRALRHDHLIVTAAGNAGCNIDE 295

EtSub1 ------------------------------------------------------------
EtSub3 ------------------------------------------------------------
EtSub5 NERCYFTPGAFRLRGLVNVMASDISGFRASFSNYGSRSVDVAAPGTLIYAGTNYRANSGN 355

EtSub1 ------------------------------------------------------------
EtSub2 ------------------------------------------------------------
EtSub5 RRCTSCYTYSDGTSFAAPIVAGMAAALWGYFERTNPIGWQESTEPASRKVEQAIMYSVTG 415

EtSub1 ------------------------------------------------------------
EtSub2 ------------------------------------------------------------
EtSub5 RRCTSCYTYSDGTSFAAPIVAGMAAALWGYFERTNPIGWQESTEPASRKVEQAIMYSVTG 415

EtSub1 ------------------------------------------------------------
EtSub2 ------------------------------------------------------------
EtSub5 SKALAGAVGTNGVVNLWRALSYYDTVPGPFTPEYPTPESGYNPTYTGGRPQCTQYMVLLL 475
```

Figure 4.1: ClustalW alignment of gametocyte-specific subtilases

The predicted protein sequences of *Et*SUB1, *Et*SUB2 and *Et*SUB5 were aligned and the level of homology measured between the three sequences. Identical amino acid residues between the three subtilase proteases are denoted '*', whilst highly similar amino acids are denoted ':' and similar amino acids are denoted '.'.The selected protein sequence for expression and purification is highlighted (yellow) and represents a region with 37.5-45% homology between the three subtilases.

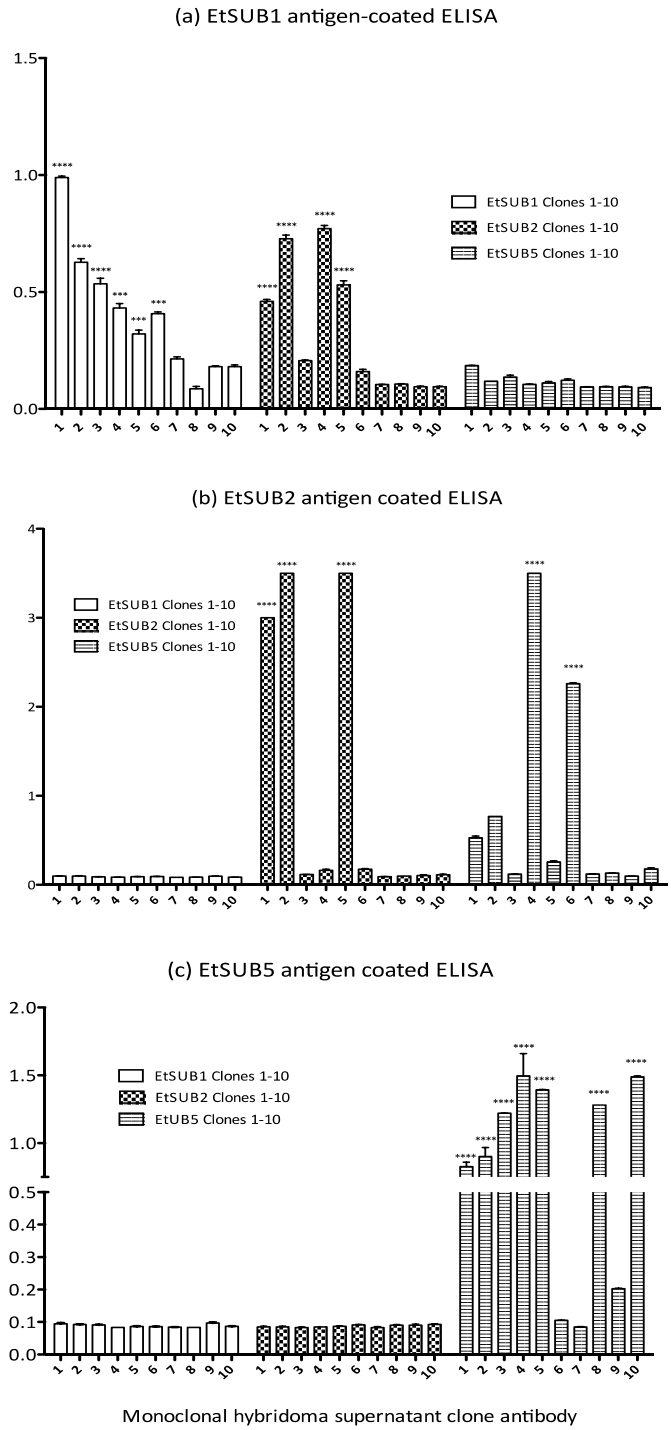


Figure 4.2: Reactivity of monoclonal antibodies to *Et*SUB1, *Et*SUB2 and *Et*SUB5

Mice were immunised with purified synthetic peptides from *Et*SUB1, *Et*SUB2 and *Et*SUB5 and spleen cells of antisera mice fused with myeloma sp2/0 cells. The hybridoma supernatants of ten clones per subtilase were then tested for reactivity with each purified antigen *Et*SUB1 (a), *Et*SUB2 (b) and *Et*SUB5 (c) by ELISA. Each clone supernatant was assayed in triplicate and the mean and SEM were calculated and compared with the pre-bleed negative control by two-way ANOVA and Bonferroni post tests. **** = $p<0.001$, *** = $p<0.005$, ** = $p<0.01$, and *$p<0.05$.

4.2.2 Production of Antibodies. 2. Polyclonal Antisera to Synthetic Peptides Derived from *Et*SUB1, *Et*SUB2 and *Et*SUB5

Kyte-Doolittle plots were produced to show the hydropathic/hydrophilic regions of each protein sequence from their amino acid sequences, using the approach described by Kyte and Doolittle (1982). The plot shows levels above the line corresponding to hydrophilic and, therefore, potentially antigenic regions of the protein, while those below the line indicate hydropathic regions, with neutral regions plotted at the midline (Figure 4.3a). A Jameson-Wolf antigenic plot was also produced to show the predicted antigenic regions of each protein, as described by Jameson and Wolf (Jameson and Wolf, 1988), where windows above the line are indicative of strength of antigenicity of the protein (Figure 4.3a). In parallel, ClustalW analysis was performed on the three subtilases in *E. tenella,* as described in section 2.2.6.3 to ensure that the peptide sequences identified by Jameson-Wolf and Kyte-Doolittle analyses were not homologous to other *E. tenella* subtilases (Figure 4.3b). In combination, these three analyses identified putative synthetic peptides, each 15 amino acids in length, located in areas of high antigenicity and hydrophilicity that were specific for each of *Et*Sub1, *Et*SUB2 and *Et*SUB5 (Figure 4.3).

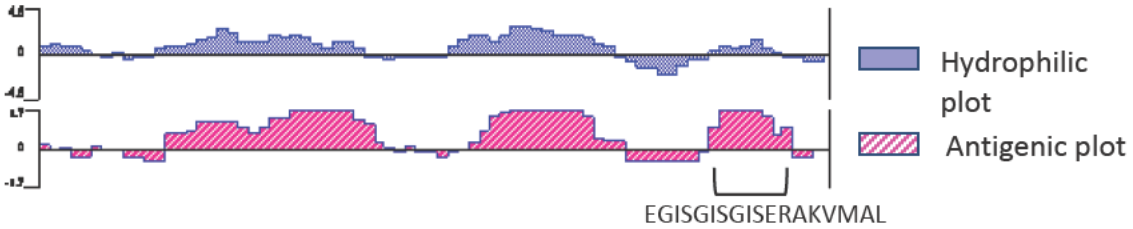
The peptides selected above were manufactured and conjugated to a carrier protein, Keyhole Limpet Haemocyanin (KLH), via a cysteine residue, in collaboration with Auspep Pty Ltd. KLH is a large protein and is attached routinely to peptides to increase the antigenicity and, therefore, the immune response to the peptide (Lateef *et al.*, 2007). Mice were immunised with 500 μg of peptide in Genscript T-Max™ adjuvant by intraperitoneal injection on days 0, 14, 35 and 56 and sera was collected on day 66 post immunisation three, following the standard Genscript protocol (Section 2.2.8.1.4).

Antigen-specific IgG antibody responses to the synthetic peptides of individual mice were analysed by ELISA (Section 2.2.8.2). Unconjugated synthetic peptide (500ng/well) was used to coat each ELISA plate and was reacted with corresponding pre-immune sera, final bleed antisera and no antisera negative control. Two-way ANOVA was performed to determine the level of reactivity and statistical significance.

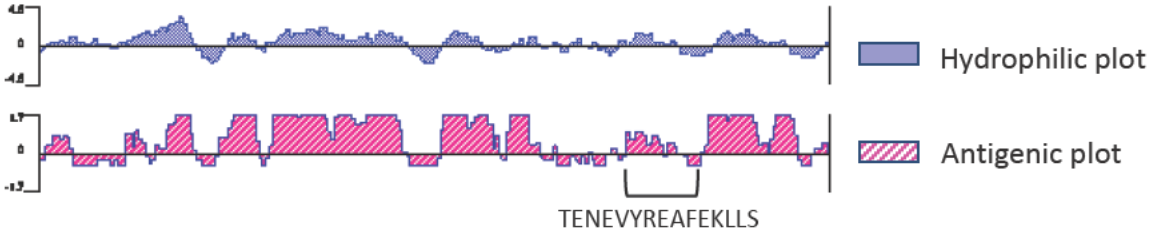
Figure 4.4a shows the level of reactivity of anti-*Et*SUB1 antibody detected in the final bleed antisera from five individual mice immunised with the synthetic peptide. Levels detected in mice A, B and D were significantly higher ($p<0.005$) than the level observed in the pre-immune sera or no antisera controls (Figure 4.4a). No reactivity was detected in final bleed antisera from mice C and E.

(a)

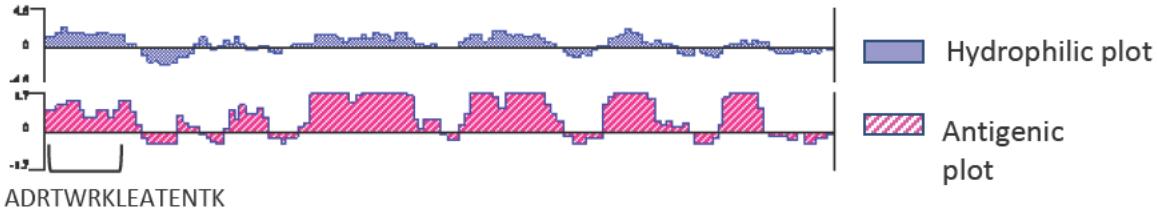
*Et*SUB1



*Et*SUB2



*Et*SUB5



(b)

```
EtSub1          ------------------------------------------------------------
EtSub2          ------------------------------------------------------------
EtSub5          PSLDFLPKIREVAKSLSSGTDEVAGALGRVITAAVACVDKISLDRKVDVQRLGSLATNSQ 60

EtSub1          ------------------------------------------------------------
EtSub2          -------------------------AAANVEDPYFPLQWHLQAAEGLYNLGADKAWQQMQ 35
EtSub5          GNMSLQEPIKGTVNDGATSPQLLRAPSGLLGDPFQMRQWYLHSFFGNFTVEADRTWRKLE 120

EtSub1          ----------SVPVAVVDTGINYLHPELS------LSMWVNRKELHGEEGVDDDGNGFVD 44
EtSub2          SAQQQQPQQQQVVVALLDTGCGPHEDINN-------ILWINSKE-NCSDGVDHDNNGYVN 87
EtSub5          ATENTKP----VVLAILDTGCYLHQDFIDDFDIAKSIFWDNAGETDCSDGIDNDGNGYID 176
                           * :*::***    .   .        :* *  * . .:*:*.*.**:::

EtSub1          DLFGWNFIQDNNNPMDDN-GHGSHVAGIVAALQNNGEGISGISERAKVMALKILDQKGEG 103
EtSub2          DCKGWDFVDDT-NDITDTYGHGTAVAALMAA-SFDKKGGRGVSPTGRVMCLRVGDQGG-V 144
EtSub5          DCFGWNFVEDNGHPFTDDSGHGTSVTSVAAARAHDGKGGRGVLPNPTVMCLRVGSERG-V 235 *
                *  **:*::*. : : *  ***: *:.: **   : :*  *:     **.*:: .: *

EtSub1          ------------------------------------------------------------ 128
EtSub2          QLSRQLMAMDYAIQQGAQVSVHPFSLATENEVYREAFEKLLSSNHLAVVAAGDNGCDLDE 204
EtSub5          WTSATIPALDYAVKMKARVSNHSYGGPGFVSAEYEAFLRALRHDHLIVTAAGNAGCNIDE 295
                  *  : *::**:.  *:*  :.:.

EtSub1          ------------------------------------------------------------
EtSub2          DT-CKSYPASFKIP---------------------------EPSGLLVVGSSTY------ 230
EtSub5          NERCYFTPGAFRLRGLVNVMASDISGFRASFSNYGSRSVDVAAPGTLIYAGTNYRANSGN 355

EtSub1          ------------------------------------------------------------
EtSub2          ------------------------------------------------------------
EtSub5          RRCTSCYTYSDGTSFAAPIVAGMAAALWGYFERTNPIGWQESTEPASRKVEQAIMYSVTG 415

EtSub1          ---------------------------------------------------
EtSub2          ---------------------------------------------------
EtSub5          SKALAGAVGTNGVVNLWRALSYYDTVPGPFTPEYPTPESGYNPTYTGGRPQ 466
```

Figure 4.3: Identification of unique antigenic sequences for *Et*SUB1, *Et*SUB2 and *Et*SUB5.

(a) The Kyte Doolittle program was used to produce the predicted hydrophilicity plot (top) and the Jameson Wolf program was used to produce the predicted antigenic plot (bottom) for *Et*SUB1, *Et*SUB2 and *Et*SUB5. The bracketed region corresponds to the selected region, with the corresponding peptide sequence of each subtilase protein sequence shown below. (b) ClustalW analysis was performed to align the gametocyte-specific subtilases in *E. tenella* and compare homology between the three translated amino acid sequences. Synthetic peptides were designed for regions that were predicted to be antigenic and hydrophilic and located in less homologous regions of each subtilase to minimise potential cross-reactivity between antisera raised in mice. Identical amino acid residues between the three subtilase proteases are denoted '*', whilst highly similar amino acids are denoted ':' and similar amino acids are denoted '.'.The amino acid sequence of each synthetic peptide is highlighted in the sequence for *Et*Sub1 (green), *Et*Sub2 (blue) and *Et*Sub5 (yellow).

A lower level of reactivity was observed with the *Et*SUB2 peptide (Figure 4.4b). However, significant reactivity was observed in mouse A ($p<0.005$) and mouse D ($p<0.01$) in the final bleed compared with the pre-bleed or no antisera controls. However, mice B, C and E showed no significant reactivity.

All five mice immunised with the *Et*SUB5 peptide showed levels of reactivity that were significantly greater ($p<0.001$) than pre-bleed or no antisera controls (Figure 4.4c).

Cross-reactivity between the antisera and the various synthetic peptides was also analysed to determine specificity of the antibodies. Only the mice with the highest level of reactivity to their respective peptide were assayed. This was mouse B for *Et*SUB1, mouse D for *Et*SUB2 and mouse C for *Et*SUB5. Plates were coated with unconjugated synthetic peptide (500ng/well) and probed with antisera from each mouse. Two-way ANOVA was performed to determine the level of reactivity and statistical significance. Results are shown in Figure 4.5.

There was a high level of reactivity of antiserum to the synthetic peptide from *Et*SUB1 with that synthetic peptide (Figure 4.5) but antisera to the synthetic peptides from *Et*SUB2 or *Et*SUB5 showed no reactivity to the *Et*SUB1 synthetic peptide. Similarly, antisera to synthetic peptides of *Et*SUB2 or α-*Et*SUB5 appeared to be specific with significant reactivity only detected with their relevant synthetic peptide antigen (Figure 4.5).

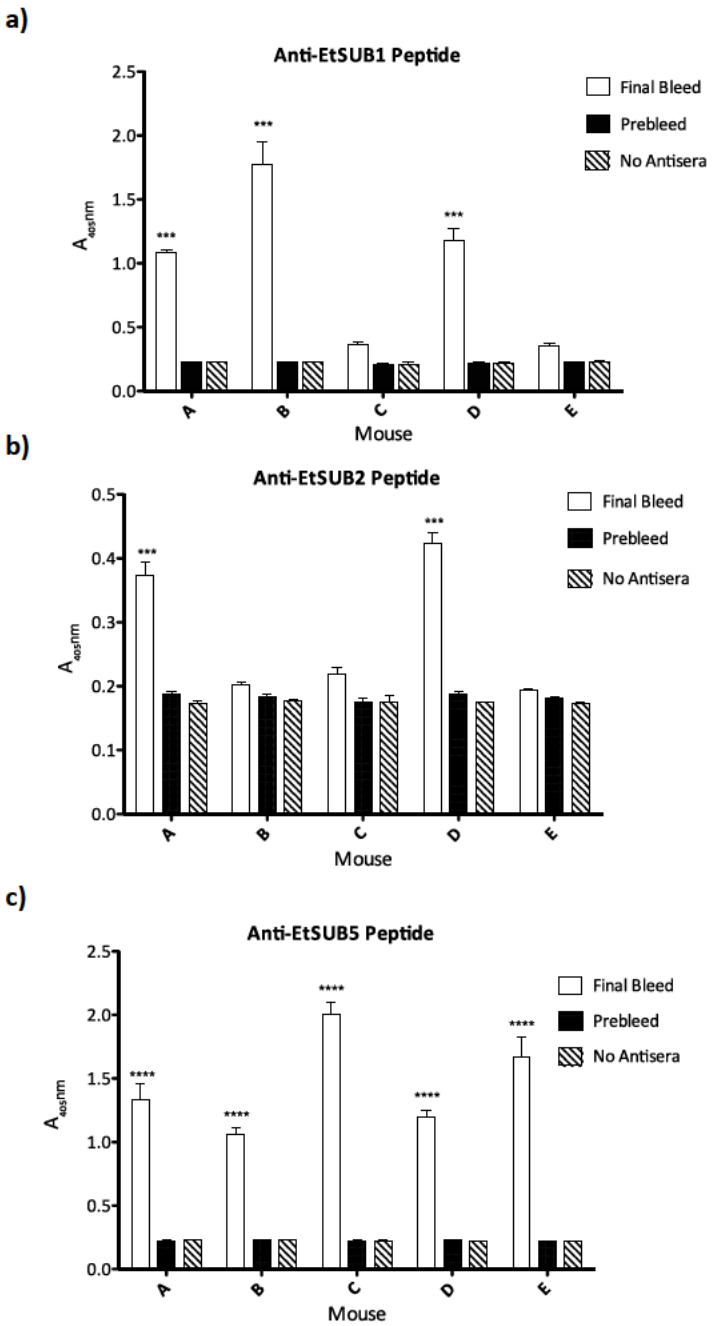


Figure 4.4: Reactivity of antibodies to synthetic peptides from *Eimeria tenella* subtilases

Levels of reactivity of antisera raised against synthetic peptides from *Et*SUB1 (a), *Et*SUB2 (b) or *Et*SUB5 (c) were assessed by ELISA. Five mice (A-E) per synthetic peptide were immunised with four separate doses of 500µg of peptide in Genscript T-Max™ adjuvant by intraperitoneal injection on days 0, 14, 35 and 56. The absorbance at 405nm was determined for each mouse and compared with the absorbance obtained from a pre-bleed sample of the same mouse and a no sera control. A two way ANOVA was performed to determine the p-value of antisera against no antisera control. '****' represents a significant difference ($p<0.001$); '***a significant difference of $p<0.005$; '**' a significant difference of $p<0.01$; and '*' a significant difference of $p<0.05$ to the no antisera control. Results represent the mean ± SE of four replicate assays per mouse.

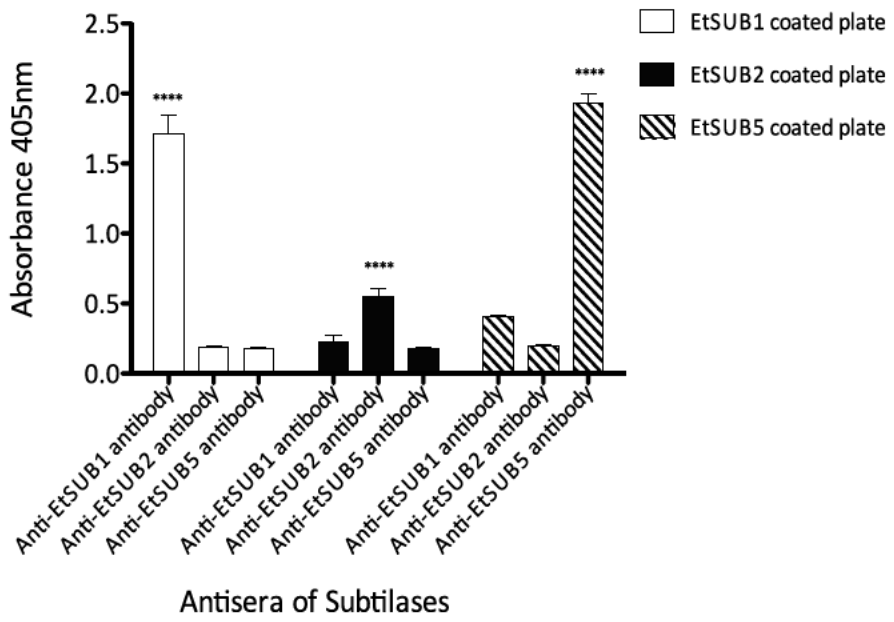


Figure 4.5: Specificity of antibodies against synthetic peptides from *Et*SUB1, *Et*SUB2 and *Et*SUB5

Five mice (A-E) were immunised with four separate doses of 500μg of synthetic peptide in Genscript T-Max™ adjuvant by intraperitoneal injection on days 0, 14, 35 and 56. Mice that showed high reactivity against their antigen were selected to test for cross-reactivity with the other subtilases by ELISA (see section 2.2.8.2). Two way ANOVA was performed to determine the p-value of antisera against no antisera control. '****' represents a significant difference ($p<0.001$); '***a significant difference of $p<0.005$; '**' a significant difference of $p<0.01$; and '*' a significant difference of $p<0.05$ to the no antisera control. Results represent the mean ± SE of four replicate assays per mouse.

4.2.3 Production of Antibodies. 3. Polyclonal Antisera to Recombinant *Et*SUB1, *Et*SUB2 and *Et*SUB5

The PCR-confirmed sequence (as described in section 3.2.3) of *Et*SUB1 was used to produce recombinant protein. The expressed protein was 86 amino acids in length, had a predicted pI of 4.41, and a predicted molecular weight of 12kDa. Vector-specific primers were designed to directionally clone EtSub1 into the pET101/D-TOPO® vector. Following expression in *E. coli* cells, the protein was purified by affinity purification using the Profinia™ protein purification system. Purified recombinant protein was analysed by Coomassie stained SDS-PAGE and immunoblot. A prominent band, 12kDa in size, was observed in the Coomassie stained SDS-PAGE gel (Figure 4.6a). There was also a smaller band, approximately 8kDa in size, observed in the gel. Probing of the immunoblot with anti-penta His antibody detected bands corresponding to these two proteins (Figure 4.6a). Given that the predicted molecular weight is 12.4kDa, the top band is most likely the correct protein, and the lower band a degradation product. Expression and purification produced approximately 1mg of protein per litre of culture. Therefore, expression was repeated a further four times to produce 4mg of purified *Et*SUB1 protein for use in the immunisation of mice for the production of antisera.

The PCR-confirmed sequence (as described in section 3.2.3) of *Et*SUB2 was used to produce recombinant protein. The expressed protein was 230 amino acids in length, with a predicted pI of 4.43, and a predicted molecular weight of 31.5kDa. Vector-specific primers ACCATGGCAGCAGCAAATGTTG and GCCCTTTGGCAGCAGCACTTAT were designed to directionally clone *etsub2* into the pET101/D-TOPO® vector and to fuse a his-tag in frame at the C-terminal end of the protein, followed by a stop codon. Following expression of gametocyte-specific subtilases in *E. coli* cells, the protein was purified by affinity purification using the Profinia™ protein purification system. Coomassie stain and immunoblot were used to confirm purification of the protein. The anti-penta His antibody detected a protein band migrating at approximately 30kDa (Figure 4.6b). However, a number of smaller bands were also observed to react to the anti-penta His antibody (Figure 4.6b) and are most probably degradation products; these could not be removed even under numerous wash conditions of increasing concentrations of imidazole. Therefore, in order to purify *Et*SUB2 for immunisation, the band of the appropriate size was excised from the gel and used to immunise mice to produce antisera (Section 2.2.8.1.4).

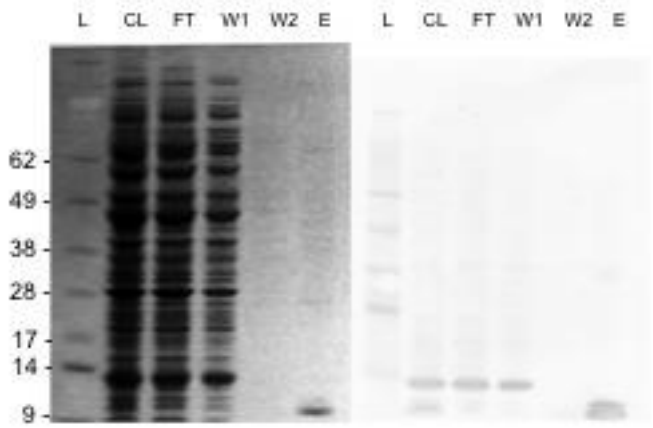
The PCR-confirmed sequence (as described in section 3.2.3) of *Et*SUB5 was used to produce recombinant protein. The expressed protein was 139 amino acids in length, had a predicted pI of 4.97, and a predicted molecular weight of 18.9kDa. Vector-specific primers ACCATGGAAGCCGACAGGAC and GCCCTTCGCGTAGTCAAGAGCAGG were designed to directionally

clone *Et*Sub5 into the pET101/D-TOPO® vector. Expressed protein was affinity purified using a nickel-chelated column as described in Section 2.2.8.1.3.3 and purification confirmed by Coomassie stain and immunoblot. A band approximately 20kDa in size was observed in the Coomassie stained gel with the anti-penta His antibody detecting the same 20kDa band (Figure 4.6c). A number of higher and lower molecular weight proteins were also recognised by the anti-penta His antibody and, although varying wash conditions were used, the bands could not be removed. Therefore, to purify *Et*SUB5 for immunisation, the band of the appropriate size was excised from the gel and used to immunise mice to produce antisera (Section 2.2.8.1.4).

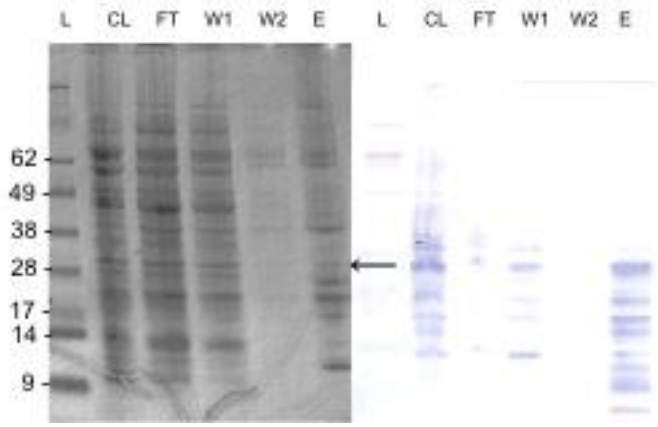
Antigen-specific IgG antibody responses to each subtilase protein were analysed by ELISA (Section 2.2.8.2) to determine the level of cross-reactivity. Purified protein (500ng/well) was used to coat each ELISA plate and pre-immune sera, final bleed antisera and no sera negative control from each mouse was tested for reactivity against the immunising protein as well as cross-reactivity to the other proteins. Two-way ANOVA was performed to determine the level of reactivity and Bonferroni post-test to determine statistical significance.

The level of reactivity to *Et*SUB1 detected in the antisera from the final bleed in mice immunised with r*Et*SUB1 was significantly higher ($p<0.001$) than the level observed in the prebleed (Figure 4.7a). However, there also appeared to be a significant level of cross reactivity ($p<0.001$) of the antisera to *Et*SUB1 with the *Et*SUB2 and *Et*SUB5 purified proteins (Figure 4.7b and c). In contrast, antisera to EtSUB2 detected r*Et*SUB2 at a significantly higher level ($p<0.05$) than the level observed in the prebleed or no antisera control (Figure 4.7b) but lacked any significant reactivity to the other recombinant proteins (Figure 4.7a and c). However, antibody to *Et*SUB5 was cross-reactive; reactivity detected in antisera from mice immunised with r*Et*SUB5 was significantly higher ($p<0.001$) than the level observed in the prebleed or no antisera control (Figure 4.7c) but, the anti-EtSUB5 antibody also appeared to have a significant level of cross-reactivity to the *Et*SUB2 purified protein (Figure 4.7b).

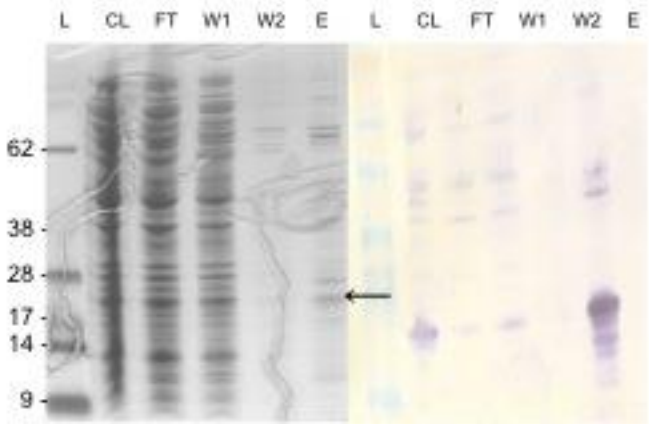
a) *Et*SUB1 protein purification



b) *Et*SUB2 protein purification



c) *Et*SUB5 protein purification



L: Ladder (kDa) SeeBlue® Plus
CL: Clear lysate
FT: Flow through
W1: Wash 1
W2: Wash 2
E: Elution

Figure 4.6. Purification of recombinant versions of *Eimeria tenella* subtilases by affinity purification from *E. coli* cell lysate

*Et*SUB1, *Et*SUB2 and *Et*SUB5 were cloned into pET101/D-TOPO® in *E. coli* BL21 Star™ and expression induced with 1mM IPTG. Recombinant protein was purified with the Profinia™ protein purification system on a Bio-Scale Mini Profinity IMAC nickel-chelating column. A Coomassie and immunoblot were performed to confirm purification of each subtilase: (a) *Et*SUB1; (b) *Et*SUB2; and (c) *Et*SUB5. Protein bands of *Et*SUB2 and *Et*SUB5 excised from protein gels are indicated with an arrow.

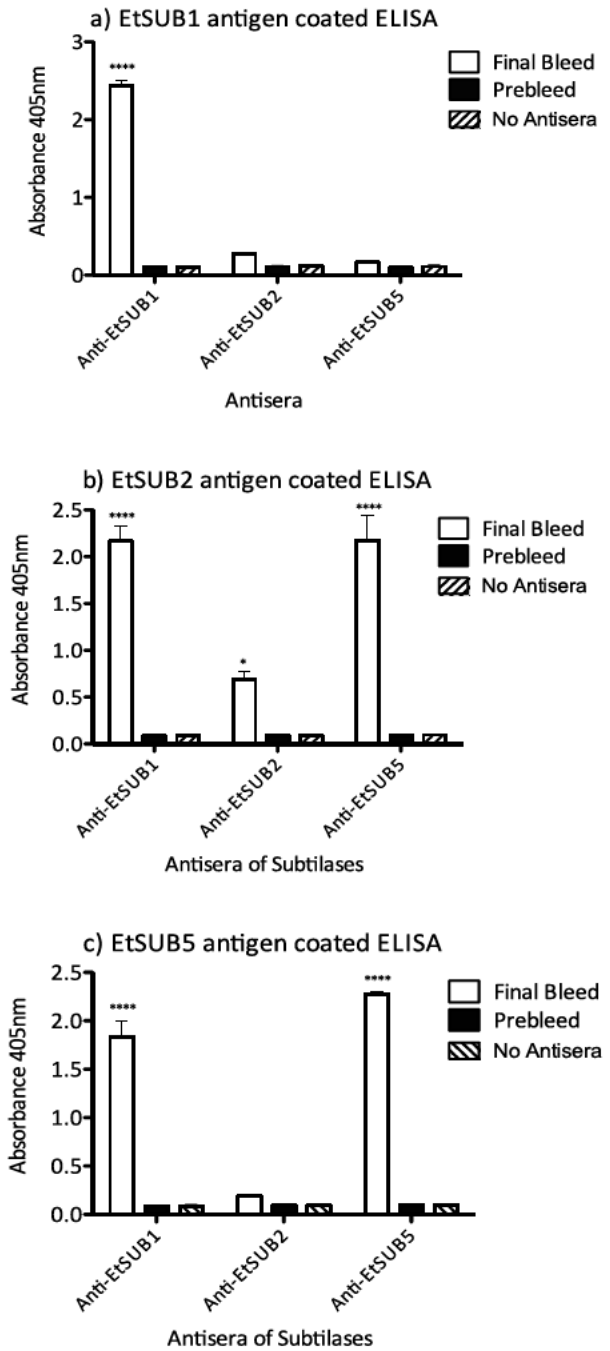


Figure 4.7: Specificity of antibodies against recombinant protein versions of *Et*SUB1, *Et*SUB2 and *Et*SUB5

Six mice per protein were immunised with four separate doses of 100µg of purified protein by subcutaneous injection on days 0, 14, 35 and 56. Antisera collected from the mice on day 66 post were tested by ELISAs (see Section 2.2.8.2) on (a) *Et*SUB1 recombinant protein, (b) *Et*SUB2 recombinant protein and (c) *Et*SUB5 recombinant protein. Reactivity was compared with pre-bleeds and no antisera. Two way ANOVA was performed to determine the p-value of antisera against no antisera control. '****' represents a significant difference ($p<0.001$) to the no antisera control. Results represent the mean ± SE of six mice with four replicate assays each.

4.3 Stage-Specific Reactivity of Antibodies

4.3.1 Stage-Specific Reactivity of Antibodies. 1. Monoclonal Antibodies to *Et*SUB1, *Et*SUB2 and *Et*SUB5

Protein lysates were prepared from a number of different lifecycle stages of *E. tenella* and standardised to concentrations of 1mg/ml using the NanoDrop spectrophotometer as described in section 2.2.8.4. Samples were reduced and alkylated as described in section 2.2.8.3. Each sample (10μg) was loaded onto separate wells of a 4-12% Bis-Tris gel and electrophoresed under reducing and alkylating conditions. Following SDS-PAGE, each membrane was probed with monoclonal antibodies (diluted 1:20) to *Et*SUB1, *Et*SUB2 or *Et*SUB5. Monoclonal supernatant 1 produced against the *Et*SUB1 protein recognised two bands approximately 60kDa-65kDa in lysates of both early and late stage gametocytes (Figure 4.8). The band in later stage gametocytes was more prominent than that observed at 134hrs p.i. No reactivity was observed in any of the other life cycle stages, or in uninfected host tissue (Figure 4.8a). Monoclonal supernatant 5 produced against *Et*SUB2 recognised three dominant bands in lysates of late stage gametocytes of approximately 70kDa, 60kDa, and 55kDa (Figure 4.8b). Two fainter bands of 55 kDa and 60 kDa were also observed in early stage gametocytes. These bands are all larger than the predicted 48 kDa. No reactivity was observed in any other lifecycle stages of *E. tenella* with anti-*Et*SUB2 monoclonal antibody. Similarly, no reactivity was observed to uninfected host intestine.. A 60kDa band was observed in both early and late stage gametocytes of *E. tenella* probed with monoclonal antibody 10 produced against the *Et*SUB5 (Figure 4.8c). A second band of approximately 70-75kDa was also observed in late stage gametocytes (144hrs p.i.). Both of these bands were smaller than the predicted size of the native protein of 82kDa. There was no reactivity observed in any of the other life cycle stages of *E. tenella,* and no reactivity observed in uninfected host intestine. As a negative control, membranes blotted with identical *E. tenella* protein lysates were also probed with pre-immune mouse sera. No reactivity was observed (data not shown).

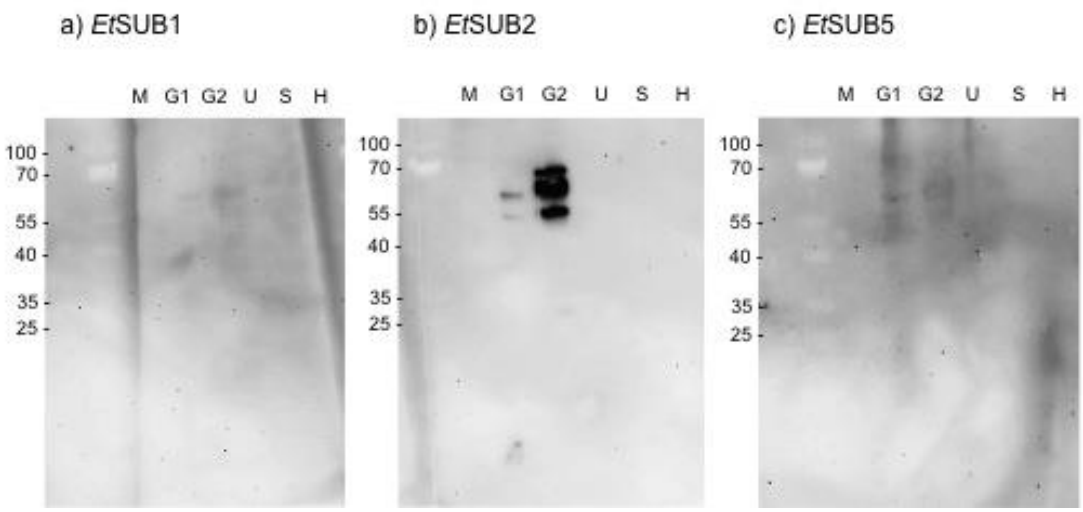


Figure 4.8: Western blot detection of different developmental stages of *Eimeria tenella* with monoclonal antibodies from mice immunised with recombinant peptides from *Et*Sub1, *Et*Sub2 and *Et*Sub5

Protein lysates were prepared from different *E. tenella* developmental stages including merozoites (M), early stage gametocytes 134hrs p.i. (G1), late stage gametocytes 144hrs p.i. (G2), unsporulated oocysts (UO) and sporulated oocysts (SO), and from uninfected host intestine (H). Protein lysates were alkylated and reduced then separated by SDS-PAGE on a 4 – 12% Bis-Tris gradient gel. Following transfer to PVDF membrane, the lysates were probed with the strongest reacting monoclonal hybridoma clone as judged by ELISA (Section 2.2.8.2) to each of (a) *Et*SUB1, (b) *Et*SUB2 and (c) *Et*SUB5. Bands were visualised by incubation of blots with secondary anti-mouse IgG HRP conjugate and then development with Chemiluminescent Peroxidase Substrate-1 (Sigma®).

4.3.2 Stage-Specific Reactivity of Antibodies. 2. Polyclonal Antibodies to Synthetic Peptides from *Et*SUB1, *Et*SUB2 and *Et*SUB5

To determine the parasite stage in which the original reference protein is found, protein lysates of various *E. tenella* lifecycle stages were immunoblotted with antisera from mice immunised with synthetic peptides from *Et*SUB1, *Et*SUB2 and *Et*SUB5. The antisera samples that were selected were the same ones analysed in Figure 4.5, ie, antisera from mouse B (*Et*SUB1), mouse D (*Et*SUB2) and mouse C (*Et*SUB5).

No protein bands were detected by antiserum to *Et*SUB1 in any parasite stage examined (Figure 4.9a). Two bands, of approximately 55kDa and 45kDa, were observed in gametocytes probed with antiserum to *Et*SUB2 (Figure 4.9b). A larger band of 198kDa was also observed in gametocytes within a large smear that extended to 62kDa when probed with antibody to *Et*SUB2, however, that same band is observed in the pre-immune sera western blot (Figure 4.10d). A band of approximately 62kDa was observed in sporulated oocysts probed with antiserum to *Et*SUB2. A smear was also observed in sporulated oocysts, starting just above the 62kDa marker, with an intense band observed at approximately 70kDa (Figure 4.9b). A band of approximately 100kDa in size was also present in unsporulated oocysts on the *Et*SUB2 blot, and this band was not detected by the other antisera.

Two bands approximately 55kDa and 45kDa were also observed when blots were probed with antiserum to *Et*SUB5, as was the non-specific 198kDa band and smear in gametocytes starting from the 62kDa marker to the top of the blot (Figure 4.9c). The 55kDa band was also observed in unsporulated oocysts in the immunoblot probed with antiserum to *Et*SUB5. Two faint bands approximately 70kDa and 80kDa in size were also observed in sporulated oocysts probed with the same antiserum.

Antiserum to GAM56 was used as a positive control to confirm purity of the gametocyte protein lysate preparation. The immunoblot (Figure 4.9e) detected a band of the predicted size 56kDa in gametocytes and a 35kDa band, the processed product of GAM56, in unsporulated oocysts (Belli *et al.*, 2002a). The antiserum to GAM56 also recognised the non-specific 198kDa band in gametocytes.

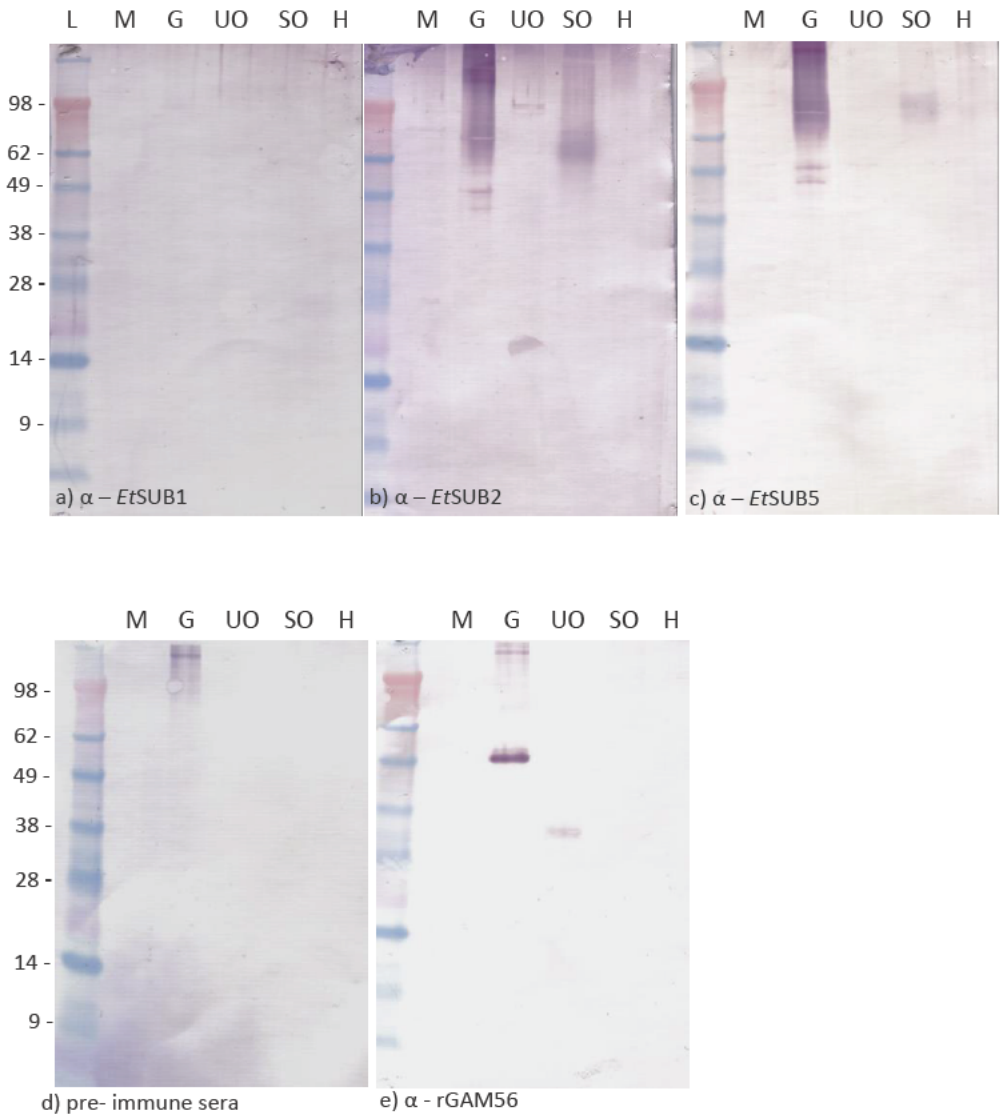


Figure 4.9: Western blot detection of *E. tenella* lifecycle stages probed with antisera from mice immunised with synthetic peptides of *Et*SUB1, *Et*SUB2 and *Et*SUB5

Protein lysates were prepared from different *E. tenella* developmental stages including merozoites (M), gametocytes 134hrs p.i. (G), unsporulated oocysts (UO) and sporulated oocysts (SO), and from uninfected host intestine (H). Protein lysates were separated by SDS-PAGE on a 4 – 12% Bis-Tris gradient gel and, following transfer to PVDF membrane, were probed with antisera to (a) synthetic peptide from *Et*SUB1, (b) synthetic peptide from *Et*SUB2 or (c) synthetic peptide from *Et*SUB5. Sera from unimmunised mice (d) and mice immunised with recombinant GAM56 (e) were included as negative and positive assay controls, respectively. Membranes were then probed with secondary anti-mouse IgG Alkaline phosphatase conjugate and visualised with alkaline phosphatase substrate Sigma Fast™ BCIP®/NBT.

4.3.3 Stage-Specific Reactivity of Antibodies. 3. Polyclonal Antibodies to Purified Recombinant Proteins of *Et*SUB1, *Et*SUB2 and *Et*SUB5

Mice were immunised with purified recombinant protein versions of *Et*SUB1, *Et*SUB2 and *Et*SUB5. Six mice per protein were immunised with 100µg of protein on day 0, with three booster shots given on days 14, 35 and 56 (Section 2.2.8.1.4). Sera were collected 10 days after the last booster shot (day 66). Western blot analysis was used to investigate if the antibodies raised recognise the immunising protein. Antisera samples used were selected based on samples that showed the highest level of reactivity judged by ELISA (data not shown).

The immunoblot probed with antisera to *Et*SUB1 showed prominent bands of approximately 50 and 55kDa in the lysates of gametocytes (Figure 4.10a). Two fainter bands of approximately 80kDa and 115kDa were also observed in this life cycle stage of the parasite. A smaller band of approximately 25kDa was seen in lysates of unsporulated oocysts.

The immunoblot of *E. tenella* protein lysates probed with antisera to *Et*SUB2 also showed a dominant band of approximately 55kDa in lysates of gametocytes as well as a larger band of 100kDa (Figure 4.10b). A 70kDa band was also observed in all the parasite lifecycle stages – merozoites, gametocytes, unsporulated and sporulated oocysts – but not in the chicken caeca sample (Figure 4.10b).

Two bands of approximately 55kDa and 60kDa were observed in lysates of gametocytes in the immunoblot probed with antisera to *Et*SUB5 antibody as well as larger bands of approximately 80kDa and 115 kDa in size (Figure 4.10c).

GAM56 protein was again used as a positive control to confirm the gametocyte stage specificity of the protein and subsequent immunoblot product, where a band of the predicted size 56kDa was observed in gametocytes (Belli *et al.*, 2002a), and 35kDa, the processed product of GAM56, was observed in unsporulated oocysts. No bands were observed in the immunoblot probed with negative antisera (representative of a number of mice tested) or in the chicken host tissue control.

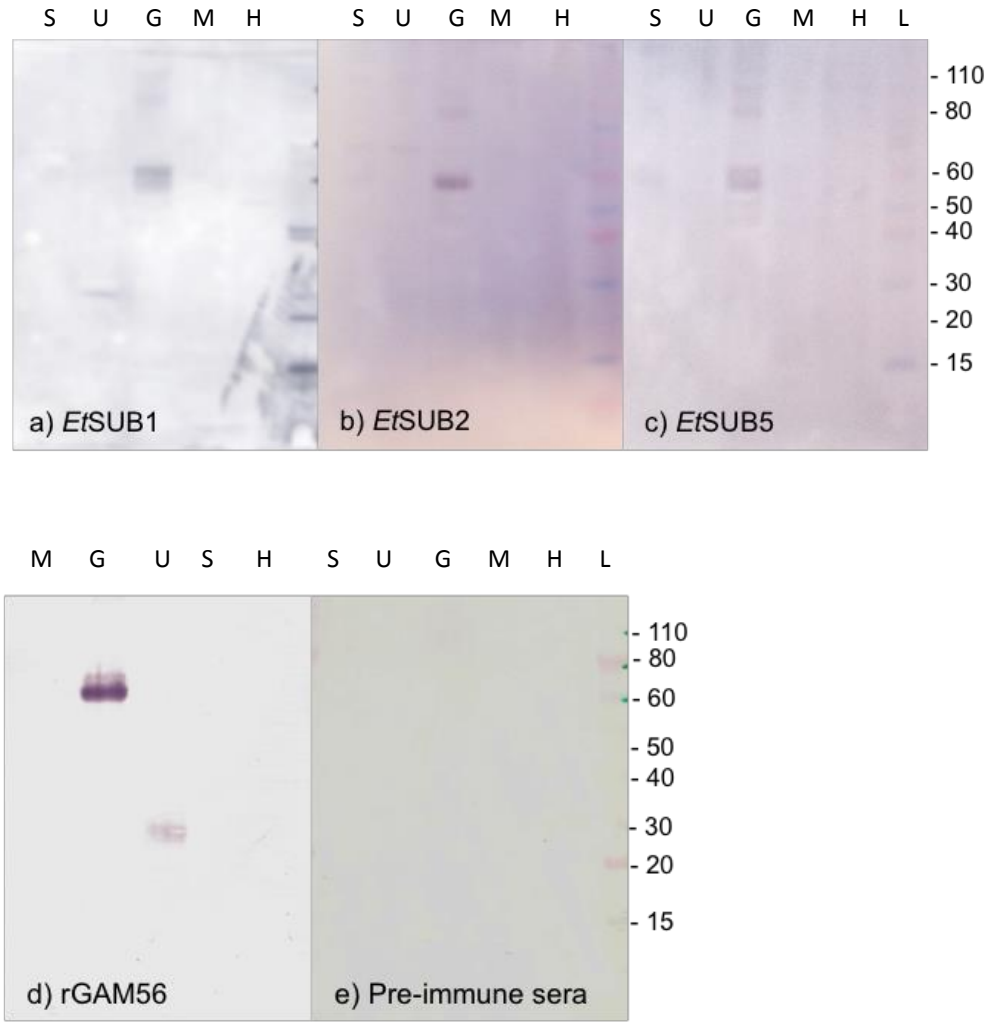


Figure 4.10: Western blots of *Eimeria tenella* lifecycle stages probed with antisera from mice immunised with purified recombinant protein versions of *Et*SUB1, *Et*SUB2 and *Et*SUB5

Protein lysates were prepared from different *E. tenella* developmental stages including merozoites (M), gametocytes 134hrs p.i. (G), unsporulated oocysts (UO) and sporulated oocysts (SO), and from uninfected host intestine (H) as described in section 2.2.8.3. Protein lysates were separated by SDS-PAGE on a 4 – 12% Bis-Tris gradient gel and, following transfer to PVDF membrane, were probed with antisera to (a) recombinant *Et*SUB1, (b) recombinant *Et*SUB2 or (c) recombinant *Et*SUB5. Sera from mice immunised with recombinant GAM56 (d) or unimmunised mice (e) were included as positive and negative assay controls, respectively. Membranes were then probed with secondary anti-mouse IgG Alkaline phosphatase conjugate and visualised with alkaline phosphatase substrate Sigma Fast™ BCIP®/NBT as described in section 2.2.8.6.

4.4 Immunolocalisation of *Et*SUB1, *Et*SUB2 and *Et*SUB5

4.4.1 Immunolocalisation of *Et*SUB1, *Et*SUB2 and *Et*SUB5 in Macrogametocytes of *Eimeria tenella.* 1. Monoclonal Antibodies

Paraffin sections of *E. tenella*-infected chicken caeca, taken at 112hrs p.i., 134hrs p.i. and 144hrs p.i., were stained with monoclonal antibodies to *Et*SUB1, *Et*SUB2 and *Et*SUB5 monoclonal antibodies at 1:20 dilution. Rabbit antisera to rGAM56, raised previously against a recombinant 56kDa gametocyte antigen (Belli *et al.*, 2002b) was used as a positive control for the identification of macrogametocyte-specific wall-forming bodies type II (Ferguson *et al.*, 2003a). Mouse and rabbit sera were used as negative control in each experiment. DAPI was used for counterstaining of *E. tenella* microgametocytes and host cell nuclei. Antisera to rGAM56 showed distinctive reactivity to WFBII of macrogametocytes at both 134hrs and 144hrs p.i. with negligible background staining (Figure 4.11b, f, and j). This not only confirmed the structure and the location of macrogametocytes within the sample, but also confirmed the success of the immunohistochemistry. Despite a positive reactivity to the *Et*SUB1 antigen observed in both ELISA and immunoblot analysis (Section 4.2.1 and 4.3.1), no reactivity was observed between the monoclonal antibody of *Et*SUB1 and macrogametocytes in caecal sections, (Figure 4.11a). Similar results were obtained for the monoclonal antibodies to *Et*SUB2 (Figure 4.11e) and *Et*SUB5 (Figure 4.11i). No reactivity of monoclonal antibodies to other developmental stage *of E. tenella* (data not shown) or to host tissue was detected (Figure 4.11 c, d, g, h, k, l).

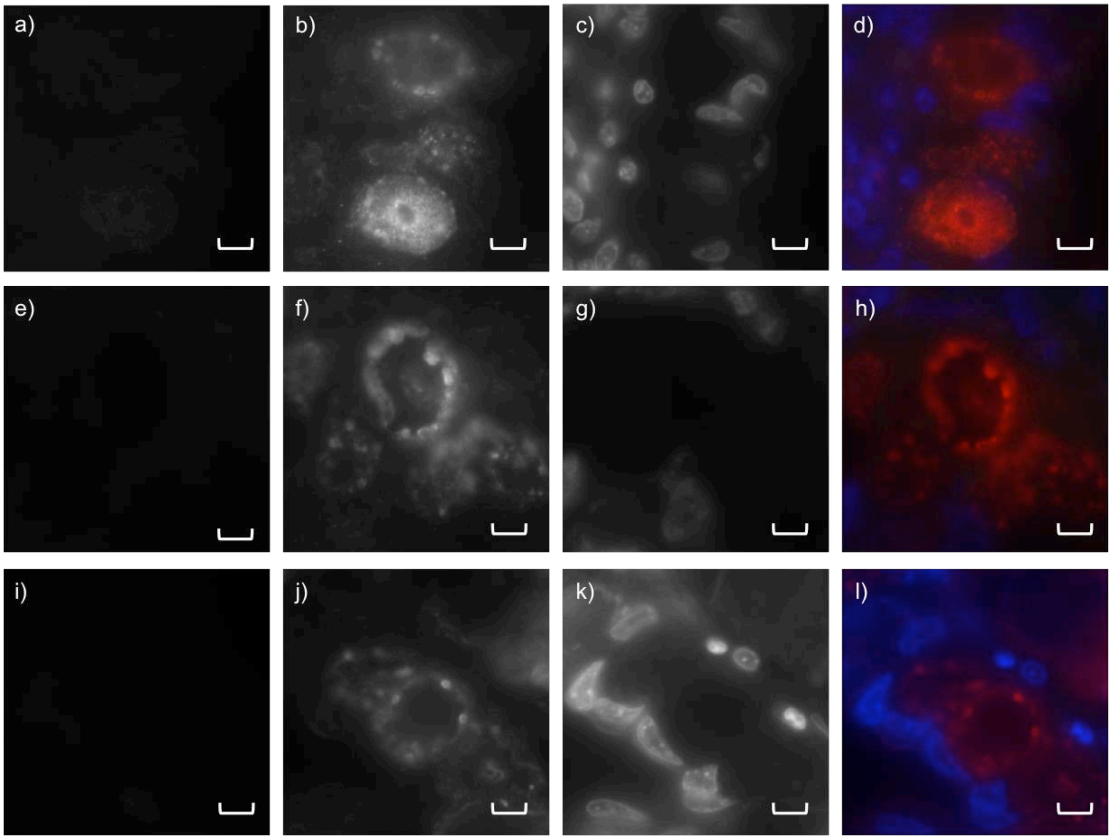


Figure 4.11: Immunolocalisation of monoclonal antibodies to *Et*SUB1, *Et*SUB2 and *Et*SUB5

Eimeria tenella gametocyte infected intestinal tissue: (a) reacted with monoclonal antibody to *Et*SUB1 visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (b) reacted with antisera to rGAM56 visualised using anti-rabbit (whole molecule) conjugated to AlexaFluor594; (c) counterstained with 4` 6 – diamidino – 2 – phenylindone (DAPI) [(d) is the overlay of (a), (b) and (c)]; (e) reacted with monoclonal antibody to *Et*SUB2 visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (f) reacted with antisera to rGAM56 visualised using anti-rabbit (whole molecule) conjugated to AlexaFluor594; (g) counterstained with DAPI [(h) is the overlay of (e), (f) and (g)]; (i) reacted with monoclonal antibody to *Et*SUB5 visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (j) reacted with antisera to rGAM56 visualised using anti-rabbit (whole molecule) conjugated to AlexaFluor594; (k) counterstained with DAPI [(l) is the overlay of (i), (j) and (k)]. Bar is 5μm.

4.4.2 Immunolocalisation of *Et*SUB1, *Et*SUB2 and *Et*SUB5 in Macrogametocytes of *Eimeria tenella.* 2. Polyclonal Antisera to Synthetic Peptides

Localisation of gametocyte-specific subtilases in *E. tenella* was investigated by immunofluorescence microscopy (IFM) employing the mouse antiserum to the synthetic peptides from *Et*SUB1, *Et*SUB2 and *Et*SUB5 analysed in Figures 4.12 and 4.13. Tissue sections were prepared from caeca taken from chickens infected with *E. tenella* 134h and/or 144h previously. Sections were double stained with polyclonal anti-rGAM56 antisera and the appropriate anti-peptide antiserum raised in mice (Figure 4.12a-i). Pre-immune sera taken from mice prior to immunisation were used as negative controls for non-specific reactivity. A polyclonal antibody (raised previously in rabbits to the recombinant GAM56 (Belli *et al.*, 2002b) that localises specifically to wall forming bodies type II (WFBII) in macrogametocytes was used as a marker of macrogametocytes in infected tissue. DAPI was used for the counterstaining of host cell nuclei. DAPI also confirmed the presence of microgametocytes in the tissue sections and allowed confirmation that none of the antisera to any of the subtilisins reacted with this form of the parasite and lack of reactivity with merozoites and oocysts was also confirmed (data not shown).

Antiserum to *Et*SUB1 did not react with macrogametocytes in caecal sections, but it did appear to react slightly and inconsistently with host cells (Figure 4.12a-d). Antiserum to peptide from *Et*SUB5 also failed to react with macrogametes in caecal sections and was non-reactive towards host cells (Figure 4.12i-l). In contrast, antiserum to *Et*SUB2 reacted strongly to the macrogametocytes of *E. tenella* (Figure 4.12e-h) with an overlay of reactivity with antiserum to GAM56 showing that *Et*SUB2 is not located in WFBIIs (Figure 4.12h); rather it appears to be located in the cytoplasm adjacent to the wall forming bodies. The same antiserum did not react with host cells even under prolonged fluorescence exposure times.

Following the positive staining of anti-*Et*SUB2 antibody in macrogametocytes of *E. tenella,* further immunolocalisation experiments were performed to try and pinpoint where in the macrogametocyte the antibody was binding, as it did not appear to co-localise with antiserum to GAM56 in WFBII. Tissue sections of *E. tenella* infected caeca were double stained with antiserum to *Et*SUB2 and mAb E1D8, a monoclonal antibody developed previously against the wall forming bodies type I (WFBI) of *E. tenella* (Mouafo *et al.*, 2002b). DAPI was used to counterstain *E. tenella* microgametocytes and host cell nuclei. Again, antiserum to *Et*SUB2 reacted strongly with the macrogametocytes of *E. tenella* (Figure 4.13a) and strong reactivity of WFB1 with mAb E1D8 was also seen (Figure 4.13b). An overlay of both images, plus DAPI (Figure 4.13c-d) again appears to show that *Et*SUB2 localises to the cytoplasm adjacent to wall forming bodies and not to the WFBI themselves.

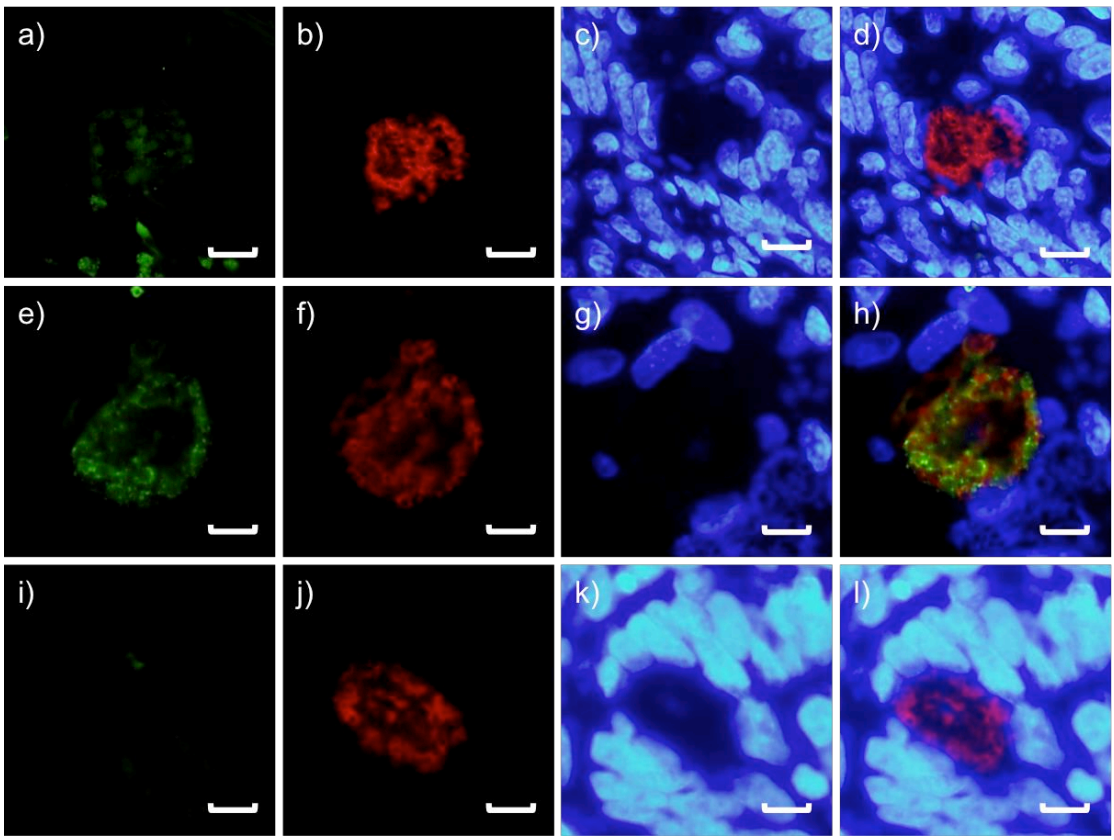


Figure 4.12: Immunofluorescence of *Eimeria tenella* gametocytes in infected chicken caeca probed with antisera to synthetic peptides from *Et*SUB1, *Et*SUB2 and *Et*SUB5

Eimeria tenella gametocyte infected caecal tissue was: (a) reacted with antiserum to *Et*SUB1, visualised using anti-mouse (λ chain specific) conjugated to isothiocyanate (FITC); (b) reacted with antiserum to GAM56 visualised using anti-rabbit (whole molecule) conjugated to tetramethylrhodamine isothiocyanate (TRITC); (c) counterstained with 4` 6 – diamidino – 2 – phenylindone (DAPI) [(d) is the overlay of (a), (b) and (c)]; (e) reacted with antiserum to *Et*SUB2 visualised using anti-mouse (λ chain specific) conjugated to FITC; (f) reacted with antiserum to GAM56 visualised using anti-rabbit (whole molecule) conjugated to TRITC; (g) counterstained with DAPI [(h) is the overlay of (e), (f) and (g)]; (i) reacted with antiserum to *Et*SUB5 visualised using anti-mouse (λ chain specific) conjugated to FITC; (j) reacted with antiserum to GAM56 visualised using anti-rabbit (whole molecule) conjugated to TRITC; (k) counterstained with DAPI [(l) is the overlay of (i), (j) and (k)]. Bar is 5μm.

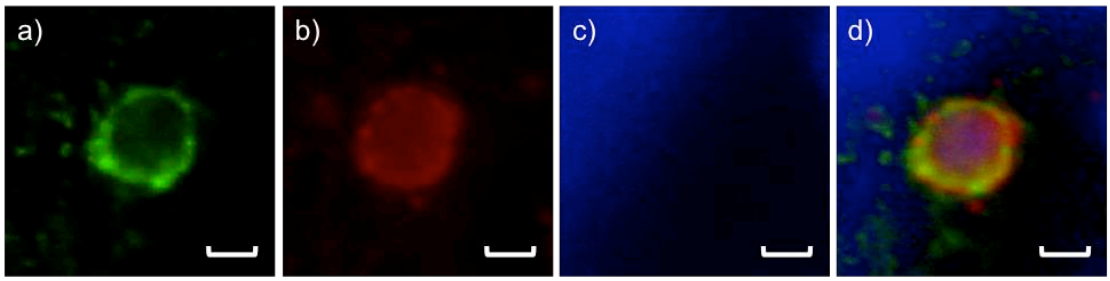


Figure 4.13: Lack of co-localisation of *Eimeria tenella* gametocyte-specific subtilase *Et*Sub2 with wall forming bodies type 1 of macrogametocytes

Eimeria tenella gametocyte-infected caeca were stained with: (a) antiserum to *Et*SUB2, visualised using anti-mouse (λ chain specific) conjugated to isothiocyanate (FITC); (b) monoclonal antibody E1D8, visualised using anti-rabbit (whole molecule) conjugated to tetramethylrhodamine isothiocyanate (TRITC); and (c) 4` 6 – diamidino – 2 – phenylindone (DAPI). Image (d) is an overlay of (a), (b) and (c). Bar is 5μm.

4.4.4 Immunolocalisation of *Et*SUB2 in Macrogametocytes of *Eimeria tenella.* 3. Polyclonal Antisera to Recombinant Proteins

The cross-reactivity observed between the subtilase antisera that were produced from mice immunised with purified protein of *Et*SUB1 and *Et*SUB5 (section 4.2.3) ruled out their use in immunofluorescence microscopy (IFM); however, the antisera raised against *Et*SUB2 was used. Tissue sections were prepared from caeca infected with *E. tenella* 134h and 144h p.i., and double stained with polyclonal antisera to rGAM56 and *Et*SUB2 (Figure 4.14). Pre-immune serum, taken prior to immunisation, was used as a negative control for non-specific activity. DAPI was used for the counterstaining of host cell nuclei.

Strong reactivity was observed in infected tissue probed with antiserum to rGAM56, localising to the WFBII in macrogametocytes (Figure 4.14). Antiserum to *Et*SUB2 also localised to the macrogametocytes of *E. tenella* (Figure 4.14) but, again, did not co-localise with antiserum to rGAM56 in WFBII; rather it localised to the cytoplasm adjacent to the WFBs in a pattern that was seen using the antiserum to the *Et*SUB2 synthtetic peptide (Figures 4.13 and 4.14). No antisera reacted with *E. tenella* merozoites or oocysts (data not shown), or host cells (counter-stained with DAPI; Figure 4.14).

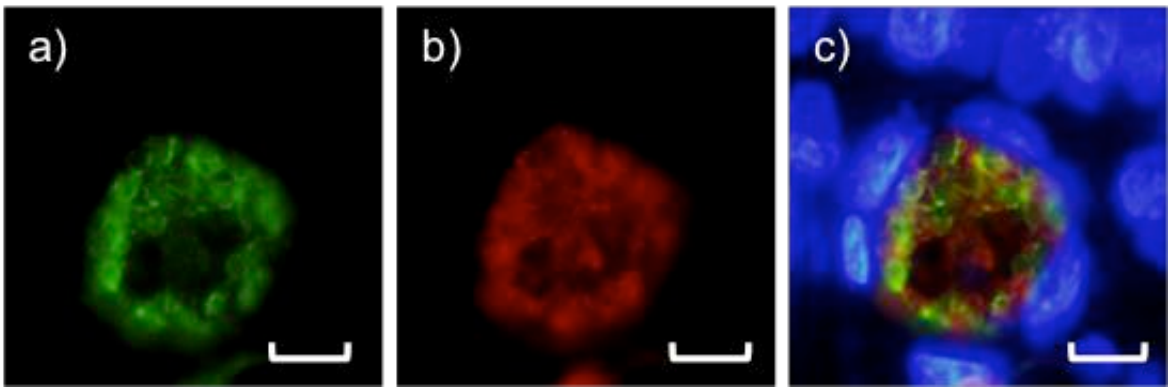


Figure 4.14: Immunofluorescence of *Eimeria tenella* gametocytes in infected chicken caeca probed with antisera to recombinant versions of *Et*SUB2

Eimeria tenella gametocyte infected caecal tissue was: (a) reacted with antiserum to *Et*SUB2, visualised using anti-mouse (λ chain specific) conjugated to isothiocyanate (FITC); (b) reacted with antiserum to GAM56 visualised using anti-rabbit (whole molecule) conjugated to tetramethylrhodamine isothiocyanate (TRITC); (c) counterstained with 4` 6 – diamidino – 2 – phenylindone (DAPI) and overlaid with (a) and (b). Bar is 5μm.

4.5 Discussion

A panel of monoclonal and polyclonal antibodies was produced successfully against various recombinant and synthetic versions of peptides derived from the predicted amino acid sequences of the putative gametocyte-specific subtilisins, *Et*SUB1, *Et*SUB2 and *Et*SUB5, from *E. tenella.* As expected, monoclonal antibodies to recombinant proteins and polyclonal sera to synthetic peptides were very specific, as confirmed by ELISAs (Figures 4.2 and 4.5). However, polyclonal sera raised against recombinant versions of *Et*SUB1 and *Et*SUB5 cross-reacted with other subtilases (Figure 4.7), limiting dramatically their utility in Western blot and immunofluorescence localisation analyses. This cross-reactivity may reflect the sequence conservation apparent around the catalytic triad of the three proteases, recognition of which was potentially maximised by this approach; and, in this context, it is perhaps noteworthy that some of the monoclonal antibodies also displayed cross-reactivity (Figure 4.2). It would be interesting, in the future, to map the epitopes recognised by this panel of monoclonal antibodies but this was not attempted in this study.

The monoclonal antibodies against *Et*SUB1, *Et*SUB2 and *Et*SUB5 proved useful in confirming the stage-specific production of these proteins; in all three cases, reactivity on Western blots was confined to gametocyte lysates (Figure 4.8). The sizes of the bands recognised by these monoclonal antibodies was rarely an exact match for their respective predicted molecular masses but, given the uncertainties of the true expressed sequence of all three of these subtilisins, this is perhaps not surprising (*eg, Et*SUB2 is predicted to be 48kDa by the current genome database but available sequence does not include code for the third serine active site, which is essential for activity (Siezen and Leunissen, 1997a)). Multiple banding was also a feature of the Western blot profile of the subtilases; post-translational processing of pro-enzymes is a characteristic of subtilisins (Siezen and Leunissen, 1997a), which may account for this observation, though this remains to be confirmed. Another possible explanation is self-degradation since no protease inhibitors were included in lysate samples meaning that proteolytic activity may have led to cleavage of the subtilisins themselves (Cheng *et al.*, 2009).

Polyclonal antisera to synthetic proteins derived from *Et*SUB1, *Et*SUB2 or *Et*SUB5 proved less useful for Western blot analyses. Antiserum to *Et*SUB1 failed to recognise any bands in any developmental stages whilst the similarity of the smeared banding pattern for antisera to both *Et*SUB2 and *Et*SUB5 suggested a degree of cross-reactivity that, despite its specificity for gametocyte lysates (Figure 4.9), obviously complicates diagnostic usefulness.

Polyclonal antisera to recombinant subtilases were quite cross-reactive, antiserum to *Et*SUB1 recognising both *Et*SUB2 and *Et*SUB5 and antiserum to *Et*SUB5 cross-reacting with *Et*SUB2 (Figure

4.7). The antiserum to *Et*SUB2, however, was very specific. Unfortunately, the cross-reactivity seen in ELISAs was reflected in Western blots (Figure 4.10), with gametocyte bands of 50-55 kDA recognized by all three sera. On balance, and taking into account the other Western blot results (Figures 4.8 and 4.9), these are most likely bands for *Et*SUB2, and rule out using the antisera for *Et*SUB1 and *Et*SUB5 in immunolocalisation experiments.

Unfortunately, the monoclonal antibodies to all three subtilisins proved non-reactive in immunocalisation investigations (Figure 4.11), failing to react with parasites within caecal sections. It is probable that the epitopes recognised in ELISAs and Western blots as a result of immunisation with the various recombinant subtilases are either not exposed or are affected by tissue processing in the caecal section samples.

Antisera to synthetic peptides of *Et*SUB1 and *Et*SUB5 also failed to react with infected caeca (Figure 4.12). However, antiserum to synthetic *Et*SUB2 reacted very clearly, producing a distinctive staining pattern (Figures 4.12 and 4.13) that was replicated with the polyclonal antisera to recombinant *Et*SUB2 (Figure 4.14). The reactivity of the antiserum indicates that *Et*SUB2 is not located within the WFBs of *Eimeria,* as proposed by (Wallach, 2010) but, rather, in the cytoplasm immediately surrounding them. This is intriguing and, arguably, makes more sense than location within the WFBs. If Belli *et al.* (2006) are correct, and the WFBs are a storage site for proteins like GAM56 and GAM82 that are destined to be processed and incorporated into the oocyst wall, then the presence of subtilisins in the WFBs could be quite detrimental, potentially leading to premature processing of GAM56 and GAM82, creating the possibility that they may cross-link within WFBs before appropriate transportation and disaggregation of the WFBs at the periphery of the macrogamete. In such a scenario, oocyst wall formation would certainly be compromised. Thus, *Et*SUB2 is a viable candidate subtilisin to catalyse the cleavage of tyrosine-rich proteins, including GAM56 and GAM82, as a prelude to oocyst wall formation. Indeed, it is at present the leading candidate for such a role because, whilst roles for *Et*SUB1 and *Et*SUB5 cannot be ruled out because evidence is absent either for or against their participation, it does seems significant that *Et*SUB2 is, thus far, the only subtilisin that has been shown to be completely gametocyte-specific by quantitative PCR (Chapter 3), protein detection and immunolocalisation. Put another way, it is in the right place at the right time to process GAM56 and GAM82 into tyrosine-rich peptides; the same cannot currently be said for other enzymes.

CHAPTER 5

BIOINFORMATIC IDENTIFICATION OF PUTATIVE OXIOREDUCTASES IN *EIMERIA TENELLA*

5.1 Introduction

Hydrogen peroxide (H_2O_2) has been implicated in a number of physiological processes, including cellular signaling pathways, thyroid biosynthesis and fertilisation (De Deken *et al.*, 2000, Donko *et al.*, 2005). It has also been shown to play a role in dityrosine cross-linking of proteins during the formation of extracellular matrices. Dual Oxidase (Duox) enzymes have been shown to provide the oxidative basis for this process in both sea urchin fertilisation membrane development (Heinecke and Shapiro, 1992, Wong *et al.*, 2004) and in protein cross-linking to stabilise the nematode cuticle of *C. elegans* (Edens *et al.*, 2001). Dityrosine and DOPA, amino acids implicated in the synthesis of extracellular matrices *via* quinone tanning (Waite, 1990b), have been detected previously in the *Eimeria* oocyst wall at levels that suggest that the formation is an enzyme-mediated reaction (Belli *et al.*, 2003b), though this remains to be determined. A study by Mai *et al.* (2011) showed that exoogenous peroxidases, myeloperoxidase and horseradish peroxidase, are able to catalyse the rapid formation of dimers, trimmers and polymers of a truncated recombinant version of GAM56. High levels of dityrosine were detected in the polymer samples (Mai *et al.*, 2011). Given this, it was predicted that an oxioreductase protein, with oxidative activity could potentially play a role in the oxidation of tyrosine-rich, or other wall proteins during oocyst wall formation of the parasite *E. tenella*. Oxidases, or oxioreductases, are the broad range of enzymes that catalyse many redox reactions. Their actions include the catalysis of the transformation of free, neutral oxygen gas into free radicals, including H_20_2, or, alternatively, they can act as free radical scavenger systems, where the by-product can be H_2O_2 (Ngoka, 2008).

The goal of this chapter was to identify oxioreductases in *E. tenella.* However, there is limited genomic, transcript and proteomic data available on the *E. tenella* genome database making it problematic for data mining of unknown genes. Therefore, the *T. gondii* genome database (Gajria *et al.*, 2008), available at http://toxodb.org/toxo/, was used to identify potential homologues of *E. tenella* oxioreductases*. Toxoplasma gondii* is an excellent model coccidian and there is a lot of publically available information on the *T. gondii* genome, transcript and proteome data accessible through this database. Much of this can be extrapolated to other coccidian parasites, such as *E. tenella.* Moreover, the observation that *Toxoplasma* oocysts autofluoresce at similar wavelengths to those of *Eimeria* (Belli *et al.*, 2006b, Lindquist *et al.*, 2003) suggests that dityrosine bonds are present in both parasites and that the mechanism of oocyst wall formation is very similar. Supporting this idea, six tyrosine-rich proteins have been identified in the oocyst proteome of *T. gondii* (Fritz *et al.*, 2012). Thus, the *Toxoplasma gondii* genome database was employed to generate a list of oxioreductase-like proteins with known expression in *T. gondii* oocysts*,* and this list used to identify homologues in *E. tenella*.

5.2 Identification of oxioreductases in *E. tenella* databases

5.2.1 *Toxoplasma gondii* database search for homologous oxioreductases

A bioinformatics strategy was developed to interrogate the *T. gondii* genome database, generate a list of oxioreductase-like proteins with known expression in *T. gondii* oocysts, and use this list as a basis to identify homologues in *E. tenella* worthy of further investigation of transcript and protein expression. This strategy, which utilises several bioinformatics tools available on the *T. gondii* genome database (Gajria *et al.*, 2008), is highlighted in Figure 5.1.

The first integrated search (Step 1, Figure 2.1) was performed as described in Section 2.2.6.6. The proteomic data used in this search was generated from sporulated, mature day 10 *T. gondii* (M4, type II) oocysts that were fractionated into oocyst walls (bleach-treated and untreated), sporocysts and sporozoites. The subsequent protein preparations were analysed by 1DE, LC-MS/MS (Fritz *et al.*, 2012). This search resulted in the identification of 997 proteins in the oocyst/sporozoite proteome of *T. gondii* including rhoptry proteins (Table S1), microneme proteins, oxioreductases, tyrosine-rich proteins, late embryogenesis abundant domain-containing (LEA) proteins and hypothetical proteins (Fritz *et al.*, 2012). Six tyrosine-rich hypothetical proteins were found to be present in oocyst wall fractions, consistent with their proposed role in oocyst wall formation (Fritz *et al.*, 2012).

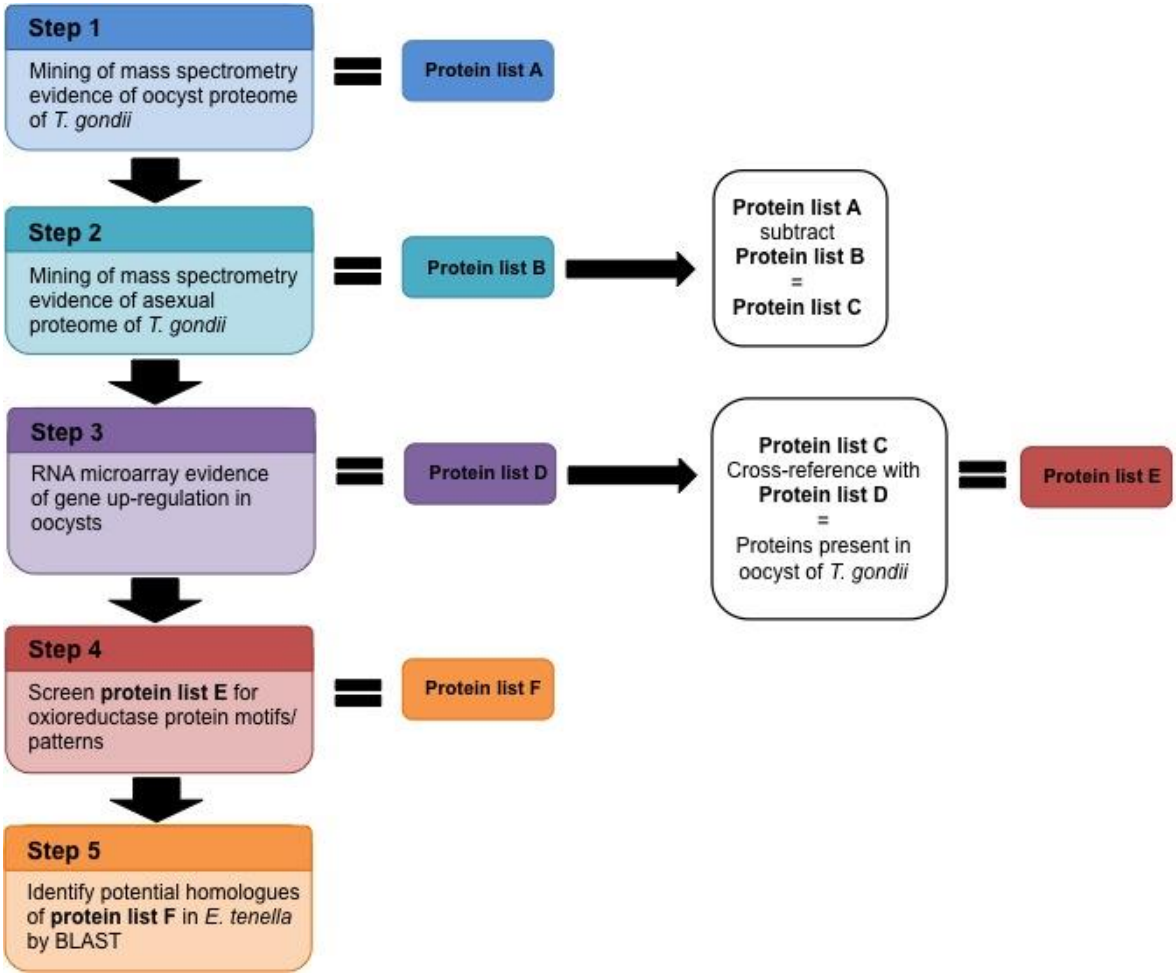


Figure 5.1: Diagrammatic representation of the bioinformatics strategy taken for data mining of *Toxoplasma gondii* genome database

A bioinformatics approach was taken to identify potential oxioreductases in *E. tenella.* The database of *T. gondii* was mined using several different searches to identify potential homologues. Step 1 involved data mining the oocyst proteome of *T. gondii* resulting in Protein List A*;* step 2 involved data mining the asexual proteome of *T. gondii* (Protein List B) and subtracting proteins found in both lists to produce Protein List C; step 3 involved data mining of RNA microarray data to identify *T. gondii* genes up-regulated in oocysts (Protein List D); this list was cross-referenced with Protein List C and proteins identified in both lists constituted Protein List E; Protein List E was screened for oxioreductase protein motifs/patterns (Step 4); and the resulting Protein List F was used to identify homologues in *E. tenella* by BLAST analysis.

A second search was conducted on the tachyzoite proteome to identify genes associated with asexual stages (Figure 2.2). For this search, everything but the oocyst proteome was chosen (Figure 2.2).

This search resulted in the identification of 9698 genes (Table S2). Using the integrated search tool on http://toxodb.org/toxo/, any genes that were present in the first search (Table S1) and also found in tachyzoite extracts (Table S2) were eliminated as candidates for oxidation of oocyst wall proteins in wall formation (Figure 2.2). After elimination of any proteins that were expressed in the tachyzoite proteome as well as the oocyst proteome, the list of possible candidates was narrowed to 120 (Table 5.1). Genes identified in this more definitive list included 53 hypothetical proteins, *Cryptosporidium parvum* oocyst wall protein (COWP) homologues, TgOWP6 and TgOWP2 (Possenti *et al.*, 2010), and seven oxioreductases.

Table 5.1 120 proteins identified in the integration of search 1 and 2 of the *Toxoplama gondii* database

Gene ID*	Product Description	Peptide Count
Hypothetical Proteins		
TGME49_094400	Hypothetical protein	6
TGME49_094600	Hypothetical protein	53
TGME49_009470	Hypothetical protein	25
TGME49_009810	Hypothetical protein	6
TGME49_097280	Hypothetical protein	4
TGME49_097290	Hypothetical protein	3
TGME49_053180	Hypothetical protein	2
TGME49_053300	Hypothetical protein	2
TGME49_120710	Hypothetical protein	2
TGME49_120530	Hypothetical protein	4
TGME49_120280	Hypothetical protein	10
TGME49_119890	Hypothetical protein	23
TGME49_065130	Hypothetical protein	3
TGME49_065120	Hypothetical protein, conserved	13
TGME49_088220	Hypothetical protein	2
TGME49_106050	Hypothetical protein	6
TGME49_106270	Hypothetical protein	4
TGME49_087250	Hypothetical protein	8
TGME49_086770	Hypothetical protein, conserved	5
TGME49_085440	Hypothetical protein	2
TGME49_038160	Hypothetical protein	3
TGME49_039360	Hypothetical protein	3
TGME49_040470	Hypothetical protein	5
TGME49_003890	Hypothetical protein	5
TGME49_003230	Hypothetical protein	9
TGME49_081590	Hypothetical protein	10
TGME49_058820	Hypothetical protein, conserved	3
TGME49_073510	Hypothetical protein	18
TGME49_072240	Hypothetical protein	15
TGME49_071490	Hypothetical protein	17
TGME49_069380	Hypothetical protein	10
TGME49_068880	Hypothetical protein	2
TGME49_067980	Hypothetical protein	3
TGME49_028240	Hypothetical protein	7
TGME49_023430	Hypothetical protein	3
TGME49_037080	Hypothetical protein	17
TGME49_014570	Hypothetical protein	20
TGME49_015140	Hypothetical protein	5
TGME49_109030	Hypothetical protein	2
TGME49_109530	Hypothetical protein	4
TGME49_109540	Hypothetical protein	4
TGME49_110680	Hypothetical protein	2

TGME49_112950	Hypothetical protein	20
TGME49_116550	Hypothetical protein	4
TGME49_048900	Hypothetical protein	2
TGME49_051910	Hypothetical protein	15
TGME49_070950	Hypothetical protein	17
TGME49_011320	Hypothetical protein	7
TGME49_106500	Hypothetical protein	15
TGME49_066700	Hypothetical protein	2
TGME49_092350	Hypothetical protein	18
TGME49_086040	Hypothetical protein	6
TGME49_09	Hypothetical protein	4
Oocyst Wall Proteins		
TGME49_009610	Oocyst wall protein COWP, putative	18
TGME49_086250	Oocyst wall protein, putative	20
TGME49_076880	Late embryogenesis abundant domain-containing protein	9
TGME49_076870	LEA1 protein, putative	10
TGME49_076860	Late embryogenesis abundant domain-containing protein	20
TGME49_076850	Late embryogenesis abundant domain-containing protein	12
Transcription, nucleic acid metabolism and regulation		
TGME49_067410	Scavenger receptor protein TgSR1, putative	2
TGME49_057290	tRNA 2` phosphotransferase, putative	2
TGME49_025120	rRNA processing protein, putative	4
TGME49_006400	DNA photolyase, putative	27
TGME49_110160	Piwi-PAZ domain-containing protein	18
TGME49_019830	High mobility group protein, putative	1
Protein processing, turnover and processing/Proteolytic Proteins		
TGME49_035680	M16 family peptidase, putative	9
TGME49_032400	PAN domain-containing protein	5
TGME49_009920	PAN domain-containg protein	55
TGME49_035200	PAN domain-containing protein	80
TGME49_035210	PAN domain-containing protein	10
TGME49_035390	PAN domain-containing protein	30
TGME49_054330	Lipase domain-containing protein	10
TGME49_066270	2-methylbutyryl-CoA dehydrogenase, putative	4
TGME49_059180	Cystathionine beta-synthase, putative	13
TGME49_112930	Cystathionine beta-lyase, putative	4
TGME49_016640	Homoserine kinase, putative	8
TGME49_047240	Ubiquitin carboxyl-terminal hydrolase isozyme L5, putative	2
TGME49_046800	Acylamino-acid-releasing enzyme, putative	6
TGME49_059670	Von Willebrand factor type A domain-containing protein	70
TGME49_065860	Nit protein, putative	2
TGME49_095640	Endothelian-converting enzyme, putative	11
TGME49_095470	C2 domain-containing protein	2
Surface and apical complex proteins		
TGME49_054430	Microneme protein, putative	16

TGME49_056040	LCCL domain-containing protein	2
TGME49_115550	Microneme protein, putative	5
TGME49_115730	Apical membrane antigen, putative	15
TGME49_058810	SRS27B	8
TGME49_058550	SRS28 (= SporoSAG)	39
TGME49_058370	Rhoptry kinase family protein ROP28	9
Oxioreductase motif/pattern containing proteins without a known function		
TGME49_026300	Oxioreductase, putative	4
TGME49_069120	Oxioreductase, putative	30
TGME49_069400	Oxioreductase, short chain dehydrogenase/reductase domain-containing protein	9
TGME49_043660	Glucose-methanol-choline domain-containing protein	2
TGME49_013340	Glucose-methanol-choline oxioreductase domain-containing protein	9
TGME49_003580	NAD-binding domain-containing protein	8
TGME49_053150	Oxioreductase, putative	45
Metabolic and antioxidant proteins		
TGME49_010260	Glycerol-3-phosphate dehydrogenase, putative	15
TGME49_106450	Short chain dehydrogenase family protein, putative	6
TGME49_089930	Phosphophenolpyruvate carboxykinase, putative	3
TGME49_031900	Acyl-CoA dehydrogenase, putative	8
TGME49_027100	Glutaredoxin, putative	13
TGME49_115260	Alanine dehydrogenase, putative	8
TGME49_115480	Acyl-CoA oxidase, putative	9
TGME49_116190	Superoxide dismutase, putative	14
TGME49_003500	Alanine dehydrogenase, putative	3
TGME49_003570	Acyltransferase, putative	14
TGME49_082230	Sulfate adenylyltransferase-adenylsulfate kinase, putative	
TGME49_038200	Alpha/beta hydrolase fold domain containing protein	
TGME49_042390	Enoyl-CoA hydratase/isomerase family protein, putative	
TGME49_034570	Peroxisomal multifunctional enzyme type 2, putative	18
TGME49_019230	Acetoacetyl-CoA synthetase, putative	10
TGME49_058200	Glutamic acid-rich protein, putative	25
TGME49_032580	Acetyl-coenzyme A synthetase, putative	2
TGME49_047500	Acyl-coenzyme A oxidase, putative	7
TGME49_071210	4-alpha-glucanotransferase, putative	19
TGME49_069110	Ornithine aminotransferase, putative	8
TGME49_003720	Vitamin K epoxide reductase complex subunit 1, putative	5
Membrane and extracellular proteins		
TGME49_085690	Notch (DSL) domain-containing protein	9
TGME49_078120	SCP-like extracellular domain-containing protein	10
TGME49_020280	SCP-like domain-containing protein	8
Signaling and trafficking proteins		
TGME49_055460	EGF-like domain-containing protein	12

*Accession numbers of genes in *T. gondii* database http://toxodb.org/toxo/

5.2.2 Identification of genes with known protein expression evidence in *Toxoplasma gondii* oocyst proteome – data from microarray of oocyst, tachyzoites and bradyzoites extracts

A third search was done using microarray data to identify genes that showed significantly up-regulated transcript levels (>3 fold increase) in oocysts compared with transcript levels in tachyzoites or bradyzoites as detailed in Section 2.2.6.6. (Step 3, Figure 5.1). A total of 673 gene transcripts were identified (Table S3). These genes were then cross-referenced with the 120 genes identified in protein list C (Table 5.1). Eighty-two genes were identified as being oocyst/sporozoite specific as well as significantly up-regulated (Table 5.2). The resulting list included a putative oocyst wall protein, homologous to *C. parvum* COWP, as well as four annotated oxioreductases and hypothetical proteins (Table 5.2).

Table 5.2: Identification of genes in the oocyst proteome of *Toxoplasma gondii* with up-regulated gene expression in sporulated oocysts

Gene ID*	Product Description	Peptide Count
Hypothetical Proteins		
TGME49_094600	hypothetical protein	53
TGME49_009470	hypothetical protein	25
TGME49_097280	hypothetical protein, conserved	6
TGME49_097290	hypothetical protein, conserved	3
TGME49_053300	hypothetical protein, conserved	2
TGME49_120710	hypothetical protein	2
TGME49_120530	hypothetical protein	4
TGME49_120280	hypothetical protein	10
TGME49_119890	hypothetical protein	23
TGME49_066700	hypothetical protein	2
TGME49_065120	hypothetical protein, conserved	2
TGME49_088220	hypothetical protein	2
TGME49_092350	hypothetical protein	18
TGME49_106050	hypothetical protein	6
TGME49_106270	hypothetical protein	4
TGME49_087250	hypothetical protein	8
TGME49_086770	hypothetical protein, conserved	5
TGME49_086040	hypothetical protein	6
TGME49_038160	hypothetical protein	3
TGME49_039360	hypothetical protein	3
TGME49_003230	hypothetical protein	9
TGME49_081590	hypothetical protein	10
TGME49_073510	hypothetical protein	18
TGME49_072240	hypothetical protein	15
TGME49_071490	hypothetical protein	17
TGME49_070950	hypothetical protein	17
TGME49_069380	hypothetical protein	10
TGME49_023430	hypothetical protein	3
TGME49_037080	hypothetical protein	17
TGME49_014570	hypothetical protein	20
TGME49_015140	hypothetical protein	5
TGME49_112950	hypothetical protein	20
TGME49_116550	hypothetical protein	4
TGME49_003890	hypothetical protein	5
TGME49_011320	hypothetical protein	7
Oocyst wall proteins		
TGME49_009610	oocyst wall protein COWP, putative	18
TGME49_086250	oocyst wall protein, putative	20
TGME49_076880	late embryogenesis abundant domain-containing protein	9
TGME49_076870	LEA1 protein, putative	10
TGME49_076860	late embryogenesis abundant domain-containing protein	20
TGME49_076850	late embryogenesis abundant domain-containing protein	12

Surface and apical complex proteins		
TGME49_058810	SRS27B	8
TGME49_058550	SRS28 (= SporoSAG)	39
TGME49_058370	Rhoptry kinase family protein ROP28	9
TGME49_115550	microneme protein, putative	5
TGME49_115730	apical membrane antigen, putative	15
TGME49_056040	LCCL domain-containing protein	2
Protein synthesis, turnover and processing/proteolytic proteins		
TGME49_009920	PAN domain-containing protein	55
TGME49_095640	endothelin-converting enzyme, putative	11
TGME49_112930	cystathionine beta-lyase, putative	4
TGME49_032400	PAN domain-containing protein	5
TGME49_059670	von Willebrand factor type A domain-containing protein	70
Proteins with oxioreductase motifs with known function		
TGME49_116190	superoxide dismutase, putative	14
TGME49_115260	alanine dehydrogenase, putative	8
TGME49_115480	acyl-CoA oxidase, putative	9
TGME49_034570	peroxisomal multifunctional enzyme type 2, putative	18
TGME49_027100	glutaredoxin, putative	13
TGME49_031900	acyl-CoA dehydrogenase, putative	8
TGME49_003500	alanine dehydrogenase, putative	3
TGME49_106450	short chain dehydrogenase family protein, putative (3-oxoacyl – [acyl-carrier-protein] reductase)	6
TGME49_010260	glycerol-3-phosphate dehydrogenase, putative	15
TGME49_047500	acyl-coenzyme A oxidase, putative	7
TGME49_003720	vitamin K epoxide reductase complex subunit 1, putative	5
Proteins with oxioreductase motifs with unknown function		
TGME49_026300	oxidoreductase, putative	4
TGME49_069120	oxidoreductase, putative	30
TGME49_013340	glucose-methanol-choline oxidoreductase domain-containing protein	9
TGME49_053150	oxidoreductase, putative	45
Transcription, nucleic acid metabolism and regulation		
TGME49_067410	scavenger receptor protein TgSR1, putative	2
TGME49_110160	piwi-PAZ domain-containing protein	18
TGME49_006400	DNA photolyase, putative	27
Signaling and trafficking proteins		
TGME49_055460	EGF-like domain-containing protein	12
Metabolic proteins		
TGME49_032580	acetyl-coenzyme A synthetase, putative	2
TGME49_071210	4-alpha-glucanotransferase, putative	19
TGME49_069110	ornithine aminotransferase, putative	8
TGME49_016640	homoserine kinase, putative	8
TGME49_019230	acetoacetyl-CoA synthetase, putative	10
TGME49_058200	glutamic acid-rich protein, putative	25
TGME49_042390	enoyl-CoA hydratase/isomerase family protein, putative	10
TGME49_082230	sulfate adenylyltransferase-adenylylsulfate kinase,	4

	putative	
TGME49_089930	phosphoenolpyruvate carboxykinase, putative	3
Membrane and extracellular proteins		
TGME49_078120	SCP-like extracellular domain-containing protein	10
TGME49_085690	notch (DSL) domain-containing protein	9

*Accession numbers of genes in *T. gondii* database http://toxodb.org/toxo/

5.2.3 Identification of oxioreductase patterns and motifs in protein sequences

The corresponding protein sequences of the 82 genes listed in Table 5.2 were all screened for motifs specific for: (1) oxioreductase activity; (2) oxidation activity, indicating potential involvement in the production of H_2O; or (3) peroxidase activity that could potentially catalyse dityrosine crosslinking. This process was carried out to identify any possible incorrect annotations in the database and to highlight homology to any oxioreductases. The amino acid sequences of all 82 queries were analysed using two online programs, ProSite and InterProScan (Section 2.2.6.4).

No motifs were identified in any of the hypothetical proteins screened using either program nor was there any BLAST homology to known oxioreductases using NCBI BLAST. InterProScan identified thirteen proteins with a motif or domain that gave them putative enzyme oxidative/reductive potential including three oxioreductases, several dehydrogenases, a glutaredoxin, superoxide dismutase and peroxisomal multifunctional enzyme list (Table 5.2).

Many of the genes with oxioreductase capabilities including acyl-CoA oxidase, acyl-CoA dehydrogenase and alanine dehydrogenase, were only expressed as proteins in the sporozoite and/or sporocyst extracts. These proteins were considered unlikely to be associated specifically with the oocyst wall and were eliminated from the list.

Other putative genes identified in this search with known oxioreductase enzymology, identified as having a known function in either *T. gondii* or another organism, were superoxide dismutase, peroxisomal multifunctional enzyme type 2, glycerol-3-phosphate dehydrogenase and short-chain dehydrogenase. A peroxisomal multifunctional enzyme type 2, or 3-hydroxyacyl-CoA dehydrogenase, is thought to play a metabolic role in *T. gondii,* the protein is usually involved in peroxisomal fatty acid beta-oxidation. A protein with a 3-hydroxyacyl-CoA dehydrogenase domain has previously been identified in lipid metabolism of the parasite (Edqvist and Blomqvist, 2006, Lige *et al.*, 2009). Similarly, glycerol-3-phosphate dehydrogenase (GPDH) has a known role in glycerolipid metabolism, and has been previously characterised in the homologous apicomplexan *Plasmodium* (Ling and Wilson, 1988). A short-chain dehydrogenase was also identified in the bioinformatics analysis of the *T. gondii* genome, classified as a 3-oxoacyl- [acyl-carrier-protein] reductase; it participates in fatty acid biosynthesis and has been previously identified in apicomplexans (Goodman and McFadden, 2007). Another protein identified in the search was a superoxide dismutase, which has previously been identified in *T. gondii,* and is thought to have a role in antioxidant protection and carbon metabolism in the parasite (Pino *et al.*, 2007). There were also no differences between the original annotations of these proteins in the *T. gondii* genome database, and the motif-analysed list indicating all proteins were annotated correctly.

With important, known metabolic and antioxidant roles in asexual phases of these organisms there was reason to believe that they were, therefore, less likely to have a role in the sexual phase oocyst wall formation. For this reason, these genes were also eliminated as potential candidates.

Oxioreductases without a known role were investigated further. Oxioreductase N- and C-terminal motifs (PFAM) and FAD/ADP binding (FAM) motifs were identified in *T. gondii* proteins TGME49_053150 oxioreductase (putative), TGME49_069120 oxioreductase (putative), TGME49_026300 oxioreductase (putative) and TGME49_ 0133405, a GMC oxioreductase domain-containing protein (putative).

5.2.4 BLAST analysis

The predicted protein database of *E. tenella* was then screened for putative oxioreductase homologues to the four *T. gondii* genes TGME49_053150, TGME49_069120, TGME49_026300 and TGME49_ 0133405, using tBLASTn and BLASTp (Section 2.2.6.5). The translated protein sequences for these four genes were used to BLAST the *E. tenella* database and identify *E. tenella* homologues of interest, detailed in Table 5.3, which includes the identity of each primary BLAST hit and its reported E-value.

BLAST analysis of both oxioreductase TGME49_053150 ($1e^{-132}$) and GMC oxioreductase TGME49_013340 ($2.9e^{-83}$) identified the same putative oxioreductase ETH_0003360 in *E. tenella*. Both E-values were strong so, to determine the true orthologue, the amino acid sequence of ETH_00033360 and BLASTx, BLASTp and BLASTn were used to BLAST the *T. gondii* protein database. TGME49_053150 (*Tg*OXIO1) was revealed as the closest homologue ($2.8e^{-144}$) in comparison to TGME49_013340 ($2.5e^{-83}$). The level of conservation was also addressed using ClustalW, where the conserved N-terminal and C-terminal GMC oxioreductase domains were highly conserved between the two proteins (Figure 5.2). Proteomic data for the *Tg*OXIO1 profile comes from washed oocyst wall extracts, and not sporozoites or sporocysts; thus, there is a high chance that the putative orthologue identified in *E. tenella* ETH_00033360, redefined as *Et*OXIO1 from this point on, could be responsible for producing H_2O_2 and inducing oocyst wall formation in *E. tenella* via dityrosine crosslinks. Therefore, this putative oxioreductase was selected for further investigation to determine its function in *E. tenella*.

Another oxioreductase, TGME49_069120, redefined here as *Tg*OXIO2, was also identified in *T. gondii* in this search, with data showing protein expression in the oocyst proteome, both in oocyst walls and sporozoites/sporocysts, cytosol and membranes. Its putative orthologue in *E. tenella,* ETH_00013835 (*Et*OXIO2), was identified by BLAST analysis of the *E. tenella* predicted protein sequence and is described as an unspecified product in the *E. tenella* database. Given the high

homology to the *T. gondii* oxioreductase, the amino acid sequence of EtOXIO2 was analysed with InterProScan and Prosite to identify what function the gene may have. Analysis showed that the amino acid sequence contained a short-chain dehydrogenase/reductase family signature (Figure 5.3). InterProScan defined the gene as a glucose/ribitol dehydrogenase according to signature motifs. Whilst analysis showed that *Tg*OXIO2 did not have a homologous reductase family signature motif as identified in the *Et*OXIO2 protein sequence, InterProScan still defined it as belonging to the short chain dehydrogenase/reductase family. ClustalW was used to assess the homology between the *T. gondii* and *E. tenella* protein sequences (Figure 5.3). Whilst a level of homology was observed between the two protein sequences, with a percent identity of 22.8% between the two sequences, there was a low level of conservation between the short-chain dehydrogenase/reductase motif highlighted in *Et*OXIO2 and *Tg*OXIO2. As a dehydrogenase, the predicted function of the *E. tenella* EtOXIO2 protein would be the oxidation of NAD^+/NADH in metabolic pathways. If it were a glucose dehydrogenase, it would have a specific role in the pentose phosphate pathway. Therefore, this gene was not selected for further analysis.

The fourth oxioreductase identified in *T. gondii* via bioinformatics analysis, TGME49_026300, was observed to have mass spectrometry data from the cytosol and membrane of sporozoites and sporocysts only, with no evidence for its presence in oocyst extracts. Therefore, it was suspected that this protein may have a function other than oocyst wall formation in *T. gondii*; thus, the gene transcript of the putative orthologue, ETH_00015230, in *E. tenella* was not investigated further.

Table 5.3: *Eimeria tenella* orthologues identified in BLAST analysis with *Toxoplasma gondii* oxioreductases

***T. gondii* protein**	**Gene ID***	***E. tenella* protein**	**Gene ID***	**E-value**
Oxioreductase	TGME49_053150	Oxioreductase, putative	ETH_00033360	$1e^{-132}$
GMC oxioreductase domain-containing protein	TGME49_013340	Oxioreductase, putative	ETH_00033360	$2.9e^{-83}$
Oxioreductase	TGME49_069120	Unspecific product	ETH_00013835	$2.6e^{-09}$
Oxioreductase	TGME49_026300	Oxioreductase, putative	ETH_00015230	$4.9e^{-20}$

*Accession numbers of genes in *T. gondii* http://toxodb.org/toxo/ and *E. tenella* database http://www.genedb.org/Homepage/Etenella

```
TgOXIO1     MAGHSSLRAFQLGKLVFLGALFLMAVSAEPGTEKAPENLASGSKFSTDAKGDRRRPKRFQ 60
EtOXIO1     MRVNSALVLSQAYDYVAWKLARAMPVLPPFGLAYKEALTATG-------------YKNYG 47
            *  :*:*   *  . *       *.* .  *        *:*              *.:

TgOXIO1     DVTVLETDEKLSPETKAALDSLLDADNEGKKPRHRSESTAEAEKEADRTRYAPPPAPTFV 120
EtOXIO1     DGDYPEPSLYVQPGYQWVGYSLFDAGNN----FQRNAS---------------------- 81
            *    *..  :.*  : .  **:**.*:     :*. *

TgOXIO1     PKFVPVDALTPTLSNSASHSEESEEDFSYDIQKEYEKARLKYRDAQVIKDKDEFDFIVVG 180
EtOXIO1     ------DALLERLPN--------------------------------------------- 90

TgOXIO1     CGCAGCPLASYLALKGQGRRVLVLERGLERTFEQTPMAMTVYGHGVAAADPTLMQVFNTR 240
EtOXIO1     -----------------------------------------------------LVVKTR 96

TgOXIO1     SGVRTHVATVMGGGTSIDLAIYVEETENYFKFMNKRFPAYKFHWPTIQRAYRYVRRIVAR 300
EtOXIO1     HTAR------------------------------------------------------- 100

TgOXIO1     RMPFDNYFGLRYQQALQWRGFKPIGGKIPADYSSSLQWIEFDLDQNPPRAKCVVYRPTAY 360
EtOXIO1     ----------------------------------SITFIDTDHGKT---ANCVVYRPTKY 123
                                              *: :*: * .:.   *:******* *

TgOXIO1     EDIKPQGTSFAASLPASSAHWGPWIGTFLAPKSVLIKSGEVDPTDFFQKKEAKFFRKVCL 420
EtOXIO1     QDIRPIG-DLWSLFPSSAADWAPWIGTFLEPTSVLVNAGVNDPT--------KKLRRVCL 174
            :**:* * .: : :*:*:*.*.******* *.***:::*  ***        * :*:***

TgOXIO1     TDRKGSQIILSAGAIHSAVILYKSGIGPIPQLKKIKTPVIVEHPYLGQRFSDRVFIPVSA 480
EtOXIO1     SNPGSGRIVVSGGAVNTPLLLYRSGIGPALQLKLLKVPQVLDSPAVGSAFSDRVFLSIPG 234
            ::  ..:*::*.**:::.::**:*****  *** :*.* ::: * :*. ******:.:..

TgOXIO1     FSKHYADELTPVPFAPIRRLSSLDPLSPKATTPRGRTRRDRHFSNKDEAEPSSGHDETKR 540
EtOXIO1     FLKHYSDSIP-----LINPSKSLRRLQANEDTGSKKVNRLQLLS-RPAFEPILTQEELEK 288
            * ***:*.:.      *.  .**  *..:  *   :..* : :* :   **   ::* ::

TgOXIO1     GIRKERRHQKGTFTAGDQVNVEGGEGKEEADQGVTRATGGISALGVSENDEDGAEVANAF 600
EtOXIO1     AITPEVR------------------------EAINRLMNQEVPVEIITVDEDDSAVSNRF 324
            .*  * *                        :.:.*  .   .: :   ***.: *:* *

TgOXIO1     NEDDQASGSSPARIFLMDLQKPIPPELSGKYLDVYVPEKGELDPPRVCQGMGIKKGGPPH 660
EtOXIO1     EP----PVILPDFLLFPGVLQYLSPALQ----------------PRVCQFMGLKYTGPRC 364
            :     .   *  ::: .: : :.* *.                ***** **:*  **

TgOXIO1     GCDKENYSIGRRTLSCSLMVAEEMSGGRAQEGIIYASRFIFPPLFRNDPLVDVIFEILRS 720
EtOXIO1     KSN----GINERNLGCSLVTMEELSGDRLAEGFIYASRYMFPTAFRKDPILDAVTEIIQS 420
             .:    .*..*.*.***:. **:**.*  **:*****::**. **:**::*.: **::*

TgOXIO1     CSEYRAPFAVAGLKPLCAIVVPIIKCFRKSLATFYFTAEPKSRGSVRLSANGMIDVNGNY 780
EtOXIO1     CSNYRSPFGIIFLKPLCVLAQPIVKCLRKAVAPFYFTSEPKSRGSIRLKPTGEVVVDPQY 480
            **:**:**.:  *****.:. **:**:**::*.****:*******:**...* : *: :*

TgOXIO1     LKDEQDLFDAVRGVSNLINAMNGDAYRGVLQPAGSLSCPVTVLNGLLDLILTVASMTSLF 840
EtOXIO1     LVDEQDLFDAIRGTATLVEQLNGDSYRGIIQEKGPQSCPLAILNGLLDILLTLASTTSPF 540
            * ********:**.:.*:: :***:***::*  *. ***:::******::**:** ** *

TgOXIO1     ITKPGNLPLIQKYLQDLLPANSRRLRRLVAVGDDVKPRRLEATDLGDLEEPGDYHQDGMN 900
EtOXIO1     LTHSSNFAEIQRYLQDLIPPESRRLRRLMVVNEGYR------------MKPAVYEDEEGN 588
            :*:..*:. **:*****:*.:*******:.*.:. :             :*. *.::  *

TgOXIO1     EPNTFILEDYSDHVIENVDEKKMEELGIKRKLEEAGFDFEAYRKHLDPNAEEDKDTVISE 960
EtOXIO1     E-----YEDYSAFVENDVDEAKIERLGIKRRLEAVGFDFDSYRAHVERTDESAKDPSLHL 643
            *      **** .* ::*** *:*.*****:** .****::** *:: . *. **. :

TgOXIO1     TANWFVSRRLQQQMEEDQHLTAEEKHTARRLMELQREKEDHAAEAAFEVSVQELADAKAA 1020
EtOXIO1     GT---LGRRLDAEIE--QILGPEKMDQLKKLADSIKSEQALRASKEDPVLKESIEDCRKP 698
             :   :.***: ::*  * * .*: .  ::* :  :.::   *.    *  :.: *.: .

TgOXIO1     CQKPCSENDACPAADVCCAAFNNSLCIRTQETEELERRLRAQMFEQKLTTGAPSRQAKRT 1080
EtOXIO1     CTKEAIQNHTCAKSDVCCTAYGNTKCIRLPGQPAFAEQQ---------------NKEPYT 743
            * * . :*.:*. :****:*:.*: ***      : .:                .:   *

TgOXIO1     EEGVANEKTAVYPEGFLGALPNALPELLNFVPSPQKSPPHPLLYEKPPPHDFPTYLNPPK 1140
EtOXIO1     SADAADSGGAVKKEEFL------------------------------------------ 760
            . ..*:.  **  * **

TgOXIO1     PEGFPDPSMFPAAPPAGTPPPDVLPYTNKQWAATFPPRLPNPYKPKEVAKYALTYMSSVW 1200
EtOXIO1     ------------APFFGLNTVSVAHADDEQWAATYPPRLPSVQNPKALAKFALTYMTSIG 808
                        **  *  . .*    ::*****:*****.  :** :**:*****:*:

TgOXIO1     HHSNTVPMGEVVNTDFEVIGTQGLSVIDASVMNQLTRLNPLATLIMFGVYGAMKKLKKAG 1260
EtOXIO1     HTYGSAPMGRVVDDKFQVKGVNGLSIVDASVLPQLTRMNPVFTIMALGRYAGEMMLGEQ- 867
            *  .:.***.**: .*:* *.:***::****: ****:**: *:: :* *..   * :
```

Figure 5.2: Conservation of *Toxoplasma gondii* TgOxio1 and *Eimeria tenella* EtOxio1 oxioreductase putative orthologues

The amino acid sequences of putative orthologues TgOxio1 and EtOxio1 were aligned using ClustalW. Prosite and InterProScan were used to analyse the amino acid sequence of each protein for signature motifs. The conserved N-terminal (blue) and C-terminal (green) GMC oxioreductase motifs are highlighted. Conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids are shown.

```
TgOXIO2        -------MECIGLSCFFNALVESPLLMIVACLGMFKIAGAVLNSLWCVVAALILRDCDFL 53
EtOXIO2        MRMLSLALGGLCLALFSRDRAEGSRHLPRQTLGFLSRP--LPSSSPEVTLSPSLRTAPQL 58
                      :  : *: * .  .*..  :    **::. .  : .*   *. :  ** .  *

TgOXIO2        TKILPEKKD------------------------------YWKDRTAWVTGACSGIGLSV 82
EtOXIO2        TPATPSAAAASDAASAAERSESTRCFAYPLTYGGGPNGGGEGEGRVALITGASRGIGLAI 118
               *   *.                                    :.*.* :***. ****::

TgOXIO2        ARLLAMRGCKIIMSSNRPDALKAAVDDIIKFCNKKGTQRKESDFLVLPCDMLHLETLDDT 142
EtOXIO2        AKTFAKGGVETIICVARDQAACDAAAAEVRALGAASEGHG--------VDVADSRSVSEL 170
               *: :*  * : *:.  * :*   *.   ::  .  .  :          *: . .::.:

TgOXIO2       VVKAKEWQGKIDFLFNNAGIIGVGALMPYKSDER--VVTIDLLAQMKLSKLVIPIMKEAG 200
EtOXIO2       CASLLEKHKKIDILVNNAGITRDNLFIRMKDNEWTDVIDTNLNSAFYFSLPLIKQMTKNR 230
               ..  * : ***:*.*****   . ::  *.:*   *:  :* : : :*  :*  *.:

TgOXIO2       FGHIIFTASMYSRFVTGGFASYCAAKHGLRAFAEGLDREFRAKGYNIAVTNICPGWVDTP 260
EtOXIO2       FGRIINMSSVVGVGGNPGQANYAASKAGLIGLTKSLAKEYANR--NVTVNAIAPGFIKSL 288
              **:**  :*: .   . * *.*.*:* ** .:::.* :*:  :  *::*. *.**::.:


TgOXIO2       IYTRAVGPAEMPQAVDYKSKKVTVGLSSDTVAAKILKSSSNQLTECWVSRNPELLIAYFN 320
EtOXIO2       MTDKMTEAAKAGALASIPAGRFGTPQEVADLAAFLASDQAGYITGKVIPIDGGMLFGSN- 347
              :  : . .*:    ..  : :. .  .   :** : ...:. :*   :. :  :*:.

TgOXIO2       YYVPFAARIFLDISARSMCAATDETVADIIAQKAVNNPRAELR 363
EtOXIO2       ------------------------------------------
```

Figure 5.3: Conservation of *Eimeria tenella* and *Toxoplasma gondii* oxioreductase orthologues

The amino acid sequences of putative orthologues TGME49_069120 (*T. gondii Tg*OXIO2) and ETH_00013835 (*E. tenella Et*OXIO2) were aligned using ClustalW. Prosite and InterProScan were used to analyse the amino acid sequence of each protein for signature motifs. A Short-chain dehydrogenase/reductase family signature motif was identified in the amino acid sequence of *E. tenella* 'unspecified product' ETH_00013835 (green). Conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids are shown.

5.2.5 Identification of putative oxioreductases within a proteome of *Eimeria ten*ella oocyst wall extracts

In addition to those genes identified through bioinformatics analysis, a number of putative oxioreductases were identified from mass spectrometry data obtained from washed oocyst wall extracts (Flowers, 2011). In work described by Flowers (2011), oocyst walls of *E. tenella* unsporulated oocysts were bleached to remove contaminants, washed with detergent, solubilised in lauryldimethylamine-oxide (LDAO) and analysed by mass spectrometry. Thirteen proteins were identified that may be associated with the oocyst wall, or oocyst wall formation of *E. tenella* and that were unlikely to act as a structural element of the oocyst wall, including four proteins with predicted oxidative function.

Of the enzymes found to have an association with the oocyst wall of *E. tenella,* two amine oxidases were identified. The first, ETH_00024850, was annotated as an unspecified product on the *E. tenella* database, however, the *T. gondii* homologue, TGME49_042420, was annotated as an amine oxidase (monoamine oxidase), flavin-containing domain protein. ClustalW was used to analyse the conservation between the two proteins. Figure 5.4 shows the level of homology between the two proteins with a 40.1% sequence identity measured. The amino acid sequence of the *E. tenella* gene, from this point on referred to as amine oxidase 1 (*Et*AO1), was also screened for protein motifs using InterProScan. An amine oxidase protein motif (Figure 5.4, highlighted in green) was identified, with a high level of conservation between the two genes around this amine oxidase motif. The *T. gondii* homologue (*Tg*AO1) was detected previously in this study in the oocyst proteome, but there was evidence to suggest it was also present in the tachyzoite proteome, and so it had been eliminated in the earlier high-stringent bioinformatics searches. However, given the proteomic evidence of a high association with the *E. tenella* oocyst wall, as shown by Flowers (2011), and the identification of an amine oxidase motif in the predicted protein, it was decided to investigate the expression of *Et*AO1 in *E. tenella* as an amine oxidase with potential involvement in oocyst wall formation.

The second amine oxidase protein identified in mass spectrometry, thought to have an association with the oocyst wall, was ETH_00028385, a putative copper-containing amiloride-sensitive amine oxidase, redefined here as *Et*AO2. Screening of the amino acid sequence for protein motifs, confirmed it as an extra-cellular protein containing a MAM domain and Cu-amine oxidase domain. BLAST analysis identified the *T. gondii* homologue for this gene, TGME49_286782, as a putative copper-containing amiloride-sensitive amine oxidase (*Tg*AO2). Mass spectrometry-based expression evidence available for this protein on the *T. gondii* database showed that it is present in the oocyst proteome (Fritz *et al.*, 2012). A high level of homology was observed

```
EtAO1          RGTSGDNLTAQEEEGDCGRSLLTETHPVRLRVEVNRR---------LRHRASHGLPQAEK 1074
TgAO1          EEEEEDREEERDKEKEGGEEEDLSLWPVRVRVRAPLETEKFKQSEGVHFRRKQGAPEGEE 1317

EtAO1          -------------------QEEPPTHVWAHAKYCVVAVPLAQL----------------- 1098
TgAO1          DLRQEASGEKKRIFEADAEENSDAIVFDVYAKFAIVAAPVSQLGESFPGAGSTDRVTAFR 1377
                                   ::. .  . .:**:.:**.*::**

EtAO1          ----ALAPLPPTRALTAEG------TVPPTLGLIDFSPPLGPDRRRALARYRMGCHNKVV 1148
TgAO1          TRKKALKTFPDDKKVDGETQEDLLPALPRSSGLLQFDPPLSLEKQQALTRIAMGVHNKVI 1437
                   ** .:*  : : .*       ::* : **::*.***. ::::**:*  ** ****:

EtAO1          IRWHPDDVFWHGRGLQLNCLDQKFQFLNLHAYGKPGCLLAHCFPPFSSGYGGLLTDQEVV 1208
TgAO1          LRWHPADIFWERTQLQLNCLDQKFQFLNLHAYGKEGCLLAHSFPPFSNGYGGLQSDDAVV 1497
               :**** *:**.   ******************* ******.*****.***** :*: **

EtAO1          AEVLALLQRMFGISDSSFPRPVDFVVSRWDEDCFSRGSYSTPGVNAYDNDLDIMRAPHPA 1268
TgAO1          AEVLAVLQRMFGLPDKKFPPPVDFIVTRWQDDPFARGSYSFPAVNAFDDDTEILRTPHPA 1557
               *****:******:.*..** ****:*:**::* *:***** *.***:*:* :*:*:****

EtAO1          NDPRVVFCGEHLSKAYFQCVDGAYDTGLRAAEAVAHECLQLPLPPREDGRNFALDYFC-L 1327
TgAO1          ENPRVLFAGEYVSKAYFQCVDGAFDTGLRAAECVAHEGLHLPLPPREDRTPLALDVFTGC 1617
               ::***:*.**::***********:********.**** *:********   :*** *

EtAO1          PPTTTVSTSARRPHASPGLEGEQREGRIDNILLSADKIEPSGESVAEPSRQPSQGLLPQC 1387
TgAO1          LSSRQADACSEFALASPAKKREANQEKREGNRAKREGKRRKRQKRRENELSLSTHMLTPC 1677
                .:  ..:.:. . ***. : * .: : :.   . :  . . :.  * . . *  :*. *

EtAO1          PFSGFPMPPLPRVLCGHYLTDQSDLGLTDAEGEHWAETGSSRRSFQTCESERRVPQVAAA 1447
TgAO1          PFTGFPVPSVPTALRGYYITDMSDAGLTDGEGSGPELSFSLSAEGRRCSAG-KRAELELC 1736
               **:***:*.:* .* *:*:** ** ****.**.    : *   . : *.:  : .::  .

EtAO1          EEYFLAHCLAFIKRGSLRLLEAFASQHP------ISSNRSRFLVEEWIDGVVDDEGFACE 1501
TgAO1          EERFLRRCLEFLEAGGEEEKGKGAKEDSGLLGGFFSFPSQRFFYDAFFGHPTLCGSEENR 1796
               ** ** :** *:: *. .     *.:..      :*   .**: : ::.  .   .   .

EtAO1          AKTAGFSGDGDDEAAAGDRPPTVPPVQCLSETASSFLQSLPTAAQRLFYRSFLSEFPLTP 1561
TgAO1          VPCSGESGRRSSFADPTVLCAVLTSALAFTREEIEEALATHARQGRSSIRGSFFSDADAS 1856
               .  :* **  .. * .    ..:... .::.   .   :  :   *   *. : . . :.

EtAO1          SPSVGDAEQGQRSTPTLLMQEQQGCTQNAPDDLQQQPQRLPLLS---------------- 1605
TgAO1          AAVEGKVEREGVELRRSALRLDEQAVEPVSEEFQELRDFLAFLGTRDPFARRRLLLPLWV 1916
               :.  *..*:   .     :: :: ..: ..:::*:  : *.:*.

EtAO1          ---------DFLDALIMCTSSNNGMKNNQESPLKETALLLSQAFARRLKRGSGGKTLNME 1656
TgAO1          AAQSARERRGQVKGLEEVDEERRGASNEDEPQETEASLARETVETERGKADDEGVASCAA 1976
                         . :..*    ....* .*::*.  .*::*  . . :.* * .. * :

EtAO1          AFLLLLAKNNEVIRDEVLKPIRDMAHAFPKTDSGVGQ----------------------P 1694
TgAO1          SSPLIERRFGAEGREQVNRRARRTATTVAEDRLSHAESELDSEGEGAERKERQRVRWVYL 2036
               :  *:  : .   *::* :  *  * :..:   . .:

EtAO1          REAEAVGSPSDPAAAVAVPLEVFVKELMLPAFPAYARNTRP-----PRMAAVSVGVLLME 1749
TgAO1          RAAEVLTRHARAAAFLPSLSPVGLASFLVDSLPRLPPKKRTGASAKSGRTADRETLRTLR 2096
               * **.:   : .** :.    * : .::: ::*  . :.*.     .  :*    :  :.

EtAO1          TLERLCAAFPPLNSKEQLHEMATVVDSLVLKNLANSYFISAIGGGSS-------LRVAER 1802
TgAO1          EKKSRNSDLAAHATRAHTHEDAENTPDLVERSRSFPPSLSREVSPECGNETRMALIVSRS 2156
                 :   : :..  :: : ** *  . .** :. : .  :*   . ..       * *:.

EtAO1          QTERNNTSEPLQHEQLCWLCRGEGDLLLCDGMPASALDSKDDMSGSPSAPAASKSCRHVF 1862
TgAO1          QAERENIRGDCGGVKGGCSEEKTEKNSRSLGVTRKVLKGARRRLPTASEFKRELENRLLH 2216
               *:**:*        :     .   .   . *:. ..*..      :.*    . . * :.

EtAO1          HIG-CAFPPPTPAELQPNAHWSCP-LCKAKAFSQRMNKQQNRGDRREQLEEDVHEERWLS 1920
TgAO1          HLARSAFPSDIEATALRELLRAAQDERDIISCCHRRQAEAGKGVAKEDLEVQGDSSCLSP 2276
               *:. .***.   *    :   :.    .  : .:* : : .:*  :*:** : ...   .
```

Figure 5.4: Conservation between *Eimeria tenella Et*AO1 and *Toxoplasma gondii Tg*AO1 homologues

ClustalW was used to analyse homology between *T. gondii* amine oxidase, flavin domain-containing protein, *Tg*AO1, and *Et*AO1. Prosite and InterProScan were used to analyse the amino acid sequence of each protein for signature motifs. Amino acid sequence is shown between 1023-1920, depicting the amine oxidase motif (green) for *E. tenella* 'unspecified product', and *T. gondii* 1257-2276 (blue). Conserved amino acids '*', highly similar amino acids ':' and similar '.' Amino acids are shown.

between the two proteins with 61.68% sequence identity measured (Figure 5.5). Prosite and InterProScan identified the MAM domain, and the copper amine oxidase signature in both protein sequences with a high level of conservation in amino acids observed between these domains (Figure 5.5). Given the high level of evidence for an association of this gene with the oocyst wall of *E. tenella* (Flowers, 2011), and the expression of the homologue in the *T. gondii* oocyst proteome, this gene was also selected as a gene of interest in oocyst wall formation.

A putative quinone oxioreductase was also identified in extracts of *E. tenella* oocyst wall that underwent the most stringent wash in the study by Flowers (2011), suggesting that the protein may have a strong association with the oocyst wall of *E. tenella.* The protein, ETH_00043030, was found to be alcohol dehydrogenase related, with a possible alcohol dehydrogenase zinc binding C-terminal motif ($1.5e^{-18}$), also denoted as a NAD(P) binding domain, using InterProScan. BLAST analysis identified a homologue in *T. gondii,* TGME49_044870 (*Tg*QO1); however, there was very little data available for this gene in *T. gondii*, with only a single peptide detected in mass spectrometry analysis of the tachyzoite phosphoproteome. InterProScan also identified this protein as related to the alcohol dehydrogenase family, however, with a higher similarity to a NAD(P) binding domain ($1.2e^{-29}$). A high level of conservation was observed with the amino acid residues that made up the noted motifs (Figure 5.6). Nevertheless, this gene, defined from this point as *Et*QO1 was selected for further analysis due to the strong association observed in mass spectrometry analysis of *E. tenella* oocyst wall extracts (Flowers, 2011).

A fourth putative oxioreductase gene identified in the mass spectrometry study of the *E. tenella* oocyst wall extracts was the oxioreductase, glutathione peroxidase. Defined as a glutathione/thioredoxin peroxidase, this gene (ETH_00011565) was found to have a glutathione peroxidase domain, identified using Prosite. BLAST analysis identified two homologues in *T. gondii.* The first was a thioredoxin-dependent peroxidase (*Tg*TPX1/2) ($6.6e^{-41}$) shown to localise to the apicoplast and mitochondrion of *T. gondii*; it is thought to play a role in the antioxidant pathway of the parasite (Pino *et al.*, 2007). The second, a putative TPX (PRX2) (TGME49_066130) ($2.2e^{-33}$), had expressed sequence tags detected only in sporulated oocysts (Pino *et al.*, 2007). A level of homology was observed between the three sequences around the glutathione peroxidase profile motif (Figure 5.7). The presence of the glutathione domain in the putative protein sequence of this gene, defined from this point on as *Et*GPX1, indicated it was important to perform further analysis, to determine expression in *E. tenella.*

Table 5.4 summarises the five oxioreductase genes selected for further investigation based on bioinformatics analysis and mass spectrometry data. All five were consider potential candidates in

oocyst wall formation due to the presence of an oxioreductase motif within the proteins, and their specificity to oocyst wall preparations in *T. gondii.* As all proteins identified in *E. tenella* have an as yet unknown role, they were considered to be candidates for having a role in the oxidation process of oocyst wall formation either by providing the source of the H_2O_2 or acting as the peroxidase needed for oocyst wall formation.

```
EtAO2     MTKFPRAPLKRPERRQRLFTLALLAGGILGSTVGSCAASFEGTSPVDPTRLLPRRLAPQE 60
TgAO2     ------------------------------------------------------------

EtAO2     CKTYVYTEIDHEPTEMSDRGSCPTGGGKRAFLLDTSPEGAGAPALLEVSPVQSQQYSCRA 120
TgAO2     ------------------------------------------------------------

EtAO2     SRPCSLVVAGFGLQAEDSIAVLKGIEDCPPTSLRNAQIITSTRHITSSSAATFNFHIRPQ 180
TgAO2     ------------------------------------------------------------

EtAO2     KAKTELPVFDLGPIQGSTFTVCYSPYLATRGEAPSAKDYHYRAGTVTIRDEMLNCDFEEG 240
TgAO2     ---------------------------------------------------MNCDFEED 8
                                                            :******.

EtAO2     MCGLMSFVIDPRMSLEWRRHSGATETAGTGPSADHTIGRRPGKRGHYLYMDSPGYMLGES 300
TgAO2     MCDLMIYRDHEDMDLYWVRHSGSTETPGTGPSFDHTYGTQ--RDGSYIYMDSPGYVAGAS 66
          **.** :  .  *.* * ****:***.***** *** * :  : * *:*******: * *

EtAO2     ASLLAPPQIYPKGLYCARMWYHMYGEDVNSLRVYMRTLEGDAEALGDWGKPYLLRVGDHG 360
TgAO2     ATLIGPPTIFPQGLFCARLWYHMWGDDVNSLRVYMHELKDDQDATGNWGKPRLLRVGDHG 126
          *:*:.** *:*:**:***:****:*:*********: *:.* :* *:**** ********

EtAO2     DDWLEAGFEFSSDGHTAQQLVIEALAGFSEKGDIAIDDLRIVSGKCPEQISRDVTQQDYI 420
TgAO2     DEWLEGGFEFQSDGKTAQQLVIESLAGFSEKGDIAVDDLRIVPGRCPEEIKRDSYNADYI 186
           *:***.****.***:********:***********:******.*:***:*.**  : ***

EtAO2     CGEVRLLTGNYAEDKQWFVEGAVSCAGKGYSLSSMQGRWVPCCVPRYGSYTVVLRDSYGD 480
TgAO2     CGEVRLLTGNWSQEKSWFVEGAVSCAGRGYSYDNLKGNWVPCCVPRYGQYTLVLHDAFGD 246
          **********::::*.***********:*** ..::*.**********.**:**:*::**

EtAO2     GWNRSRLEFRFFGDLITFGEDFN----GTEKKYKLNIGLVQIQRVEGSEDRIGIQLQVAE 536
TgAO2     GWSGSKLEFRFFDQVQEFGSDFMGIEDGAEKKYTLNIGLLQITRIEGSQDRIALEIQVAQ 306
          **. *:******.::  **.**     *:****.*****:** *:***:***.:::***:

EtAO2     PRTYVWCGAAQSGTPPPTVEILKRYGIRSQNPTSHAGEVLRMEIANKPGSRRVVMPSTTY 596
TgAO2     PNSYVWCGAALSGSPAPTVEVLKKYGIRSQEQTKKPGDRLRMEIANKPNDRRILQPATDY 366
          *.:******* **:*.****:**:******: *.:.*: *********..**:: *:* *

EtAO2     NVYCYAEASTANIPEPGGGHGTHRERMDDAQVAASRMAVTTDSTPPNLSIQGTEAHLNDV 656
TgAO2     NIYCYAEPAAYAADQ------NQVMMMDDAQVAASRVMVTTDNKPPEITVESVDEEYTEA 420
          *:*****.::    :      .:   **********: ****..**::::::.: . .:.

EtAO2     IVRLTSDEPSTVWCLARDEKKE--AAAGGEPPSVEALKKGNKIQILHGEEGSPQALSITG 714
TgAO2     SVRIRSDEVATVWCAAEEVSKVPTDDAGQSPPRSVIKKAGNKIQIVKGTENTPQILKLKN 480
           **: *** :**** *.: .*     ** .**     * ******::* *.:** *.:..

EtAO2     LTPDNHYTVYCFAQDAALPKGNGISLSQVKELSFSVQTKSKVPELSILSYKAFQKGFRVT 774
TgAO2     LMPDTKYRVYCYAEDRALPKPNFTGPEQVSAKAFDLVTPSKVPELKIESYKAFVRGFKLS 540
          * **.:* ***:*:* **** *  . .**.  :*.: * ******.* ***** :**:::

EtAO2     VQADTPAKVWCGAAMAGSAFPSREDVRRVGATTEIEDADQKTELEIRGVPPNTRYTIYCF 834
TgAO2     VKTDTPAKVWCGAAMAGSAFPKKEDVKRVGATAEVTDTTKPADLEIRGVPPNTRYSIYCM 600
           *::******************.:***:*****:*: *: : ::*************:***:

EtAO2     ASSRGGSAEMTDAQMWAKALEVTSFGRFCDMPVEPADLEVADKPTLFDPLTSTEEFLVRE 894
TgAO2     ASSRSGASEMTDQQMWDGALEVTSFGKFCDMPIEPAVLSEG-PATPFDPLTPPEEFMVRD 659
          ****.*::**** ***  ********:*****:*** *. .  .* *****..***:**:

EtAO2     FMGRQRNLNIEGVYRINLFIDKEKIIEHLDKGGPAPRRYARVRVGTCKDREGAYRQYKVG 954
TgAO2     FMNTQKNLRLEGIYRINLFLDKDAILNHFEKGAEFPKRYARVRAGVCNNGKGAYRQFKVG 719
          **. *:**.:**:******:**: *::*::**.  *:******.*.*:: :*****:***

EtAO2     PLDVRSADQMTLEPVGAVVPTDCGGYNPQGVFGRRLLAVDEEEELAEILQESFGFGFGKS 1014
TgAO2     PLDTEDAKNLKIEDIGEPVETDCGGYNPQGVFGRRLLDQEQENELDAFLHRVFGYHFGVK 779
          ***...*.::.:* :*  * *****************  ::*:**  :*:. **: ** .

EtAO2     RCQFEHKKSHKEHHEGCLEFGPMLYEEGDDKKIITIWAGLRTPTG------DKVPFYFLL 1068
TgAO2     ACVYG-------EVDNCLNFGPLIYKQKGDQ-LEEVWIGLKTPPKEQGDYEHNLPFYIVA 831
           * :        . :.**:***::*:: .*: :  :* **:**.       .::***::

EtAO2     PAPEGVDAQVLQKPDDREELTFKLDKVKGYWYDGQQFSTLREFVEAYRTGKNFKKVTPEI 1128
TgAO2     EPPVADEKVDVVNAESRDRVALLLEHVKGYSLNGVQFDSLDEMFETVDTAAN-----PEL 886
           .* . :   : :.:.*:.::: *::****  :* **.:* *:.*:  *. *     **:

EtAO2     LEERRKMENQEQEHNIQKERALALRRLAPAWLAPNPEGRVGLEYRAAPEHIEPQGRRYTI 1188
TgAO2     TVEKLRDAMSEVRNKREEQRANAVRRLAPYNIGPNPKGREGLEYRVGPEHIEPQGKRYTI 946
            *: :   .* .:: :::** *:*****  :.***:** *****..********:****

EtAO2     RQSPGQESYTINYAGWEFVINMDRDTALRLWDIKFQGQRVVFEMGMMEALAHYSVAERNW 1248
TgAO2     THNPDSTSYSINYAGWTMTVTNDRDTSLKIWNLRFKDQRYAFELGVNEALAHYSVAERHW 1006
           :.*.. **:****** :.:. ****:*::*:::*:.** .**:*: ***********:*

EtAO2     YFIDSWYGGLGSASRRLHPGIECAKTGQLLFDGGSLCIFEKDFARPLRAHWKSGTLRDGA 1308
TgAO2     YFMDSWYGGLGAAARRVHPGIECAKTGTVLFWDKSLCVFEEDKARPLRSHWKAGVLRDAV 1066
```

```
                 **:********:*:**:********** :** . ***:**:* *****:***:*.***..

EtAO2            PHMALVLRHMITVSNYDYITDYYFHVSGWFEASVSFTGELYAGVEVPWYSARQRRHGTQV 1368
TgAO2            PHYALVLRQIITVSNYDYITDYVFHQSGFFEATMAFTGELYAGVEVPWYSARQANHGTQV 1126
                 ** *****::************ ** **:***:::****************** .*****

EtAO2            SGSMRMGALHGHLAVWKVDFDLTPDYRSNSVVFSEIVHDTARPGAIKLDQWIGEKELDGY 1428
TgAO2            TGSMRMGALHAHLASWKIDFDLD-DYKSNSIAWHEIKHDTARPGALKIDTWFAETEEQAY 1185
                 :*********.*** **:****  **:***:.: ** ********:*:* *:.*.* :.*

EtAO2            IAYNSTRPIHYTIVNEDHNVYGNVGGMTVLPYPTVAIPNPQFELYTGPCAWAKYRVATTV 1488
TgAO2            YKFNDTRPLHYQVINEDHNVYGNVGGFTVLPLPTVAIPNPEFELYTGPAAWAKYRILTTL 1245
                 :*.***:** ::************:**** ********:*******.******: **:

EtAO2            RHPDESEATLPRDNKYALKPAVSLDRYLRVRSKWIALEPFHPWTKNAYATLAHEAVTNLF 1548
TgAO2            RKEDEFHGSLPRDNKYALRPAVSIDRYIRD----------------------------- 1275
                 *: ** ..:*********:****:***:*

EtAO2            QLALFPNNETIRKADLVTWISSAVWHIPVVEDMPLTLAQGNTLGWLVKPHNYYMEDMSMD 1608
TgAO2            -------NEPIRKKDVVTWVASSVWHIPVIEDMPQTLSLGNTLGFLVKPHNFFTEDPSMD 1328
                        **.*** *:***::*:******:**** **: *****:******:: ** ***

EtAO2            LHNAIGGAVQDPGTCALIRQEMPKYGDAPQN 1639
TgAO2            LHNSLGGAVQDPGTCAIIRQETEKYLQE--- 1356
                 ***::***********:****  ** :
```

Figure 5.5: Conservation between *Eimeria tenella Et*AO2 and *Toxoplasma gondii Tg*AO1 homologues

ClustalW was used to analyse homology between *T. gondii* amine oxidase, flavin domain-containing protein, *Tg*AO2, and *Et*AO2. Prosite and InterProScan were used to analyse the amino acid sequence of each protein for signature motifs. Amino acid sequence is shown between 233–408, depicting the MAM domain (green) and 1313-1326 copper amine oxidase topaquinone signature (blue) for the *Et*AO2. The amino acid sequence is also shown between 1-174, depicting the MAM domain (yellow), and 1071-1084, depicting the copper amine oxidase topaquinone signature (purple) for *Tg*AO2. Conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids are shown.

```
EtQO1          MKPLAAAQRMRAILVGPPSLPKGSSFGGPLEGPGGPPGSPGGASRGPHGGPQGAPYLYLS 60
TgQO1          MVLSAPLRRLPHMMR--AVLVDAAPLSGP----------PAVEAKN-----QQSFKLFLG 43
               *. :*:  ::   . * ..:.:.**          *.  ::.     * :  *:*.

EtQO1          ECPLPKLRP--NEMLVRVGAAGVNRMDLLQKAGKYPPPPGTTDILGPEAAGIVVSSS-VY 117
TgQO1          SAAMPQARKEKNEILIRVCAAGVNRMDLLQKRGKYPPPPGASTILGPEAAGVVASSGKRF 103
               ...:*: *   **:*:** ************ ********:: ********:*.**.  :

EtQO1          KEGTPVACLLPGGGYAEYVAADPALCFPIPSNLTIREAACIPENWITAYQLLEEVARLS- 176
TgQO1          REGDRVMALLQGGGYAEYVAVEEGLCMPIPSTLSFSQAAAIPENWLTAYQLLHFVAGLQP 163
               :**  * .** *********.: .**:****.*:: :**.*****:******. ** *.

EtQO1          ------------------------DFKANGGK-----------PSG----IKHALVYAAG 197
TgQO1          ASTPSSACFSSSPPSSATATADSLNSSEHGERQTDACPFVCAGASGRRLPIRSVLIHAAA 223
                              : . :* :           .**    *: .*::**.

EtQO1          SGVGVALLQLFRVFIPSVSVIAVAGSNEKLQRVKELGAAVCLNYKELGDSLGDEVLKATG 257
TgQO1          SGVGTALVQLCRLLAIPT-IIASAGSDEKLQLCKTLGATHLINHRAVEGRFSEAVKNATH 282
               ****.**:** *::  .. :** ***:****  * ***:  :*:: : . :.: * :**

EtQO1          GQGVDIVLDCVGASLVNQTLKSLKTDGT-------------------------------- 285
TgQO1          GLGVDLLLDPVGASFMQENAQSCALDASWVLYGALGGVRVPSFDLQLLFAKRIRLLSSTL 342
               * ***::** ****:::::. :*   *.:

EtQO1          ------------------------------------------------------------
TgQO1          RSQDLSYRESLVRSFEETILPKLADSSLRVILDSIYTASEADQAHQRLEKNENRGKVVLT 402

EtQO1          -----
TgQO1          FPCPS 407
```

Figure 5.6: Conservation between *Eimeria tenella* *Et*QO1 and *Toxoplasma gondii* *Tg*QO1 homologues

ClustalW was used to analyse homology between *E. tenella* quinone oxioreductase, *Et*QO1, and *T. gondii* TGME49_044870 (*Tg*QO1). Prosite and InterProScan were used to analyse the amino acid sequence of each protein for signature motifs. Amino acid sequence is shown between 199-285, depicting the alcohol dehydrogenase zinc binding C- terminal motif (green), for the *Et*QO1; and 225-368, depicting the NAD(P)-binding site (blue) for the *Tg*QO1. Conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids are shown.

```
EtGPX1          ------------------------------------------------------------
TgTPX1/2        MLPLCASSQMYLGQECLSCPVLPRLCLATCLGALYGRAFWRTNKRTAGHGKKWLGAHGKE 60
TgPRX2          ---------------------------------------------------------MESP 4

EtGPX1          ----------MLLTRATAAAAPRVFSRLLPRTCWAASRAPAAAAAATAPAAAAAAAAVTR 50
TgTPX1/2        KFHCFSGSFFLAKTLTSAKSFFGRFSSLVPPSVFTTPRKQSAFRFSHLIPTRSPLSAQYP 120
TgPRX2          TTLTEERADVPAANIGAALEQPDEVPHLSPVIRQVDIIEEKTSGDDNERFVKIVVKDFEG 64
                              .  :*      .. * *    .      :        .

EtGPX1          QQRTMWLFSAWRS------AAVPPELLAADLWSLAAKDPEGKEVPLKDFKGKVLLITNTA 104
TgTPX1/2        GLKFLSSFSFLRMGIGGSRAAFAPDQIPVSFSTITFNDIYGVQRSLGEWDGKVKIVVNVA 180
TgPRX2          QERGTFCSDPEVIP------VFGPGNIPAAFVTIRTRDLDGNEHVMEEYDGKVKLITNVA 118
                  :     .           .. *  :.. : ::  .*  * :  : ::.*** ::.*.*

EtGPX1          TKCGFTQQHLKQFHDLKEKYGDQGFEILAFPTLQFKQQEEKDPQKMCEVYKSFNANFPIF 164
TgTPX1/2        SNCGLTKAHNKEFIELREKIGTDAFEILAFPSRQFANQEFADIAETQQFCERVKIPFPVF 240
TgPRX2          SLSKDARKTYDELRQIYDRFLPQGFEILAFPCQQYVPEEYVHTEDIKTFLADHNVGFPVF 178
                : .  ::   .:: :: ::   :.*******  *:  :*  .  .   .    :  **:*

EtGPX1          AITEVNGPNANPIFQYCKFNTDELYS------KGQLQAIGWNFGKFLLDKEGKVFKYFGP 218
TgTPX1/2        TTSDVNGPETNPVFLYCKWNSDSFYHPVKNSKSAKLSDIGWNYGKFLVDKDNGVYKYYGP 300
TgPRX2          EMTAVNGPETHPVFLYCKWNSDEFYR------DGQLNNLTGHFGKFLLDRDNRVYKFYPP 232
                  : ****:::*:* ***:*:*.:*       ..:*. :  ::****:*::. *:*:: *

EtGPX1          RQSPSEMEEDILLLLRGEAHGKQRNKDGLLQ--------------- 249
TgTPX1/2        RTKPLEMEEDIRKLIAGQAKGMKRNAAGELKPL------------- 333
TgPRX2          DTNPLRLVEDIKKLLNGELTGKIRGPDGKLYHQPGTVSQENQTTWS 278
                  .* .: ***  *: *:  *  *.  * *
```

Figure 5.7: Conservation between *Eimeria tenella Et*GPX1 and *Toxoplasma gondii Tg*TPX1/2 and *Tg*PRX2 homologues

ClustalW was used to analyse homology between *E. tenella* glutathione peroxidase, *Et*GPX1, and *T. gondii* TgTPX1/2 and TgPRX2. Prosite and InterProScan were used to analyse the amino acid sequence of each protein for signature motifs.

Amino acid sequence is shown between 65-239, depicting the glutathione peroxidase profile in *Et*GPX1 (green), *Tg*TPX1/2 (blue) and *Tg*PRX2 (yellow). The glutathione peroxidase active site was also identified in *Et*GPX1 between amino acids 95-110 (bold). Conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids are shown.

Table 5.4: Five oxioreductases identified in *Eimeria tenella* for further analysis as candidates for a role in oxidation during oocyst wall formation.

Gene ID*	Gene description	Mass spectrometry evidence: peptide count		
		***Toxoplasma gondii* oocyst proteome**		***Eimeria tenella* oocyst proteome**
		Oocyst walls	**Sporocysts/ sporozoites membrane**	**Oocyst wall**
ETH_00033360	Oxioreductase (*Et*Oxio1)	45	13	
ETH_00024850	Unspecified product (*Et*AO1)			11
ETH_00028385	Amiloride-sensitive amine oxidase, copper-containing (*Et*AO2)			74
ETH_00011565	Glutathione peroxidase (*Et*GPX1)			17
ETH_00043030	Quinone oxioreductase (*Et*QO1)			9

*Accession numbers of genes in *E. tenella* database http://www.genedb.org/Homepage/Etenella

5.3 Discussion

Bioinformatic analysis of the *T. gondii* database was employed to identify oxioreductases that may have a role in the oxidation of tyrosine-rich proteins destined for incorporation into the oocyst wall of *E. tenella. Toxoplasma gondii*, a coccidian parasite like *E. tenella*, shares many similarities in the development of a resistant wall around the developing zygote of the parasite before it is then shed in faeces as an unsporulated oocyst (Belli 2006). There is considerable evidence for dityrosine cross-linking, as an important element in the composition and structure of the oocyst wall of *Eimeria* (Belli *et al.*, 2006b). Likewise, 3,4-dihydroxyphenylalanine (DOPA), a derivative of tyrosine implicated in the synthesis of extracellular matrices via DOPA-mediated quinone tanning, may be formed with the assembly of the oocyst wall (Belli *et al.*, 2006b, Belli *et al.*, 2003b). The observation of autofluorescence in *Toxoplasma* oocysts (Lindquist *et al.*, 2003) suggests that dityrosine bonds, known to result in autofluorescence as observed in *Eimeria*, may be present. The recent identification of six tyrosine-rich proteins in the oocyst proteome of *T. gondii* (Fritz *et al.*, 2012) further supports the idea that these proteins may be cross-linked through their tyrosine residues, suggesting that it shares a common mechanism of oocyst wall assembly with *Eimeria.* This suggests that the genes and enzymes involved in oocyst wall formation would be somewhat conserved between the two parasites. The *T. gondii* database was selected for an initial screening of oxioreductase genes as there are extensive biological datasets accessible in the one database on http://toxodb.org/toxo/ (Version 7.2); *E. tenella* homologues of these genes were then identified by BLAST analysis of the *E. tenella* genome database.

The parameters used in this chapter utilised mass spectrometry and microarray data previously reported for *T. gondii* to identify 82 proteins of interest. Within this group of proteins, several known proteins characteristic of the oocyst wall of *T. gondii* were identified, including an oocyst wall protein, TgOWP2, (TGME49_009610), localised previously to *T. gondii* oocyst walls (Possenti *et al.*, 2010) and homologous to a previously identified *Cryptosporidium* oocyst wall protein (COWP) (Ranucci *et al.*, 1993, Spano *et al.*, 1997); and also six tyrosine-rich proteins (hypothetical proteins; TGME49_120530, TGME49_11980, TGME49_087250, TGME49_081590, TGME49_037080 and TGME49_116550) that may play a similar role to *Eimeria* tyrosine-rich GAM proteins (Fritz *et al.*, 2012). Therefore highlighting, that the search parameters used for this bioinformatics analysis were capable of identifying oocyst wall specific proteins.

As previously discussed, in analogous systems of extracellular matrix formation, including cuticle formation in *C. elegans* (Donko *et al.*, 2005, Edens *et al.*, 2001) and the fertilisation membrane in sea urchin (Deits *et al.*, 1984), peroxidases and oxidase-type enzymes catalyse tyrosine crosslinks, resulting in di- and trityrosine formation similar to that observed in *Eimeria*. Thus, an *E. tenella*

oxioreductase would need to have the potential to catalyse a redox reaction, either by directly catalysing the transformation of free oxygen into free radicals, including H_2O_2; or alternatively, by acting as a free radical scavenger system, wherethe by-product can be H_2O_2 or, potentially, both.

Several putative genes with potential oxioreductase enzymology were identified in the data mining analysis of the *T. gondii* database. These genes were identified by the screening of each protein sequence identified in bioinformatics analysis, corresponding to 82 proteins (Table 5.2), for motifs specific for: (1) oxioreductase activity; (2) oxidation activity, indicating potential involvement in the production of H_2O_2; or (3) peroxidase activity with potential to catalyze dityrosine crosslinking in the oocyst wall. These motifs included GMC oxioreductase domain, FAD/ADP binding domain, peroxidase domain and short-chain dehydrogenase domain. This proved essential, as often the presence of an oxioreductase motif in a sequence was not recognised by the genome database the sequence was stored on. A putative orthologue in *E. tenella* of an oxioreductase identified in the search of T*. gondii* oocyst proteome was defined as an unspecific product. This was despite the gene, following analysis, being defined as glucose/ribitol dehydrogenase due to signature motifs present in the protein sequence. This scenario highlights the application of performing motif search using Prosite and InterProScan on all the initial proteins identified in *T. gondii* (Table 5.2) including hypothetical proteins that were identified in the final list of 82.

When the 82 proteins identified in the oocyst (and sporocyst) proteome of *T. gondii* were scanned for signature protein motifs, several oxioreductase motif and/or domain containing proteins were identified (Table 5.2). When analysed, many of the identified proteins were found to have an assigned metabolic or antioxidant function in *T. gondii* or *Plasmodium*, and, therefore, were not considered further as candidates for involvement in oocyst wall formation in *E. tenella*. Further to this, genes with oxioreductase capabilities, including acyl-CoA oxidase, acyl-CoA dehydrogenase and alanine dehydrogenase, also with known metabolic functions, were found to be expressed in sporozoites and sporocyst extracts only. Therefore, it is unlikely these proteins are associated specifically with oocyst wall formation and no searches were conducted for *E. tenella* homologues. Oxioreductase proteins identified with, as yet, unknown functions were chosen as more likely to play a role in the oxidation of the *E. tenella* oocyst wall.

A previously unidentified oxioreductase in *E. tenella, Et*OXIO1 (ETH_00033360), was identified for further investigation by the mining of the *T. gondii* database for its homologue, also a putative oxioreductase, *Tg*OXIO1 (TGME49_053150). This gene was initially identified in mass spectrometry analysis of sporulated oocysts of *T. gondii,* where extracts were enriched for oocyst walls. Highly abundant in this fraction, the protein was thought to be located in the outer layer of

the oocyst wall, as bleach-treatment of oocyst wall fractions eliminated any spectra recorded for this protein (Fritz *et al.*, 2012). It has been suggested that the peroxidase homologue oxioreductase identified in *T. gondii* could provide the catalytic machinery involved in dityrosine cross-linking that is thought to occur in the oocyst wall (Fritz *et al.*, 2012). Motif scanning of this protein identified a N- and C- terminal motif (PFAM) and FAD/ADP binding motif (FAM) in the putative protein sequence, defining it as an oxioreductase that can act as a catalyst of an oxidation-reduction (redox) reaction in which a CH-OH group acts as a hydrogen donor. BLAST analysis identified the putative orthologue of this protein in *E. tenella, Et*OXIO1, which had the conserved signature oxioreductase motifs in the protein sequence. Therefore, the investigation of the homologue in *E. tenella, Et*OXIO1, was considered vital in determining if this gene plays a catalytic role in the formation of dityrosine cross-links in oocyst wall formation. Interestingly, a second *T. gondii* oxioreductase protein, TGME49_013340, also identified the *Et*OXIO1 as a homologue. Further analysis revealed that, indeed, the *Tg*OXIO1 (TGME49_053150) protein has a higher level of similarity with *Et*OXIO1 and was, therefore, more likely its possible orthologue. Whilst *Tg*OXIO1 was found to have a higher level of conservation with *Et*OXIO1, it reflects the need to further investigate the putative protein in *E. tenella* to confirm its gene/protein sequence in *E. tenella*, and where in the life cycle of the parasite it's expression is up-regulated.

BLAST uses the comparison of nucleotide and/or protein sequences from different organisms, and can be a powerful tool to identify similarities between sequences and infer the function of newly identified sequences, predict new members of gene families and explore evolutionary relationships (Madden, 2002). In this study, BLAST was used to identify putative orthologues in *E. tenella* to *T. gondii* oxioreductases identified in oocyst proteome analysis. However, the identification of putative orthologues by similarity was limited in this study as, often there was only one sequence identified in the *E. tenella* database to two *T. gondii* sequences. Therefore, these genes need to be investigated further to confirm expression in *E. tenella,* and identify whether or not they are a true orthologue of *T. gondii,* if possible.

In previous work on the oocyst walls of *E. tenella,* mass spectrometry analysis identified several additional gene candidates for the catalytic formation of dityrosine crosslinks in oocyst wall formation of *E. tenella* (Flowers, 2011). Two amine oxidases were identified in oocyst wall extracts implying that they were closely associated with the wall itself as they remained in extracts after stringent washing with detergent. The *E. tenella* genomic database defined the first amine oxidase *Et*AO1, as an unspecified product, without assigning the protein a putative role based on its predicted protein sequence. BLAST analysis found that the putative *T. gondii* orthologue was an amine oxidase (monoamine oxidase), flavin containing domain protein, with a high level of

conservation observed between the two sequences. The *Et*AO1 sequence was then investigated for signature motifs. Motif scanning of *Et*AO1 identified an amine oxidase motif present in the protein sequence and defined the product as an amine oxidase. The second amine oxidase, *Et*AO2, identified in *E. tenella* was defined as an extra-cellular protein containing a MAM domain and cooper-amine oxidase domain.

Amine oxidases are enzymes that catalyse the oxidation of primary amines to aldehydes with concomitant production of H_2O_2 and ammonia. They have been shown to mediate the degradation of putrescine, a substrate of amine oxide produced by amino acid breakdown in plants, producing H_2O_2, which leads to stomal closure on leaves (An *et al.*, 2008). Amine oxidases are not, however, only scavengers of circulating amines, they have also been shown to play a role in signaling and stimulation of specific physiological processes, specifically the production of H_2O_2 (Marti *et al.*, 2001). Such processes include a significant function in adipose tissue including a specific effect on glucose transport (Marti *et al.*, 1998), an increase in glucose disposal (Marti *et al.*, 2001) and the inhibition of lipolysis (Visentin *et al.*, 2003). Interestingly, a copper amine oxidase in chickpea seedlings has also been shown to produce H_2O_2 in the cell wall, and appears to be integrated in the remodeling of the plant cell wall occurring during ontogenesis and wound healing. In this process, H_2O_2 is essential for peroxidase activity during catalysis of crosslinking of phenolic groups between the lignin-suberin barrier during wound healing (Rea *et al.*, 1998) and the generation of iso-dityrosine between glycoproteins in the plant cell wall (Brady and Fry, 1997, Federico and Angelini, 1986). Interestingly the cell wall of higher plants has an extracellular matrix similar to the oocyst wall of *Eimeria.* The glycoproteins, *Em*GAM56 and *Em*GAM82, and the proline and histidine-rich *Et*GAM22 are similar to proline or hydroxyproline-rich glycoproteins found in plant cell walls (Belli *et al.*, 2002b, Kieliszewski *et al.*, 1992, Kieliszewski and Shpak, 2001, Krucken *et al.*, 2008, Walker *et al.*, 2010). Given the similarities, it is appealing to speculate that, similar to the process observed in chickpea seedlings, an amine oxidase is involved in oocyst wall formation. Thus, both amine oxidase genes identified were considered to be potential participants in this process, perhaps via the production of the H_2O_2 required for dityrosine cross-linking formation.

The identification of a quinone oxioreductase was interesting, as this enzyme has been shown to up-regulate tyrosinase activity and positively regulate melanin synthesis (Yamaguchi *et al.*, 2010), whereby tyrosinase acts as an oxidase, converting tyrosine to melanin or other pigments via oxidation. In apicomplexan parasites, NADH-quinone oxioreductases have been identified as playing a role in oxidative phosphorylation (Uyemura *et al.*, 2004). Interestingly, a NAD(P) binding domain was identified in the putative sequence of the *Tg*QO1 homologue, but not the *Et*QO1

sequence. However, analysis of the two sequences by ClustalW shows that the protein sequence of *Et*QO1 is actually terminated at a threonine residue, while the *Tg*QO1 sequence, and the remaining amino acids that make up the NAD(P) binding domain continue for a further 56 residues. Therefore, it is highly likely that the annotation of the *Et*QO1 gene/protein may be incorrect in the *E. tenella* database, as has been observed previously with other proteins in this study, and that the remaining sequence of *Et*QO1, not correctly annotated, contains the remaining residues that make up the NAD(P) binding site. Therefore, it is likely that the *Et*QO1 contains a NAD(P) binding motif, and may play an important role in dityrosine formation in *E. tenella.* Whilst a tyrosinase has not been identified in *E. tenella* to date, quinone oxioreductases have been found to act independently of tyrosinase (Gong *et al.*, 2008, Uyemura *et al.*, 2004). Given that the quinone oxioreductase may be associated with the oocyst wall of *E. tenella* (Flowers, 2011), this gene needs to be analysed further and its expression investigated.

Mass spectrometry studies on oocyst wall extracts of *E. tenella* also identified an oxioreductase, glutathione/thioredoxin peroxidase (*Et*GPX1) (Flowers, 2011). Two homologues of this gene exist in *T. gondii.* The first is a thioredoxin-like dependent peroxidase (TgTPX1/2) that localises to the apicoplast and mitochondrion of *T. gondii,* where it is thought to act as an antioxidant on the H_2O_2 produced by detoxification of oxygen radicals by superoxide dismutase (Pino *et al.*, 2007). The second *T. gondii* homologue of the *E. tenella* glutathione peroxidase is another putative thioredoxin-like peroxidase, found only in sporulated oocysts (Pino *et al.*, 2007). The identification of a peroxidase domain in the putative protein sequence of *Et*GPX1, as well as homology to these two genes, suggests that this gene, in *E. tenella,* has the potential to use H_2O_2 as a substrate in oxidation reactions. Evidence that the second homologue also only exists in sporulated oocysts of *T. gondii*, suggests that this gene may be similar to the one identified in *E. tenella.* Further, it is exciting to speculate that it too may play a role in the oocyst wall of *T. gondii,* in a similar manner to that hypothesised here. *In vitro* assays showed that exogenous peroxidases, including a horseradish peroxidase and *Arthromyces* peroxidase, can catalyse the formation of dityrosine cross-links between a recombinant, tyrosine-rich 42kDa oocyst wall protein from *Eimeria* to form different size polymers (Mai *et al.*, 2011), suggesting that an as yet unidentified endogenous peroxidase is responsible for the *in vivo* formation of these cross-links in the oocyst wall. Therefore, the identification by mass spectrometry analysis of a glutathione peroxidase in oocyst wall extracts, suggests that this protein could play a role in oocyst wall formation via the oxidation of H_2O_2 to form dityrosine cross-links between oocyst wall proteins. Thus, this gene was also selected for further analysis, to determine its expression profile and localisation in *E. tenella*, bringing the total number of candidate enzymes for a role in dityrosine bond and, therefore,

oocyst wall formation to five: two amine oxidases; a glutathione/thioredoxin peroxidase; a quinone oxioreductase; and an oxioreductase.

CHAPTER 6

DEVELOPMENTAL EXPRESSION PROFILES OF OXIOREDUCTASES IN *EIMERIA TENELLA*

6.1 Introduction

We know that the strength and resilience of the oocyst wall can be attributed to the presence of dityrosine crosslinks, formed by the oxidation of small tyrosine-rich wall proteins through a deliberate, enzyme mediated reaction (Belli *et al.,* 2006). The endogenous oxidative enzyme responsible for the generation of these bonds in the *Eimeria* oocyst wall is, to date, unknown. However, several oxioreductases that have the potential to participate in the catalysis of dityrosine bond formation in the oocyst wall of *E. tenella* have been identified using bioinformatic and proteomic approaches (Chapter 5). This identification included the possession of characteristic motifs defining them as oxioreductases with the ability to catalyse redox reactions including generation of the H_2O_2 required for oxidation and/or peroxidation, using H_2O_2 as a substrate in the reaction leading to dityrosine bond formation. Thus, genes for two amine oxidases, *etao1* and *etao2*, a glutathione peroxidase, *etgpx1*, a quinone oxioreductase, *etqo1*, and a putative oxioreductase, *etoxio1*, were identified. This chapter will address the hypothesis that these candidates are gametocyte-specific proteins that, therefore, could play an enzymatic role in oocyst wall formation, specifically, the oxidation of tyrosine-rich oocyst wall proteins. The developmental gene expression profile of each putative oxioreductase will be assessed using quantitative real-time (qRT) PCR and localisation within the parasite determined using antibodies raised to recombinant versions of each protein.

6.2 Gene expression of oxidases by qRT-PCR

6.2.1 Oligonucleotide design

Quantitative reverse-transcriptase PCR (qRT-PCR) was used to examine the expression of various oxidase genes in different life-cycle stages of *E. tenella*. Primers were designed for each of the candidate oxioreductases identified in Chapter 5 using the predicted coding sequences available on the *E. tenella* genome database. They were designed to amplify single amplicons with minimal secondary structure or homo- or heterodimerisation. As well as the candidate oxioreductase genes, primers were also designed to amplify transcripts of: *etgam56*, previously detected only in gametocyte stages (Walker *et al.*, 2010); *etmic4*, detected primarily in merozoite stages (Tomley *et al.*, 2001); as well as the reference gene, *et*r18S. Table 6.1 shows the qRT-PCR primers designed from the predicted mRNA sequences of the *E. tenella* genes. Each primer pair was designed using Primer3 software (Rozen and Skaletsky, 2000) to ensure efficiency, sensitivity, self-complementarity, primer length and a melting temperature of around 60°C. BLASTn on NCBI was used to assess the complementation of each primer to other *E. tenella* genes, and to *G. gallus* genes to ensure there was no redundancy.

Each primer pair (Table 6.1) was used to amplify a gene product from *E. tenella* cDNA as described in Section 2.2.4.3. Products were verified by cloning into pGEM T-easy vector and sequencing as described in Section 2.2.5.

Table 6.1: Primers for quantitative real-time PCR of *Eimeria tenella* oxioreductases

Gene*	Primer	Sequence 5′→3′	Tm(°C)	Amplicon (Bp)
etr18s	Sense	CTGATGCATGCAACGAGTTT	60.05	234
(Supercontig 190rc)	Antisense	GACCAGCCCCACAAATAAG	60.0	
etao1	Sense	TCGAGTTCGGAGAATTGTGA	59.37	176
(ETH_00024850)	Antisense	AGAGGCTGAGATGTGCTGGT	60.0	
etao2	Sense	TCACTAAGCAGCATGCAAGG	60.16	207
(ETH_00028385)	Antisense	CTGAATCTGCACAAGACCGA	59.98	
etgpx1	Sense	CCCGAAGGAAAAGAAGTTCC	60.05	140
(ETH_00011565)	Antisense	AAGCCCTGGTCTCCGTATTT	59.96	
etqo1	Sense	CAGAAAGCAGGCAAATACCC	59.71	229
(ETH_00043030)	Antisense	TTATCCAGTTCTCCGGGATG	59.89	
etoxio1	Sense	GATGTTTGCTGCACTGCCTA	60.02	168
(ETH_33360)	Antisense	ATACGGTGTTGAGGCCAAAG	59.99	
etgam56	Sense	GACACCATGCCCGAGTTC	60.05	120
(ETH_00007320)	Antisense	GCGGTAGTTGGTGGTGTTTT	59.90	
etmic4	Sense	AGGAGGCTGAGCAAGTTGAG	59.75	109
(ETH_00024085)	Antisense	GCCCAGTAGTCGTCTTCGTC	59.87	

**Eimeria tenella* database (http://www.genedb.org/Homepage/Etenella) accession numbers for each sequence are indicated in brackets; Tm: melting temperature; Bp: base pairs.

6.2.2 Validation of the qRT-PCR assay- Primer efficiencies

To determine the relative quantification of oxioreductase genes of interest present in lifecycle stages, the amplification efficiency of primer pairs designed for each individual gene was determined. The amplification efficiency of primers designed for the qRT-PCR assay of oxioreductase genes was performed as described in Section 2.2.4.3.2. Cycle threshold (C_T) values were plotted against the log dilution of cDNA and a standard curve produced (Figure 6.1) to generate a linear standard curve with $R^2 > 0.98$ and a line of regression calculated for each primer pair. The slope of the line was then used to calculate the amplification efficiency of each primer pair (E) and presented as a percentage (Table 6.2). Primer pairs with an amplification efficiency of >85% were deemed satisfactory to determine relative transcript abundance (Pfaffl, 2001, Tichopad *et al.*, 2004). Melt curve analysis was determined by Rotor-Gene ScreenClust HRM Software (Qiagen), with one dominant peak observed for each primer pair confirming a single amplification product in the reaction.

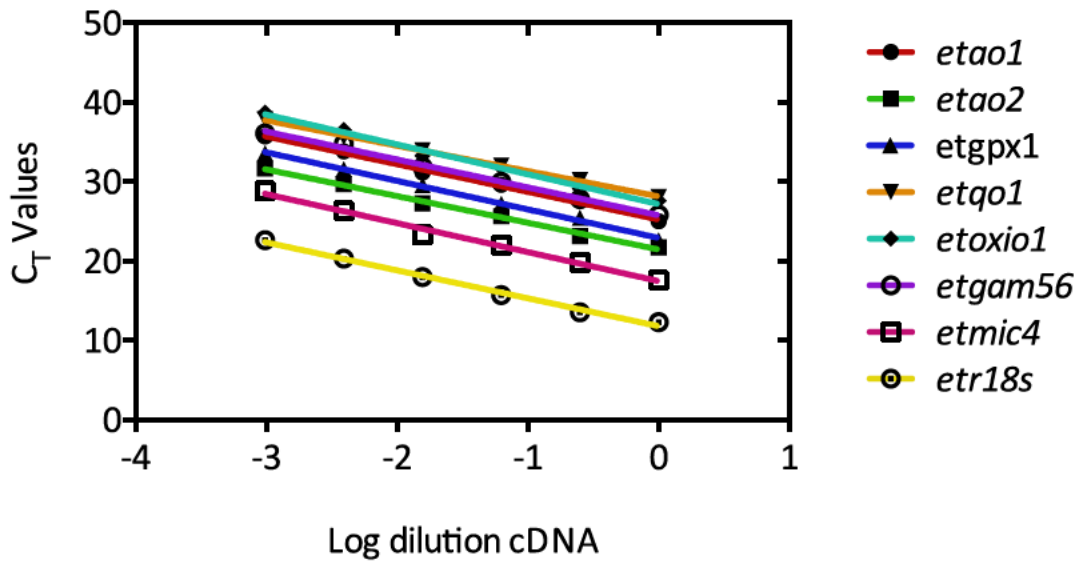


Figure 6.1: Linear regression of C_T versus template dilution cDNA of *Eimeria tenella* in qRT-PCR reactions

The C_T values of the qRT-PCR reactions were plotted against the log dilution factor of cDNA for *etao1, etao2, etgpx1, etqo1, etoxio1, etgam56, etmic4 and etr18S*, the slope of each line being used to determine amplification efficiency of each primer pair.

Table 6.2: Primer efficiency of *Eimeria tenella* oxioreductase-specific qRT-PCR primers

Gene	R^2	Slope of CT vs. log dilution cDNA	Reaction efficiency (E)	Amplification efficiency (%)
etr18s	0.9924	-3.52	1.92	92.4
etao1	0.996	-3.49	1.94	93.6
etao2	0.9962	-3.36	1.98	98.4
etgpx1	0.9964	-3.59	1.9	89.8
etqo1	0.9974	-3.19	2.05	105
etoxio1	0.9901	-3.76	1.86	86.3
etgam56	0.9945	-3.52	1.93	92.5
etmic4	0.9904	-3.65	1.88	87.8

6.2.3 Relative transcript abundance of oxioreductase genes in *E. tenella*

Relative quantitation of gene transcript abundance has been shown to be adequate for investigating gene expression levels (Pfaffl, 2001). To determine the relative transcript abundance of oxioreductase genes in the different life cycle stages of *E. tenella,* the level of transcript expression of the target gene was normalised to the transcript expression of the reference *E. tenella* ribosomal 18S gene. This ratio was then compared between the different life cycle stages to determine the stage-specific expression of each gene. Note, the term 'transcript abundance' is used here rather than 'transcription' due to known mechanisms of mRNA-storage and repressed translation in *Plasmodium* gametocytes.

RNA was extracted and relative gene quantity of oxioreductase genes in the different life cycle stages of *E. tenella* was determined by qRT-PCR as described in Section 2.2.4.3.1. Genes with known gene expression profiles were also investigated with qRT-PCR to validate relative quantities observed. Semi-quantitative analysis of *etgam56* has shown previously that expression of this gene is gametocyte specific (Walker *et al.*, 2010) and it was used here as a control for gametocyte expression. Figure 6.3 (a) shows gametocyte-specific relative transcript abundance of *etgam56*, with a significantly greater ($p<0.05$) level in early stage gametocytes ($p<0.05$) compared with merozoites, unsporulated and sporulated oocysts. Similarly, there is a significant up-regulation of *etgam56* in late stage gametocytes ($p<0.001$) compared with early gametocytes, and merozoites, unsporulated and sporulated oocysts where no transcript was detected (Figure 6.3a).

Gene *etmic4* codes for a microneme protein that has been detected in sporozoite and merozoite stages of *E. tenella* (Tomley *et al.*, 2001) and, as such, was used as a merozoite positive control. Transcript levels of *etmic4* were up-regulated significantly in merozoites in comparison with early stage gametocytes, unsporulated and sporulated oocysts ($p<0.001$), as well as late stage gametocytes ($p<0.005$). In addition, the level detected in late stage gametocytes (144hrs post infection) was significantly higher ($p<0.05$) when compared with unsporulated oocysts where no *etmic4* transcript was detected (Figure 6.3b), reflective of the inevitable contamination that can be observed of gametocyte samples with merozoite parasites

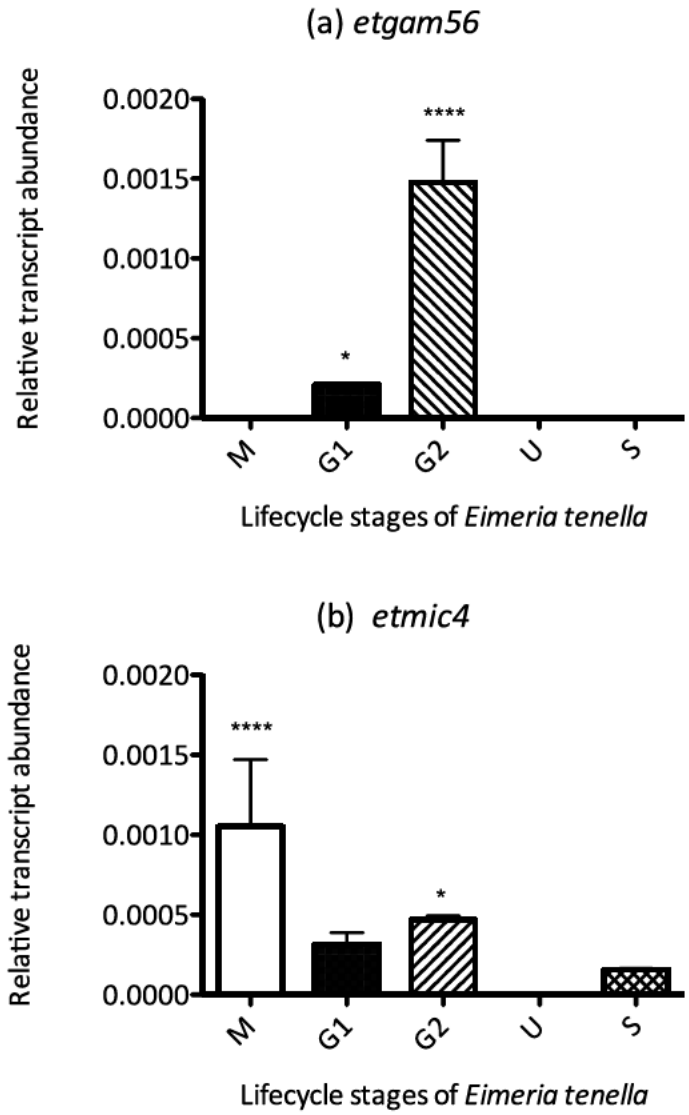


Figure 6.2 (a-b): Relative transcript abundance of *etgam56* and *etmic4* in different life cycle stages of *Eimeria tenella*

The relative transcript abundance of (a) *etgam56* and (b) *etmic4* were examined in different life cycle stages of *E. tenella.* Total RNA was reverse transcribed with random hexamer primers and the gene was amplified from cDNA using gene specific primers. Relative transcript abundance of each gene was determined for each life cycle stage by calculating the difference between the C_T value of the target gene and the C_T of the reference gene *etr18s*. This ratio was then compared between the different life cycle stages. Bars show the mean and the error bars show the standard deviation of three replicates per life cycle stage from at least two different experiments. Significant difference is represented by '****' (P<0.001); '***' (P<0.005); '**' (P<0.05) and '*' (P<0.01) between the different life cycle stages.

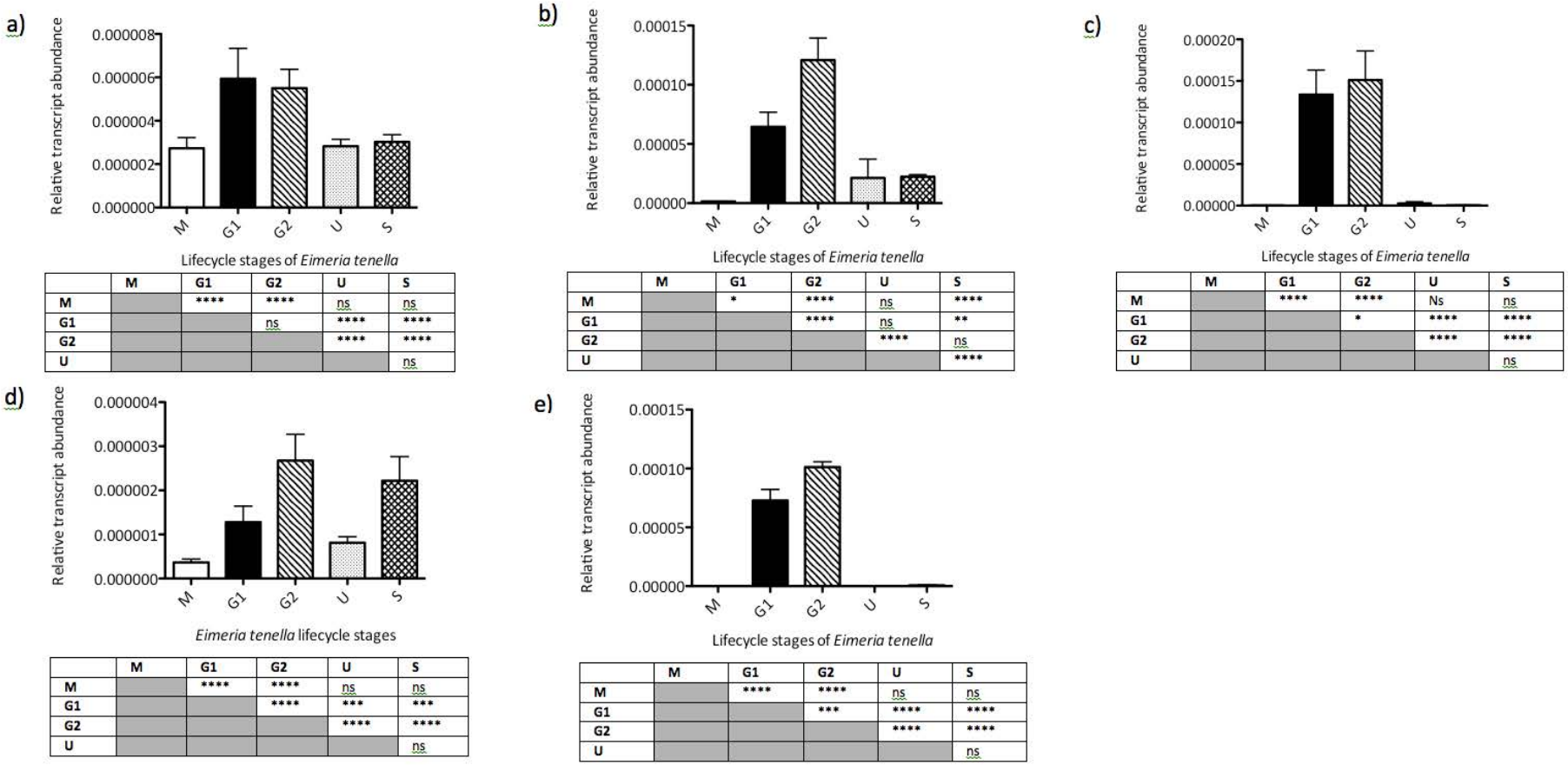
Relative transcript of oxioreductase genes in the different life cycle stages of *E. tenella* was determined by qRT-PCR (Section 2.2.4.3.1). Figure 6.4 (a) shows that the transcript for *etao1* was detected in all the life cycle stages of *E. tenella* but at low relative levels compared to *etgpx1, etao2* and *etoxio1.* The level of *etao1* transcript was slightly, albeit significantly, increased in both early and late stage gametocytes of *E. tenella* compared with levels seen in the other lifecycle stages ($p<0.001$), which were statistically different from each other. The expression of *etao1* was clearly not gametocyte-specific and could not be confidently classified as up-regulated in gametocytes, so was not studied further in this thesis.

Gene transcript of *etgpx1* was detected in gametocytes, unsporulated and sporulated oocysts but there did not appear to be any transcription of *etgpx1* in merozoites of *E. tenella* (Figure 6.4b). Transcript abundance was significantly increased in late stage gametocytes in comparison with unsporulated and sporulated oocysts ($P<0.001$) and early stage gametocytes ($p<0.001$). Moreover, whilst significantly lower than in the later stage gametocytes, glutathione peroxidase transcript abundance in early stage gametocytes was significantly increased compared with unsporulated and sporulated oocysts ($p<0.005$) (Figure 6.4b).

The relative transcript abundance of *etao2* showed gametocyte-specific stage expression in *E. tenella* (Figure 6.4c). Relative transcript abundance in both early- (G1, 134hrs post infection) and late-stage (G2, 144hrs post infection) gametocytes was increased significantly ($p<0.005$) compared with the other life cycle stages, where little or no transcription was apparent.

The relative transcript abundance of *etqo1* in the different life cycle stages of *E. tenella* is shown in Figure 6.4 (d). Like *etao1,* the relative transcript abundance of *etqo1* was low compared to other oxioreductases examined. However, there was a significant up-regulation ($p<0.005$) of relative transcript abundance observed in late stage gametocytes and sporulated oocysts compared with transcript levels in merozoites, early stage gametocytes and unsporulated oocysts (Figure 6.4d). Whilst not as abundant in sporulated oocysts or late stage gametocytes, the relative transcript abundance observed in early stage gametocytes was increased significantly in comparison with transcript detected in merozoites ($p<0.01$).

Figure 6.4 (e) shows that the *etoxio1* transcript was gametocyte-specific, with a significant increase observed in both early and late gametocytes ($P<0.001$) when compared with the other life cycle stages in *E. tenella,* where no transcripts were observed.



	M	G1	G2	U	S
M		****	****	ns	ns
G1			ns	****	****
G2				****	****
U					ns

	M	G1	G2	U	S
M		*	****	ns	****
G1			****	ns	**
G2				****	ns
U					****

	M	G1	G2	U	S
M		****	****	Ns	ns
G1			*	****	****
G2				****	****
U					ns

	M	G1	G2	U	S
M		****	****	ns	ns
G1			****	***	***
G2				****	****
U					ns

	M	G1	G2	U	S
M		****	****	ns	ns
G1			***	****	****
G2				****	****
U					ns

Figure 6.3: Relative transcript abundance of oxioreductase gene, *etao1, etgpx1, etao2, etqo1 and etoxio1* in different life cycle stages of *Eimeria tenella.*

The relative transcript abundance of oxioreductase genes, *etao1* (a)*, etgpx1* (b)*, etao2* (c)*, etqo1* (d) and *etoxio1* (e)*,* was examined in different life cycle stages of *E. tenella.* Total RNA was reverse transcribed with random hexamer primers and the gene was amplified from cDNA using gene specific primers. Relative transcript abundance of each gene was calculated for each life cycle stage as described in Section 2.2.4.3; the level of transcript expression of the target gene, measured as C_T, was normalized to the transcript expression of the reference *E. tenella* ribosomal 18s RNA gene C_T. Bars show the mean and the error bars show the standard deviation of three replicates per life cycle stage from at least two different experiments. A statistical comparison was made between each of the five life cycle stages, represented by '****' (P<0.001); '***' (P<0.005); '**' (P<0.05);'*' (P<0.01) and 'ns' (not significant). Merozoites (M), 134hr p.i. gametocytes (G1), 144hr p.i. gametocytes (G2), unsporulated oocysts (U) and sporulated oocysts (S) are as indicated.

6.2.4 Sequence analysis of oxioreductases of *Eimeria tenella* identified by qRT-PCR

Oxioreductase genes *etao2, etgpx1, etoxio1* and *etqo1* showed significant up-regulation in gametocyte stages of *E. tenella* and, therefore, were investigated further. The qRT-PCR products assayed (Section 6.2.3) were analysed by agarose gel electrophoresis (Section 2.2.4.5) and single product confirmed for each reaction and gel-purified (Section 2.2.5.2). The product was then cloned into pGEM T-easy cloning vector. Plasmid sequencing was performed with M13 vector-specific primers. Each gene was sequenced in both directions to ensure fidelity, and at least two separate cloning reactions were sequenced. Sequences of qRT-PCR products of all five oxioreductases were confirmed and shown to be as predicted by the *E. tenella* genome database using ClustalW2 software (Section 2.2.6.3). Control genes, *etgam56* and *etmic4,* were also sequenced and confirmed.

6.3 Immunolocalisation of oxioreductases with gametocyte transcript profile in *Eimeria tenella*

Following sequence confirmation, the presence and location of the oxioreductases in the various life cycle stages were analysed by Western blot and immunofluorescence microscopy. In order to differentiate between the different *E. tenella* oxioreductase proteins, polyclonal sera was raised against recombinant truncated versions of *Et*AO2, *Et*GPX1, *Et*QO1 and *Et*OXIO1.

6.3.1 Production of oxioreductase antisera

6.3.1.1 Recombinant protein sequence design and synthesis

Prior to cloning, coding sequences of each oxioreductase gene were assessed for suitability in cloning, expression and immunogenicity. The protein sequence of each gene was obtained from the *E. tenella* genome database (http://www.genedb.org/Homepage/Etenella). Each of the truncated protein sequences selected for expression was aligned to ensure that the recombinant proteins were not highly homologous to other *E. tenella* proteins. Little or no homology was observed between the five truncated protein sequences, limiting the chances that there would be cross-reactivity between the antibodies produced against the truncated protein and *E. tenella* oxioreductase paralogues.

6.3.1.1.1 Recombinant protein sequence design and synthesis of *Et*AO2

A truncated recombinant protein sequence was designed for *Et*AO2. A transmembrane domain that was predicted in the full-length protein was excluded as the presence of hydrophobic sequences may make the recombinant protein difficult to express (Hu *et al.*, 2007, Lindskog *et al.*, 2005). BLASTx sequence anlaysis and ClustalW showed a low level of homology between the 12.6kDa designed protein of *Et*AO2 and *G. gallus* amine oxidase (XP004948684.1); http://www.ncbi.nlm.nih.gov/genome/111/?project_id=10809 (Figure 6.4).

6.3.1.1.2 Recombinant protein sequence design and synthesis of *Et*GPX1

A high level of homology existed between *Et*GPX1 and *G. gallus* glutathione peroxidase (accession number NP_001156717.1; http://www.ncbi.nlm.nih.gov/genome/111/?project id=10809), therefore, a region of protein with limited homology was selected for recombinant protein expression and an 8.7kDa protein designed (Figure 6.5).

6.3.1.1.3 Recombinant protein sequence design and synthesis of *Et*QO1

Similarly, a high level of homology existed between *Et*QO1 and the *G. gallus* quinone oxioreductase (accession number NP_001073231.1; http://www.ncbi.nlm.nih.gov/genome/111/?project_id=10809 (Figure 6.6) so a region of minimal homology between the parasite and host proteins was also selected and an 8.5kDa protein designed. The sequence selected for expression is highlighted in Figure 6.6.

6.3.1.1.4 Recombinant protein sequence design and synthesis of *Et*OXIO1

After initial problems were encountered with expression of recombinant *Et*Oxio1 protein in the expression vectors E3, E6 or C5 (GenScript USA Inc.), a truncated protein sequence, 45kDa in size, was selected for expression. An alignment with the *T. gondii* homologue identified regions of the protein sequence that were conserved between the two genes and that could prove to be important in the protein function (Figure 6.7). Further analysis using Epic Antigenic, as described by Kolaskar and Tongaonkar (1990) showed that this region of protein sequence was also predicted to be antigenic (Figure 6.8). These predictions were compared with a Kyte-Doolittle plot to show the hydrophobic/hydrophilic regions of the sequence (Figure 6.8) as described by Kyte and Doolittle (1982), where hydrophilic regions of the protein sequence are also potentially antigenic. The truncated protein sequence was, therefore, also selected to include regions that were more highly antigenic in the corresponding protein sequence of *Et*OXIO1, highlighted blue (Figure 6.7 and 6.8). No homology was observed between the truncated protein sequence and the other oxioreductases to be expressed nor was any homology to the host protein observed.

```
Gallus          ------------------------------------------------------------
Eimeria         MTKFPRAPLKRPERRQRLFTLALLAGGILGSTVGSCAASFEGTSPVDPTRLLPRRLAPQE 60

Gallus          ------------------------------------------------------------
Eimeria         CKTYVYTEIDHEPTEMSDRGSCPTGGGKRAFLLDTSPEGAGAPALLEVSPVQSQQYSCRA 120

Gallus          ------------------------------------------------------------
Eimeria         SRPCSLVVAGFGLQAEDSIAVLKGIEDCPPTSLRNAQIITSTRHITSSSAATFNFHIRPQ 180

Gallus          ------------------------------------------------------------
 Eimeria        KAKTELPVFDLGPIQGSTFTVCYSPYLATRGEAPSAKDYHYRAGTVTIRDEMLNCDFEEG 240

Gallus          ------------------------------------------------------------
Eimeria         MCGLMSFVIDPRMSLEWRRHSGATETAGTGPSADHTIGRRPGKRGHYLYMDSPGYMLGES 300

Gallus          ------------------------------------------------------------
Eimeria         ASLLAPPQIYPKGLYCARMWYHMYGEDVNSLRVYMRTLEGDAEALGDWGKPYLLRVGDHG 360

Gallus          ------------------------------------------------------------
Eimeria         DDWLEAGFEFSSDGHTAQQLVIEALAGFSEKGDIAIDDLRIVSGKCPEQISRDVTQQDYI 420

Gallus          ------------------------------------------------------------
Eimeria         CGEVRLLTGNYAEDKQWFVEGAVSCAGKGYSLSSMQGRWVPCCVPRYGSYTVVLRDSYGD 480

Gallus          ------------------------------------------------------------
Eimeria         GWNRSRLEFRFFGDLITFGEDFNGTEKKYKLNIGLVQIQRVEGSEDRIGIQLQVAEPRTY 540

Gallus          ------------------------------------------------------------
Eimeria         VWCGAAQSGTPPPTVEILKRYGIRSQNPTSHAGEVLRMEIANKPGSRRVVMPSTTYNVYC 600

Gallus          ------------------------------------------------------------
Eimeria         YAEASTANIPEPGGGHGTHRERMDDAQVAASRMAVTTDSTPPNLSIQGTEAHLNDVIVRL 660

Gallus          ------------------------------------------------------------
Eimeria         TSDEPSTVWCLARDEKKEAAAGGEPPSVEALKKGNKIQILHGEEGSPQALSITGLTPDNH 720

Gallus          ------------------------------------------------------------
Eimeria         YTVYCFAQDAALPKGNGISLSQVKELSFSVQTKSKVPELSILSYKAFQKGFRVTVQADTP 780

Gallus          -------------------------------MNMKTVLVLLLLALATIFALVCVLLTRAR 29
Eimeria         AKVWCGAAMAGSAFPSREDVRRVGATTEIEDADQKTELEIRGVPPNTRYTIYCFASSRGG 840
                                                ** * :  :.  * ::: *.  :*.

Gallus          APGACQQQPSAQEEADDGHSRVF------ADLTPEELAQVVQYLRGSLGVPLADAARAKP 83
Eimeria         SAEMTDAQMWAKALEVTSFGRFCDMPVEPADLEVADKPTLFDPLTSTEEFLVREFMGRQR 900
                 .   : *  *:     ...*.       ***   : . :.: * .:  . : :    :

Gallus          SDNCIASVDVQLPA-KAEVLRFLDAGGARPPREALAVLYFGAQAEPNVTEYVVGPLPAPT 142
Eimeria         NLNIEGVYRINLFIDKEKIIEHLDKGGPAPRRYARVRVGTCKDREGAYRQYKVGPLDVRS 960
                . *  .   ::*   * :::..** **. * * * . :    : *    :* **** . :

Gallus          YHRDVTVQKYGGKVP-----------YHRRPALPAEYKQIADFFRG---RVFPSAPSFMD 188
Eimeria         AD-QMTLEPVGAVVPTDCGGYNPQGVFGRRLLAVDEEEELAEILQESFGFGFGKSRCQFE 1019
                 . ::*::  *. **           : **     * :::*::::      * .: . ::

Gallus          QVMEYDGANLAAVTTAPRGSRSGDRLTWFVMFQNVSG--------FFLHPVG-------L 233
Eimeria         HKKSHKEHHEGCLEFGPMLYEEGDDKKIITIWAGLRTPTGDKVPFYFLLPAPEGVDAQVL
                :  .:.  : ..:  .*   ..**  . :.:: .:          :** *.        *

Gallus          EVLVDHSSLDLTQWAVRRVFYNGQYYRDMSELERTYVQGG--------VHVEKVRRAPRD 285
Eimeria         QKPDDREELTFKLDKVKGYWYDGQQFSTLREFVEAYRTGKNFKKVTPEILEERRKMENQE 1139
                :   *:..* :.   *:  :*:** :  : *: .:*  *         :  *: :   ::

Gallus          GDFSSMRPRVPVAATFP--------------------LQYEPQGPRYSIRSN------LV 319
Eimeria         QEHNIQKERALALRRLAPAWLAPNPEGRVGLEYRAAPEHIEPQGRRYTIRQSPGQESYTI 1199
                 :..  : *. .   :.                     : **** **:**..       :

Gallus          IFQAWSFAFGMSVMTGMRLFDIRYNGERVVYEMSVQEALSVYGSNCPGGMSTRYMDGSFG 379
Eimeria         NYAGWEFVINMDRDTALRLWDIKFQGQRVVFEMGMMEALAHYSVAERN---WYFIDSWYG 1256
                 : .*.*.:.*.  *.:**:**:::*:***:**.: ***: *.    .     ::*. :*

Gallus          -IGRYASPLVRGVDCPYTATYVDTHTLHEGLSSSRRTAALCIFEQNLGSPLRRHYSNLQS 438
Eimeria         GLGSASRRLHPGIECAKTG-----QLLFDGGS-------LCIFEKDFARPLRAHWK--SG 1302
                 :*  :  *  *::*. *.     : *.:* *       *****:::. *** *:.  ..

Gallus          LYYGGMVNSALVLRSITTVGNYDYVWDFIFYQNGAIEGKVQATGYTSS----SFFHGNGL 494
Eimeria         TLRDGAPHMALVLRHMITVSNYDYITDYYFHVSGWFEASVSFTGELYAGVEVPWYSARQR 1362
```

```
                 .*  : ***** : **.****: *: *: .* :*..*. **   :    .:: ..

Gallus         QYGNRVWEHT-LGTIHTHFVNYKVDLDVG--EVKNSLVAHDMAFEMVQAPWDPEQQIERP 551
Eimeria        RHGTQVSGSMRMGALHGHLAVWKVDFDLTPDYRSNSVVFSEIVHDTARP--------GAI 1414
               ::*.:*     :*::* *:. :***:*:     .**:*  ::..: .:.

Gallus         RLTEKVLDMEDQAAFRLHSTMPRYLYFATNQKNKWGHQRGYRIQIISFAGDHVPEASSME 611
Eimeria        KLDQWIGEKELDGYIAYNSTRPIHYTIVNEDHNVYGNVGGMTVLPYPTVAIPNPQFELYT 1474
               :* : : : * :. :  :** * :  :..:::* :*:  *  :   . ..   *: .

Gallus         RAISWARYKLAVTRRKEEEPTSTSVYNQNDPWTPTVTFADFI------------------ 653
Eimeria        GPCAWAKYRVATTVRHPDESEATLPRDNKYALKPAVSLDRYLRVRSKWIALEPFHPWTKN 1534
                . :**:*::*.* *: :*. :*   ::: . .*:*::  ::

Gallus         --------------------NNETITNEDLVAWITAGFLHIPHSEDIPNTLTVGNSVGFL 693
Eimeria        AYATLAHEAVTNLFQLALFPNNETIRKADLVTWISSAVWHIPVVEDMPLTLAQGNTLGWL 1594
                                   ***** : ***:**::.. ***  **:* **: **::*:*

Gallus         LRPYNYYNLDPSIYSPDGVFFTSEQNFMACEVNPITCLSQTASCLPSLPPFTYDGFQNIS 753
Eimeria        VKPHNYYMEDMSMDLHNAIGGAVQD-------------PGTCALIRQEMPKYGDAPQN—  1639
               ::*:***  * *:   :.:  : ::             . *.: : .  *   *. **

Gallus         RL 755
Eimeria        --
```

Figure 6.4: ClustalW alignment of *Gallus gallus* amine oxidase and *Eimeria tenella Et*AO2 amine oxidase protein sequence

A ClustalW alignment was performed to analyse the homology between *G. gallus* amine oxidase (NP) and *E. tenella Et*AO2 to ensure a recombinant protein would not cross-react with host antigen. The transmembrane domain excluded from the designed protein is underlined. Conserved amino acids '*', highly similar amino acid ':' and similar '.' amino acids are shown.

```
Gallus          MLDPEGSVEHKP---SSSFPKVFLIP---LAMLLAITALLLLAFSATQQKE--------- 45
Eimeria         MLLTRATAAAAPRVFSRLLPRTCWAASRAPAAAAAATAPAAAAAAAAVTRQQRTMWLFSA 60
                ** ...:.   *   *  :*:.   .    *   * **    * :*:  ::

Gallus          ------------TDFYTFKVVNIRGKLVSLEKYRGSVSLVVNVASECGFTDSHYKALQQL 93
Eimeria         WRSAAVPPELLAADLWSLAAKDPEGKEVPLKDFKGKVLLITNTATKCGFTQQHLKQFHDL 120
                            :*:::: . : .** *.*:.::*.* *:.*.*::****:.* * :::*

Gallus          QKDLGPYHFNVLAFPCNQFGQQEPDTNKEIESFARKTYGASFPMFSKVAVSGAGAIPAFK 153
Eimeria         KEKYGDQGFEILAFPTLQFKQQE-EKDPQKMCEVYKSFNANFPIFAITEVNGPNANPIFQ 179
                ::. *   *::****  ** *** :.: :  . . *::.*.**:*: . *.*..* * *:

Gallus          YLIDSTGEEPT--------WNFWKYLVDPNGKVVKAWDSTVSVEEIRPHVTELVR---KI 202
Eimeria         YCKFNTDELYSKGQLQAIGWNFGKFLLDKEGKVFKYFGPRQSPSEMEEDILLLLRGEAHG 239
                *   .*.*  :        *** *:*:* :***.* :..  * .*:. .:  *:*   :

Gallus          ILKKKDEL-- 210
Eimeria         KQRNKDGLLQ 249
                  ::** *
```

Figure 6.5: ClustalW alignment of *Gallus gallus* glutathione peroxidase 7 and *Eimeria tenella* glutathione peroxidase protein sequence

A ClustalW alignment was performed to analyse the homology between *G. gallus* glutathione peroxidase 7 protein (NP_001156717.1) and *E. tenella* glutathione peroxidase to ensure a recombinant protein would not cross-react with host antigen. Conserved amino acids '*', highly similar amino acids ':' and similar '.' amino acids are shown. The protein sequence selected for protein expression is highlighted in yellow.

```
Gallus           MAAAAARGAMRAVRV--------FEFGGPE------------------------VLRLQA 28
Eimeria          MKPLAAAQRMRAILVGPPSLPKGSSFGGPLEGPGGPPGSPGGASRGPHGGPQGAPYLYLS 60
                 * . **   ***: *         .****                                :

Gallus           DVPIPSPEDAQVLIKVHACGVNPVETYIRSGNYARKPALPYTPGSDVAGVIEGVGGRVTA 88
Eimeria          ECPLPKLRPNEMLVRVGAAGVNRMDLLQKAGKYPPPPGTTDILGPEAAGIVVSS----SV 116
                 : *:*. .  ::*::* *.*** ::   ::*:*.  *. .   *.:.**:: .     :.

Gallus          FKKGDRVFTSATLSGGYADYAVAEADTVFPLSDKLSFSQGAAIAIPYFTAYRALFLKG-- 146
Eimeria          YKEG-TPVACLLPGGGYAEYVAADPALCFPIPSNLTIREAACIPENWITAYQLLEEVARL 175
                 :*:*   .:.   .****:*..*:.   **:..:*:: :.*.*.  ::***: *   .

Gallus          ---HAKAG-----ESVLIHGASGGVGIAACQIARVYG--LKVLGTAGTEQGMNVVLRNGA 196
Eimeria          SDFKANGGKPSGIKHALVYAAGSGVGVALLQLFRVFIPSVSVIAVAGSNEKLQRVKELGA 235
                    :*:.*     : .*::.*..***:*  *: **:   :.*:..**::: :: * . **

Gallus          HQAFNHREAN--YLDKIMEYTKMQGVNIIIEMLSNVNLAADLHLLSHAGRVMVVGCRGPI 254
Eimeria          AVCLNYKELGDSLGDEVLKATGGQGVDIVLDCVGASLVNQTLKSLKTDGT---------- 285
                   .:*::* .    *:::: *  ***:*::: :.   :   *: *.  *

Gallus          EINPRDTMNKESSIRGVSLFLQTEEERRECVRAVLDGIESGWLKPVVGPEYPLQEVAKAH 314
Eimeria         ------------------------------------------------------------
```

Figure 6.6: ClustalW alignment of *Gallus gallus* quinone oxioreductase and *Eimeria tenella* quinone oxioreductase protein sequence

A ClustalW alignment was performed to analyse the homology between *G. gallus* quinone oxioreductase protein (NP_001073231.1) and *E. tenella* quinone oxioreductase to ensure a recombinant protein would not cross-react with host antigen. Conserved amino acids '*', highly

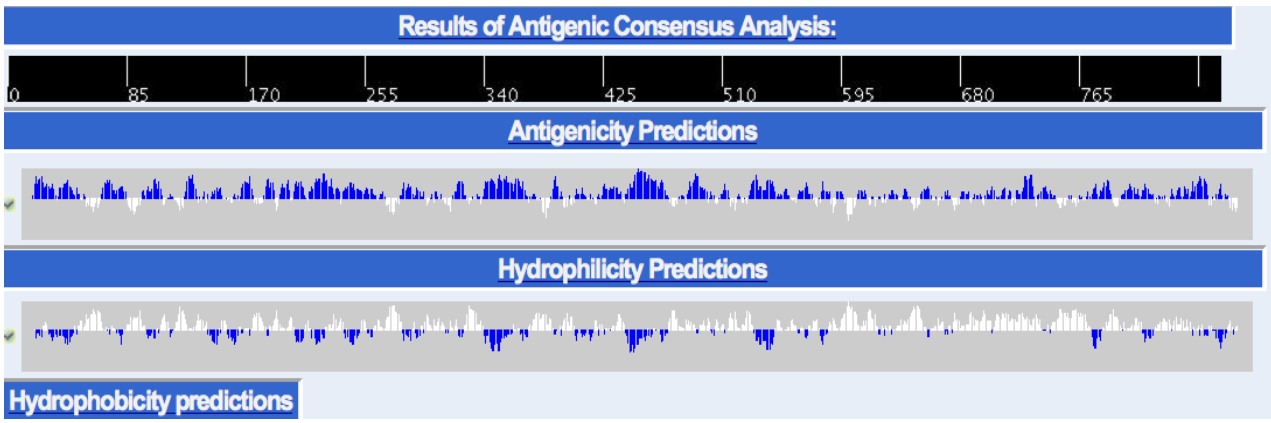
similar amino acids ‘:’ and similar ‘.’ amino acids are shown. The protein sequence selected for protein expression is highlighted in yellow.

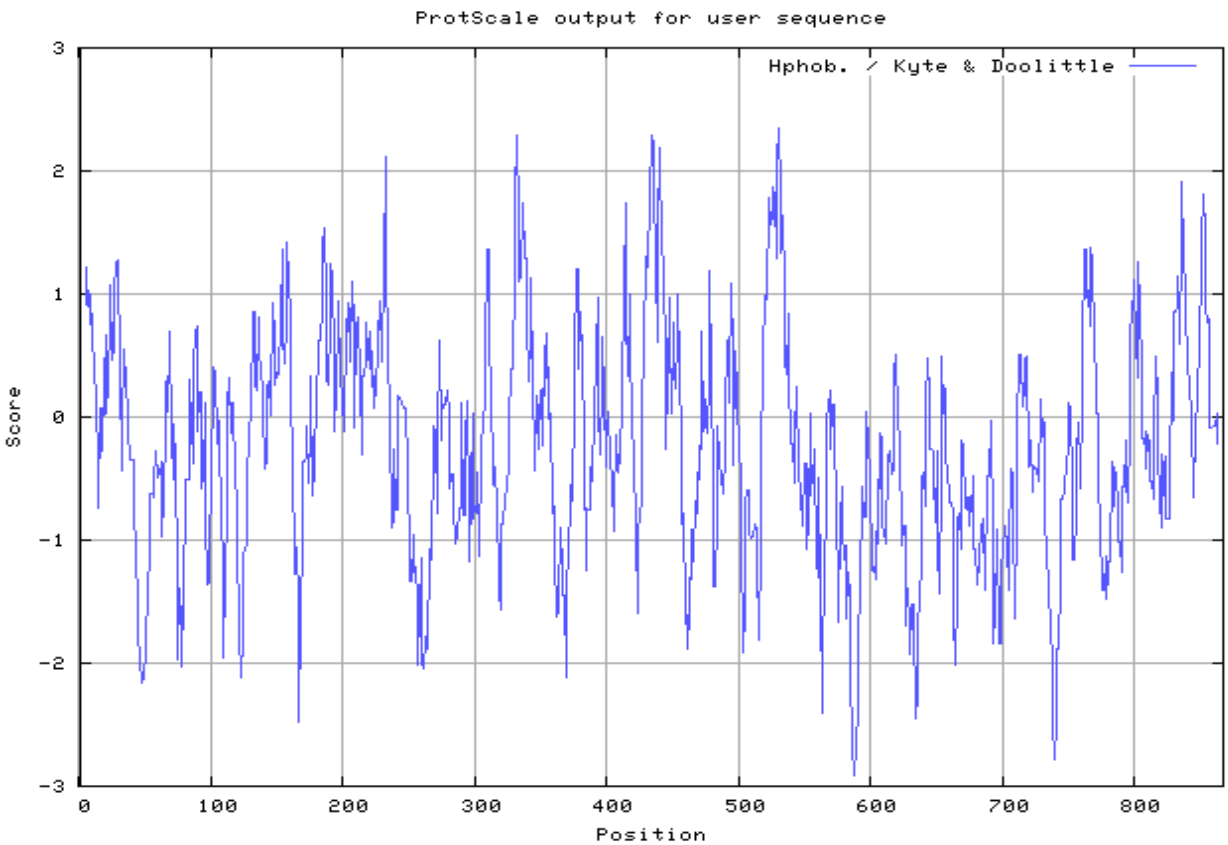
```
TgOXIO1         SGVRTHVATVMGGGTSIDLAIYVEETENYFKFMNKRFPAYKFHWPTIQRAYRYVRRIVAR 300
EtOXIO1         HTAR------------------------------------------------------- 100

TgOXIO1         RMPFDNYFGLRYQQALQWRGFKPIGGKIPADYSSSLQWIEFDLDQNPPRAKCVVYRPTAY 360
EtOXIO1         ------------------------------HTARSITFIDTDHGKT---ANCVVYRPTKY 123
                                                *: :*: * .:.   *:******* *

TgOXIO1         EDIKPQGTSFAASLPASSAHWGPWIGTFLAPKSVLIKSGEVDPTDFFQKKEAKFFRKVCL 420
EtOXIO1         QDIRPIG-DLWSLFPSSAADWAPWIGTFLEPTSVLVNAGVNDPT--------KKLRRVCL 174
                :**:* * .: : :*:*:*.*.******* *.***:::*  ***        * :*:***

TgOXIO1         TDRKGSQIILSAGAIHSAVILYKSGIGPIPQLKKIKTPVIVEHPYLGQRFSDRVFIPVSA 480
EtOXIO1         SNPGSGRIVVSGGAVNTPLLLYRSGIGPALQLKLLKVPQVLDSPAVGSAFSDRVFLSIPG 234
                ::  ..:*::*.**:::.::**:*****  *** :*.* ::: * :*. ******:.:..

TgOXIO1         FSKHYADELTPVPFAPIRRLSSLDPLSPKATTPRGRTRRDRHFSNKDEAEPSSGHDETKR 540
EtOXIO1         FLKHYSDSIP-----LINPSKSLRRLQANEDTGSKKVNRLQLLS-RPAFEPILTQEELEK 288
                * ***:*.:.      *.  .**  *..:  *   :..* : :* :   **   ::* ::

TgOXIO1         GIRKERRHQKGTFTAGDQVNVEGGEGKEEADQGVTRATGGISALGVSENDEDGAEVANAF 600
EtOXIO1         AITPEVR------------------------EAINRLMNQEVPVEIITVDEDDSAVSNRF 324
                .*  * *                        :.:.*  .   .: :   ***.: *:* *

TgOXIO1         NEDDQASGSSPARIFLMDLQKPIPPELSGKYLDVYVPEKGELDPPRVCQGMGIKKGGPPH 660
EtOXIO1         EP----PVILPDFLLFPGVLQYLSPALQ----------------PRVCQFMGLKYTGPRC 364
                :     .   *  ::: .: : :.* *.                ***** **:*  **

TgOXIO1         GCDKENYSIGRRTLSCSLMVAEEMSGGRAQEGIIYASRFIFPPLFRNDPLVDVIFEILRS 720
EtOXIO1         KSN----GINERNLGCSLVTMEELSGDRLAEGFIYASRYMFPTAFRKDPILDAVTEIIQS 420
                 .:    .*..*.*.***:. **:**.*  **:*****::**. **:**::*.: **::*

TgOXIO1         CSEYRAPFAVAGLKPLCAIVVPIIKCFRKSLATFYFTAEPKSRGSVRLSANGMIDVNGNY 780
EtOXIO1         CSNYRSPFGIIFLKPLCVLAQPIVKCLRKAVAPFYFTSEPKSRGSIRLKPTGEVVVDPQY 480
                **:**:**.:  *****.:. **:**:**::*.****:*******:**...* : *: :*

TgOXIO1         LKDEQDLFDAVRGVSNLINAMNGDAYRGVLQPAGSLSCPVTVLNGLLDLILTVASMTSLF 840
EtOXIO1         LVDEQDLFDAIRGTATLVEQLNGDSYRGIIQEKGPQSCPLAILNGLLDILLTLASTTSPF 540
                * ********:**.:.*:: :***:***::*  *. ***:::******::**:** ** *

TgOXIO1         ITKPGNLPLIQKYLQDLLPANSRRLRRLVAVGDDVKPRRLEATDLGDLEEPGDYHQDGMN 900
EtOXIO1         LTHSSNFAEIQRYLQDLIPPESRRLRRLMVVNEGYR------------MKPAVYEDEEGN 588
                :*:..*:. **:*****:*.:*******:.*.:. :             :*. *.::  *

TgOXIO1         EPNTFILEDYSDHVIENVDEKKMEELGIKRKLEEAGFDFEAYRKHLDPNAEEDKDTVISE 960
EtOXIO1         E-----YEDYSAFVENDVDEAKIERLGIKRRLEAVGFDFDSYRAHVERTDESAKDPSLHL 643
                *      **** .* ::*** *:*.*****:** .****::** *:: . *. **. :

TgOXIO1         TANWFVSRRLQQQMEEDQHLTAEEKHTARRLMELQREKEDHAAEAAFEVSVQELADAKAA 1020
EtOXIO1         GT---LGRRLDAEIE--QILGPEKMDQLKKLADSIKSEQALRASKEDPVLKESIEDCRKP 698
                 :   :.***: ::*  * * .*: .  ::* :  :.::   *.    *  :.: *.: .

TgOXIO1         CQKPCSENDACPAADVCCAAFNNSLCIRTQETEELERRLRAQMFEQKLTTGAPSRQAKRT 1080
EtOXIO1         CTKEAIQNHTCAKSDVCCTAYGNTKCIRLPGQPAFAEQQ---------------NKEPYT 743
                * * . :*.:*. :****:*:.*: ***      : .:                .:   *

TgOXIO1         EEGVANEKTAVYPEGFLGALPNALPELLNFVPSPQKSPPHPLLYEKPPPHDFPTYLNPPK 1140
EtOXIO1         SADAADSGGAVKKEEFL------------------------------------------ 760
                . ..*:.  **  * **

TgOXIO1         PEGFPDPSMFPAAPPAGTPPPDVLPYTNKQWAATFPPRLPNPYKPKEVAKYALTYMSSVW 1200
EtOXIO1         ------------APFFGLNTVSVAHADDEQWAATYPPRLPSVQNPKALAKFALTYMTSIG 808
                            **  *  . .*    ::*****:*****.  :** :**:*****:*:
```

Figure 6.7: ClustalW alignment of *Eimeria tenella* oxioreductase, *Et*Oxio1, and *Toxoplasma gondii* oxioreductase, *Tg*Oxio1

A ClustalW alignment compared the *E. tenella* oxioreductase gene *Et*Oxio1 and the oxioreductase homologue in T. *gondii* (accession number TGME49_053150; http://toxodb.org/toxo/) to identify regions of high conservation. A Kyte-Doolittle plot identified residues within this conserved region in *Et*OXIO1 that were hydrophilic, and potentially antigenic. The Conserved amino acids '*', highly

similar amino acids ':' and similar '.' amino acids are shown. The *Et*OXIO1 protein sequence selected for protein expression is highlighted in blue.





Truncated protein sequence *Et*Oxio1 134-538

Figure 6.8: Antigenicity and hydrophobicity plots of *Et*Oxio1 protein sequence

An antigenic plot was mapped of *E. tenella* oxioreductase *Et*Oxio1 using the online software Epic Antigenic (http://epic.embl.de/) to predict the antigenic regions of the protein sequence. This

corresponding data was compared to a hydrophobic plot that was mapped for the protein sequence, to determine the design of truncated protein sequence of *Et*Oxio1 for protein expression.

6.3.2 Analysis of antibody specificity with ELISA

The specificity and sensitivity of the oxioreductase antibodies was investigated by ELISAs for final bleed antiserum samples (Section 2.2.8.2) against the original purified recombinant protein antigen. Each ELISA microplate was coated with unconjugated synthetic peptide (500ng per well) and probed with antibodies from pre-immune sera (blood taken from mice prior to immunisation), final bleed antisera (blood taken 10 days after the final booster shot on day 66) and no sera negative control. Two-way ANOVA was performed to determine the level of reactivity.

The level of reactivity against recombinant *Et*AO2 (r*Et*AO2) in antisera from mice immunised with r*Et*AO2 was significantly higher ($p<0.001$) than reactivity seen in pre-immunisation samples or no antisera control (Figure 6.9a). Significant levels of cross reactivity were also seen between mice immunised with *Et*GPX1 or *Et*QO1 and r*Et*AO2. No significant cross reactivity was observed between mice immunised with *Et*OXIO1 and r*Et*AO2 (Figure 6.9a).

Similar results were seen with antisera raised against recombinant protein *Et*GPX1 (r*ET*GPX1) (Figure 6.9b). Antisera from all mice immunised with *Et*GPX1 showed a significant level of reactivity to the recombinant protein. A significant level of cross-reactivity was also observed in antisera to *Et*QO1 (mice 1-5) against r*Et*GPX1 in comparison with the no antisera control (Figure 6.9b). No significant cross-reactivity was observed in antisera from mice immunised with *Et*AO2 or *Et*OXIO1 recombinant protein.

A significant level of reactivity against recombinant protein *Et*QO1 (r*Et*QO1) was observed in the antisera from mice immunised with the same recombinant protein compared with no antisera and pre-immune sera controls (Figure 6.9c). A significant level of cross-reactivity was also observed with antisera to *Et*GPX1 and antisera to *Et*OXIO1 against the r*Et*QO1 compared with the no antisera control. No cross-reactivity was observed with antisera to *Et*AO2 from mice 1, 3, 4 and 5, however, significant cross-reactivity ($p<0.05$) was observed in antisera from mouse 2 immunised with *Et*AO2 recombinant protein (Figure 6.9c).

Similarly, a significant level of reactivity was observed for antisera to *Et*OXIO1 ($p<0.001$) against the recombinant protein *Et*OXIO1 (r*Et*OXIO1) compared with no antisera and pre-immune sera controls (Figure 6.9d). Significant cross-reactivity was detected in all mice immunised with rEtQO1 ($p<0.001$) as well as in mouse 2 immunised with *Et*AO2 ($p<0.005$) and mouse 1 immunised with r*Et*GPX ($p<0.005$) compared with no antisera control (Figure 6.9d).

Antiserum from mouse 3 immunised with *Et*AO2, antiserum from mouse 3 immunised with *Et*GPX1, antiserum from mouse 3 immunised with *Et*QO1, and antiserum from mouse 2 immunised with *Et*OXIO1 were selected for use in Western blots and immunolocalisation experiments since these represented the serum for each protein that showed the highest specific reactivity with the immunising antigen but the lowest crossreactivity with other oxioreductases.

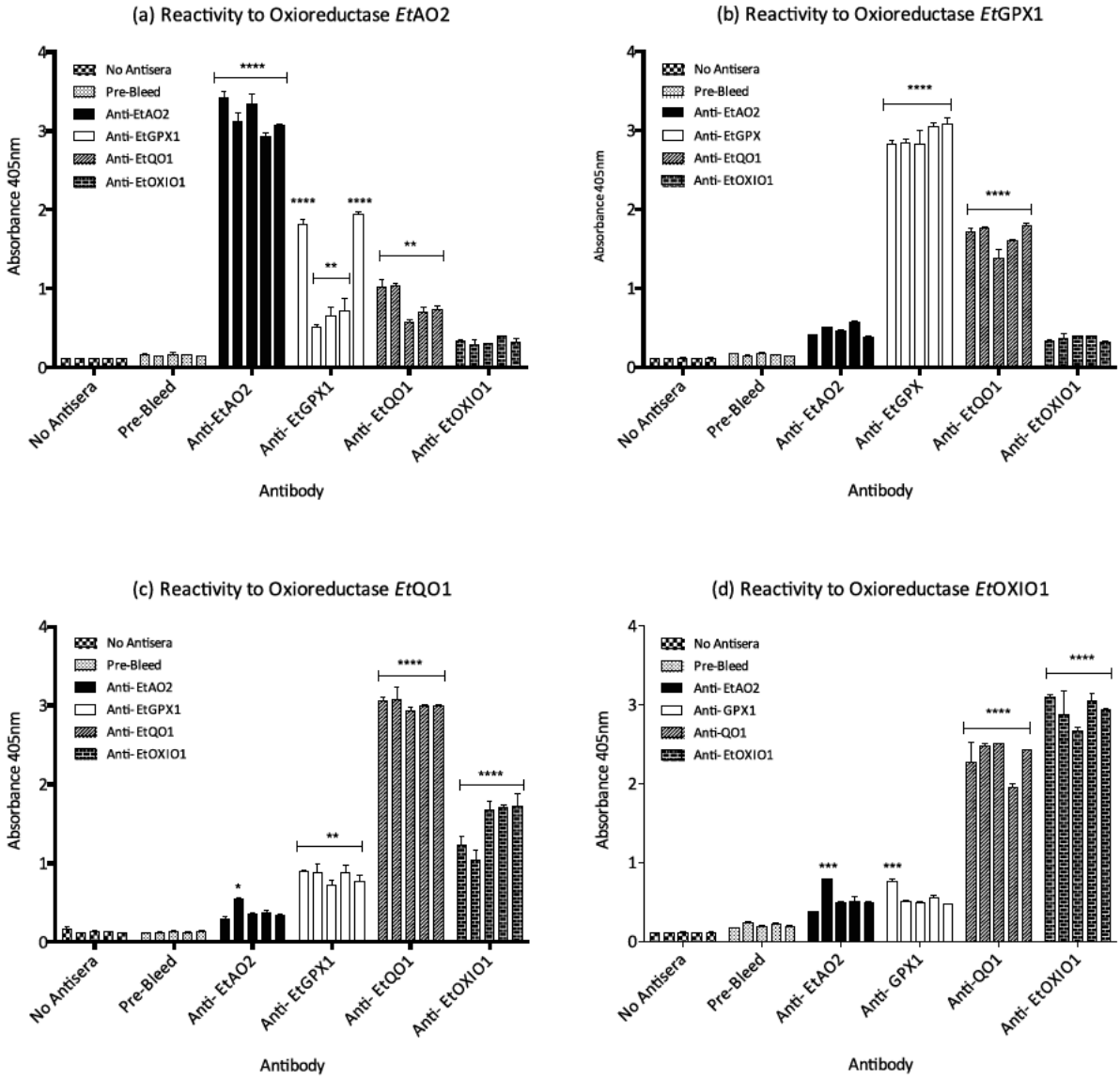


Figure 6.9: Reactivity of antisera produced from mice immunised with recombinant oxioreductase proteins, *Et*AO2, *Et*GPX1, *Et*QO1 or *Et*OXIO1

Five mice (1-5) per protein were immunised with four separate doses of 500µg of purified recombinant protein of *Et*AO2, *Et*GPX1, *Et*QO1 or *Et*OXIO1 by intraperitoneal injection on days 0, 14, 35 and 56. The level of immunological response was determined by ELISA of test bleed antisera from the mice on day 66 after initial immunisation. This was compared with pre-bleeds and no

antisera by measuring absorbance at 405nm. (a) Reactivity of antisera to recombinant *Et*AO2. (b) Reactivity of antisera to recombinant *Et*GPX1. (c) Reactivity of antisera to recombinant *Et*QO1. (d) Reactivity of antisera to recombinant *Et*OXIO1. Two-way ANOVA was performed to determine the p-value of antisera against no antisera control. '****' represents a significant difference ($p<0.001$); '***) a significant difference of $p<0.005$; '**' a significant difference of $p<0.01$; and '*' a significant difference of $p<0.05$ to the no antisera control. Results represent the mean ± SE of four samples per bleed per mouse.

6.3.3 Immunodetection of native oxioreductase proteins in *Eimeria tenella*

Recombinant proteins of *Et*AO2, *Et*GPX1, *Et*QO1 and *Et*OXIO1 were used to immunise mice. Five mice per recombinant protein were immunised with 500µg of protein, with two subsequent booster shots given on days 14, 35 and 56 following initial immunisation (Section 2.2.8.1.4) to generate antibodies. Western blot analysis was performed to determine if the native *Et*AO2, *Et*GPX1, *Et*QO1 and *Et*Oxio1 proteins of *E. tenella* could be identified using the final bleed antisera raised against their respective recombinant protein. Protein lysates were prepared from a number of different life cycle stages and standardised to 1mg/ml concentrations. Protein lysates were prepared from *E. tenella* life cycle stages as described in Section 2.2.8.3. Lysates were separated on a 4-12% Bis-Tris gel under reducing conditions and transferred to nitrocellulose membrane as described in Sections 2.2.8.5 and 2.2.8.6. Membranes were then probed with antisera raised to the various oxioreductase proteins as described in Section 2.2.8.6. Figure 6.11 shows the life cycle stage expression of *Et*AO2 (a), *Et*OXIO1 (b), *Et*QO1 (c) and *Et*GPX1 (d).

Positive control immunoblots were also performed using antiserum to *Em*GAM56 to confirm the protein lysates from *E. tenella* were not degraded (Fig 6.10a). As expected, a dominant 56kDa band was observed in the immunoblot probed with antiserum to *Em*GAM56 in early and late stage gametocytes. The processed GAM56 product, approximately 36kDa in size, was observed in late stage gametocytes and unsporulated oocysts (Figure 6.10a). A pre-immunisation antisera taken from each mouse prior to immunisation was also used as a negative control for each oxidase antibody. Pre-bleed samples from each mouse selected for each oxidase (as detailed above) were used to probe the protein lysates from each lifecycle stage of *E. tenella.* Figure 6.10 (b) shows the immunoblot probed with pre-immunisation serum prior to immunisation with *Et*AO2 recombinant protein with no bands present. This blot is representative of all four blots performed for each oxidase, where no bands were observed for the pre-immunisation sera from mice immunised with *Et*GPX1, *Et*OXIO1 or *Et*QO1 (not shown).

A dominant band, 130kDa in size, was observed in gametocytes at 144hr p.i., and in unsporulated and sporulated oocysts in the immunoblot probed with antiserum to *Et*AO2 (Figure 6.11a). The band in the unsporulated oocysts appears to be the most prominent. This band was not observed in immunoblots probed with antisera from other oxioreductases (Figure 6.9b, c, d) and pre-bleed sera (Figure 6.10b). This band, therefore, was specific to antiserum to *Et*AO2.

A number of bands were observed in both early and late stage gametocytes in the immunoblot probed with antiserum to *Et*OXIO1 (Figure 6.11b). A dominant band was seen at 96kDa in late stage gametocytes with bands of 130 kDa, 120kDa, 110 kDa, 68 kDa 55kDa and 50kDa also (Figure 6.11b). A high molecular weight band was seen in merozoites but no bands were observed in unsporulated or sporulated oocysts (Figure 6.11b).

A surprising number of bands, across all developmental stages of *E. tenella,* were observed on the immunoblot probed with antiserum to *Et*QO1 (Figure 6.11c), confirming the lack of specificity noted in ELISAs for antisera to *Et*QO1. Two dominant bands of 55kDa and 60 kDa were observed in early and late stage gametocytes, with fainter bands of 110kDa and 35kDa also identified in early and late stage gametocytes. A band of 40kDa was observed in both sporulated and unsporulated oocysts with other bands of approximately 100kDa, 90kDa and 50kDa also detected in sporulated oocysts, (Figure 6.11c). A single band of 100kDa in size was detected in merozoites when probed with antiserum to *Et*QO1 (Figure 6.11c).

In the immunoblot probed with antiserum to *Et*GPX1, dominant 55kDa and 60kDa bands were observed in early and late stage gametocyte lysates. A smaller band 30kDa was also observed in late stage gametocytes, unsporulated and sporulated oocysts (Figure 6.11d).

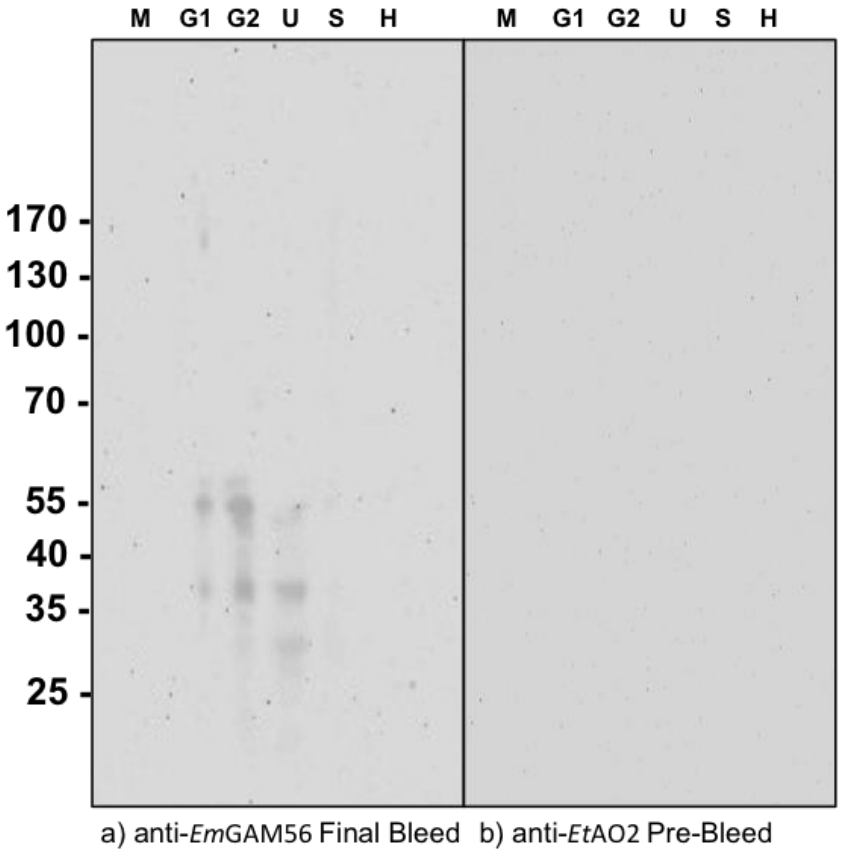


Figure 6.10: Control immunoblot of *Eimeria tenella* protein lysate probed with antisera to *Em*GAM56 and pre-immunisation sera.

Protein lysates were prepared from *E. tenella* lifecycle stages merozoites (M), early and late gametocytes (G1 and G2), unsporulated (U) and sporulated (S) oocysts, and from uninfected caeca host (H). Protein lysates were separated on a 4-12% Bis-Tris gel under reducing conditions and transferred to nitrocellulose membrane. Following the transfer, a membrane was probed with (a) antiserum to *Em*GAM56 or (b) pre-immunisation serum from a mouse immunised subsequently with *Et*AO2.

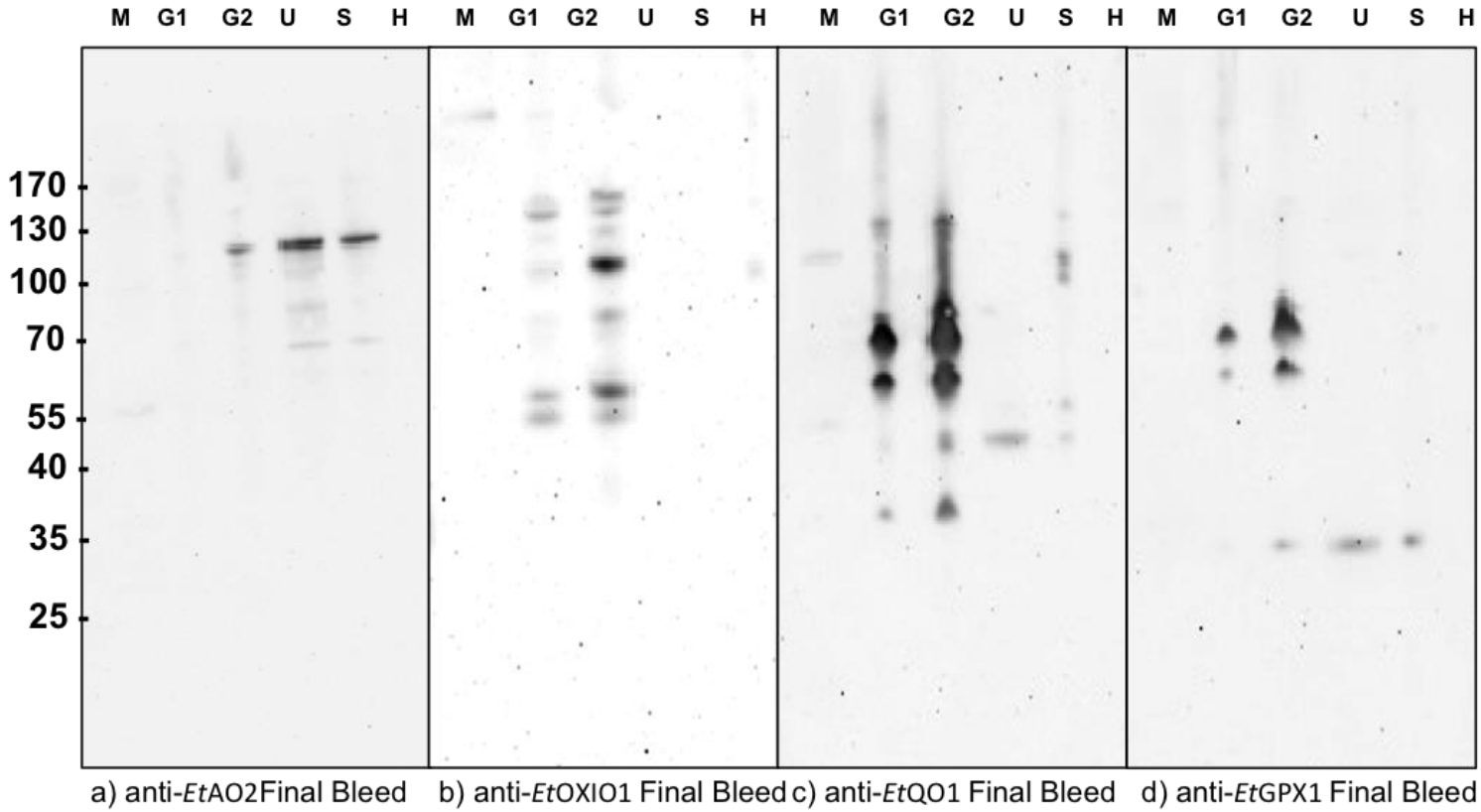


Figure 6.11: Immunoblot of *Eimeria tenella* protein lysate probed with antisera raised to oxioreductase *Et*AO2, *Et*OXIO1, *Et*QO1, and *Et*GPX1

Protein lysates were prepared from *E. tenella* lifecycle stages merozoites (M), early and late gametocytes (G1 and G2), unsporulated (U) and sporulated (S) oocysts, and from uninfected caeca host (H). Protein lysates were separated on a 4-12% Bis-Tris gel under reducing conditions and transferred to nitrocellulose membrane. Following transfer, the membrane was probed with (a) antiserum to *Et*AO2, (b) antiserum to *Et*OXIO1, (c) antiserum to *Et*QO1 or (d) antiserum to *Et*GPX1.

6.3.4 Immunolocalisation of oxioreductases in *Eimeria tenella*

Paraffin sections of *E. tenella* gametocytes were probed with antisera raised against *Et*AO2, *Et*GPX1, and *Et*Oxio1 recombinant proteins to localise the expression of the native proteins. (Since antisera to *Et*QO1 was shown to be relatively non-specific in its reactivity both in Western blots and ELISAs it was concluded that it woud not be useful or informative to use it in immunolocalisation studies). Sections of 4-6μm thickness were taken from *E. tenella* infected chicken caeca at 134hr p.i. and 144 p.i. and stained with antiserum to *Et*AO2, antiserum to *Et*GPX1, or antiserum to *Et*Oxio1, respectively, at 1:100 dilution. Mouse pre-bleeds were used as negative controls to ensure localisation observed was specific to each antibody. No reactivity was observed in tissue sections using pre-immune antisera (Figure 6.12).

Rabbit antiserum against rGAM56, raised previously against a recombinant version of the *E. maxima* 56kDa gametocyte antigen (Belli *et al.*, 2002b), was used as a positive control for the identification of macrogametocyte-specific wall-forming bodies type II (WFBII) (Ferguson *et al.*, 2003a). The antiserum to *Em*GAM56 reacted strongly with *E. tenella* macrogametocytes at both 134hrs p.i. and 144hrs p.i. (Figure 6.14b and f), but not with *E. tenella* merozoites or chicken host cells, which were DAPI-stained positive cells (Figure 6.14c and g).

Antiserum to *Et*AO2 was found to localise to macrogametocytes at 144hrs p.i. in *E. tenella* infected caeca (Figure 6.13e). Interestingly, antiserum to *Et*AO2 antibody did not appear to localise to macrogametocytes very strongly at the earlier time point of 134hrs p.i. (Figure 6.13a) where only binding of antiserum to *Em*GAM56 was observed (Figure 6.13b). An overlay of the images (Figure 6.13h) shows the antiserum to *Et*AO2 at 144hrs p.i. localising to a different region of the macrogametocyte than WFBII but with a distinctly punctate pattern within the cytoplasm of parasites Antiserum to *Et*AO2 did not recognise microgametocytes (Data not shown).

Antiserum raised to recombinant *Et*GPX1 also localised, in a punctate pattern, to the cytoplasm of macrogametocytes at both 134hrs p.i. and 144hrs p.i. (Figure 6.14a and e). As expected, the antiserum to EmrGAM56 localised to WFBII of macrogametocytes of *E. tenella* (Figure 6.14b and f); however, an overlay of the images shows the antiserum to *Et*GPX1 localising to a different region of the macrogametocyte than WFBII at both time points (Figure 6.14d and h). Antiserum to *Et*GPX1 did not appear to localise to microgametocytes (data not shown).

This pattern of staining was also observed for antiserum raised to recombinant *Et*OXIO1, which appeared to localise to macrogametocytes, and not microgametocytes (data not shown), in *E. tenella*-infected intestinal sections in both early and late stage infections (Figure 6.15a and e). Overlays of these images with the images of localisation of *Em*rGAM56 show antiserum to

*Et*OXIO1 localising to what appears to be the cytoplasm of the parasite and not to the WFBII (Figure 6.15d and h) at both 134hrs p.i. and 144hrs p.i.; again, a punctate distribution of reactivity was seen.

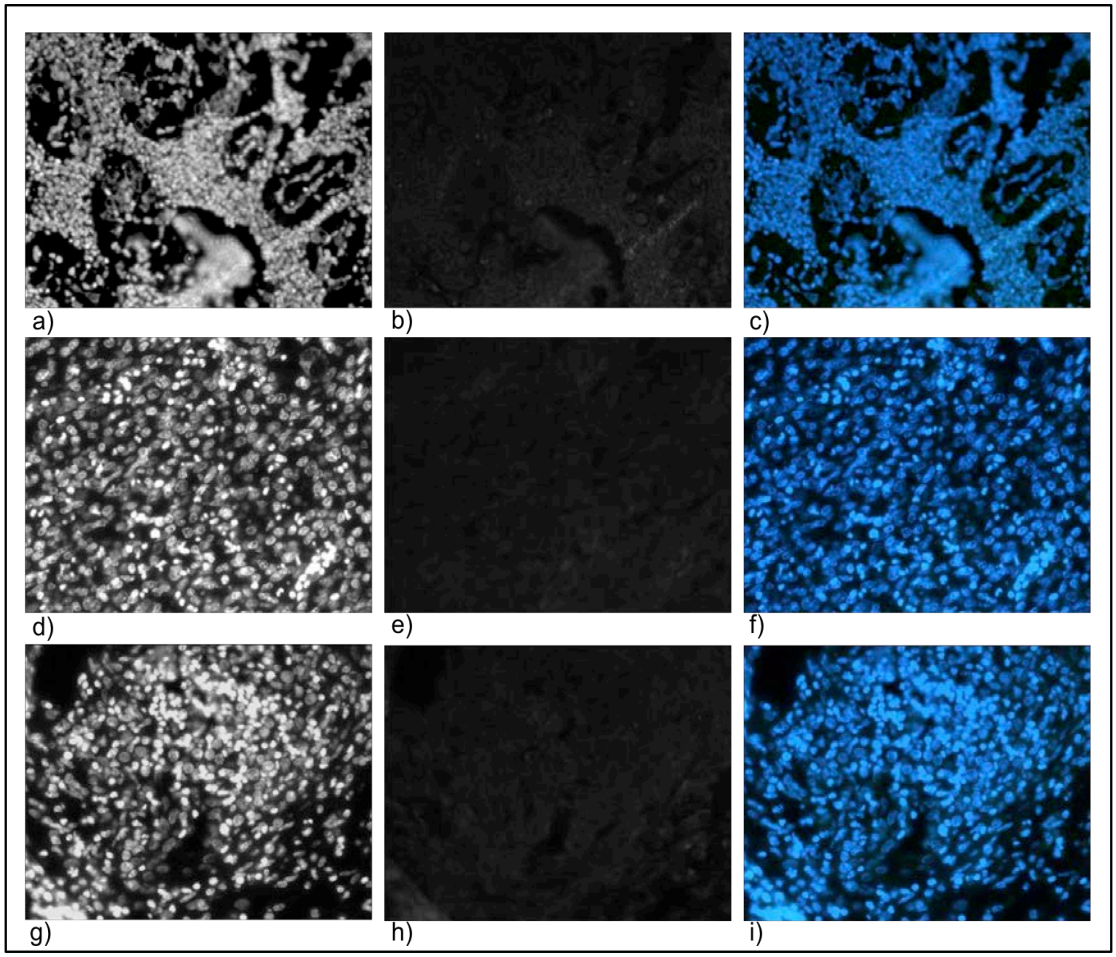


Figure 6.12: Immunofluorescence of *Eimeria tenella* infected caeca with pre-bleed sera

Immunofluorescence microscopy was used to visualise *E. tenella* infected chicken caeca probed with pre-bleed mouse sera from mice subsequently immunised with recombinant proteins *Et*AO2 (a-c), *Et*GPX1 (d-f) and *Et*OXIO1 (g-i). Tissue sections were then stained with 4`,6-diamidino-2-phenylindole (DAPI) (a), (d) and (g), highlighting the nuclei of the chicken cells, and counter-stained with Alexa Fluor® 488 (b), (e) and (h), and the images over-layed (c), (f) and (i).

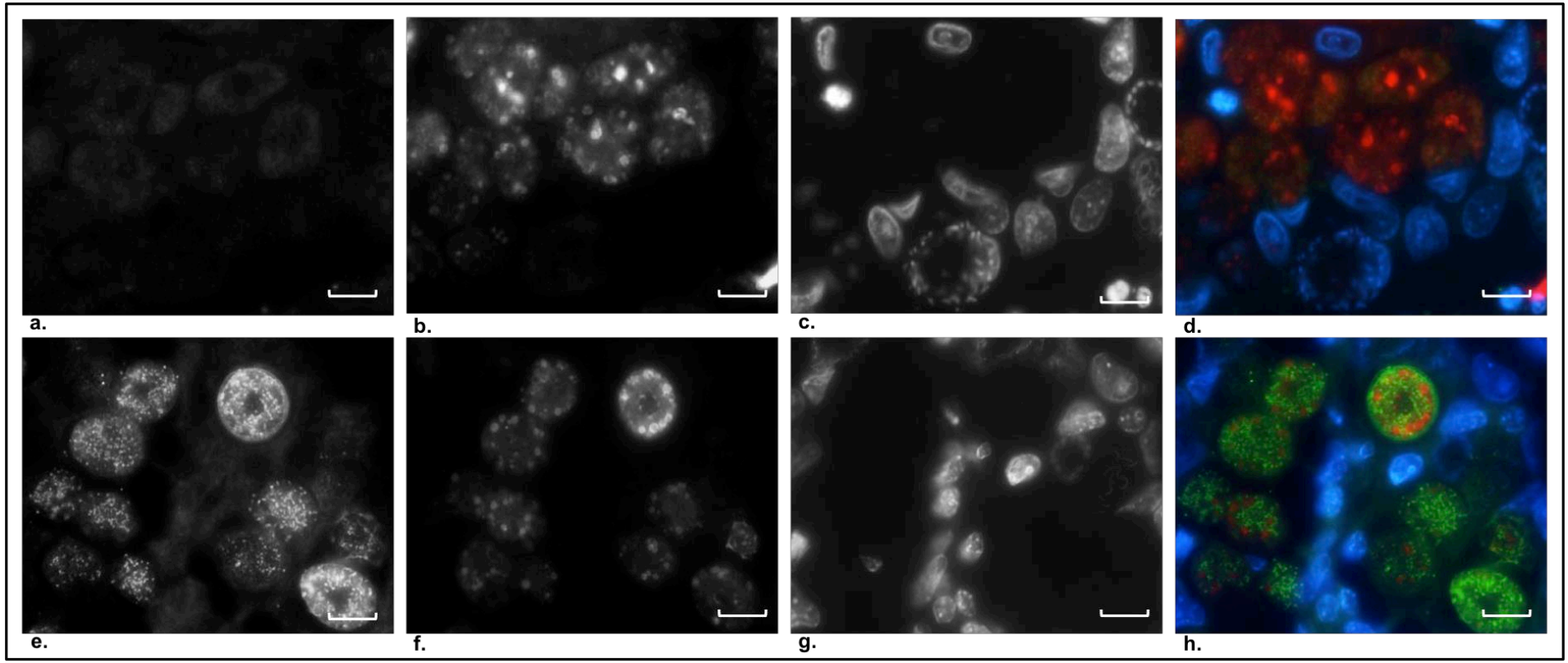


Figure 6.13: Immunolocalisation of antiserum to recombinant *Et*AO2 in *Eimeria tenella*-infected intestinal tissue at 134 hrs p.i. (a-d) and 144 hrs p.i. (e-h). *Eimeria tenella* gametocyte infected intestinal tissue stained with: (a) antiserum to *Et*AO2 visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (b) antiserum to rGAM56 visualised using anti-rabbit (whole molecule) conjugated to AlexaFluor594; (c) counterstained with 4` 6 – diamidino – 2 – phenylindone (DAPI); (d) overlay of (a), (b) and (c) at 134hrs p.i.; (e) antiserum to *Et*AO2 visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (f) antiserum to rGAM56 visualised using anti-rabbit (whole molecule) conjugated to AlexaFluor594; (g) counterstained with DAPI; and (h) overlay of (e), (f) and (g) at 144hrs p.i. Images were taken using the Zeiss Axio Imager M1 microscope at 40x and 60x. Bar is 10µm.

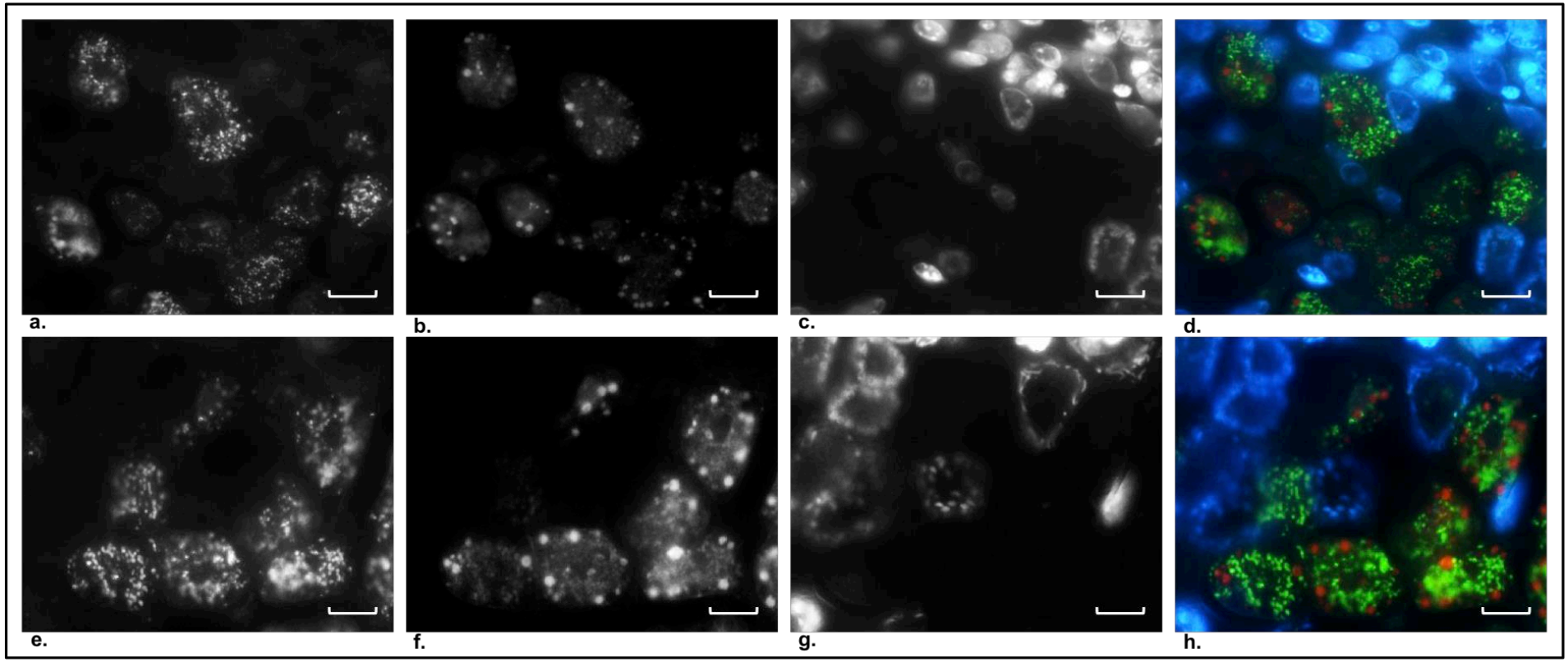


Figure 6.14: Immunolocalisation of antiserum to recombinant *Et*GPX1 in *Eimeria tenella*-infected intestinal tissue at 134 hrs p.i. (a-d) and 144 hrs p.i. (-h). *Eimeria tenella* gametocyte infected intestinal tissue stained with: (a) antiserum to *Et*GPX1 visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (b) antiserum to rGAM56 visualised using anti-rabbit (whole molecule) conjugated to AlexaFluor594; (c) counterstained with 4` 6 – diamidino – 2 – phenylindone (DAPI); (d) overlay of (a), (b) and (c) at 134hrs p.i.; (e) antiserum to *Et*GPX1 visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (f) antiserum to rGAM56 visualised using anti-rabbit (whole molecule) conjugated to AlexaFluor594; (g) counterstained with DAPI; and (h) overlay of (e), (f) and (g) at 144hrs p.i. Images were taken using the Zeiss Axio Imager M1 microscope at 40x and 60x. Bar is 10μm.

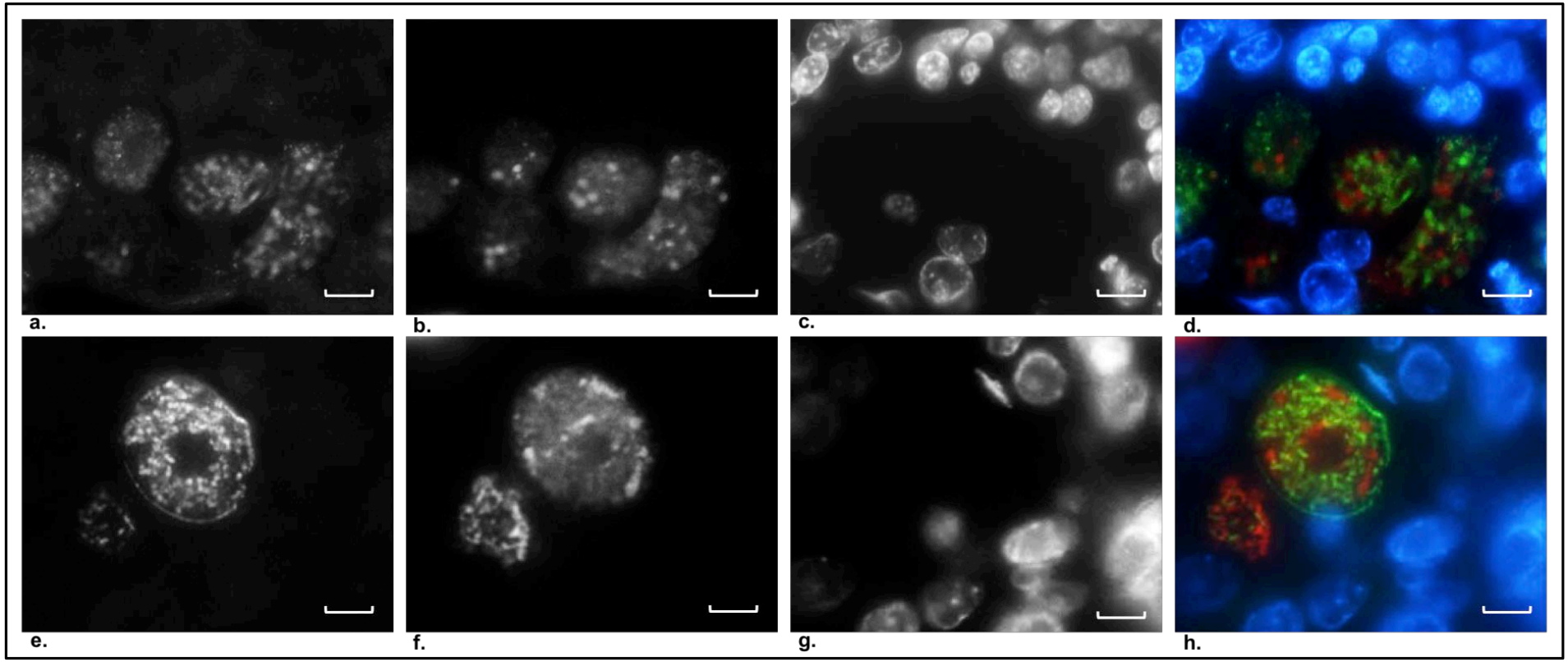


Figure 6.15: Immunolocalisation of antiserum to recombinant *Et*OXIO1 antibody in *Eimeria tenella*-infected intestinal tissue at 134 hrs p.i. (a-d) and 144 hrs p.i. (e-h)

Eimeria tenella gametocyte infected intestinal tissue stained with: (a) antiserum to *Et*OXIO1 visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (b antiserum to rGAM56 using anti-rabbit (whole molecule) conjugated to AlexaFluor594; (c) counterstained with 4` 6 – diamidino – 2 – phenylindone (DAPI); (d) overlay of (a), (b) an (c) at 134hrs p.i.; (e) antiserum to *Et*OXIO visualised using anti-mouse (λ chain specific) conjugated to AlexaFluor488; (f) antiserum to rGAM56 visualised using anti-rabbit (whol molecule) conjugated to AlexaFluor594; (g) counterstained with DAPI; and (h) overlay of (e), (f) and (g) at 144hrs p.i. Images were taken using the Zeiss Axio Imager M1 microscop at 40x and 60x. Bar is 10μm.

6.4 Discussion

Rigorous bioinformatic analyses identified five oxioreductase genes with potential to play roles in the oxidative catalysis of dityrosine bond formation and, therefore, in oocyst wall formation in *E. tenella:* two amine oxidases, *etao1* and *etao2*; a glutathione peroxidase, *etgpx1*; a quinone oxioreductase, *etqo1*; and a putative oxioreductase, *etoxio1* (Chapter 5). In this chapter, quantitative real-time (qRT) PCR was used to determine the stage-specific expression of these oxioreductase genes – relative transcript abundance was measured for each gene to determine the specificity of expression in different life cycle stages of *E. tenella,* with the view that genes showing an up-regulated transcript profile in gametocytes, especially in late stage gametocytes, were good candidates to play a role in oocyst wall formation. Transcripts of all five oxioreductases were found to be present in gametocytes of *E. tenella.* However, one of these genes, *etao1,* was expressed at relatively low levels (compared to other oxioreductase genes) and was expressed in all lifecycle stages; it was not possible to convincingly argue that its expression was up-regulated, let alone specific, in gametocytes. Whilst this does not necessarily rule it out completely as a player in oocyst wall formation, it was not investigated further here. Likewise, the quinone oxidase gene, *etqo1,* was expressed at relatively low levels and throughout parasite development. However, it does possess an intriguing pattern of expression, being relatively prominent in late gametocytes and in sporulated oocysts. This is intriguing because autofluorescence and anti-GAM56 antibody binding are apparent in the developing sporocyst wall as well as the oocyst wall (Belli *et al.*, 2003b). This enzyme is, therefore, worthy of future investigation. Unfortunately, antibodies produced against *Et*QO1 proved to be quite non-specific (Figures 6.9 and 6.11), so it could not be studied further here.

Three oxioreductase genes, *etao2, etoxio1* and *etgpx1,* were expressed prominently in gametocytes, with *etao2* and *etoxio1* appearing to be expressed exclusively in gametocytes (Figure 6.3). These two genes were equally well expressed in early and late gametocytes whereas *etgpx1* was especially prominently expressed in late gametocytes. Expression of this gene was also apparent in unsporulated and sporulated oocysts but at much lower levels than those seen in either early or late gametocytes.

To further investigate these genes, antibodies were raised to recombinant versions of each oxioreductase and reactivity assessed by ELISA. ELISAs demonstrated clearly that sera from all mice immunised with recombinant versions of oxioreductases reacted strongly to the immunising antigen (Figure 6.9). Although every effort was made to select areas specific to each oxioreductase for cloning to reduce the chances of shared epitopes, cross-reactivity was observed between some antisera and the other proteins. Antisera to *Et*QO1 was particularly promiscuous in

reactivity, as already mentioned. Several of the antisera raised to *Et*GPX1 also reacted with other oxioreductase antigens, especially *Et*QO1 (Figure 6.9b). It was, however, possible to select one antisera for each of *Et*GPX1, *Et*AO2 and *Et*OXIO1 for use in Western blots and immunolocalisation work.

Immunoblots were conducted with antisera raised to each of the five oxioreductase to determine if the antisera raised recognised specific protein bands in any of the protein lysate life cycle stages of *E. tenella.* For *Et*AO2, a single 130kDa band was observed in late stage gametocytes, unsporulated and sporulated oocysts (Figure 6.9b). This band was not observed in blots probed with antisera from the other oxioreductases. Interestingly, the predicted size of the *Et*AO2 native protein is 183kDa, somewhat larger than the band observed here. There may be several reasons to account for this. The first is that the *Et*AO2 protein may not actually migrate according to its molecular weight. An abundance of hydrophobic residues in a peptide segment, as is often observed in a transmembrane sequence of membrane proteins, can often give unclear representation of SDS-PAGE migration (Rath *et al.*, 2009). The *Et*AO2 sequence showed a 13kDa transmembrane domain at the N-terminal end of the protein, which may affect its migration. Protein tertiary structure can also affect polypeptide SDS-PAGE migration rates with the presence of di-sulphide bonds in a protein being linked to anomalously fast migration of reduced proteins (Dunker and Kenyon, 1976, Pitt-Rivers and Impiombato, 1968, Rath *et al.*, 2009). It has been suggested that this is because the di-sulphide bonds stay intact, giving the enzyme a more compact shape that allows faster migration (Dunker and Kenyon, 1976). *Et*AO2 contains a MAM domain characterised by four conserved cysteine residues that form di-sulphide bridges and this may explain the smaller than expected size observed with SDS-PAGE. Another explanation is that perhaps, *Et*AO2 is not as large as predicted on the database. Whilst peptides of the protein have been confirmed by mass spectrometry in earlier studies (Flowers, 2011), the entire protein sequence has yet to be confirmed. Given the high incidence of mis-predictions associated with the *E. tenella* genome database, described earlier in this study, it is possible that the parts of the sequence not confirmed by mass spectrometry are not present. The *Et*AO2 protein may, in fact, be smaller than the 183kDa predicted in the database.

The most intriguing aspect of the Western blot profile obtained for *Et*AO2 was the lack of detection of protein in early stage gametocytes and the clear detection of protein in oocysts; this does not reflect the qRT-PCR profile for this enzyme. On possible explanation for this is that, although gene expression may occur relatively early in gametocyte development, protein translation may be delayed. There is precedence for this idea. Thus, studies in another apicomplexan, *Plasmodium falciparum,* have suggested that there are regulatory events at the

mRNA level of stability and/or translation whereby the transcriptome and proteome in life cycle stages of the parasite do not correlate (Le Roch *et al.*, 2004). For example, mRNA transcript of six ubiquitin hydrolase proteins is detected in the gametocyte stage of *Plasmodium falciparum*, however, the protein is absent during this life-cycle stage, with message shown to be stored until subsequent developmental milestones are reached. It is thought that this delayed translation involves two RNA binding proteins belonging to the Puf family of translation and mRNA stability factors that have been identified in the *P. falciparum* genome. Similarly, *pbs21* is a sexual stage gene of *Plasmodium berghei* where the mRNA is detected in the female gametocytes but it is not until the ookinete stage that it is translated into protein (Paton *et al.*, 1993, Vervenne *et al.*, 1994).

Antibody was raised to *Et*OXIO1 using a highly hydrophilic section of the protein (Figure 6.8). When *E. tenella* protein lysates were probed with antiserum raised to this second gametocyte-specific oxioreductase a unique pattern of bands was observed (Figure 6.11c). A prominent band of 96kDa was observed in late stage gametocytes, corresponding to the predicted 96kDa of the *Et*OXIO1 in the database (Figure 6.11c). A number of other bands were also observed ranging from 130kDa to 50kDa; it is not clear what these represent. Gametocyte-specific expression and translation of this protein does, however, appear to be confirmed by these Western blot results.

A 30kDa band was observed in late stage gametocytes, unsporulated and sporulated oocysts of protein lysates probed with antisera to *Et*GPX1 (Figure 6.11e), confirming the qRT-PCR results. This is consistent with the predicted molecular weight for *Et*GPX1 is 27kDa. A further two bands, of 55kDa and 60kDa in size were observed in early and late stage gametocytes. Interestingly, these two bands were also observed in immunoblots of *E. tenella* protein lysates probed with antisera to *Et*QO1 (Figure 6.11). Again, it is unclear what these bands represent (they are not consistent with the predicted size of either *Et*GPX1 or *Et*QO1) but their presence does dictate some caution in interpretation of subsequent immunolocalistion experiments. Encouragingly, however, although some ELISA cross-reactivity of antiserum to *Et*GPX1 with *Et*AO2 was observed, no similarity in Western blotting patterns was seen.

Immunolocalisation studies showed clearly that *Et*AO2, *Et*OXIO1 and *Et*GPX1 are all present in macrogametocytes and not microgametocytes of *E. tenella.* Furthermore, all three have a punctate distribution within the cytoplasm of macrogametocytes, with no co-localisation to WFBs. In this respect, these results are similar to those obtained for *Et*SUB2 (Chapter 4), though the distribution of these oxioreductases is quite distinct from that of *Et*SUB2, which was much less "packaged" than the oxioreductases and more closely proximal to the periphery of the WFBs. Whether the oxioreductases truly are packaged remains to be proven and immune-electron microscopy will likely be required to resolve this question. However, the apparent packaging of

these enzymes makes some sense when considering their proposed function; perhaps they are stored until WFBs migrate to the periphery of macrogametocytes and disaggregate in preparation for oocyst wall formation. As argued for *Et*SUB2 in Chapter 5, premature exposure of GAM56 or GAM82 to oxioreductase activity could lead to premature dityrosine bonding between proteins with drastic effects on oocyst wall formation. In this respect, perhaps *Et*AO2 deserves special attention since mRNA for this enzyme is clearly produced in early gametocytes (Figure 6.3) but protein does not appear until later gametocytes (Figures 6.11 and 6.13) and is carried through into oocysts (Figure 6.11).

The localisation of oxioreductases in the cytoplasm of macrogametocytes of *E. tenlla* and not in WFBs contradicts previous findings showing peroxidase activity within WFBs of *E. maxima* (Belli *et al.*, 2006b, Belli *et al.*, 2003b). There is currently no satisfactory explanation for this discrepancy but, again, immune-electron microscopy may help resolve this contradiction, particularly if a means to purify WFBs can be developed, allowing combined proteomic and microscopic analysis of these organelles. The previous finding of peroxidase activity in association with the developing oocyst wall (Belli *et al.*, 2006b, Belli *et al.*, 2003b) is less difficult to explain; indeed, close examination of Figures 6.13h and 6.15h appears to reveal that at least some of the more mature macrogametocytes have *Et*AO2 and *Et*GPX1 distributed around the periphery of the parasite, perhaps in association with the developing oocyst wall.

As alluded to already, some caution needs to be exercised in the interpretation of the localisation of *Et*GPX1 because available antisera to this protein do have some tendency to cross-react with *Et*AO2, *Et*OXIO1 and *Et*QO1, at least in ELISAs (Figure 6.9). However, in favor of the localisation observed here are the facts that the antiserum to *Et*GPX1 chosen for immunolocalisation work: (1) was the least cross-reactive of the panel; (2) recognised bands in Western blots that were not recognized by antisera to *Et*AO2 or *Et*OXIO1; (3) recognised bands in all stages of development in a pattern that reflected the qRT-PCR results; and (4) localised only weakly to early gametocytes but strongly to later gametocytes, again reflecting the qRT-PCR results. The cross-reactivity seen in ELISAs may, at least partially, be explained by possible reactivity to the GenScript© tag in Expression Vector X1, which was used in the production of all recombinant oxioreductases used to immunise mice; such reactivity would not be apparent in Western blots or tissue sections. Thus, this Chapter documents further evidence supporting potential involvement of three oxioreductases in the catalysis of dityrosine bond formation as part of the development of the coccidian oocyst wall.

CHAPTER 7

RETROSPECTIVE, PERSPECTIVE and PROSPECTIVE

7.1 Retrospective

The bioinformatics, molecular and immunological investigations carried out in this thesis have identified three subtilase-like serine protease genes with qRT-PCR profiles that indicate an up-regulation of expression during the sexual development of *E. tenella.* One of these genes, *etsub2*, was exclusively expressed in gametocytes but equally strongly transcribed in early and late gametocytes. *Et*SUB2 protein, however, was much more abundant in late gametocytes implying that the message produced earlier is stored for use at a later time in gametocyte development. The intriguing and specific localisation of *Et*SUB2 around, but not within, WFBs of the macrogametocytes of *E. tenella* indicates that it perhaps lies in wait for the release of GAM56 and GAM82 (or perhaps additional as yet unidentified proteins) from the WFBs. Similarly, bioinformatics qRT-PCR analyses ultimately identified three oxioreductases, the transcription of which are very much up-regulated in gametocytes. Two of these genes (*etao2* and *etoxio1*) were transcribed equally strongly in early and late gametocytes but, like *Et*SUB2, *Et*AO2 was only translated into protein at a later stage of gametocyte development. The third oxioreductase gene, *etgpx1,* was transcribed more in late gametocytes than early gametocytes and some transcription of this gene continued through oocyst development. All three oxioreductases could be found in macrogametocytes but not microgametocytes, localising in the cytoplasm with a distinctive punctate distribution.

7.2 Perspective

It has been 7 years since Belli *et al.* (2006) proposed a model for how the tyrosine-rich wall proteins are incorporated into the coccidian oocyst wall. Describing the model in *Eimeria,* GAM56 and GAM82, stored in WFBII in macrogametocytes, are precursor proteins that are processed into small tyrosine-rich wall proteins. Subsequently, these wall proteins are oxidised and crosslinked via their tyrosine residues to form a matrix that becomes dehydrated and hardens to form the oocyst wall, with its accompanying, notorious resilience. This model has much in common with the formation of the nematode cuticle, where the assembly of cuticlins and collagens involves a number of critical catalytic pathways (reviewed by Page & Winter 2003): (1) collagens are synthesised as proproteins that are cleaved at the N-terminus by a subtilisin-like protease prior to cuticle formation; (2) the collagens (and cuticlins) are held together by di- and tri-tyrosine crosslinks; and (3) dual oxidase is the oxidative enzyme responsible for the generation of the tyrosine crosslinks. Mutations of key enzymes involved in any one of these steps results in the formation of a structurally defective cuticle lacking di- and tri-tyrosine crosslinks, and ultimately leads to parasite death (Page and Winter, 2003b). The Belli *et al.* model of oocyst wall formation predicted a very similar picture and generated several testable hypotheses: first, dityrosine bonds

will be present in the oocyst wall; second precursor proteins found in WFBs will be processed by gametocyte-specific proteases into smaller, tyrosine-rich peptides that are found in the oocyst wall; and, third, oxioreductases will catalyse cross-linking of proteins. There is already some evidence in recent literature that all these hypotheses are correct and investigations conducted in this thesis provide further evidence.

Under UV light in the excitation wavelength range of 330-385nm, oocysts of all coccidians autofluoresce blue, characterstic of dityrosine bonds present in the coccidian wall matrix (see Belli *et al.* 2006). Furthermore, the oocyst wall of *E. maxima* contains remarkable levels of dityrosine (Belli *et al.* 2003). Higher then many normal physiological situtaions, these measurable levels imply that the generation of dityrosine bonds in the oocyst wall is the result of a deliberate enzymatic process engineered by the parasite (Belli *et al.* 2006).

N-terminal sequencing of small tyrosine-rich wall proteins of 8, 10, 12 and 31 kDa shows that these proteins are the processed derivatives of proporteins GAM56 and GAM82 (Belli *et al.* 2003). Thus, antibodies to GAM56 and GAM82 also react with these smaller oocyst wall proteins and N-terminal sequencing maps the location of these proteins within GAM56 and GAM82. Katrib *et al.* (2012) recently provided evidence that the formation of these smaller proteins is dependent on serine proteases, most likely subtilase-like proteases from *Eimeria*.

As summarised above in *Retrospective,* and detailed in Chapters 3 and 4 of this thesis, we can now assign more confidently the subtilisin-like serine protease most likely to catalyse the processing of GAM56 and/or GAM82 into tyrosine-rich proteins destined for incorporation into the oocyst wall. That subtilisin is *Et*SUB2 (notwithstanding the continuing possibility that *Et*SUB1 and *Et*SUB5 may also be involved) and its addition to the Belli *et al.* model of oocyst wall formation is depicted in Figure 7.1.

Following subtilase-catalysed processing of GAM56 and GAM82, oxioreductases may catalyse the cross-linking of the small, tyrosine-rich derivatives and the formation of dityrosine bond formation. To date, an endogenous oxioreductase has not been isolated from coccidian parasites; however, an exogenous peroxidase has been shown to induce crosslinking of GAM56 *in vitro,* with accompanying formation of dityrosine bonds, as detected by HPLC with UV and visible detectors (Mai *et al.*, 2011). And, an oxidoreductase has been found in the *Toxoplasma* oocyst wall (Fritz *et al.*, 2012). In this thesis, the case for believing that three oxioreductases from *E. tenella* are candidates for catalysing dityrosine crosslinking in oocyst wall formation has been strengthened.

Thus, as summarised above in *Retrospective* and detailed in Chapters 5 and 6 of this thesis, three oxidoreductases were found at the right time and in the right place to help catalyse dityrosine

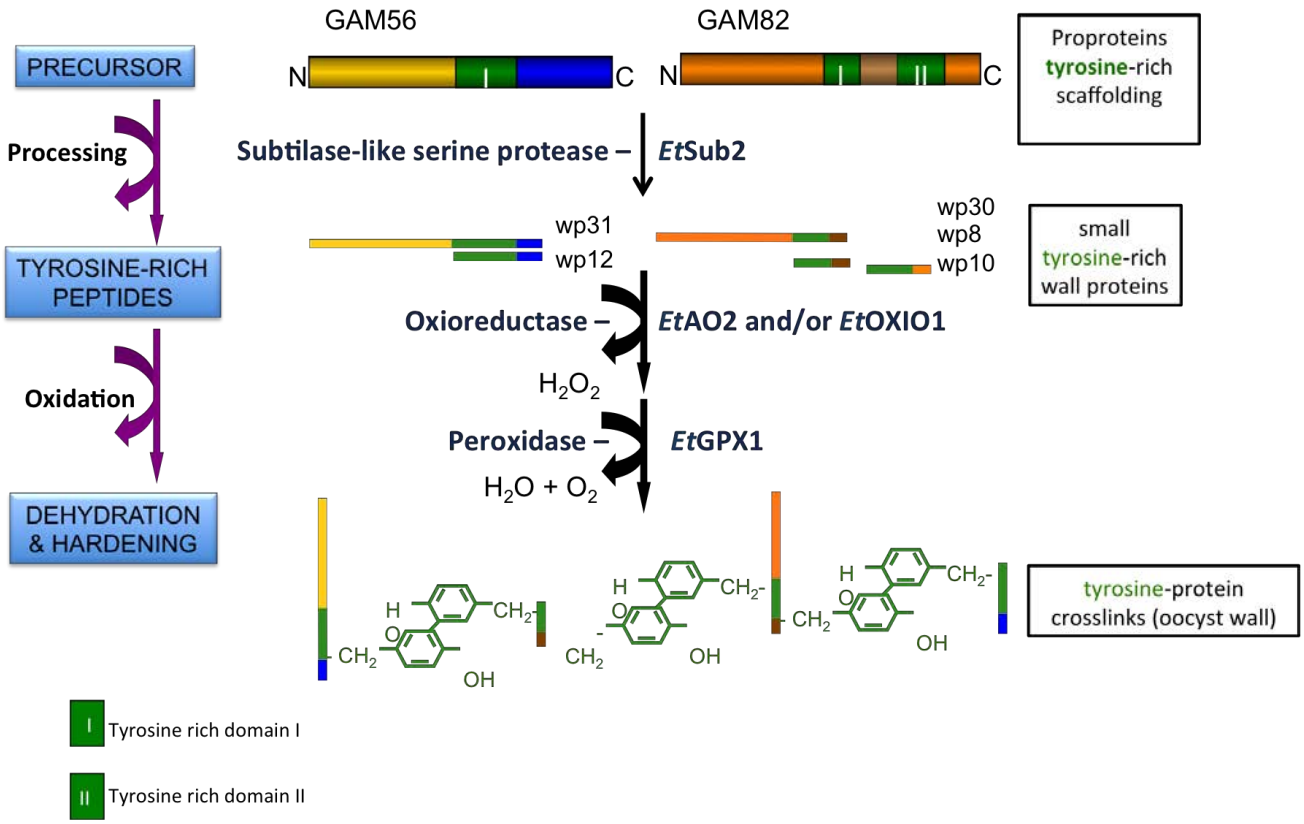


Figure 7.1: Proposed model of oocyt wall formation in *Eimeria*

Tyrosine-rich precursor proteins (GAM56 and GAM82), derived from WFB2 in macrogametocytes, are processed by a subtilase-like serine protease *Et*SUB2, into smaller tyrosine-rich proteins found in the oocyst wall. These wall proteins are then oxidised by an oxioreductase *Et*AO2 and/or *Et*OXIo1, and/or a peroxidase *Et*GPX1 to form dityrosine crosslinks. These bonds then dehydrate and form a hard, resistant oocyst wall. Modified from Belli *et al* (2006).

Abbreviations –C, C-terminal of Protein; N, N-terminal of protein; wp, wall protein;

bond formation, with *Et*AO2 producing the H_2O_2 required for dityrosine cross-linking, or conversely for bond formation between his-rich proteins such as GAM22; *Et*OXIO1, in a similar fashion, producing H_2O_2; and *Et*GPX1 being the peroxidase responsible for oxidation using H_2O_2 to produce dityrosine cross-links between oocyst wall proteins. The addition of these three enzymes to the Belli *et al.* model of oocyst wall formation is also presented in Figure 7.1.

Intriguingly, and arguably, significantly, the localisation of *Et*SUB2 and the three oxidoreductases is not, as previously predicted (Belli *et al.*, 2006b, Wallach, 2010), within WFBs. Rather, both types of enzyme are found in the cytoplasm of the macrogametocyte but have distinct distribution patterns. Thus, whilst *Et*SUB2 is found to localise around, but not within, WFBs, all three oxioreductases have a punctate distribution throughout the cytoplasm of macrogametocytes, almost as if they are packaged for later involvement than the subtilase. This all makes some sense. Premature exposure of GAM56 or GAM82 to either subtilase or oxioreductase activity, let alone both, could lead to premature dityrosine bonding and incorrect assembly of the matrix for the oocyst wall. Separating the three components may insure against this.

7.3 Prospective

The findings of this thesis generate several avenues of research to address outstanding questions on the molecular basis for coccidian oocyst wall formation. A series of experiments at the cell, protein and gene level can be imagined. It has already been raised (Chapters 4 and 6) that the localisation of *Et*SUB, *Et*AO2, *Et*OXIO1 and *Et*GPX1 needs to be better defined to help resolve just how closely *Et*SUB2 is associated with WFBs and whether or not the oxioreductases are, indeed, packaged and stored in some way within the cytoplasm of the microgametocyte. This might be achieved through use of immuno-electron microscopy (see, *eg,* Ferguson *et al.* 2003). Alternatively, new technologies like the DeltaVision OMX Blaze™ imaging system, which allows ultra-high resolution microscopy on live cells (see, *eg,* (Riglar *et al.*, 2013), may confer a decided advantage when trying to determine the genesis and movement of subtilisins and oxioreductases within developing macrogametocytes.

Another useful experiment, at the cellular level, would be one to demonstrate that antibodies and/or specific inhibitors of the active sites of *Eimeria* subtilases can inhibit oocyst formation. This would probably need to be carried out in an *in vitro* culture system because the likelihood of delivering an antibody or inhibitor to the right place, at the right time and in the right concentration in an infected chicken would seem to be extremely low. However, there are several challenges in achieving this aim *in vitro* too. First, more complete data on the sequence of these various enzymes needs to be gained so that a reasonable, valid model of the active sites can be

determined to facilitate the design of specific inhibitors and/or the production of specific antibodies. Second, a reliable *in vitro* cultivation system for the coccidian lifecycle needs to be developed. This has proved extremely challenging in the past, with even the best laboratories not able to reliably cultivate parasites beyond the first couple of phases of schizogony let alone into the sexual phase and through to the production of oocysts. However, recent breakthroughs achieved using the rodent coccidian, *Eimeria nieschulzi* (Chen *et al.*, 2013) encourage optimism on this score. It is likely that *E. nieschulzi* possesses homologues of the enzymes identified in *E. tenella* in this thesis since it is already known that it possesses a homologue of GAM56 (Hanig *et al.*, 2012).

At the protein level there are two sets of experiments that can immediately be envisaged. The first again relies on the development of a valid structural model of *Et*SUB2 so that inhibitors of its active site can be designed and synthesised. If this is achieved, these inhibitors could be used in the GAM56 processing assay developed by Katrib *et al.* (2012) to demonstrate that, indeed, *Et*SUB2 is involved in the processing of GAM56 into smaller, tyrosine-rich proteins, ready for cross-linking. The second experiment is perhaps even more challenging because it requires the production and/or purification of active *Et*AO2, *Et*OXIO1 and *Et*GPX1 so that these can be used, either alone or in combination, in the assay developed by Mai *et al.* (2011) to demonstrate that these enzymes can, in fact, cross-link a truncated version of GAM56, to generate dimers, trimers and polymers of proteins bonded together by dityrosine links.

Arguably, the most informative and important future experiments to carry out are those that involve manipulation of *Eimeria* at the level of the gene. Plasmid-mediated transient and stable transfection systems have been established for many years in apicomplexan parasites such as *Toxoplasma* and *Plasmodium*. However, the lack of regulatory DNA sequences and noted difficulty of completing the lifecycle of *Eimeria in vitro,* have meant that genetic manipulation has lagged behind in this parasite (Hao *et al.*, 2007, Kelleher and Tomley, 1998, Shi *et al.*, 2008) and we are now playing catch up with the molecular techniques. Only recently, has a stable transfection system been established for *Eimeria* (Clark *et al.*, 2008), which now allows the possibility of using transgenics to analyse the role of certain genes *in situ.*

Work has started to develop a transgenic construct of the *etsub5* gene under the control of its own promoter using the reporter gene tomato fluorescent protein as a reporter label. Promoters of both constitutive and stage-specific genes have been used previously to drive expression of the reporter gene to identify expression *in situ* in *E. tenella.* The yellow fluorescent protein, *eyfp,* marker gene was used as a reporter gene with *E. tenella* microneme protein 1 (*Et*MIC1), whereby the *Et*MIC1 promoter drives EYFP expression in sporulated but not unsporulated oocysts (Yin *et al.*,

2011). The reporter *eyfp* did not alter its expression. Rieux *et al.* (2012) recently reported the successful genetic tagging of a cathepsin protease in *E.* tenella, allowing them to map its location and movement within the sporulating oocysts. And, perhaps most significantly, in the context of the work carried out in this thesis, a reporter tag has been successfully added to the *E. nieschulzi gam56* promotor (Chen *et al.*, 2013, Hanig *et al.*, 2012). Applied to stage-specific genes *etsub1, etao2, etoxio1* genes, such techniques may allow the expression of these genes in macrogametocytes under their natural promoter to be further analysed, and the transport, location and functions of these genes elaborated. Moreover, analysis of the several subtilases and oxioreductases for which it was not possible to raise specific antibodies (see Chapters 4 and 6) using reporter genes would allow more definitive evaluation of the timing of their expression during microgametocyte and oocyst development. These experiments could conceivably be done via infection of chickens but, for optimal results and ease of analysis, such experiments require a reliable *in vitro* cultivation system, combined with the use of new technologies like the DeltaVision OMX Blaze™ imaging system to make possible high resolution microscopy on live cells, in real time.

Currently, a method for verification of gene function by a targeted gene knockout (KO) system remains elusive in *Eimeria,* complicated by low insertion efficiency and a requirement for serial *in vivo* passage to select and stabilise transgenic parasite lines (Chapman *et al.*, 2013, Clark *et al.*, 2008). However, the successful development of a gene KO system in the homologous coccidian parasite *T. gondii* has shown to be instrumental in studying gene function in the parasite (Donald and Roos 1994) and, as such, these techniques could be manipulated for use in *Eimeria*. The use of positive/negative selectable markers has been developed for homologous recombination resulting in allelic replacement and gene knockouts in *T. gondii* (Donald and Roos 1994, Fox *et al*., 1999; 2001, Radke and White, 1999). This has been shown to be successful to identify genes expressed in a stage-specific manner (Knoll and Boothroyd 1998, Matrajt *et al*., 2002b). And the recent development of *ku80* knockout strains of *T. gondii* (Fox *et al.*, 2009, Huynh and Carruthers, 2009, Rommereim *et al.*, 2013) have revolutionised the efficiency with which gene knockouts can be achieved for this coccidian parasite. Developing a knockout system in *Eimeria* based on methods currently utilised in *T. gondii* would allow a more detailed investigation into the role of genes in the parasite, not just in oocyst wall formation, but other biological pathways too. Work in our laboratory has also begun to try and develop a knockout construct of the *etsub5* subtilisin gene based on the *T. gondii* system of homologous recombination and allelic knockout (Donald and Roos 1994). A consideration in developing a knockout system in *Eimeria,* for the genes identified in this study, is that they may prove to be essential in the development of the parasite. To overcome this in other Apicomplexa a tetracycline transactivator-based inducible system has

been developed (Meissner 2002) and shown to successfully disrupt essential genes in *T. gondii* while avoiding pleitropic effects (Mital 2005, Mazumdar 2006). Further to this, the same system has since been established in *Plasmodium* (Meissner and Soldati 2005). If gene knockout technologies ultimately prove inapplicable to *Eimeria,* then the only alternative to achieve such definitive studies may be to use *T. gondii.* This, of course, carries its own challenges, requiring either infections of cats or the development of an *in vitro* cultivation system for the coccidian lifecycle of *T. gondii,* which has so far been even more elusive than for *Eimeria.* However, like *Eimeria,* recent success with the related coccidian parasite of pigs, *Isospora suis* (Worliczek *et al.*, 2013), encourages optimisim that this goal is, in fact, achievable.

The result of all these proposed experiments will be the definitive demonstration of whether or not subtilisins and oxioreductases are critical for the catalysis of dityrosine bond formation and, hence, the development of the coccidian oocyst wall. This will validate these enzymes as targets for novel drugs or vaccines to prevent transmission of coccidian parasites from host to host. The potential impact of this on animal and human well-being could be enormous.

REFERENCES

AESCHBACH, R., AMADO, R. & NEUKOM, H. 1976. Formation of dityrosine cross-links in proteins by oxidation of tyrosine residues. *Biochim Biophys Acta,* 439**,** 292-301.

ALAM, A., BHATNAGAR, R. K. & CHAUHAN, V. S. 2012. Expression and characterization of catalytic domain of Plasmodium falciparum subtilisin-like protease 3. *Molecular and biochemical parasitology,* 183**,** 84-9.

ALBANESE, A. & SMETANA, H. 1937. Studies on the effect of X-rays on the pathogenicity of *Eimeria tenella*. *Am J Hyg,* 26**,** 27-39.

ALLEN, P. & TEASDALE, D. 1994. Plasma Nitrite increases during Eimeria Acervulina infections in chickens. *Federation of American Societies for Experimental Biology Conference.*

ALLEN, P. C. 1997. Nitric Oxide Production During Eimeria tenella Infections in Chickens. *Poultry Science,* 76**,** 810-3.

ALLEN, P. C. & FETTERER, R. H. 2002. Recent advances in biology and immunobiology of Eimeria species and in diagnosis and control of infection with these coccidian parasites of poultry. *Clin Microbiol Rev,* 15**,** 58-65.

ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res,* 25**,** 3389-402.

AN, Z., JING, W., LIU, Y. & ZHANG, W. 2008. Hydrogen peroxide generated by copper amine oxidase is involved in abscisic acid-induced stomatal closure in Vicia faba. *J Exp Bot,* 59**,** 815-25.

ANDREWS, R., O'DONOGHUE, P., ADAMS, M. & PROWSE, S. 1990. Enzyme markers for the genetic characterization of avian Eimeria spp. *Parasitol Res,* 76**,** 627-9.

ARASTU-KAPUR, S., PONDER, E. L., FONOVIC´, U. P., YEOH, S., YUAN, F., FONOVIC´, M., GRAINGER, M., PHILLIPS, C. I., POWERS, J. C. & BOGYO, M. 2008. Identification of proteases that regulate erythrocyte rupture by the malaria parasite Plasmodium falciparum. *Nature Chemical Biology,* 4**,** 203-213.

AUSTRALIA, B. F. O. 2006. Organic Annual Edition 3. *In:* CORPERATION, B. F. O. A. (ed.). Biological Farmers of Australia Corperation.

BAKER, D., SHIAV, A. K. & AGARD, D. A. 1993. The role of pro-regions in protein folding. *Current Opinion in Cell Biology,* 5**,** 966-70.

BARALE, J.-C., BLISNICK, T., FJIOKA, H., ALZARI, P. M., AIKAWA, M., BRAUN-BRETON, C. & LANGSLEY, G. 1999a. *Plasmodium falciparum* subtilisin-like protease 2, a merozoite candidate for the merozoite surface protein 1-42 maturase. *Proc. Natl. Acad. Sci. USA,* 96**,** 6445-6450.

BARALE, J. C., BLISNICK, T., FUJIOKA, H., ALZARI, P. M., AIKAWA, M., BRAUN-BRETON, C. & LANGSLEY, G. 1999b. Plasmodium falciparum subtilisin-like protease 2, a merozoite candidate for the merozoite surface protein 1-42 maturase. *Proc Natl Acad Sci U S A,* 96**,** 6445-50.

BELLI, S., MAI, K., SKENE, C., GLEESON, M., WITCOMBE, D., KATRIB, M., FINGER, A., WALLACH, M. & SMITH, N. 2004a. Characterisation of the antigenic and immunogenic properties of bacterially expressed, sexual stage antigens of the coccidian parasite, Eimeria maxima. *Vaccine,* 22**,** 4316-25.

BELLI, S., WITCOMBE, D., WALLACH, M. G. & SMITH, N. C. 2002a. Functional genomics of *gam56:* characterisation of the role of a 56 kilodalton sexual stage antien in oocyst wal formation in *Eimeria maxima*. *International Journal for Parasitology,* 32**,** 1727-1737.

BELLI, S. I., FERGUSON, D. J., KATRIB, M., SLAPETOVA, I., MAI, K., SLAPETA, J., FLOWERS, S. A., MISKA, K. B., TOMLEY, F. M., SHIRLEY, M. W., WALLACH, M. G. & SMITH, N. C. 2009. Conservation of proteins involved in oocyst wall formation in Eimeria maxima, Eimeria tenella and Eimeria acervulina. *Int J Parasitol,* 39**,** 1063-70.

BELLI, S. I., LEE, M., THEBO, P., WALLACH, M. G., SCHWARTSBURD, B. & SMITH, N. C. 2002b. Biochemical characterisation of the 56 and 82 kDa immunodominant gametocyte antigens from Eimeria maxima. *Int J Parasitol,* 32**,** 805-16.

BELLI, S. I., MAI, K., SKENE, C. D., GLEESON, M. T., WITCOMBE, D. M., KATRIB, M., FINGER, A., WALLACH, M. G. & SMITH, N. C. 2004b. Characterisation of the antigenic and immunogenic properties of bacterially expressed, sexual stage antigens of the coccidian parasite, Eimeria maxima. *Vaccine,* 22**,** 4316-25.

BELLI, S. I., SMITH, N. C. & AND FERGUSON, D. J. P. 2006a. The coccidian oocyst: a tough nut to crack! *Trends in Parasitology,* 22**,** 416-423.

BELLI, S. I., SMITH, N. C. & FERGUSON, D. J. 2006b. The coccidian oocyst: a tough nut to crack! *Trends Parasitol,* 22**,** 416-23.

BELLI, S. I., WALLACH, M. G., LUXFORD, C., DAVIES, M. J. & AND SMITH, N. C. 2003a. Roles of Tyrosine-Rich Precursor Glycoproteins and Dityrosine-and 3,4- Dihydroxyphenylalanine-Mediated Protein Cross-Linking in Development of the Oocyst Wall in the Coccidian Parasite *Eimeria maxima*. *Eukaryotic Cell,* 2**,** 456-464.

BELLI, S. I., WALLACH, M. G., LUXFORD, C., DAVIES, M. J. & SMITH, N. C. 2003b. Roles of tyrosine-rich precursor glycoproteins and dityrosine- and 3,4-dihydroxyphenylalanine-mediated protein cross-linking in development of the oocyst wall in the coccidian parasite Eimeria maxima. *Eukaryot Cell,* 2**,** 456-64.

BELLI, S. I., WALLACH, M. G. & SMITH, N. C. 2003c. Cloning and characterisation of the 82 kDa tyrosine-rich sexual stage glycoprotein, GAM82, and its role in oocyst wall formation in the apicomplexan parasite, *Eimeria maxima*. *Gene,* 307**,** 201-212.

BELLI, S. I., WALLACH, M. G. & SMITH, N. C. 2003d. Cloning and characterization of the 82 kDa tyrosine-rich sexual stage glycoprotein, GAM82, and its role in oocyst wall formation in the apicomplexan parasite, Eimeria maxima. *Gene,* 307**,** 201-12.

BELLI, S. I., WITCOMBE, D., WALLACH, M. G. & SMITH, N. C. 2002c. Functional genomics of gam56: characterisation of the role of a 56 kilodalton sexual stage antigen in oocyst wall formation in Eimeria maxima. *Int J Parasitol,* 32**,** 1727-37.

BERGER, I. & NNADOZIE, J. 1993. Onchocerciasis and other eye problems in developing countries: a challenge for optometrists. *Journal of the American Optometric Association,* 64**,** 699-702.

BIGGS, P. M. 1982. The world of poultry disease. *Avian Pathol,* 11**,** 281-300.

BLACKMAN, M. 2000a. Proteases involved in Erythrocyte Invasion by the Malaria Parasite: Function and Potential as Chemotherapeutic Targets. *Current Drug Targets,* 1**,** 59-83.

BLACKMAN, M. J. 2000b. Proteases involved in erythrocyte invasion by the malaria parasite: function and potential as chemotherapeutic targets. *Curr Drug Targets,* 1**,** 59-83.

BLACKMAN, M. J., FUJIOKA, H., STAFFORD, W. H., SAJID, M., CLOUGH, B., FLECK, S. L., AIKAWA, M., GRAINGER, M. & HACKETT, F. 1998a. A subtilisin-like protein in secretory organelles of Plasmodium falciparum merozoites. *J Biol Chem,* 273**,** 23398-409.

BLACKMAN, M. J., FUJIOKA, H., STAFFORD, W. H. L., SAJID, M., CLOUH, B., FLECK, S. L., AIKAWA, M., GRAINGER, M. & AND HACKETT, F. 1998b. A Subtilisin-like Protein in Secretory Organelles of *Plasmodium falciparum* Merozoites. *The Journal of Biological Chemistry,* 273**,** 23398-23409.

BRADLEY, P. J., WARD, C., CHENG, S. J., ALEXANDER, D. L., COLLER, S., COOMBS, G. H., DUNN, J. D., FERGUSON, D. J., SANDERSON, S. J., WASTLING, J. M. & BOOTHROYD, J. C. 2005. Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in Toxoplasma gondii. *J Biol Chem,* 280**,** 34245-58.

BRADY, J. D. & FRY, S. C. 1997. Formation of Di-Isodityrosine and Loss of Isodityrosine in the Cell Walls of Tomato Cell-Suspension Cultures Treated with Fungal Elicitors or H2O2. *Plant Physiol,* 115**,** 87-92.

BRENT, M. R. 2008. Steady progress and recent breakthroughs in the accuracy of automated genome annotation. *Nat Rev Genet,* 9**,** 62-73.

BRIZA, P., BREITENBACH, M., ELLINGER, A. & SEGALL, J. 1990. Isolation of two developmentally regulated genes involved in spore wall maturation in *Saccharomyces cerevisiae*. *Gene and Development,* 4**,** 1775-1789.

CANNING, E. U. & ANWAR, M. 1968. Studies on meiotic division in coccidial and malarial parasites. *J Protozool,* 15**,** 290-8.

CARRUTHERS, V. B., SHERMAN, G. D. & SIBLEY, L. D. 2000a. The Toxoplasma adhesive protein MIC2 is proteolytically processed at multiple sites by two parasite-derived proteases. *J Biol Chem,* 275**,** 14346-53.

CARRUTHERS, V. B., SHERMAN, G. D. & SIBLEY, L. D. 2000b. The *Toxoplasma* Adhesive Protein MIC2 is Proteolytically Processed at Mutliple Sites by Two Parasite-derived Proteases. *The Journal of Biological Chemistry,* 275**,** 14346-14353.

CARRUTHERS, V. B. & SIBLEY, L. D. 1999. Mobilization of intracellular calcium stimulates microneme discharge in Toxoplasma gondii. *Molecular Microbioogy,* 31**,** 421-8.

CEVALLOS, A. M., ZHANG, X., WALDOR, M. K., JAISON, S., ZHOU, X., TZIPORI, S., NEUTRA, M. R. & WARD, H. D. 2000. Molecular cloning and expression of a gene encoding Cryptosporidium parvum glycoproteins gp40 and gp15. *Infect Immun,* 68**,** 4108-16.

CHANCE, B. & WILLIAMS, G. R. 1955. Respiratory enzymes in oxidative phosphorylation. IV. The respiratory chain. *J Biol Chem,* 217**,** 429-38.

CHANG, H.-S. & ZEPEDA, L. 2005. Consumer perceptions and demand for organic food in Australia: Focus group discussions. *Renewable Agric. Food Syst.,* 20**,** 155–167.

CHAPMAN, H. 2000. Practical use of vaccines for the control of coccidiosis in the chicken *World's Poult Sci J,* 56**,** 7-20.

CHAPMAN, H. & CHERRY, T. 1997. Eyespray vaccination: Infectivity and development of immunity to *Eimeria acervulina* and *Eimeria tenella*. *J Appl Poult Res,* 6**,** 274-278.

CHAPMAN, H. D. 1997. Biochemical, genetic and applied aspects of drug resistance in Eimeria parasites of the fowl. *Avian Pathol,* 26**,** 221-44.

CHAPMAN, H. D., BARTA, J. R., BLAKE, D., GRUBER, A., JENKINS, M., SMITH, N. C., SUO, X. & TOMLEY, F. M. 2013. A selective review of advances in coccidiosis research. *Adv Parasitol,* 83**,** 93-171.

CHAPMAN, H. D. & HACKER, A. B. 1994. Sensitivity of field isolates of Eimeria from two broiler complexes to anticoccidial drugs in the chicken. *Poult Sci,* 73**,** 1404-8.

CHEN, H., WIEDMER, S., HANIG, S., ENTZEROTH, R. & KURTH, M. 2013. Development of Eimeria nieschulzi (Coccidia, Apicomplexa) Gamonts and Oocysts in Primary Fetal Rat Cells. *Journal of parasitology research,* 2013**,** 591520.

CHEN, L.-L., REKOSH, D. M. & LOVERDE, P. T. 1992. *Schistosoma mansoni* p48 eggshell protein gene: characterization, developmentally regulated expression and comparison to the p14 eggshell protein gene. *Molecular and Biochemical Parasitology,* 52**,** 39-52.

CHENG, G., ZHAO, P., TANG, X. F. & TANG, B. 2009. Identification and characterization of a novel spore-associated subtilase from Thermoactinomyces sp. CDF. *Microbiology,* 155**,** 3661-72.

CHENNA, R., SUGAWARA, H., KOIKE, T., LOPEZ, R., GIBSON, T. J., HIGGINS, D. G. & THOMPSON, J. D. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.,* 31**,** 3497-500.

CLARK, J. D., BILLINGTON, K., BUMSTEAD, J. M., OAKES, R. D., SOON, P. E., SOPP, P., TOMLEY, F. M. & BLAKE, D. P. 2008. A toolbox facilitating stable transfection of Eimeria species. *Mol Biochem Parasitol,* 162**,** 77-86.

CONSEIL, V., SOETE, M. & DUBREMETZ, J. 1999. Serine protease Inhibitors Block Invasion of Host Cells by *Toxoplasma gondii*. *Antimicrobial Agents and Chemotherapy* 43**,** 1358-1361.

DANFORTH, H. D. 1998. Use of live oocyst vaccines in the control of avian coccidiosis: experimental studies and field trials. *Int J Parasitol,* 28**,** 1099-109.

DANFORTH, H. D., WATKINS, K., MARTIN, A. & DEKICH, M. 1997. Evaluation of the efficacy of Eimeria maxima oocyst immunization with different strains of day-old broiler and roaster chickens. *Avian Dis,* 41**,** 792-801.

DAVIES, M., FU, S., WANG, H. & WANG, R. 1999. Stable markers of oxidant damage to proteins and their application in the study of human disease. *Free Radical Biology and Medicine,* 27**,** 1151-63.

DE DEKEN, X., WANG, D., MANY, M. C., COSTAGLIOLA, S., LIBERT, F., VASSART, G., DUMONT, J. E. & MIOT, F. 2000. Cloning of two human thyroid cDNAs encoding new members of the NADPH oxidase family. *J Biol Chem,* 275**,** 23227-33.

DEITS, T., FARRANCE, M., KAY, E. S., MEDILL, L., TURNER, E. E., WEIDMAN, P. J. & SHAPIRO, B. M. 1984. Purification and properties of ovoperoxidase, the enzyme responsible for hardening the fertilization membrane of the sea urchin egg. *J Biol Chem,* 259**,** 13525-33.

DEL CACHO, E., PAGES, M., GALLEGO, M., MONTEAGUDO, L. & SANCHEZ-ACEDO, C. 2005. Synaptonemal complex karyotype of Eimeria tenella. *Int J Parasitol,* 35**,** 1445-51.

DING, Y., WANG, H., SHEN, G. & YU, R. 2005. Enzyme-catalyzed amplified immunoassay for the detection of Toxoplasma gondii-specific IgG using Faradaic impedance spectroscopy, CV and QCM. *Anal Bioanal Chem,* 382**,** 1491-9.

DLUZEWSKI, A., RANGACHARI, K., WILSON, R. & GRATZER, W. 1986. *Plasmodium falciparum:* protease inhibitors and inhibition of erythrocyte invasion. *Experimental Parasitology,* 62**,** 416-422.

DONKO, A., PETERFI, Z., SUM, A., LETO, T. & GEISZT, M. 2005. Dual oxidases. *Philos Trans R Soc Lond B Biol Sci,* 360**,** 2301-8.

DUCKERT, P., BRUNAK, S. & BLOM, N. 2004. Prediction of proprotein convertase cleavage sites. *Protein Eng Des Sel,* 17**,** 107-12.

DUNKER, A. K. & KENYON, A. J. 1976. Mobility of sodium dodecyl sulphate - protein complexes. *Biochem J,* 153**,** 191-7.

DYBAS, J. M., MADRID-ALISTE, C. J., CHE, F., NIEVES, E., RYKUNOV, D., ANGELETTI, R. H., WEISS, L. M., KIM, K. & ANDRAS FISER, A. 2008. Computational Analysis and Experimental Validation of Gene Predictions in Toxoplasma gondii. *PLoS ONE,* 3**,** e3899.

EDENS, W. A., SHARLING, L., CHENG, G., SHAPIRA, R., KINKADE, J. M., LEE, T., EDENS, H. A., TANG, X., SULLARDS, C., FLAHERTY, D. B., BENIAN, G. M. & LAMBETH, J. D. 2001. Tyrosine cross-linking of extracellular matrix is catalyzed by Duox, a multidomain oxidase/peroxidase with homology to the phagocyte oxidase subunit gp91phox. *J Cell Biol,* 154**,** 879-91.

EDQVIST, J. & BLOMQVIST, K. 2006. Fusion and fission, the evolution of sterol carrier protein-2. *J Mol Evol,* 62**,** 292-306.

ESCHENBACHER, K. H., EGGLI, P., WALLACH, M. & BRAUN, R. 1996. Characterization of a 14 kDa oocyst wall protein of Eimeria tenella and E. acervulina. *Parasitology,* 112 (Pt 2)**,** 169-76.

FEDERICO, R. & ANGELINI, R. 1986. Occurrence of diamine oxidase in the apoplast of pea epicotyls. *Planta,* 167**,** 300–302.

FELIG, P. 1975. Amino acid metabolism in man. *Annu Rev Biochem,* 44**,** 933-55.

FERGUSON, D. J., BELLI, S. I., SMITH, N. C. & WALLACH, M. G. 2003a. The development of the macrogamete and oocyst wall in Eimeria maxima: immuno-light and electron microscopy. *Int J Parasitol,* 33**,** 1329-40.

FERGUSON, D. J., BELLI, S. I., SMITH, N. C. & WALLACH, M. G. 2003b. The development of the macrogamete and oocyst wall in Eimeria maxima: immuno-light and electron microscopy. *International Journal of Parasitology,* 33**,** 1329-40.

FERGUSON, D. J., BIRCH-ANDERSEN, A., HUTCHISON, W. M. & SIIM, J. C. 1977. Ultrastructural studies on the endogenous development of Eimeria brunetti. IV. Formation and structure of the oocyst wall. *Acta Pathol Microbiol Scand B.,* 85**,** 201-211.

FERGUSON, D. J., BRECHT, S. & SOLDATI, D. 2000. The microneme protein MIC4, or an MIC4-like protein, is expressed within the macrogamete and associated with oocyst wall formation in Toxoplasma gondii. *Int J Parasitol,* 30**,** 1203-9.

FERGUSON, D. J., HUTCHISON, W. M. & SIIM, J. C. 1975. The ultrastructural development of the macrogamete and formation of the oocyst wall of Toxoplasma gondii. *Acta Pathol Microbiol Scand B. ,* 83**,** 491-505.

FERGUSON, D. J. P. 2002. *Toxoplasma gondii* and sex: essential or optional extra? *Trends in Parasitology,* 18**,** 355-259.

FERNIE, A. R., CARRARI, F. & SWEETLOVE, L. J. 2004. Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Curr Opin Plant Biol,* 7**,** 254-61.

FETTERER, R. H. & BARFIELD, R. C. 2003. Characterization of a developmentally regulated oocyst protein from *Eimeria tenella*. *Experimental Parasitology,* 89**,** 553-564.

FETTERER, R. H. & HILL, D. E. 1993. The occurrence of phenol oxidase activity in female Trichuris suis. *J Parasitol,* 79**,** 155-9.

FETTERER, R. H., MISKA, K. B., LILLEHOJ, H. & BARFIELD, R. C. 2007. Serine protease activity in developmental stages of Eimeria tenella. *The Journal of parasitology,* 93**,** 333-40.

FETTERER, R. H. & RHOADS, M. L. 1990. Tyrosine-derived cross-linking amino acids in the sheath of Haemonchus contortus infective larvae. *J Parasitol,* 76**,** 619-24.

FITZ-COY, S. 1992. Antigenic variation among strains of Eimeria maxima and E. tenella of the chicken. *Avian Dis,* 36**,** 40-43.

FITZ-COY, S. H. & EDGAR, S. A. 1992. Pathogenicity and control of Eimeria mitis infections in broiler chickens. *Avian Dis,* 36**,** 44-8.

FLEIGE, T., PFAFF, N., GROSS, U. & BOHNE, W. 2008. Localisation of gluconeogenesis and tricarboxylic acid (TCA)-cycle enzymes and first functional analysis of the TCA cycle in Toxoplasma gondii. *Int J Parasitol,* 38**,** 1121-32.

FLOWERS, S. A. 2011. *Mass spectrometric identification of structurally diverse Coccidian oocyst wall proteins.* PhD, University of Technology, Sydney.

FOERDER, C. A. & SHAPIRO, B. M. 1977a. Release of ovoperoxidase from sea urchin eggs hardens the fertililzation membrane with tyrosine crosslinks. *Proc. Natl. Acad. Sci. USA,* 74**,** 4214-4218.

FOERDER, C. A. & SHAPIRO, B. M. 1977b. Release of ovoperoxidase from sea urchin eggs hardens the fertilization membrane with tyrosine crosslinks. *Proc Natl Acad Sci U S A,* 74**,** 4214-8.

FOX, B. A., RISTUCCIA, J. G., GIGLEY, J. P. & BZIK, D. J. 2009. Efficient gene replacements in Toxoplasma gondii strains deficient for nonhomologous end joining. *Eukaryotic cell,* 8**,** 520-9.

FREEMAN, M. 2004. Proteolysis within the membrane: rhomboids revealed. *Nature Reviews. Molecular Cell Biology,* 5**,** 188-197.

FRIDOVICH, I. 1978. The biology of oxygen radicals. *Science,* 201**,** 875-80.

FRIED, M., MENCHER, D., SAR-SHALOM, O. & WALLACH, M. 1992. Developmental gene expression of a 230-kilodalton macrogamete-specific protein of the avian coccidial parasite, Eimeria maxima. *Mol Biochem Parasitol,* 51**,** 251-62.

FRITZ, H. M., BOWYER, P. W., BOGYO, M., CONRAD, P. A. & BOOTHROYD, J. C. 2012. Proteomic analysis of fractionated Toxoplasma oocysts reveals clues to their environmental resistance. *PLoS One,* 7**,** e29955.

FRY, M., HUDSON, A., RANDALL, A. & WILLIAMS, R. 1984. Potent and selective hydroxynaphthoquinone inhibitors of mitochondrial electron transport in Eimeria tenella (Apicomplexa: Coccidia). *Biochemical Pharmacology,* 33**,** 2115-22.

FRY, M. & WILLIAMS, R. 1984. Effects of decoquinate and clopidol on electron transport in mitochondria of Eimeria tenella (Apicomplexa: Coccidia). *Biochemical Pharmacology,* 33**,** 229-40.

FULLER, A. L. & MCDOUGALD, L. R. 1990. Reduction in cell entry of Eimeria tenella (Coccidia) sporozoites by protease inhibitors, and partial characterization of proteolytic activity associated with intact sporozoites and merozoites. *J Parasitol,* 76**,** 464-7.

GAJRIA, B., BAHL, A., BRESTELLI, J., DOMMER, J., FISCHER, S., GAO, X., HEIGES, M., IODICE, J., KISSINGER, J. C., MACKEY, A. J., PINNEY, D. F., ROOS, D. S., STOECKERT, C. J., JR., WANG, H. & BRUNK, B. P. 2008. ToxoDB: an integrated Toxoplasma gondii database resource. *Nucleic Acids Res,* 36**,** D553-6.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. (eds.) 2005. *Protein identification and analysis tools on the ExPASy server,* Totowa: Humana Press.

GEORGIEVA, N. & GADJEVA, V. 2005. Beneficial effects of isonicotinoylhydrazones on isoniazid-induced disruption of ecological oxidative balance. *Med Hypotheses,* 64**,** 662.

GEORGIEVA, N. V., KOINARSKI, V. & GADJEVA, V. 2006. Antioxidant status during the course of Eimeria tenella infection in broiler chickens. *Vet J,* 172**,** 488-92.

GOEL, V. K., LI, X., CHEN, H., LIU, S.-C., CHISHTI, A. H. & OH, S. S. 2003a. Band 3 is a host receptor binding merozoite surface protein 1 during the Plasmodium falciparum invasion of erythrocytes. *Proc. Natl. Acad. Sci. USA,* 100**,** 5164-9.

GOEL, V. K., LI, X., CHEN, H., LIU, S. C., CHISHTI, A. H. & OH, S. S. 2003b. Band 3 is a host receptor binding merozoite surface protein 1 during the Plasmodium falciparum invasion of erythrocytes. *Proc Natl Acad Sci U S A,* 100**,** 5164-9.

GONG, X., GUTALA, R. & JAISWAL, A. K. 2008. Quinone oxidoreductases and vitamin K metabolism. *Vitam Horm,* 78**,** 85-101.

GOODMAN, C. D. & MCFADDEN, G. I. 2007. Fatty acid biosynthesis as a drug target in apicomplexan parasites. *Curr Drug Targets,* 8**,** 15-30.

GUPTA, M. C., GRAHAM, P. L. & KRAMER, J. M. 1997. Characterization of alpha1(IV) collagen mutations in Caenorhabditis elegans and the effects of alpha1 and alpha2(IV) mutations on type IV collagen distribution. *J Cell Biol,* 137**,** 1185-96.

HACKETT, F., SAJID, M., WITHERS-MARTINEZ, C., GRAINER, M. & BLACKMAN, M. J. 1999a. PfSUB-2: a second subtilisin-like protein in *Plasmodium falciparum* merozoites. *Molecular and Biochemical Parasitology,* 103**,** 183-195.

HACKETT, F., SAJID, M., WITHERS-MARTINEZ, C., GRAINGER, M. & BLACKMAN, M. J. 1999b. PfSUB-2: a second subtilisin-like protein in Plasmodium falciparum merozoites. *Mol Biochem Parasitol,* 103**,** 183-95.

HALLIWELL, B. 1974. Superoxide dismutase, catalase and glutation peroxidase: Solutions to the problem of living with O2. *New Phytologist,* 73**,** 1075.

HAN, Y. S. & BARILLAS-MURY, C. 2002a. Implications of Time Bomb model of ookinete invasion of midgut cells. *Insect Biochemistry and Molecular Biology,* 32**,** 1311-1316.

HAN, Y. S. & BARILLAS-MURY, C. 2002b. Implications of Time Bomb model of ookinete invasion of midgut cells. *Insect Biochem Mol Biol,* 32**,** 1311-6.

HANIG, S., ENTZEROTHB, R. & KURTHA, M. 2012. Chimeric fluorescent reporter as a tool for generation of transgenicEimeria (Apicomplexa, Coccidia) strains with stage specific reporter gene expression. *Parasitology International,* 61**,** 391-8.

HAO, L. L., LIU, X. Y., ZHOU, X. Y., LI, J. D. & SUO, X. 2007. Transient transfection of Eimeria tenella using yellow or red fluorescent protein as a marker. *Molecular and Biochemical Parasitology,* 153**,** 213-215.

HARRIS, P. K., YEOH, S., DLUZEWSKI, A. R., O'DONNELL, R. A., WITHERS-MARTINEZ, C., HACKETT, F., BANNISTER, L. H., MITCHELL, G. H. & BLACKMAN, M. J. 2005. Molecular identification of a malaria merozoite surface sheddase. *PLoS Pathog,* 1**,** 241-51.

HASLAM, N. J. & GIBSON, T. J. 2010. EpiC: an open resource for exploring epitopes to aid antibody-based experiments. *J Proteome Res,* 9**,** 3759-63.

HATEFI, Y. 1985. The mitochondrial electron transport and oxidative phosphorylation system. *Annu Rev Biochem,* 54**,** 1015-69.

HEHL, A. B., LEKUTIS, C., GRIGG, M. E., BRADLEY, P. J., DUBREMETZ, J. F., ORTEGA-BARRIA, E. & BOOTHROYD, J. C. 2000. Toxoplasma gondii homologue of plasmodium apical membrane antigen 1 is involved in invasion of host cells. *Infect Immun,* 68**,** 7078-86.

HEIN, H. 1971. Pathogenic effects of Eimeria necatrix in young chickens. *Experimental parasitology,* 30**,** 321-30.

HEINECKE, J. W. & SHAPIRO, B. M. 1992. The respiratory burst oxidase of fertilization. A physiological target for regulation by protein kinase C. *J Biol Chem,* 267**,** 7959-62.

HINSHAW, W. 1952. Coccidiosis vaccine. *Vet Med,* 47**,** 429-430.

HOPKINS, T. L. 1992. Insect Cuticle Sclerotization. *Annual Reviews in Entomology,* 37**,** 273-302.

HOWELL, S. A., WITHERS-MARTINEZ, C., KOCKEN, C. H., THOMAS, A. W. & BLACKMAN, M. J. 2001a. Proteolytic Processing and Primary Structure of *Plasmodium falciparum* Apical Membrane Antigen-1. *The Journal of Biological Chemistry,* 276**,** 31311-31320.

HOWELL, S. A., WITHERS-MARTINEZ, C., KOCKEN, C. H., THOMAS, A. W. & BLACKMAN, M. J. 2001b. Proteolytic processing and primary structure of Plasmodium falciparum apical membrane antigen-1. *J Biol Chem,* 276**,** 31311-20.

HU, J., QIN, H., LI, C., SHARMA, M., CROSS, T. A. & GAO, F. P. 2007. Structural biology of transmembrane domains: efficient production and characterization of transmembrane peptides by NMR. *Protein Sci,* 16**,** 2153-65.

HUYNH, M. H. & CARRUTHERS, V. B. 2009. Tagging of endogenous genes in a Toxoplasma gondii strain lacking Ku80. *Eukaryotic cell,* 8**,** 530-9.

IKEMURA, H. & INOUYE, M. 1988. *In vitro* processing of pro-subtilisin produced in *Escherichia coli*. *The Journal of Biological Chemistry,* 263**,** 12954-12958.

JAMESON, B. A. & WOLF, H. 1988. The antigenic index: a novel algorithm for predicting antigenic determinants. *Comput Appl Biosci,* 4**,** 181-6.

JANKIEVICZ, H. & SCHOFIELD, R. 1934. The administration of heated oocysts of *Eimeria tenella* as a means of establishing resistance and immunity to cecal coccidiosis. *J Am Vet Med Assoc,* 37**,** 507-526.

JEFFERS, T. K. 1975. Attenuation of Eimeria tenella through selection for precociousness. *J Parasitol,* 61**,** 1083-90.

JENKINS, M. C. 1998. Progress on developing a recombinant coccidiosis vaccine. *Int J Parasitol,* 28**,** 1111-9.

JENKINS, M. C. 2001. Advances and prospects for subunit vaccines against protozoa of veterinary importance. *Vet Parasitol,* 101**,** 291-310.

JEURISSEN, S. H., JANSE, E. M., VERMEULEN, A. N. & VERVELDE, L. 1996. Eimeria tenella infections in chickens: aspects of host-parasite: interaction. *Veterinary immunology and immunopathology,* 54**,** 231-8.

JOHNSON, J., REID, W. M. & JEFFERS, T. K. 1979. Practical immunization of chickens against coccidiosis using an attenuated strain of Eimeria tenella. *Poult Sci,* 58**,** 37-41.

JOHNSON, W. 1932. Immunity to coccidiosis in chickens, produced by innoculation through the ration. *J Parasitol,* 19**,** 106-161.

JOHNSTONE, I. L. 2000a. Cuticle Collagen Genes expression in *Caenorhabditis elegans*. *Trends in Genetics,* 16**,** 21-27.

JOHNSTONE, I. L. 2000b. Cuticle collagen genes. Expression in Caenorhabditis elegans. *Trends Genet,* 16**,** 21-7.

JORGENSEN, W. K., ANDERSON, G. R., JESTON, P. J., BLIGHT, G. W. & MOLLOY, J. B. 2006a. Selection and characterisation of two attenuated vaccine lines of Eimeria tenella in Australia. *Aust Vet J,* 84**,** 89-94.

JORGENSEN, W. K., RICHARDS, G. & ANDERSON, G. R. 2006b. Evaluation of the most effective measures of infectivity of a precocious line of Eimeria acervulina in poultry. *Aust Vet J,* 84**,** 428-30.

JOYNER, L. P. & NORTON, C. C. 1973. The immunity arising from continuous low-level infection with Eimeria tenella. *Parasitology,* 67**,** 333-40.

KARIM, M. J., BASAK, S. C. & TREES, A. J. 1996. Characterization and immunoprotective properties of a monoclonal antibody against the major oocyst wall protein of Eimeria tenella. *Infection and Immunity,* 64**,** 1227-32.

KATRIB, M., IKIN, R. J., BROSSIER, F., ROBINSON, M., SLAPETOVA, I., SHARMAN, P. A., WALKER, R. A., BELLI, S. I., TOMLEY, F. M. & SMITH, N. C. 2012. Stage-specific expression of protease genes in the apicomplexan parasite, Eimeria tenella. *BMC Genomics,* 13**,** 685.

KAWAHARA, K., MOGI, T., TANAKA, T. Q., HATA, M., MIYOSHI, H. & KITA, K. 2009. Mitochondrial dehydrogenases in the aerobic respiratory chain of the rodent malaria parasite Plasmodium yoelii yoelii. *J Biochem,* 145**,** 229-37.

KELLEHER, M. & TOMLEY, F. M. 1998. Transient expression of beta-galactosidase in differentiating sporozoites of Eimeria tenella. *Mol Biochem Parasitol,* 97**,** 21-31.

KIELISZEWSKI, M., DE ZACKS, R., LEYKAM, J. F. & LAMPORT, D. T. 1992. A repetitive proline-rich protein from the gymnosperm douglas fir is a hydroxyproline-rich glycoprotein. *Plant Physiol,* 98**,** 919-26.

KIELISZEWSKI, M. J. & SHPAK, E. 2001. Synthetic genes for the elucidation of glycosylation codes for arabinogalactan-proteins and other hydroxyproline-rich glycoproteins. *Cell Mol Life Sci,* 58**,** 1386-98.

KIM, J. W. & DANG, C. V. 2005. Multifaceted roles of glycolytic enzymes. *Trends Biochem Sci,* 30**,** 142-50.

KIM, K. 2004. Role of proteases in host cell invasion by *Toxoplasma gondii* and other apicomplexa. *Acta Tropica,* 91**,** 69-81.

KOGUT, M. H., GORE, T. C. & LONG, P. L. 1983. Serial passage of Eimeria tenella and E. necatrix in turkey embryos. *Parasitology,* 86 (Pt 2)**,** 199-209.

KOLASKAR, A. S. & TONGAONKAR, P. C. 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett,* 276**,** 172-4.

KORICH, D. G., MEAD, J. R., MADORE, M. S., SINCLAIR, N. A. & STERLING, C. R. 1990. Effects of ozone, chlorine dioxide, chlorine, and monochloramine on Cryptosporidium parvum oocyst viability. *Applied & Environmental Microbiology,* 56**,** 1423-8.

KOUWENHOVEN, B. & KUIL, H. 1976. Demonstration of circulating antibodies to *Eimeria tenella* by the indirect immunofluorescent antibody test using sporozoites and second-stage schizonts as antigen. *Vet Parasitol,* 2.

KRUCKEN, J., HOSSE, R. J., MOUAFO, A. N., ENTZEROTH, R., BIERBAUM, S., MARINOVSKI, P., HAIN, K., GREIF, G. & WUNDERLICH, F. 2008. Excystation of Eimeria tenella sporozoites impaired by antibody recognizing gametocyte/oocyst antigens GAM22 and GAM56. *Eukaryot Cell,* 7**,** 202-11.

KUTICIC, V. & WIKERHAUSER, T. 1996. Studies of the effect of various treatments on the viability of Toxoplasma gondii tissue cysts and oocysts. *Current Topics in Microbiology & Immunology,* 219**,** 261-5.

KYTE, J. & DOOLITTLE, R. F. 1982. A simple method for displaying the hydropathic character of a protein. *J Mol Biol,* 157**,** 105-32.

LAMBETH, J. D. 2004. NOX enzymes and the biology of reactive oxygen. *Nat Rev Immunol,* 4**,** 181-9.

LANGKJAER, M. & ROEPSTORFF, A. 2008. Survival of Isospora suis oocysts under controlled environmental conditions. *Veterinary Parasitology,* 152**,** 186-93.

LATEEF, S. S., GUPTA, S., JAYATHILAKA, L. P., KRISHNANCHETTIAR, S., HUANG, J. S. & LEE, B. S. 2007. An improved protocol for coupling synthetic peptides to carrier proteins for antibody production using DMF to solubilize peptides. *J Biomol Tech,* 18**,** 173-6.

LAXER, M. A., HEALEY, M. C. & YOUSSEF, N. N. 1987. Production of monoclonal antibodies specific for Eimeria tenella microgametocytes. *J Parasitol,* 73**,** 611-6.

LE ROCH, K. G., JOHNSON, J. R., FLORENS, L., ZHOU, Y., SANTROSYAN, A., GRAINGER, M., YAN, S. F., WILLIAMSON, K. C., HOLDER, A. A., CARUCCI, D. J., YATES, J. R., 3RD & WINZELER, E. A. 2004. Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. *Genome research,* 14**,** 2308-18.

LEVY, A. D., YANG, J. & KRAMER, J. M. 1993. Molecular and genetic analyses of the Caenorhabditis elegans dpy-2 and dpy-10 collagen genes: a variety of molecular alterations affect organismal morphology. *Mol Biol Cell,* 4**,** 803-17.

LI, J., ZHANG, X., LIU, Q., YIN, J. & YANG, J. 2006. Eimeria tenella: cloning of a novel Eimeria tenella cDNA encoding a protein related to rhomboid family from F2 hybrid strain. *Exp Parasitol,* 113**,** 215-20.

LI, X., CHEN, H., OO, T. H., DALY, T. M., BERGMAN, L. W., LIU, S.-C., CHISHTI, A. H. & OH, S. S. 2004. A co-ligand complex anchors Plasmodium falciparum merozoites to the erythrocyte invasion receptor band 3. *The Journal of Biological Chemistry,* 279**,** 5765-71.

LI, Y., HU, Z., JORDAN, F. & INOUYE, M. 1995. Functional Analysis of the Propeptide of Subtilisin E as an Intramolecular Chaperone for Protein Folding. *The Journal of Biological Chemistry,* 270**,** 25127-25132.

LIGE, B., JAYABALASINGHAM, B., ZHANG, H., PYPAERT, M. & COPPENS, I. 2009. Role of an ancestral d-bifunctional protein containing two sterol-carrier protein-2 domains in lipid uptake and trafficking in Toxoplasma. *Mol Biol Cell,* 20**,** 658-72.

LILLEHOJ, H. 2005. Immune Response to Coccidia. *The IXth International Coccidiosis Conference.* Foz do Iguassu, Parana, Brazil.

LIN, S. S., GROSS, U. & BOHNE, W. 2009. Type II NADH dehydrogenase inhibitor 1-hydroxy-2-dodecyl-4(1H)quinolone leads to collapse of mitochondrial inner-membrane potential and ATP depletion in Toxoplasma gondii. *Eukaryot Cell,* 8**,** 877-87.

LINDQUIST, H. D., BENNETT, J. W., HESTER, J. D., WARE, M. W., DUBEY, J. P. & EVERSON, W. V. 2003. Autofluorescence of Toxoplasma gondii and related coccidian oocysts. *J Parasitol,* 89**,** 865-7.

LINDSAY, D. S., MITCHELL, S. M., VIANNA, M. C. & DUBEY, J. P. 2004. Sarcocystis neurona (Protozoa: Apicomplexa): description of oocysts, sporocysts, sporozoites, excystation, and early development. *Journal of Parasitology,* 90**,** 461-5.

LINDSAY, D. S., UPTON, S. J. & DUBEY, J. P. 1999a. A structural study of the Neospora caninum oocyst. *Int J Parasitol,* 29**,** 1521-3.

LINDSAY, D. S., UPTON, S. J. & DUBEY, J. P. 1999b. A structural study of the Neospora caninum oocyst. *International Journal for Parasitology,* 29**,** 1521-3.

LINDSKOG, M., ROCKBERG, J., UHLEN, M. & STERKY, F. 2005. Selection of protein epitopes for antibody production. *Biotechniques,* 38**,** 723-7.

LING, I. T. & WILSON, R. J. 1988. Glucose-6-phosphate dehydrogenase activity of the malaria parasite Plasmodium falciparum. *Mol Biochem Parasitol,* 31**,** 47-56.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

LONG, P. L. 1929. A study of Eimeria maxima Tyzzer, a coccidium of the fowl (Gallus gallus). *Ann Trop Med Parasitol.* **,** 325-333.

LONG, P. L. 1972. Eimeria tenella: reproduction, pathogenicity and immunogenicity of a strain maintained in chick embryos by serial passage. *J Comp Pathol,* 82**,** 429-37.

LONG, P. L. 1974. Further studies on the pathogenicity and immunogenicity of an embryo-adapted strain of Eimeria tenella. *Avian Pathol,* 3**,** 255-68.

LONG, P. L., JOHNSON, J., MCKENZIE, M. E., PERRY, E., CRANE, M. S. & MURRAY, P. K. 1986. Immunisation of young broiler chickens with low level infections of Eimeria tenella, E. acervulina or E. maxima. *Avian Pathol,* 15**,** 271-8.

LONG, P. L. & MILLARD, B. J. 1968. Eimeria: effect of meticlorpindol and methyl benzoquate on endogenous stages in the chicken. *Exp Parasitol,* 23**,** 331-8.

LOUIE, K. & CONRAD, P. A. 1999. Characterization of a cDNA encoding a subtilisin-like serine protease (NC-p65) of Neospora caninum. *Mol Biochem Parasitol,* 103**,** 211-23.

LOUIE, K. & CONRAD, P. A. 1999 Characterization of a cDNA encoding a subtilisin-like serine protease (NC-p65) of *Neospora caninum*. *Molecular and Biochemical Parasitology,* 103**,** 211-223.

LOWRY, O. H. & PASSONNEAU, J. V. 1964. The Relationships between Substrates and Enzymes of Glycolysis in Brain. *J Biol Chem,* 239**,** 31-42.

MADDEN, T. 2002. The BLAST Sequence Analysis Tool. *In:* MCENTYRE, J. & OSTELL, J. (eds.) *The NCBI Handbook.* Bethesda (MD): National Center for Biotechnology Information.

MAI, K., SHARMAN, P., WALKER, R., KATRIB, M., DESOUZA, D., MCCONVILLE, M., WALLACH, M., BELLI, S., FERGUSON, D. & SMITH, N. 2008 The molecular basis for oocyst wall formation in the apicomplexan parasite, Eimeria maxima. *Memorias do Instituto Oswaldo Cruz***,** (In press).

MAI, K., SMITH, N. C., FENG, Z. P., KATRIB, M., SLAPETA, J., SLAPETOVA, I., WALLACH, M. G., LUXFORD, C., DAVIES, M. J., ZHANG, X., NORTON, R. S. & BELLI, S. I. 2011. Peroxidase catalysed cross-linking of an intrinsically unstructured protein via dityrosine bonds in the oocyst wall of the apicomplexan parasite, Eimeria maxima. *Int J Parasitol,* 41**,** 1157-64.

MALENCIK, D. A. & ANDERSON, S. R. 1996. Dityrosine formation in calmodulin: cross-linking and polymerization catalyzed by Arthromyces peroxidase. *Biochemistry,* 35**,** 4375-86.

MARTI, L., ABELLA, A., CARPENE, C., PALACIN, M., TESTAR, X. & ZORZANO, A. 2001. Combined treatment with benzylamine and low dosages of vanadate enhances glucose tolerance and reduces hyperglycemia in streptozotocin-induced diabetic rats. *Diabetes,* 50**,** 2061-8.

MARTI, L., MORIN, N., ENRIQUE-TARANCON, G., PREVOT, D., LAFONTAN, M., TESTAR, X., ZORZANO, A. & CARPENE, C. 1998. Tyramine and vanadate synergistically stimulate glucose transport in rat adipocytes by amine oxidase-dependent generation of hydrogen peroxide. *J Pharmacol Exp Ther,* 285**,** 342-9.

MARTIN, A. G., DANFORTH, H. D., BARTA, J. R. & FERNANDO, M. A. 1997. Analysis of immunological cross-protection and sensitivities to anticoccidial drugs among five geographical and temporal strains of Eimeria maxima. *Int J Parasitol,* 27**,** 527-33.

MCDONALD, V., ROSE, M. E. & JEFFERS, T. K. 1986. Eimeria tenella: immunogenicity of the first generation of schizogony. *Parasitology,* 93 (Pt 1)**,** 1-7.

MCDONALD, V. & SHIRLEY, M. W. 2009. Past and future: vaccination against Eimeria. *Parasitology,* 136**,** 1477-89.

MCDONALD, V., WISHER, M. H., ROSE, M. E. & JEFFERS, T. K. 1988. Eimeria tenella: immunological diversity between asexual generations. *Parasite Immunol,* 10**,** 649-60.

MICHAEL, A. 2003. The practical use of a maternal vaccine against coccidiosis. *World Poult,* 19**,** 24-26.

MICHALSKI, W. P., CROOKS, J. K. & PROWSE, S. J. 1994. Purification and characterisation of a serine-type protease from *Eimeria tenella* oocysts. *International Journal for Parasitology,* 24**,** 189-195.

MICHALSKI, W. P. & PROWSE, S. J. 1991. Superoxide dismutases in Eimeria tenella. *Mol Biochem Parasitol,* 47**,** 189-95.

MILLER, S. A., BINDER, E. M., BLACKMAN, M. J. & CARRUTHERS, V. B. 2001a. A Conserved Subtilisin-like Protein TgSUB1 in Microneme Organelles of *Toxoplasma gondii*. *The Journal of Biological Chemistry,* 276**,** 45341-45348.

MILLER, S. A., BINDER, E. M., BLACKMAN, M. J., CARRUTHERS, V. B. & KIM, K. 2001b. A conserved subtilisin-like protein TgSUB1 in microneme organelles of Toxoplasma gondii. *J Biol Chem,* 276**,** 45341-8.

MILLER, S. A., THATHY, V., AJIOKA, J. W., BLACKMAN, M. J. & KIM, K. 2003a. TgSUB2 is a Toxoplasma gondii rhoptry organelle processing proteinase. *Mol Microbiol,* 49**,** 883-94.

MILLER, S. A., THATHY, V., AJIOKA, J. W., BLACKMAN, M. J. & KIM, K. 2003b. TgSUB2 is a *Toxoplasma gondii* rhoptry organelle processing proteinase. *Molecular Microbiology,* 49**,** 883-94.

MISKA, K. B., FETTERER, R. H. & BARFIELD, R. C. 2004. Analysis of Transcripts Expressed by Eimeria tenella Oocysts Using Subtractive Hybridisation Methods. *Journal of Parasitology,* 90**,** 1245-1252.

MONNÉ, L. & G, H. 1954. On the properties of the shells of coccidian oocysts. *Arkiv för Zoologi,* 7**,** 251-256.

MONTERO, E., GONZALEZ, L. M., RODRIGUEZ, M., OKSOV, Y. & BLACKMAN, M. J. 2006a. A Conserved Subtilisin Protease Identified in *Babesia diverens* Merozoites. *The Journal of Biological Chemistry,* 281**,** 35717-35726.

MONTERO, E., GONZALEZ, L. M., RODRIGUEZ, M., OKSOV, Y., BLACKMAN, M. J. & LOBO, C. A. 2006b. A conserved subtilisin protease identified in Babesia divergens merozoites. *J Biol Chem,* 281**,** 35717-26.

MOUAFO, A. N., WECK-HEIMANN, A., DUBREMETZ, J. F. & ENTZEROTH, R. 2002a. Monoclonal antibodies specific for the two types of wall-forming bodies of Eimeria tenella macrogametes (Coccidia, Apicomplexa). *Parasitology Research,* 88**,** 217-24.

MOUAFO, A. N., WECK-HEIMANN, A., DUBREMETZ, J. F. & ENTZEROTH, R. 2002b. Monoclonal antibodies specific for the two types of wall-forming bodies of Eimeria tenella macrogametes (Coccidia, Apicomplexa). *Parasitol Res,* 88**,** 217-24.

MULLER, S. 2004. Redox and antioxidant systems of the malaria parasite Plasmodium falciparum. *Mol Microbiol,* 53**,** 1291-305.

NEBL, T., PRIETO, J. H., KAPP, E., SMITH, B. J., WILLIAMS, M. J., YATES, J. R., 3RD, COWMAN, A. F. & TONKIN, C. J. 2011. Quantitative in vivo analyses reveal calcium-dependent phosphorylation sites and identifies a novel component of the Toxoplasma invasion motor complex. *PLoS Pathog,* 7**,** e1002222.

NGOKA, L. C. 2008. Dramatic down-regulation of oxidoreductases in human hepatocellular carcinoma hepG2 cells: proteomics and gene ontology unveiling new frontiers in cancer enzymology. *Proteome Sci,* 6**,** 29.

NICKEL, C., RAHLFS, S., DEPONTE, M., KONCAREVIC, S. & BECKER, K. 2006. Thioredoxin networks in the malarial parasite Plasmodium falciparum. *Antioxid Redox Signal,* 8**,** 1227-39.

NOVELLI, J., AHMED, S. & HODGKIN, J. 2004. Gene interactions in Caenorhabditis elegans define DPY-31 as a candidate procollagen C-proteinase and SQT-3/ROL-4 as its predicted major target. *Genetics,* 168**,** 1259-73.

OPITZ, C., CRISTINA, M. D., REISS, M., RUPPERT, T., CRISANTI, A. & SOLDATI, D. 2002a. Intramembrane cleavage of microneme proteins at the surface of the apicomplexan parasite Toxoplasma gondii. *The EMBO Journal,* 21**,** 1577-1585.

OPITZ, C., DI CRISTINA, M., REISS, M., RUPPERT, T., CRISANTI, A. & SOLDATI, D. 2002b. Intramembrane cleavage of microneme proteins at the surface of the apicomplexan parasite Toxoplasma gondii. *EMBO J,* 21**,** 1577-85.

OVINGTON, K. S. & SMITH, N. C. 1992. Cytokines, free radicals and resistance to Eimeria. *Parasitol Today,* 8**,** 422-6.

PAGE, A. & WINTER, A. 2003a. Enzymes involved in the biogenesis of the nematode cuticle. *Advances in Parasitology,* 53**,** 85-148.

PAGE, A. P. & WINTER, A. D. 2003b. Enzymes involved in the biogenesis of the nematode cuticle. *Adv Parasitol,* 53**,** 85-148.

PARISE, G. & BAZZICALUPO, P. 1997a. Assembly of nematode cuticle: role of hydrophobic interactions in CUT-2 cross-linking. *Biochim Biophys Acta,* 1337**,** 295-301.

PARISE, G. & BAZZICALUPO, P. 1997b. Assmebly of nematode cuticle: role of hydrophobic interactions in CUT-2 cross-linking. *Biochemica et Biophysica Acta,* 1337**,** 295-301.

PATON, M. G., BARKER, G. C., MATSUOKA, H., RAMESAR, J., JANSE, C. J., WATERS, A. P. & SINDEN, R. E. 1993. Structure and expression of a post-transcriptionally regulated malaria gene encoding a surface protein from the sexual stages of Plasmodium berghei. *Mol Biochem Parasitol,* 59**,** 263-75.

PFAFFL, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.,* 29**,** e45.

PINO, P., FOTH, B. J., KWOK, L. Y., SHEINER, L., SCHEPERS, R., SOLDATI, T. & SOLDATI-FAVRE, D. 2007. Dual targeting of antioxidant and metabolic enzymes to the mitochondrion and the apicoplast of Toxoplasma gondii. *PLoS Pathog,* 3**,** e115.

PITT-RIVERS, R. & IMPIOMBATO, F. S. 1968. The binding of sodium dodecyl sulphate to various proteins. *Biochem J,* 109**,** 825-30.

PITTILO, R. M. & BALL, S. J. 1980. The ultrastructural development of the oocyst wall of Eimeria maxima. *Parasitology,* 81**,** 115-22.

POLONAIS, V. & SOLDATI-FAVRE, D. 2010. Versatility in the acquisition of energy and carbon sources by the Apicomplexa. *Biol Cell,* 102**,** 435-45.

POOLE, C. B., JIN, J. & MCREYNOLDS, L. A. 2003. Cloning and Biochemical Charaterisation of Blisterase, a Subtilisin-like Convertase from the Filarial Parasite, *Onchocerca volulus*. *The Journal of Biological Chemistry,* 278**,** 36183-36190.

POSSENTI, A., CHERCHI, S., BERTUCCINI, L., POZIO, E., DUBEY, J. P. & SPANO, F. 2010. Molecular characterisation of a novel family of cysteine-rich proteins of Toxoplasma gondii and ultrastructural evidence of oocyst wall localisation. *Int J Parasitol*.

POUT, D. D. 1967. Villous atrophy and coccidiosis. *Nature,* 213**,** 306-7.

PUGATSCH, T., MENCHER, D. & WALLACH, M. 1989. Eimeria maxima: isolation of gametocytes and their immunogenicity in mice, rabbits, and chickens. *Exp Parasitol,* 68**,** 127-34.

QUEVILLON, E., SILVENTOINEN, V., PILLAI, S., HARTE, N., MULDER, N., APWEILER, R. & LOPEZ, R. 2005. InterProScan: protein domains identifier. *Nucleic Acids Res,* 33**,** W116-20.

QUINN, M. T. & GAUSS, K. A. 2004. Structure and regulation of the neutrophil respiratory burst oxidase: comparison with nonphagocyte oxidases. *J Leukoc Biol,* 76**,** 760-81.

RABENAU, K. E., SOHRABI, A., TRIPATHY, A., REITTER, C., AJIOKA, J. W., TOMLEY, F. M. & CARRUTHERS, V. B. 2001a. TgM2AP participates in Toxoplasma gondii invasion of host cells and is tightly associated with the adhesive protein TgMIC2. *Mol Microbiol,* 41**,** 537-47.

RABENAU, K. E., SOHRABI, A., TRIPATHY, A., REITTER, C., AJIOKA, J. W., TOMLEY, F. M. & CARRUTHERS, V. B. 2001b. TgM2AP participates in Toxoplasma gondii invasion of host cells and is tightly associated with the adhesive protein TgMIC2. *Molecular Microbiology,* 41**,** 1-12.

RANUCCI, L., MULLER, H. M., LA ROSA, G., RECKMANN, I., MORALES, M. A., SPANO, F., POZIO, E. & CRISANTI, A. 1993. Characterization and immunolocalization of a Cryptosporidium protein containing repeated amino acid motifs. *Infect Immun,* 61**,** 2347-56.

RATH, A., GLIBOWICKA, M., NADEAU, V. G., CHEN, G. & DEBER, C. M. 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Natl Acad Sci U S A,* 106**,** 1760-5.

RAWLINGS, N. D. & BARRETT, A. J. 1994. Families of Serine Peptidases. *Methods in Enzymology,* 244**,** 19-61.

REA, G., LAURENZI, M., TRANQUILLI, E., D'OVIDIO, R., FEDERICO, R. & ANGELINI, R. 1998. Developmentally and wound-regulated expression of the gene encoding a cell wall copper amine oxidase in chickpea seedlings. *FEBS Lett,* 437**,** 177-82.

REID, W. M. 1990. History of avian medicine in the United States. X. Control of coccidiosis. *Avian Dis,* 34**,** 509-25.

RICE, P., LONGDEN, I. & BLEASBY, A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet,* 16**,** 276-7.

RICHARDS, G. & WOODS, W. 2001. The use of four species live Eimeria vaccines in broilers and layers in Australia. *VIIIth International Coccidiosis Conference.* Palm Cove, Sydney, Australia.

RIEUX, A., GRAS, S., LECAILLE, F., NIEPCERON, A., KATRIB, M., SMITH, N. C., LALMANACH, G. & BROSSIER, F. 2012. Eimeripain, a cathepsin B-like cysteine protease, expressed throughout sporulation of the apicomplexan parasite Eimeria tenella. *PloS one,* 7**,** e31914.

RIGLAR, D. T., ROGERS, K. L., HANSSEN, E., TURNBULL, L., BULLEN, H. E., CHARNAUD, S. C., PRZYBORSKI, J., GILSON, P. R., WHITCHURCH, C. B., CRABB, B. S., BAUM, J. & COWMAN, A. F. 2013. Spatial association with PTEX complexes defines regions for effector export into Plasmodium falciparum-infected erythrocytes. *Nature communications,* 4**,** 1415.

ROMMEREIM, L. M., HORTUA TRIANA, M. A., FALLA, A., SANDERS, K. L., GUEVARA, R. B., BZIK, D. J. & FOX, B. A. 2013. Genetic manipulation in Deltaku80 strains for functional genomic analysis of Toxoplasma gondii. *Journal of visualized experiments : JoVE***,** e50598.

ROSE, M. E. 1967. Immunity to Eimeria tenella and Eimeria necatrix infections in the fowl. I. Influence of the site of infection and the stage of the parasite. II. Cross-protection. *Parasitology,* 57**,** 567-83.

ROSE, M. E. 1971. Immunity to coccidiosis: protective effect of transferred serum in Eimeria maxima infections. *Parasitology,* 62**,** 11-25.

ROSE, M. E. 1972. Immunity to coccidiosis: maternal transfer in Eimeria maxima infections. *Parasitology,* 65**,** 273-82.

ROSE, M. E. 1974. The early development of immunity to Eimeria maxima in comparison with that to Eimeria tenella. *Parasitology,* 68**,** 35-45.

ROSE, M. E. 1987. Immunity to Eimeria infections. *Vet Immunol Immunopathol,* 17**,** 333-43.

ROSE, M. E. & HESKETH, P. 1976. Immunity to coccidiosis: stages of the life-cycle of Eimeria maxima which induce, and are affected by, the response of the host. *Parasitology,* 73**,** 25-57.

ROZEN, S. & SKALETSKY, H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol.,* 132**,** 365-86.

RYLEY, J. F. 1969. Ultrastructural studies on the sporozoite of Eimeria tenella. *Parasitology,* 59**,** 67-72.

SAJID, M., WITHERS-MARTINEZ, C. & BLACKMAN, M. J. 2000a. Maturation and Specificity of *Plasmodium falciparum* Subtilisin-like Protease-1, a Malaria Merozoite Subtilisin-like Serine Protease. *The Journal of Biological Chemistry,* 275**,** 631-641.

SAJID, M., WITHERS-MARTINEZ, C. & BLACKMAN, M. J. 2000b. Maturation and specificity of Plasmodium falciparum subtilisin-like protease-1, a malaria merozoite subtilisin-like serine protease. *J Biol Chem,* 275**,** 631-41.

SCHAAP, D., ARTS, G., VAN POPPEL, N. F. & VERMEULEN, A. N. 2005. De novo ribosome biosynthesis is transcriptionally regulated in Eimeria tenella, dependent on its life cycle stage. *Mol Biochem Parasitol. ,* 139**,** 239-48.

SCHNITZLER, B. B. & SHIRLEY, M. W. 1999. Immunological aspects of infections with *Eimeria maxima*: a short review. *Avian Pathology,* 28**,** 537-543.

SCHOLTYSECK, E., MEHLHORN, H. & HAMMOND, D. M. 1971. Fine structure of macrogametes and oocysts of Coccidia and related organisms. *Z Parasitenkd.,* 37**,** 1-43.

SHANER, N. C., CAMPBELL, R. E., STEINBACH, P. A., GIEPMANS, B. N., PALMER, A. E. & TSIEN, R. Y. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol,* 22**,** 1567-72.

SHI, T. Y., LIU, X. Y., HAO, L. L., LI, J. D., GH, A. N., ABDILLE, M. H. & SUO, X. 2008. Transfected Eimeria tenella could complete its endogenous development in vitro. *J Parasitol,* 94**,** 978-80.

SHIRLEY, M. W. 1980. Eimeria necatrix: the development and characteristics of an egg-adapted (attenuated) line. *Parasitology,* 81**,** 525-35.

SHIRLEY, M. W. 1989. Development of a live attenuated vaccine against coccidiosis of poultry. *Parasite Immunol,* 11**,** 117-24.

SHIRLEY, M. W. & BEDRNIK, P. 1997. Live attenuated vaccines against avian coccidiosis: Success with precocious and egg-adapted lines of Eimeria. *Parasitol Today,* 13**,** 481-4.

SHIRLEY, M. W., BUSHELL, A. C., BUSHELL, J. E., MCDONALD, V. & ROBERTS, B. 1995. A live attenuated vaccine for the control of avian coccidiosis: trials in broiler breeders and replacement layer flocks in the United Kingdom. *Vet Rec,* 137**,** 453-7.

SHIRLEY, M. W., MCDONALD, V. & BALLINGALL, S. 1981. Eimeria spp. from the chicken: from merozoites to oocysts in embryonated eggs. *Parasitology,* 83**,** 259-67.

SHIRLEY, M. W. & MILLARD, B. J. 1986. Studies on the immunogenicity of seven attenuated lines of Eimeria given as a mixture to chickens. *Avian Pathol,* 15**,** 629-38.

SHIRLEY, M. W., SMITH, A. L. & BLAKE, D. P. 2007. Challenges in the successful control of the avian coccidia. *Vaccine,* 25**,** 5540-7.

SHIRLEY, M. W., SMITH, A. L. & TOMLEY, F. M. 2005. The biology of avian Eimeria with an emphasis on their control by vaccination. *Adv Parasitol,* 60**,** 285-330.

SIEZEN, R., VOS, W. D., LEUNISSEN, J. & DIJSTRA, B. 1991. Homology Modelling and Protein Engineering Strategy of Subtilases, the family of Subtilisin-like Serine Proteinases. *Protein Engnineering,* 4**,** 719-737.

SIEZEN, R. J. & LEUNISSEN, J. A. 1997a. Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci,* 6**,** 501-23.

SIEZEN, R. J. & LEUNISSEN, J. A. M. 1997b. Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Science,* 6**,** 501- 523.

SIEZEN, R. J., RENCKENS, B. & BOEKHORST, J. 2007a. Evolution of prokaryotic subtilases: genome-wide analysis reveals novel subfamilies with different catalytic residues. *Proteins,* 67**,** 681-94.

SIEZEN, R. J., RENCKENS, B. & BOEKHORST, J. 2007b. Evolution of Prokaryotic Subtilases: Genome-Wide Analysis Reveals Novel Subfamilies With Different Catalytic Residues. *Proteins: Structure, Function and Bioinformatics,* 67**,** 681-694.

SIGRIST, C. J., CERUTTI, L., HULO, N., GATTIKER, A., FALQUET, L., PAGNI, M., BAIROCH, A. & BUCHER, P. 2002. PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform,* 3**,** 265-74.

SMAIL, E. H., BRIZA, P., PANAGOS, A. & BERENFELD, L. 1995. *Candida albicans* Cell Walls Contain the Flourescent Cross-Linking Amino Acid Dityrosine. *Infection and Immunity,* 63**,** 4078-4083.

SMITH, N. C., WALLACH, M., MILLER, C. M., BRAUN, R. & ECKERT, J. 1994a. Maternal transmission of immunity to Eimeria maxima: western blot analysis of protective antibodies induced by infection. *Infect Immun,* 62**,** 4811-7.

SMITH, N. C., WALLACH, M., MILLER, C. M., MORGENSTERN, R., BRAUN, R. & ECKERT, J. 1994b. Maternal transmission of immunity to Eimeria maxima: enzyme-linked immunosorbent assay analysis of protective antibodies induced by infection. *Infect Immun,* 62**,** 1348-57.

SMITH, N. C., WALLACH, M., PETRACCA, M., BRAUN, R. & ECKERT, J. 1994c. Maternal transfer of antibodies induced by infection with Eimeria maxima partially protects chickens against challenge with Eimeria tenella. *Parasitology,* 109 (Pt 5)**,** 551-7.

SPANO, F., PURI, C., RANUCCI, L., PUTIGNANI, L. & CRISANTI, A. 1997. Cloning of the entire COWP gene of Cryptosporidium parvum and ultrastructural localization of the protein during sexual parasite development. *Parasitology,* 114 (Pt 5)**,** 427-37.

SPEER, C. A., CLARK, S. & DUBEY, J. P. 1998. Ultrastructure of the oocysts, sporocysts, and sporozoites of Toxoplasma gondii. *J Parasitol,* 84**,** 505-12.

SPEER, C. A. & DUSZYNKI, D. W. 1975. Fine Structure of the Oocyst Walls of *Isospora serini* and *Isospora canaria* and excystation of *Isospora serini* from the canary *Serinus canarius*. *L.J. Protozool,* 22**,** 476-481.

STEPHEN, B., ROMMEL, M., DAUGSCHIES, A. & HABERKORN, A. 1997. Studies of resistance to anticoccidials in Eimeria field isolates and pure Eimeria strains. *Vet Parasitol,* 69**,** 19-29.

STOTISH, R. L., WANG, C. C. & MEYENHOFER, M. 1978a. Structure and composition of the oocyst wall of Eimeria tenella. *J Parasitol,* 64**,** 1074-81.

STOTISH, R. L., WANG, M. & MEYENHOFER, M. 1978b. Structure and Composition of the Oocyst Wall of *Eimeria tenella*. *The Journal of Parasitology,* 64**,** 1074-1081.

TAYLOR, M. A., COOP, R. L. & WALL, R. L. 2007. Parasites of Poultry and Game Birds. *In:* ANDERSON, J. M. & MACFADYEN, A. (eds.) *Veterinary Parasitology.* Iowa State, USA: Blackwell Publishing.

THACKER, C., PETERS, K., SRAYKO, M. & ROSE, A. M. 1995. The bli-4 locus of *Caenorhabditis elegans* encodes structually distinct kex2/subtilisin-like endoproteases essential for early development and adult morphology. *Genes and Development,* 9**,** 956-971.

THACKER, C. & ROSE, A. M. 2000. A look at the *Caenorhabditis elegans* Kex2/Subtilisi-like proprotein convertase family. *Bioessays,* 22**,** 545-553.

THOMPSON, J. D., HIGGINS, D. G. & GIBSON, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res,* 22**,** 4673-80.

TICHOPAD, A., DIDIER, A. & PFAFFL, M. W. 2004. Inhibition of real-time RT-PCR quantification due to tissue-specific contaminants. *Mol Cell Probes,* 18**,** 45-50.

TOMLEY, F. M. 1994. Characterization of rhoptry proteins of Eimeria tenella sporozoites: antigenic diversity of rhoptry epitopes within species of the genus Eimeria and among three asexual generations of a single species, E. tenella. *Infect Immun,* 62**,** 4656-8.

TOMLEY, F. M., BILLINGTON, K. J., BUMSTEAD, J. M., CLARK, J. D. & MONAGHAN, P. 2001. EtMIC4: a microneme protein from Eimeria tenella that contains tandem arrays of epidermal growth factor-like repeats and thrombospondin type-I repeats. *Int J Parasitol,* 31**,** 1303-10.

TOMLEY, F. M., BUMSTEAD, J. M., BILLINGTON, K. J. & DUNN, P. P. 1996. Molecular cloning and characterization of a novel acidic microneme protein (Etmic-2) from the apicomplexan protozoan parasite, Eimeria tenella. *Mol Biochem Parasitol,* 79**,** 195-206.

TOMLEY, F. M., CLARKE, L. E., KAWAZOE, U., DIJKEMA, R. & KOK, J. J. 1991. Sequence of the gene encoding an immunodominant microneme protein of Eimeria tenella. *Mol Biochem Parasitol,* 49**,** 277-88.

TREECK, M., SANDERS, J. L., ELIAS, J. E. & BOOTHROYD, J. C. 2011. The phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries. *Cell Host Microbe,* 10**,** 410-9.

TYZZER, E. 1929. Coccidiosis in gallinaceous birds. *Am J Hyg,* 10**,** 269-383.

TYZZER, E., THEILER, H. & JONES, E. 1932. Coccidiosis in gallinaceous birds II. A comparative study of species of *Eimeria* of the chicken. *Am J Hyg,* 15**,** 319-393.

TYZZER, E. E. 1932. Criteria and Methods in the Investigation of Avian Coccidiosis. *Science,* 75**,** 324-8.

UYEMURA, S. A., LUO, S., VIEIRA, M., MORENO, S. N. & DOCAMPO, R. 2004. Oxidative phosphorylation and rotenone-insensitive malate- and NADH-quinone oxidoreductases in Plasmodium yoelii yoelii mitochondria in situ. *J Biol Chem,* 279**,** 385-93.

UZUREAU, P., BARALE, J. C., JANSE, C. J., WATERS, A. P. & BRETON, C. B. 2004a. Gene targeting demonstrates that the Plasmodium berghei subtilisin PbSUB2 is essential for red cell invasion and reveals spontaneous genetic recombination events. *Cell Microbiol,* 6**,** 65-78.

UZUREAU, P., JC, J. B., JANSE, C., WATERS, A. & BRETON, C. 2004b. Gene targeting demonstrates that the Plasmodium berghei subtilisin PbSUB2 is essential for red cell invasion and reveals spontaneous genetic recombination events. *Cellular Microbiology,* 6**,** 65-78.

VELOT, C., MIXON, M. B., TEIGE, M. & SRERE, P. A. 1997. Model of a quinary structure between Krebs TCA cycle enzymes: a model for the metabolon. *Biochemistry,* 36**,** 14271-6.

VERCESI, A. E., RODRIGUES, C. O., UYEMURA, S. A., ZHONG, L. & MORENO, S. N. 1998. Respiration and oxidative phosphorylation in the apicomplexan parasite Toxoplasma gondii. *J Biol Chem,* 273**,** 31040-7.

VERMEULEN, A. N. 1998. Progress in recombinant vaccine development against coccidiosis. A review and prospects into the next millennium. *Int J Parasitol,* 28**,** 1121-30.

VERMEULEN, A. N. 2005. Vaccination against coccidia parasites. The method of choice? *IXth International Coccidiosis Conference.* Foz do Iguasso, Parana, Brazil.

VERMEULEN, A. N., KOK, J. J., VAN DEN BOOGAART, P., DIJKEMA, R. & CLAESSENS, J. A. 1993. Eimeria refractile body proteins contain two potentially functional characteristics: transhydrogenase and carbohydrate transport. *FEMS Microbiol Lett,* 110**,** 223-9.

VERVENNE, R. A., DIRKS, R. W., RAMESAR, J., WATERS, A. P. & JANSE, C. J. 1994. Differential expression in blood stages of the gene coding for the 21-kilodalton surface protein of ookinetes of Plasmodium berghei as detected by RNA in situ hybridisation. *Mol Biochem Parasitol,* 68**,** 259-66.

VISENTIN, V., PREVOT, D., MARTI, L. & CARPENE, C. 2003. Inhibition of rat fat cell lipolysis by monoamine oxidase and semicarbazide-sensitive amine oxidase substrates. *Eur J Pharmacol,* 466**,** 235-43.

WAITE, J. H. 1990a. The Phylogeny and Chemical Diversity of Quinone-Tanned Glues and Varnishes. *Comparative Biochemistry and Physiology,* 97B**,** 19-29.

WAITE, J. H. 1990b. The phylogeny and chemical diversity of quinone-tanned glues and varnishes. *Comp Biochem Physiol B,* 97**,** 19-29.

WALKER, R. A., SLAPETOVA, I., SLAPETA, J., MILLER, C. M. & SMITH, N. C. 2010. The glycosylation pathway of Eimeria tenella is upregulated during gametocyte development and may play a role in oocyst wall formation. *Eukaryot Cell,* 9**,** 127-35.

WALLACH, M. 1997. The importance of transmission-blocking immunity in the control of infections by apicomplexan parasites. *International Journal for Parasitology,* 27**,** 1159-1167.

WALLACH, M. 2003. The development of a novel vaccine against coccidiosis. *World Poult,* 19**,** 24-26.

WALLACH, M. 2010. Role of antibody in immunity and control of chicken coccidiosis. *Trends Parasitol,* 26**,** 382-7.

WALLACH, M., HALABI, A., PILLEMER, G., SAR-SHALOM, O., MENCHER, D., GILAD, M., BENDHEIM, U., DANFORTH, H. D. & AUGUSTINE, P. C. 1992. Maternal immunization with gametocyte antigens as a means of providing protective immunity against Eimeria maxima in chickens. *Infect Immun,* 60**,** 2036-9.

WALLACH, M., PILLEMER, G., YARUS, S., HALABI, A., PUGATSCH, T. & MENCHER, D. 1990. Passive immunization of chickens against Eimeria maxima infection with a monoclonal antibody developed against a gametocyte antigen. *Infect Immun,* 58**,** 557-62.

WALLACH, M., SMITH, N. C., BRAUN, R. & ECKERT, J. 1995a. Potential Control of Chicken Coccidiosis by Maternal Immunization. *Parasitology Today,* 11**,** 262-265.

WALLACH, M., SMITH, N. C., PETRACCA, M., MILLER, C. M., ECKERT, J. & BRAUN, R. 1995b. Eimeria maxima gametocyte antigens: potential use in a subunit maternal vaccine against coccidiosis in chickens. *Vaccine,* 13**,** 347-54.

WALLACH, M. G., ASHASH, U., MICHAEL, A. & SMITH, N. C. 2008. Field application of a subunit vaccine against an enteric protozoan disease. *PLoS One,* 3**,** e3948.

WALLACH, M. G., MENCHER, D., YARUS, S., PILLEMER, G., HALABI, A. & PUGATSCH, T. 1989. Eimeria maxima: identification of gametocyte protein antigens. *Exp Parasitol,* 68**,** 49-56.

WANG, C. C. 1975. Studies of the mitochondria from Eimeria tenella and inhibition of the electron transport by quinolone coccidiostats. *Biochim Biophys Acta,* 396**,** 210-9.

WANG, C. C., SIMASHKEVICH, P. M. & STOTISH, R. L. 1979. Mode of anticoccidial action of arprinocid. *Biochem Pharmacol,* 28**,** 2241-8.

WATERHOUSE, A. M., PROCTER, J. B., MARTIN, D. M., CLAMP, M. & BARTON, G. J. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* 25**,** 1189-91.

WEIDMAN, P. J., KAY, E. S. & SHAPIRO, B. M. 1985. Assembly of the Sea Urchin Fertilization Membrane: Isolation of Proteoliasin, a Calcium-dependent Ovoperoxidase Binding Protein. *The Journal of Cell Biology,* 100**,** 938-946.

WILLIAMS, R. B. 1994. Safety of the attenuated anticoccidial vaccine 'Paracox' in broiler chickens isolated from extraneous coccidial infection. *Vet Res Commun,* 18**,** 189-98.

WILLIAMS, R. B. 1995. Epidemiological studies of coccidiosis in the domesticated fowl (Gallus gallus): II. Physical condition and survival of Eimeria acervulina oocysts in poultry-house litter. *Applied parasitology,* 36**,** 90.

WILLIAMS, R. B. 1998. Epidemiological aspects of the use of live anticoccidial vaccines for chickens. *Int J Parasitol,* 28**,** 1089-98.

WILLIAMS, R. B. 1999. Three enzymes newly identified from the genus Eimeria and two more newly identified from E. maxima, leading to the discovery of some aliphatic acids with activity against coccidia of the domesticated fowl. *Vet Res Commun,* 23**,** 151-63.

WILLIAMS, R. B. 2002. Anticoccidial vaccines for broiler chickens: pathways to success. *Avian Pathol,* 31**,** 317-53.

WILLIAMS, R. B., CARLYLE, W. W. H., BOND, D. R. & BROWN, I. A. G. 1999. The efficacy and economic benefits of Paracox (R), a live attenuated anticoccidial vaccine, in commercial trials with standard broiler chickens in the United Kingdom. *International Journal for Parasitology,* 29**,** 341-355.

WILLIAMS, R. B., MARSHALL, R. N., PAGES, M., DARDI, M. & DEL CACHO, E. 2009. Pathogenesis of Eimeria praecox in chickens: virulence of field strains compared with laboratory strains of E. praecox and Eimeria acervulina. *Avian Pathol,* 38**,** 359-66.

WITHERS-MARTINEZ, C., SALDANHA, J. W., ELY, B., HACKETT, F., O'CONNOR, T. & BLACKMAN, M. J. 2002. Expression of Recombinant *Plasmodium falciparum* Subtilisin-like Protease-1 in Insect Cells. *The Journal of Biological Chemistry,* 277**,** 29698-29709.

WONG, J. L., CRETON, R. & WESSEL, G. M. 2004. The oxidative burst at fertilization is dependent upon activation of the dual oxidase Udx1. *Dev Cell,* 7**,** 801-14.

WORLICZEK, H., L ,, RUTTKOWSKI, B., SCHWARZ, L., WITTER, K. & TSCHULENK, W. 2013. Isospora suis in an Epithelial Cell Culture System – An In Vitro Model for Sexual Development in Coccidia. *PLoS One,* 8**,** e69797.

XIA, D., SANDERSON, S. J., JONES, A. R., PRIETO, J. H., YATES, J. R., BROMLEY, E., TOMLEY, F. M., LAL, K., SINDEN, R. E., BRUNK, B. P., ROOS, D. S. & WASTLING, J. M. 2008. The proteome of Toxoplasma gondii: integration with the genome provides novel insights into gene expression and annotation. *Genome Biol,* 9**,** R116.

XIE, M., CAI, J., LI, A. & PENG, X. Coccidiosis of domestic fowl in China. VIIIth International Coccidiosis Conference, 2001 Palm Cove, Sydney, Australia. 153-154.

YABUTA, Y., TAKAGI, H., INOUYE, M. & SHINDE, U. 2001. Folding Pathway Mediated by an Intramolecular Chaperone : Propeptide Release Modulates Activation Precision of Pro-Subtilisin. *The Journal of Biological Chemistry,* 276**,** 44427-44434.

YAMAGUCHI, Y., HEARING, V. J., MAEDA, A. & MORITA, A. 2010. NADPH:quinone oxidoreductase-1 as a new regulatory enzyme that increases melanin synthesis. *J Invest Dermatol,* 130**,** 645-7.

YANG, J. & KRAMER, J. M. 1994. In Vitro Mutagenesis of *Caenorhabditis elegans* Cuticle Collagens Identifies a Potential Subtilisin-like Protease Cleavage Site and Demonstrates that carboxyl Domain Disulfide Bonding is Required for Normal Function but Not Assembly. *Molecular and Cellular Biology,* 14**,** 2722-2730.

YEOH, S., O'DONNELL, R., KOUSSIS, K., DLUZEWSKI, A., ANSELL, K., OSBORNE, S., HACKETT, F., WITHERS-MARTINEZ, C., MITCHELL, G., BANNISTER, L., BRYANS, J., KETTLEBOROUGH, C. & BLACKMAN, M. 2007. Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. *Cell. ,* 131**,** 1072-83.

YIN, G., LIU, X., ZOU, J., HUANG, X. & SUO, X. 2011. Co-expression of reporter genes in the widespread pathogen Eimeria tenella using a double-cassette expression vector strategy. *Int J Parasitol,* 41**,** 813-6.

ZHANG, H., COMPAORE, M. K., LEE, E. G., LIAO, M., ZHANG, G., SUGIMOTO, C., K, K. F., NISHIKAWA, Y. & XUAN, X. 2007. Apical membrane antigen 1 is a cross-reactive antigen between Neospora caninum and Toxoplasma gondii, and the anti-NcAMA1 antibody inhibits host cell invasion by both parasites. *Molecular and Biochemical Parasitology,* 151**,** 205-12.

ZHOU, C., CONFALONIERI, F., M, J., R, P., LI, Z. & JANIN, J. 2001. Silk fibroin: structural implications of a remarkable amino acid sequence. *Proteins,* 44**,** 119-22.

ZHOU, X. W., BLACKMAN, M. J., HOWELL, S. A. & CARRUTHERS, V. B. 2004a. Proteomic Analysis of Cleavage Events Reveals a Dynamic Two-step Mechanism for Proteolysis of a Key Parasite Adhesive Complex. *Molecullar and Cellular Proteomics,* 3**,** 565-576.

ZHOU, X. W., BLACKMAN, M. J., HOWELL, S. A. & CARRUTHERS, V. B. 2004b. Proteomic analysis of cleavage events reveals a dynamic two-step mechanism for proteolysis of a key parasite adhesive complex. *Mol Cell Proteomics,* 3**,** 565-76.

ZHOU, X. W., KAFSACK, B. F., COLE, R. N., BECKETT, P., SHEN, R. F. & CARRUTHERS, V. B. 2005. The opportunistic pathogen Toxoplasma gondii deploys a diverse legion of invasion and survival proteins. *J Biol Chem,* 280**,** 34233-44.

APPENDIX

Table S1: Identification of 997 proteins in the oocyst/sporocyst proteome of *Toxoplasma gondii*

[Gene ID]	[Number of Peptide Sequences]	[Number of Spectra]	[Product Description]
TGME49_035210	10	183	PAN domain-containing protein
TGME49_097470	10	87	myosin light chain 2, putative
TGME49_043660	2	6	glucose-methanol-choline domain-containing protein
TGME49_053180	2	4	hypothetical protein
TGME49_100100	16	320	rhoptry neck protein 2
TGME49_119090	2	4	hypothetical protein
TGME49_026410	6	27	elongation factor 1-beta, putative
TGME49_071490	17	564	hypothetical protein
TGME49_046690	8	86	1,4-alpha-glucan branching enzyme, putative
TGME49_114410	3	9	hypothetical protein
TGME49_120670	2	4	hypothetical protein, conserved
TGME49_035680	9	74	M16 family peptidase, putative
TGME49_108710	19	665	hypothetical protein, conserved
TGME49_108600	5	13	hypothetical protein
TGME49_117700	2	4	enoyl-CoA hydratase/isomerase family domain-containing protein
TGME49_063060	17	489	26S proteasome non-ATPase regulatory subunit 1, putative
TGME49_058740	2	4	eukaryotic translation initiation factor 2A, putative
TGME49_038250	2	4	60s ribosomal protein L36, putative
TGME49_111360	11	106	AGC kinase TgPKG1
TGME49_058060	5	20	hypothetical protein, conserved
TGME49_064080	2	4	acyl carrier protein
TGME49_116190	14	360	superoxide dismutase, putative
TGME49_090960	9	45	phosphoglucan, water dikinase protein, putative
TGME49_092130	9	70	60s ribosomal protein L13a, putative
TGME49_088650	19	942	hypothetical protein
TGME49_057990	14	121	heat shock protein, putative
TGME49_030850	3	12	mucin, putative
TGME49_006510	10	120	peptidase M16 domain containing protein
TGME49_097970	2	4	aspartyl aminopeptidase, putative
TGME49_030410	4	15	peroxiredoxin 3
TGME49_019800	5	17	vacuolar ATP synthase subunit B, putative
TGME49_047450	4	16	hypothetical protein, conserved
TGME49_101120	4	16	acetyil CoA acetyltransferase/thiolase, putative
TGME49_118580	9	89	phosphoglucomutase, putative
TGME49_077000	6	31	transport protein Sec24, putative
TGME49_066700	2	4	hypothetical protein
TGME49_047960	2	4	hypothetical protein
TGME49_105850	4	20	RNA binding protein, putative
TGME49_004530	9	60	microneme protein MIC11
TGME49_115480	9	118	acyl-CoA oxidase, putative
TGME49_114810	5	33	60S ribosomal protein L7, putative
TGME49_093590	4	13	3-oxoacyl-[acyl-carrier-protein] synthase, putative
TGME49_023070	5	66	hypothetical protein

TGME49_016590	4	13	transportin, putative
TGME49_121620	6	40	dynamin-like protein, putative
TGME49_015260	7	45	carbamoyl phosphate synthetase II
TGME49_097060	11	76	phosphoglycerate mutase 1, putative
TGME49_028210	2	4	26S proteasome regulatory ATPase subunit, putative
TGME49_000360	2	4	hypothetical protein
TGME49_035930	7	40	KH domain-containing protein
TGME49_119890	23	1181	hypothetical protein
TGME49_053290	6	40	valyl-tRNA synthetase, putative
TGME49_065530	4	16	RNA binding motif-containing protein
TGME49_097430	5	42	hypothetical protein, conserved
TGME49_091180	4	8	hypothetical protein, conserved
TGME49_035580	37	1999	hypothetical protein
TGME49_060500	2	4	hypothetical protein, conserved
TGME49_036580	2	4	60S ribosomal protein L7a, putative
TGME49_013350	4	26	40S ribosomal protein S15, putative
TGME49_120630	6	45	choline kinase, putative
TGME49_031410	3	16	hypothetical protein, conserved
TGME49_045460	5	77	40S ribosomal protein S8, putative
TGME49_113230	4	16	eukaryotic translation initiation factor 2 alpha subunit, putative
TGME49_088460	8	45	hypothetical protein
TGME49_061600	2	4	X-prolyl aminopeptidase, putative
TGME49_018780	4	25	phosphoserine aminotransferase, putative
TGME49_035700	2	4	sedoheptulose-1,7 bisphosphatase, putative
TGME49_006400	27	1291	DNA photolyase, putative
TGME49_034190	2	4	glycine hydroxymethyltransferase, putative
TGME49_039020	4	13	ATP-binding cassette protein subfamily B member 2
TGME49_061780	5	22	microneme protein 7
TGME49_046130	18	260	serine protease inhibitor, putative
TGME49_042330	3	21	40S ribosomal protein S5, putative
TGME49_093180	7	109	NADP-specific glutamate dehydrogenase, putative
TGME49_121650	8	65	protein kinase
TGME49_007830	5	22	MORN repeat-containing protein
TGME49_113390	8	54	60s ribosomal protein l6, putative
TGME49_005710	3	8	lysyl-tRNA synthetase, putative
TGME49_001410	2	4	protein phosphatase 2C, putative
TGME49_112200	5	20	serine/threonine protein phosphatase, putative
TGME49_094350	16	192	DEAD/DEAH box helicase, putative
TGME49_005010	6	46	splicing factor 3B subunit 1, putative
TGME49_056760	14	405	pyruvate kinase, putative
TGME49_016450	4	16	proteasome subunit alpha type 3, putative
TGME49_047930	2	4	syntaxin, putative
TGME49_062670	8	60	60S ribosomal protein L18a, putative
TGME49_118430	8	96	malate dehydrogenase, putative
TGME49_005340	5	18	40S ribosomal protein S12, putative

TGME49_109600	4	29	hypothetical protein
TGME49_063530	2	6	chaperonin CPN10, mitochondrial, putative
TGME49_019170	1	4	hypothetical protein
TGME49_062620	4	17	Gbp1p protein, putative
TGME49_110860	2	4	hypothetical protein, conserved
TGME49_030210	14	212	hypothetical protein
TGME49_061950	10	145	ATP synthase beta chain, putative
TGME49_001700	3	9	sec 13, putative
TGME49_030510	2	4	hypothetical protein
TGME49_016690	2	4	ankyrin repeat domain-containing protein
TGME49_053430	2	4	asparagine synthase, putative
TGME49_099810	4	13	cysteinyl-tRNA synthetase, putative
TGME49_063520	6	55	hypothetical protein, conserved
TGME49_090940	3	16	endomembrane domain70-containing protein
TGME49_090670	38	2533	cytosol aminopeptidase
TGME49_070320	2	4	protein phosphatase 2C, putative
TGME49_029180	22	413	importin beta-3 subunit, putative
TGME49_011420	3	13	hypothetical protein
TGME49_115770	5	41	cytochrome p450, putative
TGME49_110780	7	132	hypothetical protein
TGME49_093870	5	30	thioredoxin, putative
TGME49_112630	94	18488	hypothetical protein
TGME49_032600	4	16	patatin-like phospholipase domain-containing protein
TGME49_109210	7	40	peroxidoxin 2
TGME49_068980	2	4	hypothetical protein
TGME49_042390	10	138	enoyl-CoA hydratase/isomerase family protein, putative
TGME49_049670	2	4	cysteine proteinase, putative
TGME49_065330	3	16	CMGC kinase, GSK family TgPK3
TGME49_100000	1	4	60S ribosomal protein L18, putative
TGME49_089650	2	4	phosphoenolpyruvate carboxykinase
TGME49_063850	11	237	hypothetical protein
TGME49_001780	25	1224	microneme protein 2
TGME49_055890	4	33	monodehydroascorbate reductase, putative
TGME49_002390	7	49	hypothetical protein
TGME49_080380	25	822	non-transmembrane antigen
TGME49_017460	3	9	glutaminyl-tRNA synthetase, putative
TGME49_118510	2	4	N-ethylmaleimide-sensitive factor
TGME49_115130	3	8	protein-L-isoaspartate O-methyltransferase, putative
TGME49_003580	8	87	NAD-binding domain-containing protein
TGME49_105780	5	18	exoribonuclease, putative
TGME49_053300	2	4	hypothetical protein, conserved
TGME49_019850	10	100	prolyl-tRNA synthetase, putative
TGME49_074060	7	45	mitochondrial 2-oxoglutarate/malate carrier protein, putative
TGME49_095040	5	25	importin subunit beta-1, putative
TGME49_106050	6	41	hypothetical protein
TGME49_111470	11	119	hypothetical protein

TGME49_071430	4	8	hypothetical protein
TGME49_073520	5	13	proteasome PCI domain-containing protein
TGME49_066640	2	4	acetyl-coenzyme A synthetase, putative
TGME49_089820	2	4	TBC domain-containing protein
TGME49_061440	2	4	armadillo/beta-catenin-like repeat-containing protein
TGME49_067350	2	4	small nuclear ribonucleoprotein, putative
TGME49_078740	2	4	diaminopimelate decarboxylase, putative
TGME49_062920	7	40	trypsin, putative
TGME49_052360	5	42	Rhoptry kinase family protein ROP24 (incomplete catalytic triad)
TGME49_110490	5	24	ribosomal protein L22, putative
TGME49_019830	1	4	high mobility group protein, putative
TGME49_015040	10	75	hypothetical protein
TGME49_008030	10	101	microneme protein MIC4
TGME49_078990	5	18	phosphate carrier protein, putative
TGME49_088220	2	6	hypothetical protein
TGME49_078870	12	113	myosin F (TgMyoF) protein
TGME49_113380	9	314	hypothetical protein, conserved
TGME49_120210	4	22	WD-40 repeat protein, putative
TGME49_106910	3	9	hypothetical protein
TGME49_012290	4	13	40s ribosomal protein s19, putative
TGME49_118230	19	560	phosphoglycerate kinase, putative
TGME49_075650	6	34	p97 protein
TGME49_002380	5	25	rab GDP dissociation inhibitor alpha, putative
TGME49_063320	2	4	hypothetical protein
TGME49_110440	2	4	phosphatidylinositol-4-phosphate 5-kinase, putative
TGME49_069980	5	29	protein transport protein Sec61 alpha subunit isoform 1, putative
TGME49_120570	3	8	G1 to S phase transition protein, putative
TGME49_049780	2	4	hypothetical protein
TGME49_119350	20	905	SRS domain containing protein
TGME49_082230	4	17	sulfate adenylyltransferas-adenylylsulfate kinase, putative
TGME49_035490	3	13	hypothetical protein
TGME49_094200	5	13	glucose-6-phosphate dehydrogenase
TGME49_073870	4	13	chromatin remodelling complex protein SNF2L, putative
TGME49_027360	3	9	60S ribosomal protein L3, putative
TGME49_032300	3	25	40S ribosomal protein S3, putative
TGME49_033460	13	332	SRS29B (= SAG1, P30)
TGME49_116520	3	9	1,4-alpha-glucan branching enzyme, putative
TGME49_040090	4	25	Rhoptry kinase family protein ROP34, putative
TGME49_029010	13	246	hypothetical protein
TGME49_057090	2	4	NAC domain containing protein
TGME49_020100	4	13	phosphoribosylpyrophosphate synthetase, putative
TGME49_110150	19	225	long-chain-fatty-acid-CoA ligase, putative
TGME49_060820	4	42	hypothetical protein
TGME49_075660	21	562	hypothetical protein
TGME49_045500	3	9	dipeptidyl peptidase IV domain-containing protein
TGME49_067050	3	9	hypothetical protein

TGME49_118440	3	9	DEAH-box RNA/DNA helicase, putative
TGME49_040280	4	40	hypothetical protein
TGME49_088500	4	19	malate:quinone oxidoreductase, putative
TGME49_054470	2	4	hypothetical protein
TGME49_039760	2	4	60S ribosomal protein L22, putative
TGME49_009060	2	4	thrombospondin type 1 domain-containing protein
TGME49_048900	2	4	hypothetical protein
TGME49_049590	4	16	proteasome subunit alpha type 5, putative
TGME49_078660	11	93	P-type Ca(2+)-ATPase, putative
TGME49_026300	4	13	oxidoreductase, putative
TGME49_036070	4	16	pyrroline-5-carboxylate reductase, putative
TGME49_090160	6	17	sortilin, putative
TGME49_095360	5	30	hypothetical protein
TGME49_014080	23	998	toxofilin
TGME49_031770	2	6	tubulin alpha chain, putative
TGME49_111240	4	16	DnaJ domain-containing protein
TGME49_072640	2	4	translation initiation factor eIF-2B epsilon subunit, putative
TGME49_089310	5	25	cullin family protein
TGME49_118310	11	146	transketolase, putative
TGME49_076850	12	506	late embryogenesis abundant domain-containing protein
TGME49_021670	3	9	transcription elongation factor FACT 140 kDa, putative
TGME49_104490	5	20	hypothetical protein
TGME49_048340	18	222	GTP-binding nuclear protein RAN/TC4, putative
TGME49_044200	9	78	2-oxoglutarate dehydrogenase, putative
TGME49_066960	14	581	tubulin beta chain
TGME49_110460	1	4	RAB6 protein
TGME49_110800	2	16	CRAL/TRIO domain-containing protein
TGME49_068950	6	41	la domain-containing protein
TGME49_097420	2	4	beta-tubulin cofactor D, putative
TGME49_077790	2	4	hypothetical protein
TGME49_017740	1	4	oxoacyl-ACP reductase, putative
TGME49_027960	6	36	proteasome PCI domain-containing protein
TGME49_013900	12	153	regulator of chromosome condensation, putative
TGME49_089330	29	791	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_094710	2	4	hypothetical protein
TGME49_097160	22	743	hypothetical protein
TGME49_011290	5	13	hypothetical protein
TGME49_020400	11	397	actin depolymerizing factor
TGME49_009910	2	9	histone H2B variant 1
TGME49_111400	8	42	hypothetical protein, conserved
TGME49_032350	31	4602	lactate dehydrogenase
TGME49_100060	3	9	signal peptidase subunit, putative
TGME49_014260	23	421	alpha-glucan water dikinase 1, putative
TGME49_114780	6	17	myosin G
TGME49_097960	4	13	hypothetical protein, conserved

TGME49_014490	22	482	M16 family peptidase, putative
TGME49_002930	2	4	hypothetical protein
TGME49_109370	2	4	adaptin, putative
TGME49_022430	15	136	ubiquitin-transferase domain containing protein
TGME49_016660	2	4	hypothetical protein, conserved
TGME49_108000	2	4	Gpi16 subunit, GPI transamidase domain-containing protein
TGME49_080740	3	8	signal peptidase, putative
TGME49_023960	6	20	26S proteasome non-ATPase regulatory subunit 4, putative
TGME49_025940	3	8	hypothetical protein, conserved
TGME49_068760	6	18	hypothetical protein
TGME49_109560	3	29	hypothetical protein, conserved
TGME49_083780	13	121	glucose-6-phosphate isomerase
TGME49_046580	3	8	hypothetical protein
TGME49_077720	6	58	nucleoside-triphosphatase, putative
TGME49_089580	6	44	strictosidine synthase domain-containing protein
TGME49_108840	31	2207	SRS51 (= SRS3)
TGME49_092080	8	52	leucyl-tRNA synthetase, putative
TGME49_071970	2	4	hypothetical protein
TGME49_022660	2	4	hypothetical protein
TGME49_066130	3	16	glutathione/thioredoxin peroxidase, putative
TGME49_053400	4	16	hypothetical protein
TGME49_068850	21	776	enolase 2
TGME49_004130	24	1021	membrane-attack complex / perforin domain-containing protein
TGME49_105270	5	20	hypothetical protein
TGME49_013770	2	4	RNA helicase, putative
TGME49_024900	6	29	adenylate kinase, putative
TGME49_023430	3	10	hypothetical protein
TGME49_039360	3	12	hypothetical protein
TGME49_111720	29	1988	heat shock protein 70, putative
TGME49_086420	24	1291	elongation factor 1-alpha, putative
TGME49_115730	15	366	apical membrane antigen, putative
TGME49_013340	9	31	glucose-methanol-choline oxidoreductase domain-containing protein
TGME49_058200	25	641	glutamic acid-rich protein, putative
TGME49_063050	3	8	60S ribosomal protein L13, putative
TGME49_068890	4	20	citrate synthase, putative
TGME49_092210	5	12	hypothetical protein
TGME49_052510	4	25	hypothetical protein, conserved
TGME49_069380	10	96	hypothetical protein
TGME49_025930	8	68	triosephosphate isomerase, putative
TGME49_018530	2	4	hypothetical protein
TGME49_114260	3	26	hypothetical protein
TGME49_113410	19	304	26S proteasome non-ATPase regulatory subunit 2, putative
TGME49_093570	4	18	translocation protein Sec62, putative
TGME49_001840	4	17	eukaryotic aspartyl protease, putative
TGME49_058550	39	7826	SRS28 (= SporoSAG)
TGME49_015590	3	8	succinate dehydrogenase (ubiquinone) flavoprotein subunit, mitochondrial,

TGME49_013010	2	4	hypothetical protein
TGME49_106500	15	101	hypothetical protein
TGME49_049170	2	4	ras family domain-containing protein
TGME49_111210	5	41	hypothetical protein
TGME49_019810	6	20	hypothetical protein
TGME49_059180	13	185	cystathionine beta-synthase, putative
TGME49_115260	8	100	alanine dehydrogenase, putative
TGME49_054390	3	8	hypothetical protein
TGME49_023590	7	64	proteasome component PRE3 precursor, putative
TGME49_025060	5	29	nnucleoredoxin, putative
TGME49_028470	2	4	ribosomal protein L15, putative
TGME49_095710	2	4	ubiquitin-transferase domain-containing protein
TGME49_010840	5	30	arginyl-tRNA synthetase, putative
TGME49_019820	5	54	polyubiquitin, putative
TGME49_034370	4	8	SRS42
TGME49_019710	9	69	hypothetical protein
TGME49_093770	5	20	chitinase class I, putative
TGME49_112150	8	40	hypothetical protein
TGME49_036040	24	780	fructose-1,6-bisphosphate aldolase
TGME49_066060	7	46	40S ribosomal protein S0-A, putative
TGME49_017510	2	4	hypothetical protein
TGME49_120710	2	4	hypothetical protein
TGME49_028360	2	4	FK506-binding protein 1, putative
TGME49_010410	1	4	high mobility group protein
TGME49_094220	2	4	hypothetical protein
TGME49_066420	2	4	hypothetical protein
TGME49_067060	2	4	60S ribosomal protein L14, putative
TGME49_093340	3	16	ran-specific GTPase-activating protein, putative
TGME49_046450	2	4	hypothetical protein
TGME49_009600	4	16	hypothetical protein
TGME49_023680	2	4	ubiquitin, putative
TGME49_002870	4	8	mitochondrial carrier domain-containing protein
TGME49_039100	5	26	40S ribosomal protein S7, putative
TGME49_013870	3	8	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_110680	2	4	hypothetical protein
TGME49_049530	10	89	exportin, putative
TGME49_033770	5	29	Ca2+-ATPase, putative
TGME49_036050	3	24	fructose-bisphosphate aldolase, putative
TGME49_088210	8	52	phospholipase A2 activating protein, putative
TGME49_080710	2	4	proteasome A-type and B-type domain-containing protein
TGME49_088360	24	776	tryptophanyl-tRNA synthetase, putative
TGME49_072550	2	4	hypothetical protein
TGME49_091890	13	129	microneme protein MIC1
TGME49_001800	2	4	RNA-binding protein, putative
TGME49_080700	17	300	lysine decarboxylase, putative

TGME49_089300	4	13	methionyl-tRNA synthetase, putative
TGME49_003230	9	78	hypothetical protein
TGME49_093900	3	9	hypothetical protein
TGME49_058950	6	25	lectin-domain protein
TGME49_026970	4	13	40S ribosomal protein S11, putative
TGME49_115220	4	47	rhoptry protein, putative
TGME49_060170	3	9	elongation factor G, putative
TGME49_094820	19	230	type I fatty acid synthase, putative
TGME49_053690	8	62	hypothetical protein
TGME49_038040	6	34	thioredoxin, putative
TGME49_004520	10	124	hypothetical protein
TGME49_031080	2	4	hypothetical protein
TGME49_005560	11	81	nascent polypeptide-associated complex alpha chain, putative
TGME49_014190	21	568	SRS46
TGME49_006590	2	4	CAM kinase, CDPK family
TGME49_040890	12	113	phosphofructokinase, putative
TGME49_067070	9	231	aquaporin, putative
TGME49_115210	2	4	rhoptry protein, putative
TGME49_090200	2	4	opine dehydrogenase, putative
TGME49_032180	2	4	UBA/TS-N domain-containing protein
TGME49_106330	10	66	hypothetical protein, conserved
TGME49_046800	6	38	acylamino-acid-releasing enzyme, putative
TGME49_062450	2	4	hypothetical protein
TGME49_062400	13	89	lipase domain-containing protein
TGME49_043580	2	4	HIT domain-containing protein
TGME49_002500	1	4	hypothetical protein
TGME49_052290	3	9	importin alpha, putative
TGME49_109820	5	20	ribosomal protein L11, putative
TGME49_015470	5	46	60S ribosomal protein L10a, putative
TGME49_048460	4	17	ubiquitin, putative
TGME49_011680	23	950	protein disulfide isomerase
TGME49_001680	8	50	eukaryotic translation initiation factor 3 subunit 10, putative
TGME49_038050	6	20	tudor / staphylococcal nuclease domain-containing protein
TGME49_026380	12	148	hypothetical protein
TGME49_003310	28	3716	dense granule protein 7
TGME49_097290	3	12	hypothetical protein, conserved
TGME49_004400	5	20	ATP synthase alpha chain, putative
TGME49_110040	3	10	ubiquitin-conjugating enzyme E2, putative
TGME49_019630	2	4	NADPH-cytochrome p450 reductase, putative
TGME49_024270	3	9	hypothetical protein, conserved
TGME49_061570	2	4	60S ribosomal protein L7a, putative
TGME49_060630	5	18	DnaJ domain-containing protein
TGME49_058580	13	208	Rhoptry kinase family protein ROP17
TGME49_021870	3	9	hypothetical protein
TGME49_086720	11	173	heat shock protein 28
TGME49_039820	12	105	D-3-phosphoglycerate dehydrogenase, putative

TGME49_113140	6	29	isocitrate dehydrogenase, putative
TGME49_030340	11	101	high molecular mass nuclear antigen, putative
TGME49_032230	4	13	60S ribosomal protein L30, putative
TGME49_015980	6	60	hypothetical protein
TGME49_014760	2	4	D-3-phosphoglycerate dehydrogenase, putative
TGME49_002460	2	4	diacylglycerol kinase, putative
TGME49_092020	9	36	cysteine repeat modular protein, putative
TGME49_073510	18	248	hypothetical protein
TGME49_042290	8	81	proteasome subunit alpha type 6, putative
TGME49_106270	4	20	hypothetical protein
TGME49_011040	4	14	Sec61beta family protein
TGME49_003810	2	4	kelch motif domain-containing protein
TGME49_109760	14	169	hypothetical protein
TGME49_049850	2	4	hypothetical protein, conserved
TGME49_031140	5	20	R1 protein
TGME49_048390	2	6	60S ribosomal protein L26, putative
TGME49_025150	3	9	hypothetical protein
TGME49_038100	2	4	emp24/gp25L/p24 family domain-containing, transmembrane protein, putative
TGME49_019540	8	50	alanyl-tRNA synthetase, putative
TGME49_016260	3	9	eukaryotic initiation factor-2B gamma subunit, putative
TGME49_009810	6	80	hypothetical protein
TGME49_011690	2	4	ATP-dependent helicase, putative
TGME49_058660	6	88	rhoptry protein 6
TGME49_032710	9	140	40S ribosomal protein S3a, putative
TGME49_009690	3	9	U2 small nuclear ribonucleoprotein, putative
TGME49_027800	3	9	hypothetical protein, conserved
TGME49_051780	16	231	heat shock protein 70, putative
TGME49_104670	2	8	internalin, putative
TGME49_015220	11	46	hypothetical protein, conserved
TGME49_058150	8	64	proteasome subunit alpha type 7, putative
TGME49_028240	7	70	hypothetical protein
TGME49_031000	2	4	phosphatidylcholine transfer protein, putative
TGME49_040650	17	170	coatomer alpha subunit, putative
TGME49_116250	6	87	hypothetical protein
TGME49_047510	2	4	fructose-1,6-bisphosphatase, putative
TGME49_105030	5	20	ribokinase, putative
TGME49_085440	2	6	hypothetical protein
TGME49_040470	5	20	hypothetical protein
TGME49_025310	4	8	ARF1-directed GTPase-activating protein, putative
TGME49_002770	2	4	RNA-binding protein, putative
TGME49_028630	2	4	hypothetical protein
TGME49_025860	3	9	hypothetical protein
TGME49_026860	2	15	hypothetical protein
TGME49_020890	4	18	hypothetical protein
TGME49_026910	58	4314	glycogen debranching enzyme, putative

TGME49_005700	10	98	20 kDa cyclophilin precursor
TGME49_049270	7	42	thioredoxin, putative
TGME49_040060	5	24	hypothetical protein
TGME49_106450	6	64	short chain dehydrogenase family protein, putative
TGME49_085690	9	106	notch (DSL) domain-containing protein
TGME49_108020	45	5913	SRS57 (= SAG3, P43)
TGME49_061210	5	25	26S proteasome non-ATPase regulatory subunit, putative
TGME49_105340	5	24	hypothetical protein
TGME49_049990	4	29	microtubule-binding protein, putative
TGME49_095470	2	4	C2 domain-containing protein
TGME49_066970	2	4	hypothetical protein
TGME49_108580	2	4	lon protease, putative
TGME49_110830	4	33	delta3,5-delta2,4-dienoyl-CoA isomerase, putative
TGME49_025080	3	23	40S ribosomal protein S18, putative
TGME49_071050	7	88	SRS34A (= SAG2A, P22)
TGME49_108080	29	1753	Rhoptry kinase family protein ROP5 (incomplete catalytic triad)
TGME49_051850	2	8	serine/threonine protein phosphatase 5, putative
TGME49_075690	2	4	chaperone clpB 1 protein, putative
TGME49_106030	4	13	hypothetical protein
TGME49_110640	8	72	uridine phosphorylase, putative
TGME49_058800	2	4	Rhoptry kinase family protein ROP31
TGME49_035160	9	55	hypothetical protein
TGME49_105930	3	9	hypothetical protein
TGME49_062460	2	4	PX domain-containing protein
TGME49_048910	4	13	hypothetical protein
TGME49_087270	4	16	hypothetical protein
TGME49_095640	11	41	endothelin-converting enzyme, putative
TGME49_113000	5	41	hypothetical protein
TGME49_105520	7	69	40S ribosomal protein S2, putative
TGME49_010370	8	53	RNA helicase-1
TGME49_090660	4	24	polypyrimidine track-binding protein, putative
TGME49_097070	5	72	duplicated carbonic anhydrase
TGME49_053730	9	76	importin-alpha re-exporter, putative
TGME49_049610	6	20	nuclear cap binding protein, putative
TGME49_094620	11	89	eukaryotic translation initiation factor 3 subunit 8, putative
TGME49_031970	13	104	pre-mRNA splicing factor PRP8, putative
TGME49_035470	60	8820	myosin A, putative
TGME49_119560	6	69	microneme protein MIC3
TGME49_063090	14	150	14-3-3 protein, putative
TGME49_047760	10	125	long chain acyl-CoA synthetase, putative
TGME49_009610	18	529	oocyst wall protein COWP, putative
TGME49_047520	5	43	hypothetical protein, conserved
TGME49_016880	6	38	receptor for activated C kinase, RACK protein, putative
TGME49_114400	8	112	branched-chain alpha-keto acid dehydrogenase E1 component beta chain, putative
TGME49_019750	6	17	cytochrome c, putative

TGME49_065130	3	25	hypothetical protein
TGME49_031600	14	164	importin, putative
TGME49_063040	10	116	40S ribosomal protein S16, putative
TGME49_071930	7	41	hypothetical protein, conserved
TGME49_068880	2	9	hypothetical protein
TGME49_039260	10	151	histone H4, putative
TGME49_056060	2	4	nucleosome assembly protein, putative
TGME49_105860	10	79	CAM kinase, CDPK family TgCDPK1_2 (TGTPK4)
TGME49_008530	2	4	nicotinate phosphoribosyltransferase, putative
TGME49_025120	4	13	rRNA processing protein, putative
TGME49_095700	8	22	UBA/TS-N domain-containing protein
TGME49_015140	5	41	hypothetical protein
TGME49_038200	8	62	alpha/beta hydrolase fold domain containing protein
TGME49_032400	5	21	PAN domain-containing protein
TGME49_063500	2	4	vacuolar protein sorting 26, putative
TGME49_092220	10	148	26S proteasome non-ATPase regulatory subunit, putative
TGME49_118530	11	70	hypothetical protein
TGME49_092920	2	4	heat shock protein 90, putative
TGME49_003990	3	8	hypothetical protein, conserved
TGME49_073090	23	976	cell division protein 48, putative
TGME49_011030	3	8	hypothetical protein
TGME49_009290	2	6	40S ribosomal protein S28, putative
TGME49_032410	14	408	hypothetical protein
TGME49_057680	5	42	myosin light chain TgMLC1
TGME49_060190	59	8777	microneme protein, putative
TGME49_111310	3	9	serine/threonine protein phosphatase, putative
TGME49_111230	6	29	hypothetical protein, conserved
TGME49_005470	35	2695	elongation factor 2, putative
TGME49_094360	2	4	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_089680	2	4	Ras family domain-containing protein
TGME49_019310	27	776	heat shock protein, putative
TGME49_080600	5	20	histidyl-tRNA synthetase, putative
TGME49_025200	2	4	trichohyalin, putative
TGME49_110080	5	25	long-chain-fatty-acid-CoA ligase, putative
TGME49_092350	18	561	hypothetical protein
TGME49_058360	4	20	hypothetical protein
TGME49_014270	4	18	translation initiation factor IF-2, putative
TGME49_007620	1	4	pyridine nucleotide-disulphide oxidoreductase domain-containing protein
TGME49_033500	3	9	triosephosphate isomerase, putative
TGME49_062150	5	34	kelch motif domain-containing protein
TGME49_029640	2	4	hypothetical protein, conserved
TGME49_003570	14	290	acyltransferase, putative
TGME49_105050	3	20	calmodulin, putative
TGME49_018520	5	34	microneme protein MIC6
TGME49_109530	4	60	hypothetical protein
TGME49_013670	2	4	hypothetical protein

TGME49_087250	8	141	hypothetical protein
TGME49_121530	2	12	cathepsin L-like thiolproteinase, putative
TGME49_121420	9	74	kelch motif domain-containing protein
TGME49_074190	7	49	eukaryotic initiation factor 2B epsilon subunit, putative
TGME49_120050	5	55	ribosomal protein L5, putative
TGME49_065860	2	4	nit protein 2, putative
TGME49_111370	5	25	methylmalonate-semialdehyde dehydrogenase, putative
TGME49_009470	25	1167	hypothetical protein
TGME49_040980	4	8	hypothetical protein
TGME49_073370	11	122	coatomer gamma 2-subunit protein, putative
TGME49_047500	7	45	acyl-coenzyme A oxidase, putative
TGME49_014570	20	1356	hypothetical protein
TGME49_002680	6	38	mitochondrial-processing peptidase alpha subunit, putative
TGME49_095350	6	81	nucleoside diphosphate kinase, putative
TGME49_043950	4	16	prohibitin, putative
TGME49_083850	3	9	peptidyl-prolyl isomerase, putative
TGME49_058810	8	89	SRS27B
TGME49_022380	8	60	exportin 7, putative
TGME49_030350	3	9	hypothetical protein, conserved
TGME49_064610	2	4	heterogeneous nuclear ribonucleoprotein A3, putative
TGME49_032940	10	285	small heat shock protein 20
TGME49_069400	9	81	oxidoreductase, short chain dehydrogenase/reductase domain-containing protein
TGME49_023050	2	13	40s ribosomal protein S20, putative
TGME49_030450	5	25	GMP synthase, putative
TGME49_042570	2	4	hypothetical protein
TGME49_026730	35	2524	aconitate hydratase, putative
TGME49_064040	4	33	hypothetical protein, conserved
TGME49_026030	2	4	AGC kinase
TGME49_061750	12	128	hypothetical protein
TGME49_073760	42	3938	heat shock protein 70, putative
TGME49_110010	5	22	sushi domain-containing protein / SCR repeat-containing protein
TGME49_076140	4	13	ADP ribosylation factor 1
TGME49_049900	10	151	ADP/ATP carrier, putative
TGME49_114750	3	8	hypothetical protein, conserved
TGME49_057480	9	58	succinate-semialdehyde dehydrogenase, putative
TGME49_070950	17	1673	hypothetical protein
TGME49_035390	30	1243	PAN domain-containing protein
TGME49_052430	3	9	hypothetical protein, conserved
TGME49_076860	20	524	late embryogenesis abundant domain-containing protein
TGME49_116550	4	30	hypothetical protein
TGME49_105160	7	84	histone H2B, putative
TGME49_039740	15	188	hypothetical protein
TGME49_119530	2	4	splicing factor, putative
TGME49_091600	6	24	interferon gamma-inducible protein 30
TGME49_093320	2	4	hypothetical protein

TGME49_058820	3	16	hypothetical protein, conserved
TGME49_051810	5	46	translation initiation factor eIF-5A, putative
TGME49_086080	8	49	U5 small nuclear ribonucleoprotein, putative
TGME49_107850	12	238	6-phosphogluconate dehydrogenase, putative
TGME49_099070	2	4	pyruvate kinase, putative
TGME49_046080	6	91	hypothetical protein
TGME49_053150	45	4666	oxidoreductase, putative
TGME49_035970	8	97	eukaryotic translation initiation factor 2 gamma subunit, putative
TGME49_035340	3	9	hypothetical protein
TGME49_106930	3	9	proteasome subunit beta type 7, putative
TGME49_058410	3	9	hypothetical protein
TGME49_120220	11	170	ubiquinol-cytochrome c reductase domain-containing protein
TGME49_037250	4	33	hypothetical protein
TGME49_006470	6	28	dihydrolipoyl dehydrogenase, putative
TGME49_039490	11	234	mitochondrial branched-chain alpha-keto acid dehydrogenase E1,
TGME49_047330	3	9	hypothetical protein
TGME49_089000	2	4	hypothetical protein
TGME49_016000	22	1272	inner membrane complex protein IMC3
TGME49_109120	13	129	60s ribosomal protein L4, putative
TGME49_100260	9	78	threonyl-tRNA synthetase, putative
TGME49_094600	53	8917	hypothetical protein
TGME49_000310	3	8	hypothetical protein, conserved
TGME49_031850	9	163	serine-threonine phosophatase 2C
TGME49_090040	7	79	macrophage migration inhibitory factor, putative
TGME49_038010	6	46	ribosomal protein L23a, putative
TGME49_094670	1	4	eukaryotic translation initiation factor 3, putative
TGME49_086240	3	12	leucine-zipper-like transcriptional regulator 1, putative
TGME49_112090	7	60	60S ribosomal protein L23, putative
TGME49_097220	15	287	very long-chain acyl-CoA synthetase, putative
TGME49_055260	15	261	apical membrane antigen 1, putative
TGME49_108890	3	9	hypothetical protein, conserved
TGME49_050710	6	94	microneme protein 10
TGME49_049250	3	9	60s ribosomal protein L33-A, putative
TGME49_040870	8	53	beta adaptin protein, putative
TGME49_042070	4	25	cAMP-dependent protein kinase regulatory subunit, putative
TGME49_066270	14	234	2-methylbutyryl-CoA dehydrogenase, putative
TGME49_027620	4	8	28 kDa antigen
TGME49_113670	6	29	gamma-adaptin, putative
TGME49_071350	2	4	folylpolyglutamate synthase, putative
TGME49_042890	22	215	hypothetical protein
TGME49_110160	18	405	piwi-PAZ domain-containing protein
TGME49_078840	2	17	hypothetical protein
TGME49_100140	6	29	elongation factor 1-gamma, putative
TGME49_069860	11	262	hypothetical protein
TGME49_071210	19	400	4-alpha-glucanotransferase, putative
TGME49_032280	15	259	hypothetical protein

TGME49_109590	6	45	rhoptry protein, putative
TGME49_009720	3	13	hypothetical protein
TGME49_036890	2	4	hypothetical protein
TGME49_106060	83	7230	hypothetical protein
TGME49_063630	2	4	hypothetical protein
TGME49_097810	2	4	hypothetical protein, conserved
TGME49_063080	2	4	hypothetical protein
TGME49_081590	10	400	hypothetical protein
TGME49_010720	2	4	hypothetical protein
TGME49_014770	4	12	RAS small GTpase, putative
TGME49_025260	2	4	DNA-directed RNA polymerase II largest subunit, putative
TGME49_040600	2	4	TCP-1/cpn60 chaperonin family protein, putative
TGME49_069110	8	89	ornithine aminotransferase, putative
TGME49_088380	50	3199	heat shock protein 90
TGME49_031210	3	9	EH protein, putative
TGME49_115550	5	24	microneme protein, putative
TGME49_016620	4	13	EF hand domain-containing protein
TGME49_024280	3	9	hypothetical protein
TGME49_099050	4	25	60S ribosomal protein L17, putative
TGME49_003890	5	36	hypothetical protein
TGME49_110670	56	4583	glycogen phosphorylase family protein, putative
TGME49_051680	1	4	translationally-controlled tumor protein, putative
TGME49_021320	3	8	acetyl-CoA carboxylase, putative
TGME49_049030	4	13	endonuclease/exonuclease/phosphatase domain-containing protein
TGME49_059670	70	18857	von Willebrand factor type A domain-containing protein
TGME49_009980	26	1664	Rhoptry kinase family protein ROP42 (incomplete catalytic triad)
TGME49_032660	2	8	58 kDa phosphoprotein, putative
TGME49_069120	30	2795	oxidoreductase, putative
TGME49_054440	3	10	60S ribosomal protein L12, putative
TGME49_069950	16	332	hypothetical protein
TGME49_030520	2	4	cyclophilin, putative
TGME49_047350	2	4	thioredoxin, putative
TGME49_064000	6	54	hypothetical protein
TGME49_044110	2	4	nucleosome assembly protein, putative
TGME49_010690	5	50	40s ribosomal protein S6, putative
TGME49_098630	2	4	SPX domain-containing protein
TGME49_002530	2	4	aspartyl-tRNA synthetase, putative
TGME49_111890	7	29	hypothetical protein
TGME49_026960	25	766	phosphofructokinase, putative
TGME49_105810	2	4	hypothetical protein
TGME49_023700	3	9	LCCL domain-containing protein / F5/8 type C domain-containing protein
TGME49_043540	3	9	U4/U6 small nuclear ribonucleoprotein, putative
TGME49_110750	12	368	hypothetical protein, conserved
TGME49_090950	50	1629	clathrin heavy chain, putative
TGME49_020280	8	61	SCP-like domain-containing protein
TGME49_039480	4	8	RNB-like protein domain containing protein

TGME49_030360	2	4	hypothetical protein
TGME49_050880	1	4	adenosine kinase
TGME49_011320	7	28	C protein immunoglobulin-A-binding beta antigen
TGME49_003500	3	13	alanine dehydrogenase, putative
TGME49_063720	4	10	HMG box domain-containing protein
TGME49_023540	2	4	hypothetical protein, conserved
TGME49_073740	9	108	acetyl-CoA acyltransferase B, putative
TGME49_093430	11	118	hypothetical protein
TGME49_015910	10	62	hypothetical protein
TGME49_086580	28	1355	lysophospholipase, putative
TGME49_029360	9	90	transaldolase, putative
TGME49_049390	8	41	NAD-specific glutamate dehydrogenase, putative
TGME49_112110	4	13	nucleoredoxin, putative
TGME49_042420	2	4	amine oxidase, flavin-containing domain-containing protein
TGME49_066990	10	68	coatomer beta subunit, putative
TGME49_032830	8	36	vacuolar proton-translocating ATPase subunit, putative
TGME49_011090	2	4	cysteine desulfurase, putative
TGME49_042660	12	157	vacuolar sorting protein 35, putative
TGME49_068590	5	17	rhomboid-like protease 4
TGME49_004020	5	17	60S ribosomal protein L8, putative
TGME49_055190	8	57	myosin C
TGME49_116680	2	4	RNA pseudouridylate synthase domain-containing protein
TGME49_015290	16	654	saccharopine dehydrogenase, putative
TGME49_059010	5	78	vacuoloar ATP synthase subunit D, putative
TGME49_016860	7	48	ATP-dependent RNA helicase, putative
TGME49_094900	2	4	DnaK family domain containing protein
TGME49_017530	16	303	hypothetical protein
TGME49_029000	10	139	kelch motif domain-containing protein
TGME49_018240	2	16	hypothetical protein
TGME49_015060	2	4	small GTP-binding protein sar1, putative
TGME49_029930	4	19	p25-alpha domain-containing protein
TGME49_066070	2	9	60s ribosomal protein L31, putative
TGME49_008450	15	213	serine proteinase inhibitor, putative
TGME49_090290	15	147	ubiquitin-activating enzyme E1, putative
TGME49_109810	6	91	60S acidic ribosomal protein P2, putative
TGME49_021180	10	235	transmembrane domain-containing protein
TGME49_108860	4	24	hypothetical protein
TGME49_034450	5	29	40S ribosomal protein S15a, putative
TGME49_017680	5	29	hypothetical protein, conserved
TGME49_109540	4	35	hypothetical protein
TGME49_084540	4	8	ATP synthase subunit O, putative
TGME49_112950	20	902	hypothetical protein
TGME49_087040	9	147	hypothetical protein
TGME49_010700	2	4	hypothetical protein
TGME49_032250	11	103	peroxisomal catalase

TGME49_016670	10	69	KH domain-containing protein
TGME49_023940	10	226	hypothetical protein
TGME49_060540	3	9	hypothetical protein
TGME49_008370	11	98	myosin heavy chain, putative
TGME49_100200	1	4	histone H2A
TGME49_089970	1	6	hypothetical protein
TGME49_082050	2	4	hypothetical protein
TGME49_071890	9	53	3-ketoacyl-CoA reductase, putative
TGME49_091680	8	53	protein transport protein Sec23, putative
TGME49_025160	8	49	hypothetical protein
TGME49_021210	2	21	18 kDa cyclophilin
TGME49_012240	23	2230	tubulin beta chain, putative
TGME49_071610	2	4	pyrroline-5-carboxylase reductase, putative
TGME49_105010	2	4	RNA binding protein, putative
TGME49_032580	2	4	acetyl-coenzyme A synthetase, putative
TGME49_115610	2	8	hypothetical protein
TGME49_088810	2	4	hypothetical protein
TGME49_023670	2	4	inositol monophosphatase domain-containing protein
TGME49_065120	13	285	hypothetical protein, conserved
TGME49_112270	19	734	hypothetical protein
TGME49_040860	2	4	1-acyl-sn-glycerol-3-phosphate acyltransferase, putative
TGME49_007440	3	17	hypothetical protein
TGME49_031630	14	566	membrane skeletal protein IMC1, putative
TGME49_032170	7	34	hypothetical protein
TGME49_037080	17	372	hypothetical protein
TGME49_068830	6	36	hypothetical protein, conserved
TGME49_118410	4	13	TCP-1/cpn60 family chaperonin, putative
TGME49_119920	9	87	dihydrolipoamide branched chain transacylase, E2 subunit, putative
TGME49_036210	7	137	mitochondrial-processing peptidase beta subunit, putative
TGME49_087500	2	4	TCP-1/cpn60 family chaperonin, putative
TGME49_006690	5	37	hypothetical protein, conserved
TGME49_070930	2	4	hypothetical protein
TGME49_116340	2	4	hypothetical protein
TGME49_065370	2	4	hypothetical protein
TGME49_084560	5	49	60S ribosomal protein L9, putative
TGME49_024350	39	1542	aminopeptidase N, putative
TGME49_029990	3	9	TCP-1/cpn60 family chaperonin, putative
TGME49_027100	13	1020	glutaredoxin, putative
TGME49_026070	2	4	hypothetical protein, conserved
TGME49_063300	7	43	porin, putative
TGME49_014290	6	40	intracellular protease, putative
TGME49_054330	10	84	lipase domain-containing protein
TGME49_089930	3	9	phosphoenolpyruvate carboxykinase, putative
TGME49_088720	7	72	60S ribosomal protein L10, putative
TGME49_024890	6	27	hypothetical protein, conserved
TGME49_094400	6	40	hypothetical protein

TGME49_021620	3	9	tubulin beta chain, putative
TGME49_021470	13	123	hypothetical protein
TGME49_066460	3	9	ubiquitin-like protein SMT3 precursor, putative
TGME49_045550	2	4	hypothetical protein
TGME49_106390	2	4	hypothetical protein
TGME49_000320	5	44	hypoxanthine-xanthine-guanine phosphoribosyl transferase
TGME49_056970	16	319	vacuolar ATP synthase catalytic subunit A, putative
TGME49_031480	19	147	translational activator, putative
TGME49_069180	11	66	MIF4G domain-containing protein
TGME49_016640	8	50	homoserine kinase, putative
TGME49_016650	12	195	hypothetical protein, conserved
TGME49_018410	8	45	60S acidic ribosomal protein P0
TGME49_058830	2	4	hypothetical protein
TGME49_058210	2	4	DNA-directed RNA polymerase II subunit RPB2, putative
TGME49_069890	9	54	M16 family peptidase, putative
TGME49_066120	6	71	glutathione/thioredoxin peroxidase, putative
TGME49_035590	4	13	hypothetical protein
TGME49_121680	5	13	nuclear RNA binding protein, putative
TGME49_020360	4	13	FAD-depdendent monooxygenase, putative
TGME49_059660	2	4	orotate phosphoribosyltransferase, putative
TGME49_060310	3	8	ATP-binding cassette protein subfamily B member 1
TGME49_055420	16	369	hypothetical protein
TGME49_067330	10	65	fumarase, putative
TGME49_115670	2	4	HEAT repeat containing protein
TGME49_023840	2	4	ATP-citrate synthase, putative
TGME49_012300	6	29	hypothetical protein
TGME49_047240	2	4	ubiquitin carboxyl-terminal hydrolase isozyme L5, putative
TGME49_031900	8	36	acyl-CoA dehydrogenase, putative
TGME49_004340	9	42	hypothetical protein
TGME49_070250	4	57	dense granule protein 1 / major antigenp24
TGME49_057290	2	4	tRNA 2' phosphotransferase, putative
TGME49_044250	13	250	hypothetical protein
TGME49_077760	3	8	hypothetical protein, conserved
TGME49_089750	7	58	ubiquitin / ribosomal protein CEP52 fusion protein, putative
TGME49_043960	3	9	ras-GTPase-activating protein binding protein, putative
TGME49_009260	2	4	cytochrome c oxidase, putative
TGME49_115920	2	4	RNA polymerase II subunit RPB11, putative
TGME49_091960	8	53	Rhoptry kinase family protein ROP40 (incomplete catalytic triad)
TGME49_035020	7	52	coatomer protein complex subunit beta, putative
TGME49_024210	2	4	pseudouridylate synthase 1, putative
TGME49_031640	35	2638	membrane skeletal protein IMC1
TGME49_027810	12	189	Rhoptry kinase family protein ROP11 (incomplete catalytic triad)
TGME49_003720	5	42	vitamin K epoxide reductase complex subunit 1, putative
TGME49_094790	2	4	hypothetical protein
TGME49_056040	2	4	LCCL domain-containing protein
TGME49_094330	26	1012	EGF-like domain-containing protein

TGME49_054870	2	4	hypothetical protein
TGME49_065250	4	13	alpha-1 type II collagen, putative
TGME49_031350	4	13	glucosamine--fructose-6-phosphate aminotransferase (isomerizing), putative
TGME49_086120	4	13	prolyl endopeptidase, putative
TGME49_050770	10	148	eukaryotic translation initiation factor 4A
TGME49_046340	2	4	DnaJ domain-containing protein
TGME49_043710	4	16	TCP-1/cpn60 family chaperonin, putative
TGME49_031920	5	13	sepiapterin reductase, putative
TGME49_086780	51	2573	amiloride-sensitive amine oxidase, copper-containing, putative
TGME49_043250	6	29	myosin, putative
TGME49_109730	9	57	thioredoxin reductase, putative
TGME49_052630	2	4	hypothetical protein
TGME49_101400	2	4	hypothetical protein
TGME49_056770	7	44	ATP-dependent helicase, putaive
TGME49_101440	1	4	CAM kinase, CDPK family TgCDPK1
TGME49_025880	2	4	UDP-glucose 4-epimerase, putative
TGME49_048740	13	374	hypothetical protein
TGME49_048670	7	26	H+-translocating inorganic pyrophosphatase TVP, putative
TGME49_014940	9	90	MIC2-associated protein M2AP
TGME49_017890	15	614	peroxiredoxin
TGME49_009950	4	13	thioredoxin, putative
TGME49_053030	52	2880	alpha-glucosidase II, putative
TGME49_069190	2	4	glyceraldehyde-3-phosphate dehydrogenase
TGME49_065450	10	96	hexokinase
TGME49_061080	2	4	hypothetical protein, conserved
TGME49_120280	10	148	hypothetical protein
TGME49_043930	24	437	hypothetical protein
TGME49_054720	4	24	dense granule protein GRA8
TGME49_097720	3	9	trehalose-6-phosphate synthase domain-containing protein
TGME49_044850	6	45	GMC oxidoreductase, putative
TGME49_063730	4	8	glycerol-3-phosphate dehydrogenase, putative
TGME49_049480	2	4	TPR domain-containing protein
TGME49_048990	3	10	hypothetical protein
TGME49_032130	6	20	hypothetical protein, conserved
TGME49_023920	61	3846	hypothetical protein
TGME49_050820	7	28	hypothetical protein
TGME49_027970	3	10	DNA-binding protein HU, putative
TGME49_093690	5	57	profilin family protein
TGME49_033220	8	55	hypothetical protein
TGME49_063190	4	13	adenylosuccinate lyase, putative
TGME49_019230	10	141	acetoacetyl-CoA synthetase, putative
TGME49_101210	4	16	NAD(P) transhydrogenase, alpha subunit, putative
TGME49_120600	2	10	glycine-rich protein 2, putative
TGME49_073560	6	42	kinesin heavy chain, putative
TGME49_110420	2	4	hypothetical protein

TGME49_031160	2	4	hypothetical protein
TGME49_015460	2	4	40S ribosomal protein S24, putative
TGME49_084190	21	315	pyruvate carboxylase, putative
TGME49_022950	1	4	14 kDa phosphohistidine phosphatase, putative
TGME49_053820	6	40	hypothetical protein, conserved
TGME49_018820	4	18	hypothetical protein, conserved
TGME49_085980	5	12	phosphoglucomutase/parafusin related protein 1, putative
TGME49_052380	2	4	hypothetical protein
TGME49_112930	4	24	cystathionine beta-lyase, putative
TGME49_076870	10	276	LEA1 protein, putative
TGME49_007930	4	13	phosphatidylethanolamine-binding protein, putative
TGME49_109100	2	4	hypothetical protein
TGME49_046980	2	4	hypothetical protein
TGME49_116400	22	1164	tubulin alpha chain
TGME49_114280	2	4	hypothetical protein
TGME49_035200	80	16254	PAN domain-containing protein
TGME49_093360	4	18	hypothetical protein
TGME49_110000	3	17	lung seven transmembrane receptor domain-containing protein
TGME49_077570	32	1647	GPI transamidase subunit PIG-U, putative
TGME49_051910	15	180	hypothetical protein
TGME49_097280	6	55	hypothetical protein, conserved
TGME49_056990	3	9	glycyl-tRNA synthetase, putative
TGME49_036540	2	4	RRM domain-containing protein
TGME49_048880	4	13	Ras family domain-containing protein
TGME49_045680	2	4	ribosomal protein L21, putative
TGME49_015280	5	25	iron-sulfur subunit of succinate dehydrogenase, putative
TGME49_003240	27	915	hypothetical protein
TGME49_039500	2	4	proteasome subunit alpha type 4, subunit
TGME49_058160	4	16	hypothetical protein
TGME49_034570	18	582	peroxisomal multifunctional enzyme type 2, putative
TGME49_089080	2	4	adaptin N terminal region domain-containing protein
TGME49_032050	4	16	DnaJ domain-containing protein
TGME49_066760	8	80	isocitrate dehydrogenase, putative
TGME49_045490	4	17	microneme protein 8
TGME49_121500	5	16	CELF family protein, putative
TGME49_073360	2	4	ABC transporter, putative
TGME49_090890	5	21	carbonyl reductase, putative
TGME49_026250	8	52	ATP-dependent RNA helicase, putative
TGME49_090580	2	4	ABC transporter, putative
TGME49_006670	3	9	hypothetical protein
TGME49_081580	2	4	protein phosphatase 2C, putative
TGME49_029480	4	16	membrane-associated calcum-binding protein, related
TGME49_117720	4	17	eukaryotic translation initiation factor 3 subunit 7, putative
TGME49_099780	5	20	hypothetical protein
TGME49_073460	6	34	eukaryotic translation initiation factor 3 subunit 6 interacting protein, putative
TGME49_093390	2	4	cytochrome c heme lyase, putative

TGME49_007170	2	4	hypothetical protein
TGME49_086750	9	88	MA3 domain protein
TGME49_024870	2	4	hypothetical protein
TGME49_049650	5	25	apolipoprotein A-I binding protein, putative
TGME49_086450	23	1287	dense granule protein 5 precursor
TGME49_007840	4	16	40S ribosomal protein S17, putative
TGME49_023390	13	121	sec63 domain-containing DEAD/DEAH box helicase, putative
TGME49_010780	2	4	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_048810	6	26	hypothetical protein
TGME49_058460	2	4	hypothetical protein
TGME49_112640	2	4	hypothetical protein
TGME49_058370	9	82	Rhoptry kinase family protein ROP28
TGME49_052220	3	16	Hsc70/Hsp90-organizing protein, putative
TGME49_075440	8	106	granule antigen protein GRA6
TGME49_090700	6	116	hypothetical protein
TGME49_086250	20	729	oocyst wall protein, putative
TGME49_025850	15	158	hypothetical protein
TGME49_112480	5	36	uracil phosphoribosyltransferase
TGME49_010260	15	254	glycerol-3-phosphate dehydrogenase, putative
TGME49_120530	4	13	hypothetical protein
TGME49_054430	16	295	microneme protein, putative
TGME49_019320	21	1032	acid phosphatase, putative
TGME49_120460	2	4	ABC transporter protein, putative
TGME49_097500	3	17	TCP-1/cpn60 family chaperonin, putative
TGME49_026430	5	56	reticulon domain-containing protein
TGME49_078830	63	7288	glucose-6-phosphate dehydrogenase, putative
TGME49_115490	7	33	hypothetical protein
TGME49_009160	2	4	glycylpeptide N-tetradecanoyltransferase, putative
TGME49_061240	6	45	histone H3
TGME49_033300	2	4	rhoGAP protein, putative
TGME49_022860	5	32	eukaryotic translation initiation factor 3 subunit 9, putative
TGME49_013680	4	16	MmgE/PrpD family domain containing protein
TGME49_094550	16	165	dynein heavy chain, putative
TGME49_027950	5	18	zinc metalloprotease 2, putative
TGME49_030960	10	93	splicing factor 3B subunit 3, putative
TGME49_003970	4	16	dolichyl-di-phosphooligosaccharide-protein glycotransferase, putative
TGME49_063180	25	1022	myosin D, putative
TGME49_030690	2	4	lipin, putative
TGME49_114970	8	33	hypothetical protein
TGME49_089690	33	2094	glyceraldehyde-3-phosphate dehydrogenase
TGME49_009030	31	3146	actin
TGME49_044560	32	1638	heat shock protein 90, putative
TGME49_087210	3	8	proteasome subunit alpha type 2, putative
TGME49_060140	4	13	mannose-6-phosphate isomerase, putative
TGME49_029670	3	10	ribosomal protein S23, putative
TGME49_030420	12	127	calcium-transporting ATPase, putative

TGME49_060260	1	8	60S acidic ribosomal protein P1, putative
TGME49_058870	5	21	hypothetical protein
TGME49_076880	9	318	late embryogenesis abundant domain-containing protein
TGME49_115320	14	181	SRS52A
TGME49_069870	2	4	hypothetical protein, conserved
TGME49_028170	44	2110	serine/threonine protein phosphatase, putative
TGME49_065010	2	4	glutamate 5-kinase, putative
TGME49_031960	3	8	ppg3, putative
TGME49_033100	3	9	prohibitin, putative
TGME49_047550	17	389	heat shock protein 60
TGME49_047220	9	78	hypothetical protein
TGME49_069440	3	9	calmodulin, putative
TGME49_007240	3	9	hypothetical protein
TGME49_100310	2	4	26S protease regulatory subunit 6a, putative
TGME49_057380	4	54	hypothetical protein
TGME49_005130	5	18	glyoxalase, putative
TGME49_063870	6	36	glutamyl-tRNA synthetase, putative
TGME49_076990	2	4	cytochrome b5-like Heme/Steroid binding domain containing protein
TGME49_078120	10	102	SCP-like extracellular domain-containing protein
TGME49_057160	2	4	hypothetical protein
TGME49_064440	4	13	signal recognition particle receptor subunit beta, putative
TGME49_067980	3	9	hypothetical protein
TGME49_118520	14	267	hypothetical protein
TGME49_067810	2	4	Rab 5
TGME49_101420	4	13	hypothetical protein, conserved
TGME49_011710	3	9	TB2/DP1, HVA22 domain-containing protein
TGME49_097910	7	43	hypothetical protein
TGME49_027030	3	9	hypothetical protein
TGME49_086040	6	40	hypothetical protein
TGME49_009920	55	8413	PAN domain-containing protein
TGME49_005360	7	17	hypothetical protein
TGME49_005250	20	937	Rhoptry kinase family protein ROP18
TGME49_109030	2	4	hypothetical protein, conserved
TGME49_020270	8	100	hypothetical protein, conserved
TGME49_020510	4	13	hypothetical protein
TGME49_063700	9	60	40s ribosomal protein S14, putative
TGME49_062040	8	53	eukaryotic translation initiation factor 3 subunit 11, putative
TGME49_035140	4	8	hypothetical protein
TGME49_029030	3	10	hypothetical protein, conserved
TGME49_079450	2	4	adenylosuccinate synthetase, putative
TGME49_067410	2	4	scavenger receptor protein TgSR1, putative
TGME49_044370	3	12	hypothetical protein
TGME49_079400	3	24	glutaredoxin, putative
TGME49_087980	3	9	forkhead-aasociated domain-containing protein
TGME49_088000	4	16	hypothetical protein

TGME49_113080	8	50	hypothetical protein
TGME49_013280	22	3460	hypothetical protein
TGME49_067430	3	15	DnaJ domain-containing protein
TGME49_081620	2	4	aldo/keto reductase family oxidoreductase, putative
TGME49_027890	2	4	hypothetical protein
TGME49_064830	10	88	hypothetical protein
TGME49_086770	5	28	hypothetical protein, conserved
TGME49_038160	3	34	hypothetical protein
TGME49_024850	10	89	polyadenylate-binding protein, putative
TGME49_118160	5	26	major sperm protein domain-containing protein
TGME49_067710	2	4	hypothetical protein
TGME49_067740	4	36	hypothetical protein
TGME49_070380	6	50	40S ribosomal protein S13, putative
TGME49_043800	6	20	long-chain fatty acid CoA ligase, putative
TGME49_043600	5	20	acetyltransferase domain-containing protein
TGME49_005570	2	4	hypothetical protein, conserved
TGME49_059900	46	2122	hypothetical protein, conserved
TGME49_055460	12	242	EGF-like domain-containing protein
TGME49_072240	15	407	hypothetical protein
TGME49_067680	58	7503	microneme protein, putative
TGME49_080490	6	36	ubiquitin conjugation factor, putative
TGME49_067400	3	13	ribosomal protein L32, putative
TGME49_079420	11	91	hypothetical protein
TGME49_014220	3	16	hypothetical protein
TGME49_026830	13	148	DnaK family protein
TGME49_022160	23	346	aldehyde dehydrogenase, putative
TGME49_067660	4	16	hypothetical protein
TGME49_090300	4	39	hypothetical protein
TGME49_078050	3	9	proteasome subunit alpha type 1, putative
TGME49_051690	3	9	seryl-tRNA synthetase, putative
TGME49_013730	5	40	lanthionine synthetase C-like protein 1, putative
TGME49_010090	4	25	Rhoptry kinase family protein ROP43 (incomplete catalytic triad)
TGME49_000230	25	1407	PAN domain-containing protein
TGME49_075460	2	4	hypothetical protein
TGME49_000250	4	13	PAN domain-containing protein
TGME49_005320	2	4	hypothetical protein
TGME49_079100	9	112	hypothetical protein
TGME49_075300	10	167	Rhoptry kinase family protein ROP2B (incomplete catalytic triad)
TGME49_015780	26	2895	Rhoptry kinase family protein ROP2A (incomplete catalytic triad)
TGME49_121700	9	79	hypothetical protein
TGME49_080170	13	200	hypothetical protein
TGME49_061250	10	324	histone H2A, putative
TGME49_010110	10	368	Rhoptry kinase family protein ROP44 (incomplete catalytic triad)

Table S2: Identification of 997 proteins in the oocyst/sporocyst proteome of *Toxoplasma gondii*

[Gene ID]	[Number of Peptide Sequences]	[Number of Spectra]	[Product Description]
TGME49_096950	21	360	hypothetical protein
TGGT1_097200	6	69	conserved hypothetical protein
TGGT1_097180	3	51	conserved hypothetical protein
TGME49_023150	2	2	hypothetical protein
TGGT1_097190	1	3	conserved hypothetical protein
TGGT1_037720	12	311	conserved hypothetical protein
TGME49_086300	11	278	hypothetical protein
TGGT1_037710	1	2	conserved hypothetical protein
TGME49_003140	2	4	dynein beta chain, flagellar outer arm, putative
TGME49_034920	6	142	hypothetical protein
TGGT1_071190	4	96	conserved hypothetical protein
TGGT1_071200	36	2060	PHD-finger domain-containing protein, putative
TGME49_034700	25	1141	PHD-finger domain-containing protein
TGME49_058970	67	2171	hypothetical protein
TGGT1_010900	28	992	conserved hypothetical protein
TGME49_058960	15	89	hypothetical protein
TGGT1_010910	6	60	hypothetical protein
TGGT1_032690	2	6	conserved hypothetical protein
TGME49_090460	2	6	hypothetical protein
TGME49_078090	1	12	hypothetical protein
TGME49_078100	1	12	hypothetical protein
TGME49_106020	92	11684	hypothetical protein
TGGT1_041340	8	1058	hypothetical protein
TGGT1_041330	2	316	hypothetical protein
TGME49_009790	13	176	radical SAM domain-containing protein
TGGT1_022830	1	3	Radical SAM domain-containing protein, putative
TGME49_014230	11	251	dopey, N-terminal domain-containing protein
TGGT1_124550	10	232	hypothetical protein
TGGT1_124530	2	4	hypothetical protein
TGME49_112300	11	283	TBC domain-containing protein
TGGT1_088550	7	182	TBC domain-containing protein, putative
TGME49_097470	26	2439	myosin light chain 2, putative
TGGT1_096740	3	95	hypothetical protein
TGME49_112610	15	155	hypothetical protein
TGME49_112620	14	717	hypothetical protein, conserved
TGGT1_088910	12	609	conserved hypothetical protein
TGGT1_088900	9	139	conserved hypothetical protein
TGME49_101450	15	171	T-cell immunomodulatory protein, putative
TGGT1_059900	3	29	conserved hypothetical protein
TGGT1_059920	1	2	conserved hypothetical protein
TGGT1_036630	5	48	conserved hypothetical protein
TGME49_084920	4	45	hypothetical protein

TGME49_085040	2	28	hypothetical protein
TGME49_004440	3	181	cyclic nucleotide-binding domain containing protein
TGGT1_021150	5	46	multidrug resistance-associated protein, putative
TGME49_008050	3	27	ABC transporter, putative
TGME49_008060	1	5	ABC transporter, putative
TGGT1_021160	1	2	ABC transporter, putative
TGGT1_110880	3	35	conserved hypothetical protein
TGME49_071880	3	34	B-box zinc finger protein, putative
TGGT1_110890	1	4	B-box zinc finger protein, putative
TGGT1_046220	1	1	conserved hypothetical protein
TGME49_043650	1	1	hypothetical protein
TGGT1_048910	4	15	conserved hypothetical protein
TGME49_095480	1	4	hypothetical protein
TGME49_058110	4	25	N-acetylglucosaminylphosphatidylinositol deacetylase, putative
TGGT1_012100	2	14	N-acetylglucosaminylphosphatidylinositol deacetylase, putative
TGME49_003320	32	1237	hypothetical protein
TGGT1_017530	29	1521	ubiquitin ligase E3 alpha, putative
TGGT1_014280	5	50	conserved hypothetical protein
TGME49_055960	5	37	hypothetical protein
TGME49_006430	23	1296	formin homology 2 domain-containing protein
TGGT1_062340	3	100	conserved hypothetical protein
TGGT1_062350	1	47	conserved hypothetical protein
TGGT1_062360	1	42	hypothetical protein
TGGT1_002260	3	10	conserved hypothetical protein
TGGT1_002150	2	23	conserved hypothetical protein
TGME49_027270	2	12	hypothetical protein
TGGT1_082660	2	10	conserved hypothetical protein
TGME49_085830	13	212	hypothetical protein
TGGT1_037230	4	24	conserved hypothetical protein
TGGT1_125320	1	2	hypothetical protein
TGGT1_117800	1	1	DNA polymerase theta, putative
TGME49_013390	48	2814	plectin, putative
TGGT1_039260	9	299	conserved hypothetical protein
TGGT1_039270	2	210	hypothetical protein
TGME49_045430	1	2	hypothetical protein
TGME49_041190	10	71	hypothetical protein
TGGT1_049360	8	46	conserved hypothetical protein
TGME49_041300	1	1	hypothetical protein
TGME49_078130	13	543	hypothetical protein
TGGT1_104420	2	12	conserved hypothetical protein
TGGT1_104430	1	21	conserved hypothetical protein
TGME49_088700	3	18	chromosome segregation protein smc1, putative
TGGT1_034570	1	6	conserved hypothetical protein

TGGT1_034690	1	3	structural maintenance of chromosomes smc1, putative
TGGT1_093550	25	928	EF hand domain-containing protein, putative
TGME49_115860	19	523	EF hand domain-containing protein
TGME49_115870	1	9	hypothetical protein
TGME49_003540	9	194	RNA binding protein, putative
TGGT1_017290	8	207	RNA binding protein, putative
TGME49_036200	1	1	DNA mismatch repair protein, putative
TGGT1_007110	4	10	conserved hypothetical protein
TGME49_062880	4	6	hypothetical protein
TGME49_077910	15	174	thrombospondin type 1 domain-containing protein
TGGT1_104740	8	136	thrombospondin type 1 domain-containing protein, putative
TGME49_011250	9	83	hypothetical protein
TGGT1_123670	8	81	conserved hypothetical protein
TGME49_100100	130	68347	rhoptry neck protein 2
TGGT1_095290	35	3562	conserved hypothetical protein
TGME49_052180	1	1	hypothetical protein
TGGT1_088960	7	212	conserved hypothetical protein
TGME49_112660	7	212	hypothetical protein
TGGT1_004130	2	5	histone acetyltransferase GCN5, putative
TGME49_054550	1	4	histone acetyltransferase GCN5
TGME49_089050	24	2034	FIK kinase (incomplete catalytic triad)
TGGT1_034200	22	1942	serine/threonine-protein kinase, putative
TGGT1_034210	3	54	hypothetical protein
TGGT1_063400	15	435	conserved hypothetical protein
TGME49_005500	12	305	hypothetical protein, conserved
TGME49_087840	14	428	hypothetical protein
TGME49_087950	13	349	virulent strain associated lipoprotein
TGME49_018720	4	25	CAM kinase, CDPK family
TGGT1_029570	2	8	calcium-dependent protein kinase, putative
TGGT1_029680	1	1	hypothetical protein
TGGT1_108930	1	2	DNA replication licensing factor MCM2, putative
TGGT1_010820	1	2	hypothetical protein
TGME49_059080	1	2	hypothetical protein
TGGT1_010810	1	1	hypothetical protein
TGME49_088420	2	10	hypothetical protein
TGME49_057330	2	8	hypothetical protein
TGGT1_013470	1	3	conserved hypothetical protein
TGGT1_053330	2	10	conserved hypothetical protein
TGME49_010270	2	10	hypothetical protein, conserved
TGME49_078440	12	163	transcription regulatory protein SNF2, putative
TGGT1_104110	11	146	transcription regulatory protein SNF2, putative
TGME49_050690	23	1545	zinc finger (CCCH type) protein, putative
TGGT1_102120	11	509	hypothetical protein
TGGT1_102130	3	94	hypothetical protein

TGGT1_122010	8	159	conserved hypothetical protein
TGME49_119090	3	35	hypothetical protein
TGME49_044040	10	191	hypothetical protein, conserved
TGGT1_045910	2	8	hypothetical protein
TGME49_029140	19	616	peroxisomal multifunctional enzyme, putative
TGGT1_118690	3	72	conserved hypothetical protein
TGGT1_102640	1	1	conserved hypothetical protein
TGME49_026410	49	4323	elongation factor 1-beta, putative
TGGT1_081580	15	1986	elongation factor 1-beta, putative
TGME49_038410	3	9	hypothetical protein, conserved
TGME49_116730	7	53	hypothetical protein
TGGT1_094570	1	9	hypothetical protein
TGGT1_000850	2	6	conserved hypothetical protein
TGME49_052060	2	6	hypothetical protein
TGGT1_038710	16	1693	hypothetical protein
TGME49_013790	14	1267	hypothetical protein
TGGT1_024540	22	5152	conserved hypothetical protein
TGME49_048420	18	3712	hypothetical protein, conserved
TGME49_048410	6	1282	hypothetical protein
TGME49_046690	6	23	1,4-alpha-glucan branching enzyme, putative
TGGT1_026420	3	15	1,4-alpha-glucan branching enzyme, putative
TGGT1_026410	1	1	conserved hypothetical protein
TGGT1_026860	7	82	conserved hypothetical protein
TGME49_046190	6	59	hypothetical protein, conserved
TGME49_114410	42	1629	hypothetical protein
TGGT1_090990	20	713	conserved hypothetical protein
TGGT1_091200	7	80	conserved hypothetical protein
TGME49_077940	9	223	hypothetical protein
TGGT1_104710	3	41	hypothetical protein
TGGT1_104590	1	16	hypothetical protein
TGME49_033960	4	34	hypothetical protein
TGME49_120670	5	8	hypothetical protein, conserved
TGGT1_120610	2	2	vcl1, putative
TGME49_105590	23	413	ABC transporter, putative
TGGT1_041890	12	413	ATP-binding cassette, sub-family C, member 1
TGGT1_122990	17	240	WD-repeat protein, putative
TGME49_118260	10	113	hypothetical protein
TGME49_118250	3	42	transcription initiation factor TFIID subunit 5, putative
TGME49_005620	8	125	hypothetical protein
TGGT1_063300	1	17	conserved hypothetical protein
TGME49_080540	30	688	hypothetical protein
TGGT1_072770	23	667	conserved hypothetical protein
TGGT1_126660	2	3	cytidine and deoxycytidylate deaminase zinc-binding region domain-containing protein, putative

TGME49_015750	1	1	cytidine and deoxycytidylate deaminase zinc-binding region domain-containing protein
TGGT1_097910	9	119	zinc finger (C3HC4 RING finger) protein, putative
TGME49_017220	7	76	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_005680	2	2	conserved hypothetical protein
TGME49_081790	2	2	Rhoptry kinase family protein ROP45 (incomplete catalytic triad)
TGGT1_015370	2	6	conserved hypothetical protein
TGME49_040840	2	4	hypothetical protein
TGGT1_125710	7	41	conserved hypothetical protein
TGME49_015080	2	7	hypothetical protein
TGME49_011010	2	4	hypothetical protein
TGME49_111290	6	108	protein tyrosine phosphatase-like domian-containing protein
TGGT1_083330	6	82	conserved hypothetical protein
TGME49_028070	2	48	hypothetical protein
TGME49_108710	46	4259	hypothetical protein, conserved
TGGT1_055640	19	2609	conserved hypothetical protein
TGME49_108600	11	884	hypothetical protein
TGGT1_055650	3	66	ankyrin repeat domain protein, putative
TGME49_108990	2	4	hypothetical protein
TGGT1_056030	1	3	conserved hypothetical protein
TGME49_109000	1	2	sodium:solute symporter family domain-containing protein
TGME49_024980	20	680	hypothetical protein
TGGT1_079910	4	51	hypothetical protein
TGGT1_105590	1	1	enoyl-CoA hydratase, putative
TGME49_117700	1	1	enoyl-CoA hydratase/isomerase family domain-containing protein
TGME49_021330	13	75	DNA gyrase subunit A, putative
TGME49_121660	3	9	hypothetical protein
TGGT1_064700	1	18	glycosyltransferase, putative
TGME49_057080	18	1093	cAMP-specific 3',5'-cyclic phosphodiesterase, putative
TGGT1_013630	7	460	hypothetical protein
TGGT1_013640	2	62	high-affinity cGMP-specific 3',5'-cyclic phosphodiesterase, putative
TGME49_105250	4	66	hypothetical protein
TGGT1_042260	3	46	conserved hypothetical protein
TGME49_118880	5	68	hypothetical protein
TGGT1_122320	5	26	conserved hypothetical protein
TGGT1_096240	15	1617	conserved hypothetical protein
TGME49_097870	14	1609	hypothetical protein
TGME49_097860	3	7	hypothetical protein
TGME49_121630	5	9	hypothetical protein
TGME49_063060	63	5740	26S proteasome non-ATPase regulatory subunit 1, putative
TGGT1_043150	31	2716	26S proteasome non-ATPase regulatory subunit, putative

TGME49_058740	23	604	eukaryotic translation initiation factor 2A, putative
TGGT1_011450	16	537	eukaryotic translation initiation factor 2A, putative
TGGT1_126560	5	117	zinc finger (C3HC4 type, RING finger) protein
TGME49_015640	3	24	zinc finger (C3HC4 type RING finger) protein, putative
TGME49_015630	2	90	hypothetical protein
TGME49_030490	13	759	phosphatidylinositol-4-phosphate 5-kinase, putative
TGGT1_117470	12	870	phosphatidylinositol-4-phosphate 5-kinase, putative
TGGT1_065930	24	896	conserved hypothetical protein
TGME49_022100	24	681	hypothetical protein
TGME49_009450	12	208	TLD domain-containing protein
TGGT1_022390	9	206	TLD domain-containing protein, putative
TGME49_018760	16	641	alpha-soluble NSF attachment protein, putative
TGGT1_029730	13	487	soluble nsf attacment protein, putative
TGME49_032420	3	19	hypothetical protein
TGME49_072910	46	3746	TCP-1/cpn60 family chaperonin, putative
TGGT1_112090	23	2609	chaperonin containing t-complex protein 1, delta subunit, tcpd, putative
TGME49_033810	12	97	hypothetical protein
TGGT1_113520	9	118	conserved hypothetical protein
TGME49_038250	11	395	60s ribosomal protein L36, putative
TGGT1_052600	3	63	60S ribosomal protein L36, putative
TGME49_111360	57	4740	AGC kinase TgPKG1
TGGT1_087710	30	1839	cGMP-dependent protein kinase, putative
TGGT1_125260	8	44	conserved hypothetical protein
TGME49_014830	7	33	hypothetical protein
TGME49_003790	3	7	hypothetical protein
TGGT1_016940	2	6	conserved hypothetical protein
TGGT1_017430	5	31	conserved hypothetical protein
TGME49_003400	5	19	hypothetical protein
TGME49_032260	5	58	hypothetical protein
TGGT1_115380	4	31	conserved hypothetical protein
TGGT1_056070	12	540	conserved hypothetical protein
TGME49_109040	12	540	hypothetical protein
TGGT1_075380	2	5	conserved hypothetical protein
TGME49_095030	2	5	hypothetical protein
TGME49_058060	20	1695	hypothetical protein, conserved
TGGT1_012150	9	218	conserved hypothetical protein
TGME49_105770	8	47	ABC transporter, putative
TGGT1_041810	6	43	ABC transporter, putative
TGME49_086790	16	1055	pinin/SDK/memA/ domain-containing protein
TGGT1_038180	15	1377	pinin/SDK/memA/ domain-containing protein, putative
TGGT1_092750	1	1	conserved hypothetical protein
TGGT1_013850	16	845	DEAD/DEAH box helicase, putative
TGME49_056900	7	294	DEAD/DEAH box helicase, putative

TGGT1_050550	7	98	conserved hypothetical protein
TGME49_040300	7	98	hypothetical protein, conserved
TGGT1_007560	3	30	calcium/calmodulin-dependent protein kinase, putative
TGME49_062540	2	15	CAM kinase
TGGT1_066430	1	3	conserved hypothetical protein
TGME49_004100	19	286	Wee kinase
TGGT1_016600	8	249	eukaryotic translation initiation factor 2-alpha kinase, putative
TGGT1_071560	7	160	hypothetical protein
TGME49_034350	6	99	hypothetical protein
TGGT1_087140	3	18	step II splicing factor SLU7, putative
TGME49_110820	2	9	step II splicing factor SLU7, putative
TGGT1_028860	1	1	conserved hypothetical protein
TGME49_001710	10	137	U3 small nucleolar ribonucleoprotein complex-associated protein, putative
TGGT1_019720	8	134	U3 small nucleolar ribonucleoprotein complex-associated protein, putative
TGGT1_064500	55	6209	conserved hypothetical protein
TGME49_121450	46	4602	myb-like DNA-binding domain-containing protein
TGME49_021590	6	27	dual-specificity phosphatase, catalytic domain-containing protein
TGGT1_065460	3	11	dual-specificity phosphatase, catalytic domain-containing protein, putative
TGGT1_104100	2	3	conserved hypothetical protein
TGME49_078450	1	2	hypothetical protein
TGGT1_100490	19	899	phospholipase C, putative
TGME49_048830	17	869	phospholipase C delta 1
TGME49_064080	16	383	acyl carrier protein
TGGT1_044320	8	239	acyl carrier protein, putative
TGME49_007720	28	734	hypothetical protein
TGGT1_020600	22	811	nucleolar RNA-associated protein, putative
TGGT1_087330	2	4	conserved hypothetical protein
TGME49_111020	1	1	hypothetical protein
TGME49_045750	1	9	WD domain-containing protein
TGGT1_116540	15	275	structural maintenance of chromosomes smc4, putative
TGME49_031170	12	161	chromosome condensation protein, putative
TGGT1_032160	9	231	starch binding domain-containing protein, putative
TGME49_090960	8	215	phosphoglucan, water dikinase protein, putative
TGME49_114090	33	1338	proteasome subunit beta type 3, putative
TGGT1_090730	7	371	proteasome subunit beta type, putative
TGME49_092130	19	2254	60s ribosomal protein L13a, putative
TGGT1_031440	12	555	60S ribosomal protein L13A, putative
TGME49_088650	88	537471	hypothetical protein
TGGT1_034740	23	3393	conserved hypothetical protein
TGME49_091930	29	1786	nucleolar phosphoprotein nucleolin, putative

TGGT1_031630	20	1794	nucleolar phosphoprotein, putative
TGME49_057990	67	6593	heat shock protein, putative
TGGT1_012210	43	5326	heat shock protein, putative
TGME49_073490	39	1634	glutamine synthetase, putative
TGGT1_112570	23	1160	glutamine synthetase, putative
TGGT1_028970	11	311	immunoglobulin-binding protein, putative
TGME49_018250	11	242	immunoglobulin-binding protein, putative
TGME49_030850	26	1707	mucin, putative
TGGT1_116990	9	359	mucin, putative
TGME49_006510	80	9243	peptidase M16 domain containing protein
TGGT1_062250	39	4108	metalloprotease, putative
TGGT1_069810	1	5	conserved hypothetical protein
TGME49_035960	1	5	hypothetical protein
TGME49_085210	9	230	hypothetical protein
TGGT1_036710	7	226	conserved hypothetical protein
TGME49_060390	12	346	hypothetical protein
TGGT1_009820	11	371	conserved hypothetical protein
TGGT1_069290	1	3	conserved hypothetical protein
TGME49_036550	1	3	hypothetical protein
TGGT1_006060	11	230	pyrophosphate-fructose 6-phosphate 1-phosphotransferase beta subunit, putative
TGME49_081390	8	183	pyrophosphate-fructose 6-phosphate 1-phosphotransferase beta subunit, putative
TGME49_082170	8	38	hypothetical protein
TGGT1_005300	5	32	conserved hypothetical protein
TGME49_097970	21	460	aspartyl aminopeptidase, putative
TGGT1_096140	16	369	aspartyl aminopeptidase, putative
TGME49_046620	5	24	G10 protein, putative
TGGT1_026480	4	20	conserved hypothetical protein
TGGT1_065550	3	16	conserved hypothetical protein
TGME49_021680	3	15	hypothetical protein
TGME49_057350	43	5803	eukaryotic translation intiation factor, putative
TGGT1_013450	32	6254	eukaryotic initiation factor 4f, putative
TGME49_030410	26	1692	peroxiredoxin 3
TGGT1_117560	8	212	peroxiredoxin, putative
TGGT1_097000	3	26	conserved hypothetical protein
TGME49_097230	3	26	hypothetical protein, conserved
TGME49_064480	3	30	hypothetical protein, conserved
TGGT1_116480	6	470	conserved hypothetical protein
TGME49_031220	3	186	hypothetical protein
TGME49_019800	46	4387	vacuolar ATP synthase subunit B, putative
TGGT1_030800	25	2828	vacuolar ATP synthase subunit B, putative
TGME49_053640	3	27	hypothetical protein
TGGT1_002810	3	23	conserved hypothetical protein
TGME49_032020	8	131	hypothetical protein
TGGT1_115740	3	17	conserved hypothetical protein

TGME49_088570	15	736	hypothetical protein
TGGT1_034820	10	706	conserved hypothetical protein
TGGT1_120680	3	27	conserved hypothetical protein
TGME49_120610	3	27	hypothetical protein
TGME49_047450	62	17291	hypothetical protein, conserved
TGGT1_025460	59	18046	conserved hypothetical protein
TGGT1_030880	24	1125	DNA replication licensing factor, putative
TGME49_019870	24	1028	DNA replication licensing factor, putative
TGME49_053750	204	47453	hypothetical protein
TGGT1_002920	141	44739	conserved hypothetical protein
TGGT1_009590	1	3	conserved hypothetical protein
TGME49_060590	1	3	hypothetical protein, conserved
TGGT1_104900	2	10	conserved hypothetical protein
TGME49_077750	2	10	hypothetical protein
TGME49_101120	17	330	acetyil CoA acetyltransferase/thiolase, putative
TGGT1_059540	10	182	acetyl-CoA acetyltransferase, putative
TGME49_032090	10	309	3-hydroxybutyryl-CoA dehydrogenase, putative
TGGT1_115670	9	268	3-hydroxybutyryl-CoA dehydrogenase, putative
TGME49_118580	46	1819	phosphoglucomutase, putative
TGGT1_122530	22	1170	phosphoglucomutase, putative
TGGT1_066380	1	3	conserved hypothetical protein
TGME49_022870	1	3	hypothetical protein
TGGT1_014300	1	2	conserved hypothetical protein
TGME49_055920	1	2	hypothetical protein
TGME49_026580	12	150	hypothetical protein
TGGT1_081730	9	274	conserved hypothetical protein
TGME49_003520	3	21	hypothetical protein
TGGT1_050370	5	98	conserved hypothetical protein
TGME49_040500	4	80	hypothetical protein
TGGT1_093860	1	1	protein kinase domain-containing protein, putative
TGME49_116150	1	1	ULK kinase
TGME49_113210	2	4	hypothetical protein
TGME49_077000	39	2456	transport protein Sec24, putative
TGGT1_105320	22	1763	hypothetical protein
TGGT1_110070	1	2	conserved hypothetical protein
TGME49_071200	1	2	hypothetical protein
TGGT1_024960	2	8	conserved hypothetical protein
TGME49_047960	2	7	hypothetical protein
TGGT1_113480	3	21	conserved hypothetical protein
TGME49_033850	3	21	hypothetical protein
TGME49_105850	33	1409	RNA binding protein, putative
TGGT1_041620	14	844	RNA binding protein, putative
TGGT1_090970	13	527	conserved hypothetical protein
TGME49_114390	5	195	hypothetical protein
TGME49_070540	5	47	hypothetical protein

TGME49_030150	21	221	hypothetical protein
TGGT1_117750	8	164	conserved hypothetical protein
TGGT1_016150	20	4293	hypothetical protein
TGME49_004530	15	2077	microneme protein MIC11
TGGT1_024830	10	407	conserved hypothetical protein
TGME49_048110	7	226	hypothetical protein, conserved
TGME49_111300	2	2	cAMP-dependent protein kinase regulatory subunit, putative
TGME49_025540	20	204	hypothetical protein
TGGT1_080770	6	63	conserved hypothetical protein
TGGT1_018610	3	16	conserved hypothetical protein
TGME49_002480	3	16	hypothetical protein
TGME49_113470	2	4	helix-hairpin-helix motif-containing protein
TGME49_027420	6	24	lytB domain-containing protein
TGGT1_082920	2	8	lytB domain-containing protein, putative
TGME49_062720	7	120	eukaryotic translation initiation factor 1A, putative
TGGT1_007370	5	105	eukaryotic translation initiation factor 1A, putative
TGGT1_109740	2	24	conserved hypothetical protein
TGME49_070890	1	9	hypothetical protein
TGGT1_116710	10	462	transcription initiation factor iie, alpha subunit, putative
TGME49_031010	6	193	hypothetical protein
TGGT1_016850	3	26	conserved hypothetical protein
TGME49_003870	2	15	isy1-like splicing family domain-containing protein
TGME49_094730	32	1125	hypothetical protein
TGGT1_075050	26	1055	conserved hypothetical protein
TGME49_057180	1	25	DNA repair protein Rad50, putative
TGME49_114810	32	4346	60S ribosomal protein L7, putative
TGGT1_092100	9	476	riboxomal protein L30/L7E / wx protein, putative
TGGT1_121280	9	203	proliferating cell nuclear antigen, putative
TGME49_120110	9	203	proliferating cell nuclear antigen, putative
TGME49_035660	4	34	hypothetical protein, conserved
TGGT1_070110	3	26	conserved hypothetical protein
TGGT1_082800	2	6	conserved hypothetical protein
TGME49_027310	1	2	hypothetical protein
TGGT1_035170	2	7	GRIP domain-containing protein, putative
TGME49_088240	2	7	GRIP domain-containing protein
TGGT1_065210	2	16	conserved hypothetical protein
TGME49_021360	1	7	beclin 1, putative
TGGT1_103080	1	8	conserved hypothetical protein
TGME49_051710	1	8	hypothetical protein
TGGT1_111530	3	14	conserved hypothetical protein
TGME49_072340	3	14	hypothetical protein
TGME49_047200	4	7	hypothetical protein
TGGT1_025800	1	1	conserved hypothetical protein
TGME49_093590	20	700	3-oxoacyl-[acyl-carrier-protein] synthase,

			putative
TGGT1_073980	14	559	beta-ketoacyl synthase, putative
TGME49_023070	15	536	hypothetical protein
TGGT1_066610	11	519	conserved hypothetical protein
TGME49_059240	1	12	40S ribosomal protein S21, putative
TGGT1_010570	1	6	40S ribosomal protein S21, putative
TGME49_049010	7	106	hypothetical protein
TGGT1_100680	5	102	conserved hypothetical protein
TGME49_047770	12	993	hypothetical protein, conserved
TGGT1_025040	3	54	hypothetical protein
TGGT1_014160	1	5	protein kinase, putative
TGME49_056070	1	5	CMGC kinase
TGME49_016590	33	2025	transportin, putative
TGGT1_098660	27	2099	transportin, putative
TGGT1_108480	4	21	conserved hypothetical protein
TGME49_069940	4	21	putative zinc finger motif, C2HC5-type domain-containing protein
TGME49_089600	16	1113	Hsp20/alpha crystallin domain-containing protein
TGGT1_033610	12	686	Hsp20/alpha crystallin domain-containing protein, putative
TGME49_121620	65	9543	dynamin-like protein, putative
TGGT1_064650	47	8708	dynamin, putative
TGGT1_101380	3	22	conserved hypothetical protein
TGME49_049710	2	18	hypothetical protein
TGME49_062900	2	2	hypothetical protein
TGGT1_007090	1	1	conserved hypothetical protein
TGME49_097780	6	36	DNA gyrase subunit B, putative
TGME49_086920	28	833	hypothetical protein
TGGT1_038200	21	667	conserved hypothetical protein
TGGT1_008830	8	194	homeobox-containing protein, putative
TGME49_061410	7	189	hypothetical protein
TGME49_015260	63	4972	carbamoyl phosphate synthetase II
TGGT1_126000	48	4847	carbamoyl phosphate synthetase II, putative
TGGT1_020090	2	9	kinesin motor domain-containing protein, putative
TGME49_001230	1	4	kinesin motor domain-containing protein
TGME49_025690	9	700	hypothetical protein
TGGT1_080820	8	699	conserved hypothetical protein
TGGT1_069480	3	9	conserved hypothetical protein
TGME49_036270	3	9	hypothetical protein, conserved
TGGT1_112080	2	11	DNA repair protein, putative
TGME49_072900	2	11	DNA repair protein, putative
TGME49_018280	36	11269	eukaryotic porin domain-containing protein
TGGT1_029010	13	306	eukaryotic porin domain-containing protein, putative
TGME49_106220	10	62	formin binding protein, putative
TGGT1_041210	8	52	formin binding protein, putative

TGME49_080520	4	16	hypothetical protein
TGGT1_072790	2	12	conserved hypothetical protein
TGME49_097060	47	3142	phosphoglycerate mutase 1, putative
TGGT1_097170	14	914	phosphoglycerate mutase, putative
TGGT1_035580	20	610	conserved hypothetical protein
TGME49_083560	15	440	hypothetical protein
TGME49_028210	25	1040	26S proteasome regulatory ATPase subunit, putative
TGGT1_083680	18	879	26S proteasome regulatory ATPase subunit, putative
TGME49_021170	9	944	peptidase family M48 domain-containing protein
TGGT1_064880	5	98	caax prenyl protease ste24, putative
TGME49_019670	3	8	zinc finger (CCCH type) protein, putative
TGME49_000360	7	146	hypothetical protein
TGGT1_060580	6	187	conserved hypothetical protein
TGME49_035930	49	3838	KH domain-containing protein
TGGT1_069840	28	2466	KH domain-containing protein, putative
TGME49_093860	5	51	hypothetical protein
TGGT1_074250	4	46	conserved hypothetical protein
TGME49_042700	8	146	hypothetical protein, conserved
TGGT1_047220	3	46	conserved hypothetical protein
TGME49_058050	19	406	actin-like family protein ARP4a, putative
TGGT1_012160	14	444	actin, putative
TGGT1_038900	1	2	ABC1 domain-containing protein, putative
TGME49_013620	1	2	ABC1 domain-containing protein
TGME49_054690	30	1317	phospholipase/carboxylesterase domain containing protein
TGGT1_004240	17	1000	hypothetical protein
TGME49_097800	22	288	structural maintenance of chromosomes protein, putative
TGGT1_096300	9	141	structural maintenance of chromosome protein, putative
TGME49_053290	50	3980	valyl-tRNA synthetase, putative
TGGT1_002280	35	3568	valyl-tRNA synthetase, putative
TGME49_018220	11	113	CMGC kinase, CDK family TgPK2
TGGT1_028940	9	95	casein kinase II alpha, putative
TGME49_094050	6	58	hypothetical protein
TGGT1_074450	4	33	conserved hypothetical protein
TGME49_053990	4	6262	hypothetical protein
TGGT1_003270	4	34	conserved hypothetical protein
TGGT1_123050	1	2	phosphoglycerate mutase, putative
TGME49_118190	1	2	phosphoglycerate mutase, putative
TGME49_093050	3	5	sybindin-like family domain-containing protein
TGGT1_073390	2	4	synbindin, putative
TGGT1_098740	1	1	hypothetical protein
TGME49_016500	1	1	tRNA synthetase, putative
TGME49_003940	5	13	ABC1 domain-containing protein
TGGT1_016770	1	1	ABC1 domain-containing protein, putative
TGME49_065530	37	2344	RNA binding motif-containing protein

TGGT1_057460	22	1739	RNA binding motif-containing protein, putative
TGME49_110470	14	962	cytochrome c oxidase subunit II, putative
TGGT1_086690	3	51	cytochrome C oxidase subunit II, putative
TGGT1_108140	23	1356	conserved hypothetical protein
TGME49_069710	21	1150	hypothetical protein
TGME49_097430	16	865	hypothetical protein, conserved
TGGT1_096800	9	562	conserved hypothetical protein
TGME49_089890	1	4	hypothetical protein
TGGT1_065220	2	10	conserved hypothetical protein
TGME49_021370	1	4	hypothetical protein
TGME49_091180	19	351	hypothetical protein, conserved
TGGT1_031930	16	408	vacuolar sorting protein, putative-associated protein, putative
TGME49_047730	19	878	hypothetical protein
TGGT1_025080	16	957	conserved hypothetical protein
TGME49_035580	4	14	hypothetical protein
TGGT1_044080	4	133	conserved hypothetical protein
TGME49_063860	4	133	hypothetical protein
TGME49_030880	17	604	glutamic acid-rich protein, putative
TGGT1_116970	2	47	conserved hypothetical protein
TGME49_049910	4	14	mitochondria-associated granulocyte macrophage CSF signaling molecule, putative
TGGT1_101600	2	4	mitochondria-associated granulocyte macrophage CSF signaling molecule, putative
TGGT1_051970	13	225	conserved hypothetical protein
TGME49_039010	6	75	hypothetical protein
TGME49_060500	13	789	hypothetical protein, conserved
TGGT1_009700	5	100	conserved hypothetical protein
TGGT1_074960	28	1767	ribonucleoside-diphosphate reductase, large subunit, putative
TGME49_094640	28	1767	ribonucleoside-diphosphate reductase, large subunit, putative
TGME49_019200	7	29	DEAD/DEAH box RNA helicase, putative
TGGT1_030160	4	24	DEAD/DEAH box RNA helicase, putative
TGGT1_053730	43	16529	conserved hypothetical protein
TGME49_020900	35	12763	hypothetical protein
TGME49_082180	18	2262	hypothetical protein
TGGT1_005290	7	145	NBP2B protein, putative
TGME49_105820	11	269	calcyclin binding protein, putative
TGGT1_041650	11	265	calicylin binding protein, putative
TGGT1_036570	1	1	conserved hypothetical protein
TGME49_084650	1	1	hypothetical protein
TGME49_094190	10	48	hypothetical protein, conserved
TGGT1_074490	1	3	conserved hypothetical protein
TGGT1_037060	6	39	DEAD/DEAH box helicase domain-containing protein, putative

TGME49_085660	6	34	DEAD/DEAH box helicase domain-containing protein
TGGT1_113720	2	6	RNA polymerase, putative
TGME49_033720	2	6	RNA polymerase, putative
TGME49_004090	6	19	hypothetical protein
TGGT1_016610	2	7	conserved hypothetical protein
TGME49_036580	23	1005	60S ribosomal protein L7a, putative
TGGT1_069260	4	128	60S ribosomal protein L7A, putative
TGME49_013350	27	1799	40S ribosomal protein S15, putative
TGGT1_039300	8	379	40S ribosomal protein S15, putative
TGGT1_013090	1	4	conserved hypothetical protein
TGGT1_092570	30	1470	WD domain, G-beta repeat-containing protein, putative
TGME49_115140	25	1095	WD domain, G-beta repeat-containing protein
TGGT1_039810	6	246	conserved hypothetical protein
TGME49_012940	5	200	hypothetical protein
TGME49_023860	5	10	RNA-binding protein, putative
TGGT1_078230	4	9	RNA-binding protein, putative
TGGT1_120660	10	205	choline kinase, putative
TGME49_120630	10	205	choline kinase, putative
TGGT1_097140	9	169	pyridoxine kinase, putative
TGME49_097080	8	152	hypothetical protein
TGGT1_115020	3	24	conserved hypothetical protein
TGME49_032500	3	24	hypothetical protein
TGME49_031410	44	8351	hypothetical protein, conserved
TGGT1_116260	17	861	conserved hypothetical protein
TGME49_045460	22	3149	40S ribosomal protein S8, putative
TGGT1_027600	14	894	40S ribosomal protein S8, putative
TGGT1_037040	6	74	conserved hypothetical protein
TGME49_085540	4	44	hypothetical protein, conserved
TGGT1_038190	4	41	conserved hypothetical protein
TGME49_086800	2	15	hypothetical protein
TGME49_113230	15	976	eukaryotic translation initiation factor 2 alpha subunit, putative
TGGT1_089620	14	898	eukaryotic translation initiation factor 2 alpha subunit, putative
TGGT1_114060	1	2	WD domain, G-beta repeat-containing protein, putative
TGME49_033410	1	2	WD domain, G-beta repeat-containing protein
TGGT1_021310	1	1	pinA, putative
TGME49_008310	1	1	pinA, putative
TGME49_059020	3	15	bradyzoite antigen, putative
TGME49_088460	25	659	hypothetical protein
TGGT1_034930	21	661	conserved hypothetical protein
TGGT1_028470	1	1	ribosomal protein L17, putative
TGME49_017590	1	1	hypothetical protein
TGGT1_015190	9	224	RNase H, putative
TGME49_020470	7	161	RNase H, putative

TGME49_031370	8	43	patatin-like phospholipase domain-containing protein
TGGT1_116300	7	62	conserved hypothetical protein
TGME49_119870	43	2900	ubiquitin-conjugating enzyme domain-containing protein
TGGT1_121520	4	173	ubiquitin-conjugating enzyme E2, putative
TGGT1_020780	1	2	conserved hypothetical protein
TGGT1_065450	6	39	conserved hypothetical protein
TGME49_021580	6	39	ftsJ-like methyltransferase domain-containing protein
TGGT1_087910	3	48	SET domain-containing protein, putative
TGME49_111660	2	27	SET domain-containing protein
TGGT1_071600	8	212	conserved hypothetical protein
TGME49_034290	7	195	hypothetical protein
TGME49_059250	2	4	ATP-dependent DNA helicase, putative
TGGT1_010560	1	6	ATP-dependent DNA helicase, putative
TGME49_047480	4	26	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_021890	2	12	conserved hypothetical protein
TGME49_008860	2	12	hypothetical protein
TGME49_061600	36	1094	X-prolyl aminopeptidase, putative
TGGT1_008630	22	840	X-prolyl aminopeptidase, putative
TGGT1_034880	1	2	conserved hypothetical protein
TGME49_088510	1	2	hypothetical protein
TGGT1_014720	2	6	conserved hypothetical protein
TGME49_020120	1	4	hypothetical protein
TGGT1_082210	2	8	conserved hypothetical protein
TGGT1_043030	6	118	pantothenate kinase, putative
TGME49_107770	5	83	pantothenate kinase, putative
TGME49_010400	2	4	hypothetical protein
TGGT1_076890	2	3	GTP-binding protein HFLX, putative
TGME49_055470	1	2	GTP-binding protein HFLX, putative
TGME49_113430	90	5618	nucleoporin FG repeat-containing protein
TGGT1_089830	31	3016	nucleoporin FG repeat-containing protein, putative
TGME49_044210	24	197	hypothetical protein
TGGT1_045730	10	192	conserved hypothetical protein
TGME49_010360	15	190	DEAD/DEAH box helicase, putative
TGGT1_053270	9	132	DEAD/DEAH box helicase, putative
TGME49_018780	37	2546	phosphoserine aminotransferase, putative
TGGT1_029750	20	2000	phosphoserine aminotransferase, putative
TGGT1_089690	4	58	hypothetical protein
TGME49_113300	4	58	hypothetical protein, conserved
TGGT1_030320	17	350	conserved hypothetical protein
TGME49_019440	13	241	hypothetical protein
TGME49_075410	4	7	hypothetical protein, conserved
TGGT1_068090	2	3	conserved hypothetical protein
TGGT1_017320	5	44	zinc finger (C3HC4 RING finger) protein
TGME49_003510	5	37	zinc finger (C3HC4 RING finger) protein, putative

TGGT1_069830	1	2	conserved hypothetical protein
TGME49_035940	1	2	hypothetical protein
TGGT1_108170	3	9	DNA-binding protein, putative
TGME49_069740	3	9	DNA-binding protein, putative
TGME49_035700	15	272	sedoheptulose-1,7 bisphosphatase, putative
TGGT1_070070	10	231	sedoheptulose-1,7 bisphosphatase, putative
TGGT1_029860	1	1	conserved hypothetical protein
TGME49_018890	1	1	hypothetical protein
TGME49_008560	36	1297	mitochondrial carrier domain-containing protein
TGGT1_021590	17	741	mitochondrial carrier domain-containing protein
TGGT1_017300	2	6	conserved hypothetical protein
TGME49_003530	2	6	RNA recognition motif domain containing protein
TGME49_068580	11	116	hypothetical protein, conserved
TGGT1_106650	7	100	conserved hypothetical protein
TGME49_015450	1	14	aquaporin
TGME49_037500	4	15	protein phosphatase 2C, putative
TGGT1_068400	3	6	protein phosphatase 2C, putative
TGME49_066260	3	12	hypothetical protein, conserved
TGGT1_056600	2	8	conserved hypothetical protein
TGGT1_009750	2	8	DEAD/DEAH box helicase, putative
TGME49_060450	1	4	DEAD/DEAH box helicase, putative
TGME49_025230	6	42	hypothetical protein
TGGT1_080340	3	21	conserved hypothetical protein
TGME49_034190	75	5679	glycine hydroxymethyltransferase, putative
TGGT1_071690	20	2180	glycine hydroxymethyltransferase, putative
TGGT1_022690	1	2	conserved hypothetical protein
TGME49_009630	1	2	hypothetical protein
TGME49_030070	4	38	BolA-like domain containing protein
TGGT1_117810	2	34	conserved hypothetical protein
TGGT1_063180	1	8	conserved hypothetical protein
TGME49_005730	1	8	hypothetical protein
TGGT1_032740	20	1223	conserved hypothetical protein
TGME49_090310	15	715	hypothetical protein
TGME49_034420	51	4642	ATPase, AAA family domain-containing protein
TGGT1_071470	19	949	conserved hypothetical protein
TGME49_110370	3	33	hypothetical protein
TGGT1_086590	2	8	conserved hypothetical protein
TGME49_039020	4	116	ATP-binding cassette protein subfamily B member 2
TGGT1_052610	6	75	bystin, putative
TGME49_038240	6	75	bystin, putative
TGME49_062560	3	23	hypothetical protein
TGGT1_007540	3	19	conserved hypothetical protein
TGME49_022220	48	13219	articulin 4
TGGT1_066030	33	4353	articulin, putative

TGME49_025030	2	4	myosin heavy chain, putative
TGGT1_011560	1	4	adenylate kinase, putative
TGME49_058640	1	4	adenylate kinase, putative
TGME49_061780	17	643	microneme protein 7
TGGT1_008350	11	323	microneme protein, putative
TGME49_054500	13	185	hypothetical protein, conserved
TGGT1_004070	11	296	conserved hypothetical protein
TGME49_075680	10	130	hypothetical protein, conserved
TGGT1_000650	9	126	conserved hypothetical protein
TGME49_052270	4	10	50S ribosomal protein L1, putative
TGGT1_001180	2	6	50S ribosomal protein L1p, putative
TGME49_046130	23	945	serine protease inhibitor, putative
TGGT1_026920	17	771	serine protease inhibitor, putative
TGME49_046940	18	971	hypothetical protein
TGGT1_026170	14	990	conserved hypothetical protein
TGME49_110060	8	242	oligoribonuclease, putative
TGGT1_086205	7	193	oligoribonuclease protein, putative
TGGT1_118290	6	70	NC domain-containing protein, putative
TGME49_029380	5	83	NC domain-containing protein
TGGT1_030670	10	206	pyruvate dehydrogenase, putative
TGME49_019680	8	97	hypothetical protein, conserved
TGME49_051550	3	7	acyl-CoA-binding protein, putative
TGGT1_102950	2	3	acyl-CoA-binding protein, putative
TGGT1_118020	9	98	conserved hypothetical protein
TGME49_029770	6	63	hypothetical protein
TGME49_046220	6	24	hypothetical protein, conserved
TGGT1_026720	2	9	lipopolysaccharide-induced transcription factor regulating tumor necrosis factor alpha, putative
TGME49_008200	7	593	PHD-finger domain-containing protein
TGGT1_021290	7	592	PHD-finger domain-containing protein, putative
TGGT1_018090	13	326	conserved hypothetical protein
TGME49_002850	10	272	hypothetical protein
TGME49_073900	21	860	hypothetical protein
TGGT1_112990	19	928	conserved hypothetical protein
TGME49_029700	1	2	hypothetical protein, conserved
TGME49_023510	10	144	hypothetical protein
TGGT1_077780	8	122	conserved hypothetical protein
TGME49_042330	38	7763	40S ribosomal protein S5, putative
TGGT1_047500	13	931	40S ribosomal protein S5, putative
TGME49_093180	31	1885	NADP-specific glutamate dehydrogenase, putative
TGGT1_073530	17	1352	NADP-specific glutamate dehydrogenase, putative
TGME49_095400	6	108	hypothetical protein
TGGT1_054540	4	106	conserved hypothetical protein
TGME49_121650	38	961	protein kinase
TGGT1_064690	20	627	conserved hypothetical protein

TGME49_007830	14	363	MORN repeat-containing protein
TGGT1_020810	12	381	MORN repeat-containing protein, putative
TGME49_068810	7	137	casein kinase II beta chain, putative
TGGT1_107000	7	128	casein kinase II beta chain, putative
TGGT1_076220	1	2	conserved hypothetical protein
TGME49_066750	1	2	hypothetical protein, conserved
TGME49_033110	53	2733	inosine-5'-monophosphate dehydrogenase, putative
TGGT1_114400	25	1438	inosine-5'-monophosphate dehydrogenase, putative
TGME49_113390	36	3988	60s ribosomal protein l6, putative
TGGT1_089790	13	788	60S ribosomal protein l6, putative
TGME49_026390	10	76	hypothetical protein
TGGT1_081560	3	14	conserved hypothetical protein
TGGT1_112870	4	62	conserved hypothetical protein
TGME49_073790	1	1	hypothetical protein
TGME49_048600	34	1035	aspartate aminotransferase, putative
TGGT1_024250	10	284	hypothetical protein
TGME49_116120	2	4	GTP-binding protein, putative
TGGT1_021610	5	29	DNA ligase 1 precursor, putative
TGME49_008580	4	21	DNA ligase 1 precursor, putative
TGME49_020930	14	1436	hypothetical protein, conserved
TGGT1_053750	5	80	conserved hypothetical protein
TGGT1_089670	13	274	WD-repeat protein, putative
TGME49_113280	8	147	WD-40 repeat protein, putative
TGGT1_043370	5	76	conserved hypothetical protein
TGME49_063260	3	40	major facilitator superfamily domain-containing protein
TGME49_068860	5	53	enolase 1
TGGT1_038330	3	67	sulfate transporter, putative
TGME49_087230	3	36	sulfate transporter family domain-containing protein
TGME49_061580	18	1092	histone H2A, putative
TGGT1_008650	4	78	histone H2A, putative
TGME49_005740	59	11296	hypothetical protein
TGGT1_063160	13	304	conserved hypothetical protein
TGGT1_034840	4	30	conserved hypothetical protein
TGME49_088550	4	30	hypothetical protein
TGGT1_075820	32	3449	protein phosphatase 2C, putative
TGME49_067100	29	3023	protein phosphatase 2C, putative
TGME49_005710	44	2866	lysyl-tRNA synthetase, putative
TGGT1_063200	21	1860	lysyl-tRNA synthetase, putative
TGGT1_015160	9	132	ribonuclease H, putative
TGME49_020450	6	50	ribonuclease H, putative
TGGT1_017960	3	26	glycine cleavage system H protein, putative
TGME49_002960	2	14	glycine cleavage system H protein, putative
TGME49_092300	6	26	RNA polymerase Rpb7, N-terminal domain-containing protein

TGGT1_031270	4	22	RNA polymerase Rpb7, N-terminal domain-containing protein, putative
TGGT1_099370	3	14	trichohyalin, putative
TGME49_015890	3	14	trichohyalin, putative
TGGT1_088470	2	9	mitochondrial import inner membrane translocase subunit TIM17, putative
TGME49_112220	2	4	mitochondrial import inner membrane translocase subunit TIM17, putative
TGME49_001640	1	9	hypothetical protein
TGME49_016210	29	988	hypothetical protein
TGGT1_099050	27	983	conserved hypothetical protein
TGGT1_077970	21	1108	conserved hypothetical protein
TGME49_023610	17	936	hypothetical protein
TGGT1_033710	18	784	conserved hypothetical protein
TGME49_089520	17	612	hypothetical protein
TGGT1_029380	1	2	conserved hypothetical protein
TGME49_018550	1	2	PIK3R4 kinase-related protein
TGGT1_046330	1	1	PPR repeat-containing protein, putative
TGGT1_044050	9	90	ATP-dependent RNA helicase, putative
TGME49_063820	7	67	ATP-dependent RNA helicase, putative
TGME49_094560	7	23	Rhoptry kinase family protein ROP37 (incomplete catalytic triad)
TGGT1_074860	2	7	conserved hypothetical protein
TGGT1_044040	3	54	conserved hypothetical protein
TGME49_063810	1	14	hypothetical protein
TGGT1_020010	1	1	protein phosphatase 2C, putative
TGME49_001410	1	1	protein phosphatase 2C, putative
TGME49_105980	24	1093	dihydrolipoyl dehydrogenase protein, putative
TGGT1_041490	16	766	dihydrolipoamid dehydrogenase, putative
TGGT1_025950	12	563	conserved hypothetical protein
TGME49_047050	9	407	hypothetical protein
TGME49_013070	10	194	hypothetical protein, conserved
TGGT1_039660	7	168	conserved hypothetical protein
TGGT1_038050	1	1	conserved hypothetical protein
TGME49_086620	1	1	30S ribosomal protein S1, putative
TGGT1_069210	1	2	conserved hypothetical protein
TGME49_036630	1	2	hypothetical protein
TGME49_063150	9	87	hypothetical protein
TGGT1_043260	7	104	conserved hypothetical protein
TGME49_059690	15	317	orotidine-monophosphate-decarboxylase, putative
TGGT1_010340	11	304	conserved hypothetical protein
TGME49_046210	6	36	hypothetical protein
TGGT1_026830	5	35	conserved hypothetical protein
TGME49_086050	7	29	hypothetical protein
TGGT1_037450	4	16	conserved hypothetical protein
TGGT1_047570	15	517	conserved hypothetical protein

TGME49_042260	15	517	hypothetical protein
TGGT1_048330	4	21	zinc finger (CCCH type) protein
TGME49_042090	4	21	zinc finger (CCCH type) protein, putative
TGGT1_114700	3	35	zinc finger protein, putative
TGME49_032820	3	35	tim10/DDP zinc finger domain-containing protein
TGGT1_027520	25	1278	conserved hypothetical protein
TGME49_045560	22	1050	hypothetical protein
TGME49_107780	1	1	hypothetical protein
TGME49_116300	5	14	DEAD/DEAH box helicase, putative
TGGT1_093990	3	10	DEAD/DEAH box helicase, putative
TGME49_109740	6	95	small nuclear ribonucleoprotein, putative
TGGT1_085700	5	84	small nuclear ribonucleoprotein, putative
TGGT1_109900	1	1	conserved hypothetical protein
TGME49_071020	1	1	hypothetical protein, conserved
TGGT1_096680	2	6	conserved hypothetical protein
TGME49_097510	1	1	hypothetical protein
TGME49_093300	6	71	WW domain containing protein
TGGT1_073660	5	69	WW domain containing protein, putative
TGGT1_113430	1	11	conserved hypothetical protein
TGME49_033910	1	11	hypothetical protein
TGME49_112200	19	530	serine/threonine protein phosphatase, putative
TGGT1_088450	16	432	hypothetical protein
TGGT1_020910	4	19	conserved hypothetical protein
TGME49_007920	3	16	hypothetical protein
TGME49_092160	8	50	hypothetical protein
TGGT1_031400	6	47	conserved hypothetical protein
TGGT1_040910	1	1	conserved hypothetical protein
TGME49_106440	1	1	hypothetical protein
TGGT1_050770	4	33	RNA recognition motif domain-containing protein, putative
TGME49_040070	4	33	RNA recognition motif domain-containing protein
TGGT1_073840	2	14	DNA ligase, putative
TGME49_093460	2	14	DNA ligase, putative
TGME49_095090	8	43	hypothetical protein
TGGT1_075450	3	18	hypothetical protein
TGME49_094350	99	10723	DEAD/DEAH box helicase, putative
TGGT1_074650	63	8950	DEAD/DEAH box helicase, putative
TGME49_005010	50	2464	splicing factor 3B subunit 1, putative
TGGT1_064040	26	1884	splicing factor 3B subunit, putative
TGME49_032430	2	7	hypothetical protein
TGGT1_115200	1	6	conserved hypothetical protein
TGGT1_104330	3	23	prenyltransferase and squalene oxidase repeat domain-containing protein, putative
TGME49_078230	3	23	geranylgeranyl transferase type II beta subunit, putative
TGME49_056760	71	18560	pyruvate kinase, putative

TGGT1_014000	34	13719	pyruvate kinase, putative
TGGT1_035130	4	18	conserved hypothetical protein
TGME49_088270	3	12	hypothetical protein
TGME49_016450	27	861	proteasome subunit alpha type 3, putative
TGGT1_098790	11	534	20S proteasome subunit alpha, putative
TGME49_078700	3	9	hypothetical protein
TGGT1_016430	13	257	protein kinase domain-containing protein, putative
TGME49_004280	10	198	CMGC kinase, Dyrk family
TGME49_017790	2	6	GTPase regulator, putative
TGME49_075630	2	6	ubiquitin-transferase domain-containing protein
TGGT1_000700	1	2	ubiquitin-transferase domain-containing protein, putative
TGGT1_038590	7	77	conserved hypothetical protein
TGME49_013890	6	60	hypothetical protein, conserved
TGME49_025050	65	7377	adenosylhomocysteinase, putative
TGGT1_079990	26	4210	adenosylhomocysteinase, putative
TGME49_028000	11	128	splicing factor 3A subunit 2, putative
TGGT1_083280	6	90	splicing factor 3A subunit, putative
TGME49_047930	12	162	syntaxin, putative
TGGT1_024990	6	67	hypothetical protein
TGME49_062670	17	2090	60S ribosomal protein L18a, putative
TGGT1_007420	8	372	60S ribosomal protein L18A, putative
TGGT1_095120	5	48	conserved hypothetical protein
TGME49_100180	3	32	hypothetical protein
TGME49_118430	26	4808	malate dehydrogenase, putative
TGGT1_122700	20	3063	malate dehydrogenase, putative
TGGT1_087990	1	2	conserved hypothetical protein
TGME49_111740	1	2	hypothetical protein
TGGT1_011480	13	457	conserved hypothetical protein
TGME49_058710	13	457	hypothetical protein
TGME49_047470	43	2790	nucleolar protein 5A, putative
TGGT1_025440	20	1597	nucleolar protein 5A, putative
TGME49_005340	32	2836	40S ribosomal protein S12, putative
TGGT1_063570	11	916	40S ribosomal protein S12, putative
TGME49_109600	60	3588	hypothetical protein
TGGT1_085560	19	653	conserved hypothetical protein
TGGT1_118340	1	1	conserved hypothetical protein
TGME49_029320	1	1	haloacid dehalogenase-like hydrolase domain-containing protein
TGGT1_001160	2	8	ATPase domain-containing cell division control protein, putative
TGME49_052240	1	4	ATPase domain-containing cell division control protein, putative
TGGT1_015070	11	441	conserved hypothetical protein
TGME49_020370	11	335	200 kDa antigen p200, putative
TGGT1_059930	1	3	conserved hypothetical protein

TGME49_101460	1	3	hypothetical protein
TGME49_063530	7	132	chaperonin CPN10, mitochondrial, putative
TGGT1_043750	4	70	heat shock protein HSP10, putative
TGME49_009050	3	7	Tyrosine kinase-like (TKL) protein
TGGT1_021990	1	3	protein kinase, putative
TGGT1_071630	9	332	conserved hypothetical protein
TGME49_034260	8	199	hypothetical protein
TGME49_044440	12	202	adenosine transporter
TGGT1_045500	6	58	adenosine transporter, putative
TGME49_032540	11	182	FAD-depdendent monooxygenase, putative
TGGT1_114980	4	29	hypothetical protein
TGGT1_104760	5	104	hypothetical protein
TGME49_077890	5	55	hypothetical protein, conserved
TGME49_039120	1	1	hypothetical protein
TGME49_113610	2	4	hypothetical protein
TGME49_019170	11	1186	hypothetical protein
TGGT1_030130	9	1145	conserved hypothetical protein
TGGT1_025470	1	13	conserved hypothetical protein
TGGT1_022180	8	82	anti-silencing protein, putative
TGME49_009140	4	26	anti-silencing protein 1, putative
TGGT1_068500	8	293	protein phosphatase 2C, putative
TGME49_037400	8	293	protein phosphatase 2C, putative
TGGT1_021880	4	12	SRS11 (= SAG2B)
TGME49_008850	3	8	SRS11 (= SAG2B)
TGME49_089730	9	68	vacuolar protein sorting 18, putative
TGGT1_033490	4	32	vacuolar sorting protein, putative
TGME49_088710	6	18	hypothetical protein
TGGT1_034560	2	8	conserved hypothetical protein
TGGT1_086500	12	384	conserved hypothetical protein
TGME49_110270	11	317	hypothetical protein, conserved
TGME49_024840	6	64	cAMP phosphodiesterase, putative
TGGT1_013500	7	332	nuclear movement domain-containing protein, putative
TGME49_057310	6	285	nuclear movement domain-containing protein
TGME49_035920	3	26	dynein 1-beta heavy chain, flagellar inner arm I1 complex, putative
TGME49_115110	11	220	prefoldin subunit 5, putative
TGGT1_092410	7	183	prefoldin subunit, putative
TGME49_007370	35	1723	hypothetical protein
TGGT1_020330	29	1638	conserved hypothetical protein
TGME49_110480	2	6	hypothetical protein
TGGT1_026350	20	780	conserved hypothetical protein
TGME49_046740	17	624	hypothetical protein, conserved
TGGT1_083150	8	66	conserved hypothetical protein
TGME49_027850	6	40	RNA recognition motif-containing protein
TGME49_062620	37	3926	Gbp1p protein, putative

TGGT1_007470	16	1093	gbp1p protein, putative
TGME49_048750	7	18	hypothetical protein
TGGT1_100290	1	1	conserved hypothetical protein
TGME49_110860	14	360	hypothetical protein, conserved
TGGT1_087180	12	326	U5 snRNP-specific 40 kDa protein, putative
TGGT1_050250	8	209	protein kinase, putative
TGME49_040630	6	141	ULK kinase
TGGT1_091210	2	8	conserved hypothetical protein
TGME49_114420	2	8	AP-4 complex subunit sigma-1, putative
TGME49_083700	19	339	phosphatidylinositol 3- and 4-kinase domain-containing protein
TGME49_030210	49	17407	hypothetical protein
TGGT1_117680	24	2830	conserved hypothetical protein
TGGT1_111760	9	155	serine-threonine protein kinase, putative
TGME49_072470	9	155	Tyrosine kinase-like (TKL) protein
TGME49_024480	23	1426	CMGC kinase, Lammer family
TGGT1_079100	9	562	dual specificity protein kinase, putative
TGGT1_006810	33	3613	arginine/serine-rich splicing factor, putative
TGME49_104760	33	3126	hypothetical protein, conserved
TGME49_061950	89	367609	ATP synthase beta chain, putative
TGGT1_008270	32	7470	ATP synthase beta chain, putative
TGME49_003290	7	75	hypothetical protein
TGGT1_017570	6	71	conserved hypothetical protein
TGME49_001700	14	355	sec 13, putative
TGGT1_019730	8	296	protein transport protein sec13, putative
TGGT1_121880	3	15	conserved hypothetical protein
TGME49_119320	3	15	hypothetical protein
TGGT1_020120	4	16	zinc finger (CCCH type) protein
TGME49_001200	3	9	zinc finger (CCCH type) protein, putative
TGGT1_043540	23	1087	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_063420	21	871	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_095900	36	2172	cyclic nucleotide-binding domain-containing protein
TGGT1_054220	28	2272	cyclic nucleotide-binding domain-containing protein, putative
TGGT1_010270	16	352	conserved hypothetical protein
TGME49_059860	10	197	hypothetical protein
TGGT1_022810	1	2	DNA repair helicase, putative
TGME49_009770	1	2	DNA repair helicase, putative
TGGT1_038380	3	11	conserved hypothetical protein
TGME49_087260	3	11	hypothetical protein
TGME49_030510	11	302	hypothetical protein
TGGT1_117450	7	119	conserved hypothetical protein
TGME49_089830	30	2087	eukaryotic translation initiation factor 3 delta subunit, putative
TGGT1_033400	19	1486	eukaryotic translation initiation factor 3 delta subunit, putative

TGME49_095000	13	179	sec1 family domain-containing protein
TGGT1_075350	11	170	vacuolar sorting protein, putative
TGME49_049300	8	109	hypothetical protein
TGGT1_100970	2	7	conserved hypothetical protein
TGME49_017010	11	233	CD2 antigen cytoplasmic tail-binding protein 2, putative
TGGT1_098120	8	216	CD2 antigen cytoplasmic tail-binding protein, putative
TGME49_066620	8	239	thioredoxin, putative
TGGT1_076350	5	99	thioredoxin, putative
TGGT1_098450	4	26	ankyrin repeat-containing protein, putative
TGME49_016690	3	18	ankyrin repeat domain-containing protein
TGGT1_095530	2	62	conserved hypothetical protein
TGME49_099980	2	62	hypothetical protein
TGGT1_065860	4	33	conserved hypothetical protein
TGME49_022030	3	32	hypothetical protein
TGME49_071110	2	12	hypothetical protein
TGGT1_109990	1	8	conserved hypothetical protein
TGGT1_003370	6	128	conserved hypothetical protein
TGME49_054090	6	128	hypothetical protein, conserved
TGME49_027820	30	783	hypothetical protein
TGGT1_083120	19	1131	conserved hypothetical protein
TGME49_019300	19	721	zinc finger, putative
TGGT1_030270	11	364	zinc finger protein, putative
TGGT1_038520	7	114	conserved hypothetical protein
TGME49_013950	6	100	hypothetical protein, conserved
TGME49_090970	2	2	serine palmitoyltransferase, putative
TGGT1_032150	1	1	serine palmitoyltransferase, putative
TGGT1_070030	5	23	IMP dehydrogenase, putative
TGME49_035740	5	23	IMP dehydrogenase / GMP reductase domain containing protein
TGME49_053430	64	6808	asparagine synthase, putative
TGGT1_002510	27	4520	asparagine synthase (glutamine-hydrolyzing), putative
TGME49_099810	42	2198	cysteinyl-tRNA synthetase, putative
TGGT1_095800	26	2004	cysteinyl-tRNA synthetase, putative
TGME49_040440	2	6	hypothetical protein
TGME49_110700	33	1295	serine/threonine protein phosphatase, putative
TGGT1_086920	14	303	serine/threonine protein phosphatase, putative
TGME49_044180	13	67	microneme protein, putative
TGGT1_045760	1	1	microneme protein, putative
TGGT1_034260	11	401	conserved hypothetical protein
TGME49_089010	8	234	hypothetical protein, conserved
TGGT1_017580	1	2	conserved hypothetical protein
TGME49_003280	1	2	hypothetical protein
TGME49_063520	44	7830	hypothetical protein, conserved
TGGT1_043740	24	2479	conserved hypothetical protein
TGME49_078900	8	124	protein kinase, BUD32 homolog

TGGT1_103460	5	74	O-sialoglycoprotein endopeptidase, putative
TGME49_073840	11	170	ribosome biogenesis protein SSF1, putative
TGGT1_112920	8	171	ribosome biogenesis protein SSF1, putative
TGGT1_007570	3	9	hypothetical protein
TGME49_062530	3	9	hypothetical protein
TGME49_078640	5	301	protein inhibitor of activated STAT protein, putative
TGGT1_103920	4	163	protein inhibitor of activated STAT protein, putative
TGME49_024960	8	89	hypothetical protein
TGGT1_079880	3	16	conserved hypothetical protein
TGGT1_081110	16	796	protein kinase, putative
TGME49_025960	11	605	STE kinase
TGME49_090940	9	98	endomembrane domain70-containing protein
TGGT1_032180	3	8	endomembrane domain-containing protein, putative
TGME49_081440	11	72	hypothetical protein
TGGT1_005910	1	2	conserved hypothetical protein
TGGT1_116870	6	44	conserved hypothetical protein
TGME49_030950	3	9	hypothetical protein
TGME49_010790	23	1562	dihydroorotate dehydrogenase
TGGT1_124080	8	138	dihydroorotate dehydrogenase, putative
TGME49_090670	75	12725	cytosol aminopeptidase
TGGT1_032360	40	8629	cytosol aminopeptidase, putative
TGME49_070320	35	1944	protein phosphatase 2C, putative
TGGT1_108850	18	942	protein phosphatase 2C, putative
TGGT1_116960	6	66	PHD-finger domain-containing protein, putative
TGME49_030890	6	66	PHD-finger domain-containing protein
TGME49_029180	83	10293	importin beta-3 subunit, putative
TGGT1_118650	46	7288	importin beta-3 subunit, putative
TGME49_011420	15	572	hypothetical protein
TGGT1_123370	7	235	arginine/serine-rich splicing factor, putative
TGGT1_024790	6	37	conserved hypothetical protein
TGME49_048150	2	7	hypothetical protein
TGME49_115770	19	1322	cytochrome p450, putative
TGGT1_093470	4	51	cytochrome P450, putative
TGGT1_112380	4	26	conserved hypothetical protein
TGME49_073320	1	3	hypothetical protein
TGGT1_043790	2	9	trichohyalin, putative
TGME49_063570	2	9	trichohyalin, putative
TGGT1_076070	2	21	conserved hypothetical protein
TGME49_066890	2	21	hypothetical protein
TGGT1_103110	22	2097	conserved hypothetical protein
TGME49_051740	17	1512	hypothetical protein
TGME49_044290	20	521	adaptin N terminal region domain-containing protein
TGGT1_045650	18	512	conserved hypothetical protein
TGME49_037830	3	7	DNA polymerase I, putative

TGGT1_058440	1	3	DNA polymerase I, putative
TGME49_056090	13	146	glycerophosphoryl diester phosphodiesterase protein, putative
TGGT1_014140	5	43	glycerophosphoryl diester phosphodiesterase, putative
TGME49_120490	11	81	hypothetical protein
TGGT1_120900	5	23	outer membrane protein romA, putative
TGGT1_089390	14	347	signal recognition particle 54 kda protein, putative
TGME49_113100	13	346	signal recognition particle 54 kda protein, putative
TGME49_110780	24	1247	hypothetical protein
TGGT1_087090	12	1131	conserved hypothetical protein
TGGT1_098770	2	7	conserved hypothetical protein
TGME49_016470	2	7	hypothetical protein
TGME49_037150	10	444	signal peptide peptidase domain-containing protein
TGGT1_068770	4	75	signal peptide peptidase domain-containing protein, putative
TGME49_093870	22	1816	thioredoxin, putative
TGGT1_074260	7	700	conserved hypothetical protein
TGGT1_065240	25	3184	conserved hypothetical protein
TGME49_021390	24	2855	hypothetical protein
TGME49_040590	9	103	DNA-directed RNA polymerases I, II, and III subunit RPABC1, putative
TGGT1_050280	7	88	DNA-directed RNA polymerase II 23 kDa, putative
TGME49_039700	5	91	hypothetical protein
TGGT1_051160	4	90	regulator of chromosome condensation, putative
TGME49_073730	7	25	hypothetical protein
TGGT1_112810	1	1	conserved hypothetical protein
TGGT1_088920	184	53392	conserved hypothetical protein
TGME49_112630	147	40870	hypothetical protein
TGME49_032600	46	5510	patatin-like phospholipase domain-containing protein
TGGT1_114920	27	3059	conserved hypothetical protein
TGGT1_064910	5	64	pre-mRNA cleavage factor im, 25kD subunit, putative
TGME49_021190	4	52	mRNA cleavage factor-like protein, putative
TGGT1_017340	18	550	GTP-binding protein-animal, putative
TGME49_003490	16	376	GTPase domain containing protein
TGME49_109210	25	2673	peroxidoxin 2
TGGT1_056360	21	2308	peroxidoxin, putative
TGGT1_123270	3	7	conserved hypothetical protein
TGME49_011620	3	7	hypothetical protein
TGME49_068980	7	63	hypothetical protein
TGGT1_107170	6	70	conserved hypothetical protein
TGGT1_023920	8	89	conserved hypothetical protein
TGME49_048720	6	57	hypothetical protein
TGME49_010730	25	716	ATP-dependent protease ATP-binding subunit, putative
TGGT1_124140	15	465	ATP-dependent protease ATP-binding subunit, putative

TGGT1_027460	8	78	conserved hypothetical protein
TGME49_045610	8	77	hypothetical protein
TGGT1_045820	26	3558	conserved hypothetical protein
TGME49_044130	20	2297	hypothetical protein
TGGT1_107950	17	1708	conserved hypothetical protein
TGME49_069620	15	1364	hypothetical protein
TGGT1_101330	13	366	cysteine proteinase, putative
TGME49_049670	13	341	cysteine proteinase, putative
TGGT1_091220	5	25	protein-tyrosine phosphatase, putative
TGME49_114430	4	16	dual specificity protein phosphatase CDC14A, putative
TGME49_065330	53	2901	CMGC kinase, GSK family TgPK3
TGGT1_057780	17	753	protein kinase, putative
TGGT1_083320	1	2	conserved hypothetical protein
TGME49_028060	1	2	hypothetical protein
TGGT1_034790	11	148	conserved hypothetical protein
TGME49_088600	11	100	hypothetical protein
TGME49_015670	1	9	SRS47A
TGME49_100000	12	2110	60S ribosomal protein L18, putative
TGGT1_095500	3	97	60S ribosomal protein L18, putative
TGME49_000470	8	79	hypothetical protein
TGGT1_060690	7	75	conserved hypothetical protein
TGME49_062650	12	258	hypothetical protein
TGGT1_007440	9	252	conserved hypothetical protein
TGGT1_024840	2	9	hypothetical protein
TGME49_048100	2	9	synaptobrevin domain-containing protein
TGME49_055510	13	275	ankyrin repeat-containing protein
TGGT1_076930	12	271	ankyrin repeat-containing protein, putative
TGME49_017020	9	78	ATPase, AFG1 family domain-containing protein
TGGT1_098110	3	21	ATPase n2b, putative
TGME49_011340	5	71	hypothetical protein
TGGT1_123460	4	70	conserved hypothetical protein
TGME49_057530	8	940	hypothetical protein, conserved
TGGT1_013150	6	1262	conserved hypothetical protein
TGGT1_064720	4	17	conserved hypothetical protein
TGME49_121690	4	11	hypothetical protein
TGME49_089650	79	10542	phosphoenolpyruvate carboxykinase
TGGT1_033560	36	7189	phosphoenolpyruvate carboxykinase, putative
TGGT1_019090	12	973	conserved hypothetical protein
TGME49_002120	9	607	hypothetical protein
TGME49_018940	16	722	hypothetical protein
TGGT1_029910	7	204	conserved hypothetical protein
TGME49_063850	26	1554	hypothetical protein
TGGT1_044070	15	870	conserved hypothetical protein
TGGT1_069890	1	2	apoptosis-regulating basic protein, putative
TGME49_035880	1	2	apoptosis-regulating basic protein, putative

TGGT1_097040	2	8	ribosomal protein L17, putative
TGME49_097170	2	8	ribosomal protein L17, putative
TGGT1_078500	1	1	conserved hypothetical protein
TGME49_024020	1	1	hypothetical protein, conserved
TGME49_047020	5	49	hypothetical protein, conserved
TGGT1_025980	5	47	axonemal dynein beta heavy chain 9, putative
TGME49_016930	9	87	cholinephosphate cytidylyltransferase, putative
TGGT1_098200	7	82	cholinephosphate cytidylyltransferase, putative
TGGT1_016970	6	58	conserved hypothetical protein
TGME49_003760	4	36	hypothetical protein
TGME49_036910	11	151	U2 snRNP auxiliary factor small subunit, putative
TGGT1_069020	7	122	U2 snRNP auxiliary factor small subunit, putative
TGME49_051500	43	3305	eukaryotic translation initiation factor 3 subunit 3, putative
TGGT1_102900	23	2188	eukaryotic translation initiation factor 3 subunit, putative
TGGT1_044330	19	873	conserved hypothetical protein
TGME49_064090	19	648	hypothetical protein
TGME49_031200	6	84	hypothetical protein
TGGT1_116510	5	83	conserved hypothetical protein
TGME49_065510	7	174	hypothetical protein
TGGT1_057480	7	162	conserved hypothetical protein
TGME49_058020	4	16	hypothetical protein
TGME49_095410	8	54	hypothetical protein
TGGT1_054530	6	45	conserved hypothetical protein
TGME49_030590	9	52	beta-1,4-mannosyltransferase, putative
TGGT1_117260	5	35	hypothetical protein
TGGT1_112770	5	53	protein kinase, putative
TGME49_073690	5	53	Wee kinase
TGME49_001780	50	67016	microneme protein 2
TGGT1_019450	22	2850	microneme protein, putative
TGME49_055890	18	313	monodehydroascorbate reductase, putative
TGGT1_014330	12	392	pyridine nucleotide-disulphide oxidoreductase, putative
TGGT1_088060	4	17	conserved hypothetical protein
TGME49_111810	3	10	hypothetical protein
TGGT1_048870	7	89	C2 domain-containing protein, putative
TGME49_041840	5	66	C2 domain-containing protein
TGGT1_115730	9	613	conserved hypothetical protein
TGME49_032030	8	443	hypothetical protein
TGGT1_114730	3	63	conserved hypothetical protein
TGME49_032790	2	42	hypothetical protein
TGME49_097460	6	18	hypothetical protein
TGGT1_096760	4	21	conserved hypothetical protein
TGGT1_100630	3	50	conserved hypothetical protein
TGME49_048960	3	50	hypothetical protein
TGGT1_034700	3	9	DnaJ domain-containing protein, putative

TGME49_088690	3	9	DnaJ domain-containing protein
TGGT1_117710	10	117	conserved hypothetical protein
TGME49_030180	7	25	hypothetical protein
TGGT1_051170	2	49	conserved hypothetical protein
TGME49_039690	2	49	hypothetical protein
TGGT1_065800	1	2	conserved hypothetical protein
TGME49_021970	1	2	hypothetical protein
TGGT1_017200	1	2	conserved hypothetical protein
TGME49_114740	35	5724	splicing factor 3B subunit 2, putative
TGGT1_092030	25	5533	splicing factor 3B subunit, putative
TGGT1_034420	2	4	conserved hypothetical protein
TGME49_088840	1	1	hypothetical protein
TGGT1_117000	1	1	conserved hypothetical protein
TGME49_095070	9	72	ATP-dependent helicase, putative
TGGT1_075430	7	68	ATP-dependent helicase, putative
TGGT1_026900	1	1	zinc finger (C3HC4 RING finger) protein
TGME49_046160	1	1	zinc finger (C3HC4 RING finger) protein, putative
TGME49_068710	16	193	elongation factor Tu GTP-binding domain-containing protein
TGGT1_106900	12	203	GTP-binding domain containing protein, putative
TGME49_089630	5	71	hypothetical protein
TGME49_058690	2	2	5-aminolevulinic acid synthase, putative
TGGT1_011500	1	1	5-aminolevulinic acid synthase, putative
TGME49_057450	6	33	SCO1/SenC domain-containing protein
TGGT1_013250	1	3	SCO1/SenC domain-containing protein, putative
TGGT1_030520	1	2	conserved hypothetical protein
TGME49_019640	1	2	hypothetical protein
TGGT1_121640	3	12	3' exoribonuclease, putative
TGME49_119650	3	12	3' exoribonuclease, putative
TGGT1_037140	1	1	TBC domain-containing protein, putative
TGME49_085730	1	1	TBC domain-containing protein
TGME49_002390	43	4088	hypothetical protein
TGGT1_018690	16	906	conserved hypothetical protein
TGME49_080380	41	2016	non-transmembrane antigen
TGGT1_072940	20	900	non-transmembrane antigen, putative
TGGT1_047560	2	3	NUDIX hydrolase domain-containing protein, putative
TGME49_042270	2	3	NUDIX hydrolase domain-containing protein
TGGT1_043430	1	1	conserved hypothetical protein
TGGT1_088080	2	10	conserved hypothetical protein
TGME49_111830	2	10	hypothetical protein
TGME49_017460	94	8517	glutaminyl-tRNA synthetase, putative
TGGT1_028630	38	4079	hypothetical protein
TGME49_015610	5	22	hypothetical protein, conserved
TGGT1_126540	4	28	hypothetical protein
TGME49_054490	23	411	hypothetical protein

TGGT1_004060	2	4	conserved hypothetical protein
TGME49_111690	15	526	hypothetical protein
TGGT1_087940	13	503	conserved hypothetical protein
TGME49_002820	9	149	ubiquitin-conjugating enzyme E2, putative
TGGT1_018120	5	90	ubiquitin-conjugating enzyme E2, putative
TGME49_089370	1	1	hypothetical protein
TGME49_018840	19	177	mismatch repair protein, putative
TGGT1_029810	11	156	conserved hypothetical protein
TGME49_067360	7	34	Sir2 domain-containing protein
TGGT1_075680	4	19	chromatin regulatory protein sir2, putative
TGGT1_008720	1	1	conserved hypothetical protein
TGME49_061510	1	1	hypothetical protein
TGME49_059640	37	1260	nucleoporin, putative
TGGT1_010380	28	1096	nucleoporin, putative
TGGT1_065700	2	12	conserved hypothetical protein
TGME49_021910	2	12	AN1-like Zinc finger domain-containing protein
TGGT1_013120	15	869	WD domain, G-beta repeat-containing protein, putative
TGME49_057560	12	513	WD domain, G-beta repeat-containing protein
TGGT1_076810	3	9	conserved hypothetical protein
TGME49_055390	3	9	hypothetical protein
TGGT1_088700	1	1	conserved hypothetical protein
TGME49_112410	1	1	hypothetical protein
TGGT1_003820	2	6	hypothetical protein
TGME49_054380	2	6	ribosomal protein L11, putative
TGME49_107010	13	419	jmjC domain-containing protein
TGGT1_055200	12	418	hypothetical protein
TGME49_118510	28	702	N-ethylmaleimide-sensitive factor
TGGT1_122610	16	506	N-ethylmaleimide-sensitive factor, putative
TGME49_091130	1	1	glutamic acid-rich protein, putative
TGME49_093380	10	201	hypothetical protein
TGGT1_073740	8	192	histone acetyltransferase, putative
TGGT1_072920	1	4	conserved hypothetical protein
TGME49_080400	1	4	hypothetical protein
TGGT1_052840	2	4	cyclophilin, putative
TGME49_038000	1	1	cyclophilin, putative
TGME49_115130	3	8	protein-L-isoaspartate O-methyltransferase, putative
TGGT1_092560	1	3	protein-L-isoaspartate O-methyltransferase, putative
TGME49_105780	32	1339	exoribonuclease, putative
TGGT1_041790	27	1361	5'->3' exoribonuclease, putative
TGME49_048120	13	199	hypothetical protein
TGGT1_024820	10	153	hypothetical protein
TGME49_043570	7	109	40S ribosomal protein S26, putative
TGGT1_046280	4	68	40S ribosomal protein S26, putative
TGGT1_124900	6	65	protamine P1, putative
TGME49_014500	5	45	protamine P1, putative

TGGT1_045700	2	17	n6-DNA-methyltransferase, putative
TGME49_044240	1	9	hypothetical protein
TGGT1_052690	2	4	conserved hypothetical protein
TGME49_038170	2	4	hypothetical protein
TGGT1_083130	7	85	import inner membrane translocase subunit TIM44, putative
TGME49_027830	5	50	mitochondrial import inner membrane translocase TIM44, putative
TGME49_022110	6	28	UBA/TS-N domain-containing protein
TGGT1_065940	2	8	hypothetical protein
TGME49_021710	19	393	TBC domain-containing protein
TGGT1_065590	11	205	hypothetical protein
TGME49_091140	44	3583	CCR4-Not complex component, Not1 domain-containing protein, conserved
TGGT1_031980	40	3889	CCR4-not transcription complex, putative
TGME49_053470	6	43	hypothetical protein, conserved
TGGT1_002640	5	40	conserved hypothetical protein
TGME49_113370	57	8380	hypothetical protein
TGGT1_089770	50	10734	conserved hypothetical protein
TGME49_030910	4	18	hypothetical protein
TGGT1_116920	2	2	conserved hypothetical protein
TGGT1_087230	70	17304	hypothetical protein
TGME49_110900	61	12118	hypothetical protein
TGME49_019850	108	18992	prolyl-tRNA synthetase, putative
TGGT1_030860	50	10452	prolyl-tRNA synthetase, putative
TGME49_074060	38	22653	mitochondrial 2-oxoglutarate/malate carrier protein, putative
TGGT1_113250	14	579	mitochondrial 2-oxoglutarate/malate carrier protein, putative
TGME49_106660	194	26696	RNA pseudouridine synthase domain containing protein
TGGT1_040670	115	26897	RNA pseudouridylate synthase, putative
TGGT1_040940	27	3822	conserved hypothetical protein
TGME49_106400	25	2996	hypothetical protein, conserved
TGGT1_085840	13	534	conserved hypothetical protein
TGME49_109890	9	345	hypothetical protein
TGME49_095040	42	1918	importin subunit beta-1, putative
TGGT1_075390	23	1278	importin beta-1, putative
TGME49_094010	27	1131	yrdC domain protein
TGGT1_074400	21	1087	yrdC domain protein, putative
TGGT1_064530	2	6	SRS12B
TGGT1_032090	4	54	conserved hypothetical protein
TGME49_091030	4	54	hypothetical protein
TGME49_008830	17	319	hypothetical protein
TGGT1_021860	16	423	conserved hypothetical protein
TGGT1_004220	3	15	conserved hypothetical protein
TGME49_054670	2	8	hypothetical protein

TGGT1_056610	1	1	sterol-regulatory element binding protein site 2 protease, putative
TGGT1_114230	9	142	GTP-binding protein, putative
TGME49_033260	9	142	GTP-binding protein, putative
TGME49_111470	53	2852	hypothetical protein
TGGT1_087820	18	1260	conserved hypothetical protein
TGGT1_027070	7	96	conserved hypothetical protein
TGME49_045990	5	77	hypothetical protein
TGME49_040370	2	4	hypothetical protein
TGME49_046500	27	948	surp module domain-containing protein
TGGT1_026620	14	684	surp module domain-containing protein, putative
TGME49_090920	28	782	2OG-Fe(II) oxygenase family protein
TGGT1_032210	21	661	prolyl 4-hydroxylase alpha subunit, putative
TGME49_070220	11	260	hypothetical protein
TGGT1_108750	8	260	conserved hypothetical protein
TGME49_016680	15	298	ankyrin repeat-containing protein
TGGT1_098460	7	38	ankyrin repeat-containing protein, putative
TGGT1_110300	10	176	conserved hypothetical protein
TGME49_071430	8	158	hypothetical protein
TGGT1_095460	1	7	ubiquitin-conjugating enzyme E2, putative
TGME49_100040	1	7	ubiquitin-conjugating enzyme E2, putative
TGGT1_088130	3	54	conserved hypothetical protein
TGME49_111880	3	54	hypothetical protein
TGME49_042640	9	96	hypothetical protein
TGGT1_047280	7	111	conserved hypothetical protein
TGME49_073520	52	4666	proteasome PCI domain-containing protein
TGGT1_112600	29	2288	proteasome PCI domain-containing protein, putative
TGGT1_108180	3	10	conserved hypothetical protein
TGME49_069750	1	3	hypothetical protein
TGGT1_076330	33	4927	acetyl-coenzyme A synthetase, putative
TGME49_066640	22	3588	acetyl-coenzyme A synthetase, putative
TGME49_089820	6	66	TBC domain-containing protein
TGGT1_033410	6	55	TBC domain-containing protein, putative
TGME49_100150	7	195	hypothetical protein
TGGT1_095140	7	193	conserved hypothetical protein
TGME49_028180	6	34	cytochrome C oxidase assembly factor COX15, putative
TGGT1_029690	3	83	conserved hypothetical protein
TGME49_018730	3	83	hypothetical protein
TGGT1_111480	4	88	Radical SAM domain-containing protein, putative
TGME49_072270	2	26	radical SAM domain-containing protein
TGME49_062100	23	473	TPR domain-containing protein
TGGT1_008010	17	392	TPR domain-containing protein, putative
TGME49_061480	8	896	phosphatidylserine synthase, putative
TGGT1_008750	7	971	phosphatidylserine synthase, putative

TGME49_061440	21	967	armadillo/beta-catenin-like repeat-containing protein
TGGT1_008800	8	241	conserved hypothetical protein
TGGT1_115430	3	17	conserved hypothetical protein
TGME49_032210	3	17	hypothetical protein
TGME49_067350	14	190	small nuclear ribonucleoprotein, putative
TGGT1_075690	3	19	small nuclear ribonucleoprotein, putative
TGME49_002490	23	587	hypothetical protein
TGGT1_018600	16	615	conserved hypothetical protein
TGGT1_036620	1	4	conserved hypothetical protein
TGME49_084800	1	4	hypothetical protein
TGME49_024920	4	59	hypothetical protein, conserved
TGGT1_079830	4	26	conserved hypothetical protein
TGME49_078740	7	27	diaminopimelate decarboxylase, putative
TGGT1_103830	3	23	pyridoxal-dependent decarboxylase, putative
TGME49_062920	55	2745	trypsin, putative
TGGT1_007070	25	1898	trypsin, putative
TGGT1_078010	2	8	conserved hypothetical protein
TGME49_023660	2	8	50S ribosomal protein L4, putative
TGGT1_100560	8	72	actin, putative
TGME49_048890	8	72	actin-like protein 3b, putative
TGME49_105190	8	33	corA-like Mg2+ transporter domain-containing protein
TGME49_052360	26	976	Rhoptry kinase family protein ROP24 (incomplete catalytic triad)
TGGT1_001270	14	594	conserved hypothetical protein
TGGT1_064330	16	1043	conserved hypothetical protein
TGME49_121300	12	709	hypothetical protein
TGGT1_058110	10	96	ribosome biogenesis protein tsr1, putative
TGME49_065110	10	96	hypothetical protein
TGGT1_108860	1	2	cyclin dependent kinase, putative
TGGT1_085890	9	139	microtubule-associated protein, putative
TGME49_109950	8	116	microtubule-associated protein, putative
TGME49_037280	55	10525	TLD domain-containing protein
TGGT1_068620	51	12179	TLD domain-containing protein, putative
TGGT1_078590	7	69	conserved hypothetical protein
TGME49_024130	7	69	hypothetical protein
TGME49_067300	8	136	hypothetical protein
TGGT1_075740	2	19	conserved hypothetical protein
TGME49_110490	12	496	ribosomal protein L22, putative
TGGT1_086710	4	74	ribosomal protein L22, putative
TGME49_096010	22	1591	phosphatidylinositol 3- and 4-kinase domain-containing protein
TGGT1_054000	22	1590	phosphatidylinositol 3- and 4-kinase domain-containing protein, putative
TGME49_080660	11	271	hypothetical protein
TGGT1_072650	10	270	ubiquitin-protein ligase, putative

TGME49_015040	29	953	hypothetical protein
TGGT1_125670	24	936	tip120, putative
TGGT1_068340	1	4	KH domain-containing protein, putative
TGME49_037550	1	4	KH domain-containing protein
TGME49_046510	10	58	serine/threonine protein phosphatase 2A B56 delta subunit, putative
TGGT1_026610	6	36	serine/threonine protein phosphatase 2A B56 delta subunit, putative
TGME49_095380	8	80	hypothetical protein
TGGT1_054560	5	55	conserved hypothetical protein
TGGT1_045190	3	39	conserved hypothetical protein
TGME49_044640	3	39	hypothetical protein
TGGT1_085250	3	5	conserved hypothetical protein
TGME49_109380	3	5	nuf2 family domain-containing protein
TGME49_089990	11	89	hypothetical protein
TGGT1_033150	7	83	conserved hypothetical protein
TGME49_008030	58	25480	microneme protein MIC4
TGGT1_021130	21	2370	microneme protein, putative
TGME49_078990	79	58600	phosphate carrier protein, putative
TGGT1_103360	14	797	phosphate carrier protein, putative
TGME49_073450	14	76	sufB/sufD domain-containing protein
TGGT1_112530	2	6	sufB/sufD domain-containing protein, putative
TGME49_042830	28	492	5'-3' exonuclease, putative
TGGT1_047080	20	496	conserved hypothetical protein
TGME49_072600	9	97	hypothetical protein
TGGT1_111890	8	95	conserved hypothetical protein
TGGT1_094260	2	4	longevity-assurance LAG1 domain-containing protein
TGME49_116450	2	4	longevity-assurance domain-containing protein
TGGT1_110060	1	1	bicoid-interacting (Bin3) domain-containing protein
TGME49_071190	1	1	Bin3 domain-containing protein
TGME49_118370	4	12	hypothetical protein
TGGT1_122870	1	3	conserved hypothetical protein
TGME49_091020	3	5	IQ calmodulin-binding motif domain-containing protein
TGGT1_032100	1	1	calmodulin-binding domain containing protein, putative
TGGT1_033230	4	30	conserved hypothetical protein
TGME49_089910	4	30	hypothetical protein, conserved
TGME49_043620	12	130	hypothetical protein
TGGT1_046250	5	30	conserved hypothetical protein
TGME49_033940	1	1	leucine rich repeat protein, putative
TGME49_118710	34	2236	ATP-binding cassette sub-family F member 1
TGGT1_122410	23	2070	ABC transporter, putative
TGGT1_079950	3	26	conserved hypothetical protein
TGME49_025010	3	26	hypothetical protein
TGGT1_015020	6	91	conserved hypothetical protein
TGME49_020320	4	77	hypothetical protein

TGGT1_097020	7	72	hypothetical protein
TGME49_097200	5	38	hypothetical protein
TGME49_078870	154	42767	myosin F (TgMyoF) protein
TGGT1_103490	58	9770	conserved hypothetical protein
TGGT1_009570	1	1	conserved hypothetical protein
TGME49_060610	1	1	hypothetical protein
TGME49_113380	32	16409	hypothetical protein, conserved
TGGT1_089780	18	1669	conserved hypothetical protein
TGME49_063650	7	24	pre-mRNA splicing factor RNA helicase, putative
TGGT1_043900	3	15	hypothetical protein
TGME49_070270	2	10	hypothetical protein
TGGT1_108810	2	8	conserved hypothetical protein
TGME49_114950	7	34	NAD-dependent DNA ligase OB-fold domain containing protein
TGME49_120210	34	1870	WD-40 repeat protein, putative
TGGT1_121080	25	1464	WD-repeat protein, putative
TGGT1_102710	17	327	eukaryotic translation initiation factor 3 subunit, putative
TGME49_051410	15	267	hypothetical protein
TGME49_043390	9	139	hypothetical protein
TGGT1_046470	4	10	conserved hypothetical protein
TGGT1_100690	3	6	kinesin motor domain-containing protein, putative
TGME49_049020	2	5	kinesin motor domain-containing protein
TGME49_014620	12	174	hypothetical protein, conserved
TGGT1_125030	4	23	conserved hypothetical protein
TGME49_106910	9	34	hypothetical protein
TGGT1_055500	4	14	conserved hypothetical protein
TGME49_012290	20	1386	40s ribosomal protein s19, putative
TGGT1_072140	10	603	40S ribosomal protein s19, putative
TGME49_118230	85	17504	phosphoglycerate kinase, putative
TGGT1_123010	32	6978	phosphoglycerate kinase, putative
TGGT1_003920	1	3	conserved hypothetical protein
TGME49_054460	1	3	hypothetical protein
TGGT1_065510	4	49	conserved hypothetical protein
TGME49_021640	2	22	ppg3, putative
TGGT1_007410	5	22	conserved hypothetical protein
TGME49_062680	4	17	hypothetical protein
TGGT1_065520	20	642	ATP-dependent RNA helicase, putative
TGME49_021650	17	472	ATP-dependent RNA helicase, putative
TGGT1_105120	1	8	conserved hypothetical protein
TGGT1_028660	5	41	protease inhibitor, putative
TGME49_017430	4	25	protease inhibitor, putative
TGME49_088750	25	1603	ubiquinol-cytochrome C reductase complex 14 kDa protein, putative
TGGT1_034520	5	62	conserved hypothetical protein
TGME49_042840	5	42	hypothetical protein

TGGT1_047070	1	2	conserved hypothetical protein
TGME49_075650	82	13344	p97 protein
TGGT1_000680	52	8752	p97 protein, putative
TGME49_099190	3	17	B-box zinc finger protein, putative
TGME49_014410	5	28	hypothetical protein, conserved
TGGT1_124820	3	24	conserved hypothetical protein
TGGT1_124120	1	2	phenylalanyl-tRNA synthetase, putative
TGME49_010750	1	2	hypothetical protein
TGME49_120200	7	58	SRS16A
TGGT1_121090	3	15	SRS16A
TGME49_002380	39	1683	rab GDP dissociation inhibitor alpha, putative
TGGT1_018700	14	763	RAB GDP dissociation inhibitor alpha, putative
TGME49_018920	19	380	proteasome subunit beta type 5, putative
TGGT1_029890	6	87	proteasome subunit beta type, putative
TGME49_063320	18	462	hypothetical protein
TGGT1_043440	12	310	conserved hypothetical protein
TGME49_021230	8	53	hypothetical protein
TGGT1_065070	6	48	conserved hypothetical protein
TGME49_110440	17	357	phosphatidylinositol-4-phosphate 5-kinase, putative
TGGT1_086660	13	300	phosphatidylinositol-4-phosphate 5-kinase, putative
TGGT1_085270	3	8	structural maintenance of chromosomes 6 smc6, putative
TGME49_109400	1	3	chromosome segregation protein, putative
TGME49_051950	5	160	ribosomal protein L7/L12, putative
TGGT1_103330	4	38	ribosomal protein L7/L12, putative
TGME49_025700	3	135	hypothetical protein
TGGT1_080830	2	130	conserved hypothetical protein
TGGT1_109880	2	9	conserved hypothetical protein
TGME49_071000	1	4	hypothetical protein
TGGT1_036750	18	699	conserved hypothetical protein
TGME49_085250	18	699	hypothetical protein
TGGT1_036640	6	198	conserved hypothetical protein
TGME49_085150	6	198	hypothetical protein
TGME49_039400	52	12442	hypothetical protein
TGGT1_051470	48	15658	conserved hypothetical protein
TGGT1_001120	1	1	DHHC domain-containing protein, putative
TGME49_052200	1	1	zinc finger DHHC domain-containing protein
TGGT1_047090	7	124	conserved hypothetical protein
TGME49_042820	6	95	hypothetical protein
TGME49_069980	25	3695	protein transport protein Sec61 alpha subunit isoform 1, putative
TGGT1_108520	9	362	protein transport protein Sec61 alpha subunit isoform, putative
TGGT1_080880	5	51	conserved hypothetical protein
TGME49_025750	5	31	hypothetical protein
TGME49_112280	18	349	ATP-dependent RNA helicase, putative

TGGT1_088540	17	345	ATP-dependent RNA helicase, putative
TGME49_120570	49	3126	G1 to S phase transition protein, putative
TGGT1_120820	26	2464	G1 to S phase transition protein, putative
TGGT1_101460	42	2428	conserved hypothetical protein
TGME49_049780	40	2311	hypothetical protein
TGME49_014950	2	63	hypothetical protein
TGGT1_125370	1	61	conserved hypothetical protein
TGME49_070120	6	70	thioredoxin, putative
TGGT1_108660	4	54	thioredoxin, putative
TGME49_085950	17	185	hypothetical protein
TGGT1_037340	9	204	conserved hypothetical protein
TGME49_020250	7	67	chloride channel, nucleotide-sensitive, 1A, putative
TGGT1_014930	6	63	methylosome subunit pICln, putative
TGME49_035420	53	3488	hypothetical protein
TGGT1_070460	43	3624	conserved hypothetical protein
TGME49_027560	15	421	hypothetical protein
TGGT1_082970	14	370	conserved hypothetical protein
TGME49_115100	2	6	hypothetical protein
TGGT1_028620	1	2	conserved hypothetical protein
TGME49_017470	1	2	hypothetical protein
TGGT1_071840	3	41	molybdenum cofactor synthesis protein, putative
TGME49_012100	2	23	molybdenum cofactor synthesis protein, putative
TGME49_000460	4	32	hypothetical protein
TGGT1_060680	3	31	conserved hypothetical protein
TGGT1_012960	1	2	methionine aminopeptidase, putative
TGME49_057730	1	2	methionine aminopeptidase, putative
TGGT1_050380	3	7	conserved hypothetical protein
TGME49_040490	3	7	hypothetical protein
TGME49_026570	9	2059	hypothetical protein
TGGT1_081720	4	1766	conserved hypothetical protein
TGGT1_058210	1	3	choline/ethanolamine kinase, putative
TGME49_065000	1	3	choline/ethanolamine kinase, putative
TGME49_030100	6	25	hypothetical protein
TGGT1_117790	2	8	conserved hypothetical protein
TGME49_112600	10	331	small heat shock protein 21
TGGT1_088890	9	427	small heat shock protein / bRadyzoite-specific protein, putative
TGGT1_041310	9	158	zinc finger (CHY type) protein
TGME49_106040	9	94	zinc finger (CHY type) protein, putative
TGME49_026350	8	65	hypothetical protein
TGGT1_081520	6	90	conserved hypothetical protein
TGME49_097150	7	41	MORN repeat-containing protein
TGGT1_097070	6	40	MORN repeat-containing protein, putative
TGME49_119350	26	3131	SRS domain containing protein
TGGT1_121850	13	576	SRS domain containing-protein

TGME49_046970	19	370	exosome component 10, putative
TGGT1_026040	16	591	3'-5' exonuclease domain-containing protein, putative
TGGT1_021930	5	20	conserved hypothetical protein
TGME49_008990	4	13	hypothetical protein
TGME49_057430	21	772	hypothetical protein
TGGT1_013270	20	766	conserved hypothetical protein
TGGT1_009210	5	25	conserved hypothetical protein
TGME49_061060	4	16	putative diphthamide synthesis domain-containing protein
TGGT1_122890	1	2	pumilio, putative
TGME49_118350	1	2	pumilio-family RNA binding repeat domain-containing protein
TGME49_114480	21	379	hypothetical protein
TGGT1_091270	17	368	conserved hypothetical protein
TGGT1_070380	25	1187	conserved hypothetical protein
TGME49_035490	23	836	hypothetical protein
TGME49_094180	11	92	tRNA (guanine-N(7)-)-methyltransferase, putative
TGGT1_074480	7	52	tRNA (guanine-N(7)-)-methyltransferase, putative
TGME49_061540	5	38	DNA-directed RNA polymerases I and III subunit RPAC2, putative
TGGT1_008690	4	21	DNA-directed RNA polymerase, putative
TGME49_094200	55	11143	glucose-6-phosphate dehydrogenase
TGGT1_074500	24	1880	glucose-6-phosphate dehydrogenase, putative
TGGT1_125280	3	15	ribosomal protein L9, N-terminal domain protein, putative
TGME49_014870	2	8	ribosomal protein L9, N-terminal domain protein
TGME49_112160	5	77	hypothetical protein
TGGT1_088410	1	4	conserved hypothetical protein
TGME49_073870	40	1377	chromatin remodelling complex protein SNF2L, putative
TGGT1_112960	32	1360	chromatin remodelling complex protein SNF2L, putative
TGGT1_001970	8	238	engulfment and cell motility, putative
TGME49_053000	7	162	ELMO/CED-12 family domain-containing protein
TGGT1_121450	1	1	synaptosomal-associated protein, putative
TGME49_119940	1	1	synaptosomal-associated protein 25, putative
TGGT1_112580	11	187	signal transduction protein, putative
TGME49_073500	11	153	signal transduction protein, putative
TGGT1_102140	30	1767	conserved hypothetical protein
TGME49_050700	28	1525	hypothetical protein
TGME49_022350	11	125	hypothetical protein
TGGT1_066160	8	77	conserved hypothetical protein
TGGT1_114900	10	158	hypothetical protein
TGME49_032620	8	105	hypothetical protein
TGME49_027360	47	9619	60S ribosomal protein L3, putative
TGGT1_082860	21	1967	60S ribosomal protein L3, putative
TGGT1_023010	29	1488	glycan synthetase, putative

TGME49_009960	28	1418	glycan synthetase, putative
TGME49_109110	2	26	tRNA methyltransferase domain-containing protein
TGME49_119500	1	1	hypothetical protein, conserved
TGME49_020260	8	136	hypothetical protein, conserved
TGGT1_014940	7	117	conserved hypothetical protein
TGME49_008720	2	4	phosphatase, putative
TGME49_032300	37	3395	40S ribosomal protein S3, putative
TGGT1_115350	13	1110	40S ribosomal protein S3, putative
TGGT1_086200	10	218	RNA binding protein, putative
TGME49_110050	9	189	RNA binding protein, putative
TGME49_033460	85	1124216	SRS29B (= SAG1, P30)
TGGT1_114010	69	14134	SRS29B (= SAG1, P30)
TGGT1_094330	14	239	1,4-alpha-glucan branching enzyme, putative
TGME49_116520	13	199	1,4-alpha-glucan branching enzyme, putative
TGGT1_047300	13	186	conserved hypothetical protein
TGME49_042620	12	137	hypothetical protein
TGME49_008970	12	493	hypothetical protein
TGGT1_021910	12	492	serine/arginine rich splicing factor, putative
TGME49_040090	12	305	Rhoptry kinase family protein ROP34, putative
TGGT1_050750	8	111	protein kinase, putative
TGGT1_020170	22	905	copper-transporting ATPase, putative
TGME49_001150	21	629	copper-transporting ATPase, putative
TGME49_029010	78	16530	hypothetical protein
TGGT1_118730	38	6748	conserved hypothetical protein
TGME49_057090	20	1529	NAC domain containing protein
TGGT1_013620	9	1042	NAC domain containing protein, putative
TGME49_020100	47	5158	phosphoribosylpyrophosphate synthetase, putative
TGGT1_014600	18	851	phosphoribosylpyrophosphate synthetase, putative
TGME49_058470	27	4806	hypothetical protein
TGGT1_011740	15	1808	conserved hypothetical protein
TGGT1_106200	1	1	conserved hypothetical protein
TGGT1_058260	2	3	conserved hypothetical protein
TGME49_064960	1	1	hypothetical protein
TGGT1_075310	4	45	conserved hypothetical protein
TGME49_094960	4	45	hypothetical protein
TGGT1_040870	1	5	isoprenylcysteine carboxylmethyltransferase, putative
TGME49_106470	1	5	isoprenylcysteine carboxylmethyltransferase, putative
TGME49_085200	13	104	cleavage and polyadenylation specificity factor, putative
TGGT1_036700	9	106	cleavage and polyadenylation specificity factor, putative
TGGT1_006020	1	1	serine/threonine-protein kinase, putative
TGME49_081430	1	1	protein kinase, PfPK7 homolog
TGME49_049720	1	24	ATP synthase lipid-binding protein, putative
TGGT1_050470	5	28	hypothetical protein

TGME49_040400	5	28	hypothetical protein
TGGT1_045805	10	191	hypothetical protein
TGME49_044150	9	142	glycerate kinase, putative
TGME49_016960	9	53	elongation factor Tu GTP-binding domain-containing protein
TGGT1_098170	6	44	GTP-binding domain containing protein, putative
TGGT1_054440	1	3	conserved hypothetical protein
TGME49_095590	1	3	hypothetical protein
TGGT1_044250	3	76	nucleic acid binding protein, putative
TGME49_064010	3	76	nucleic acid binding protein, putative
TGME49_091310	1	1	vacuolar ATP synthase 22 kDa proteolipid subunit, putative
TGME49_019530	4	6	hypothetical protein, conserved
TGGT1_102350	17	469	conserved hypothetical protein
TGME49_050800	17	361	hypothetical protein
TGME49_098020	23	693	DEAD-box ATP-dependent RNA helicase, putative
TGGT1_096090	22	827	DEAD-box helicase family protein
TGME49_110150	65	62198	long-chain-fatty-acid-CoA ligase, putative
TGGT1_086370	23	1167	long-chain-fatty-acid--CoA ligase, putative
TGGT1_033430	4	36	ATP-dependent heat shock protein, putative
TGME49_089780	4	36	ATP-dependent heat shock protein, putative
TGME49_112530	16	363	splicing factor protein, putative
TGGT1_088820	12	383	splicing factor protein, putative
TGME49_037230	6	135	hypothetical protein
TGGT1_068680	5	112	conserved hypothetical protein
TGME49_090170	19	189	protein serine/threonine phosphatase, putative / sortilin
TGGT1_032960	9	115	bsu-protein phosphatase, putative
TGGT1_092130	7	93	conserved hypothetical protein
TGME49_114840	6	73	ubiquitin carboxyl-terminal hydrolase domain-containing protein
TGME49_009440	9	105	hypothetical protein
TGGT1_022380	8	73	conserved hypothetical protein
TGGT1_042140	14	263	acetyltransferase domain-containing protein, putative
TGME49_105450	12	182	acetyltransferase domain-containing protein
TGGT1_069500	7	144	protein phosphatase 2C, putative
TGME49_036260	3	32	protein phosphatase 2C, putative
TGGT1_102910	2	3	conserved hypothetical protein
TGME49_051510	1	2	ulp1 protease family, C-terminal catalytic domain-containing protein
TGGT1_082430	19	609	conserved hypothetical protein
TGME49_026950	19	334	hypothetical protein, conserved
TGME49_105800	10	845	hypothetical protein, conserved
TGGT1_041770	10	832	conserved hypothetical protein
TGGT1_018230	8	143	conserved hypothetical protein
TGME49_002740	6	78	hypothetical protein
TGME49_002970	2	4	hypothetical protein

TGGT1_055490	9	384	conserved hypothetical protein
TGME49_106920	8	316	hypothetical protein
TGME49_060820	14	518	hypothetical protein
TGGT1_009340	9	336	conserved hypothetical protein
TGME49_073660	6	41	hypothetical protein
TGGT1_112740	5	39	conserved hypothetical protein
TGME49_075660	1	1	hypothetical protein
TGME49_027000	35	1218	hypothetical protein
TGGT1_082490	17	850	conserved hypothetical protein
TGGT1_057700	8	224	conserved hypothetical protein
TGME49_065400	4	128	hypothetical protein
TGME49_077500	28	1102	26S proteasome AAA-ATPase subunit RPT1, putative
TGGT1_105160	18	906	26S proteasome AAA-ATPase subunit RPT1, putative
TGGT1_016760	10	73	conserved hypothetical protein
TGME49_003950	4	17	myb-like DNA-binding / high mobility group box domain-containing protein
TGGT1_027560	14	275	dipeptidyl peptidase IV, putative
TGME49_045500	10	159	dipeptidyl peptidase IV domain-containing protein
TGME49_009080	1	1	transport protein particle component Bet3 domain-containing protein
TGME49_036530	9	73	hypothetical protein
TGGT1_069320	7	69	conserved hypothetical protein
TGGT1_090310	7	89	conserved hypothetical protein
TGME49_113810	7	89	hypothetical protein
TGME49_067050	14	151	hypothetical protein
TGGT1_075880	9	111	conserved hypothetical protein
TGGT1_093430	8	146	SRS54
TGME49_115740	4	29	SRS54
TGGT1_063440	2	7	hypothetical protein
TGME49_005460	2	7	AN1-like Zinc finger-containing protein
TGME49_044550	7	75	hypothetical protein
TGGT1_045390	4	66	conserved hypothetical protein
TGME49_005690	2	8	hypothetical protein
TGGT1_063220	1	3	hypothetical protein
TGGT1_017970	18	574	conserved hypothetical protein
TGME49_002950	18	413	hypothetical protein, conserved
TGGT1_002710	1	3	conserved hypothetical protein
TGME49_085180	19	302	hypothetical protein
TGGT1_036680	11	274	conserved hypothetical protein
TGME49_036950	10	281	hypothetical protein
TGGT1_068970	5	132	conserved hypothetical protein
TGME49_100280	4	34	small nuclear ribonucleoprotein, putative
TGGT1_095010	3	28	small nuclear ribonucleoprotein, putative
TGME49_118440	123	9754	DEAH-box RNA/DNA helicase, putative
TGGT1_122680	56	6238	DEAH-box RNA/DNA helicase, putative

TGME49_040280	26	1230	hypothetical protein
TGGT1_050570	12	210	conserved hypothetical protein
TGGT1_056640	3	18	ATP-dependent RNA helicase, putative
TGME49_066110	3	18	ATP-dependent RNA helicase, putative
TGME49_035130	10	359	hypothetical protein
TGGT1_070970	8	348	conserved hypothetical protein
TGME49_088500	55	9365	malate:quinone oxidoreductase, putative
TGGT1_034890	17	1012	malate:quinone oxidoreductase, putative
TGGT1_083720	10	113	conserved hypothetical protein
TGME49_028230	8	74	hypothetical protein
TGME49_044910	18	1359	MIZ zinc finger protein, putative
TGGT1_044900	16	1355	MIZ zinc finger protein, putative
TGME49_054470	25	618	hypothetical protein
TGGT1_004030	18	431	conserved hypothetical protein
TGGT1_045950	1	2	conserved hypothetical protein
TGME49_044010	1	2	hypothetical protein
TGME49_057740	15	494	UMP-CMP kinase, putative
TGGT1_012950	13	455	UMP-CMP kinase, putative
TGGT1_047390	2	9	hypothetical protein
TGME49_042430	2	9	hypothetical protein
TGME49_115170	7	132	soluble liver antigen/liver pancreas antigen domain-containing protein
TGGT1_092600	5	194	soluble liver antigen/liver pancreas antigen domain-containing protein, putative
TGME49_039760	10	611	60S ribosomal protein L22, putative
TGGT1_051070	5	170	60S ribosomal protein L22, putative
TGGT1_044380	14	576	conserved hypothetical protein
TGME49_064140	13	477	hypothetical protein
TGME49_068200	37	3221	RNA recognition motif domain-containing protein
TGGT1_106270	19	2160	RNA recognition motif domain-containing protein, putative
TGME49_048480	20	1167	40S ribosomal protein S9, putative
TGGT1_024480	8	312	40S ribosomal protein S9, putative
TGGT1_101700	1	1	conserved hypothetical protein
TGME49_050010	1	1	hypothetical protein, conserved
TGME49_084660	1	1	hypothetical protein, conserved
TGME49_009060	8	98	thrombospondin type 1 domain-containing protein
TGGT1_022000	6	93	thrombospondin type 1 domain-containing protein, putative
TGME49_080410	21	663	3',5'-cyclic phosphodiesterase, putative
TGGT1_072910	12	294	3',5'-cyclic phosphodiesterase, putative
TGME49_110760	12	44	protein phosphatase 2C, putative
TGGT1_087070	5	25	protein phosphatase 2C, putative
TGGT1_000440	3	13	conserved hypothetical protein
TGME49_076100	2	5	hypothetical protein
TGME49_017820	35	2253	proteasome PCI domain-containing protein
TGGT1_028210	16	1327	proteasome PCI domain-containing protein,

			putative
TGGT1_002800	2	6	conserved hypothetical protein
TGME49_053630	2	6	hypothetical protein
TGME49_092110	5	26	formate/nitrite transporter, putative
TGGT1_031460	4	25	formate/nitrite transporter, putative
TGME49_039590	8	111	WD domain, G-beta repeat-containing protein
TGGT1_051270	6	75	WD domain, G-beta repeat-containing protein, putative
TGME49_026020	11	884	transmembrane domain-containing protein
TGGT1_081160	9	875	transmembrane domain-containing protein, putative
TGME49_062930	2	29	hypothetical protein
TGGT1_007060	1	2	conserved hypothetical protein
TGME49_072660	24	1060	hypothetical protein
TGGT1_111950	12	548	conserved hypothetical protein
TGGT1_088790	5	62	conserved hypothetical protein
TGME49_112500	5	62	hypothetical protein
TGGT1_063150	1	1	conserved hypothetical protein
TGME49_005750	1	1	histone deacetylase interacting domain-containing protein
TGME49_028400	21	667	hypothetical protein
TGGT1_083970	19	758	conserved hypothetical protein
TGME49_049590	30	1802	proteasome subunit alpha type 5, putative
TGGT1_101260	14	1028	proteasome subunit alpha type, putative
TGGT1_078750	89	24908	PHD-finger domain-containing protein, putative
TGME49_024260	79	17094	PHD-finger domain-containing protein
TGGT1_082280	6	62	conserved hypothetical protein
TGME49_026900	6	43	hypothetical protein
TGME49_094780	8	45	hypothetical protein
TGGT1_075100	4	8	conserved hypothetical protein
TGGT1_050420	5	49	maf protein, putative
TGME49_040450	5	49	maf-like protein, putative
TGGT1_108060	7	350	conserved hypothetical protein
TGME49_069630	4	207	hypothetical protein
TGME49_110550	1	2	hypothetical protein
TGME49_059000	7	37	hypothetical protein
TGGT1_010870	4	24	conserved hypothetical protein
TGGT1_113190	8	59	hypothetical protein
TGME49_074010	5	25	hypothetical protein
TGME49_078660	80	27174	P-type Ca(2+)-ATPase, putative
TGGT1_103910	49	9020	P-type ATPase, putative
TGME49_068640	3	9	WD domain, G-beta repeat-containing protein
TGGT1_106820	2	5	WD domain, G-beta repeat-containing protein, putative
TGME49_104970	47	1427	CAM kinase, CDPK family TgPK1
TGGT1_042520	21	967	protein kinase domain-containing protein, putative
TGME49_036070	11	220	pyrroline-5-carboxylate reductase, putative
TGGT1_069690	6	130	pyrroline-5-carboxylate reductase, putative

TGGT1_043960	1	1	conserved hypothetical protein
TGME49_063690	1	1	hypothetical protein
TGME49_101160	10	78	SRS19C (= SRS7)
TGGT1_059580	4	13	SRS19C (= SRS7)
TGGT1_021510	4	18	conserved hypothetical protein
TGME49_008490	4	18	hypothetical protein
TGME49_090160	66	13712	sortilin, putative
TGGT1_032970	28	2067	BNR/Asp-box repeat domain-containing protein, putative
TGME49_030010	10	92	hypothetical protein
TGGT1_117870	9	109	conserved hypothetical protein
TGGT1_009320	1	1	splicing factor u2af large subunit, putative
TGME49_060840	1	1	hypothetical protein
TGGT1_079290	1	4	conserved hypothetical protein
TGME49_024570	1	4	hypothetical protein
TGME49_063480	4	59	Na+/H+ exchanger, putative
TGME49_095360	31	7891	hypothetical protein
TGGT1_054580	21	4836	conserved hypothetical protein
TGGT1_010950	4	30	conserved hypothetical protein
TGME49_058920	2	10	hypothetical protein
TGGT1_110250	9	139	conserved hypothetical protein
TGME49_071380	6	86	hypothetical protein
TGME49_032160	6	20	zinc finger (C3HC4 RING finger) protein, putative
TGME49_011060	8	137	hypothetical protein
TGGT1_123870	2	26	conserved hypothetical protein
TGME49_014080	16	1634	toxofilin
TGGT1_124310	13	1094	toxofilin, putative
TGGT1_112450	3	8	conserved hypothetical protein
TGME49_073380	3	7	hypothetical protein
TGME49_031770	19	148	tubulin alpha chain, putative
TGGT1_116000	1	1	tubulin alpha chain, putative
TGGT1_109620	25	1499	conserved hypothetical protein
TGME49_070870	23	1440	hypothetical protein
TGGT1_035080	1	2	adenylate kinase, putative
TGME49_088320	1	2	adenylate kinase, putative
TGME49_097730	12	133	hypothetical protein, conserved
TGGT1_096450	6	118	conserved hypothetical protein
TGME49_111240	50	4583	DnaJ domain-containing protein
TGGT1_087560	26	2490	DnaJ domain-containing protein, putative
TGGT1_101010	7	40	conserved hypothetical protein
TGME49_049340	6	33	hypothetical protein
TGME49_099110	36	3279	cleft lip and palate transmembrane protein 1, putative
TGGT1_006340	12	408	conserved hypothetical protein
TGME49_033080	7	18	tuftelin interacting protein 11, putative
TGGT1_114430	6	21	tuftelin interacting protein, putative
TGGT1_108090	5	129	helicase, putative

TGME49_069660	3	56	TFIIH basal transcription factor complex helicase XPB subunit, putative
TGGT1_041480	17	497	conserved hypothetical protein
TGME49_106000	16	393	hypothetical protein
TGGT1_118180	21	1463	TPR domain-containing protein, putative
TGME49_029490	18	1176	TPR domain-containing protein
TGME49_075670	6	60	hypothetical protein, conserved
TGGT1_000660	2	27	conserved hypothetical protein
TGME49_113920	11	309	hypothetical protein
TGGT1_090560	10	424	conserved hypothetical protein
TGME49_072640	19	1083	translation initiation factor eIF-2B epsilon subunit, putative
TGGT1_111930	16	1158	translation initiation factor eIF-2B epsilon subunit, putative
TGGT1_033450	1	1	biotin protein ligase, putative
TGME49_089760	1	1	biotin-[acetyl-CoA-carboxylase] ligase, putative
TGGT1_031900	3	6	conserved hypothetical protein
TGME49_091320	1	1	hypothetical protein
TGGT1_031740	3	17	ADP-ribosylation factor, arf, putative
TGME49_091800	3	12	ADP-ribosylation factor-like protein 1, putative
TGGT1_007390	8	70	TPR domain containing protein, putative
TGME49_062700	7	62	TPR domain containing protein
TGGT1_033910	12	157	conserved hypothetical protein
TGME49_089310	12	157	cullin family protein
TGME49_017570	3	26	40s ribosomal protein S27, putative
TGGT1_028490	1	7	40S ribosomal protein S27, putative
TGME49_030700	9	106	hypothetical protein, conserved
TGGT1_117050	8	125	conserved hypothetical protein
TGGT1_113700	1	4	conserved hypothetical protein
TGME49_007710	2	18	phosphatidylinositol synthase, putative
TGGT1_020590	1	8	phosphatidylinositol synthase, putative
TGGT1_010260	18	977	conserved hypothetical protein
TGME49_059870	16	874	hypothetical protein
TGME49_051770	22	235	hypothetical protein
TGGT1_103140	6	66	conserved hypothetical protein
TGGT1_009940	2	8	conserved hypothetical protein
TGGT1_089010	9	122	conserved hypothetical protein
TGME49_112820	8	102	hypothetical protein
TGGT1_035220	2	6	conserved hypothetical protein
TGME49_088190	2	6	hypothetical protein
TGGT1_000710	6	33	conserved hypothetical protein
TGME49_075620	5	27	hypothetical protein
TGME49_019280	20	718	actin, putative
TGGT1_030250	15	640	actin, putative
TGME49_118310	64	5593	transketolase, putative
TGGT1_122940	37	4162	transketolase, putative
TGME49_024050	15	240	hypothetical protein

TGGT1_078530	13	308	conserved hypothetical protein
TGGT1_030340	22	604	diphteria toxin resistance protein, putative
TGME49_019460	16	322	hypothetical protein
TGGT1_098860	1	2	RNA methyltransferase, putative
TGME49_016390	1	2	RNA methyltransferase, putative
TGGT1_052000	1	1	retinitis pigmentosa GTPase regulator, putative
TGME49_038980	1	1	retinitis pigmentosa GTPase regulator, putative
TGGT1_035990	16	2617	DEAH-box RNA/DNA helicase, putative
TGME49_084050	15	1890	DEAH-box RNA/DNA helicase, putative
TGGT1_106280	2	36	phosphoinositide-dependent protein kinase, putative
TGME49_068210	2	36	AGC kinase
TGME49_021670	49	2992	transcription elongation factor FACT 140 kDa, putative
TGGT1_065540	38	2538	cell division control protein, putative
TGGT1_025320	3	26	conserved hypothetical protein
TGME49_047590	3	26	methyltransferase domain containing protein
TGGT1_068710	7	83	serine-threonine protein kinase, putative
TGME49_037210	6	79	Tyrosine kinase-like (TKL) protein
TGGT1_113440	4	22	serine/threonine-protein kinase, putative
TGME49_033900	4	22	CAM kinase, SNF1 family
TGME49_105290	19	860	vacuolar ATP synthase subunit E, putative
TGGT1_042210	11	322	vacuolar ATP synthase subunit E, putative
TGME49_116700	7	26	hypothetical protein
TGGT1_094540	6	31	conserved hypothetical protein
TGGT1_041530	4	13	peptidyl-prolyl cis-trans isomerase, putative
TGME49_105940	4	13	peptidyl-prolyl cis-trans isomerase, putative
TGME49_104490	4	20	hypothetical protein
TGGT1_006630	2	6	conserved hypothetical protein
TGME49_086510	7	19	hypothetical protein
TGGT1_037920	4	4	conserved hypothetical protein
TGGT1_054940	1	2	uracil-DNA glycosylase, putative
TGME49_107650	1	2	uracil-DNA glycosylase, putative
TGGT1_059670	8	149	conserved hypothetical protein
TGME49_101250	7	113	hypothetical protein
TGGT1_003870	2	4	conserved hypothetical protein
TGME49_054420	2	4	patatin-like phospholipase domain-containing protein
TGME49_024220	5	25	serine/threonine protein phosphatase, putative
TGGT1_078720	4	24	serine/threonine protein phosphatase, putative
TGGT1_019800	2	3	protein phosphatase 2C, putative
TGME49_001520	2	3	protein phosphatase 2C, putative
TGGT1_125910	8	88	serine/threonine protein phosphatase, putative
TGME49_015170	7	74	serine/threonine protein phosphatase, putative
TGME49_026980	8	338	hypothetical protein
TGGT1_082460	7	300	conserved hypothetical protein
TGME49_048340	44	4498	GTP-binding nuclear protein RAN/TC4, putative
TGGT1_024610	17	1831	GTP-binding nuclear protein RAN/TC4, putative

TGGT1_086750	1	2	calcium/calmodulin-dependent 3', 5'-cyclic nucleotide phosphodiesterase, putative
TGME49_110520	1	2	calcium/calmodulin-dependent 3', 5'-cyclic nucleotide phosphodiesterase, putative
TGGT1_037900	1	2	conserved hypothetical protein
TGME49_086490	1	2	hypothetical protein
TGME49_044200	74	11491	2-oxoglutarate dehydrogenase, putative
TGGT1_045740	39	6143	2-oxoglutarate dehydrogenase, putative
TGME49_046530	8	41	phosphatidylglycerophosphate synthase, putative
TGGT1_026590	5	36	phospholipase D active site motif domain-containing protein, putative
TGME49_066960	121	417632	tubulin beta chain
TGGT1_075980	8	636	tubulin beta chain, putative
TGME49_105560	13	235	hypothetical protein
TGGT1_041920	13	231	hypothetical protein
TGGT1_073590	2	14	conserved hypothetical protein
TGME49_093240	2	14	hypothetical protein
TGME49_053940	8	44	CAM kinase (incomplete catalytic triad)
TGGT1_003220	3	22	calcium/calmodulin-dependent protein kinase, putative
TGME49_078960	5	85	hypothetical protein
TGGT1_103390	1	56	conserved hypothetical protein
TGME49_070200	4	118	hypothetical protein
TGGT1_108730	3	114	conserved hypothetical protein
TGGT1_114870	9	145	conserved hypothetical protein
TGME49_032650	9	145	hypothetical protein, conserved
TGGT1_069620	10	148	signal recognition particle domain-containing protein, putative
TGME49_036140	10	110	signal recognition particle domain-containing protein
TGME49_110460	11	94	RAB6 protein
TGGT1_086680	5	34	RAB6 protein, putative
TGGT1_065870	4	23	mog1 protein, putative
TGME49_022040	4	23	mog1 protein, putative
TGGT1_087120	13	674	CRAL/TRIO domain-containing protein, putative
TGME49_110800	7	277	CRAL/TRIO domain-containing protein
TGME49_068950	42	3084	la domain-containing protein
TGGT1_107140	26	2378	la domain-containing protein, putative
TGME49_005510	46	4022	nucleolar protein NOP5, putative
TGGT1_063390	28	1980	nucleolar protein NOP5, putative
TGGT1_096810	22	766	beta-tubulin cofactor D, putative
TGME49_097420	21	1048	beta-tubulin cofactor D, putative
TGGT1_035780	14	315	mRNA processing protein, putative
TGME49_083740	12	267	mRNA processing protein, putative
TGGT1_104860	14	639	conserved hypothetical protein
TGME49_077790	11	338	hypothetical protein
TGME49_087170	17	4582	hypothetical protein

TGGT1_038270	15	4577	conserved hypothetical protein
TGGT1_076790	1	1	conserved hypothetical protein
TGME49_055370	1	1	hypothetical protein
TGGT1_071960	10	177	conserved hypothetical protein
TGME49_012220	10	177	hypothetical protein, conserved
TGME49_094690	5	93	rhomboid-like protease 5
TGGT1_075010	3	23	rhomboid-like protease 5
TGME49_017740	31	2263	oxoacyl-ACP reductase, putative
TGGT1_028300	15	1141	oxidoreductase, putative
TGME49_022210	22	1448	SPFH domain / Band 7 family domain-containing protein
TGGT1_066020	8	121	conserved hypothetical protein
TGGT1_123030	15	627	conserved hypothetical protein
TGME49_118210	15	616	hypothetical protein
TGGT1_071400	1	1	conserved hypothetical protein
TGME49_034490	1	1	hypothetical protein
TGGT1_010100	4	529	conserved hypothetical protein
TGME49_060020	3	372	hypothetical protein
TGME49_104740	23	512	Rhoptry kinase family protein ROP35
TGGT1_006790	9	187	conserved hypothetical protein
TGGT1_088940	3	8	conserved hypothetical protein
TGME49_112650	2	4	hypothetical protein
TGME49_025020	19	316	hypothetical protein
TGGT1_079960	5	28	conserved hypothetical protein
TGGT1_101760	5	28	DNA-directed RNA polymerase I, putative
TGME49_050060	5	28	RNA polymerase, putative
TGME49_027960	29	1068	proteasome PCI domain-containing protein
TGGT1_083240	17	759	proteasome PCI domain-containing protein, putative
TGGT1_118150	1	2	conserved hypothetical protein
TGME49_029620	1	2	hypothetical protein, conserved
TGGT1_112720	2	8	conserved hypothetical protein
TGGT1_089950	6	32	conserved hypothetical protein
TGME49_113540	5	22	hypothetical protein
TGGT1_092170	2	8	conserved hypothetical protein
TGGT1_082270	20	816	conserved hypothetical protein
TGME49_026890	17	656	hypothetical protein
TGGT1_022290	13	698	conserved hypothetical protein
TGME49_009250	11	536	hypothetical protein
TGME49_060510	9	77	ubiquitin thiolesterase protein, putative
TGME49_023800	8	140	hypothetical protein
TGGT1_078160	6	136	conserved hypothetical protein
TGME49_013900	32	1140	regulator of chromosome condensation, putative
TGGT1_038580	24	1503	regulator of chromosome condensation, putative
TGME49_002830	10	153	C2 domain-containing protein
TGGT1_018110	6	107	C2 domain-containing protein, putative

TGGT1_033890	58	8810	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_089330	56	7569	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_001790	7	135	forkhead-associated domain-containing protein
TGGT1_019440	5	133	forkhead-associated domain-containing protein
TGGT1_086630	9	164	WD-repeat proein, putative
TGME49_110410	6	72	WD-repeat proein, putative
TGGT1_123720	12	434	conserved hypothetical protein
TGME49_011210	11	376	hypothetical protein
TGGT1_071230	1	2	actin, putative
TGME49_034670	1	2	hypothetical protein
TGME49_094710	37	2655	hypothetical protein
TGGT1_075030	29	2509	conserved hypothetical protein
TGGT1_046870	46	3700	conserved hypothetical protein
TGME49_043200	44	3491	hypothetical protein
TGGT1_079840	5	158	jmjC domain-containing protein, putative
TGME49_024930	2	54	jmjC domain-containing protein
TGGT1_025330	1	2	glutaredoxin, GRX, putative
TGME49_039890	25	3303	SCP-like extracellular domain-containing protein
TGGT1_050930	3	30	conserved hypothetical protein
TGME49_069850	7	13	hypothetical protein
TGME49_018590	22	416	protein phosphatase 2C, putative
TGGT1_029430	7	124	protein phosphatase 2C, putative
TGGT1_025380	4	37	conserved hypothetical protein
TGME49_047530	2	24	hypothetical protein
TGME49_097160	8	105	hypothetical protein
TGGT1_097050	2	20	conserved hypothetical protein
TGGT1_075920	1	1	conserved hypothetical protein
TGME49_067020	1	1	hypothetical protein
TGGT1_123520	33	1953	conserved hypothetical protein
TGME49_011290	21	1618	hypothetical protein
TGME49_060180	13	329	hypothetical protein, conserved
TGGT1_010050	6	94	conserved hypothetical protein
TGGT1_125230	3	22	conserved hypothetical protein
TGME49_014800	3	22	hypothetical protein
TGGT1_066000	5	29	conserved hypothetical protein
TGME49_022190	4	28	hypothetical protein
TGGT1_031550	4	38	transcription initiation factor IIB, putative
TGME49_092010	4	38	transcription initiation factor IIB, putative
TGME49_083760	3	55	hypothetical protein
TGGT1_035790	1	17	hypothetical protein
TGGT1_088590	1	8	conserved hypothetical protein
TGME49_030170	5	31	hypothetical protein, conserved
TGGT1_117720	3	27	conserved hypothetical protein
TGGT1_015110	20	964	actin depolymerizing factor, putative
TGME49_020400	12	511	actin depolymerizing factor

TGME49_009910	15	1199	histone H2B variant 1
TGGT1_022960	7	272	histone H2B variant, putative
TGME49_119900	44	1135	hypothetical protein
TGGT1_121490	19	990	conserved hypothetical protein
TGGT1_022500	1	1	conserved hypothetical protein
TGME49_009540	1	1	hypothetical protein
TGME49_111400	61	5820	hypothetical protein, conserved
TGGT1_087750	40	4623	WD-40 repeat-containing protein, putative
TGME49_063290	1	16	rhomboid-like protease TgROM2
TGGT1_043410	1	7	rhomboid 2 protease
TGGT1_123220	9	64	40S ribosomal protein S1, putative
TGME49_011670	9	64	S1 RNA-binding domain containing protein
TGGT1_060500	6	73	WD-repeat protein, putative
TGME49_000280	4	38	WD-repeat protein, putative
TGGT1_107330	5	27	transport protein particle component Bet3 domain-containing protein
TGME49_069140	5	27	transport protein particle component Bet3 domain-containing protein
TGME49_042460	2	5	mitochondrial carrier domain-containing protein
TGGT1_047360	2	4	mitochondrial carrier domain-containing protein
TGME49_047600	18	933	hypothetical protein, conserved
TGGT1_025210	9	237	dynein light intermediate chain, putative
TGGT1_069780	3	13	IBR domain-containing protein, putative
TGME49_035980	3	13	IBR domain-containing protein
TGME49_049600	5	42	hypothetical protein
TGGT1_101270	3	36	conserved hypothetical protein
TGME49_032350	100	38695	lactate dehydrogenase
TGGT1_115280	27	7469	lactate dehydrogenase, putative
TGGT1_029770	5	58	conserved hypothetical protein
TGME49_018800	4	42	hypothetical protein
TGME49_038090	3	14	hypothetical protein
TGGT1_052760	1	5	conserved hypothetical protein
TGGT1_045640	7	139	conserved hypothetical protein
TGME49_044300	4	80	hypothetical protein
TGGT1_121750	1	1	conserved hypothetical protein
TGME49_119540	1	1	hypothetical protein
TGME49_100060	3	40	signal peptidase subunit, putative
TGGT1_121310	13	314	conserved hypothetical protein
TGME49_120080	13	314	hypothetical protein
TGGT1_080660	2	28	conserved hypothetical protein
TGME49_025450	2	28	hypothetical protein
TGGT1_107760	1	1	polyprenyl synthetase domain-containing protein, putative
TGME49_069430	1	1	hexaprenyl pyrophosphate synthetase protein, putative
TGME49_116540	15	1136	hypothetical protein
TGGT1_094350	7	195	conserved hypothetical protein

TGGT1_123960	1	1	conserved hypothetical protein
TGME49_010990	1	1	hypothetical protein
TGME49_048530	3	5	phosphatidylinositol 3- and 4-kinase domain-containing protein
TGGT1_024330	1	1	phosphatidylinositol 3- and 4-kinase domain-containing protein, putative
TGME49_014260	33	1630	alpha-glucan water dikinase 1, putative
TGGT1_124670	31	1867	alpha-glucan water dikinase, putative
TGME49_073110	5	180	SRS30D
TGGT1_112290	3	14	SRS30D
TGGT1_086480	1	3	conserved hypothetical protein
TGME49_110250	1	3	hypothetical protein
TGGT1_110970	8	164	conserved hypothetical protein
TGME49_071960	8	164	hypothetical protein
TGME49_025920	1	1	hypothetical protein
TGME49_114780	60	1956	myosin G
TGGT1_092070	21	470	myosin, putative
TGGT1_075200	15	248	conserved hypothetical protein
TGME49_094860	15	248	hypothetical protein
TGGT1_098760	10	191	NMD3 protein, putative
TGME49_016480	8	134	NMD3 protein
TGGT1_010930	1	2	acylphosphatase, putative
TGME49_058940	1	2	acylphosphatase, putative
TGME49_097960	43	3299	hypothetical protein, conserved
TGGT1_096150	29	2180	conserved hypothetical protein
TGME49_028490	21	774	hypothetical protein, conserved
TGGT1_084050	17	819	conserved hypothetical protein
TGME49_024190	6	23	phospholipid-transporting ATPase, P-type, putative
TGGT1_078660	3	19	phospholipid-transporting ATPase, putative
TGME49_043440	29	1710	histone acetyltransferase GCN5, putative
TGGT1_046420	28	1760	histone acetyltransferase GCN5, putative
TGME49_089880	49	8422	hypothetical protein
TGGT1_033260	32	8695	conserved hypothetical protein
TGME49_014490	48	2811	M16 family peptidase, putative
TGGT1_124890	28	2401	conserved hypothetical protein
TGME49_006550	3	5	hypothetical protein
TGME49_034500	29	1277	phenylalanyl tRNA synthetase isoform, putative
TGGT1_071390	20	1054	phenylalanyl tRNA synthetase isoform, putative
TGGT1_018380	2	3	conserved hypothetical protein
TGME49_002590	2	3	hypothetical protein, conserved
TGGT1_093150	10	257	conserved hypothetical protein
TGME49_115580	9	212	hypothetical protein
TGGT1_014340	9	100	conserved hypothetical protein
TGME49_055880	7	76	hypothetical protein
TGME49_105200	3	5	hypothetical protein

TGGT1_042300	1	1	conserved hypothetical protein
TGME49_017690	31	376	hypothetical protein
TGGT1_028350	5	38	conserved hypothetical protein
TGME49_049700	7	50	mitochondrial carrier domain-containing protein
TGME49_112580	13	560	hypothetical protein
TGGT1_088870	10	545	conserved hypothetical protein
TGGT1_018000	3	8	conserved hypothetical protein
TGME49_002930	3	8	hypothetical protein
TGGT1_106860	5	102	conserved hypothetical protein
TGME49_068680	4	75	hypothetical protein
TGGT1_041930	2	12	conserved hypothetical protein
TGME49_105550	1	5	hypothetical protein, conserved
TGGT1_054880	4	67	CBS domain multi-pass transmembrane protein, putative
TGME49_107580	2	33	CBS domain multi-pass transmembrane protein, putative
TGGT1_085240	16	395	conserved hypothetical protein
TGME49_109370	16	310	adaptin, putative
TGGT1_066660	3	12	conserved hypothetical protein
TGME49_023130	2	6	hypothetical protein
TGME49_034560	18	447	hypothetical protein
TGGT1_071330	15	406	conserved hypothetical protein
TGGT1_066240	24	3049	ubiquitin ligase E3a, putative
TGME49_022430	18	2209	ubiquitin-transferase domain containing protein
TGME49_007590	4	10	T-cell activation protein phosphatase 2C, putative
TGGT1_020480	2	6	protein phosphatase 2C, putative
TGME49_031910	37	2989	ATP synthase gama chain, putative
TGGT1_115850	16	1008	ATP synthase gama chain, putative
TGGT1_063700	6	349	conserved hypothetical protein
TGME49_005300	5	258	hypothetical protein
TGME49_075750	6	33	small nuclear ribonucleoprotein E, putative
TGGT1_000580	2	16	small nuclear ribonucleoprotein E, putative
TGME49_015420	3	21	SNARE protein, putative
TGGT1_126270	3	16	SNARE protein ykt6, putative
TGME49_052480	5	13	hypothetical protein
TGGT1_001370	2	4	conserved hypothetical protein
TGME49_016660	8	90	hypothetical protein, conserved
TGGT1_059870	2	6	septum formation protein maf, putative
TGME49_101430	2	6	septum formation protein maf, putative
TGGT1_022890	3	38	conserved hypothetical protein
TGME49_009850	3	38	hypothetical protein
TGME49_108000	22	801	Gpi16 subunit, GPI transamidase domain-containing protein
TGGT1_042810	11	220	Gpi16 subunit, GPI transamidase domain-containing protein, putative
TGME49_036560	48	11416	hypothetical protein
TGGT1_069280	3	1067	hypothetical protein

TGGT1_068780	9	209	ethylene inducible protein, putative
TGME49_037140	6	109	ethylene inducible protein, putative
TGGT1_051000	3	15	conserved hypothetical protein
TGME49_039830	3	15	hypothetical protein
TGME49_080740	8	229	signal peptidase, putative
TGGT1_072560	5	87	conserved hypothetical protein
TGME49_113310	6	66	DnaJ domain-containing protein
TGGT1_089700	6	64	DnaJ domain-containing protein, putative
TGGT1_110570	4	26	conserved hypothetical protein
TGME49_071780	1	6	hypothetical protein
TGGT1_076850	4	19	conserved hypothetical protein
TGME49_055430	4	17	hypothetical protein
TGME49_023960	25	1731	26S proteasome non-ATPase regulatory subunit 4, putative
TGGT1_078340	20	1520	26S proteasome non-ATPase regulatory subunit, putative
TGME49_022910	8	200	phosphoglycerate mutase family protein
TGGT1_066440	5	80	conserved hypothetical protein
TGME49_026710	5	57	hypothetical protein
TGGT1_081890	3	44	conserved hypothetical protein
TGME49_025940	15	228	hypothetical protein, conserved
TGGT1_081090	6	91	conserved hypothetical protein
TGGT1_109550	3	28	conserved hypothetical protein
TGME49_070790	2	16	hypothetical protein, conserved
TGME49_053340	1	1	hypothetical protein
TGGT1_024580	4	34	prefoldin subunit, putative
TGME49_048370	4	34	prefoldin subunit, putative
TGME49_033520	37	2051	ATP-dependent RNA helicase, putative
TGGT1_113930	25	1777	ATP-dependent RNA helicase, putative
TGGT1_039060	1	1	conserved hypothetical protein
TGME49_013570	1	1	hypothetical protein
TGME49_049320	11	232	NADPH-dependent oxidoreductase, putative
TGGT1_100990	4	35	NADPH-dependent oxidoreductase, putative
TGGT1_022440	34	4155	conserved hypothetical protein
TGME49_009500	30	3475	hypothetical protein
TGME49_068760	13	158	hypothetical protein
TGGT1_106940	4	60	conserved hypothetical protein
TGGT1_079130	1	2	conserved hypothetical protein
TGME49_024510	1	2	hypothetical protein
TGME49_109560	10	6461	hypothetical protein, conserved
TGGT1_085520	4	109	conserved hypothetical protein
TGME49_053850	5	130	hypothetical protein
TGGT1_003010	4	126	conserved hypothetical protein
TGGT1_113170	3	21	conserved hypothetical protein
TGME49_073980	1	5	hypothetical protein
TGGT1_041020	5	15	conserved hypothetical protein
TGME49_106320	5	15	hypothetical protein

TGME49_001830	2	4	hypothetical protein, conserved
TGME49_014930	12	352	hypothetical protein, conserved
TGGT1_125350	5	60	conserved hypothetical protein
TGGT1_038780	2	12	WD repeat-containing protein, putative
TGME49_013710	2	12	WD repeat-containing protein
TGGT1_051830	6	33	protein kinase, putative
TGME49_039130	6	33	Tyrosine kinase-like (TKL) protein
TGME49_026370	1	1	dgat2l1-prov protein
TGGT1_085540	3	37	conserved hypothetical protein
TGME49_109580	3	37	hypothetical protein, conserved
TGGT1_090880	1	1	conserved hypothetical protein
TGME49_114300	1	1	hypothetical protein, conserved
TGGT1_010200	1	1	conserved hypothetical protein
TGME49_059920	1	1	hypothetical protein
TGGT1_065560	13	315	hypothetical protein
TGME49_021690	9	164	hypothetical protein, conserved
TGGT1_044940	2	15	quinone oxidoreductase, putative
TGME49_044870	1	8	quinone oxidoreductase, putative
TGGT1_045100	34	4526	trichohyalin, putative
TGME49_044710	32	4358	trichohyalin, putative
TGME49_008790	25	9356	mitochondrial carrier domain-containing protein
TGGT1_021820	7	92	mitochondrial carrier domain-containing protein
TGME49_083780	77	6605	glucose-6-phosphate isomerase
TGGT1_035800	22	1529	glucose-6-phosphate isomerase, putative
TGGT1_009310	1	2	zinc finger protein, putative
TGME49_046580	22	1728	hypothetical protein
TGGT1_026530	7	196	conserved hypothetical protein
TGGT1_034320	1	2	conserved hypothetical protein
TGME49_088950	1	2	hypothetical protein
TGME49_090900	4	19	NUDIX domain-containing protein
TGGT1_032240	2	13	conserved hypothetical protein
TGGT1_093460	39	5810	conserved hypothetical protein
TGME49_115760	39	5405	hypothetical protein, conserved
TGME49_057750	11	352	homocysteine S-methyltransferase-1
TGGT1_012940	7	279	homocysteine S-methyltransferase-1, putative
TGME49_057670	4	7	hypothetical protein
TGGT1_013020	2	3	conserved hypothetical protein
TGME49_073950	10	105	replication factor C, putative
TGGT1_113030	7	99	replication factor C, putative
TGGT1_042610	1	1	conserved hypothetical protein
TGME49_104880	1	1	hypothetical protein
TGME49_077720	26	712	nucleoside-triphosphatase, putative
TGGT1_104930	17	706	nucleoside-triphosphatase, putative
TGGT1_034180	8	182	cation-transporting ATPase, putative
TGME49_089070	7	158	P-Type cation-transporting ATPase, putative

TGGT1_086790	1	4	hypothetical protein
TGME49_110560	1	4	hypothetical protein
TGGT1_115260	21	936	zinc finger (CW-type) protein
TGME49_032370	21	789	CW-type zinc finger domain-containing protein
TGME49_035560	1	1	hypothetical protein
TGME49_019430	2	4	threonyl-tRNA synthetase, putative
TGME49_033130	7	213	adenosine transporter, putative
TGGT1_114390	4	170	adenosine transporter, putative
TGME49_089580	42	2912	strictosidine synthase domain-containing protein
TGGT1_033630	16	800	strictosidine synthase, putative
TGGT1_006980	12	345	conserved hypothetical protein
TGME49_062990	9	259	hypothetical protein
TGGT1_095000	6	60	SNARE domain-containing protein, putative
TGME49_100290	4	25	SNARE domain-containing protein
TGME49_024520	4	50	hypothetical protein
TGGT1_079140	4	45	conserved hypothetical protein
TGGT1_006040	12	202	phosphofructokinase, putative
TGME49_081410	11	180	pyrophosphate dependent phosphofructokinase, putative
TGGT1_120520	13	201	conserved hypothetical protein
TGME49_120750	13	201	hypothetical protein, conserved
TGME49_108840	42	62738	SRS51 (= SRS3)
TGGT1_055870	25	1607	SRS51 (= SRS3)
TGGT1_121570	13	357	conserved hypothetical protein
TGME49_119720	13	357	hypothetical protein
TGME49_092080	98	13933	leucyl-tRNA synthetase, putative
TGGT1_031480	58	9320	leucyl-tRNA synthetase, putative
TGGT1_089090	2	7	conserved hypothetical protein
TGME49_112890	1	4	hypothetical protein
TGGT1_110620	12	346	lipoate-protein ligase A, putative
TGME49_071820	12	315	lipoate-protein ligase A, putative
TGME49_041820	2	44	hypothetical protein
TGGT1_048900	2	9	conserved hypothetical protein
TGGT1_006400	1	4	SNARE protein ykt6, putative
TGME49_099180	1	4	hypothetical protein, conserved
TGME49_051590	11	77	protease, putative
TGGT1_102980	4	24	protease, putative
TGGT1_088430	10	187	conserved hypothetical protein
TGME49_112190	10	155	hypothetical protein
TGME49_071970	9	2469	hypothetical protein
TGGT1_110980	2	1563	conserved hypothetical protein
TGME49_036930	38	1296	hypothetical protein
TGGT1_069000	27	1380	conserved hypothetical protein
TGGT1_055510	2	5	PUA domain-containing, nucleolar protein, putative
TGME49_106900	2	5	NOL1/NOP2/sun family protein
TGGT1_079790	6	74	kinesin motor domain-containing protein,

			putative
TGME49_024880	5	57	kinesin motor domain-containing protein
TGGT1_027430	10	297	conserved hypothetical protein
TGME49_045640	10	273	hypothetical protein
TGME49_056820	11	61	hypothetical protein
TGGT1_013930	6	32	conserved hypothetical protein
TGME49_072000	7	78	hypothetical protein
TGGT1_111020	2	2	conserved hypothetical protein
TGGT1_125950	1	4	F-box domain-containing protein, putative
TGME49_015210	1	4	F-box domain-containing protein
TGME49_022660	5	35	hypothetical protein
TGGT1_066260	1	3	hypothetical protein
TGME49_016040	1	1	ribosomal protein S15, putative
TGME49_059710	11	223	hypothetical protein, conserved
TGGT1_010320	9	167	conserved hypothetical protein
TGGT1_103960	10	212	conserved hypothetical protein
TGME49_078600	7	196	hypothetical protein, conserved
TGME49_066130	1	1	glutathione/thioredoxin peroxidase, putative
TGME49_053400	4	10	hypothetical protein
TGGT1_002480	1	1	conserved hypothetical protein
TGME49_018400	2	4	NEK kinase
TGGT1_029220	1	2	protein kinase, putative
TGGT1_005260	10	112	conserved hypothetical protein
TGME49_082220	8	88	hypothetical protein
TGME49_068850	129	50399	enolase 2
TGGT1_107040	32	12902	enolase, putative
TGME49_121570	3	28	(3R)-hydroxymyristoyl ACP dehydrase, putative
TGGT1_064610	3	14	beta-hydroxyacyl-ACP dehydratase, putative
TGME49_039420	7	187	protein kinase
TGGT1_051450	6	106	calcium/calmodulin-dependent protein kinase II, putative
TGGT1_084270	7	77	phosphatidylinositol 4-kinase, putative
TGME49_028690	5	53	phosphatidylinositol 3- and 4-kinase domain-containing protein
TGME49_004130	52	4425	membrane-attack complex / perforin domain-containing protein
TGGT1_016570	22	1631	membrane-attack complex /perforin domain-containing protein
TGGT1_115980	15	299	conserved hypothetical protein
TGME49_031790	14	266	hypothetical protein
TGME49_105270	8	85	hypothetical protein
TGGT1_042230	5	48	conserved hypothetical protein
TGME49_061770	1	1	hypothetical protein
TGME49_069200	28	187	crooked neck-like protein 1, putative
TGGT1_107400	9	94	crooked neck protein, putative
TGGT1_025960	22	3046	conserved hypothetical protein
TGME49_047040	19	2551	hypothetical protein

TGME49_013020	3	9	hypothetical protein
TGME49_043690	9	94	hypothetical protein
TGGT1_046180	4	55	conserved hypothetical protein
TGGT1_034530	12	700	hypothetical protein
TGME49_088740	9	298	asparaginase, putative
TGME49_031420	1	1	BT1 transmembrane domain-containing protein
TGME49_035170	29	894	replication factor C subunit, putative
TGGT1_070940	24	889	replication factor C subunit, putative
TGGT1_004390	5	67	conserved hypothetical protein
TGME49_054830	2	14	hypothetical protein
TGME49_013770	23	643	RNA helicase, putative
TGGT1_038730	19	620	hypothetical protein
TGME49_024900	34	10896	adenylate kinase, putative
TGGT1_079810	13	470	adenylate kinase, putative
TGGT1_013940	1	1	DnaJ domain-containing protein, putative
TGME49_056810	1	1	DnaJ domain-containing protein
TGME49_025500	7	45	hypothetical protein
TGGT1_080710	5	41	conserved hypothetical protein
TGME49_069770	16	214	WD-repeat membrane protein, putative
TGGT1_108200	11	173	WD-repeat membrane protein, putative
TGGT1_005280	8	170	ADP-ribose pyrophosphatase, putative
TGME49_082190	8	170	hydrolase, NUDIX family domain containing protein
TGGT1_005480	6	67	hypothetical protein
TGME49_081990	5	50	pyrazinamidase/nicotinamidase, putative
TGME49_049310	4	38	hypothetical protein
TGGT1_100980	4	37	conserved hypothetical protein
TGME49_115700	4	9	hypothetical protein, conserved
TGGT1_093380	2	6	conserved hypothetical protein
TGGT1_122950	1	1	conserved hypothetical protein
TGME49_118300	1	1	hypothetical protein
TGGT1_009030	5	51	transcription initiation protein, putative
TGME49_061220	5	51	transcription elongation factor SPT4, putative
TGME49_003220	11	187	ATP-dependent RNA helicase, putative
TGGT1_017650	10	156	hypothetical protein
TGGT1_081620	1	1	conserved hypothetical protein
TGME49_026460	1	1	hypothetical protein
TGGT1_006470	1	2	conserved hypothetical protein
TGME49_099240	1	2	hypothetical protein
TGGT1_117040	7	48	ATPase, AAA family protein
TGME49_030710	5	30	ATPase, AAA family domain-containing protein
TGME49_111720	129	782557	heat shock protein 70, putative
TGGT1_087970	48	20991	heat shock protein, putative
TGME49_105540	4	23	hypothetical protein
TGGT1_041940	3	19	conserved hypothetical protein
TGME49_110950	23	1106	hypothetical protein

TGGT1_087270	18	992	conserved hypothetical protein
TGME49_086420	171	1229074	elongation factor 1-alpha, putative
TGGT1_037840	36	20271	elongation factor 1-alpha, putative
TGGT1_098310	4	279	transmembrane domian-containing protein, putative
TGME49_016820	2	134	transmembrane domian-containing protein
TGGT1_021530	13	539	protein phosphatase 2C, putative
TGME49_008500	8	328	protein phosphatase 2C, putative
TGGT1_081850	30	3751	conserved hypothetical protein
TGME49_026680	29	3686	hypothetical protein
TGME49_027060	10	140	hypothetical protein
TGGT1_082540	7	147	conserved hypothetical protein
TGGT1_041630	2	6	conserved hypothetical protein
TGME49_105840	2	6	hypothetical protein
TGGT1_108670	3	8	mitosis protein, putative
TGME49_070140	3	8	mitosis protein, putative
TGME49_016050	22	710	ankyrin repeat containing protein / TPR domain-containing protein
TGGT1_099200	17	661	ankyrin repeat containing protein / TPR domain-containing protein, putative
TGME49_003370	23	897	hypothetical protein
TGGT1_017460	22	978	conserved hypothetical protein
TGGT1_082990	1	1	amino acid transporter, putative
TGME49_027580	1	1	hypothetical protein
TGGT1_065200	4	36	conserved hypothetical protein
TGME49_021350	3	26	hypothetical protein
TGME49_026060	5	31	transmembrane amino acid transporter, putative
TGGT1_081210	2	14	transmembrane amino acid transporter, putative
TGME49_016220	22	1036	hypothetical protein, conserved
TGGT1_099040	21	1018	conserved hypothetical protein
TGGT1_112010	10	122	mRNA (guanine-7-)methyltransferase, putative
TGME49_072720	10	47	mRNA capping enzyme, large subunit family
TGME49_063050	26	2093	60S ribosomal protein L13, putative
TGGT1_006930	8	152	60S ribosomal protein L13, putative
TGGT1_017480	2	2	conserved hypothetical protein
TGME49_003350	2	2	hypothetical protein
TGME49_033790	2	6	AGC kinase
TGGT1_113550	1	5	serine/threonine-protein kinase, putative
TGGT1_089660	51	12100	conserved hypothetical protein
TGME49_113270	41	8315	hypothetical protein
TGME49_033140	15	581	deoxyuridine 5'-triphosphate nucleotidohydrolase, putative
TGGT1_114370	7	308	deoxyuridine 5'-triphosphate nucleotidohydrolase, putative
TGME49_068890	41	2738	citrate synthase, putative
TGGT1_107080	14	808	citrate synthase, putative
TGGT1_101580	1	1	conserved hypothetical protein

TGME49_049890	1	1	hypothetical protein
TGGT1_045970	2	15	conserved hypothetical protein
TGME49_043990	2	15	NLI interacting factor-like phosphatase domain-containing protein
TGME49_054650	10	217	RRM / zinc finger (CCCH type) domain-containing protein
TGGT1_004200	10	212	lysine-specific histone demethylase, putative
TGGT1_039320	5	44	conserved hypothetical protein
TGME49_013330	2	14	hypothetical protein
TGME49_047460	21	951	proliferating cell nuclear antigen 1
TGGT1_025450	16	878	proliferating cell nuclear antigen, putative
TGME49_090850	12	206	WD-40 repeat protein, putative
TGGT1_032300	11	190	WD-repeat protein, putative
TGGT1_054340	6	46	WD-repeat protein, putative
TGME49_095680	6	46	periodic tryptophan protein PWP2, putative
TGGT1_017720	26	3712	conserved hypothetical protein
TGME49_003150	19	2702	hypothetical protein
TGME49_093530	7	50	hypothetical protein
TGGT1_073910	3	37	conserved hypothetical protein
TGGT1_031350	21	844	conserved hypothetical protein
TGME49_092210	15	518	hypothetical protein
TGGT1_020890	4	21	transcription factor IIIB subunit, putative
TGME49_007900	3	14	transcription factor IIIB subunit, putative
TGGT1_028850	9	140	N6-adenosine-methyltransferase 70 kDa subunit, putative
TGME49_017350	8	118	N6-adenosine-methyltransferase 70 kDa subunit, putative
TGME49_013750	3	8	hypothetical protein
TGGT1_038740	1	1	hypothetical protein
TGME49_011330	3	6	methionine aminopeptidase, putative
TGGT1_123470	1	1	methionine aminopeptidase, putative
TGME49_065460	11	432	ppg3, putative
TGGT1_057630	9	428	conserved hypothetical protein
TGME49_032720	4	37	hypothetical protein
TGGT1_114800	3	33	conserved hypothetical protein
TGME49_019150	11	100	krr1 family, zinc finger-containing protein
TGGT1_030110	7	55	conserved hypothetical protein
TGGT1_089740	6	231	conserved hypothetical protein
TGME49_113340	5	194	hypothetical protein
TGGT1_037210	2	3	conserved hypothetical protein
TGME49_085810	2	3	hypothetical protein
TGGT1_032510	19	692	conserved hypothetical protein
TGME49_090630	18	406	hypothetical protein, conserved
TGME49_046180	3	34	hypothetical protein
TGGT1_026880	2	28	conserved hypothetical protein
TGME49_052510	15	194	hypothetical protein, conserved
TGGT1_001400	12	206	conserved hypothetical protein
TGME49_022400	13	672	hypothetical protein

TGGT1_066210	10	655	hypothetical protein
TGME49_049840	150	10472	dynein heavy chain domain containing protein
TGGT1_072680	15	464	conserved hypothetical protein
TGME49_080630	15	464	hypothetical protein
TGGT1_007550	2	12	conserved hypothetical protein
TGME49_062550	2	12	hypothetical protein
TGGT1_045680	1	1	conserved hypothetical protein
TGME49_044260	1	1	hypothetical protein
TGME49_070240	52	12252	cyst matrix protein
TGGT1_108770	21	1320	cyst matrix protein, putative
TGME49_025930	53	6169	triosephosphate isomerase, putative
TGGT1_081080	24	2796	triosephosphate isomerase, putative
TGGT1_045090	6	48	hypothetical protein
TGME49_044720	4	24	hypothetical protein
TGGT1_021450	1	1	protease inhibitor, putative
TGME49_008430	1	1	serine proteinase inhibitor TgPI-2, putative
TGGT1_012200	2	6	hypothetical protein
TGME49_058000	2	6	hypothetical protein
TGME49_057040	8	222	hypothetical protein, conserved
TGGT1_013680	4	152	conserved hypothetical protein
TGME49_070860	14	339	adenylate cyclase, putative
TGGT1_109610	12	335	adenylate cyclase, putative
TGGT1_013110	1	1	conserved hypothetical protein
TGME49_057570	1	1	hypothetical protein
TGGT1_091230	1	1	conserved hypothetical protein
TGME49_114440	1	1	hypothetical protein
TGGT1_087670	3	26	conserved hypothetical protein
TGME49_111320	3	18	hypothetical protein
TGME49_114850	3	9	hypothetical protein, conserved
TGGT1_092140	1	1	conserved hypothetical protein
TGGT1_055190	3	16	conserved hypothetical protein
TGME49_107020	3	16	hypothetical protein, conserved
TGGT1_111750	2	5	conserved hypothetical protein
TGME49_072460	2	5	hypothetical protein
TGME49_057360	6	45	hypothetical protein
TGME49_076130	15	121	cathepsin C2 (TgCPC2)
TGGT1_000410	6	61	cathepsin C, putative
TGME49_018530	6	63	hypothetical protein
TGGT1_029360	5	141	conserved hypothetical protein
TGME49_054580	6	188	UDP-galactose transporter protein, putative
TGGT1_004150	3	79	conserved hypothetical protein
TGME49_031830	11	122	tRNA-splicing endonuclease positive effector, putative
TGME49_114260	5	27	hypothetical protein
TGGT1_090820	2	8	conserved hypothetical protein
TGME49_020350	9	55	lysyl-tRNA synthetase, putative

TGGT1_015050	7	47	lysyl-tRNA synthetase, putative
TGME49_019600	24	468	hypothetical protein
TGGT1_030490	8	158	conserved hypothetical protein
TGGT1_088000	4	27	conserved hypothetical protein
TGME49_111750	4	27	hypothetical protein
TGME49_061450	23	1263	hypothetical protein
TGGT1_008790	20	1359	conserved hypothetical protein
TGME49_014820	12	224	DNA repair enzyme, putative
TGGT1_125250	11	210	DNA repair enzyme, putative
TGGT1_122840	8	86	conserved hypothetical protein
TGME49_118390	8	86	hypothetical protein
TGME49_026420	9	40	oligoendopeptidase F, putative
TGGT1_081590	5	42	oligoendopeptidase F, putative
TGME49_113410	66	6605	26S proteasome non-ATPase regulatory subunit 2, putative
TGGT1_089810	44	5931	26S proteasome non-ATPase regulatory subunit, putative
TGGT1_126430	8	248	conserved hypothetical protein
TGME49_015490	6	164	transporter, major facilitator family domain containing protein
TGME49_113860	34	3279	hypothetical protein, conserved
TGGT1_090380	32	3086	hypothetical protein
TGME49_093570	30	2435	translocation protein Sec62, putative
TGGT1_073950	12	484	translocation protein Sec62, putative
TGME49_060230	3	6	hypothetical protein
TGGT1_009990	1	2	conserved hypothetical protein
TGME49_023970	3	16	elongation factor G, putative
TGGT1_065440	2	3	conserved hypothetical protein
TGME49_021570	2	3	hypothetical protein
TGME49_094770	11	144	hypothetical protein
TGGT1_075090	9	119	conserved hypothetical protein
TGME49_089530	14	615	60s ribosomal protein L19, putative
TGGT1_033700	7	156	60S ribosomal protein L19, putative
TGME49_001840	24	1468	eukaryotic aspartyl protease, putative
TGGT1_019390	7	112	eukaryotic aspartyl protease, putative
TGME49_110810	3	5	apyrase, putative
TGGT1_086600	5	59	U3 small nucleolar ribonucleoprotein, putative
TGME49_110380	5	59	U3 small nucleolar ribonucleoprotein protein, putative
TGGT1_008890	10	119	hypothetical protein
TGME49_061260	10	101	jmjC domain-containing protein, conserved
TGME49_059260	38	1317	cell division protein, putative
TGGT1_010550	17	603	cell division protein, putative
TGGT1_044400	1	7	ATP-dependent RNA helicase, putative
TGME49_064160	1	7	ATP-dependent RNA helicase, putative
TGGT1_080230	1	1	hypothetical protein
TGGT1_021180	1	2	5-formyltetrahydrofolate cyclo-ligase domain-containing protein, putative

TGME49_008090	1	2	5-formyltetrahydrofolate cyclo-ligase domain-containing protein
TGME49_025990	12	206	malonyl CoA-acyl carrier protein transacylase, putative
TGGT1_081130	10	214	malonyl CoA-acyl carrier protein transacylase, putative
TGGT1_029440	2	6	conserved hypothetical protein
TGME49_018600	1	2	hypothetical protein
TGGT1_014285	2	2	CSL zinc finger domain-containing protein
TGME49_055940	2	2	CSL zinc finger domain protein
TGGT1_062260	1	1	conserved hypothetical protein
TGME49_006500	1	1	hypothetical protein
TGME49_015590	70	9149	succinate dehydrogenase (ubiquinone) flavoprotein subunit, mitochondrial, putative
TGGT1_126520	27	3580	succinate dehydrogenase, putative
TGME49_013010	14	192	hypothetical protein
TGGT1_039730	7	74	conserved hypothetical protein
TGGT1_093160	1	2	ganglioside induced differentiation associated protein, putative
TGME49_115590	1	2	Appr-1-p processing enzyme family domain-containing protein
TGME49_049170	23	680	ras family domain-containing protein
TGGT1_100730	11	282	small GTP-binding protein domain-containing protein
TGME49_055910	3	6	hypothetical protein, conserved
TGGT1_014310	1	1	conserved hypothetical protein
TGME49_002540	28	853	3',5'-cyclic nucleotide phosphodiesterase, putative
TGGT1_018550	15	919	3',5'-cyclic nucleotide phosphodiesterase, putative
TGGT1_122830	1	2	conserved hypothetical protein
TGME49_118400	1	2	hypothetical protein
TGME49_105070	8	103	hypothetical protein
TGGT1_042430	3	39	conserved hypothetical protein
TGME49_050730	4	19	ribosomal protein L7, putative
TGGT1_102170	3	18	ribosomal protein L7, putative
TGME49_003980	38	1998	hypothetical protein
TGGT1_016720	33	2048	conserved hypothetical protein
TGME49_111210	44	1672	hypothetical protein
TGGT1_087520	11	161	conserved hypothetical protein
TGME49_019810	12	122	hypothetical protein
TGGT1_030810	11	156	conserved hypothetical protein
TGGT1_036720	7	50	conserved hypothetical protein
TGME49_085220	5	28	CAP-Gly domain-containing protein
TGME49_054390	23	1529	hypothetical protein
TGGT1_003830	20	1383	conserved hypothetical protein
TGME49_009200	8	111	hypothetical protein, conserved
TGGT1_022240	5	122	conserved hypothetical protein
TGME49_023590	22	917	proteasome component PRE3 precursor, putative

TGGT1_077950	11	501	proteasome component PRE3 precursor / proteasome subunit beta type 6 precursor, putative
TGME49_025060	15	320	nnucleoredoxin, putative
TGGT1_080000	8	172	AhpC/TSA family domain-containing protein
TGME49_028470	14	1674	ribosomal protein L15, putative
TGGT1_084040	4	69	ribosomal protein L15, putative
TGME49_106290	8	54	DNA-directed RNA polymerase III largest subunit, putative
TGGT1_041040	6	60	DNA-directed RNA polymerase III largest subunit, putative
TGGT1_070230	4	32	PHD-finger domain-containing protein, putative
TGME49_035550	4	32	PHD-finger domain-containing protein
TGGT1_051040	15	466	hypothetical protein
TGME49_039790	15	436	pescadillo family protein
TGGT1_098380	8	275	conserved hypothetical protein
TGME49_016750	8	199	hypothetical protein
TGME49_042750	1	2	hypothetical protein
TGGT1_074100	8	91	conserved hypothetical protein
TGME49_093710	8	91	hypothetical protein
TGGT1_054310	54	6774	ubiquitin-transferase domain-containing protein, putative
TGME49_095710	54	6774	ubiquitin-transferase domain-containing protein
TGGT1_033670	4	28	20 kD nuclear cap binding protein, putative
TGME49_089560	4	28	nuclear cap-binding protein, putative
TGGT1_051890	4	21	conserved hypothetical protein
TGME49_039070	4	18	hypothetical protein
TGME49_100250	4	112	mtN3/saliva family domain-containing protein
TGGT1_095040	3	111	conserved hypothetical protein
TGME49_093230	8	63	hypothetical protein
TGGT1_073580	6	39	conserved hypothetical protein
TGME49_010840	47	1813	arginyl-tRNA synthetase, putative
TGGT1_124020	22	1091	arginyl-tRNA synthetase, putative
TGGT1_042000	13	256	conserved hypothetical protein
TGME49_105470	10	176	hypothetical protein
TGME49_010390	5	29	WD domain, G-beta repeat-containing protein
TGGT1_053240	4	28	WD domain, G-beta repeat-containing protein, putative
TGGT1_078550	9	193	conserved hypothetical protein
TGME49_024070	5	92	hypothetical protein
TGME49_019820	10	2069	polyubiquitin, putative
TGGT1_030830	4	26	polyubiquitin, putative
TGGT1_086930	1	2	conserved hypothetical protein
TGME49_110710	1	2	hypothetical protein
TGGT1_071520	1	1	SRS42
TGME49_034370	1	1	SRS42
TGGT1_071150	13	512	protein kinase, putative
TGME49_034970	12	459	Tyrosine kinase-like (TKL) protein
TGME49_062170	5	12	hypothetical protein

TGGT1_007940	1	1	conserved hypothetical protein
TGGT1_125530	1	7	conserved hypothetical protein
TGME49_015000	1	7	hypothetical protein
TGME49_020390	2	15	hypothetical protein
TGGT1_015090	2	10	conserved hypothetical protein
TGGT1_029800	12	180	conserved hypothetical protein
TGME49_018830	10	118	hypothetical protein
TGME49_075810	19	3085	ribosomal protein S10, putative
TGGT1_000530	7	444	ribosomal protein S10, putative
TGME49_050780	1	4	zinc-finger in ubiquitin-hydrolases domain-containing protein
TGGT1_023340	6	62	hypothetical protein
TGME49_007100	6	62	hypothetical protein
TGGT1_064070	3	17	hypothetical protein
TGME49_004880	3	17	hypothetical protein
TGME49_019710	53	2629	hypothetical protein
TGGT1_030720	45	3675	conserved hypothetical protein
TGGT1_072660	17	465	conserved hypothetical protein
TGME49_080650	16	458	hypothetical protein
TGME49_042670	2	5	hypothetical protein
TGGT1_047250	1	4	conserved hypothetical protein
TGME49_093770	31	1386	chitinase class I, putative
TGGT1_074150	18	973	NBP2B protein, putative
TGME49_112150	36	4681	hypothetical protein
TGGT1_088400	15	506	conserved hypothetical protein
TGGT1_125890	14	220	conserved hypothetical protein
TGME49_015150	13	205	hypothetical protein
TGGT1_064370	20	1040	conserved hypothetical protein
TGME49_121340	18	857	hypothetical protein, conserved
TGGT1_072520	5	31	dihydrouridine synthase domain-containing protein, putative
TGME49_080780	5	31	dihydrouridine synthase domain-containing protein
TGME49_042320	2	4	B-box zinc finger protein, putative
TGME49_036040	61	40962	fructose-1,6-bisphosphate aldolase
TGGT1_069710	16	2634	hypothetical protein
TGME49_066060	35	7005	40S ribosomal protein S0-A, putative
TGGT1_056700	15	1668	40S ribosomal protein SA, putative
TGGT1_076580	14	301	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_055180	12	144	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_070510	59	10667	asparaginyl-tRNA synthetase, putative
TGGT1_109050	33	5628	asparaginyl-tRNA synthetase, putative
TGGT1_020050	13	354	conserved hypothetical protein
TGME49_001270	10	206	hypothetical protein
TGGT1_080680	1	2	peptide methionine sulfoxide reductase, putative
TGME49_025470	1	2	peptide methionine sulfoxide reductase, putative
TGME49_017510	22	1864	hypothetical protein

TGGT1_028580	12	820	conserved hypothetical protein
TGME49_002690	9	190	DNA-directed RNA polymerase II subunit, putative
TGGT1_018280	8	163	DNA-directed RNA polymerase II subunit, putative
TGME49_032010	2	5	protein phosphatase 2C, putative
TGGT1_115750	2	2	protein phosphatase 2C, putative
TGME49_064600	2	3	hypothetical protein
TGGT1_105700	1	2	conserved hypothetical protein
TGME49_015690	2	3	hypothetical protein
TGGT1_126610	1	2	conserved hypothetical protein
TGGT1_126210	1	1	conserved hypothetical protein
TGME49_015370	1	1	hypothetical protein
TGGT1_074000	1	2	Dihydroorotase, putative
TGME49_093610	1	2	dihydroorotase protein, putative
TGME49_007180	13	331	indole-3-glycerol phosphate synthase domain containing protein
TGME49_105040	19	679	hypothetical protein
TGGT1_042460	19	667	importin beta-3, putative
TGGT1_020150	1	1	conserved hypothetical protein
TGME49_001170	1	1	hypothetical protein
TGME49_048400	7	110	lactoylglutathione lyase, putative
TGGT1_024550	7	99	lactoylglutathione lyase, putative
TGME49_120470	1	1	hypothetical protein, conserved
TGGT1_090960	2	6	conserved hypothetical protein
TGME49_114370	2	6	hypothetical protein
TGME49_004080	6	26	histidine acid phosphatase domain containing protein
TGGT1_016620	1	2	lysophosphatidic acid phosphatase, putative
TGGT1_017860	4	44	RNA recognition motif-containing protein, putative
TGME49_003080	3	28	RNA recognition motif-containing protein
TGGT1_012170	14	393	DNA polymerase delta catalytic subunit, putative
TGME49_058030	13	279	DNA polymerase delta catalytic subunit, putative
TGME49_073930	11	122	SWIB/MDM2 domain-containing protein
TGGT1_113010	4	37	SWIB/MDM2 domain-containing protein, putative
TGME49_113640	11	371	hypothetical protein
TGGT1_090150	8	233	NBP2B protein, putative
TGGT1_118000	9	114	conserved hypothetical protein
TGME49_029790	8	95	hypothetical protein
TGGT1_088020	25	3173	platelet binding protein GspB, putative
TGME49_111770	25	3027	platelet binding protein GspB, putative
TGME49_032310	14	131	endonuclease/exonuclease/phosphatase domain-containing protein
TGGT1_115340	7	95	endonuclease/exonuclease/phosphatase domain-containing protein, putative
TGME49_019660	2	10	membrane attachment protein, putative
TGGT1_083930	13	898	FK506-binding protein, putative

TGME49_028360	12	821	FK506-binding protein 1, putative
TGME49_090980	9	121	aminotransferase domain-containing protein, conserved
TGGT1_032140	7	96	aminotransferase, putative
TGME49_104720	23	654	SWIM zinc finger domain-containing protein
TGGT1_006770	21	646	SWIM zinc finger protein domain-containing protein, putative
TGGT1_012980	4	36	conserved hypothetical protein
TGME49_057710	3	24	actin-like family protein ARP6, putative
TGME49_120030	26	881	hypothetical protein
TGGT1_121360	9	254	conserved hypothetical protein
TGGT1_063750	1	2	conserved hypothetical protein
TGME49_005260	1	2	hypothetical protein
TGGT1_049910	6	23	conserved hypothetical protein
TGME49_040850	6	23	PAK1 interacting protein, putative
TGGT1_053220	2	14	high mobility group protein, putative
TGME49_010410	2	14	high mobility group protein
TGGT1_014610	2	6	conserved hypothetical protein
TGME49_020110	1	2	hypothetical protein
TGME49_030230	2	2	hypothetical protein
TGGT1_117650	1	1	hypothetical protein
TGME49_044310	18	448	developmentally regulated GTP-binding protein 1, putative
TGGT1_045630	11	347	conserved hypothetical protein
TGGT1_057680	6	257	conserved hypothetical protein
TGME49_065420	6	257	hypothetical protein
TGGT1_117700	3	10	hypothetical protein
TGME49_030190	3	10	hypothetical protein
TGME49_101380	3	7	GTP-binding protein TypA, putative
TGGT1_059800	2	6	GTP-binding protein, putative
TGGT1_037000	9	360	conserved hypothetical protein
TGME49_085500	8	279	hypothetical protein
TGME49_053590	1	1	hypothetical protein
TGME49_037220	34	1532	DNA replication licensing factor, putative
TGGT1_068700	33	1528	DNA replication licensing factor, putative
TGGT1_096910	10	229	conserved hypothetical protein
TGME49_097330	8	95	hypothetical protein
TGGT1_069230	1	1	conserved hypothetical protein
TGME49_036610	1	1	hypothetical protein
TGME49_094220	14	175	hypothetical protein
TGGT1_074520	8	102	conserved hypothetical protein
TGME49_016710	2	4	hypothetical protein, conserved
TGGT1_087440	2	5	conserved hypothetical protein
TGME49_111110	2	5	hypothetical protein
TGME49_066420	24	275	hypothetical protein
TGGT1_076450	6	42	conserved hypothetical protein
TGGT1_056390	6	156	CorA-like Mg2+ transporter domain-containing protein

TGME49_109240	4	84	corA-like Mg2+ transporter domain-containing protein
TGME49_067060	14	1251	60S ribosomal protein L14, putative
TGGT1_075870	5	151	60S ribosomal protein L14, putative
TGGT1_107690	2	10	conserved hypothetical protein
TGME49_069390	2	10	hypothetical protein
TGME49_081930	5	33	surface antigen, putative
TGGT1_005540	2	7	surface protein, putative
TGGT1_057830	5	64	conserved hypothetical protein
TGME49_065280	4	39	hypothetical protein
TGME49_000290	6	50	rhomboid family domain-containing protein
TGGT1_060510	5	46	Rhomboid-like protease 1
TGME49_093340	27	979	ran-specific GTPase-activating protein, putative
TGGT1_073700	10	465	ran-specific GTPase-activating protein, putative
TGGT1_026670	2	12	conserved hypothetical protein
TGME49_046450	2	12	hypothetical protein
TGGT1_000610	1	1	conserved hypothetical protein
TGME49_075720	1	1	hypothetical protein
TGGT1_010390	8	780	conserved hypothetical protein
TGME49_059630	6	655	hypothetical protein, conserved
TGME49_009600	41	3662	hypothetical protein
TGGT1_022650	23	2793	conserved hypothetical protein
TGME49_023680	12	678	ubiquitin, putative
TGGT1_078030	12	647	ubiquitin, putative
TGME49_044160	7	48	hypothetical protein
TGGT1_045800	2	10	conserved hypothetical protein
TGME49_054480	24	492	WD-40 repeat-containing protein
TGGT1_004040	18	527	WD-40 repeat-containing protein, putative
TGGT1_092720	9	134	conserved hypothetical protein
TGME49_115290	9	134	hypothetical protein
TGME49_091330	7	214	hypothetical protein
TGGT1_031890	7	209	RNA and export factor binding protein, putative
TGGT1_081390	1	4	zinc finger protein, putative
TGME49_026240	1	4	zinc finger protein, putative
TGGT1_121030	3	13	transmembrane domain-containing protein, putative
TGME49_120260	3	13	transmembrane domain-containing protein
TGGT1_081790	4	15	Yippee family protein, putative
TGME49_026640	4	15	yippee-like protein, putative
TGME49_061500	18	569	hypothetical protein
TGGT1_008730	15	599	hypothetical protein
TGME49_085230	11	132	PRP38 family domain-containing protein
TGGT1_036730	9	128	conserved hypothetical protein
TGGT1_083070	4	19	conserved hypothetical protein
TGME49_027670	3	12	hypothetical protein
TGME49_019140	30	1986	elongation factor 1, putative
TGGT1_030100	12	1223	elongation factor, putative

TGGT1_063360	17	903	ribosomal protein S6 kinase / cAMP-dependent protein kinase catalytic subunit, putative
TGME49_005550	14	807	AGC kinase
TGGT1_011870	47	4459	chromodomain helicase DNA binding protein, putative
TGME49_058240	44	3834	chromodomain helicase DNA binding protein, putative
TGME49_073480	5	170	axonemal beta dynein heavy chain, putative
TGGT1_112560	1	1	dynein heavy chain, putative
TGME49_002870	19	726	mitochondrial carrier domain-containing protein
TGGT1_018080	11	635	conserved hypothetical protein
TGME49_050340	14	164	caltractin, putative
TGGT1_101950	8	84	caltractin, putative
TGME49_062710	7	104	ctr copper transporter domain-containing protein
TGGT1_007380	6	88	ctr copper transporter, putative
TGME49_056000	4	27	endoplasmic reticulum retention receptor, putative
TGGT1_070060	2	4	conserved hypothetical protein
TGME49_035710	2	4	hypothetical protein
TGGT1_050440	1	1	glyoxalase, putative
TGME49_040430	1	1	glyoxalase, putative
TGGT1_008510	6	144	conserved hypothetical protein
TGME49_061720	6	144	zinc transporter, putative
TGGT1_044550	1	2	conserved hypothetical protein
TGME49_064210	1	2	hypothetical protein
TGGT1_089840	7	144	conserved hypothetical protein
TGME49_113440	6	135	hypothetical protein
TGME49_039100	40	7336	40S ribosomal protein S7, putative
TGGT1_051860	19	1899	40S ribosomal protein S7, putative
TGGT1_045180	14	271	translation initiation factor, putative
TGME49_044650	13	256	eukaryotic translation initiation factor 5, putative
TGME49_064060	6	32	proteasome PCI domain-containing protein
TGGT1_044300	4	24	proteasome PCI domain-containing protein, putative
TGME49_013870	39	1917	ubiquitin carboxyl-terminal hydrolase, putative
TGGT1_038610	30	2099	ubiquitin carboxyl-terminal hydrolase, putative
TGGT1_056080	10	364	conserved hypothetical protein
TGME49_109050	9	269	hypothetical protein
TGME49_009680	9	57	hypothetical protein
TGGT1_022740	4	22	conserved hypothetical protein
TGME49_049530	45	2626	exportin, putative
TGGT1_101180	30	2224	exportin, putative
TGME49_093670	19	715	transcription elongation factor TFIIS, putative
TGGT1_074060	19	696	transcription elongation factor, putative
TGGT1_063540	2	6	conserved hypothetical protein
TGME49_005370	1	2	hypothetical protein
TGME49_033770	10	122	Ca2+-ATPase, putative
TGGT1_079650	4	39	conserved hypothetical protein

TGME49_024730	4	39	hypothetical protein
TGME49_100380	12	168	endoplasmic reticulum oxidoreductin, putative
TGGT1_094910	3	17	endoplasmic reticulum oxidoreductin, putative
TGGT1_012930	2	13	SET domain-containing protein, putative
TGME49_057770	2	13	SET domain-containing protein
TGGT1_005820	3	19	conserved hypothetical protein
TGME49_081550	3	19	hypothetical protein, conserved
TGGT1_084310	1	1	conserved hypothetical protein
TGME49_028730	1	1	hypothetical protein
TGME49_036050	28	2677	fructose-bisphosphate aldolase, putative
TGGT1_069700	9	740	fructose-bisphosphate aldolase, putative
TGME49_088210	20	507	phospholipase A2 activating protein, putative
TGGT1_035200	16	490	phospholipase A-2-activating protein, putative
TGME49_034280	14	515	adenosine monophosphate deaminase, putative
TGGT1_071610	13	610	adenosine monophosphate deaminase, putative
TGGT1_046580	6	114	conserved hypothetical protein
TGME49_043290	2	40	hypothetical protein
TGGT1_059530	11	127	protein phsophatase-2A, putative
TGME49_101010	11	127	hypothetical protein
TGME49_006420	27	1842	myosin head motor domain-containing protein
TGGT1_062370	2	2	myosin head motor domain-containing protein, putative
TGME49_028160	2	6	serine/threonine protein phosphatase, putative
TGGT1_083630	1	2	serine/threonine protein phosphatase, putative
TGME49_019700	44	3537	DNA replication licensing factor, putative
TGGT1_030690	25	1868	DNA replication licensing factor, putative
TGME49_062390	38	5045	hypothetical protein
TGGT1_007720	32	6007	hypothetical protein
TGME49_025320	17	273	hypothetical protein, conserved
TGGT1_080520	10	181	conserved hypothetical protein
TGME49_120480	16	1131	Rab 11b, putative
TGGT1_120910	6	43	rab11, putative
TGGT1_046350	8	72	conserved hypothetical protein
TGME49_043510	8	48	OTU-like cysteine protease domain-containing protein
TGGT1_081370	16	1692	conserved hypothetical protein
TGME49_026220	14	1097	hypothetical protein
TGME49_118570	8	213	SFT2-like domain-containing protein
TGGT1_122550	7	212	conserved hypothetical protein
TGME49_107980	3	6	GTP-binding protein lepA, putative
TGME49_005440	55	4817	TCP-1/cpn60 family chaperonin, putative
TGGT1_063460	28	3340	chaperonin containing t-complex protein 1, gamma subunit, tcpg, putative
TGME49_080710	21	976	proteasome A-type and B-type domain-containing protein
TGGT1_072590	12	749	proteasome A-type and B-type domain-containing protein, putative

TGME49_088360	154	51864	tryptophanyl-tRNA synthetase, putative
TGGT1_035040	46	12308	tryptophanyl-tRNA synthetase, putative
TGME49_075350	20	540	hypothetical protein, conserved
TGGT1_068160	19	564	conserved hypothetical protein
TGGT1_056300	9	254	conserved hypothetical protein
TGME49_109170	7	167	hypothetical protein
TGGT1_111840	18	474	conserved hypothetical protein
TGME49_072550	15	362	hypothetical protein
TGGT1_069860	2	8	hypothetical protein
TGME49_035910	2	8	hypothetical protein
TGME49_091890	48	36533	microneme protein MIC1
TGGT1_031650	31	3359	microneme protein, putative
TGME49_035620	18	276	deoxyhypusine synthase, putative
TGGT1_070160	12	226	deoxyhypusine synthase, putative
TGME49_012150	5	18	hypothetical protein
TGGT1_071890	3	9	conserved hypothetical protein
TGME49_105480	7	26	elongator complex protein 3, putative
TGGT1_041990	3	11	Radical SAM domain-containing protein, putative
TGGT1_031610	29	2694	RNA binding motif-containing protein, putative
TGME49_091950	28	2293	RNA binding motif-containing protein
TGME49_069080	1	4	hypothetical protein
TGGT1_033420	1	1	conserved hypothetical protein
TGME49_089790	1	1	hypothetical protein
TGME49_090990	3	11	hypothetical protein
TGGT1_032130	1	7	conserved hypothetical protein
TGME49_100320	1	1	dimethyladenosine synthase, putative
TGME49_048160	47	133589	ATP-dependent DNA helicase II, 70 kDa subunit, putative
TGGT1_024780	4	5	ATP-dependent DNA helicase II, 70 kDa subunit, putative
TGME49_001800	19	1110	RNA-binding protein, putative
TGGT1_019430	7	118	RNA-binding protein, putative
TGME49_080700	27	1410	lysine decarboxylase, putative
TGGT1_072600	19	1203	lysine decarboxylase, putative
TGME49_080390	35	437	hypothetical protein
TGGT1_072930	6	92	conserved hypothetical protein
TGME49_089300	69	6126	methionyl-tRNA synthetase, putative
TGGT1_033920	36	4016	methionyl-tRNA synthetase, putative
TGME49_070830	8	147	small nuclear ribonucleoprotein Sm D2, putative
TGGT1_109590	5	104	small nuclear ribonucleoprotein Sm D2, putative
TGGT1_050630	2	20	phosphate transporter, putative
TGME49_040210	2	20	phosphate transporter, putative
TGGT1_038030	10	517	hypothetical protein
TGME49_086600	6	320	hypothetical protein, conserved
TGME49_053580	1	1	CMGC kinase, CDK family
TGME49_020090	13	1912	hypothetical protein

TGGT1_014590	11	2384	conserved hypothetical protein
TGGT1_072530	5	32	regulator of chromosome condensation domain-containing protein, putative
TGME49_080770	5	32	regulator of chromosome condensation domain-containing protein
TGGT1_030200	2	6	conserved hypothetical protein
TGME49_019240	2	6	hypothetical protein, conserved
TGME49_093900	19	527	hypothetical protein
TGGT1_074290	11	196	conserved hypothetical protein
TGGT1_107460	16	1469	conserved hypothetical protein
TGME49_069260	14	1240	hypothetical protein
TGME49_058950	40	5916	lectin-domain protein
TGGT1_010920	13	525	lectin-domain protein, putative
TGGT1_039080	1	1	conserved hypothetical protein
TGME49_013540	1	1	hypothetical protein
TGME49_026970	23	1618	40S ribosomal protein S11, putative
TGGT1_082450	10	412	40S ribosomal protein S11, putative
TGME49_038380	16	715	hypothetical protein
TGGT1_052570	13	707	conserved hypothetical protein
TGME49_067080	35	1877	26S proteasome subunit 4, putative
TGGT1_075850	18	1025	26S proteasome subunit, putative
TGME49_072160	7	73	hypothetical protein
TGGT1_111380	6	111	conserved hypothetical protein
TGME49_115220	32	2558	rhoptry protein, putative
TGGT1_092650	9	129	conserved hypothetical protein
TGME49_069410	7	48	hypothetical protein
TGGT1_107720	6	40	conserved hypothetical protein
TGGT1_018930	12	333	conserved hypothetical protein
TGME49_002190	7	188	hypothetical protein
TGME49_060170	10	91	elongation factor G, putative
TGGT1_010060	8	87	elongation factor G, putative
TGME49_007680	10	114	S-phase kinase-associated protein, putative
TGGT1_020570	4	64	S-phase kinase-associated protein, putative
TGGT1_013830	4	32	conserved hypothetical protein
TGME49_056920	4	32	1-phosphatidylinositol-4-phosphate 5-kinase, putative
TGME49_094820	20	2250	type I fatty acid synthase, putative
TGME49_053690	1	21	hypothetical protein
TGGT1_002860	1	5	conserved hypothetical protein
TGGT1_125700	4	20	ferredoxin, putative
TGME49_015070	4	20	ferredoxin
TGGT1_003640	1	1	conserved hypothetical protein
TGGT1_042160	2	4	cyclin, N-terminal domain-containing protein, putative
TGGT1_075360	1	2	ATP-dependent RNA helicase, putative
TGME49_095010	1	2	ATP-dependent RNA helicase, putative
TGME49_065390	2	4	erythroblast macrophage protein EMP

TGGT1_057720	1	3	erythroblast macrophage protein EMP, putative
TGGT1_062220	6	80	conserved hypothetical protein
TGME49_006540	6	80	hypothetical protein
TGGT1_122910	1	3	MYST-type acetyltransferase, putative
TGME49_118330	1	3	MYST-family histone acetyltransferase-A
TGME49_077050	11	60	hypothetical protein
TGGT1_105280	7	75	conserved hypothetical protein
TGGT1_008120	6	58	gorasp2-prov protein, putative
TGME49_061980	3	45	gorasp2-prov protein
TGGT1_017040	16	192	conserved hypothetical protein
TGME49_003690	14	134	hypothetical protein
TGME49_113930	8	285	hypothetical protein, conserved
TGGT1_090570	6	251	conserved hypothetical protein
TGME49_044880	38	1068	DNA-directed RNA polymerase I largest subunit, putative
TGGT1_044930	28	1047	RNA polymerase Rpb1 domain-containing protein, putative
TGGT1_054250	3	28	hypothetical protein
TGME49_095770	3	28	hypothetical protein
TGME49_036850	22	490	hypothetical protein
TGGT1_069090	15	513	conserved hypothetical protein
TGME49_020330	2	2	hypothetical protein
TGGT1_015030	1	1	conserved hypothetical protein
TGGT1_108470	1	2	conserved hypothetical protein
TGME49_069930	1	2	hypothetical protein, conserved
TGME49_006490	5	18	metacaspase 1 precursor, putative
TGGT1_062280	2	9	conserved hypothetical protein
TGME49_110960	6	38	vacuolar ATP synthase subunit F, putative
TGGT1_087280	4	28	vacuolar ATP synthase subunit F, putative
TGME49_001390	3	115	hypothetical protein
TGGT1_020030	1	9	conserved hypothetical protein
TGME49_038040	83	13651	thioredoxin, putative
TGGT1_052810	28	999	thioredoxin, putative
TGGT1_042420	1	2	conserved hypothetical protein
TGME49_105080	1	2	hypothetical protein
TGGT1_009490	2	6	DNA primase small subunit, putative
TGME49_060680	2	6	DNA primase small subunit, putative
TGGT1_101360	5	21	conserved hypothetical protein
TGME49_049690	5	21	hypothetical protein
TGGT1_092180	1	2	ubiquitin-activating enzyme E1, putative
TGME49_078530	6	137	multiprotein bridging factor type 1, putative
TGGT1_104030	5	131	multiprotein bridging factor type, putative
TGME49_022020	26	777	phosphoglycerate kinase, putative
TGGT1_065850	17	667	phosphoglycerate kinase, putative
TGME49_004520	2	4	hypothetical protein
TGME49_033830	12	129	hypothetical protein, conserved
TGGT1_113500	4	40	conserved hypothetical protein

TGME49_069000	9	41	ABC transporter, putative
TGGT1_107190	1	2	ABC transporter, putative
TGGT1_018660	2	8	conserved hypothetical protein
TGME49_002420	1	3	hypothetical protein
TGME49_066930	3	9	hypothetical protein
TGGT1_007990	7	139	conserved hypothetical protein
TGME49_062120	7	139	hypothetical protein
TGGT1_026010	25	2527	TPR domain-containing protein, putative
TGME49_047000	20	1728	TPR domain-containing protein
TGGT1_006200	2	19	LSM domain-containing protein, putative
TGME49_098970	2	19	LSM domain-containing protein
TGGT1_063320	5	68	zinc finger (C3HC4 RING finger) protein
TGME49_005600	2	12	zinc finger (C3HC4 RING finger) protein, putative
TGME49_031080	7	139	hypothetical protein
TGGT1_116640	5	55	60S ribosomal protein L38, putative
TGME49_049950	4	24	MAK16 protein, putative
TGGT1_101630	3	15	MAK16 protein, putative
TGGT1_118350	18	399	conserved hypothetical protein
TGME49_029310	14	244	hypothetical protein
TGGT1_050330	2	4	conserved hypothetical protein
TGME49_040540	2	4	hypothetical protein
TGGT1_009240	2	7	NADPH:adrenodoxin oxidoreductase, putative
TGME49_061030	2	7	mitochondrial NADPH:adrenodoxin oxidoreductase, putative
TGGT1_025110	51	7522	conserved hypothetical protein
TGME49_047700	39	4737	hypothetical protein
TGME49_046460	9	134	hypothetical protein
TGGT1_026660	8	182	conserved hypothetical protein
TGGT1_065750	7	112	splicing factor 3A protein, putative
TGME49_021940	6	100	splicing factor 3a protein, putative
TGGT1_057840	1	2	conserved hypothetical protein
TGME49_065270	1	2	hypothetical protein
TGGT1_086460	24	1875	hypothetical protein
TGME49_110220	24	1875	guanylate binding protein, putative
TGGT1_092060	1	2	tRNA (guanine-N(1)-)-methyltransferase, putative
TGME49_114770	1	2	tRNA (guanine-N1)-methyltransferase domain containing protein
TGME49_030380	5	41	protein translocation complex, SEC61 gamma subunit, putative
TGGT1_117590	1	1	protein translocation complex, SEC61 gamma subunit, putative
TGGT1_060160	1	1	conserved hypothetical protein
TGME49_101690	1	1	hypothetical protein
TGGT1_124510	19	464	WD-repeat protein, putative
TGME49_014200	14	268	WD-40 repeat protein, putative
TGGT1_052130	5	28	conserved hypothetical protein

TGME49_038870	5	28	hypothetical protein
TGGT1_007030	4	34	enterophilin-2L, putative
TGME49_062950	4	21	enterophilin-2L, putative
TGME49_097900	2	5	hypothetical protein
TGGT1_096200	1	1	conserved hypothetical protein
TGME49_005560	83	6979	nascent polypeptide-associated complex alpha chain, putative
TGGT1_063350	26	3868	hypothetical protein
TGME49_074110	4	10	hypothetical protein, conserved
TGME49_089540	56	3876	hypothetical protein, conserved
TGGT1_033690	28	1595	conserved hypothetical protein
TGGT1_069660	2	6	conserved hypothetical protein
TGME49_036100	2	6	hypothetical protein
TGGT1_036690	5	38	conserved hypothetical protein
TGME49_085190	4	28	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_082670	23	563	dense granule protein, putative
TGME49_027280	18	5189	dense granule protein 3
TGGT1_042320	4	18	sodium/hydrogen exchanger, putative
TGME49_105180	3	12	sodium/hydrogen exchanger, putative
TGGT1_047290	10	146	conserved hypothetical protein
TGME49_042630	8	100	hypothetical protein
TGGT1_109120	1	2	conserved hypothetical protein
TGME49_070570	1	2	hypothetical protein
TGME49_054510	12	141	hypothetical protein, conserved
TGGT1_004080	11	172	conserved hypothetical protein
TGME49_059990	9	113	SAC3/GANP domain-containing protein
TGGT1_010130	9	109	SAC3/GANP domain-containing protein, putative
TGGT1_000730	8	74	protein kinase, putative
TGME49_075610	7	52	protein kinase
TGGT1_124500	1	1	SRS46
TGME49_014190	1	1	SRS46
TGGT1_028280	5	30	conserved hypothetical protein
TGME49_017770	4	12	hypothetical protein
TGME49_057960	18	524	mannose-1-phosphate guanylyltransferase, putative
TGGT1_012240	17	520	mannose-1-phosphate guanylyltransferase, putative
TGGT1_102940	14	489	GRA9 protein, putative
TGME49_051540	14	456	GRA9 protein, putative
TGME49_006590	61	8303	CAM kinase, CDPK family
TGGT1_062170	33	5664	calmodulin-domain protein kinase, putative
TGGT1_037620	4	32	NAD dependent epimerase/dehydratase, putative
TGME49_086220	3	14	GDP-L-fucose synthetase, putative
TGME49_040890	35	1438	phosphofructokinase, putative
TGGT1_049870	25	1372	phosphofructokinase, putative
TGGT1_051620	1	2	conserved hypothetical protein

TGME49_039350	1	2	hypothetical protein
TGME49_119370	4	22	hypothetical protein
TGGT1_121830	3	21	conserved hypothetical protein
TGGT1_029780	3	7	histidyl tRNA synthetase, putative
TGME49_018810	3	7	histidyl tRNA synthetase 2
TGME49_067070	4	43	aquaporin, putative
TGGT1_075860	1	1	conserved hypothetical protein
TGME49_046020	3	16	hypothetical protein
TGGT1_070430	11	424	ubiquitin-conjugating enzyme, putative
TGME49_035450	11	376	ubiquitin-conjugating enzyme, putative
TGME49_049000	8	46	regulator of chromosome condensation domain-containing protein
TGGT1_100670	6	31	regulator of chromosome condensation domain-containing protein, putative
TGGT1_082640	2	5	serine/threonine-protein kinase rio2, putative
TGME49_027260	2	5	RIO1 family domain-containing protein
TGGT1_093660	2	6	Radical SAM domain-containing protein, putative
TGME49_115960	2	6	radical SAM domain-containing protein
TGGT1_039880	15	349	conserved hypothetical protein
TGME49_012880	13	262	hypothetical protein
TGME49_088830	44	6122	pyridine nucleotide-disulphide oxidoreductase, putative
TGGT1_034430	23	1101	pyridine nucleotide-disulphide oxidoreductase, putative
TGME49_115210	25	479	rhoptry protein, putative
TGGT1_092640	2	8	conserved hypothetical protein
TGME49_090200	43	3061	opine dehydrogenase, putative
TGGT1_032940	25	2363	conserved hypothetical protein
TGGT1_088040	2	6	conserved hypothetical protein
TGME49_111790	1	4	hypothetical protein
TGGT1_118680	3	22	conserved hypothetical protein
TGME49_029150	2	9	sugar permease, putative
TGGT1_010740	1	2	conserved hypothetical protein
TGME49_059170	1	2	hypothetical protein
TGGT1_083260	9	78	conserved hypothetical protein
TGME49_027980	9	67	hypothetical protein
TGME49_032180	87	3300	UBA/TS-N domain-containing protein
TGGT1_115460	17	285	ubiquitin-associated domain-containing protein, putative
TGGT1_106910	4	18	conserved hypothetical protein
TGME49_068720	4	18	hypothetical protein
TGME49_101270	51	3139	Tyrosine kinase-like (TKL) protein
TGGT1_059690	43	3001	protein kinase domain-containing protein, putative
TGME49_072610	17	575	hypothetical protein
TGGT1_111900	14	558	conserved hypothetical protein
TGGT1_035760	3	15	phosphotyrosyl phosphatase activator, putative
TGME49_083720	1	3	phosphotyrosyl phosphatase activator, putative

TGME49_106330	37	1615	hypothetical protein, conserved
TGGT1_041010	18	1115	conserved hypothetical protein
TGGT1_076280	2	10	conserved hypothetical protein
TGME49_066690	2	10	hypothetical protein
TGGT1_123400	6	93	sumo-1 activating enzyme subunit, putative
TGME49_011390	6	93	hypothetical protein
TGME49_021510	13	966	hypothetical protein, conserved
TGGT1_065370	7	254	conserved hypothetical protein
TGGT1_007660	19	473	conserved hypothetical protein
TGME49_062450	14	280	hypothetical protein
TGGT1_079930	3	43	conserved hypothetical protein
TGME49_024990	3	43	hypothetical protein
TGGT1_022030	2	6	conserved hypothetical protein
TGME49_009090	2	6	hypothetical protein
TGGT1_003230	4	27	XAP5 protein, putative
TGME49_053950	4	27	XAP5 protein, putative
TGGT1_054490	5	80	conserved hypothetical protein
TGME49_095450	5	80	hypothetical protein
TGGT1_034470	4	24	endonuclease/exonuclease/phosphatase domain-containing protein, putative
TGME49_088800	4	24	endonuclease/exonuclease/phosphatase domain-containing protein
TGGT1_014130	1	1	conserved hypothetical protein
TGME49_056100	1	1	hypothetical protein
TGGT1_104940	13	465	conserved hypothetical protein
TGME49_077710	10	343	hypothetical protein
TGGT1_063920	2	12	hypothetical protein
TGME49_005120	1	6	hypothetical protein
TGME49_032390	5	30	RNA 3' terminal phosphate cyclase, putative
TGGT1_115240	4	29	RNA 3' terminal phosphate cyclase, putative
TGME49_073920	8	203	aldo/keto reductase family oxidoreductase, putative
TGGT1_113000	6	195	hypothetical protein
TGGT1_103440	1	4	conserved hypothetical protein
TGME49_078920	1	4	hypothetical protein, conserved
TGGT1_043490	5	81	conserved hypothetical protein
TGME49_063370	5	81	hypothetical protein
TGME49_062400	9	65	lipase domain-containing protein
TGGT1_007710	6	64	lipase domain-containing protein, putative
TGME49_097650	30	1266	serine/threonine protein phosphatase, putative
TGGT1_096530	12	311	serine/threonine protein phosphatase, putative
TGGT1_070620	3	39	conserved hypothetical protein
TGME49_035370	3	39	protein kinase
TGME49_090720	22	419	vacuolar proton-translocating ATPase subunit, putative
TGGT1_032330	10	169	vacuolar proton-translocating ATPase subunit, putative
TGME49_023110	6	162	glycolipid transfer protein, putative

TGGT1_066650	6	153	glycolipid transfer protein, putative
TGME49_052400	3	3	HIT zinc finger domain-containing protein
TGGT1_001310	2	2	HIT zinc finger protein domain-containing protein, putative
TGME49_048630	16	487	actin, putative
TGGT1_024220	14	457	actin, putative
TGGT1_121390	5	34	conserved hypothetical protein
TGME49_120000	5	34	SCY kinase (incomplete catalytic triad)
TGGT1_042740	21	1352	conserved hypothetical protein
TGME49_108060	21	1352	hypothetical protein
TGGT1_066420	24	4320	phosphoserine phosphatase, putative
TGME49_022890	22	3511	phosphoserine phosphatase, putative
TGME49_043580	24	819	HIT domain-containing protein
TGGT1_046290	7	302	HIT domain-containing protein, putative
TGME49_002500	9	93	hypothetical protein
TGGT1_018590	3	17	conserved hypothetical protein
TGME49_086410	1	1	hypothetical protein
TGME49_110360	7	241	hypothetical protein, conserved
TGGT1_086580	4	54	conserved hypothetical protein
TGGT1_020180	18	847	conserved hypothetical protein
TGME49_001140	17	820	proline rich protein, putative
TGME49_108940	1	1	hypothetical protein
TGME49_023710	21	731	N-terminal acetyltransferase complex subunit NARG1, putative
TGGT1_078070	18	660	TPR domain-containing protein, putative
TGME49_055490	1	2	hypothetical protein
TGGT1_125000	34	4316	conserved hypothetical protein
TGME49_014600	31	2698	hypothetical protein
TGME49_083540	22	833	hypothetical protein
TGGT1_035560	9	297	conserved hypothetical protein
TGME49_025790	11	313	hypothetical protein
TGGT1_080930	6	41	conserved hypothetical protein
TGGT1_124920	3	17	tfiih, polypeptide, putative
TGME49_014520	3	11	hypothetical protein
TGGT1_057010	17	561	pantoate--beta-alanine ligase, putative
TGME49_065870	15	454	pantoate--beta-alanine ligase, putative
TGME49_061680	12	296	hypothetical protein
TGGT1_008550	1	1	conserved hypothetical protein
TGGT1_005430	2	6	conserved hypothetical protein
TGME49_082040	2	6	hypothetical protein
TGME49_092990	7	18	hypothetical protein
TGGT1_073230	3	8	conserved hypothetical protein
TGME49_113700	3	5	hypothetical protein
TGGT1_090220	1	1	conserved hypothetical protein
TGME49_052290	49	4246	importin alpha, putative
TGGT1_001200	26	2700	importin alpha, putative

TGME49_001130	8	95	Rhoptry kinase family protein ROP33, putative
TGGT1_020190	4	46	serine/threonine protein kinase, putative
TGME49_109820	18	1812	ribosomal protein L11, putative
TGGT1_085780	7	206	ribosomal protein L11, putative
TGME49_011350	15	319	CBS domain-containing protein
TGGT1_123450	13	313	CBS domain-containing protein, putative
TGGT1_090290	6	36	conserved hypothetical protein
TGME49_113770	5	18	hypothetical protein
TGGT1_050460	1	3	map kinase kinase, putative
TGGT1_020370	7	81	conserved hypothetical protein
TGME49_007390	7	81	clathrin coat assembly protein, putative
TGME49_015470	29	6929	60S ribosomal protein L10a, putative
TGGT1_126410	13	852	60S ribosomal protein L10A, putative
TGME49_094380	13	356	aminotransferase class V domain-containing protein, conserved
TGGT1_074680	3	16	PP-loop domain-containing protein, putative
TGME49_048460	17	1414	ubiquitin, putative
TGGT1_024500	7	261	conserved hypothetical protein
TGME49_011680	119	196630	protein disulfide isomerase
TGGT1_123210	38	10985	protein disulfide isomerase, putative
TGGT1_065820	4	25	conserved hypothetical protein
TGME49_021990	4	25	hypothetical protein
TGME49_060520	7	40	hypothetical protein
TGGT1_009670	6	39	conserved hypothetical protein
TGGT1_028090	11	177	DNA polymerase alpha catalytic subunit, putative
TGME49_017910	10	137	DNA polymerase alpha catalytic subunit
TGME49_007130	2	11	SRS49A (= SAG2Y)
TGGT1_023370	1	3	SRS49A (= SAG2Y)
TGGT1_088510	1	3	RNA polymerase III, putative
TGME49_112250	1	3	DNA-directed RNA polymerase III subunit, putative
TGME49_021920	12	120	hypothetical protein
TGGT1_065730	8	104	hypothetical protein
TGGT1_025640	1	1	lipid phosphate phosphohydrolase, putative
TGME49_047360	1	1	lipid phosphate phosphohydrolase 3, putative
TGGT1_125870	1	1	conserved hypothetical protein
TGME49_015130	1	1	adaptor-related protein complex 3, sigma 2 subunit, putative
TGGT1_074830	3	10	conserved hypothetical protein
TGME49_094430	3	10	hypothetical protein, conserved
TGME49_069970	4	8	hypothetical protein
TGGT1_108510	2	4	conserved hypothetical protein
TGME49_063580	36	1896	26S proteasome non-ATPase regulatory subunit, putative
TGGT1_043800	22	1869	ankyrin repeat-containing protein, putative
TGGT1_062410	10	158	conserved hypothetical protein
TGME49_006380	7	86	hypothetical protein

TGGT1_047120	16	685	trichohyalin, putative
TGME49_042790	12	469	trichohyalin, putative
TGGT1_046760	20	1471	conserved hypothetical protein
TGME49_043210	17	1184	hypothetical protein
TGGT1_112490	1	2	tyrosyl-tRNA synthetase, putative
TGME49_001680	82	10692	eukaryotic translation initiation factor 3 subunit 10, putative
TGGT1_019750	53	8072	eukaryotic translation initiation factor 3 subunit, putative
TGME49_110430	28	769	heat shock protein 90, putative
TGGT1_086650	17	609	heat shock protein, putative
TGGT1_124110	1	2	conserved hypothetical protein
TGME49_010760	1	2	hypothetical protein
TGGT1_006420	2	4	transport protein particle component Bet3, putative
TGME49_099200	2	4	trafficking protein particle complex subunit 3, putative
TGGT1_009790	2	6	hypothetical protein
TGME49_060420	1	2	HEC/Ndc80p family protein, putative
TGGT1_017560	13	508	conserved hypothetical protein
TGME49_003300	12	457	hypothetical protein
TGGT1_110230	4	80	conserved hypothetical protein
TGME49_071360	2	35	hypothetical protein
TGME49_112470	3	16	hypothetical protein
TGGT1_088760	1	3	conserved hypothetical protein
TGME49_048320	16	244	mitochondrial carrier family protein
TGGT1_024630	3	16	mitochondrial carrier domain-containing protein
TGME49_069290	103	5594	hypothetical protein
TGGT1_107490	34	1885	conserved hypothetical protein
TGGT1_011510	3	6	TATA-box binding protein, putative
TGME49_058680	2	2	TATA-box binding protein, putative
TGME49_060460	1	1	oxidoreductase, putative
TGME49_076970	6	33	hypothetical protein
TGGT1_105350	2	8	hypothetical protein
TGGT1_092190	7	63	hypothetical protein
TGME49_114900	7	46	hypothetical protein
TGGT1_044270	7	74	aminotransferase, putative
TGME49_064030	6	58	aminotransferase, putative
TGME49_004050	71	25763	subtilase family serine protease, putative
TGGT1_016650	40	8699	subtilase family protein
TGME49_038050	75	13173	tudor / staphylococcal nuclease domain-containing protein
TGGT1_052800	48	8352	staphylococcal nuclease domain-containing protein
TGME49_026380	17	383	hypothetical protein
TGGT1_081550	13	286	conserved hypothetical protein
TGME49_068650	5	22	clp ATP-binding chain B1, putative
TGGT1_106830	4	21	Clp ATP-binding chain B1, putative
TGME49_080800	20	789	SNF2 family N-terminal domain-containing protein

TGGT1_072500	17	780	hypothetical protein
TGGT1_002060	8	149	ATP-dependent RNA helicase, putative
TGME49_053090	8	149	ATP-dependent RNA helicase, putative
TGGT1_028650	16	558	conserved hypothetical protein
TGME49_017440	12	337	hypothetical protein
TGME49_003310	46	577487	dense granule protein 7
TGGT1_017550	25	10173	dense granule protein, putative
TGGT1_052010	1	1	glutamic acid-rich protein, putative
TGME49_038970	1	1	glutamic acid-rich protein, putative
TGME49_066470	2	18	hypothetical protein, conserved
TGGT1_076390	2	12	conserved hypothetical protein
TGGT1_040810	3	8	conserved hypothetical protein
TGME49_106530	1	2	hypothetical protein
TGGT1_114170	2	7	prolidase, putative
TGME49_033310	2	7	prolidase, putative
TGGT1_065120	16	706	conserved hypothetical protein
TGME49_021280	12	519	hypothetical protein
TGGT1_098350	12	189	conserved hypothetical protein
TGME49_016780	8	95	hypothetical protein
TGGT1_102460	3	12	conserved hypothetical protein
TGME49_050900	3	12	hypothetical protein
TGGT1_045670	3	18	ABC transporter, putative
TGME49_044270	2	11	ABC transporter, putative
TGME49_002570	20	303	ribophorin I, putative
TGGT1_018510	3	24	ribophorin I, putative
TGGT1_030760	10	247	N-terminal acetyltransferase complex ARD1 subunit, putative
TGME49_019760	9	184	N-terminal acetyltransferase complex subunit ARD1, putative
TGGT1_058270	4	17	conserved hypothetical protein
TGME49_064950	3	7	hypothetical protein
TGGT1_104790	2	4	conserved hypothetical protein
TGME49_077860	2	4	hypothetical protein
TGME49_021950	21	506	splicing factor 3a protein, putative
TGGT1_065780	18	531	splicing factor 3A, putative
TGME49_006610	50	4742	biotin requiring domain-containing protein / 2-oxo acid dehydrogenases acyltransferase catalytic domain-containing protein
TGGT1_062130	25	1968	biotin requiring domain-containing protein / 2-oxo acid dehydrogenases acyltransferase catalytic domain-containing protein, putative
TGGT1_058030	16	551	conserved hypothetical protein
TGME49_065190	12	404	hypothetical protein, conserved
TGGT1_090490	9	82	zinc finger (C3HC4 type, RING finger) protein
TGME49_113870	8	49	zinc finger (C3HC4 type RING finger) protein, putative
TGME49_049810	7	86	activating signal cointegrator 1 complex subunit 3, putative

TGGT1_101490	6	72	activating signal cointegrator 1 complex subunit 3, helc1, putative
TGME49_046720	18	742	hypothetical protein
TGGT1_026370	17	1039	conserved hypothetical protein
TGME49_004400	64	132096	ATP synthase alpha chain, putative
TGGT1_016310	24	2374	ATP synthase alpha chain, putative
TGME49_027600	6	239	ribosomal protein L34, putative
TGGT1_083010	2	10	ribosomal protein L34, putative
TGGT1_078120	16	644	conserved hypothetical protein
TGME49_023760	11	460	hypothetical protein
TGGT1_034950	1	4	protein kinase, putative
TGME49_088440	1	4	NEK kinase
TGGT1_077190	2	6	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, putative
TGME49_055680	2	6	YbaK / prolyl-tRNA synthetases associated domain containing protein
TGME49_110040	8	243	ubiquitin-conjugating enzyme E2, putative
TGGT1_086080	7	190	ubiquitin-conjugating enzyme E2, putative
TGME49_120440	2	4	hypothetical protein
TGME49_088670	44	572	hypothetical protein
TGGT1_034720	2	10	conserved hypothetical protein
TGME49_014310	1	1	hypothetical protein
TGME49_014350	49	3301	GTP binding protein, putative
TGGT1_124770	19	1574	conserved hypothetical protein
TGGT1_038280	16	438	conserved hypothetical protein
TGME49_087180	15	335	hypothetical protein
TGGT1_108830	1	5	conserved hypothetical protein
TGME49_019630	38	4663	NADPH-cytochrome p450 reductase, putative
TGGT1_030510	23	1876	NADPH-cytochrome P450, putative
TGGT1_079280	10	424	conserved hypothetical protein
TGME49_024560	6	264	hypothetical protein
TGME49_042030	8	44	hypothetical protein
TGGT1_048380	7	69	conserved hypothetical protein
TGGT1_078760	32	2580	vacuolar sorting protein, putative vps41, putative
TGME49_024270	10	793	hypothetical protein, conserved
TGME49_048540	1	1	hypothetical protein
TGGT1_031580	7	41	ubiquitin-transferase domain-containing protein
TGME49_091980	6	31	ubiquitin-transferase domain-containing protein
TGGT1_083050	14	618	conserved hypothetical protein
TGME49_027650	13	579	microtubule-associated protein RP/EB family member 3, putative
TGME49_061570	53	12107	60S ribosomal protein L7a, putative
TGGT1_008660	20	1391	60S ribosomal protein L7A, putative
TGGT1_118360	6	59	kelch motif domain-containing protein, putative
TGME49_029290	4	28	kelch motif domain-containing protein

TGME49_015010	5	42	hypothetical protein
TGGT1_125540	3	18	conserved hypothetical protein
TGME49_060630	36	2979	DnaJ domain-containing protein
TGGT1_009550	12	432	DnaJ domain-containing protein, putative
TGME49_030830	1	1	dynein heavy chain, putative
TGGT1_071430	2	3	conserved hypothetical protein
TGME49_034460	1	2	hypothetical protein
TGME49_018270	7	245	hypothetical protein
TGGT1_028990	2	25	conserved hypothetical protein
TGME49_013690	1	1	hypothetical protein
TGME49_058580	60	7115	Rhoptry kinase family protein ROP17
TGGT1_011620	30	3475	protein kinase domain-containing protein, putative
TGGT1_002130	3	42	conserved hypothetical protein
TGME49_053160	2	26	hypothetical protein
TGGT1_045140	7	39	conserved hypothetical protein
TGME49_044680	5	24	hypothetical protein
TGME49_022830	6	42	hypothetical protein
TGGT1_066340	4	33	conserved hypothetical protein
TGME49_021870	52	3244	hypothetical protein
TGGT1_065640	40	2376	conserved hypothetical protein
TGME49_104630	51	6169	hypothetical protein
TGGT1_006680	2	247	hypothetical protein
TGME49_014120	3	13	hypothetical protein
TGGT1_124350	2	22	conserved hypothetical protein
TGME49_016070	7	106	hypothetical protein
TGGT1_099180	5	96	conserved hypothetical protein
TGME49_088620	10	97	erv1-alr family domain-containing protein, conserved
TGGT1_034770	7	79	alr/erv, putative
TGME49_091630	3	3	hypothetical protein
TGGT1_031810	1	1	conserved hypothetical protein
TGGT1_101480	2	9	conserved hypothetical protein
TGME49_049800	2	9	hypothetical protein
TGME49_086720	30	5452	heat shock protein 28
TGGT1_038120	15	496	heat-shock protein, putative
TGGT1_043780	7	69	conserved hypothetical protein
TGME49_063560	7	69	hypothetical protein
TGME49_074120	12	162	hypothetical protein, conserved
TGGT1_113300	1	6	conserved hypothetical protein
TGME49_039820	63	9718	D-3-phosphoglycerate dehydrogenase, putative
TGGT1_051010	34	5774	D-3-phosphoglycerate dehydrogenase, putative
TGGT1_031510	9	168	EF hand domain-containing protein, putative
TGME49_092050	8	117	EF hand domain-containing protein
TGGT1_035150	2	6	ADP-ribosylation factor, arf, putative
TGME49_088260	1	2	hypothetical protein
TGME49_019210	11	207	ankyrin repeat domain-containing protein

TGGT1_030170	9	200	ankyrin repeat-containing protein, putative
TGGT1_094380	13	910	conserved hypothetical protein
TGME49_116570	12	844	hypothetical protein, conserved
TGGT1_114970	6	66	conserved hypothetical protein
TGME49_032550	6	66	hypothetical protein
TGME49_105260	14	2084	acylglycerol lipase, putative
TGGT1_042240	11	2068	conserved hypothetical protein
TGME49_044020	5	29	hypothetical protein
TGGT1_045940	5	28	conserved hypothetical protein
TGGT1_078330	2	8	conserved hypothetical protein
TGME49_023950	2	8	hypothetical protein
TGGT1_115830	12	251	conserved hypothetical protein
TGME49_031930	11	196	hypothetical protein
TGGT1_110010	1	2	SRS33
TGGT1_035610	2	10	DnaJ domain-containing protein, putative
TGME49_083580	2	10	DnaJ domain-containing protein
TGME49_113140	45	4895	isocitrate dehydrogenase, putative
TGGT1_089430	22	2814	isocitrate dehydrogenase, putative
TGGT1_039340	1	1	conserved hypothetical protein
TGME49_013310	1	1	hypothetical protein
TGME49_030340	66	8048	high molecular mass nuclear antigen, putative
TGGT1_117640	32	1921	high molecular mass nuclear antigen, putative
TGME49_032230	9	635	60S ribosomal protein L30, putative
TGGT1_115410	5	215	60S ribosomal protein L30, putative
TGME49_091060	3	10	hypothetical protein
TGME49_015980	17	646	hypothetical protein
TGGT1_099280	10	301	conserved hypothetical protein
TGME49_014760	9	56	D-3-phosphoglycerate dehydrogenase, putative
TGGT1_125180	8	85	D-3-phosphoglycerate dehydrogenase, putative
TGGT1_102240	9	308	conserved hypothetical protein
TGME49_050790	7	227	hypothetical protein, conserved
TGME49_015440	7	112	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_126290	6	111	zinc finger (C3HC4 RING finger) protein
TGGT1_094300	6	84	conserved hypothetical protein
TGME49_116490	5	80	hypothetical protein
TGGT1_074030	2	5	conserved hypothetical protein
TGME49_093640	2	5	hypothetical protein
TGGT1_069530	1	2	cop9 signalosome complex subunit, putative
TGME49_036220	1	2	hypothetical protein
TGME49_002460	18	652	diacylglycerol kinase, putative
TGGT1_018630	13	588	diacylglycerol kinase, putative
TGME49_092020	172	16670	cysteine repeat modular protein, putative
TGGT1_031540	66	5652	conserved hypothetical protein
TGME49_042290	30	1914	proteasome subunit alpha type 6, putative
TGGT1_047540	18	1360	proteasome subunit alpha type, putative

TGGT1_113280	2	6	conserved hypothetical protein
TGME49_074100	2	6	hypothetical protein
TGGT1_101340	9	74	conserved hypothetical protein
TGME49_049680	9	74	hypothetical protein
TGGT1_043300	10	258	conserved hypothetical protein
TGME49_063200	9	225	hypothetical protein
TGGT1_011770	2	7	conserved hypothetical protein
TGME49_058440	2	7	hypothetical protein, conserved
TGGT1_041050	2	2	conserved hypothetical protein
TGME49_101390	23	838	ribosome biogenesis protein BOP1, putative
TGGT1_059820	19	813	WD domain, G-beta repeat containing protein, putative
TGME49_049360	9	38	hypothetical protein, conserved
TGGT1_101030	6	29	conserved hypothetical protein
TGGT1_037970	3	48	exosome complex exonuclease rrp43, putative
TGME49_086550	2	26	hypothetical protein, conserved
TGGT1_012020	3	58	conserved hypothetical protein
TGME49_058190	3	58	hypothetical protein
TGGT1_108620	1	1	conserved hypothetical protein
TGME49_070090	1	1	hypothetical protein
TGGT1_094950	3	11	conserved hypothetical protein
TGME49_100340	2	8	CAF1 family ribonuclease domain containing protein
TGGT1_039750	9	298	ctl transporter, putative
TGME49_012990	4	150	hypothetical protein, conserved
TGME49_011040	11	1900	Sec61beta family protein
TGGT1_123890	8	1331	conserved hypothetical protein
TGME49_013800	7	62	protein phosphatase 2B regulatory subunit
TGGT1_038700	5	46	protein phosphatase 2B regulatory subunit, putative
TGGT1_082260	6	54	tubulin gamma chain, putative
TGME49_026870	6	54	tubulin gamma chain, putative
TGME49_074070	10	91	thiF family domain-containing protein
TGGT1_113260	8	82	app binding protein, putative
TGME49_003810	12	75	kelch motif domain-containing protein
TGGT1_016920	9	59	kelch motif domain-containing protein, putative
TGGT1_031530	3	18	conserved hypothetical protein
TGME49_092030	1	4	hypothetical protein
TGGT1_043340	1	1	conserved hypothetical protein
TGME49_063230	1	1	hypothetical protein
TGME49_100120	29	1218	cysteine desulfurase, putative
TGGT1_095170	25	1323	cysteine desulfurase, putative
TGGT1_037670	5	31	conserved hypothetical protein
TGME49_086260	5	31	hypothetical protein
TGME49_050740	1	12	yip1 domain-containing protein
TGGT1_102190	1	3	YIP1 domain-containing protein, putative
TGME49_109760	24	551	hypothetical protein
TGGT1_085720	13	251	conserved hypothetical protein

TGME49_049850	18	24129	hypothetical protein, conserved
TGGT1_101540	3	189	conserved hypothetical protein
TGGT1_038290	6	60	conserved hypothetical protein
TGME49_087190	5	42	hypothetical protein
TGME49_031140	19	708	R1 protein
TGGT1_116570	10	374	conserved hypothetical protein
TGME49_043750	2	6	hypothetical protein
TGME49_043270	9	57	BRCA2 repeat-containing protein
TGME49_031890	9	76	3-oxoacyl-(acyl-carrier-protein) synthase III family protein, putative
TGGT1_115870	5	63	3-oxoacyl-[acyl-carrier-protein] synthase 3, putative
TGGT1_007690	7	39	NBP2B protein, putative
TGME49_062420	6	30	myb-like DNA-binding domain-containing protein / Zinc finger, ZZ type domain-containing protein
TGME49_048390	15	1070	60S ribosomal protein L26, putative
TGGT1_024560	9	192	60S ribosomal protein L26, putative
TGME49_083830	14	205	soluble inorganic pyrophosphatase
TGGT1_035860	13	220	hypothetical protein
TGME49_025150	22	475	hypothetical protein
TGGT1_080250	15	436	conserved hypothetical protein
TGME49_038100	6	205	emp24/gp25L/p24 family domain-containing, transmembrane protein, putative
TGGT1_052750	2	9	glycoprotein 25l, putative
TGGT1_025750	12	289	retinoblastoma-binding protein, putative
TGME49_047250	11	270	WD domain, G-beta repeat domain containing protein
TGGT1_035160	8	104	conserved hypothetical protein
TGME49_088250	5	56	hypothetical protein
TGME49_009930	1	16	hypothetical protein
TGGT1_092420	1	2	conserved hypothetical protein
TGME49_115120	1	2	hypothetical protein
TGME49_014910	1	5	hypothetical protein
TGME49_097140	5	26	u6 snRNA-associated sm-like protein Lsm2, putative
TGGT1_097080	3	17	conserved hypothetical protein
TGGT1_024590	12	378	conserved hypothetical protein
TGME49_048360	10	319	hypothetical protein
TGGT1_012130	24	1966	conserved hypothetical protein
TGME49_058080	22	1395	hypothetical protein
TGGT1_035910	4	58	RNA binding motif-containing protein, putative
TGME49_083880	3	40	RNA binding motif-containing protein
TGME49_083710	4	47	longevity-assurance (LAG1) domain-containing protein
TGGT1_035750	3	14	longevity-assurance LAG1 domain-containing protein
TGME49_069990	71	2446	hypothetical protein, conserved
TGGT1_108530	21	1559	conserved hypothetical protein
TGME49_080750	6	57	enhancer of rudimentary, putative

TGGT1_072550	3	30	enhancer of rudimentary, putative
TGGT1_088690	2	8	conserved hypothetical protein
TGME49_112400	2	8	hypothetical protein
TGME49_042690	1	30	hypothetical protein
TGME49_078850	9	157	zinc finger DHHC domain-containing protein
TGGT1_103510	7	106	DHHC domain-containing protein, putative
TGME49_019540	122	23480	alanyl-tRNA synthetase, putative
TGGT1_030420	50	13353	alanyl-tRNA synthetase, putative
TGME49_016260	18	517	eukaryotic initiation factor-2B gamma subunit, putative
TGGT1_099000	16	492	translation initiation factor eIF-2B gamma subunit, putative
TGGT1_122560	3	25	3-methyl-2-oxobutanoate dehydrogenase, putative
TGME49_118560	3	25	3-methyl-2-oxobutanoate dehydrogenase (lipoamide) kinase, putative
TGME49_060870	4	63	hypothetical protein
TGGT1_009300	3	18	conserved hypothetical protein
TGGT1_027050	1	1	conserved hypothetical protein
TGME49_046010	1	1	hypothetical protein
TGGT1_036930	7	48	conserved hypothetical protein
TGME49_085430	5	24	hypothetical protein
TGME49_077260	12	163	hypothetical protein
TGGT1_105200	4	48	conserved hypothetical protein
TGGT1_096880	1	2	conserved hypothetical protein
TGME49_097350	1	2	hypothetical protein
TGME49_066080	31	723	hypothetical protein
TGGT1_056680	12	283	conserved hypothetical protein
TGME49_104770	15	604	ubiquitin-conjugating enzyme E2, putative
TGGT1_006820	12	531	ubiquitin-conjugating enzyme E2, putative
TGGT1_096890	3	37	conserved hypothetical protein
TGME49_097340	3	37	hypothetical protein
TGME49_011690	36	1927	ATP-dependent helicase, putative
TGGT1_123200	28	1609	conserved hypothetical protein
TGGT1_077940	2	6	conserved hypothetical protein
TGME49_023580	2	6	hypothetical protein
TGME49_106570	3	5	hypothetical protein
TGME49_058660	11	238	rhoptry protein 6
TGGT1_011540	4	66	conserved hypothetical protein
TGME49_015570	1	2	hypothetical protein
TGME49_016180	4	1642	hypothetical protein
TGGT1_099080	2	9	hypothetical protein
TGGT1_007220	16	428	poly(ADP-ribose) glycohydrolase, putative
TGME49_062760	14	226	poly(ADP-ribose) glycohydrolase, putative
TGGT1_120530	8	110	conserved hypothetical protein
TGME49_120740	8	110	hypothetical protein
TGGT1_088640	3	13	RNA pseudouridylate synthase, putative
TGME49_112360	2	6	RNA pseudouridylate synthase, putative

TGME49_004310	29	746	hypothetical protein
TGGT1_016400	12	224	conserved hypothetical protein
TGGT1_070640	6	86	conserved hypothetical protein
TGME49_035350	6	86	hypothetical protein
TGGT1_036990	5	82	A/G-specific adenine glycosylase/endonuclease III, putative
TGME49_085490	5	82	hypothetical protein
TGGT1_114890	1	20	conserved hypothetical protein
TGME49_032630	1	20	hypothetical protein
TGME49_047030	3	304	hypothetical protein
TGGT1_025970	2	14	conserved hypothetical protein
TGME49_106540	24	1001	choline/ethanolamine kinase domain-containing protein
TGGT1_040800	19	1109	choline/ethanolamine kinase domain-containing protein, putative
TGGT1_013860	8	119	protein kinase domain-containing protein, putative
TGME49_056890	6	79	AGC kinase
TGGT1_045370	2	10	conserved hypothetical protein
TGME49_044570	2	10	hypothetical protein
TGME49_032710	50	8233	40S ribosomal protein S3a, putative
TGGT1_114810	21	2732	40S ribosomal protein S3A, putative
TGME49_121400	3	35	microsomal signal peptidase subunit SPCS1 domain-containing protein
TGGT1_064440	2	15	conserved hypothetical protein
TGGT1_002970	4	29	retinitis pigmentosa GTPase regulator protein, putative
TGME49_053810	2	14	hypothetical protein
TGME49_032190	50	5237	sec7 domain-containing protein
TGGT1_115450	48	5903	guanyl-nucleotide exchange factor, putative
TGGT1_118210	9	149	helicase, putative
TGME49_029460	8	118	SNF2 family helicase, putative
TGME49_028250	2	11	elongation factor Tu GTP-binding domain-containing protein
TGGT1_083740	2	6	GTP-binding domain containing protein, putative
TGME49_062690	14	488	ribosomal protein L27, putative
TGGT1_007400	6	73	ribosomal protein L27, putative
TGGT1_020380	4	29	conserved hypothetical protein
TGME49_007400	3	13	hypothetical protein, conserved
TGGT1_039870	4	22	conserved hypothetical protein
TGME49_012890	2	11	hypothetical protein
TGGT1_122900	1	1	conserved hypothetical protein
TGME49_118340	1	1	hypothetical protein
TGME49_101170	4	51	SRS19D
TGGT1_059590	3	14	SRS19D
TGME49_009690	10	68	U2 small nuclear ribonucleoprotein, putative
TGGT1_022750	4	30	u2 small nuclear ribonucleoprotein, putative
TGGT1_018360	2	8	protein phosphatase 2C, putative
TGME49_002610	1	4	protein phosphatase 2C, putative

TGGT1_029410	11	104	conserved hypothetical protein
TGME49_018580	10	71	RNA methyltransferase, TrmH family domain-containing protein
TGME49_068960	7	123	hypothetical protein
TGGT1_107150	7	122	hypothetical protein
TGME49_027800	35	1797	hypothetical protein, conserved
TGGT1_083100	23	1506	conserved hypothetical protein
TGME49_075470	5	84	hypothetical protein
TGGT1_068050	2	9	conserved hypothetical protein
TGGT1_086760	1	2	hypothetical protein
TGME49_110530	1	2	helicase, putative
TGGT1_048320	4	51	conserved hypothetical protein
TGME49_042100	4	51	hypothetical protein
TGME49_051780	72	43062	heat shock protein 70, putative
TGGT1_103150	35	5451	chaperone protein DNAK / heat shock protein, putative
TGME49_104670	17	654	internalin, putative
TGGT1_006720	6	70	internalin, putative
TGME49_109140	37	1767	transducin beta-like protein 1 (TgTBL1)
TGGT1_056280	23	1566	WD domain, G-beta repeat-containing protein, putative
TGME49_001120	4	18	hypothetical protein, conserved
TGGT1_020200	2	14	engulfment and cell motility, putative
TGME49_060370	6	147	PH domain-containing protein
TGGT1_009840	5	129	PH domain-containing protein, putative
TGME49_119590	3	13	hypothetical protein
TGGT1_121700	1	5	conserved hypothetical protein
TGGT1_099220	9	169	conserved hypothetical protein
TGME49_016030	9	169	hypothetical protein
TGGT1_082230	13	205	ecotropic viral integration site, putative
TGME49_026850	12	195	TBC domain-containing protein
TGGT1_082890	13	305	conserved hypothetical protein
TGME49_027390	12	279	hypothetical protein
TGME49_116240	8	75	dynactin subunit 4, putative
TGGT1_093930	8	55	dynactin, putative
TGGT1_042590	1	6	conserved hypothetical protein
TGME49_104900	1	6	hypothetical protein
TGME49_110180	5	37	hypothetical protein, conserved
TGGT1_086410	1	3	conserved hypothetical protein
TGME49_025090	4	16	hypothetical protein, conserved
TGGT1_080030	2	12	conserved hypothetical protein
TGGT1_104530	2	14	conserved hypothetical protein
TGME49_078020	2	14	trichohyalin, putative
TGGT1_008740	33	5206	hypothetical protein
TGME49_061490	25	3748	hypothetical protein
TGGT1_016390	2	12	conserved hypothetical protein
TGME49_004320	2	12	hypothetical protein

TGME49_115150	15	377	eukaryotic translation initiation factor 4E, putative
TGGT1_092580	14	446	hypothetical protein
TGME49_015220	23	1266	hypothetical protein, conserved
TGGT1_125960	17	994	hypothetical protein
TGME49_005040	29	810	hypothetical protein
TGGT1_064010	16	278	conserved hypothetical protein
TGME49_119520	11	148	hypothetical protein
TGME49_104950	21	813	hypothetical protein
TGGT1_042540	16	502	conserved hypothetical protein
TGGT1_068720	4	110	membrane fatty acid desaturase, putative
TGME49_037200	4	89	membrane fatty acid desaturase
TGME49_058150	29	1395	proteasome subunit alpha type 7, putative
TGGT1_012060	17	1050	proteasome subunit alpha type, putative
TGME49_076110	3	32	cytochrome b5, putative
TGGT1_000430	3	28	cytochrome B5, putative
TGGT1_093000	9	111	cnh domain cotaining protein, putative
TGME49_115530	6	54	hypothetical protein
TGGT1_041980	3	17	programmed cell death protein, putative
TGME49_105490	2	13	programmed cell death protein, putative
TGME49_031000	19	456	phosphatidylcholine transfer protein, putative
TGGT1_116720	9	348	phosphatidylcholine transfer protein, putative
TGGT1_069270	18	1115	lysine decarboxylase domain-containing protein, putative
TGME49_036570	18	896	lysine decarboxylase domain-containing protein
TGME49_040650	51	4426	coatomer alpha subunit, putative
TGGT1_050230	39	4050	hypothetical protein
TGME49_040550	3	16	cytochrome C oxidase copper chaperone, putative
TGGT1_050320	3	12	hypothetical protein
TGGT1_125730	4	24	conserved hypothetical protein
TGME49_015090	3	15	conserved hypothetical ATP-binding domain-containing protein
TGME49_116250	15	348	hypothetical protein
TGGT1_093940	5	68	conserved hypothetical protein
TGME49_047510	39	2687	fructose-1,6-bisphosphatase, putative
TGGT1_025400	17	1561	fructose-1,6-bisphosphatase, putative
TGGT1_003050	10	741	conserved hypothetical protein
TGME49_053880	8	605	GNS1/SUR4 family domain-containing protein
TGME49_039930	12	112	nuclear associated protein, putative
TGGT1_050900	8	99	nuclear associated protein, putative
TGGT1_021790	1	1	conserved hypothetical protein
TGME49_008760	1	1	hypothetical protein
TGGT1_042470	13	276	ribokinase, putative
TGME49_105030	13	238	ribokinase, putative
TGGT1_078570	9	228	adhesion regulating molecule, putative
TGME49_024110	7	115	adhesion regulating molecule, putative
TGME49_055210	20	84	peroxisomal-type ATPase, putative

TGGT1_076610	1	1	peroxisomal-type ATPase, putative
TGME49_020170	4	16	hypothetical protein
TGGT1_014770	3	15	hypothetical protein
TGME49_012270	10	135	hypothetical protein, conserved
TGGT1_072120	6	90	conserved hypothetical protein
TGME49_090020	3	10	hypothetical protein
TGGT1_068080	4	64	lysine-specific histone demethylase, putative
TGME49_075420	3	50	flavin-containing amine oxidase domain-containing protein
TGGT1_035030	2	6	hypothetical protein
TGME49_088370	1	2	hypothetical protein
TGGT1_076080	1	2	dihydrouridine synthase domain-containing protein, putative
TGME49_066880	1	2	dihydrouridine synthase domain-containing protein
TGGT1_106680	5	30	conserved hypothetical protein
TGME49_068600	4	20	DNA polymerase alpha subunit, putative
TGME49_088400	18	243	hypothetical protein, conserved
TGGT1_035000	10	207	leucine zipper-ef-hand containing transmembrane protein, putative
TGME49_025310	36	4898	ARF1-directed GTPase-activating protein, putative
TGGT1_080510	32	5917	ADP-ribosylation factor GTPase-activating protein, putative
TGME49_008740	3	8	microneme protein, putative
TGGT1_021770	1	2	microneme protein, putative
TGME49_002770	32	2092	RNA-binding protein, putative
TGGT1_018190	18	1704	hypothetical protein
TGGT1_042380	2	7	sodium/proline symporter, putative
TGME49_105120	2	7	sodium:solute symporter family domain-containing protein
TGGT1_065530	3	7	conserved hypothetical protein
TGME49_021660	3	7	hypothetical protein
TGGT1_031790	1	4	conserved hypothetical protein
TGME49_091650	1	4	hypothetical protein
TGME49_060480	9	39	leucine rich repeat protein, putative
TGGT1_110150	3	17	hypothetical protein
TGME49_071280	3	17	60S ribosome subunit biogenesis protein NIP7, putative
TGME49_028630	16	820	hypothetical protein
TGGT1_084200	11	372	conserved hypothetical protein
TGME49_049570	1	1	hypothetical protein
TGME49_047190	14	117	thrombospondin type 1 domain-containing protein
TGGT1_025910	12	139	thrombospondin type 1 domain-containing protein, putative
TGGT1_081000	21	782	conserved hypothetical protein
TGME49_025860	18	452	hypothetical protein
TGGT1_025930	3	11	conserved hypothetical protein
TGME49_047170	3	7	hypothetical protein
TGME49_020080	2	5	hypothetical protein
TGME49_064590	2	35	hypothetical protein

TGGT1_072540	3	14	ubiquitin-conjugating enzyme E2, putative
TGME49_080760	3	14	ubiquitin-conjugating enzyme E2, putative
TGME49_065260	16	1271	hypothetical protein
TGGT1_057850	15	1266	conserved hypothetical protein
TGME49_026860	6	199	hypothetical protein
TGGT1_082240	4	42	conserved hypothetical protein
TGME49_020890	15	589	hypothetical protein
TGGT1_053720	5	68	conserved hypothetical protein
TGME49_045470	29	998	mitotic checkpoint protein BUB3, putative
TGGT1_027590	12	539	mitotic checkpoint protein BUB3, putative
TGME49_113130	12	92	hypothetical protein
TGGT1_089420	10	76	conserved hypothetical protein
TGME49_072290	30	1091	pyruvate dehydrogenase E1 beta subunit, putative
TGGT1_111500	14	540	transketolase, putative
TGGT1_118060	1	2	conserved hypothetical protein
TGME49_029720	1	2	hypothetical protein, conserved
TGME49_026910	58	4946	glycogen debranching enzyme, putative
TGGT1_082290	56	5858	glycogen debranching enzyme, putative
TGME49_061710	10	155	ankyrin repeat-containing protein
TGGT1_008520	8	138	ankyrin repeat-containing protein, putative
TGGT1_103100	7	45	conserved hypothetical protein
TGME49_051730	6	39	hypothetical protein
TGME49_040570	20	524	hypothetical protein
TGGT1_050300	9	219	conserved hypothetical protein
TGME49_005700	19	2574	20 kDa cyclophilin precursor
TGGT1_063210	13	707	20k cyclophilin, putative
TGGT1_081060	9	114	conserved hypothetical protein
TGME49_025910	6	87	hypothetical protein
TGME49_113290	29	1565	phosphatidylinositol-4-phosphate 5-kinase, putative
TGGT1_089680	12	737	phosphatidylinositol-4-phosphate 5-kinase, putative
TGME49_049970	23	1098	hypothetical protein
TGGT1_101660	11	519	conserved hypothetical protein
TGME49_054370	58	5806	adenylate and guanylate cyclase catalytic domain-containing protein
TGGT1_003810	17	1526	guanylate cyclase, putative
TGME49_022840	34	1179	serine/threonine protein phosphatase, putative
TGGT1_066350	12	444	serine/threonine protein phosphatase, putative
TGGT1_086850	1	1	conserved hypothetical protein
TGME49_110620	1	1	hypothetical protein
TGME49_018260	20	319	histone H3.3 variant
TGGT1_028980	1	6	histone H3.3 variant, putative
TGGT1_103000	1	1	conserved hypothetical protein
TGME49_051610	1	1	hypothetical protein
TGGT1_111000	19	1338	conserved hypothetical protein
TGME49_071990	19	1078	hypothetical protein

TGME49_049270	56	4760	thioredoxin, putative
TGGT1_100850	20	1568	thioredoxin, putative
TGGT1_018150	1	1	dihydrouridine synthase domain-containing protein, putative
TGME49_070530	12	528	ubiquitin fusion degradation domain-containing protein
TGGT1_109070	8	350	ubiquitin fusion degradation UFD1 domain-containing protein
TGGT1_102200	12	345	RNA m5u methyltransferase, putative
TGME49_050750	8	189	hypothetical protein
TGGT1_050780	20	2381	conserved hypothetical protein
TGME49_040060	19	1318	hypothetical protein
TGGT1_104390	1	2	conserved hypothetical protein
TGME49_078160	1	2	hypothetical protein
TGME49_072390	5	19	hypothetical protein
TGGT1_111580	3	19	hypothetical protein
TGME49_088860	17	523	ruvB-like 2 protein, putative
TGGT1_034400	16	506	conserved hypothetical protein
TGME49_063130	19	232	citrate synthase, putative
TGGT1_043240	7	70	citrate synthase, putative
TGGT1_069680	6	48	replication factor, putative
TGME49_036080	6	48	hypothetical protein
TGME49_009110	17	941	mRNA decapping enzyme, putative
TGGT1_022050	15	877	conserved hypothetical protein
TGGT1_001110	1	1	conserved hypothetical protein
TGME49_052190	1	1	hypothetical protein, conserved
TGGT1_037240	8	123	conserved hypothetical protein
TGME49_085840	4	42	hypothetical protein
TGME49_053800	4	20	ribosomal protein L15, putative
TGGT1_002960	1	1	ribosomal protein L15, putative
TGGT1_117760	30	3379	RAB GDP/GTP exchange factor, putative
TGME49_030140	28	2659	vacuolar sorting protein 9 domain-containing protein
TGGT1_098240	1	1	conserved hypothetical protein
TGME49_016890	1	1	hypothetical protein, conserved
TGGT1_010280	1	2	conserved hypothetical protein
TGME49_059850	1	2	hypothetical protein, conserved
TGGT1_040700	4	53	vesicle transport v-SNARE domain-containing protein
TGME49_106640	4	37	vesicle transport v-SNARE domain-containing protein
TGGT1_018890	11	178	histone deactylase, putative
TGME49_002230	9	134	histone deactylase, putative
TGME49_108020	26	9277	SRS57 (= SAG3, P43)
TGGT1_042790	19	1227	SRS57 (= SAG3, P43)
TGME49_113600	15	369	hypothetical protein
TGGT1_090010	6	59	conserved hypothetical protein
TGME49_004140	2	4	PHD-zinc finger (C3HC4 type) protein, putative
TGME49_107870	18	243	hypothetical protein
TGGT1_042940	5	25	conserved hypothetical protein

TGME49_041170	31	1605	KH domain-containing protein
TGGT1_049490	29	1618	KH domain-containing protein, putative
TGGT1_124720	2	3	conserved hypothetical protein
TGME49_014300	2	3	hypothetical protein
TGME49_061210	37	1806	26S proteasome non-ATPase regulatory subunit, putative
TGGT1_009050	21	1225	26S proteasome non-ATPase regulatory subunit, putative
TGGT1_013180	3	62	aminoalcoholphosphotransferase, putative
TGME49_057510	3	62	aminoalcoholphosphotransferase, putative
TGME49_077930	2	4	hypothetical protein
TGGT1_010370	5	60	hypothetical protein
TGME49_059650	4	50	zinc finger (C2H2 type) protein, putative
TGME49_110300	8	49	hypothetical protein
TGGT1_086530	3	12	conserved hypothetical protein
TGME49_056830	3	8	hypothetical protein
TGGT1_013920	2	10	conserved hypothetical protein
TGGT1_118040	3	12	conserved hypothetical protein
TGME49_029740	3	8	hypothetical protein, conserved
TGME49_105340	87	12705	hypothetical protein
TGGT1_042150	71	11509	kelch repeat domain, putative
TGME49_039330	3	36	60s ribosomal protein l37, putative
TGGT1_051640	2	12	60S ribosomal protein l37, putative
TGGT1_090540	4	25	calmodulin-binding domain containing protein, putative
TGME49_113900	1	8	IQ calmodulin-binding motif domain-containing protein
TGME49_114860	20	1095	zinc knuckle domain-containing protein
TGGT1_092160	16	1141	zinc knuckle domain-containing protein, putative
TGME49_049990	63	6246	microtubule-binding protein, putative
TGGT1_101680	20	770	microtubule-binding protein, putative
TGME49_073960	22	691	10 kDa chaperonin, putative
TGGT1_113140	11	288	10 kDa chaperonin, putative
TGME49_005350	7	230	fatty acid elongation protein, putative
TGGT1_063560	3	12	conserved hypothetical protein
TGME49_094840	7	65	zinc finger (CCCH type) protein, putative
TGGT1_075170	6	64	zinc finger CCCH type) protein
TGME49_120060	8	101	hypothetical protein
TGME49_106980	2	7	hypothetical protein
TGGT1_055430	1	6	conserved hypothetical protein
TGGT1_021320	13	259	polynucleotide kinase-3''''-phosphatase, putative
TGME49_008320	13	259	polynucleotide kinase-3'-phosphatase, putative
TGME49_036110	10	156	autophagocytosis associated protein, putative
TGGT1_069650	4	134	autophagocytosis associated protein, putative
TGME49_068900	9	95	dense granular protein GRA10
TGGT1_107090	7	91	conserved hypothetical protein
TGGT1_050760	3	8	conserved hypothetical protein

TGME49_040080	1	2	transporter, major facilitator family domain-containing protein
TGGT1_090740	27	2422	conserved hypothetical protein
TGME49_114100	25	2098	hypothetical protein
TGME49_028640	1	1	hypothetical protein, conserved
TGGT1_082050	2	2	cGMP-inhibited 3',5'-cyclic phosphodiesterase, putative
TGME49_026770	2	2	cGMP-inhibited 3',5'-cyclic phosphodiesterase, putative
TGGT1_044360	1	5	snap190, putative
TGME49_064120	1	5	hypothetical protein
TGME49_038070	14	229	PKC-interacting cousin of thioredoxin, putative
TGGT1_052780	4	32	PKC-interacting cousin of thioredoxin, putative
TGGT1_028530	20	2289	splicing factor, putative
TGME49_017540	17	1417	splicing factor, putative
TGGT1_117220	1	1	conserved hypothetical protein
TGME49_030630	1	1	hypothetical protein
TGGT1_080790	21	658	hypothetical protein
TGME49_025560	15	455	hypothetical protein
TGME49_066970	13	268	hypothetical protein
TGGT1_075970	9	204	conserved hypothetical protein
TGGT1_072800	2	24	conserved hypothetical protein
TGME49_080510	2	24	Dpy-30 motif domain-containing protein
TGGT1_032680	3	22	conserved hypothetical protein
TGME49_090570	3	22	hypothetical protein, conserved
TGGT1_071880	7	64	conserved hypothetical protein
TGME49_012140	7	64	hypothetical protein
TGGT1_087610	1	1	conserved hypothetical protein
TGME49_111280	1	1	hypothetical protein
TGGT1_032520	41	6219	hypothetical protein
TGME49_090620	38	5229	hypothetical protein
TGGT1_117570	10	284	conserved hypothetical protein
TGME49_030400	10	284	hypothetical protein
TGME49_104990	62	3139	guanylate binding protein, putative
TGGT1_042510	14	408	guanylate binding protein, putative
TGGT1_087870	6	50	protein kinase domain-containing protein, putative
TGME49_111510	6	50	PIK3R4 kinase-related protein (incomplete catalytic triad)
TGME49_015990	64	502	helicase, putative
TGGT1_099270	1	1	conserved hypothetical protein
TGME49_104510	6	83	hypothetical protein
TGGT1_006660	5	82	conserved hypothetical protein
TGME49_110870	4	6	hypothetical protein, conserved
TGGT1_087190	1	1	conserved hypothetical protein
TGME49_090600	34	1524	succinyl-CoA ligase alpha subunit, putative
TGGT1_032540	14	503	succinyl-CoA ligase alpha subunit, putative
TGGT1_114850	4	22	conserved hypothetical protein

TGME49_032670	4	22	hypothetical protein, conserved
TGGT1_113175	4	28	cobW/P47K family domain-containing protein
TGME49_073990	4	21	cobW/P47K family domain-containing protein
TGGT1_125170	17	744	conserved hypothetical protein
TGME49_014750	17	715	hypothetical protein
TGGT1_012030	1	1	conserved hypothetical protein
TGME49_058180	1	1	hypothetical protein
TGGT1_058220	6	80	conserved hypothetical protein
TGME49_064990	6	80	hypothetical protein, conserved
TGME49_032440	39	3817	hypothetical protein
TGGT1_115180	22	3027	NBP2B protein, putative
TGGT1_007700	8	119	WD domain, G-beta repeat-containing protein, putative
TGME49_062410	7	81	WD domain, G-beta repeat-containing protein
TGME49_055340	24	1011	TPR domain-containing protein
TGGT1_076760	23	1421	TPR domain-containing protein, putative
TGGT1_016190	4	23	conserved hypothetical protein
TGME49_004490	4	23	hypothetical protein
TGME49_108580	6	28	lon protease, putative
TGGT1_055620	5	57	lon protease, putative
TGME49_110830	1	1	delta3,5-delta2,4-dienoyl-CoA isomerase, putative
TGME49_058390	43	3342	DnaJ protein, putative
TGGT1_011820	25	1646	DnaJ domain-containing protein, putative
TGME49_061430	15	1020	hypothetical protein
TGGT1_008810	12	1263	conserved hypothetical protein
TGGT1_044200	1	1	PPR repeat-containing protein, putative
TGME49_063980	1	1	PPR repeat-containing protein
TGME49_077970	12	290	dolichol-phosphate mannosyltransferase, putative
TGGT1_104570	1	2	dolichol-phosphate mannosyltransferase, putative
TGGT1_095180	5	31	cle7, putative
TGME49_100110	4	23	hypothetical protein
TGME49_025080	15	1008	40S ribosomal protein S18, putative
TGGT1_080020	9	262	40S ribosomal protein S18, putative
TGGT1_109930	6	372	SRS34A (= SAG2A, P22)
TGME49_071050	4	230	SRS34A (= SAG2A, P22)
TGGT1_050350	4	36	conserved hypothetical protein
TGME49_040520	4	36	hypothetical protein
TGME49_108080	120	862446	Rhoptry kinase family protein ROP5 (incomplete catalytic triad)
TGGT1_042710	34	6728	rhoptry antigen, putative
TGGT1_084300	1	1	conserved hypothetical protein
TGME49_028720	1	1	hypothetical protein
TGGT1_038100	2	8	conserved hypothetical protein
TGME49_086700	2	8	hypothetical protein
TGGT1_009830	8	352	conserved hypothetical protein

TGME49_060380	8	352	hypothetical protein, conserved
TGME49_057780	3	22	hypothetical protein
TGGT1_012920	2	21	conserved hypothetical protein
TGME49_089500	5	22	hypothetical protein
TGGT1_033720	4	18	conserved hypothetical protein
TGGT1_104010	3	12	GTP-binding domain containing protein, putative
TGME49_078550	2	6	elongation factor Tu GTP-binding domain-containing protein
TGME49_072490	10	128	protoporphyrinogen oxidase, putative
TGGT1_111770	3	15	conserved hypothetical protein
TGGT1_045890	2	8	ankyrin repeat-containing protein, putative
TGME49_044060	1	2	ankyrin repeat-containing protein
TGME49_010440	5	23	hypothetical protein, conserved
TGGT1_053190	3	19	conserved hypothetical protein
TGME49_051850	25	463	serine/threonine protein phosphatase 5, putative
TGGT1_103220	17	373	serine/threonine protein phosphatase, putative
TGGT1_049710	10	268	conserved hypothetical protein
TGME49_040960	7	158	hypothetical protein
TGGT1_000640	46	4394	Clp ATP-binding chain B1, putative
TGME49_075690	44	3365	chaperone clpB 1 protein, putative
TGME49_026040	5	30	hypothetical protein, conserved
TGGT1_081180	4	29	calmodulin, putative
TGGT1_092810	1	4	conserved hypothetical protein
TGME49_048640	4	8	regulator of chromosome condensation domain-containing protein
TGGT1_046400	46	5987	conserved hypothetical protein
TGME49_043460	42	5179	hypothetical protein
TGME49_106030	23	1246	hypothetical protein
TGGT1_041320	9	153	conserved hypothetical protein
TGME49_062980	6	114	hypothetical protein
TGGT1_006990	6	113	conserved hypothetical protein
TGME49_044600	3	4	hypothetical protein
TGGT1_045230	2	3	conserved hypothetical protein
TGGT1_083170	3	51	conserved hypothetical protein
TGME49_027870	3	51	hypothetical protein
TGME49_119660	10	218	hypothetical protein
TGGT1_121630	6	202	nucleolar complex protein, putative
TGGT1_051060	1	2	conserved hypothetical protein
TGME49_039770	1	2	hypothetical protein
TGGT1_041570	2	6	conserved hypothetical protein
TGME49_105900	2	6	hypothetical protein
TGGT1_086400	17	621	hypothetical protein
TGME49_110170	15	547	hypothetical protein
TGME49_078030	27	940	hypothetical protein
TGGT1_104520	18	779	conserved hypothetical protein
TGME49_031100	10	55	hypothetical protein

TGGT1_116620	3	12	conserved hypothetical protein
TGME49_089770	11	225	mu1 adaptin
TGGT1_033440	10	216	mu1 adaptin, putative
TGME49_110640	34	3964	uridine phosphorylase, putative
TGGT1_086870	20	3236	purine nucleoside phosphorylase, putative
TGGT1_098300	1	10	conserved hypothetical protein
TGGT1_021170	18	1653	inositol polyphosphate kinase domain-containing protein, putative
TGME49_008070	15	1264	inositol polyphosphate kinase domain-containing protein
TGME49_070070	5	25	vesicle trafficking protein, putative
TGGT1_108600	4	15	vesicle trafficking protein, putative
TGGT1_122370	1	2	deoxyribose-phosphate aldolase, putative
TGME49_118750	1	2	deoxyribose-phosphate aldolase, putative
TGGT1_098260	5	99	excision repair protein Rad15, putative
TGME49_016870	5	99	excision repair protein rad15, putative
TGME49_058170	13	668	hypothetical protein
TGGT1_012040	11	659	conserved hypothetical protein
TGGT1_073380	2	9	hypothetical protein
TGME49_093040	2	9	hypothetical protein
TGME49_112050	20	376	Ras family domain-containing protein
TGGT1_088300	8	69	RAB-2,4,14, putative
TGME49_039610	8	113	hypothetical protein
TGGT1_051250	6	78	conserved hypothetical protein
TGME49_019190	3	14	nuclear movement domain-containing protein
TGGT1_030150	2	10	nuclear movement domain-containing protein, putative
TGGT1_084230	1	2	conserved hypothetical protein
TGGT1_026850	4	24	hypothetical protein
TGME49_046200	2	4	zinc finger (CCCH type) protein, putative
TGGT1_017600	1	2	conserved hypothetical protein
TGME49_003260	1	2	hypothetical protein
TGME49_028100	23	561	hypothetical protein
TGGT1_083370	22	636	conserved hypothetical protein
TGGT1_125040	2	8	hypothetical protein
TGME49_014630	1	3	hypothetical protein
TGGT1_031730	2	12	conserved hypothetical protein
TGME49_091810	2	12	hypothetical protein
TGME49_034250	1	1	neurofilament triplet M protein, putative
TGME49_033930	5	17	poly(A)-specific ribonuclease, putative
TGME49_057520	2	10	synaptobrevin-like protein, putative
TGME49_066100	2	12	Rhoptry kinase family protein ROP41
TGGT1_056660	1	5	protein kinase domain-containing protein, putative
TGME49_058800	2	4	Rhoptry kinase family protein ROP31
TGME49_035160	30	843	hypothetical protein
TGGT1_070950	20	765	conserved hypothetical protein
TGGT1_037150	7	123	dihydropteridine reductase, putative

TGME49_085750	7	82	hypothetical protein
TGME49_093190	18	504	endonuclease/exonuclease/phosphatase domain-containing protein
TGGT1_073540	7	219	endonuclease/exonuclease/phosphatase domain-containing protein, putative
TGME49_105930	7	97	hypothetical protein
TGGT1_041540	4	47	conserved hypothetical protein
TGME49_009830	1	1	myosin heavy chain, putative
TGGT1_007640	15	259	PX domain-containing protein, putative
TGME49_062460	14	184	PX domain-containing protein
TGGT1_072510	3	22	conserved hypothetical protein
TGME49_080790	3	22	N-acetylglucosaminyl transferase component domain containing protein
TGGT1_100580	3	23	conserved hypothetical protein
TGME49_048910	2	12	hypothetical protein
TGME49_108070	3	5	hypothetical protein
TGGT1_042730	1	1	conserved hypothetical protein
TGME49_026550	10	85	hypothetical protein, conserved
TGGT1_081700	7	76	conserved hypothetical protein
TGME49_087270	16	297	hypothetical protein
TGGT1_038390	8	106	conserved hypothetical protein
TGGT1_086820	2	3	conserved hypothetical protein
TGME49_110600	2	3	hypothetical protein
TGME49_110450	25	374	myosin heavy chain, putative
TGGT1_086670	11	225	conserved hypothetical protein
TGME49_113000	2	5	hypothetical protein
TGGT1_098390	9	441	riboflavin kinase/fmn adenylyltransferase, putative
TGME49_016740	4	143	hypothetical protein
TGME49_053870	4	8	hypothetical protein
TGGT1_079340	1	2	conserved hypothetical protein
TGME49_024620	1	2	hypothetical protein
TGGT1_045590	1	2	DnaJ domain-containing protein, putative
TGME49_044350	1	2	DnaJ domain-containing protein
TGME49_030110	6	29	hypothetical protein
TGGT1_117780	4	8	conserved hypothetical protein
TGGT1_041560	11	138	3-mercaptopyruvate sulfurtransferase, putative
TGME49_105910	9	95	3-mercaptopyruvate sulfurtransferase, putative
TGME49_088960	2	4	ATP-binding surface antigen, putative
TGME49_105520	31	5057	40S ribosomal protein S2, putative
TGGT1_041950	15	1216	40S ribosomal protein S2, putative
TGGT1_124660	14	257	bromodomain-containing protein, putative
TGME49_014240	12	163	bromodomain-containing protein
TGGT1_076130	28	2579	guanyl-nucleotide exchange factor, putative
TGME49_066830	28	2415	sec7 domain-containing protein
TGME49_109980	1	1	dynein-1-alpha heavy chain, flagellar inner arm I1 complex, putative

TGME49_119340	10	181	kelch motif domain-containing protein
TGGT1_121860	8	170	kelch motif domain-containing protein, putative
TGME49_010370	28	882	RNA helicase-1
TGGT1_053260	13	589	RNA helicase-1, putative
TGME49_090660	47	3436	polypyrimidine track-binding protein, putative
TGGT1_032370	16	2602	polypyrimidine track-binding protein, putative
TGGT1_058230	3	12	conserved hypothetical protein
TGME49_064980	3	12	hypothetical protein
TGME49_010800	15	205	hypothetical protein, conserved
TGGT1_124070	9	177	conserved hypothetical protein
TGME49_055160	16	211	hypothetical protein
TGGT1_076560	1	8	conserved hypothetical protein
TGME49_097070	56	20489	duplicated carbonic anhydrase
TGGT1_097160	16	1076	duplicated carbonic anhydrase, putative
TGME49_053730	31	2109	importin-alpha re-exporter, putative
TGGT1_002900	29	2080	importin-alpha re-exporter, putative
TGME49_049610	24	502	nuclear cap binding protein, putative
TGGT1_101280	20	771	nuclear cap binding protein, putative
TGGT1_096490	21	1534	conserved hypothetical protein
TGME49_097690	18	1167	hypothetical protein, conserved
TGME49_113560	11	76	60S ribosomal protein L7a, putative
TGGT1_089970	3	24	hypothetical protein
TGGT1_010070	2	7	williams-beuren syndrome critical region protein, putative
TGME49_060160	1	5	hypothetical protein
TGME49_101130	5	529	hypothetical protein
TGGT1_059550	5	528	conserved hypothetical protein
TGGT1_002740	1	3	conserved hypothetical protein
TGME49_053570	1	3	hypothetical protein
TGME49_046920	8	72	glutathione reductase, putative
TGGT1_026180	7	71	glutathione reductase, putative
TGGT1_014290	3	49	conserved hypothetical protein
TGME49_055930	2	36	hypothetical protein
TGME49_027450	40	3743	mRNA decapping protein, putative
TGGT1_082960	38	3775	conserved hypothetical protein
TGGT1_105740	7	90	conserved hypothetical protein
TGME49_064460	7	80	SIN-like domain-containing protein
TGGT1_105970	2	6	conserved hypothetical protein
TGME49_067970	2	6	helicase conserved C-terminal domain-containing protein
TGME49_030860	20	485	autophagy protein, putative
TGGT1_116985	13	461	autophagy protein, putative
TGME49_094620	49	4128	eukaryotic translation initiation factor 3 subunit 8, putative
TGGT1_074940	27	3296	eukaryotic translation initiation factor 3 subunit, putative
TGGT1_031870	6	60	conserved hypothetical protein
TGME49_091350	3	21	hypothetical protein

TGME49_078220	4	17	hypothetical protein
TGGT1_104340	2	8	hypothetical protein
TGME49_111410	14	220	U3 small nucleolar ribonucleoprotein protein MPP10, putative
TGGT1_087760	12	231	hypothetical protein
TGME49_115240	1	1	hypothetical protein
TGGT1_044430	14	510	tRNA pseudouridine synthase D, putative
TGME49_064190	13	438	hypothetical protein
TGME49_048700	29	5422	hypothetical protein, conserved
TGGT1_023950	16	2778	conserved hypothetical protein
TGME49_043490	13	89	bcs1 protein, putative
TGGT1_046370	3	18	bcs1 protein, putative
TGGT1_103270	4	33	conserved hypothetical protein
TGME49_051890	4	17	hypothetical protein
TGME49_027350	7	46	hypothetical protein
TGGT1_082850	6	63	conserved hypothetical protein
TGGT1_024750	3	11	conserved hypothetical protein
TGME49_048190	1	4	hypothetical protein
TGME49_031970	91	7918	pre-mRNA splicing factor PRP8, putative
TGGT1_115790	54	6065	pre-mRNA splicing factor PRP8, putative
TGME49_035470	145	2105940	myosin A, putative
TGGT1_070410	75	36042	myosin A, putative
TGGT1_071860	2	12	jmjC domain-containing protein, putative
TGME49_012120	1	5	jmjC domain-containing protein
TGGT1_005520	1	2	conserved hypothetical protein
TGME49_081960	1	2	hypothetical protein
TGME49_119560	47	18694	microneme protein MIC3
TGGT1_121730	25	4283	microneme protein, putative
TGME49_066320	20	281	hypothetical protein, conserved
TGGT1_056560	6	42	glycosyltransferase sugar-binding region containing DXD motif-containing protein, putative
TGME49_063090	90	20000	14-3-3 protein, putative
TGGT1_043200	22	6157	conserved hypothetical protein
TGME49_047760	38	1917	long chain acyl-CoA synthetase, putative
TGGT1_025050	31	3093	long chain acyl-CoA synthetase, putative
TGME49_093060	15	124	SPRY domain-containing protein
TGGT1_073410	10	90	SPRY domain-containing protein, putative
TGGT1_007980	8	143	conserved hypothetical protein
TGME49_062130	5	76	hypothetical protein
TGGT1_106320	4	26	WD-repeat protein, putative
TGME49_068250	4	26	WD-40 repeat protein, putative
TGGT1_025390	31	1399	conserved hypothetical protein
TGME49_047520	9	262	hypothetical protein, conserved
TGME49_029950	18	730	26S protease regulatory subunit 6b, putative
TGGT1_117930	11	585	26S protease regulatory subunit 6B, putative
TGME49_050810	1	20	ribosomal protein L35, putative

TGGT1_088660	3	9	TPR domain-containing protein, putative
TGME49_112380	3	9	TPR domain-containing protein
TGME49_016880	57	9934	receptor for activated C kinase, RACK protein, putative
TGGT1_098250	21	3235	receptor for activated C kinase, RACK protein, putative
TGME49_001380	20	805	chorismate synthase, putative
TGGT1_020040	17	751	chorismate synthase, putative
TGGT1_114880	2	7	tRNA splicing 2' phosphotransferase, putative
TGME49_032640	1	1	tRNA splicing 2' phosphotransferase 1, putative
TGME49_011700	13	166	SAC3/GANP family protein
TGGT1_123180	8	145	80 kD MCM3-associated protein, putative
TGME49_078780	8	480	hypothetical protein
TGGT1_103680	4	30	conserved hypothetical protein
TGME49_093440	28	2037	hypothetical protein
TGGT1_073820	18	1014	conserved hypothetical protein
TGME49_095460	2	5	got1-like family domain-containing protein
TGME49_106310	5	7	structural maintenance of chromosome domain-containing protein
TGGT1_041030	1	1	structural maintenance of chromosome domain-containing protein
TGME49_057490	11	264	prefoldin subunit 3, putative
TGGT1_013210	7	166	prefoldin subunit, putative
TGME49_114400	52	2576	branched-chain alpha-keto acid dehydrogenase E1 component beta chain, putative
TGGT1_090980	8	204	branched-chain alpha-keto acid dehydrogenase, putative
TGME49_058070	19	611	signal recognition particle, putative
TGGT1_012140	15	436	conserved hypothetical protein
TGGT1_068950	61	7843	helicase, putative
TGME49_036970	61	7072	SNF2 family N-terminal domain containing protein
TGGT1_018650	3	31	conserved hypothetical protein
TGME49_002430	3	31	hypothetical protein
TGGT1_035110	8	89	conserved hypothetical protein
TGME49_088290	7	74	hypothetical protein
TGGT1_003600	1	2	glycyl-tRNA synthetase, putative
TGME49_054200	1	2	glycyl-tRNA synthetase, putative
TGME49_019750	19	715	cytochrome c, putative
TGGT1_030750	8	194	conserved hypothetical protein
TGGT1_086720	1	2	conserved hypothetical protein
TGME49_110500	1	2	hypothetical protein, conserved
TGME49_031600	60	4711	importin, putative
TGGT1_116070	36	4081	importin beta-4, putative
TGGT1_003350	2	17	conserved hypothetical protein
TGME49_054070	2	17	hypothetical protein
TGME49_063040	17	1011	40S ribosomal protein S16, putative
TGGT1_006940	8	384	40S ribosomal protein S16, putative
TGME49_092120	19	943	MORN repeat-containing protein

TGGT1_031450	14	904	MORN repeat-containing protein, putative
TGME49_071930	16	1803	hypothetical protein, conserved
TGGT1_110940	13	943	conserved hypothetical protein
TGME49_118460	36	730	cation-transporting ATPase, putative
TGGT1_122660	12	263	cation-transporting ATPase, putative
TGGT1_116520	16	441	arsenical pump-driving ATPase, putative
TGME49_031190	9	262	arsenical pump-driving ATPase, putative
TGME49_066450	2	2	lysine decarboxylase domain-containing protein
TGME49_067030	1	25	helicase, putative
TGME49_085870	35	8943	SRS20A
TGGT1_037270	11	423	SRS20A
TGME49_039260	27	8805	histone H4, putative
TGGT1_051710	13	722	histone H4, putative
TGGT1_068820	8	136	conserved hypothetical protein
TGME49_037090	7	104	hypothetical protein
TGGT1_062090	18	1019	AT hook motif-containing protein, putative
TGME49_006640	11	497	AT hook motif-containing protein
TGME49_002040	3	5	hypothetical protein, conserved
TGME49_110400	5	32	glycogenin-1, putative
TGGT1_086620	4	28	glycogenin-1, putative
TGGT1_083670	18	2022	conserved hypothetical protein
TGME49_028200	18	2022	vacuolar (H+)-ATPase G subunit domain-containing protein
TGGT1_011390	10	1498	conserved hypothetical protein
TGME49_058780	8	1239	OTU-like cysteine protease domain-containing protein
TGGT1_080620	4	33	bis(5'-nucleosyl)-tetraphosphatase, putative
TGME49_025420	4	33	bis(5'-nucleosyl)-tetraphosphatase, putative
TGME49_010980	8	74	hypothetical protein, conserved
TGGT1_123980	7	65	conserved hypothetical protein
TGME49_056060	24	539	nucleosome assembly protein, putative
TGGT1_014170	16	525	nucleosome assembly protein, putative
TGME49_115230	3	4	hypothetical protein
TGGT1_092660	2	2	conserved hypothetical protein
TGME49_062960	15	276	hypothetical protein
TGGT1_007020	12	226	conserved hypothetical protein
TGGT1_039790	15	359	conserved hypothetical protein
TGME49_012960	15	293	hypothetical protein, conserved
TGGT1_059570	2	9	SRS19B (= SRS 8)
TGME49_101150	2	9	SRS19B (= SRS 8)
TGME49_097880	16	1661	hypothetical protein
TGGT1_096230	8	321	dense-granule antigen DG32, putative
TGGT1_022260	18	971	conserved hypothetical protein
TGME49_009220	13	400	hypothetical protein
TGGT1_014890	2	6	conserved hypothetical protein
TGME49_020200	1	2	hypothetical protein, conserved
TGME49_034510	6	43	acyl-CoA-binding protein, putative

TGGT1_071380	4	16	acyl-CoA-binding protein, putative
TGGT1_034010	1	4	50S ribosomal protein L12, putative
TGME49_089230	1	4	50s ribosomal protein L12, putative
TGME49_105860	54	4307	CAM kinase, CDPK family TgCDPK1_2 (TGTPK4)
TGGT1_041610	23	934	calmodulin-domain protein kinase, putative
TGGT1_102990	1	4	RNA processing factor, putative
TGME49_051600	1	4	RNA processing factor, putative
TGME49_019130	5	24	glutathione reductase, putative
TGGT1_030090	3	20	glutathione reductase, putative
TGME49_008530	22	517	nicotinate phosphoribosyltransferase, putative
TGGT1_021560	16	512	conserved hypothetical protein
TGGT1_080100	1	7	rRNA processing protein, putative
TGGT1_034040	1	1	conserved hypothetical protein
TGME49_089200	1	1	hypothetical protein, conserved
TGME49_097240	8	91	hypothetical protein, conserved
TGGT1_033160	13	168	4-nitrophenylphosphatase, putative
TGME49_089980	13	168	4-nitrophenylphosphatase, putative
TGGT1_077990	4	24	conserved hypothetical protein
TGME49_023630	2	10	hypothetical protein, conserved
TGME49_095700	71	8615	UBA/TS-N domain-containing protein
TGGT1_054320	60	6914	ubiquitin-associated domain-containing protein, putative
TGGT1_117580	12	511	conserved hypothetical protein
TGME49_030390	8	291	hypothetical protein
TGGT1_108640	1	1	aldo-keto reductase, putative
TGME49_070100	1	1	hypothetical protein
TGME49_061460	11	163	structure specific recognition protein I, putative
TGGT1_008780	10	153	structure specific recognition protein, putative
TGGT1_019360	1	8	conserved hypothetical protein
TGME49_001860	1	8	hypothetical protein
TGGT1_025770	4	22	hypothetical protein
TGME49_047230	4	22	caltractin (centrin), putative
TGME49_045450	7	363	hypothetical protein
TGGT1_027610	4	47	conserved hypothetical protein
TGME49_107810	31	3145	hypothetical protein
TGGT1_042990	20	3402	conserved hypothetical protein
TGME49_004370	3	16	hypothetical protein
TGGT1_103570	1	1	collagen alpha chain, putative
TGME49_078800	1	1	hypothetical protein
TGGT1_123070	2	7	conserved hypothetical protein
TGME49_118170	2	7	iron-sulfur cluster assembly accessory domain containing protein
TGME49_045730	33	2685	phosphatidylinositol-4-phosphate 5-Kinase, putative
TGGT1_027340	31	3570	phosphatidylinositol-4-phosphate 5-kinase, putative
TGME49_055250	20	449	NOL1/NOP2/sun family domain-containing protein

TGGT1_076670	17	331	conserved hypothetical protein
TGGT1_011700	5	45	conserved hypothetical protein
TGME49_058500	5	45	hypothetical protein
TGGT1_124870	8	330	sentrin/sumo-specific protease, putative
TGME49_014470	7	286	ulp1 protease family, C-terminal catalytic domain-containing protein
TGME49_002620	8	156	hypothetical protein
TGGT1_018350	6	150	hypothetical protein
TGGT1_090620	2	4	nep1, putative
TGME49_113980	1	1	nucleolar essential protein 1, putative
TGME49_057060	8	53	translation initiation factor SUI1 domain-containing protein
TGGT1_013660	3	24	translation initiation factor SUI1 domain-containing protein, putative
TGGT1_121550	1	2	conserved hypothetical protein
TGME49_119740	1	2	hypothetical protein
TGME49_042760	1	1	hypothetical protein
TGGT1_087590	4	28	conserved hypothetical protein
TGME49_111260	4	28	myosin light chain TgMLC1, putative
TGME49_072690	12	123	hypothetical protein
TGGT1_111990	6	50	conserved hypothetical protein
TGGT1_062100	2	20	conserved hypothetical protein
TGME49_006630	2	20	hypothetical protein
TGME49_063500	16	297	vacuolar protein sorting 26, putative
TGGT1_043620	13	272	vacuolar sorting protein, putative
TGME49_037890	16	400	CAM kinase, CDPK family
TGGT1_058370	8	284	calmodulin-domain protein kinase, putative
TGME49_092220	26	1096	26S proteasome non-ATPase regulatory subunit, putative
TGGT1_031340	17	793	26S proteasome non-ATPase regulatory subunit, putative
TGGT1_014880	9	122	conserved hypothetical protein
TGME49_020190	8	121	hypothetical protein
TGME49_115720	9	93	smg-4/UPF3 family domain protein
TGME49_121590	2	70	hypothetical protein
TGGT1_064630	1	2	conserved hypothetical protein
TGGT1_118080	13	175	autophagy protein, putative
TGME49_029690	13	136	thiF family domain-containing protein
TGME49_021310	10	34	aminopeptidase N, putative
TGGT1_006070	10	210	conserved hypothetical protein
TGME49_081380	6	138	hypothetical protein
TGME49_023000	8	225	dynein light chain, putative
TGGT1_066540	6	185	dynein light chain, putative
TGGT1_117820	8	135	acetyltransferase domain-containing protein, putative
TGME49_030060	8	135	acetyltransferase domain-containing protein
TGME49_084530	2	4	hypothetical protein
TGGT1_036350	1	3	conserved hypothetical protein
TGME49_055660	4	57	hypothetical protein

TGGT1_077180	3	30	conserved hypothetical protein
TGGT1_099400	1	1	BTB/POZ domain-containing protein / kelch motif domain-containing protein, putative
TGME49_098600	1	1	leucine zipper-like transcriptional regulator
TGGT1_002020	17	574	actin, putative
TGME49_053040	16	458	actin-like family protein ARP4a, putative
TGME49_055410	4	113	hypothetical protein
TGGT1_076830	2	60	conserved hypothetical protein
TGGT1_011640	1	1	conserved hypothetical protein
TGME49_118530	3	5	hypothetical protein
TGGT1_121140	2	20	GTP-binding protein, putative
TGME49_094280	12	103	E3 ubiquitin-protein ligase, putative
TGGT1_074570	7	89	ubiquitin-transferase domain-containing protein, putative
TGME49_092920	60	6828	heat shock protein 90, putative
TGGT1_073160	40	5100	heat shock protein, putative
TGME49_066800	9	330	integral membrane protein, putative
TGGT1_076170	2	14	integral membrane protein, putative
TGGT1_108130	18	427	RNA polymerase II phosphatase, putative
TGME49_069700	18	392	NLI interacting factor-like phosphatase domain-containing protein
TGGT1_114760	5	44	conserved hypothetical protein
TGME49_032760	5	44	protein phosphatase inhibitor 2, putative
TGGT1_026970	19	878	conserved hypothetical protein
TGME49_046090	19	447	hypothetical protein
TGGT1_048600	11	180	conserved hypothetical protein
TGME49_042000	11	180	hypothetical protein
TGME49_027910	23	920	hypothetical protein
TGGT1_083210	4	36	conserved hypothetical protein
TGGT1_015210	4	21	conserved hypothetical protein
TGME49_020490	3	13	hypothetical protein
TGME49_076170	35	6661	phosphatidylinositol 3- and 4-kinase domain-containing protein
TGGT1_000370	33	6978	phosphatidylinositol 3- and 4-kinase domain-containing protein, putative
TGGT1_106400	27	1502	conserved hypothetical protein
TGME49_068330	21	1014	hypothetical protein
TGGT1_021570	10	203	ATP-dependent RNA helicase, putative
TGME49_008540	7	74	ATP-dependent RNA helicase, putative
TGGT1_088860	11	409	mitogen-activated protein kinase, putative
TGME49_112570	10	360	CMGC kinase, MAPK family (ERK) TgMAPK-1
TGME49_003990	21	432	hypothetical protein, conserved
TGGT1_016710	9	256	conserved hypothetical protein
TGGT1_029390	7	72	NBP2B protein, putative
TGME49_018560	5	50	acetyl-coA carboxylase
TGGT1_013810	2	6	conserved hypothetical protein
TGME49_056940	1	2	hypothetical protein

TGME49_049370	2	2	translation initiation factor SUI1, putative
TGME49_115690	23	329	DnaJ domain-containing protein
TGGT1_093370	12	276	DnaJ domain-containing protein, putative
TGGT1_036390	2	7	conserved hypothetical protein
TGME49_084570	1	2	hypothetical protein
TGGT1_069720	1	1	conserved hypothetical protein
TGME49_036030	1	1	hypothetical protein, conserved
TGME49_002080	2	2	hypothetical protein
TGGT1_019130	1	1	conserved hypothetical protein
TGGT1_077490	3	9	conserved hypothetical protein
TGME49_023420	3	9	DnaJ domain-containing protein
TGME49_121520	20	512	P23 co-chaperone, putative
TGGT1_064560	9	315	HSP90 co-chaperone, putative
TGGT1_109810	8	76	conserved hypothetical protein
TGME49_070960	7	58	hypothetical protein
TGME49_113150	15	100	hypothetical protein
TGGT1_089440	3	14	conserved hypothetical protein
TGGT1_063520	20	2390	hypothetical protein
TGME49_005380	19	1785	fructose-1,6-bisphosphatase, putative
TGME49_065180	19	444	hypothetical protein
TGGT1_058050	16	432	conserved hypothetical protein
TGGT1_065830	9	165	ribosome biogenesis protein brix, putative
TGME49_022000	6	91	ribosome biogenesis protein Brix, putative
TGGT1_092710	13	1030	conserved hypothetical protein
TGME49_115280	13	802	hypothetical protein, conserved
TGME49_039460	1	1	hypothetical protein, conserved
TGME49_030370	4	8	hypothetical protein, conserved
TGME49_008440	5	85	hypothetical protein, conserved
TGGT1_021460	1	5	conserved hypothetical protein
TGME49_049230	1	1	hypothetical protein
TGME49_039250	29	2815	diacylglycerol kinase, putative
TGGT1_051720	27	3106	diacylglycerol kinase, putative
TGGT1_104270	5	38	WD-40 repeat-containing protein, putative
TGME49_078280	2	12	WD-40 repeat-containing protein
TGME49_073090	113	29101	cell division protein 48, putative
TGGT1_112260	52	13057	cell division protein, putative
TGGT1_101290	12	210	histone deacetylase, putative
TGME49_049620	9	146	histone deacetylase, putative
TGME49_011030	31	1785	hypothetical protein
TGGT1_123900	14	836	conserved hypothetical protein
TGGT1_113270	1	4	zinc finger protein, putative
TGME49_074090	1	4	tim10/DDP zinc finger domain-containing protein
TGME49_107570	17	273	glycerol-3-phosphate dehydrogenase, putative
TGGT1_054860	11	213	glycerol-3-phosphate dehydrogenase, putative
TGME49_009290	7	717	40S ribosomal protein S28, putative

TGGT1_022330	6	145	conserved hypothetical protein
TGME49_085510	15	1024	hypothetical protein
TGGT1_037010	8	270	conserved hypothetical protein
TGGT1_054740	3	15	conserved hypothetical protein
TGME49_095320	3	15	hypothetical protein, conserved
TGME49_040510	18	444	hypothetical protein
TGGT1_050360	13	325	conserved hypothetical protein
TGME49_054770	4	63	serine/threonine protein phosphatase, putative
TGGT1_004320	3	30	serine/threonine protein phosphatase, putative
TGGT1_103800	2	6	conserved hypothetical protein
TGME49_078770	1	2	hypothetical protein
TGGT1_047130	17	677	conserved hypothetical protein
TGME49_042780	15	514	hypothetical protein
TGME49_032410	39	19940	hypothetical protein
TGGT1_115220	15	1073	conserved hypothetical protein
TGME49_051880	15	528	tyrosyl-tRNA synthetase, putative
TGGT1_103260	12	504	tyrosyl-tRNA synthetase, putative
TGGT1_018170	4	25	conserved hypothetical protein
TGME49_002780	3	13	Rhoptry kinase family protein ROP25
TGME49_057680	35	59278	myosin light chain TgMLC1
TGGT1_013010	11	1850	myosin light chain, putative
TGME49_056020	8	108	hypothetical protein
TGME49_089320	14	265	Casein kinase one (CK1) TgCK1b
TGGT1_033900	6	104	casein kinase I epsilon, putative
TGME49_075480	24	910	myb-like DNA-binding domain-containing protein
TGGT1_068040	20	870	cell division control protein, putative
TGGT1_010040	9	132	conserved hypothetical protein
TGME49_060190	9	119	microneme protein, putative
TGME49_111310	36	1139	serine/threonine protein phosphatase, putative
TGGT1_087650	17	699	serine/threonine protein phosphatase, putative
TGME49_016540	7	126	dolichyl-phosphate beta-glucosyltransferase, putative
TGGT1_098710	5	122	dolichyl-phosphate beta-glucosyltransferase, putative
TGME49_014780	4	14	bis(5'-nucleosyl)-tetraphosphatase (asymmetrical), putative
TGGT1_125210	1	6	Bis(5'-nucleosyl)-tetraphosphatase (asymmetrical), putative
TGME49_111230	353	162575	hypothetical protein, conserved
TGGT1_087540	106	60712	hypothetical protein
TGME49_023790	26	447	hypothetical protein, conserved
TGGT1_078150	15	437	conserved hypothetical protein
TGGT1_010590	5	59	ATP-dependent nucleolar RNA helicase, putative
TGME49_059220	4	58	ATP-dependent nucleolar RNA helicase, putative
TGGT1_002940	1	2	GTP cyclohydrolase I, putative
TGME49_053780	1	2	GTP cyclohydrolase I, putative
TGME49_027790	9	33	glycosyl transferase, group 1 domain containing protein

TGGT1_083090	2	8	alpha-1,3-mannosyltransferase, putative
TGME49_055700	35	1852	hypothetical protein
TGGT1_077210	25	1838	conserved hypothetical protein
TGGT1_005900	2	14	hypothetical protein
TGME49_081450	2	14	CAM kinase, CDPK family
TGME49_024090	14	172	3-hydroxyisobutyryl-CoA hydrolase, mitochondrial precursor, putative
TGGT1_078560	11	183	conserved hypothetical protein
TGGT1_034060	1	2	thioredoxin, putative
TGME49_089180	1	2	hypothetical protein, conserved
TGME49_063470	22	372	ubiquitin carboxyl-terminal hydrolase isozyme, putative
TGGT1_043590	8	268	ubiquitin carboxyl-terminal hydrolase isozyme, putative
TGME49_076930	7	75	hypothetical protein
TGGT1_105390	6	71	conserved hypothetical protein
TGGT1_090550	24	2356	RNA recognition motif 2 domain-containing protein, putative
TGME49_113910	24	2229	RNA recognition motif 2 domain-containing protein
TGME49_019690	6	692	hypothetical protein
TGGT1_030680	6	680	conserved hypothetical protein
TGME49_019550	25	1195	dihydrolipoamide succinyltransferase component of 2-oxoglutaratedehydrogenase complex, putative
TGGT1_030440	14	750	dihydrolipoamide succinyltransferase component of 2-oxoglutaratedehydrogenase, putative
TGGT1_045200	2	3	tRNA pseudouridine synthase D, putative
TGME49_044630	2	3	hypothetical protein
TGME49_111680	2	25	FUN14 family domain-containing protein
TGGT1_087930	2	16	conserved hypothetical protein
TGGT1_082480	5	164	conserved hypothetical protein
TGME49_026990	4	122	hypothetical protein
TGME49_005470	89	54781	elongation factor 2, putative
TGGT1_063430	52	25528	elongation factor, putative
TGGT1_107030	4	52	N6-adenosine-methyltransferase subunit, putative
TGME49_068840	4	52	mRNA (2'-O-methyladenosine-N(6)-)-methyltransferase, putative
TGME49_115750	8	269	hypothetical protein
TGGT1_093440	7	265	conserved hypothetical protein
TGME49_020160	4	13	hypothetical protein
TGGT1_014760	3	9	conserved hypothetical protein
TGME49_086560	7	53	u6 snRNA-associated Sm-like protein, putative
TGGT1_037980	3	33	conserved hypothetical protein
TGME49_114040	5	47	cytochrome c heme lyase, putative
TGGT1_090680	5	45	cytochrome C heme lyase, putative
TGME49_094360	25	946	ubiquitin carboxyl-terminal hydrolase, putative
TGGT1_074660	20	876	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_014440	22	1796	4'-phosphopantetheinyl transferase, putative

TGGT1_124840	20	1864	4'-phosphopantetheinyl transferase, putative
TGME49_009890	13	283	hypothetical protein, conserved
TGGT1_022940	11	267	conserved hypothetical protein
TGME49_089680	22	7093	Ras family domain-containing protein
TGGT1_033530	8	173	rab11, putative
TGME49_058010	15	912	CAM kinase
TGGT1_012190	14	444	protein kinase, putative
TGGT1_113330	1	1	conserved hypothetical protein
TGME49_074150	1	1	hypothetical protein, conserved
TGME49_033230	5	144	RNA binding motif-containing protein
TGGT1_114260	5	143	RNA binding motif-containing protein, putative
TGME49_019310	72	12006	heat shock protein, putative
TGGT1_030280	45	8431	heat shock protein, putative
TGGT1_118240	6	42	conserved hypothetical protein
TGME49_029430	6	42	hypothetical protein, conserved
TGGT1_101550	15	403	centrosomal protein, putative
TGME49_049860	15	403	centrosomal protein, putative
TGME49_020570	4	100	hypothetical protein
TGGT1_015310	3	88	conserved hypothetical protein
TGME49_113070	15	258	hypothetical protein
TGGT1_089280	6	93	conserved hypothetical protein
TGME49_093510	13	357	poly(ADP)-ribose polymerase domain-containing protein
TGGT1_073890	10	315	conserved hypothetical protein
TGGT1_039910	7	83	conserved hypothetical protein
TGME49_012850	4	34	hypothetical protein
TGGT1_100610	3	23	zinc knuckle domain-containing protein, putative
TGME49_048940	2	16	zinc knuckle domain-containing protein
TGME49_077550	3	9	uvrD/REP helicase domain-containing protein
TGME49_080600	79	7363	histidyl-tRNA synthetase, putative
TGGT1_072710	42	6010	histidyl-tRNA synthetase, putative
TGGT1_037420	3	26	conserved hypothetical protein
TGME49_086020	3	26	hypothetical protein
TGGT1_010300	3	18	diacylglycerol kinase, putative
TGME49_059830	1	7	diacylglycerol kinase, putative
TGME49_111430	36	850	fibrillarin, putative
TGGT1_087780	10	427	fibrillarin, putative
TGME49_025200	18	156	trichohyalin, putative
TGGT1_080300	11	109	trichohyalin, putative
TGGT1_037890	5	84	cAMP-dependent protein kinase catalytic subunit, putative
TGME49_086470	5	84	AGC kinase
TGME49_110080	64	14492	long-chain-fatty-acid-CoA ligase, putative
TGGT1_086220	35	3147	long-chain-fatty-acid--CoA ligase, putative
TGME49_042720	3	5	hypothetical protein
TGGT1_047190	1	1	conserved hypothetical protein
TGME49_025330	11	116	hypothetical protein

TGGT1_080530	7	117	conserved hypothetical protein
TGGT1_047410	1	7	conserved hypothetical protein
TGME49_042410	1	7	hypothetical protein
TGME49_115620	14	178	vacuolar ATP synthase subunit C, putative
TGGT1_093300	9	139	vacuolar ATP synthase subunit C, putative
TGME49_069330	4	25	hypothetical protein, conserved
TGGT1_107630	4	24	conserved hypothetical protein
TGME49_054030	8	131	hypothetical protein, conserved
TGGT1_003310	2	16	conserved hypothetical protein
TGME49_059200	45	16808	Na+/H+ antiporter, putative
TGGT1_010610	43	19086	conserved hypothetical protein
TGME49_014790	31	2130	mitochondrial glycoprotein domain-containing protein
TGGT1_125220	7	537	conserved hypothetical protein
TGGT1_114130	5	120	nuclear transport factor, putative
TGME49_033350	5	120	nuclear transport factor 2, putative
TGME49_012410	1	4	dense granule protein GRA11, putative
TGME49_037800	1	3	hypothetical protein
TGGT1_009960	1	1	cyclin H, putative
TGME49_060250	1	1	cyclin H, putative
TGME49_044320	10	789	60S ribosomal protein L24, putative
TGGT1_045620	7	123	hypothetical protein
TGGT1_065500	19	1320	hypothetical protein
TGME49_021630	17	960	ppg3, putative
TGGT1_078830	12	442	proteasome activator subunit, putative
TGME49_024330	9	253	proteasome activator subunit, putative
TGGT1_045560	15	372	conserved hypothetical protein
TGME49_044380	14	327	hypothetical protein, conserved
TGME49_054880	50	6666	alpha-galactosidase A, putative
TGGT1_004450	16	688	alpha-galactosidase A, putative
TGME49_002370	25	2998	TCP-1/cpn60 family chaperonin, putative
TGGT1_018740	21	2748	chaperonin containing t-complex protein 1, epsilon subunit, tcpe, putative
TGME49_058360	27	1130	hypothetical protein
TGGT1_011850	14	367	conserved hypothetical protein
TGGT1_104070	5	41	conserved hypothetical protein
TGME49_078490	5	41	hypothetical protein, conserved
TGGT1_066110	14	371	conserved hypothetical protein
TGME49_022300	14	229	hypothetical protein
TGME49_054520	15	327	hypothetical protein
TGGT1_004090	10	247	conserved hypothetical protein
TGGT1_117660	9	77	conserved hypothetical protein
TGME49_030220	9	65	hypothetical protein
TGME49_090220	2	2	hypothetical protein, conserved
TGGT1_065290	4	52	latency associated antigen, putative
TGME49_021440	3	33	hypothetical protein

TGME49_073970	3	4	corA-like Mg2+ transporter domain-containing protein
TGGT1_113160	2	3	CorA-like Mg2+ transporter domain-containing protein
TGGT1_039980	3	57	conserved hypothetical protein
TGME49_012800	3	57	ppg3, putative
TGGT1_013610	6	35	transcription factor BTF3, putative
TGME49_057100	6	35	transcription factor BTF3, putative
TGGT1_066400	2	6	conserved hypothetical protein
TGME49_022880	2	6	hypothetical protein
TGME49_113990	10	80	hypothetical protein
TGGT1_090640	8	71	conserved hypothetical protein
TGGT1_025940	14	243	conserved hypothetical protein
TGME49_047060	13	218	hypothetical protein
TGGT1_126200	1	1	conserved hypothetical protein
TGME49_015360	1	1	hypothetical protein
TGME49_058450	13	382	hypothetical protein, conserved
TGGT1_011760	3	32	conserved hypothetical protein
TGGT1_051460	37	4192	RNA recognition motif-containing protein, putative
TGME49_039410	35	2864	RRM domain-containing protein
TGME49_025780	3	9	hypothetical protein
TGGT1_090280	6	85	conserved hypothetical protein
TGME49_113760	3	57	hypothetical protein
TGGT1_050220	3	16	conserved hypothetical protein
TGME49_040660	3	9	hypothetical protein
TGGT1_081050	32	2556	PWWP domain-containing protein, putative
TGME49_025890	32	2358	PWWP domain-containing protein
TGME49_020500	19	591	UBX domain-containing protein
TGGT1_015230	13	600	UBX domain-containing protein, putative
TGGT1_042970	2	6	conserved hypothetical protein
TGME49_107830	2	6	hypothetical protein
TGME49_014270	47	3576	translation initiation factor IF-2, putative
TGGT1_124680	35	4443	translation initiation factor IF-2, putative
TGME49_044050	4	8	hypothetical protein
TGGT1_045900	2	6	conserved hypothetical protein
TGGT1_021940	75	15555	ubiquitin-protein ligase, putative
TGME49_009000	64	11058	HECT-domain (ubiquitin-transferase) containing protein
TGME49_043500	1	1	CAM kinase, SNF1/AMK1 family ToxPK1
TGME49_029420	11	212	cytochrome c, putative
TGGT1_118250	6	164	cytochrome C, putative
TGME49_007620	61	7717	pyridine nucleotide-disulphide oxidoreductase domain-containing protein
TGGT1_020510	31	2666	pyridine nucleotide-disulphide oxidoreductase domain-containing protein, putative
TGME49_113020	8	292	sulfate transporter, putative
TGGT1_089240	4	139	sulfate transporter, putative
TGGT1_073690	1	2	conserved hypothetical protein

TGME49_093330	1	2	hypothetical protein
TGME49_033500	17	342	triosephosphate isomerase, putative
TGGT1_113950	6	79	triosephosphate isomerase, putative
TGGT1_084330	15	611	protein kinase-PH domain-containing protein, putative
TGME49_028750	15	521	CAM kinase, RAD family
TGME49_106670	19	1621	hypothetical protein
TGGT1_040660	8	86	conserved hypothetical protein
TGME49_025340	6	55	hypothetical protein, conserved
TGGT1_080540	5	73	conserved hypothetical protein
TGME49_110630	16	422	hypothetical protein
TGGT1_086860	13	310	conserved hypothetical protein
TGGT1_033200	7	73	uroporphyrinogen decarboxylase, putative
TGME49_089940	7	57	uroporphyrinogen decarboxylase, putative
TGME49_037000	8	70	hypothetical protein, conserved
TGGT1_068910	5	41	conserved hypothetical protein
TGME49_062150	28	916	kelch motif domain-containing protein
TGGT1_007960	18	604	kelch motif domain-containing protein, putative
TGGT1_118130	40	5428	conserved hypothetical protein
TGME49_029640	34	4222	hypothetical protein, conserved
TGGT1_018760	3	16	conserved hypothetical protein
TGME49_002360	3	16	hypothetical protein
TGME49_105050	14	1588	calmodulin, putative
TGGT1_042450	6	298	calmodulin, putative
TGME49_024600	5	24	GTP-binding protein, putative
TGGT1_079320	4	20	GTP-binding protein, putative
TGME49_065470	5	29	hypothetical protein, conserved
TGGT1_057520	3	21	hypothetical protein
TGME49_018520	29	4864	microneme protein MIC6
TGGT1_029350	15	1151	microneme protein, putative
TGGT1_004430	4	8	hypothetical protein
TGME49_054860	2	2	hypothetical protein, conserved
TGME49_047610	6	74	snRNP protein Lsm5, putative
TGGT1_025200	3	37	snRNP protein Lsm5, putative
TGME49_091640	19	273	aspartate carbamoyltransferase
TGGT1_031800	8	192	aspartate carbamoyltransferase, putative
TGGT1_038320	2	4	conserved hypothetical protein
TGME49_087220	2	4	hypothetical protein, conserved
TGME49_081520	6	15	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_005840	3	5	zinc finger (C3HC4 RING finger) protein
TGGT1_098840	5	74	conserved hypothetical protein
TGME49_016410	5	74	hypothetical protein, conserved
TGME49_053700	39	5781	transporter, major facilitator family domain containing protein
TGGT1_002870	36	5534	conserved hypothetical protein
TGME49_072030	3	164	hypothetical protein

TGGT1_111150	3	163	conserved hypothetical protein
TGME49_013670	7	36	hypothetical protein
TGGT1_038840	4	26	conserved hypothetical protein
TGME49_063000	21	424	hypothetical protein, conserved
TGGT1_006970	19	572	hypothetical protein
TGME49_016010	28	208	50S ribosomal protein L24, putative
TGME49_008730	1	1	hypothetical protein
TGGT1_046520	1	4	hypothetical protein
TGME49_043340	1	4	Atypical MEK-related kinase (incomplete catalytic triad)
TGGT1_095840	5	80	conserved hypothetical protein
TGME49_099670	4	56	hypothetical protein
TGGT1_014180	6	140	signal recognition particle 14 kDa protein, putative
TGME49_056050	6	111	signal recognition particle 14 kDa protein, putative
TGME49_121530	6	72	cathepsin L-like thiolproteinase, putative
TGGT1_064570	5	68	cysteine protease, putative
TGGT1_125020	21	700	conserved hypothetical protein
TGME49_014610	21	343	hypothetical protein
TGGT1_116080	6	60	SGS domain-containing protein, putative
TGME49_031590	5	40	SGS domain-containing protein
TGME49_068780	24	537	hypothetical protein
TGGT1_106960	16	549	conserved hypothetical protein
TGME49_121420	14	245	kelch motif domain-containing protein
TGGT1_064470	12	258	kelch motif domain-containing protein, putative
TGGT1_069200	1	2	zinc finger (C3HC4 RING finger) protein
TGME49_036640	1	2	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_068180	2	10	conserved hypothetical protein
TGME49_075330	2	10	hypothetical protein
TGGT1_108720	8	76	conserved hypothetical protein
TGME49_070190	5	50	hypothetical protein
TGGT1_098130	9	509	conserved hypothetical protein
TGME49_017000	4	197	hypothetical protein
TGME49_111250	16	692	hypothetical protein
TGGT1_087570	15	782	conserved hypothetical protein
TGGT1_008560	4	79	conserved hypothetical protein
TGME49_061670	3	73	hypothetical protein
TGGT1_081780	10	72	conserved hypothetical protein
TGME49_026630	8	48	hypothetical protein
TGGT1_126220	13	548	conserved hypothetical protein
TGME49_015380	12	567	hypothetical protein
TGGT1_113360	20	1111	eukariotic translation initiation factor 2B, epsilon subunit, putative
TGME49_074190	19	725	eukaryotic initiation factor 2B epsilon subunit, putative
TGME49_062490	6	55	hypothetical protein
TGGT1_007610	1	1	conserved hypothetical protein

TGGT1_006250	9	206	RNA recognition motif 2 domain-containing protein, putative
TGME49_099030	5	105	RNA recognition motif 2 domain-containing protein
TGGT1_110100	1	1	conserved hypothetical protein
TGME49_071230	1	1	hypothetical protein
TGGT1_112370	2	14	conserved hypothetical protein
TGME49_073310	2	14	hypothetical protein
TGGT1_109790	2	37	conserved hypothetical protein
TGME49_070940	2	37	hypothetical protein, conserved
TGME49_120050	38	7899	ribosomal protein L5, putative
TGGT1_121340	17	2380	ribosomal protein L5, putative
TGGT1_100450	6	41	conserved hypothetical protein
TGME49_048800	3	16	hypothetical protein
TGME49_031820	1	25	hypothetical protein, conserved
TGGT1_092380	19	1706	conserved hypothetical protein
TGME49_115080	18	1408	hypothetical protein
TGGT1_114030	7	96	conserved hypothetical protein
TGME49_033440	7	74	hypothetical protein
TGGT1_106950	3	30	dual specificity phosphatase, catalytic domain-containing protein, putative
TGME49_010710	11	76	hypothetical protein
TGGT1_124160	4	21	conserved hypothetical protein
TGME49_011240	3	8	peptidyl-tRNA hydrolase PTH2 domain-containing protein
TGGT1_123680	1	3	peptidyl-tRNA hydrolase, putative
TGME49_070880	28	672	RNA binding motif-containing protein
TGME49_048650	2	6	guanidine nucleotide exchange factor, putative
TGME49_024760	6	25	SRS40E (= SRS4)
TGME49_038390	41	5214	hypothetical protein, conserved
TGGT1_052560	39	5755	conserved hypothetical protein
TGGT1_056090	2	3	conserved hypothetical protein
TGME49_109060	2	3	ubiquitin carboxyl-terminal hydrolase domain-containing protein
TGGT1_013650	7	234	major AMPullate spidroin, putative
TGME49_057070	7	234	major ampullate spidroin 2
TGGT1_065260	6	43	conserved hypothetical protein
TGME49_021410	6	43	actin, putative
TGME49_032330	4	8	hypothetical protein
TGME49_020600	16	1514	hypothetical protein, conserved
TGGT1_015340	11	1394	conserved hypothetical protein
TGME49_031810	7	25	hypothetical protein
TGME49_024060	38	4415	thioredoxin, putative
TGGT1_078540	24	2449	thioredoxin, putative
TGME49_094610	32	4067	hypothetical protein, conserved
TGGT1_074930	26	4041	SET and MYND domain cotaining protein, putative
TGME49_116480	10	171	XRN 5'-3' exonuclease N-terminus domain-containing protein

TGGT1_094290	1	2	XRN 5'-3' exonuclease N-terminus domain-containing protein, putative
TGGT1_049720	1	7	conserved hypothetical protein
TGME49_040950	1	7	hypothetical protein, conserved
TGME49_111370	4	15	methylmalonate-semialdehyde dehydrogenase, putative
TGGT1_049680	50	13188	conserved hypothetical protein
TGME49_040980	46	11229	hypothetical protein
TGGT1_010310	13	246	conserved hypothetical protein
TGME49_059720	11	175	hypothetical protein, conserved
TGGT1_107810	2	6	adaptor complexes medium subunit domain containing protein, putative
TGME49_069480	2	6	adaptor complexes medium subunit domain containing protein
TGME49_051930	32	1047	enoyl-acyl carrier reductase
TGGT1_103310	15	566	enoyl-acyl carrier reductase, putative
TGME49_118720	8	39	proline synthetase co-transcribed protein, putative
TGGT1_122400	3	10	proline synthetase associated protein, putative
TGME49_073370	54	4427	coatomer gamma 2-subunit protein, putative
TGGT1_112440	36	3924	coatomer protein gamma 2-subunit, putative
TGME49_025550	23	556	phosphatidylserine decarboxylase proenzyme, putative
TGGT1_080780	5	19	phosphatidylserine decarboxylase proenzyme, putative
TGME49_032080	193	27029	hypothetical protein, conserved
TGGT1_115680	100	21521	conserved hypothetical protein
TGME49_050000	3	16	hypothetical protein
TGME49_104520	11	65	hypothetical protein
TGGT1_006670	9	61	conserved hypothetical protein
TGME49_037110	7	75	replication factor C small subunit, putative
TGGT1_068790	6	68	replication factor C / DNA polymerase III gamma-tau subunit, putative
TGME49_026540	7	52	protein kinase
TGGT1_081690	1	4	calmodulin-domain protein kinase, putative
TGGT1_023020	4	16	gamma-tubulin complex component, putative
TGME49_009970	4	16	gamma-tubulin complex component 4, putative
TGGT1_116360	3	15	chromatin assembly factor 1 subunit, putative
TGGT1_108230	15	430	NAD synthase and hydrolase domain-containing protein, putative
TGME49_069800	12	320	glutamine-dependent NAD(+) synthetase protein, putative
TGME49_071300	5	24	DNA-directed RNA polymerase II subunit RPB7, putative
TGGT1_110170	3	15	DNA-directed RNA polymerase II 19 kDa polypeptide, putative
TGME49_022850	12	193	hypothetical protein
TGGT1_066360	4	77	conserved hypothetical protein
TGME49_002680	60	37390	mitochondrial-processing peptidase alpha subunit, putative
TGGT1_018290	18	1270	mitochondrial processing peptidase alpha subunit, putative

TGGT1_020950	2	4	conserved hypothetical protein
TGME49_007960	1	1	hypothetical protein, conserved
TGME49_115950	34	1660	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_093640	25	1976	zinc finger (C3HC4 RING finger) protein
TGME49_054660	2	4	ankyrin repeat-containing protein
TGME49_095350	20	1010	nucleoside diphosphate kinase, putative
TGGT1_054590	9	404	nucleoside diphosphate kinase, putative
TGGT1_114930	2	3	gamma-glutamylcysteine synthetase, putative
TGME49_032590	2	3	gamma-glutamylcysteine synthetase, putative
TGGT1_043040	7	99	conserved hypothetical protein
TGME49_107760	5	76	tubulin-tyrosine ligase family protein
TGGT1_121350	4	30	rar1, putative
TGME49_120040	4	30	rar1
TGME49_093280	1	1	hypothetical protein
TGGT1_078080	1	1	conserved hypothetical protein
TGME49_023720	1	1	hypothetical protein
TGME49_015320	4	17	hypothetical protein
TGME49_118140	30	1099	hypothetical protein
TGGT1_123110	22	1063	conserved hypothetical protein
TGGT1_056580	14	224	conserved hypothetical protein
TGME49_066280	12	164	beta-adaptin-like protein, putative
TGME49_012820	39	8061	ubiquitin family domain-containing protein
TGGT1_039940	38	9312	plant ubiquilin, putative
TGGT1_112310	3	12	Radical SAM domain-containing protein, putative
TGME49_073140	3	12	radical SAM domain-containing protein
TGGT1_016300	10	202	endonuclease/exonuclease/phosphatase domain-containing protein, putative
TGME49_004410	9	164	endonuclease/exonuclease/phosphatase domain-containing protein
TGGT1_116880	3	41	conserved hypothetical protein
TGME49_030940	3	41	hypothetical protein
TGGT1_102920	1	1	conserved hypothetical protein
TGME49_051520	1	1	hypothetical protein
TGME49_043950	36	10542	prohibitin, putative
TGGT1_046010	10	423	hypothetical protein
TGGT1_090600	3	12	conserved hypothetical protein
TGME49_113960	3	12	ribosomal protein L19, putative
TGME49_110850	10	77	zinc finger MYND domain-containing protein
TGGT1_087170	4	22	MYND domain containing protein, putative
TGGT1_035880	39	2843	peptidyl-prolyl isomerase, putative
TGME49_083850	28	2554	peptidyl-prolyl isomerase, putative
TGGT1_068880	2	3	exonuclease domain-containing protein, putative
TGME49_037020	2	3	exonuclease domain-containing protein
TGME49_007450	9	75	hypothetical protein, conserved
TGGT1_020430	6	57	conserved hypothetical protein

TGME49_048770	5	38	hypothetical protein
TGGT1_100310	1	2	conserved hypothetical protein
TGME49_119550	7	138	transmembrane protein, putative
TGGT1_121740	5	128	transmembrane protein, putative
TGME49_002760	14	512	hypothetical protein
TGGT1_018200	11	487	conserved hypothetical protein
TGME49_022380	40	1648	exportin 7, putative
TGGT1_066190	22	1155	exportin, putative
TGME49_020860	29	909	DEAD/DEAH box helicase, putative
TGGT1_053690	23	798	DEAD/DEAH box helicase, putative
TGME49_104500	13	363	hypothetical protein
TGGT1_006650	3	12	conserved hypothetical protein
TGME49_073540	5	24	phosphatidylserine synthase, putative
TGGT1_112620	2	10	phosphatidylserine synthase I, putative
TGME49_028190	21	1304	eukaryotic translation initiation factor 3 subunit 5, putative
TGGT1_083660	14	996	eukaryotic translation initiation factor 3 subunit, putative
TGGT1_063260	18	618	conserved hypothetical protein
TGME49_005650	17	593	hypothetical protein
TGME49_030350	20	468	hypothetical protein, conserved
TGGT1_117630	6	53	conserved hypothetical protein
TGME49_064610	23	5385	heterogeneous nuclear ribonucleoprotein A3, putative
TGGT1_105690	16	4296	hypothetical protein
TGME49_032940	35	70714	small heat shock protein 20
TGGT1_114580	15	1116	heat-shock protein, putative
TGME49_038940	9	65	GDP-mannose 4,6 dehydratase, putative
TGGT1_052040	5	20	GDP-mannose 4,6 dehydratase, putative
TGME49_023050	15	697	40s ribosomal protein S20, putative
TGGT1_066590	9	212	40S ribosomal protein S20, putative
TGME49_104640	6	16	hypothetical protein
TGGT1_006690	2	6	conserved hypothetical protein
TGME49_071760	18	721	thioredoxin domain-containing protein
TGGT1_110550	4	28	protein disulfide isomerase, putative
TGGT1_038850	3	42	zinc finger (CCCH type) protein
TGME49_013660	3	42	zinc finger (CCCH type) protein, putative
TGME49_030450	50	3983	GMP synthase, putative
TGGT1_117510	29	3174	GMP synthase, putative
TGGT1_072610	3	8	DNA polymerase epsilon subunit B, putative
TGME49_080690	3	8	DNA polymerase epsilon subunit, putative
TGGT1_125290	29	5418	conserved hypothetical protein
TGME49_014880	25	4023	hypothetical protein
TGME49_053890	12	84	M16 family peptidase, putative
TGGT1_003060	6	48	peptidase, putative
TGGT1_037120	4	46	XPA-binding protein, putative
TGME49_085720	3	29	XPA-binding protein, putative
TGGT1_086560	8	262	conserved hypothetical protein

TGME49_110330	7	199	hypothetical protein
TGGT1_047350	29	4159	conserved hypothetical protein
TGME49_042570	29	3530	hypothetical protein
TGGT1_038530	6	72	CHCH domain-containing protein, putative
TGME49_013940	3	105	CHCH domain-containing protein
TGME49_026730	96	25142	aconitate hydratase, putative
TGGT1_082010	33	4501	aconitate hydratase, putative
TGGT1_057490	1	2	chloride channel, putative
TGME49_065500	1	2	voltage gated chloride channel domain-containing protein
TGME49_104710	25	1174	eukaryotic peptide chain release factor subunit, putative
TGGT1_006760	19	1048	eukaryotic peptide chain release factor subunit, putative
TGGT1_056330	7	149	conserved hypothetical protein
TGME49_109180	4	56	hypothetical protein
TGME49_077060	1	1	glutamic acid-rich protein, putative
TGME49_052370	1	1	hypothetical protein
TGGT1_015350	9	227	protein phosphatase 2C, putative
TGME49_020610	9	163	protein phosphatase 2C, putative
TGME49_097100	1	1	hypothetical protein
TGME49_074000	15	352	hypothetical protein
TGGT1_113180	13	424	conserved hypothetical protein
TGGT1_123160	5	43	conserved hypothetical protein
TGME49_011720	5	43	hypothetical protein
TGME49_084790	14	339	hypothetical protein
TGGT1_036610	10	262	conserved hypothetical protein
TGGT1_083900	3	12	RNA polymerase II phosphatase, putative
TGME49_028330	3	7	NLI interacting factor-like phosphatase domain-containing protein
TGME49_028760	2	7	hypothetical protein
TGGT1_084340	2	6	conserved hypothetical protein
TGGT1_103470	7	69	conserved hypothetical protein
TGME49_078890	4	33	hypothetical protein
TGME49_013610	17	448	hypothetical protein
TGGT1_038920	14	568	conserved hypothetical protein
TGGT1_039700	23	2452	conserved hypothetical protein
TGME49_013030	23	2166	hypothetical protein
TGME49_053330	2	4	Rhoptry kinase family protein, truncated (incomplete catalytic triad)
TGME49_018950	5	31	hypothetical protein
TGGT1_029920	4	27	conserved hypothetical protein
TGGT1_051180	4	26	conserved hypothetical protein
TGME49_039680	4	26	hypothetical protein, conserved
TGME49_064040	17	755	hypothetical protein, conserved
TGGT1_044280	8	156	conserved hypothetical protein
TGME49_111480	30	2079	hypothetical protein
TGGT1_087840	18	1828	conserved hypothetical protein

TGME49_090730	22	254	hypothetical protein, conserved
TGGT1_032320	6	106	conserved hypothetical protein
TGME49_026030	29	3924	AGC kinase
TGGT1_081170	5	175	cAMP-dependent protein kinase catalytic subunit, putative
TGME49_037820	8	177	hypothetical protein
TGGT1_058450	4	45	conserved hypothetical protein
TGGT1_000420	8	179	MYND domain containing protein, putative
TGME49_076120	5	118	zinc finger MYND domain-containing protein
TGME49_061750	32	1720	hypothetical protein
TGGT1_008380	18	999	conserved hypothetical protein
TGGT1_046040	15	542	DNA replication licensing factor, putative
TGME49_043920	15	542	DNA replication licensing factor, putative
TGGT1_113310	19	995	conserved hypothetical protein
TGME49_074130	16	779	hypothetical protein
TGME49_018490	1	9	hypothetical protein
TGME49_022330	2	7	hypothetical protein
TGGT1_066140	1	6	conserved hypothetical protein
TGGT1_011860	8	124	conserved hypothetical protein
TGME49_058350	8	124	hypothetical protein
TGME49_073760	142	194396	heat shock protein 70, putative
TGGT1_112840	54	15536	heat shock protein, putative
TGGT1_008610	1	4	conserved hypothetical protein
TGME49_061620	1	4	hypothetical protein
TGME49_110010	75	6106	sushi domain-containing protein / SCR repeat-containing protein
TGGT1_086050	30	3027	sushi domain-containing protein / SCR repeat-containing protein, putative
TGGT1_031240	4	33	conserved hypothetical protein
TGME49_092330	4	33	hypothetical protein
TGGT1_007970	7	41	conserved hypothetical protein
TGME49_062140	4	21	protein phosphatase methylesterase 1, putative
TGME49_049240	20	536	calmodulin
TGGT1_100820	3	36	calmodulin, putative
TGGT1_123690	5	38	hypothetical protein
TGME49_011230	5	38	translation initiation factor eIF-2B subunit alpha, putative
TGGT1_101240	5	109	conserved hypothetical protein
TGME49_049580	4	78	hypothetical protein, conserved
TGGT1_042410	2	15	cgi-121, putative
TGME49_105090	2	15	hypothetical protein
TGME49_076140	27	2151	ADP ribosylation factor 1
TGGT1_000400	15	1235	ADP-ribosylation factor, arf, putative
TGME49_064130	16	1018	hypothetical protein, conserved
TGGT1_044370	14	1014	conserved hypothetical protein
TGME49_062610	11	102	hypothetical protein
TGGT1_007490	6	66	conserved hypothetical protein

TGME49_104650	55	9186	hypothetical protein
TGGT1_006700	45	9119	conserved hypothetical protein
TGGT1_065320	4	22	conserved hypothetical protein
TGME49_021460	3	15	phosphoglycerate mutase family protein, putative
TGME49_108830	1	1	dual specificity phosphatase, catalytic domain containing protein
TGME49_015710	10	83	hypothetical protein, conserved
TGGT1_126630	2	6	hypothetical protein
TGME49_010310	8	82	hypothetical protein
TGME49_049900	68	365838	ADP/ATP carrier, putative
TGGT1_101590	25	2508	ADP/ATP carrier, putative
TGME49_114750	22	526	hypothetical protein, conserved
TGGT1_092040	13	306	conserved hypothetical protein
TGME49_086730	5	25	signal recognition particle 19 kDa protein, putative
TGGT1_038130	4	20	signal recognition particle 19 kDa protein, putative
TGME49_072570	8	68	hypothetical protein
TGME49_061370	11	73	hypothetical protein
TGGT1_008870	3	48	conserved hypothetical protein
TGME49_060440	23	1103	46 kDa FK506-binding nuclear protein, putative
TGGT1_009760	13	870	46 kDa FK506-binding nuclear protein, putative
TGGT1_118090	1	6	conserved hypothetical protein
TGME49_029680	1	6	hypothetical protein
TGGT1_099010	3	6	SMC protein, putative
TGME49_016250	3	6	SMC protein, putative
TGME49_050840	8	30	hypothetical protein
TGGT1_102400	3	21	conserved hypothetical protein
TGME49_057480	17	235	succinate-semialdehyde dehydrogenase, putative
TGGT1_013220	11	169	succinate-semialdehyde dehydrogenase, putative
TGME49_040640	15	247	Casein kinase one (CK1) TgCK1a
TGGT1_050240	6	100	casein kinase, putative
TGGT1_046560	1	7	conserved hypothetical protein
TGME49_043300	1	7	hypothetical protein
TGGT1_051980	1	4	conserved hypothetical protein
TGME49_039000	1	4	hypothetical protein
TGME49_026700	11	141	nuclease, putative
TGGT1_081870	7	127	conserved hypothetical protein
TGME49_113970	11	124	hypothetical protein
TGGT1_090610	2	8	hypothetical protein
TGGT1_050180	10	122	ubiquitin protein, putative
TGME49_040700	8	71	ubiquitin protein, putative
TGME49_016900	18	538	hypothetical protein, conserved
TGGT1_098230	14	464	conserved hypothetical protein
TGME49_097180	18	374	hypothetical protein
TGGT1_097030	5	41	conserved hypothetical protein

TGGT1_017220	1	20	conserved hypothetical protein
TGME49_003620	1	20	hypothetical protein
TGGT1_001130	27	1586	PPR repeat-containing protein, putative
TGME49_052210	24	1338	PPR repeat-containing protein
TGME49_052430	38	4483	hypothetical protein, conserved
TGGT1_001330	26	4542	conserved hypothetical protein
TGGT1_052500	1	1	conserved hypothetical protein
TGME49_038430	1	1	hypothetical protein
TGME49_019270	15	16417	hypothetical protein
TGGT1_030240	4	1443	conserved hypothetical protein
TGGT1_040850	16	946	conserved hypothetical protein
TGME49_106490	14	798	hypothetical protein
TGGT1_107200	20	2105	conserved hypothetical protein
TGME49_069010	17	1196	hypothetical protein
TGGT1_104060	7	47	conserved hypothetical protein
TGME49_078500	6	36	ftsJ-like methyltransferase domain-containing protein
TGGT1_116290	3	20	conserved hypothetical protein
TGME49_031380	3	20	DNA-directed RNA polymerase II, putative
TGME49_063220	6	47	Rhoptry kinase family protein ROP21
TGGT1_043320	4	38	conserved hypothetical protein
TGGT1_069060	2	35	conserved hypothetical protein
TGME49_036870	1	27	hypothetical protein
TGME49_015740	9	86	hypothetical protein, conserved
TGGT1_126650	8	82	WD-repeat protein, putative
TGME49_105160	15	1602	histone H2B, putative
TGGT1_042340	8	824	histone H2B, putative
TGME49_039740	33	5634	hypothetical protein
TGGT1_051100	16	1904	conserved hypothetical protein
TGME49_031150	11	106	hypothetical protein
TGGT1_116560	1	2	conserved hypothetical protein
TGGT1_026930	4	63	conserved hypothetical protein
TGME49_046120	4	63	hypothetical protein
TGGT1_006800	2	7	hypothetical protein
TGME49_104750	2	7	hypothetical protein
TGME49_119530	25	1832	splicing factor, putative
TGGT1_121760	14	1658	hypothetical protein
TGME49_061420	6	88	hypothetical protein, conserved
TGGT1_008820	5	58	conserved hypothetical protein
TGME49_091600	11	111	interferon gamma-inducible protein 30
TGGT1_031840	5	79	interferon gamma-inducible protein, putative
TGME49_036840	12	241	hypothetical protein
TGGT1_069100	11	260	zinc finger protein-containing protein p48znf, putative
TGME49_026400	5	25	lipoic acid synthase
TGGT1_081570	3	16	lipoic acid synthase, putative
TGME49_111220	16	614	hypothetical protein, conserved

TGGT1_087530	14	970	conserved hypothetical protein
TGME49_060640	18	144	autophagy protein APG9, putative
TGGT1_009540	4	27	autophagy protein APG9, putative
TGME49_093320	18	264	hypothetical protein
TGGT1_073680	11	204	conserved hypothetical protein
TGME49_081630	12	214	hydroxyacylglutathione hydrolase, putative
TGGT1_005720	8	162	hydroxyacylglutathione hydrolase, putative
TGVEG_088620	2	4	hydroxyacylglutathione hydrolase, putative
TGGT1_009640	2	9	conserved hypothetical protein
TGME49_060550	2	9	acetlytransferase, GNAT family domain containing protein
TGME49_063070	28	2293	CMGC kinase, CK2 family
TGGT1_043160	18	1748	casein kinase II alpha subunit, putative
TGGT1_001230	4	63	nuclear DNA-binding protein, putative
TGME49_052320	4	63	nuclear DNA-binding protein, putative
TGME49_051480	11	92	ATP-dependent RNA helicase, putative
TGGT1_102880	8	79	ATP-dependent RNA helicase, putative
TGGT1_025680	1	5	conserved hypothetical protein
TGME49_047320	1	5	hypothetical protein
TGGT1_084320	6	236	EF hand domain-containing protein, putative
TGME49_028740	4	97	EF hand domain-containing protein
TGGT1_011350	1	16	conserved hypothetical protein
TGME49_036130	2	21	Signal recognition particle 9 kDa protein, putative
TGGT1_069630	2	15	signal recognition particle 9 kDa protein, putative
TGME49_051810	21	3284	translation initiation factor eIF-5A, putative
TGGT1_103180	11	1691	translation initiation factor 5A2
TGME49_086080	51	3549	U5 small nuclear ribonucleoprotein, putative
TGGT1_037490	30	2427	conserved hypothetical protein
TGME49_113580	16	497	cytochrome b5-like heme/steroid binding domain-containing protein
TGGT1_089990	12	559	flavohemoprotein B5/b5r, putative
TGGT1_095020	6	124	hypothetical protein
TGME49_100270	5	85	hypothetical protein
TGME49_023140	22	588	multisynthetase complex auxiliary component p43, putative
TGGT1_066670	13	393	multisynthetase complex auxiliary component p43, putative
TGGT1_038540	4	30	exosome complex exonuclease, putative
TGME49_013930	3	20	exosome complex exonuclease, putative
TGME49_029250	7	310	60S ribosomal protein L28, putative
TGGT1_118500	5	100	60S ribosomal protein L28, putative
TGME49_111030	50	1796	hypothetical protein
TGGT1_087340	38	1538	bap28, putative
TGGT1_103970	1	6	conserved hypothetical protein
TGME49_078590	1	6	hypothetical protein
TGGT1_005870	5	56	Pdx2 protein, putative

TGME49_081490	5	56	glutamine amidotransferase, SNO family domain-containing protein
TGME49_107850	29	825	6-phosphogluconate dehydrogenase, putative
TGGT1_042950	18	745	6-phosphogluconate dehydrogenase, putative
TGGT1_108540	1	1	conserved hypothetical protein
TGME49_070000	1	1	hypothetical protein
TGGT1_021830	1	1	conserved hypothetical protein
TGME49_008800	1	1	hypothetical protein, conserved
TGGT1_046430	9	330	conserved hypothetical protein
TGME49_043430	9	258	OTU-like cysteine protease domain-containing protein
TGME49_033750	1	3	hypothetical protein
TGME49_099070	32	1568	pyruvate kinase, putative
TGGT1_006300	26	1564	pyruvate kinase, putative
TGGT1_022370	1	4	conserved hypothetical protein
TGME49_009430	1	4	U3 small nucleolar ribonucleoprotein protein IMP3, putative
TGME49_035540	23	1078	translation initiation factor 2 beta, putative
TGGT1_070240	17	1026	translation initiation factor 2 beta, putative
TGGT1_034130	21	1054	glutamic acid-rich protein, putative
TGME49_089120	16	670	glutamic acid-rich protein, putative
TGGT1_026980	17	929	conserved hypothetical protein
TGME49_046080	8	442	hypothetical protein
TGGT1_011490	3	6	conserved hypothetical protein
TGME49_058700	3	6	hypothetical protein
TGME49_093560	6	100	hypothetical protein
TGME49_095660	24	459	hypothetical protein
TGGT1_054360	13	294	ubiquitin ligase E3 alpha, putative
TGGT1_109510	2	8	hypothetical protein
TGME49_070750	2	8	hypothetical protein
TGME49_082060	6	138	hypothetical protein
TGGT1_005410	5	121	conserved hypothetical protein
TGGT1_125650	2	7	conserved hypothetical protein
TGME49_015020	1	3	hypothetical protein, conserved
TGME49_035970	34	2770	eukaryotic translation initiation factor 2 gamma subunit, putative
TGGT1_069800	19	1890	eukaryotic translation initiation factor 2 gamma subunit, putative
TGME49_073850	10	160	hypothetical protein
TGGT1_112930	8	129	conserved hypothetical protein
TGME49_035340	10	142	hypothetical protein
TGGT1_070650	6	62	conserved hypothetical protein
TGGT1_002820	4	43	DnaJ domain-containing protein, putative
TGME49_053650	4	43	DnaJ domain-containing protein
TGME49_112990	11	142	ATP-dependent RNA helicase, putative
TGGT1_089210	8	128	ATP-dependent RNA helicase, putative
TGME49_106930	16	926	proteasome subunit beta type 7, putative

TGGT1_055480	13	821	proteasome subunit beta type, putative
TGME49_058410	20	2071	hypothetical protein
TGGT1_011800	9	722	conserved hypothetical protein
TGME49_018790	6	27	elongation factor Tu GTP-binding domain-containing protein
TGGT1_029760	1	2	GTP-binding domain containing protein, putative
TGME49_120220	37	20839	ubiquinol-cytochrome c reductase domain-containing protein
TGGT1_121070	11	486	ubiquinol-cytochrome C reductase iron-sulfur subunit, putative
TGME49_037250	14	2440	hypothetical protein
TGGT1_068660	6	105	conserved hypothetical protein
TGME49_080640	35	2396	hypothetical protein
TGGT1_072670	24	2306	conserved hypothetical protein
TGGT1_108680	4	29	conserved hypothetical protein
TGME49_070150	4	20	hypothetical protein
TGGT1_004280	4	27	conserved hypothetical protein
TGME49_054730	4	27	hypothetical protein, conserved
TGGT1_102960	8	250	conserved hypothetical protein
TGME49_051560	8	250	hypothetical protein
TGGT1_039630	1	4	conserved hypothetical protein
TGME49_013100	1	4	YGGT family domain protein
TGGT1_063240	1	2	conserved hypothetical protein
TGME49_005670	1	2	SF-assemblin, putative
TGGT1_049520	1	1	DEAD/DEAH box helicase, putative
TGME49_041140	1	1	DEAD/DEAH box helicase, putative
TGGT1_032230	10	402	conserved hypothetical protein
TGME49_090910	5	126	hypothetical protein
TGME49_086710	8	59	zinc finger (C2H2 type) protein, putative
TGGT1_038110	6	48	zinc finger (C2H2 type) protein
TGME49_006470	53	4751	dihydrolipoyl dehydrogenase, putative
TGGT1_062300	22	1372	dihydrolipoyl dehydrogenase, putative
TGGT1_035100	1	1	conserved hypothetical protein
TGME49_088300	1	1	hypothetical protein, conserved
TGME49_049980	6	31	leucine rich repeat protein, putative
TGGT1_063890	6	145	conserved hypothetical protein
TGME49_005150	4	93	hypothetical protein
TGME49_080590	17	376	hypothetical protein
TGGT1_072720	13	368	conserved hypothetical protein
TGME49_026270	4	90	hypothetical protein
TGGT1_081430	2	83	small seven transmembrane domain-containing protein, putative
TGGT1_056340	12	1556	conserved hypothetical protein
TGME49_109190	11	1364	hypothetical protein
TGGT1_124960	8	253	conserved hypothetical protein
TGME49_014560	7	243	hypothetical protein
TGME49_091850	3	201	40S ribosomal protein S30, putative

TGGT1_031690	1	2	40S ribosomal protein S30, putative
TGME49_039490	61	3292	mitochondrial branched-chain alpha-keto acid dehydrogenase E1, putative
TGGT1_051380	16	1094	2-oxoisovalerate dehydrogenase, putative
TGME49_005090	3	5	hypothetical protein
TGGT1_054910	1	2	elongation factor TS, putative
TGME49_107610	1	2	elongation factor TS, putative
TGGT1_076060	2	12	cyclin, N-terminal domain-containing protein, putative
TGME49_066900	1	3	cyclin, N-terminal domain-containing protein
TGME49_042400	1	1	CAM kinase, CDPK family
TGGT1_078140	5	48	conserved hypothetical protein
TGME49_023780	5	47	hypothetical protein
TGME49_115410	8	67	SRS53F
TGGT1_092860	5	42	SRS53F
TGME49_004270	2	4	hypothetical protein
TGME49_088390	1	1	glycosyl transferase, group 2 family domain-containing protein
TGME49_016120	8	70	hypothetical protein
TGGT1_099140	6	66	conserved hypothetical protein
TGME49_106560	7	25	hypothetical protein
TGGT1_040780	4	16	conserved hypothetical protein
TGGT1_026360	6	188	conserved hypothetical protein
TGME49_046730	5	176	hypothetical protein
TGME49_012130	15	444	patatin-like phospholipase domain-containing protein
TGGT1_065340	4	648	conserved hypothetical protein
TGME49_021480	2	168	hypothetical protein
TGGT1_055880	1	2	conserved hypothetical protein
TGGT1_046410	4	22	small GTP-binding protein domain-containing protein
TGME49_043450	4	22	Ras family domain-containing protein
TGGT1_000590	4	42	conserved hypothetical protein
TGME49_075740	2	23	hypothetical protein, conserved
TGME49_093810	8	69	2-methylisocitrate lyase, putative
TGGT1_074190	5	39	carboxyvinyl-carboxyphosphonate phosphorylmutase, putative
TGGT1_010110	7	96	acetyltransferase domain-containing protein, putative
TGME49_060010	6	56	acetyltransferase domain-containing protein
TGGT1_103900	4	32	ubiquitin-conjugating enzyme E2, putative
TGME49_078670	4	26	ubiquitin-conjugating enzyme E2, putative
TGME49_094420	9	146	programmed cell death protein, putative
TGGT1_074820	8	145	programmed cell death protein, putative
TGME49_047330	14	185	hypothetical protein
TGGT1_025670	13	261	conserved hypothetical protein
TGME49_089000	17	242	hypothetical protein
TGGT1_034270	11	180	conserved hypothetical protein
TGME49_016000	59	25870	inner membrane complex protein IMC3

TGGT1_099250	15	4085	inner membrane complex protein, putative
TGGT1_047110	2	7	conserved hypothetical protein
TGME49_042800	2	7	TGF-beta-inducible nuclear protein 1, putative
TGGT1_065970	1	5	DEAD/DEAH box helicase, putative
TGME49_022140	1	5	DEAD/DEAH box helicase, putative
TGME49_116230	5	42	SAC1 phosphoinositide phosphatase, putative
TGGT1_093920	4	31	conserved hypothetical protein
TGME49_070760	16	174	hypothetical protein
TGGT1_109520	10	98	asparagine synthetase, putative
TGME49_045620	7	70	ubiquitin / ribosomal protein S27a fusion protein, putative
TGGT1_027450	4	62	ubiquitin / ribosomal protein S27A fusion protein, putative
TGGT1_123150	14	187	SET domain-containing protein, putative
TGME49_011730	14	187	SET domain-containing protein
TGGT1_076880	13	322	conserved hypothetical protein
TGME49_055450	12	287	hypothetical protein
TGME49_109120	61	21569	60s ribosomal protein L4, putative
TGGT1_056160	23	3466	60S ribosomal protein L4, putative
TGME49_093470	33	2066	hypothetical protein
TGGT1_073850	23	1145	conserved hypothetical protein
TGGT1_035540	4	54	conserved hypothetical protein
TGME49_083520	3	28	hypothetical protein
TGME49_100260	100	10488	threonyl-tRNA synthetase, putative
TGGT1_095030	46	6547	threonyl-tRNA synthetase, putative
TGME49_093630	21	1145	hypothetical protein
TGGT1_074020	18	1128	conserved hypothetical protein
TGGT1_027510	7	368	origin recognition complex subunit, putative
TGME49_045570	6	299	origin recognition complex subunit 2, putative
TGGT1_016860	6	39	conserved hypothetical protein
TGME49_003860	5	30	hypothetical protein
TGGT1_110530	26	1866	conserved hypothetical protein
TGME49_071740	25	1362	hypothetical protein
TGGT1_115270	8	151	RNA polymerase Rpb1 C-terminal repeat domain-containing protein / exonuclease domain-containing protein, putative
TGME49_032360	4	55	RNA polymerase Rpb1 C-terminal repeat domain-containing protein / exonuclease domain-containing protein
TGME49_092950	45	3479	hypothetical protein
TGGT1_073190	34	3392	conserved hypothetical protein
TGME49_119310	1	9	hypothetical protein
TGGT1_031230	7	205	zinc finger (C3HC4 RING finger) protein
TGME49_092340	6	170	zinc finger (C3HC4 RING finger) protein, putative
TGME49_054570	60	4181	hypothetical protein, conserved
TGGT1_004140	50	4302	conserved hypothetical protein
TGME49_063340	2	3	rRNA methylase, putative
TGME49_000310	7	88	hypothetical protein, conserved
TGGT1_104510	3	27	G-patch domain-containing protein, putative

TGME49_078040	2	18	G-patch domain-containing protein
TGME49_089210	7	72	prefoldin subunit, putative
TGGT1_034030	5	63	prefoldin subunit, putative
TGME49_116330	15	253	iron-containing superoxide dismutase
TGGT1_094140	2	30	hypothetical protein
TGME49_115810	31	361	tRNA nucleotidyltransferase, putative
TGGT1_093510	14	334	poly(A) polymerase, putative
TGME49_048570	4	16	M protein repeat-containing protein
TGGT1_005700	2	9	conserved hypothetical protein
TGME49_065540	2	11	protein phosphatase 2C, putative
TGME49_031850	85	13062	serine-threonine phosophatase 2C
TGGT1_115910	26	5544	serine-threonine phosophatase 2C, putative
TGME49_095420	20	765	hypothetical protein
TGGT1_054520	7	227	conserved hypothetical protein
TGGT1_027350	7	38	helicase, putative
TGME49_045720	4	14	helicase conserved C-terminal domain-containing protein
TGME49_065360	2	6	hypothetical protein
TGGT1_057750	1	2	conserved hypothetical protein
TGME49_113010	22	988	ATP-dependent RNA helicase, putative
TGGT1_089230	16	818	ATP-dependent RNA helicase, putative
TGGT1_065920	2	12	conserved hypothetical protein
TGME49_022090	2	12	hypothetical protein
TGGT1_047340	4	22	iron only hydrogenase large subunit, C-terminal domain-containing protein, putative
TGME49_042580	3	17	iron only hydrogenase large subunit, C-terminal domain-containing protein
TGGT1_047310	2	4	WD domain, G-beta repeat-containing protein, putative
TGME49_042610	2	4	WD domain, G-beta repeat-containing protein
TGGT1_120940	5	42	conserved hypothetical protein
TGME49_120450	5	42	ribosome biogenesis regulatory protein, putative
TGME49_004060	9	137	SNARE domain-containing protein
TGGT1_016640	7	128	conserved hypothetical protein
TGGT1_058430	1	1	conserved hypothetical protein
TGME49_037840	1	1	hypothetical protein
TGME49_019510	8	60	GTPase domain-containing protein
TGME49_111440	4	7	SRS50
TGGT1_086510	12	551	phosphoethanolamine cytidylyltransferase, putative
TGME49_110280	10	389	phosphoethanolamine cytidylyltransferase, putative
TGME49_038130	2	123	hypothetical protein
TGGT1_052730	2	122	conserved hypothetical protein
TGGT1_077980	4	35	membrin, putative
TGME49_023620	4	35	golgi SNARE protein, putative
TGGT1_052620	4	25	serine/threonine protein phosphatase, putative
TGME49_038230	4	25	serine/threonine protein phosphatase, putative

TGME49_025730	8	456	hypothetical protein
TGGT1_080860	3	20	conserved hypothetical protein
TGME49_090040	13	389	macrophage migration inhibitory factor, putative
TGGT1_032990	7	147	macrophage migration inhibitory factor, putative
TGGT1_118720	7	70	protein kinase, putative
TGME49_029020	7	70	CMGC kinase
TGGT1_063940	2	5	conserved hypothetical protein
TGME49_005100	2	3	hypothetical protein
TGME49_019520	27	807	arginine N-methyltransferase 1
TGGT1_030400	19	715	protein arginine N-methyltransferase, putative
TGGT1_064600	11	247	conserved hypothetical protein
TGME49_121560	11	246	hypothetical protein
TGME49_034930	9	704	SRS43
TGGT1_071180	3	11	SRS43
TGGT1_034350	13	307	CCAAT-box DNA binding protein subunit B, putative
TGME49_088930	9	202	hypothetical protein
TGGT1_057330	1	2	conserved hypothetical protein
TGME49_065650	1	2	hypothetical protein
TGGT1_043450	2	18	conserved hypothetical protein
TGME49_063330	2	18	hypothetical protein
TGME49_113060	10	390	eukaryotic translation initiation factor 2B, putative
TGGT1_089270	5	135	eukaryotic translation initiation factor 2B, putative
TGGT1_028950	6	51	SET domain containing protein, putative
TGME49_018230	4	25	SET domain containing protein
TGGT1_003360	5	24	zinc transporter, putative
TGME49_054080	4	17	zinc transporter ZIP domain-containing protein
TGGT1_010090	1	1	conserved hypothetical protein
TGME49_060030	1	1	Atypical MEK-related kinase (incomplete catalytic triad)
TGME49_090030	20	864	hypothetical protein, conserved
TGGT1_033000	5	59	conserved hypothetical protein
TGGT1_125840	1	2	cell cycle protein mesj, putative
TGGT1_099170	12	169	conserved hypothetical protein
TGME49_016080	12	169	hypothetical protein, conserved
TGGT1_042440	2	8	tubulin-specific chaperone, putative
TGME49_105060	2	7	tubulin-specific chaperone, putative
TGGT1_080960	3	30	conserved hypothetical protein
TGME49_025820	3	30	hypothetical protein
TGME49_025360	9	56	hypothetical protein
TGGT1_080560	8	59	conserved hypothetical protein
TGGT1_021950	3	12	elongation factor TS, putative
TGME49_009010	2	6	elongation factor TS, putative
TGME49_038010	13	977	ribosomal protein L23a, putative
TGGT1_052830	9	268	ribosomal protein L23A, putative
TGME49_028460	9	94	hypothetical protein

TGGT1_084030	4	22	predicted protein
TGGT1_026160	12	337	conserved hypothetical protein
TGME49_046950	11	258	hypothetical protein
TGME49_094670	32	2166	eukaryotic translation initiation factor 3, putative
TGGT1_074990	19	1475	eukaryotic translation initiation factor, putative
TGGT1_085260	2	4	conserved hypothetical protein
TGME49_109390	2	4	hypothetical protein
TGGT1_063370	15	206	RNA helicase, putative
TGME49_005540	14	157	RNA helicase, putative
TGME49_109930	7	26	alpha-N-acetylgalactosaminidase, putative
TGGT1_085870	3	12	alpha-N-acetylgalactosaminidase, putative
TGME49_062380	37	2691	elongation factor Tu, putative
TGGT1_007730	20	1482	elongation factor Tu, putative
TGGT1_072850	2	29	conserved hypothetical protein
TGME49_080460	2	29	hypothetical protein
TGGT1_037640	6	68	kelch repeat protein, putative
TGME49_086240	6	68	leucine-zipper-like transcriptional regulator 1, putative
TGME49_112090	23	1493	60S ribosomal protein L23, putative
TGGT1_088330	11	548	60S ribosomal protein L23, putative
TGGT1_097010	4	24	very long-chain acyl-CoA synthetase, putative
TGME49_097220	4	24	very long-chain acyl-CoA synthetase, putative
TGME49_112590	4	34	hypothetical protein, conserved
TGGT1_088880	2	20	conserved hypothetical protein
TGME49_055260	67	23056	apical membrane antigen 1, putative
TGGT1_076680	17	965	apical membrane antigen, putative
TGGT1_007240	9	139	conserved hypothetical protein
TGME49_062750	8	112	hypothetical protein
TGGT1_105150	9	180	dynein intermediate chain, putative
TGME49_077510	8	133	dynein intermediate chain, putative
TGME49_016140	11	109	ankyrin repeat-containing protein
TGGT1_099120	8	100	ankyrin repeat-containing protein, putative
TGGT1_103840	6	69	conserved hypothetical protein
TGME49_078730	6	69	hypothetical protein, conserved
TGGT1_066090	1	2	conserved hypothetical protein
TGME49_022280	1	2	hypothetical protein
TGGT1_083020	48	7968	conserved hypothetical protein
TGME49_027610	44	6849	hypothetical protein, conserved
TGME49_015660	2	4	hypothetical protein
TGGT1_126580	1	2	conserved hypothetical protein
TGME49_108890	57	4132	hypothetical protein, conserved
TGGT1_055920	54	4543	suppressor of ty, putative
TGGT1_042820	2	28	conserved hypothetical protein
TGGT1_007620	3	15	Roadblock/LC7 domain containing protein, putative
TGME49_062480	3	15	dynein light chain roadblock-type 2, putative
TGME49_114070	17	378	hypothetical protein

TGGT1_090710	14	286	conserved hypothetical protein
TGGT1_047490	1	1	40S ribosomal protein S29, putative
TGME49_042340	1	1	40S ribosomal protein S29, putative
TGGT1_052660	18	547	DNA-directed RNA polymerase II, putative
TGME49_038190	15	435	DNA-directed RNA polymerase II, putative
TGME49_112260	20	271	hypothetical protein
TGGT1_088520	16	338	conserved hypothetical protein
TGME49_044690	29	1848	hypothetical protein
TGGT1_045120	27	2011	hypothetical protein
TGME49_050710	24	8173	microneme protein 10
TGGT1_102150	17	1580	microneme protein, putative
TGME49_049250	18	319	60s ribosomal protein L33-A, putative
TGGT1_100830	7	156	60S ribosomal protein L33-A, putative
TGGT1_102110	9	146	conserved hypothetical protein
TGME49_050680	8	88	TBC domain-containing kinase (incomplete catalytic triad)
TGGT1_083520	55	14500	conserved hypothetical protein
TGME49_028150	54	11211	hypothetical protein
TGME49_104460	8	229	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_006600	8	220	zinc finger (C3HC4 RING finger) protein
TGGT1_032800	3	130	conserved hypothetical protein
TGME49_090240	3	77	hypothetical protein, conserved
TGME49_101180	12	274	SRS19F
TGGT1_059600	4	20	SRS19F
TGGT1_015080	1	1	conserved hypothetical protein
TGME49_020380	1	1	hypothetical protein
TGME49_115790	26	220	hypothetical protein
TGGT1_006610	2	2	oxidoreductase, putative
TGME49_104470	2	2	oxidoreductase, putative
TGGT1_065480	6	27	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_021610	3	6	ubiquitin carboxyl-terminal hydrolase, putative
TGGT1_076000	3	17	calcium/calmodulin-dependent protein kinase II, putative
TGME49_066950	3	17	Atypical MEK-related kinase (incomplete catalytic triad)
TGME49_040870	40	1833	beta adaptin protein, putative
TGGT1_049890	28	1616	beta adaptin protein, putative
TGME49_077080	23	26634	microneme TgMIC5 protein
TGGT1_105250	9	637	H4, putative
TGGT1_087400	5	60	two-pore calcium channel, putative
TGME49_111080	4	46	cation channel family domain-containing protein
TGGT1_022460	2	5	conserved hypothetical protein
TGME49_009520	2	5	hypothetical protein, conserved
TGME49_114220	5	199	hypothetical protein
TGGT1_053890	6	39	conserved hypothetical protein
TGME49_030160	15	2745	hypothetical protein
TGGT1_117730	14	1926	conserved hypothetical protein

TGME49_042070	38	11870	cAMP-dependent protein kinase regulatory subunit, putative
TGGT1_048350	32	3882	cAMP-dependent protein kinase regulatory subunit, putative
TGME49_101350	4	10	hypothetical protein
TGGT1_059760	3	9	conserved hypothetical protein
TGGT1_108580	7	108	conserved hypothetical protein
TGME49_070060	7	79	hypothetical protein
TGGT1_016990	1	7	conserved hypothetical protein
TGME49_003740	1	7	hypothetical protein
TGME49_041830	5	47	cyclophilin, putative
TGGT1_048880	3	43	cyclophilin, putative
TGME49_027620	27	19554	28 kDa antigen
TGGT1_083030	9	170	28 kDa antigen, putative
TGME49_063280	9	115	hypothetical protein, conserved
TGGT1_043400	3	15	conserved hypothetical protein
TGGT1_028110	6	84	conserved hypothetical protein
TGME49_017900	6	84	hypothetical protein, conserved
TGGT1_070170	6	43	spermatogenesis associated factor, putative
TGME49_035610	6	37	ATPase, AAA family domain-containing protein
TGME49_014130	6	52	hypothetical protein
TGGT1_124360	5	53	conserved hypothetical protein
TGME49_095610	5	55	SET domain-containing protein
TGGT1_054420	4	54	SET domain-containing protein, putative
TGME49_015350	7	256	hypothetical protein
TGGT1_126190	6	144	conserved hypothetical protein
TGME49_100030	4	24	hypothetical protein
TGGT1_095470	3	15	ubiquitin carrier protein, putative
TGGT1_089250	2	26	conserved hypothetical protein
TGME49_113040	1	15	hypothetical protein
TGME49_075860	14	206	hypothetical protein
TGGT1_000480	10	218	conserved hypothetical protein
TGGT1_005320	1	1	conserved hypothetical protein
TGME49_082150	1	1	hypothetical protein
TGME49_050830	16	195	26S proteasome regulatory particle non-ATPase subunit, putative
TGGT1_102380	5	76	NBP2B protein, putative
TGME49_034220	5	26	hypothetical protein
TGGT1_071670	4	28	conserved hypothetical protein
TGGT1_051230	7	141	cytidine deaminase, putative
TGME49_039630	5	58	cytidine deaminase, putative
TGME49_040670	2	4	ferredoxin, putative
TGME49_113670	32	2466	gamma-adaptin, putative
TGGT1_090180	29	2393	gamma-adaptin, putative
TGGT1_070100	5	26	conserved hypothetical protein
TGME49_035670	5	26	hypothetical protein
TGME49_065350	1	1	acyl-CoA thioester hydrolase, putative

TGME49_071350	15	250	folylpolyglutamate synthase, putative
TGGT1_110220	13	281	folylpolyglutamate synthase, putative
TGGT1_047010	25	2580	conserved hypothetical protein
TGME49_042890	21	1529	hypothetical protein
TGGT1_072110	18	674	conserved hypothetical protein
TGME49_012260	18	570	hypothetical protein
TGGT1_066040	28	2819	DNA repair/transcription protein met18/mms19, putative
TGME49_022230	18	1671	hypothetical protein
TGGT1_060590	2	9	geranylgeranyl transferase type II beta subunit, putative
TGME49_000370	2	9	protein farnesyltranstransferase beta subunit, putative
TGGT1_009280	14	356	DNA mismatch repair protein, putative
TGME49_061000	13	319	DNA mismatch repair protein, putative
TGGT1_107440	1	1	conserved hypothetical protein
TGME49_069240	1	1	actin-like family protein, putative
TGGT1_008080	6	67	calmodulin, putative
TGME49_062010	5	50	calmodulin, putative
TGGT1_075290	1	3	conserved hypothetical protein
TGME49_094940	1	3	hypothetical protein
TGME49_093540	2	4	hypothetical protein
TGGT1_073920	1	1	conserved hypothetical protein
TGGT1_047180	8	204	guanylate kinase, putative
TGME49_042730	4	31	guanylate kinase, putative
TGGT1_121510	5	53	MORN repeat-containing protein, putative
TGME49_119880	5	53	MORN repeat-containing protein
TGME49_078840	1	5	hypothetical protein
TGME49_119850	29	1827	hypothetical protein, conserved
TGGT1_121540	19	1495	conserved hypothetical protein
TGME49_015160	4	29	hypothetical protein
TGME49_100140	82	10131	elongation factor 1-gamma, putative
TGGT1_095150	19	2440	elongation factor 1-gamma, putative
TGGT1_006490	1	2	conserved hypothetical protein
TGME49_099270	1	2	hypothetical protein
TGGT1_027500	20	892	hypothetical protein
TGME49_045580	17	659	hypothetical protein
TGME49_016840	3	12	hypothetical protein
TGGT1_098290	1	8	conserved hypothetical protein
TGGT1_025070	2	3	RNA m5u methyltransferase, putative
TGME49_047740	1	2	RNA methyltransferase, TrmA family, putative
TGME49_039710	3	24	phosphomannomutase 2, putative
TGGT1_051150	3	23	phosphomannomutase, putative
TGGT1_027410	8	68	conserved hypothetical protein
TGME49_045660	7	60	hypothetical protein, conserved
TGGT1_053320	3	7	serine/threonine-protein kinase, putative
TGME49_010280	3	7	PEK kinase
TGME49_069860	4	22	hypothetical protein

TGME49_095670	8	64	hypothetical protein
TGGT1_054350	5	56	conserved hypothetical protein
TGME49_060470	12	89	C2 domain-containing protein
TGGT1_003840	1	1	conserved hypothetical protein
TGME49_054400	1	1	50S ribosomal protein l2, putative
TGME49_111060	6	31	hypothetical protein
TGGT1_087370	5	30	conserved hypothetical protein
TGGT1_005390	17	1260	conserved hypothetical protein
TGME49_082080	14	923	hypothetical protein
TGGT1_013130	9	310	conserved hypothetical protein
TGME49_057550	7	228	hypothetical protein, conserved
TGGT1_051900	1	1	conserved hypothetical protein
TGME49_039060	1	1	hypothetical protein
TGGT1_017020	37	3824	conserved hypothetical protein
TGME49_003710	36	3375	hypothetical protein
TGGT1_121440	1	2	conserved hypothetical protein
TGME49_119950	1	2	rRNA-processing protein FCF1, putative
TGGT1_028190	2	20	conserved hypothetical protein
TGME49_017840	1	8	hypothetical protein
TGGT1_020640	9	153	DnaJ domain-containing protein, putative
TGME49_007760	8	116	DnaJ domain-containing protein
TGGT1_117030	3	48	transmembrane domain-containing protein, putative
TGME49_030820	2	22	transmembrane domain-containing protein
TGME49_041150	11	117	hypothetical protein
TGGT1_049510	9	113	conserved hypothetical protein
TGGT1_045470	37	5454	conserved hypothetical protein
TGME49_044460	36	4704	hypothetical protein
TGGT1_068030	24	809	conserved hypothetical protein
TGME49_075490	24	701	hypothetical protein
TGGT1_023290	13	418	ribonucleotide-diphosphate reductase, small subunit, putative
TGME49_007060	13	326	ribonucleotide-diphosphate reductase, small subunit, putative
TGME49_076980	1	1	hypothetical protein
TGME49_032280	42	10028	hypothetical protein
TGGT1_115370	23	2031	conserved hypothetical protein
TGGT1_045110	1	2	ATP-NAD kinase domain-containing protein, putative
TGME49_044700	1	2	ATP-NAD kinase domain-containing protein
TGME49_054050	6	34	OPA3 domain-containing protein
TGGT1_003330	2	4	conserved hypothetical protein
TGGT1_043550	5	36	NAD binding domain of 6-phosphogluconate dehydrogenase, putative
TGME49_063430	4	27	NAD binding domain of 6-phosphogluconate dehydrogenase, putative
TGGT1_085550	34	2045	surface protein rhoptry, putative
TGME49_109590	19	1107	rhoptry protein, putative

TGGT1_018810	4	21	O-sialoglycoprotein endopeptidase, putative
TGME49_002310	2	7	glycoprotease family domain-containing protein
TGGT1_111940	8	141	calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase, putative
TGME49_072650	8	123	calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase, putative
TGGT1_112240	2	16	GTPase activating protein for Arf domain-containing protein, putative
TGME49_073070	2	16	ARF1-directed GTPase-activating protein, putative
TGME49_023040	26	2608	hypothetical protein, conserved
TGGT1_066580	6	130	conserved hypothetical protein
TGGT1_120960	13	176	cell cycle control protein, putative
TGME49_120430	9	128	cell cycle control protein, putative
TGME49_023410	13	300	eukaryotic translation initiation factor 4E, putative
TGGT1_077480	8	129	eukaryotic translation initiation factor 4E, putative
TGME49_009720	9	893	hypothetical protein
TGGT1_022780	4	97	conserved hypothetical protein
TGME49_036890	6	313	hypothetical protein
TGGT1_069040	5	281	conserved hypothetical protein
TGME49_106060	232	79763	hypothetical protein
TGGT1_041290	83	14430	conserved hypothetical protein
TGGT1_104780	7	42	conserved hypothetical protein
TGME49_077870	7	34	hypothetical protein, conserved
TGGT1_079510	4	32	conserved hypothetical protein
TGME49_024690	4	32	hypothetical protein
TGME49_048690	3	4	hypothetical protein
TGGT1_023960	1	2	conserved hypothetical protein
TGME49_119930	26	809	hypothetical protein
TGGT1_121460	17	781	conserved hypothetical protein
TGME49_063630	13	449	hypothetical protein
TGGT1_043870	4	60	conserved hypothetical protein
TGME49_097810	28	4348	hypothetical protein, conserved
TGGT1_096290	14	890	conserved hypothetical protein
TGME49_069170	1	1	hypothetical protein
TGGT1_109970	3	24	conserved hypothetical protein
TGME49_071090	3	24	hypothetical protein
TGGT1_112880	22	1023	conserved hypothetical protein
TGME49_073800	14	566	hypothetical protein, conserved
TGME49_047370	4	6	hypothetical protein
TGGT1_025630	2	4	conserved hypothetical protein
TGME49_047390	13	184	ATPase, AAA family domain-containing protein
TGGT1_025510	6	123	hypothetical protein
TGME49_063080	13	923	hypothetical protein
TGGT1_043170	7	120	conserved hypothetical protein
TGME49_009880	17	798	glutamic acid-rcih protein, putative

TGME49_097850	1	1	aminotransferase domain-containing protein
TGME49_088630	31	2674	nucleosome assembly domain-containing protein
TGGT1_034760	21	2172	nucleosome assembly domain-containing protein, putative
TGME49_046060	10	131	DNA-dependent RNA polymerase, putative
TGGT1_027000	4	12	DNA-dependent RNA polymerase, putative
TGME49_054120	8	217	microtubial-binding protein, putative
TGGT1_003400	3	51	microtubial-binding protein, putative
TGME49_078690	1	1	hypothetical protein
TGGT1_037570	27	2053	conserved hypothetical protein
TGME49_086160	17	1013	hypothetical protein
TGGT1_098320	14	351	nucleotidase, putative
TGME49_016810	14	351	nucleotidase, putative
TGME49_100020	6	54	ftsH protease, putative
TGGT1_066100	2	3	conserved hypothetical protein
TGME49_022290	2	3	U6 snRNA-associated Sm-like protein LSm1, putative
TGGT1_031430	14	348	protein kinase, putative
TGME49_092140	14	268	NEK kinase
TGME49_120140	6	58	ubiquinol-cytochrome C reductase hinge protein, putative
TGGT1_121250	5	27	ubiquinol-cytochrome C reductase, putative
TGME49_035530	1	1	hypothetical protein, conserved
TGME49_010720	30	401	hypothetical protein
TGGT1_124150	9	321	conserved hypothetical protein
TGGT1_038090	2	9	KIN protein, putative
TGME49_086690	2	9	hypothetical protein
TGME49_014770	14	1295	RAS small GTpase, putative
TGGT1_125200	7	216	small GTP-binding protein domain-containing protein
TGGT1_017990	1	2	conserved hypothetical protein
TGGT1_114500	19	1004	NOT2/NOT3/NOT5 domain-containing protein, putative
TGME49_033020	19	838	NOT2/NOT3/NOT5 domain-containing protein
TGME49_025760	3	8	hypothetical protein
TGGT1_080890	2	8	conserved hypothetical protein
TGME49_025260	27	850	DNA-directed RNA polymerase II largest subunit, putative
TGGT1_080360	24	881	DNA-directed RNA polymerase II largest subunit, putative
TGGT1_071700	2	10	conserved hypothetical protein
TGME49_034180	2	10	hypothetical protein
TGME49_042380	6	175	elongation of very long chain fatty acids 4 protein, putative
TGGT1_047450	3	26	elongation of very long chain fatty acids 4 protein, putative
TGME49_006520	1	2	hypothetical protein, conserved
TGME49_040600	55	9221	TCP-1/cpn60 chaperonin family protein, putative
TGGT1_050270	27	3201	chaperonin-60kD, ch60, putative
TGGT1_094480	5	40	cullin homog, putative
TGME49_116660	4	28	cullin homog, putative

TGME49_019790	10	144	U4/U6 small nuclear ribonucleoprotein, putative
TGGT1_030790	9	162	U4/U6 small nuclear ribonucleoprotein, putative
TGME49_027290	16	268	histone deacetylase, putative
TGGT1_082680	10	190	histone deacetylase, putative
TGGT1_080350	2	6	50S ribosomal protein L13, putative
TGME49_025240	2	6	50s ribosomal protein L13, putative
TGGT1_046100	1	2	conserved hypothetical protein
TGME49_043760	1	2	hypothetical protein, conserved
TGME49_031760	9	98	hypothetical protein
TGGT1_116010	1	14	conserved hypothetical protein
TGGT1_077360	1	1	conserved hypothetical protein
TGME49_023280	1	1	hypothetical protein
TGME49_080730	7	68	nucleotide-binding protein, putative
TGGT1_072570	5	56	nucleotide-binding protein, putative
TGGT1_096220	5	36	conserved hypothetical protein
TGME49_097890	4	21	hypothetical protein
TGME49_023910	16	304	acyltransferase domain containing protein
TGGT1_078280	5	45	conserved hypothetical protein
TGGT1_003610	14	250	conserved hypothetical protein
TGME49_054210	10	158	hypothetical protein
TGGT1_101110	1	2	WD domain, G-beta repeat-containing protein, putative
TGME49_049460	1	2	WD domain, G-beta repeat-containing protein
TGGT1_101570	1	4	conserved hypothetical protein
TGME49_049880	1	4	hypothetical protein
TGGT1_035070	1	1	hypothetical protein
TGME49_088330	1	1	hypothetical protein
TGGT1_004490	3	42	conserved hypothetical protein
TGME49_054910	3	42	hypothetical protein
TGME49_039470	23	509	hypothetical protein
TGME49_114830	9	70	splicing factor 3b subunit 10, putative
TGGT1_092120	4	22	splicing factor 3B subunit, putative
TGGT1_026470	5	19	conserved hypothetical protein
TGME49_046630	5	19	hypothetical protein
TGME49_088680	53	1658	endonuclease V, putative
TGGT1_034710	18	1127	AP endonuclease, putative
TGME49_110970	5	43	hypothetical protein
TGGT1_087290	4	40	conserved hypothetical protein
TGME49_040810	24	1953	hypothetical protein, conserved
TGGT1_049960	5	324	hypothetical protein
TGGT1_047100	2	4	conserved hypothetical protein
TGME49_042810	1	2	hypothetical protein
TGGT1_039860	1	9	conserved hypothetical protein
TGME49_012900	1	9	hypothetical protein
TGGT1_010790	1	1	conserved hypothetical protein
TGME49_059120	1	1	hypothetical protein

TGGT1_009890	17	424	conserved hypothetical protein
TGME49_060320	17	364	hypothetical protein, conserved
TGGT1_035890	3	14	leucine rich repeat protein, putative
TGME49_083860	2	10	leucine rich repeat protein, putative
TGGT1_079870	8	192	calcium-dependent protein kinase, putative
TGME49_024950	8	192	CAM kinase, CDPK family TgTOXPK2
TGME49_065320	3	9	hypothetical protein
TGGT1_035930	2	3	origin recognition complex subunit, putative
TGME49_083900	2	3	origin recognition complex 1 protein, putative
TGGT1_050450	1	1	CDKl1/4, putative
TGME49_040420	1	1	protein kinase (incomplete catalytic triad)
TGME49_013050	6	31	hypothetical protein
TGGT1_039680	6	24	conserved hypothetical protein
TGGT1_076740	5	41	mRNA turnover protein 4 mrt4, putative
TGME49_055320	4	28	hypothetical protein, conserved
TGME49_039380	5	65	PIK3R4 kinase-related protein (incomplete catalytic triad)
TGGT1_051500	2	21	hypothetical protein
TGME49_088380	141	77328	heat shock protein 90
TGGT1_035020	65	27173	heat shock protein, putative
TGME49_069090	1	4	hypothetical protein
TGGT1_020060	1	5	transporter, putative
TGME49_001260	1	5	sugar transporter, putative
TGME49_025370	2	4	hypothetical protein
TGME49_044170	6	41	hypothetical protein
TGME49_014180	11	204	EPN3 protein
TGGT1_124490	10	194	EPN3 protein, putative
TGME49_029470	19	1228	hypothetical protein
TGGT1_118200	16	1219	conserved hypothetical protein
TGGT1_010470	46	4315	hydroxymethyldihydropterin pyrophosphokinase-dihydropteroate synthase, putative
TGME49_059550	31	3378	hydroxymethyldihydropterin pyrophosphokinase-dihydropteroate synthase
TGME49_056030	13	437	hypothetical protein
TGGT1_014200	7	55	tubulin polymerization promoting protein, putative
TGME49_093890	1	1	hypothetical protein
TGGT1_057880	8	67	hypothetical protein
TGME49_065240	8	67	hypothetical protein
TGGT1_082420	1	6	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_026940	1	6	ubiquitin carboxyl-terminal hydrolase, putative
TGGT1_042360	20	974	conserved hypothetical protein
TGME49_105140	16	730	patatin-like phospholipase domain-containing protein
TGGT1_078260	38	2822	zinc finger (C3HC4 RING finger) protein
TGME49_023890	33	2032	zinc finger (C3HC4 RING finger) protein, putative
TGME49_068930	1	1	hypothetical protein
TGGT1_002110	20	754	conserved hypothetical protein

TGME49_053140	18	592	hypothetical protein
TGME49_018540	17	410	hypothetical protein, conserved
TGGT1_029370	9	116	hypothetical protein
TGME49_052490	11	165	vacuolar protein sorting 29, putative
TGGT1_001380	7	144	vacuolar sorting protein, putative
TGGT1_073930	7	228	conserved hypothetical protein
TGME49_093550	7	228	hypothetical protein
TGGT1_016960	2	4	conserved hypothetical protein
TGME49_003770	2	4	hypothetical protein
TGME49_031210	33	1156	EH protein, putative
TGGT1_116490	20	843	EH protein, putative
TGME49_048180	8	46	hypothetical protein
TGGT1_024760	5	29	conserved hypothetical protein
TGGT1_121650	1	1	conserved hypothetical protein
TGME49_119640	1	1	hypothetical protein
TGGT1_001220	3	22	conserved hypothetical protein
TGME49_052310	3	22	DNA polymerase epsilon p17 subunit, putative
TGME49_051640	2	2	ubiquitin-conjugating enzyme E2, putative
TGGT1_103030	1	1	ubiquitin-conjugating enzyme E2, putative
TGGT1_103300	2	18	conserved hypothetical protein
TGME49_051920	2	18	hypothetical protein
TGGT1_095080	1	1	conserved hypothetical protein
TGGT1_018050	13	610	conserved hypothetical protein
TGME49_002890	11	494	hypothetical protein
TGME49_069160	18	317	hypothetical protein, conserved
TGME49_016620	140	12118	EF hand domain-containing protein
TGGT1_098520	36	2475	EF hand domain-containing protein, putative
TGGT1_098670	16	431	conserved hypothetical protein
TGME49_016580	15	269	hypothetical protein
TGME49_044390	13	323	coatomer epsilon subunit, putative
TGGT1_045550	8	238	coatomer epsilon subunit, putative
TGGT1_024740	1	1	dimethyladenosine transferase, putative
TGME49_048200	1	1	dimethyladenosine transferase, putative
TGGT1_065570	1	2	hypothetical protein
TGME49_021700	1	2	hypothetical protein
TGME49_024280	23	298	hypothetical protein
TGGT1_078770	14	196	conserved hypothetical protein
TGME49_113480	19	669	hypothetical protein
TGGT1_089890	14	769	conserved hypothetical protein
TGME49_099050	21	6704	60S ribosomal protein L17, putative
TGGT1_006280	6	163	60S ribosomal protein L17, putative
TGME49_044660	2	4	hypothetical protein
TGME49_110670	68	8790	glycogen phosphorylase family protein, putative
TGGT1_086890	43	7032	glycogen phosphorylase, putative
TGGT1_039230	1	1	hypothetical protein

TGME49_013410	1	1	small nuclear ribonucleoprotein F, putative
TGGT1_098510	1	2	conserved hypothetical protein
TGME49_016630	1	2	hypothetical protein, conserved
TGME49_071440	18	529	hypothetical protein, conserved
TGGT1_110310	16	520	nuclear protein localization, putative
TGME49_069780	2	4	ADP-ribosylation factor, putative
TGGT1_069520	4	26	cyclin G-associated kinase, putative
TGME49_036240	4	26	Tyrosine kinase-like (TKL) protein
TGME49_039620	30	4071	nucleotidase, putative
TGGT1_051240	27	5034	nucleotidase, putative
TGGT1_107250	1	2	conserved hypothetical protein
TGME49_091670	2	4	RNA helicase, putative
TGME49_003630	5	75	ribosomal protein L41
TGGT1_017210	5	37	ribosomal protein L41, putative
TGME49_051680	15	1129	translationally-controlled tumor protein, putative
TGGT1_103060	7	382	conserved hypothetical protein
TGME49_085970	2	4	hypothetical protein, conserved
TGGT1_002070	2	22	conserved hypothetical protein
TGME49_053100	2	22	hypothetical protein
TGGT1_048610	1	4	conserved hypothetical protein
TGME49_041890	1	4	hypothetical protein
TGGT1_031570	39	2472	conserved hypothetical protein
TGME49_091990	35	1955	hypothetical protein
TGME49_012310	8	603	vacuolar ATP synthase 16 kDa proteolipid subunit, putative
TGGT1_072160	1	1	vacuolar ATP synthase 16 kDa proteolipid subunit, putative
TGME49_070690	21	412	hypothetical protein, conserved
TGGT1_109440	13	313	conserved hypothetical protein
TGGT1_118470	1	2	conserved hypothetical protein
TGME49_029280	1	2	hypothetical protein
TGGT1_039350	15	212	hypothetical protein
TGME49_013300	13	95	hypothetical protein
TGME49_018420	15	407	pleiotropic regulator 1, putative
TGGT1_029240	10	366	pleiotropic regulator, putative
TGGT1_110160	4	39	conserved hypothetical protein
TGME49_071290	4	39	splicing factor, arginine/serine-rich 8, putative
TGME49_034270	39	2414	hypothetical protein
TGGT1_071620	36	3097	hypothetical protein
TGGT1_028890	20	698	conserved hypothetical protein
TGME49_018070	15	302	hypothetical protein
TGGT1_097100	5	48	conserved hypothetical protein
TGME49_097120	5	48	hypothetical protein
TGME49_062060	18	130	protein kinase, putative
TGGT1_008050	4	36	conserved hypothetical protein
TGME49_021320	104	8013	acetyl-CoA carboxylase, putative
TGGT1_065170	50	5227	acetyl-CoA carboxylase, putative

TGME49_048130	6	27	hypothetical protein
TGGT1_024810	5	64	conserved hypothetical protein
TGGT1_028730	1	2	conserved hypothetical protein
TGME49_017360	1	2	hypothetical protein
TGME49_049030	31	1508	endonuclease/exonuclease/phosphatase domain-containing protein
TGGT1_100700	11	477	hypothetical protein
TGGT1_112030	8	131	DNA-directed RNA polymerase III subunit, putative
TGME49_072740	8	81	DNA-directed RNA polymerase III subunit, putative
TGME49_009980	27	8033	Rhoptry kinase family protein ROP42 (incomplete catalytic triad)
TGGT1_023030	9	204	protein kinase domain-containing protein, putative
TGME49_115930	7	194	hypothetical protein
TGGT1_093620	3	15	conserved hypothetical protein
TGME49_068570	10	496	zinc finger (CCCH type) protein, putative
TGGT1_106630	10	492	zinc knuckle domain-containing protein, putative
TGGT1_108160	8	195	protein kinase, putative
TGME49_069730	7	167	protein kinase
TGGT1_021780	6	79	RWD domain-containing protein, putative
TGME49_008750	6	79	RWD domain-containing protein
TGME49_058840	5	30	hypothetical protein, conserved
TGGT1_041780	12	219	conserved hypothetical protein
TGME49_105790	12	219	hypothetical protein
TGGT1_075160	1	12	conserved hypothetical protein
TGME49_094830	1	12	hypothetical protein
TGGT1_019460	2	6	cullin, putative
TGME49_001770	2	6	cullin 3, putative
TGME49_032660	16	610	58 kDa phosphoprotein, putative
TGGT1_114860	13	762	hypothetical protein
TGME49_014970	32	1741	DNA replication licensing factor, putative
TGGT1_125490	30	1737	DNA replication licensing factor, putative
TGGT1_083390	88	15979	conserved hypothetical protein
TGME49_028120	79	12885	hypothetical protein
TGME49_041180	2	5	hypothetical protein
TGGT1_049470	1	1	conserved hypothetical protein
TGME49_054440	12	425	60S ribosomal protein L12, putative
TGGT1_003900	6	213	L12 ribosomal protein, putative
TGME49_071060	21	256	vacuolar protein sorting-associated protein, putative
TGGT1_109940	10	148	vacuolar sorting protein, putative-associated protein, putative
TGME49_069950	36	2642	hypothetical protein
TGGT1_108490	29	2278	conserved hypothetical protein
TGME49_118700	16	269	eukaryotic translation initiation factor 6, putative
TGGT1_122420	7	176	eukaryotic translation initiation factor, putative

TGME49_066630	22	1231	hypothetical protein, conserved
TGGT1_076340	9	83	conserved hypothetical protein
TGME49_030520	36	1241	cyclophilin, putative
TGGT1_117440	7	169	cyclophilin, putative
TGGT1_120600	5	32	gamma-soluble nsf attachment protein, putative
TGME49_120690	5	32	hypothetical protein
TGGT1_090270	2	3	conserved hypothetical protein
TGME49_113750	2	3	hypothetical protein, conserved
TGGT1_088620	1	1	conserved hypothetical protein
TGME49_112340	1	1	hypothetical protein
TGGT1_110920	2	6	DnaJ domain-containing protein, putative
TGME49_071910	2	2	DnaJ domain-containing protein
TGME49_078270	6	48	nucleolar RNA-binding domain-containing protein
TGGT1_104280	3	30	nucleolar RNA-binding domain-containing protein, putative
TGME49_033340	12	213	hypothetical protein
TGGT1_114140	11	204	conserved hypothetical protein
TGME49_086440	9	89	hypothetical protein
TGGT1_037860	8	75	conserved hypothetical protein
TGME49_069650	18	301	hypothetical protein
TGGT1_108080	10	208	conserved hypothetical protein
TGGT1_021810	3	21	ubiquitin-conjugating enzyme E2, putative
TGME49_008780	3	21	ubiquitin-conjugating enzyme E2, putative
TGGT1_069470	1	4	conserved hypothetical protein
TGME49_036280	1	4	hypothetical protein
TGME49_061940	10	129	hydrolase, alpha/beta fold family domain containing protein
TGGT1_008290	8	125	valacyclovir hydrolase, putative
TGGT1_118170	2	8	conserved hypothetical protein
TGME49_029500	2	8	hypothetical protein, conserved
TGGT1_004180	1	2	conserved hypothetical protein
TGME49_054630	1	2	CMGC kinase
TGME49_047350	28	3182	thioredoxin, putative
TGGT1_025650	6	140	thioredoxin, putative
TGME49_094800	164	1208803	elongation factor 1-alpha, putative
TGGT1_022870	2	12	syntaxin, putative
TGME49_009820	2	12	syntaxin, putative
TGGT1_079330	1	2	calmodulin-binding domain containing protein, putative
TGME49_024610	1	2	IQ calmodulin-binding motif domain-containing protein
TGME49_094630	20	1879	PH domain-containing protein
TGGT1_074950	16	1842	PH domain-containing protein, putative
TGME49_064000	7	53	hypothetical protein
TGGT1_044240	4	36	conserved hypothetical protein
TGME49_010960	14	163	replication factor C subunit, putative
TGGT1_124000	8	84	replication factor C subunit, putative

TGME49_063270	10	115	glycerophosphoryl diester phosphodiesterase family domain containing protein
TGGT1_043390	7	179	glycerophosphoryl diester phosphodiesterase, putative
TGME49_071770	1	1	hypothetical protein
TGME49_038150	7	177	hypothetical protein
TGGT1_052700	2	15	conserved hypothetical protein
TGME49_060490	3	9	hypothetical protein
TGGT1_064410	1	1	conserved hypothetical protein
TGGT1_011730	8	98	conserved hypothetical protein
TGME49_058480	8	98	hypothetical protein, conserved
TGGT1_112050	8	89	conserved hypothetical protein
TGME49_072770	8	89	hypothetical protein, conserved
TGME49_009020	1	8	hypothetical protein
TGGT1_021960	1	4	conserved hypothetical protein
TGME49_044110	34	2812	nucleosome assembly protein, putative
TGGT1_045840	15	1982	nucleosome assembly protein, putative
TGME49_011260	46	2379	Rhoptry kinase family protein ROP26 (incomplete catalytic triad)
TGGT1_123660	15	1109	conserved hypothetical protein
TGME49_072360	2	4	hypothetical protein
TGME49_010690	33	7508	40s ribosomal protein S6, putative
TGGT1_124180	21	1659	40S ribosomal protein S6, putative
TGGT1_014740	3	24	calmodulin, putative
TGME49_020140	3	24	calmodulin, putative
TGME49_061380	3	12	30S ribosomal protein S8, putative
TGGT1_062490	8	110	conserved hypothetical protein
TGME49_006300	6	68	hypothetical protein
TGGT1_091320	15	192	conserved hypothetical protein
TGME49_114530	14	170	hypothetical protein
TGME49_053900	28	1163	delta-aminolevulinic acid dehydratase, putative
TGGT1_003180	19	1015	delta-aminolevulinic acid dehydratase, putative
TGME49_030030	7	37	hypothetical protein
TGGT1_117850	6	34	conserved hypothetical protein
TGME49_098630	39	1388	SPX domain-containing protein
TGGT1_099430	16	435	SPX domain-containing protein, putative
TGME49_002530	42	2112	aspartyl-tRNA synthetase, putative
TGGT1_018560	23	1348	aspartyl-tRNA synthetase, putative
TGME49_111890	19	930	hypothetical protein
TGGT1_088140	16	971	conserved hypothetical protein
TGME49_026960	316	182377	phosphofructokinase, putative
TGGT1_082440	65	24743	hypothetical protein
TGME49_072010	7	109	nucleolar protein family A, putative
TGGT1_111030	3	51	conserved hypothetical protein
TGME49_057580	6	164	hypothetical protein
TGGT1_013100	3	75	conserved hypothetical protein

TGME49_089150	11	159	hypothetical protein
TGGT1_034090	6	40	conserved hypothetical protein
TGME49_019120	1	1	hypothetical protein
TGGT1_085830	4	16	conserved hypothetical protein
TGME49_109880	4	16	hypothetical protein
TGME49_105810	10	182	hypothetical protein
TGGT1_041760	5	143	conserved hypothetical protein
TGME49_105240	20	401	XPA-binding protein, putative
TGGT1_042270	15	364	XPA-binding protein, putative
TGME49_116280	3	24	hypothetical protein
TGGT1_093970	1	17	conserved hypothetical protein
TGGT1_078060	1	1	LCCL domain-containing protein / F5/8 type C domain-containing protein, putative
TGME49_023700	1	1	LCCL domain-containing protein / F5/8 type C domain-containing protein
TGME49_106960	26	1246	phenylalanyl-tRNA synthetase beta chain, putative
TGGT1_055450	19	1052	phenylalanyl-tRNA synthetase beta chain, putative
TGGT1_038690	3	18	iron-sulfur cluster assembly accessory protein, putative
TGME49_013810	2	10	iron-sulfur cluster assembly accessory protein, putative
TGGT1_036980	3	12	conserved hypothetical protein
TGME49_085480	3	8	hypothetical protein
TGME49_016790	39	2185	ABC transporter, putative
TGGT1_098340	25	1464	ABC transporter, putative
TGME49_043540	21	543	U4/U6 small nuclear ribonucleoprotein, putative
TGGT1_046320	14	392	U4/U6 small nuclear ribonucleoprotein, putative
TGGT1_051910	3	15	myosin tail 1 protein, putative
TGME49_039050	3	15	hypothetical protein, conserved
TGGT1_123910	10	259	RNA-binding protein, putative
TGME49_011020	8	205	RNA-binding protein, putative
TGME49_094250	10	113	hypothetical protein
TGGT1_074540	6	53	conserved hypothetical protein
TGGT1_016750	6	44	conserved hypothetical protein
TGME49_003960	2	10	hypothetical protein
TGGT1_022900	10	211	nucleolar phosphoprotein p130, putative
TGME49_009860	8	116	nucleolar phosphoprotein p130, putative
TGME49_023500	4	12	hypothetical protein
TGGT1_077770	2	8	conserved hypothetical protein
TGGT1_079660	4	29	conserved hypothetical protein
TGME49_024740	1	2	hypothetical protein
TGME49_059980	9	173	DnaJ domain-containing protein
TGGT1_010140	6	165	DnaJ domain-containing protein, putative
TGGT1_010880	3	14	bromodomain-containing protein, putative
TGGT1_085850	2	5	conserved hypothetical protein
TGME49_109910	1	2	hypothetical protein

TGME49_110750	36	9225	hypothetical protein, conserved
TGGT1_087060	19	1949	conserved hypothetical protein
TGME49_015030	3	8	helicase associated domain-containing protein
TGGT1_125660	1	4	helicase, putative
TGME49_090950	139	36536	clathrin heavy chain, putative
TGGT1_032170	75	17862	clathrin heavy chain, putative
TGME49_022360	25	936	hypothetical protein
TGGT1_066170	17	418	conserved hypothetical protein
TGGT1_099150	4	24	conserved hypothetical protein
TGME49_016110	4	24	hypothetical protein
TGME49_066000	4	13	phosphatidylinositol 3- and 4-kinase domain-containing protein
TGGT1_056870	1	4	hypothetical protein
TGME49_039480	6	33	RNB-like protein domain containing protein
TGGT1_051390	3	30	ribonuclease, putative
TGME49_021850	2	5	hypothetical protein
TGGT1_065630	1	1	conserved hypothetical protein
TGGT1_008490	3	16	conserved hypothetical protein
TGME49_061740	2	4	hypothetical protein
TGME49_049380	6	121	zinc finger DHHC domain-containing protein
TGGT1_101050	5	152	DHHC domain-containing protein, putative
TGGT1_000360	13	256	bromodomain domain-containing protein, putative
TGME49_076180	12	172	bromodomain-containing protein
TGME49_111500	12	86	ubiquitin-activating enzyme, putative
TGGT1_087860	9	89	ubiquitin-activating enzyme E1b, putative
TGME49_054300	19	179	hypothetical protein
TGGT1_003720	16	191	conserved hypothetical protein
TGGT1_054750	6	46	tubulin-specific chaperone A, putative
TGME49_095310	6	46	tubulin-specific chaperone A, putative
TGME49_081420	2	4	histone deacetylase
TGGT1_096180	1	2	conserved hypothetical protein
TGME49_097930	1	2	HesB-like domain containing protein
TGGT1_118220	2	30	conserved hypothetical protein
TGME49_029450	1	3	hypothetical protein
TGGT1_008370	4	104	ethanolaminephosphotransferase, putative
TGME49_061760	3	96	ethanolaminephosphotransferase, putative
TGME49_030360	25	1660	hypothetical protein
TGGT1_117620	12	273	conserved hypothetical protein
TGGT1_014910	10	134	leucine rich repeat protein, putative
TGME49_020230	9	108	leucine rich repeat protein, putative
TGME49_078980	11	57	hypothetical protein, conserved
TGGT1_122960	1	4	conserved hypothetical protein
TGME49_118290	1	4	hypothetical protein, conserved
TGME49_050880	25	708	adenosine kinase
TGGT1_102440	11	420	adenosine kinase, putative

TGME49_048510	25	958	hypothetical protein
TGGT1_024340	21	1343	conserved hypothetical protein
TGGT1_126480	1	2	conserved hypothetical protein
TGME49_015550	1	2	hypothetical protein
TGGT1_056060	2	6	cancer-associated gene protein, putative
TGME49_109020	2	6	PP-loop domain-containing protein
TGME49_019250	16	209	acetyltransferase domain-containing protein
TGGT1_030210	10	149	acetyltransferase domain-containing protein, putative
TGME49_036650	28	2710	DEAD/DEAH box helicase, putative
TGGT1_069190	25	2247	DEAD/DEAH box helicase, putative
TGGT1_011380	3	8	zinc finger (C3HC4 RING finger) protein
TGME49_058790	3	8	zinc finger (C3HC4 type RING finger) protein, putative
TGME49_046550	30	1018	eukaryotic aspartyl protease, putative
TGGT1_026570	11	348	eukaryotic aspartyl protease, putative
TGME49_046910	7	29	histone-lysine N-methyltransferase, putative
TGGT1_026190	3	21	SET and MYND domain cotaining protein, putative
TGGT1_023250	1	2	conserved hypothetical protein
TGME49_007020	1	2	hypothetical protein, conserved
TGGT1_017700	10	182	replication factor, putative
TGME49_003170	8	129	replication factor-A C terminal domain-containing protein
TGME49_120100	51	6933	RNA recognition motif-containing protein
TGGT1_121290	42	6640	conserved hypothetical protein
TGME49_063720	4	29	HMG box domain-containing protein
TGGT1_043980	3	21	HMG box domain-containing protein, putative
TGGT1_101100	16	679	conserved hypothetical protein
TGME49_049440	15	678	hypothetical protein
TGME49_032680	16	166	hypothetical protein
TGGT1_114840	13	157	conserved hypothetical protein
TGME49_034440	4	43	aminotransferase, putative
TGGT1_071450	3	7	aminotransferase, putative
TGME49_070550	20	1044	gamma-glutamyl phosphate reductase, putative
TGGT1_109100	19	1030	gamma-glutamyl phosphate reductase, putative
TGGT1_043480	8	203	WD repeat domain, putative
TGME49_063360	7	162	hypothetical protein
TGME49_023540	33	1596	hypothetical protein, conserved
TGGT1_077900	25	1483	conserved hypothetical protein
TGME49_003480	17	1007	hypothetical protein, conserved
TGGT1_017350	16	1091	conserved hypothetical protein
TGME49_113420	2	9	hypothetical protein
TGME49_089380	26	713	hypothetical protein
TGGT1_033840	11	254	conserved hypothetical protein
TGGT1_076720	42	2830	conserved hypothetical protein
TGME49_055300	38	2440	hypothetical protein
TGME49_116620	2	10	WD-40 repeats-containing protein

TGGT1_010160	1	1	adenosine/guanosine diphosphatase, putative
TGME49_059960	1	1	GDA1/CD39 (nucleoside phosphatase) family domain containing protein
TGME49_070800	22	449	GAF domain-containing protein
TGGT1_109560	1	2	hypothetical protein
TGGT1_052110	30	1678	transcription initiation factor ia, putative
TGME49_038880	28	1435	hypothetical protein
TGGT1_081840	4	14	conserved hypothetical protein
TGME49_026670	2	5	hypothetical protein
TGGT1_122520	6	73	mrp protein, putative
TGME49_118590	6	73	mrp protein, putative
TGME49_113360	25	266	hypothetical protein
TGGT1_089760	6	78	conserved hypothetical protein
TGME49_110290	12	124	hypothetical protein, conserved
TGGT1_086520	8	142	regulator of chromosome condensation, putative
TGME49_026660	44	1319	hypothetical protein
TGGT1_081820	1	16	hypothetical protein
TGME49_053610	1	1	hypothetical protein
TGME49_022410	16	321	hypothetical protein, conserved
TGGT1_066220	6	53	conserved hypothetical protein
TGGT1_114720	26	1107	conserved hypothetical protein
TGME49_032800	26	883	hypothetical protein
TGGT1_116210	1	3	conserved hypothetical protein
TGME49_031460	1	3	hypothetical protein
TGGT1_118010	3	22	galactokinase, putative
TGME49_029780	2	11	galactokinase, putative
TGGT1_017390	2	16	anaphase-promoting complex, putative
TGME49_003440	2	16	anaphase-promoting complex subunit 10 domain-containing protein
TGME49_093730	25	813	DHHC zinc finger domain-containing protein
TGGT1_074120	15	740	conserved hypothetical protein
TGME49_058770	24	1019	UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase T2
TGGT1_011410	8	125	UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase T2, putative
TGGT1_115810	2	5	conserved hypothetical protein
TGME49_031950	2	5	hypothetical protein
TGGT1_032260	2	4	DnaJ domain-containing protein, putative
TGME49_090880	1	1	DnaJ domain-containing protein
TGGT1_103380	16	664	conserved hypothetical protein
TGME49_078970	14	568	ICE-like protease (caspase) p20 domain-containing protein
TGME49_089620	17	593	cathepsin C
TGGT1_033590	9	202	cathepsin C, putative
TGME49_024530	10	425	inner membrane complex associated protein 4
TGGT1_079150	9	421	membrane skeletal protein IMC1, putative
TGGT1_099330	1	3	conserved hypothetical protein

TGME49_015930	1	3	hypothetical protein
TGGT1_082510	4	22	NAD-dependent deacetylase, putative
TGME49_027020	4	22	NAD-dependent deacetylase, putative
TGME49_073740	1	4	acetyl-CoA acyltransferase B, putative
TGME49_078240	2	8	zinc-finger-Ran binding domain-containing protein
TGGT1_104320	1	4	zinc finger protein-Ran binding domain-containing protein, putative
TGGT1_117270	16	353	conserved hypothetical protein
TGME49_030580	13	167	hypothetical protein
TGME49_036010	3	3	prenylcysteine oxidase, putative
TGGT1_069750	2	2	prenylcysteine oxidase, putative
TGGT1_083850	4	95	conserved hypothetical protein
TGME49_028280	3	69	hypothetical protein
TGME49_093430	20	1795	hypothetical protein
TGGT1_073810	11	188	conserved hypothetical protein
TGME49_039560	12	204	myosin E, putative
TGME49_036000	4	19	2Fe-2S iron-sulfur cluster binding domain containing protein
TGGT1_069760	3	15	2Fe-2S iron-sulfur cluster binding domain containing protein, putative
TGGT1_033930	17	824	conserved hypothetical protein
TGME49_089290	17	824	hypothetical protein
TGME49_037470	9	65	aminomethyltransferase, mitochondrial, putative
TGGT1_068430	6	82	aminomethyltransferase, putative
TGME49_058980	4	12	cell differentiation protein Rcd1, putative
TGGT1_010890	2	8	conserved hypothetical protein
TGME49_061800	3	6	helicase, putative
TGGT1_008320	1	2	pom1, putative
TGME49_049920	8	52	diphthine synthase, putative
TGGT1_101610	5	43	tetrapyrrole (corrin/porphyrin) methylase domain-containing protein, putative
TGME49_083590	2	12	hypothetical protein
TGGT1_035620	2	6	conserved hypothetical protein
TGME49_039540	7	82	LEM3 / CDC50 family protein
TGGT1_051320	6	78	LEM3 / CDC50 domain-containing protein
TGME49_044580	2	6	50S ribosomal protein L1, putative
TGGT1_045350	1	2	hypothetical protein
TGME49_042060	1	25	hypothetical protein
TGME49_095630	6	443	hypothetical protein
TGGT1_054390	6	439	conserved hypothetical protein
TGGT1_068670	2	7	conserved hypothetical protein
TGME49_037240	2	7	hypothetical protein
TGME49_005680	13	467	transmembrane domain-containing protein
TGGT1_063230	5	149	transmembrane domain-containing protein, putative
TGME49_015910	39	1407	hypothetical protein
TGGT1_099350	20	550	conserved hypothetical protein

TGME49_034380	37	17220	hypothetical protein
TGGT1_071510	19	1361	conserved hypothetical protein
TGGT1_116680	5	22	exosome complex exonuclease, putative
TGME49_031040	5	22	exosome complex exonuclease, putative
TGME49_028420	2	14	AGC kinase
TGME49_034390	12	46	hypothetical protein
TGGT1_071500	5	27	conserved hypothetical protein
TGME49_014320	29	63749	facilitative glucose transporter, putative
TGGT1_124740	14	2052	facilitative glucose transporter, putative
TGME49_086580	114	127459	lysophospholipase, putative
TGGT1_038000	41	6990	lysophospholipase, putative
TGGT1_094940	5	40	hypothetical protein
TGME49_100350	5	40	molybdopterin cofactor sulfurase, putative
TGME49_055220	8	43	hypothetical protein
TGGT1_076630	2	13	hypothetical protein
TGME49_069690	48	1910	hypothetical protein
TGGT1_108120	19	799	conserved hypothetical protein
TGME49_029360	46	3480	transaldolase, putative
TGGT1_118310	21	2098	transaldolase, putative
TGME49_002510	6	22	hypothetical protein
TGGT1_018580	1	1	conserved hypothetical protein
TGME49_047690	22	560	p-type ATPase2, putative
TGGT1_025120	11	202	ATPase, putative
TGGT1_052580	1	6	conserved hypothetical protein
TGME49_038270	1	6	hypothetical protein
TGME49_049390	85	24431	NAD-specific glutamate dehydrogenase, putative
TGGT1_101070	58	16863	NAD-specific glutamate dehydrogenase, putative
TGME49_112110	52	5916	nucleoredoxin, putative
TGGT1_088350	30	2238	nucleoredoxin, putative
TGGT1_089000	8	82	conserved hypothetical protein
TGME49_112700	6	72	hypothetical protein
TGME49_042420	8	60	amine oxidase, flavin-containing domain-containing protein
TGGT1_047400	6	72	amine oxidase, putative
TGME49_070770	18	262	hypothetical protein
TGGT1_109530	13	283	rbm25 protein, putative
TGGT1_113490	2	6	conserved hypothetical protein
TGME49_033840	1	2	hypothetical protein
TGME49_054940	100	42218	hypothetical protein
TGGT1_004520	98	47476	eukaryotic translation initiation factor 4g, putative
TGGT1_034380	5	34	conserved hypothetical protein
TGME49_088890	3	22	hypothetical protein
TGGT1_092930	1	1	conserved hypothetical protein
TGGT1_110600	9	144	conserved hypothetical protein
TGME49_071800	9	144	hypothetical protein

TGME49_121640	60	2599	cell division protein 48, putative
TGGT1_064680	28	1272	cell division protein, putative
TGGT1_098980	3	9	conserved hypothetical protein
TGME49_016270	3	9	hypothetical protein
TGGT1_055170	25	702	AROM polypeptide/ shikimate-quinate 5-dehydrogenase, putative
TGME49_107040	25	551	AROM polypeptide/ shikimate-quinate 5-dehydrogenase
TGGT1_080940	3	15	ABC transporter, putative
TGME49_025800	2	8	ABC transporter, putative
TGME49_063710	8	60	sterol O-acyltransferase, putative
TGGT1_043970	6	115	dihydrolipoamide succinyltransferase component of 2-oxoglutarate dehydrogenase, putative
TGME49_000410	1	1	hypothetical protein
TGGT1_094470	42	10008	hypothetical protein
TGME49_116650	33	7942	hypothetical protein, conserved
TGME49_066990	36	1679	coatomer beta subunit, putative
TGGT1_075950	34	1818	coatomer beta subunit, putative
TGGT1_042750	2	6	conserved hypothetical protein
TGME49_108050	2	6	hypothetical protein, conserved
TGGT1_110390	1	1	conserved hypothetical protein
TGME49_017480	20	459	hypothetical protein
TGGT1_028610	16	418	conserved hypothetical protein
TGME49_032830	66	8088	vacuolar proton-translocating ATPase subunit, putative
TGGT1_114690	21	1319	vacuolar proton-translocating ATPase subunit, putative
TGME49_011090	10	148	cysteine desulfurase, putative
TGGT1_123840	8	144	cysteine desulfurase, putative
TGME49_026280	2	4	hypothetical protein, conserved
TGME49_089570	10	63	phosphatidylinositol transfer protein beta isoform, putative
TGGT1_033660	4	49	phosphatidylinositol transfer protein, putative
TGGT1_033640	1	3	oxysterol-binding protein, putative
TGME49_031120	2	2	ribosomal protein S11, putative
TGGT1_116600	1	1	ribosomal protein S11, putative
TGME49_002060	16	421	hypothetical protein
TGGT1_019160	9	161	hypothetical protein
TGGT1_019150	6	60	hypothetical protein
TGME49_046100	14	328	hypothetical protein
TGGT1_026950	14	327	hypothetical protein
TGGT1_089470	8	258	prp4, putative
TGME49_113180	6	163	CMGC kinase, Dyrk family
TGGT1_045410	5	45	conserved hypothetical protein
TGME49_044530	5	45	hypothetical protein
TGME49_051400	3	18	transmembrane protein, putative
TGGT1_102700	2	14	transmembrane protein, putative
TGGT1_077230	1	2	conserved hypothetical protein
TGME49_055720	1	2	hypothetical protein

TGME49_042660	25	979	vacuolar sorting protein 35, putative
TGGT1_047260	17	792	vacuolar sorting protein, putative
TGME49_068590	24	1661	rhomboid-like protease 4
TGGT1_106670	14	984	rhomboid-like protease 4
TGGT1_075440	4	32	conserved hypothetical protein
TGME49_095080	4	32	hypothetical protein
TGME49_004020	31	4679	60S ribosomal protein L8, putative
TGGT1_016680	11	497	60S ribosomal protein L8, putative
TGME49_022270	17	390	hypothetical protein
TGGT1_066080	9	220	conserved hypothetical protein
TGME49_055190	98	217040	myosin C
TGGT1_076590	35	1675	myosin C, putative
TGME49_105300	1	1	DNA-repair endonuclease subunit, putative
TGGT1_021550	7	81	conserved hypothetical protein
TGME49_008520	6	75	hypothetical protein
TGGT1_094510	32	2380	RNA pseudouridylate synthase domain-containing protein, putative
TGME49_116680	24	1763	RNA pseudouridylate synthase domain-containing protein
TGME49_006480	4	29	hypothetical protein, conserved
TGGT1_062290	3	23	conserved hypothetical protein
TGME49_112830	4	23	hypothetical protein
TGGT1_089020	3	22	conserved hypothetical protein
TGME49_015290	13	81	saccharopine dehydrogenase, putative
TGGT1_126030	7	60	saccharopine dehydrogenase, putative
TGME49_089250	5	55	cyclophilin, putative
TGME49_045710	3	5	MORN repeat protein, putative
TGGT1_027360	1	1	MORN repeat protein, putative
TGME49_039750	63	3491	hypothetical protein
TGGT1_051090	28	2401	galactosyltransferase, putative
TGME49_091040	4	36	lactate dehydrogenase
TGGT1_114240	6	519	conserved hypothetical protein
TGME49_033250	6	362	hypothetical protein
TGGT1_010830	2	12	acetyltransferase domain-containing protein, putative
TGME49_059070	1	6	acetyltransferase domain-containing protein
TGGT1_010290	3	6	conserved hypothetical protein
TGME49_059840	1	2	hypothetical protein
TGME49_042020	20	395	mannosyl-oligosaccharide glucosidase, putative
TGGT1_048390	15	317	mannosyl-oligosaccharide glucosidase, putative
TGME49_059010	40	3345	vacuoloar ATP synthase subunit D, putative
TGGT1_010860	17	769	vacuoloar ATP synthase subunit D, putative
TGME49_016860	57	6350	ATP-dependent RNA helicase, putative
TGGT1_098270	21	2923	ATP-dependent RNA helicase, putative
TGME49_076890	41	1332	hypothetical protein
TGGT1_105430	12	216	conserved hypothetical protein
TGME49_106620	7	40	platelet binding protein GspB, putative

TGME49_023750	2	4	hypothetical protein, conserved
TGME49_024660	24	1151	transcription factor S-II central domain-containing protein
TGGT1_079480	23	1142	transcription factor TFIIS central domain-containing protein
TGME49_026600	14	156	syntaxin, putative
TGGT1_081750	5	40	syntaxin, putative
TGGT1_069770	7	56	hypothetical protein
TGME49_035990	2	10	hypothetical protein
TGME49_038180	8	166	26S proteasome non-ATPase regulatory subunit, putative
TGGT1_052680	4	74	26S proteasome non-ATPase regulatory subunit, putative
TGGT1_039670	27	2099	WD domain-containing protein, putative
TGME49_013060	16	1049	WD domain-containing protein
TGGT1_025710	3	5	conserved hypothetical protein
TGME49_047290	2	2	hypothetical protein
TGGT1_098920	1	2	conserved hypothetical protein
TGME49_016330	1	2	hypothetical protein
TGME49_094900	36	2731	DnaK family domain containing protein
TGGT1_075250	30	2622	heat shock protein 70kD, putative
TGME49_005220	16	434	U5 snRNP-associated 102 kDa protein, putative
TGGT1_063800	14	428	U5 snRNP-associated 102 kDa protein, putative
TGGT1_034150	20	741	merozoite surface protein, putative
TGME49_089100	19	586	hypothetical protein, conserved
TGME49_017530	16	215	hypothetical protein
TGGT1_028540	7	102	conserved hypothetical protein
TGGT1_034730	1	10	conserved hypothetical protein
TGME49_088660	1	10	hypothetical protein
TGME49_065150	19	437	hypothetical protein
TGGT1_058070	15	367	conserved hypothetical protein
TGME49_029000	46	2195	kelch motif domain-containing protein
TGGT1_118740	18	805	kelch motif domain-containing protein, putative
TGGT1_010840	5	120	conserved hypothetical protein
TGME49_059040	5	104	hypothetical protein
TGGT1_099070	6	123	conserved hypothetical protein
TGME49_016190	5	88	hypothetical protein
TGGT1_099020	8	101	conserved hypothetical protein
TGME49_016240	8	101	hypothetical protein
TGME49_018900	4	41	hypothetical protein
TGGT1_008710	4	32	nucleolar GTP-binding protein NGB, putative
TGME49_061520	4	32	nucleolar GTP-binding protein, putative
TGGT1_028180	2	11	conserved hypothetical protein
TGME49_017850	2	11	hypothetical protein
TGME49_035180	2	4	hypothetical protein
TGME49_014960	11	244	hypothetical protein
TGGT1_125480	5	45	conserved hypothetical protein
TGGT1_123040	4	59	conserved hypothetical protein

TGME49_118200	4	59	hypothetical protein
TGGT1_087250	16	1819	conserved hypothetical protein
TGME49_110930	13	1517	hypothetical protein
TGME49_018240	31	916	hypothetical protein
TGGT1_028960	25	1172	conserved hypothetical protein
TGME49_090690	5	9	hypothetical protein
TGGT1_065470	3	9	conserved hypothetical protein
TGME49_021600	2	4	hypothetical protein, conserved
TGME49_081980	11	625	cytidine diphosphate-diacylglycerol synthase, putative
TGGT1_005490	10	612	hypothetical protein
TGME49_015060	18	652	small GTP-binding protein sar1, putative
TGGT1_125690	13	565	small GTP-binding protein sar1, putative
TGME49_024580	9	118	splicing factor, putative
TGGT1_079300	5	96	splicing factor, putative
TGGT1_032710	13	396	conserved hypothetical protein
TGME49_090340	13	349	PBS lyase HEAT-like repeat domain-containing protein
TGGT1_071290	14	464	conserved hypothetical protein
TGME49_034610	13	350	hypothetical protein
TGME49_001760	21	843	hypothetical protein
TGGT1_019570	13	677	conserved hypothetical protein
TGME49_029930	27	1151	p25-alpha domain-containing protein
TGGT1_117960	14	544	p25-alpha domain-containing protein, putative
TGGT1_118670	2	18	DHHC domain-containing protein, putative
TGME49_029160	2	18	zinc finger DHHC domain-containing protein
TGGT1_126620	16	501	phosphatidylinositol 3- and 4-kinase domain-containing protein
TGME49_015700	15	366	phosphatidylinositol 3-kinase, putative
TGME49_066070	18	3299	60s ribosomal protein L31, putative
TGGT1_056690	3	63	60S ribosomal protein L31, putative
TGME49_069250	14	559	26S proteasome regulatory subunit, putative
TGGT1_107450	10	393	26S proteasome regulatory subunit, putative
TGME49_120170	3	9	SRS16E (= SRS6)
TGME49_081640	14	115	hypothetical protein
TGGT1_005710	4	30	conserved hypothetical protein
TGGT1_123320	3	19	conserved hypothetical protein
TGME49_011470	3	19	hypothetical protein
TGME49_112100	81	16621	Ca2+-ATPase
TGGT1_088340	30	3128	calcium dependent ATPase, putative
TGME49_080560	56	3743	selenophosphate synthetase, putative
TGGT1_072750	42	3392	selenophosphate synthase, putative
TGGT1_086470	1	2	conserved hypothetical protein
TGME49_110240	1	2	hypothetical protein
TGME49_008450	15	509	serine proteinase inhibitor, putative
TGGT1_021470	9	316	hypothetical protein
TGGT1_108440	10	92	WD-repeat protein, putative

TGME49_069900	9	81	WD-repeat protein, putative
TGGT1_096830	5	43	conserved hypothetical protein
TGME49_097400	5	43	hypothetical protein
TGME49_090290	78	9574	ubiquitin-activating enzyme E1, putative
TGGT1_032760	43	7399	hypothetical protein
TGGT1_051990	5	65	conserved hypothetical protein
TGME49_038990	5	65	hypothetical protein, conserved
TGME49_109850	1	1	hypothetical protein
TGME49_022720	6	39	glycogen synthase, putative
TGGT1_066330	2	3	glycogen synthase, putative
TGME49_109810	30	4979	60S acidic ribosomal protein P2, putative
TGGT1_085770	10	1025	60S acidic ribosomal protein P2, putative
TGGT1_043310	7	143	conserved hypothetical protein
TGME49_063210	4	66	hypothetical protein
TGME49_039300	32	815	hypothetical protein
TGGT1_051670	9	109	conserved hypothetical protein
TGGT1_017230	4	16	3'-5' exoribonuclease csl4, putative
TGME49_003610	3	10	3'-5' exoribonuclease csl4, putative
TGME49_027010	4	29	Rhoptry kinase family protein ROP30
TGME49_050050	1	2	hypothetical protein
TGGT1_064900	18	2939	transmembrane domain-containing protein, putative
TGME49_021180	18	2135	transmembrane domain-containing protein
TGME49_115510	14	192	hypothetical protein
TGGT1_092980	9	131	conserved hypothetical protein
TGME49_108860	47	11444	hypothetical protein
TGGT1_055890	27	5256	conserved hypothetical protein
TGGT1_124030	9	122	serine/threonine-protein kinase rio1, putative
TGME49_010830	9	122	RIO1 family domain-containing protein
TGME49_106460	10	284	hypothetical protein
TGGT1_040880	9	275	bromodomain cotaining protein, putative
TGME49_034450	18	1192	40S ribosomal protein S15a, putative
TGGT1_071440	7	301	40S ribosomal protein S15A, putative
TGGT1_052100	5	30	conserved hypothetical protein
TGME49_038890	5	30	hypothetical protein
TGME49_059940	3	12	hypothetical protein
TGGT1_010180	2	15	conserved hypothetical protein
TGME49_026320	17	1412	hypothetical protein
TGGT1_081490	6	87	conserved hypothetical protein
TGME49_017680	97	6215	hypothetical protein, conserved
TGGT1_028360	42	3058	conserved hypothetical protein
TGGT1_021100	1	3	conserved hypothetical protein
TGME49_008010	1	3	hypothetical protein
TGGT1_009020	7	212	ankyrin repeat-containing protein, putative
TGME49_061230	6	174	ankyrin repeat-containing protein
TGME49_005290	5	11	dual-specificity phosphatase laforin, putative

TGGT1_063710	2	2	dual-specificity phosphatase laforin, putative
TGGT1_021370	3	25	conserved hypothetical protein
TGME49_008360	2	6	hypothetical protein
TGGT1_117950	5	40	peptidyl-prolyl cis-trans isomerase, cyclophylin protein, putative
TGME49_029940	5	40	peptidyl-prolyl cis-trans isomerase, cyclophylin protein, putative
TGME49_058400	1	1	LCCL domain-containing protein
TGGT1_109170	5	32	conserved hypothetical protein
TGME49_070610	2	8	hypothetical protein
TGGT1_114710	2	6	conserved hypothetical protein
TGME49_032810	1	2	erv1 / Alr family domain-containing protein
TGGT1_026540	1	2	conserved hypothetical protein
TGME49_046570	1	2	hypothetical protein
TGME49_007140	3	12	SRS49B (= SAG2X)
TGGT1_023380	2	10	SRS49B (= SAG2X)
TGME49_057170	4	15	hypothetical protein
TGGT1_013540	1	2	conserved hypothetical protein
TGGT1_094160	5	18	erythrocyte-binding protein, putative
TGME49_116350	5	18	erythrocyte-binding protein, putative
TGME49_084540	15	733	ATP synthase subunit O, putative
TGGT1_036360	8	155	ATP synthase, putative
TGGT1_081860	6	43	hypothetical protein
TGME49_026690	5	36	hypothetical protein
TGGT1_038240	1	3	conserved hypothetical protein
TGME49_087040	1	3	hypothetical protein
TGGT1_041550	10	124	endonuclease III, putative
TGME49_105920	7	43	endonuclease III-like protein 1, putative
TGME49_047280	4	7	hypothetical protein
TGGT1_025720	1	2	conserved hypothetical protein
TGGT1_093680	1	2	conserved hypothetical protein
TGME49_115980	1	2	hypothetical protein
TGGT1_001250	17	693	glycerol uptake protein, putative
TGME49_052340	3	93	MBOAT domain-containing protein
TGME49_090010	25	1241	proteasome subunit beta type 1, putative
TGGT1_033030	12	936	proteasome subunit beta type, putative
TGME49_009180	1	1	PAN domain-containing protein, putative
TGME49_095340	16	853	UV excision repair protein rhp23, putative
TGGT1_054610	10	682	UV excision repair protein rad23, putative
TGME49_010700	81	3980	hypothetical protein
TGGT1_124170	37	3072	vacuolar sorting protein, putative-associated protein, putative
TGGT1_115390	34	4004	peroxisomal catalase, putative
TGME49_032250	31	4385	peroxisomal catalase
TGME49_118150	14	429	major facilitator superfamily domain-containing protein
TGGT1_123100	7	416	conserved hypothetical protein

TGME49_016670	122	35909	KH domain-containing protein
TGGT1_098470	65	21575	KH domain-containing protein, putative
TGGT1_114350	29	2481	conserved hypothetical protein
TGME49_033160	27	1648	hypothetical protein
TGGT1_018780	1	1	50S ribosomal protein L21, putative
TGME49_002350	1	1	hypothetical protein
TGME49_026440	18	290	SNF2/RAD54 helicase family protein
TGGT1_081610	15	278	DNA repair helicase rad5,16, putative
TGGT1_036600	9	269	conserved hypothetical protein
TGME49_084780	9	246	hypothetical protein
TGGT1_068330	7	87	nifU protein, putative
TGME49_037560	7	87	nifU protein, putative
TGGT1_105260	2	10	hypothetical protein
TGME49_077070	2	10	SNF2 family N-terminal domain-containing protein
TGME49_115710	2	4	hypothetical protein, conserved
TGME49_023940	37	6859	hypothetical protein
TGGT1_078320	20	4828	conserved hypothetical protein
TGME49_058730	2	6	hypothetical protein
TGGT1_011460	1	2	conserved hypothetical protein
TGME49_060540	35	2502	hypothetical protein
TGGT1_009650	15	925	conserved hypothetical protein
TGGT1_032480	3	15	conserved hypothetical protein
TGME49_090650	3	15	hypothetical protein, conserved
TGGT1_028270	1	1	conserved hypothetical protein
TGME49_017780	1	1	hypothetical protein
TGGT1_123340	2	4	conserved hypothetical protein
TGME49_011450	2	4	hypothetical protein, conserved
TGME49_008370	87	20772	myosin heavy chain, putative
TGGT1_021380	17	1286	myosin heavy chain, putative
TGME49_111160	10	122	PWI domain-containing protein
TGGT1_087470	5	99	PWI domain-containing protein, putative
TGGT1_111370	3	76	conserved hypothetical protein
TGME49_072150	2	52	hypothetical protein
TGGT1_074050	1	1	conserved hypothetical protein
TGME49_093660	1	1	hypothetical protein
TGME49_057110	6	18	hypothetical protein
TGGT1_013600	3	5	kub3-prov protein, putative
TGGT1_041070	3	23	conserved hypothetical protein
TGME49_106250	2	14	hypothetical protein
TGME49_110130	7	46	Spc97 / Spc98 family domain-containing protein
TGGT1_086250	5	27	gamma-tubulin complex component, putative
TGGT1_021360	4	29	RNA recognition motif-containing protein, putative
TGME49_008350	4	29	RNA recognition motif-containing protein
TGGT1_057470	13	311	conserved hypothetical protein
TGME49_065520	13	311	hypothetical protein

TGME49_113570	3	18	regulator of chromosome condensation domain-containing protein
TGGT1_089980	2	12	UVb-resistance protein uvr8, putative
TGGT1_053250	1	1	conserved hypothetical protein
TGME49_010380	1	1	hypothetical protein
TGGT1_071830	19	693	conserved hypothetical protein
TGME49_012090	17	662	hypothetical protein
TGGT1_005340	3	29	conserved hypothetical protein
TGME49_082130	2	15	hypothetical protein
TGME49_100200	21	1697	histone H2A
TGGT1_095100	4	100	histone H2A, putative
TGME49_090840	27	588	trypsin, putative
TGGT1_032310	15	451	trypsin, putative
TGME49_047210	5	25	hypothetical protein
TGGT1_025790	3	18	conserved hypothetical protein
TGME49_089970	10	1056	hypothetical protein
TGGT1_033170	8	683	conserved hypothetical protein
TGME49_090270	6	49	SPRY domain-containing protein
TGGT1_032770	4	41	SPRY domain-containing protein, putative
TGGT1_005420	6	129	hypothetical protein
TGME49_082050	4	128	hypothetical protein
TGME49_002300	9	110	inosine triphosphate pyrophosphatase, putative
TGGT1_018820	5	86	inosine triphosphate pyrophosphatase, putative
TGME49_053440	23	1403	CMCG kinase (incomplete catalytic triad)
TGGT1_002520	23	1399	srpk, putative
TGME49_071890	34	2029	3-ketoacyl-CoA reductase, putative
TGGT1_110900	12	304	oxidoreductase, putative
TGGT1_086420	15	402	conserved hypothetical protein
TGME49_110190	13	179	PIK3R4 kinase-related protein (incomplete catalytic triad)
TGGT1_106310	4	36	conserved hypothetical protein
TGME49_068240	2	14	hypothetical protein
TGGT1_122640	10	287	DNA repair protein, putative
TGME49_118480	10	287	DNA repair protein, putative
TGME49_054010	7	51	serine carboxypeptidase S28, putative
TGGT1_003290	6	47	serine carboxypeptidase, putative
TGME49_097830	18	185	armadillo/beta-catenin-like repeat-containing protein
TGGT1_096270	13	212	conserved hypothetical protein
TGGT1_052450	24	1532	conserved hypothetical protein
TGME49_038510	24	1459	hypothetical protein, conserved
TGME49_061010	25	844	26S protease regulatory subunit 8, putative
TGGT1_009260	14	621	26S protease regulatory subunit, putative
TGME49_091680	30	1425	protein transport protein Sec23, putative
TGGT1_031760	26	1420	protein transport protein Sec23, putative
TGGT1_064490	16	298	transcription regulatory protein SNF2, putative

TGME49_121440	15	265	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin, putative
TGME49_016510	15	569	hypothetical protein
TGGT1_098730	2	6	conserved hypothetical protein
TGGT1_100650	3	9	nuclear movement domain-containing protein, putative
TGME49_048980	3	9	nuclear movement domain-containing protein
TGGT1_080260	7	204	conserved hypothetical protein
TGME49_025160	6	156	hypothetical protein
TGGT1_099130	4	59	ubiquitin-conjugating enzyme E2, putative
TGME49_016130	4	59	ubiquitin-conjugating enzyme E2, putative
TGGT1_078170	2	10	conserved hypothetical protein
TGGT1_037590	7	148	conserved hypothetical protein
TGME49_086190	6	101	hypothetical protein
TGGT1_010170	12	249	carbonic anhydrase, putative
TGME49_059950	11	206	carbonic anhydrase domain containing protein
TGGT1_008030	4	128	conserved hypothetical protein
TGME49_062080	3	88	hypothetical protein
TGGT1_017450	20	680	DnaJ / zinc finger (C2H2 type) domain-containing protein
TGME49_003380	20	601	DnaJ domain-containing protein
TGME49_118730	21	1110	UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase T3, putative
TGGT1_122390	7	52	UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase T3, putative
TGME49_021210	37	2651	18 kDa cyclophilin
TGGT1_065030	12	684	18 kDa cyclophilin, putative
TGME49_120300	28	872	SNF2 family N-terminal domain-containing protein
TGGT1_120990	15	775	transcription regulatory protein SNF2, putative
TGME49_012240	97	98068	tubulin beta chain, putative
TGGT1_071990	1	2	tubulin beta chain, putative
TGGT1_078610	13	151	conserved hypothetical protein
TGME49_024150	10	96	hypothetical protein
TGME49_106080	6	41	DEAD/DEAH box helicase, putative
TGGT1_041270	5	40	DEAD/DEAH box helicase, putative
TGME49_086590	21	447	hypothetical protein, conserved
TGGT1_038020	11	223	conserved hypothetical protein
TGGT1_112980	3	15	conserved hypothetical protein
TGME49_073890	1	3	hypothetical protein
TGME49_070370	2	11	clathrin-associated adaptor complex small chan, putative
TGGT1_108900	2	10	clathrin-associated adaptor complex small chan, putative
TGGT1_010020	1	2	conserved hypothetical protein
TGME49_062520	5	17	cyclophilin, putative
TGME49_065770	26	1246	hypothetical protein
TGGT1_057210	21	1290	conserved hypothetical protein
TGME49_071610	7	36	pyrroline-5-carboxylase reductase, putative

TGGT1_110400	5	41	pyrroline-5-carboxylase reductase, putative
TGME49_112230	24	885	DNA topoisomerase II, putative
TGGT1_088480	21	940	DNA topoisomerase II, putative
TGGT1_052740	2	9	conserved hypothetical protein
TGME49_038110	2	9	replication factor A protein 3 domain-containing protein
TGME49_110350	10	52	hypothetical protein
TGGT1_086570	1	2	conserved hypothetical protein
TGGT1_043730	16	358	gamma-tubulin complex component, putative
TGME49_063510	16	358	gamma-tubulin complex component 2, putative
TGGT1_045660	10	836	conserved hypothetical protein
TGME49_044280	9	659	hypothetical protein, conserved
TGGT1_087880	12	227	WD repeat domain-containing protein, putative
TGME49_111620	11	193	WD repeat domain-containing protein
TGGT1_118230	1	2	zinc finger (C3HC4 type, RING finger) protein
TGME49_029440	1	2	zinc finger (C3HC4 type RING finger) protein, putative
TGGT1_003850	5	65	protein phosphatase 2C, putative
TGME49_054410	2	17	protein phosphatase 2C, putative
TGGT1_024680	2	4	zinc finger (CCCH type) protein
TGME49_048270	2	4	zinc finger (CCCH type) protein, putative
TGME49_115660	2	2	hypothetical protein
TGGT1_026020	2	3	conserved hypothetical protein
TGME49_046990	1	2	hypothetical protein
TGGT1_003700	3	48	conserved hypothetical protein
TGME49_054280	2	31	hypothetical protein
TGGT1_115030	1	2	conserved hypothetical protein
TGME49_032490	1	2	hypothetical protein
TGGT1_020800	4	50	mitogen-activated protein kinase, putative
TGME49_007820	4	50	CMGC kinase, MAPK family
TGGT1_031620	3	10	conserved hypothetical protein
TGME49_091940	1	3	hypothetical protein
TGGT1_049940	3	6	valacyclovir hydrolase, putative
TGME49_040830	1	2	hydrolase, alpha/beta fold family domain-containing protein
TGGT1_058140	4	25	conserved hypothetical protein
TGME49_065080	4	25	tubulin-tyrosine ligase family domain containing protein
TGME49_043350	3	17	gamma-glutamyl hydrolase, putative
TGGT1_046510	2	10	gamma-glutamyl hydrolase, putative
TGME49_032380	14	151	WD-40 repeat protein, putative
TGGT1_115250	12	168	WD-repeat protein, putative
TGME49_094980	5	27	hypothetical protein
TGGT1_075330	4	21	hypothetical protein
TGGT1_065600	1	3	conserved hypothetical protein
TGME49_021720	1	3	hypothetical protein, conserved
TGME49_094040	8	824	hypothetical protein
TGGT1_074440	4	24	conserved hypothetical protein

TGGT1_041830	9	154	GTP-binding protein, putative
TGME49_105750	6	55	nucleolar GTP-binding protein NOG2, putative
TGME49_105010	6	60	RNA binding protein, putative
TGGT1_042490	6	59	conserved hypothetical protein
TGME49_106410	6	28	hypothetical protein
TGGT1_040930	4	24	conserved hypothetical protein
TGME49_109010	7	37	hypothetical protein
TGGT1_056050	3	12	conserved hypothetical protein
TGME49_023440	2	6	ATP-dependent RNA helicase, putative
TGME49_115610	11	364	hypothetical protein
TGGT1_093190	4	95	conserved hypothetical protein
TGME49_085850	15	131	FK506-binding nuclear protein, putative
TGGT1_037250	8	108	FK506-binding nuclear protein, putative
TGGT1_003430	5	79	conserved hypothetical protein
TGME49_054150	2	31	hypothetical protein
TGME49_011220	6	108	hypothetical protein
TGGT1_123710	5	106	conserved hypothetical protein
TGGT1_030260	2	7	F-actin capping protein beta subunit, putative
TGME49_019290	2	7	F-actin capping protein beta subunit, putative
TGME49_114470	7	49	hypothetical protein
TGGT1_091260	5	62	conserved hypothetical protein
TGME49_022310	6	30	hypothetical protein, conserved
TGGT1_066130	3	21	conserved hypothetical protein
TGGT1_034460	7	201	hypothetical protein
TGME49_088810	6	163	hypothetical protein
TGGT1_016980	11	186	conserved hypothetical protein
TGME49_003750	10	182	hypothetical protein
TGGT1_088720	1	1	conserved hypothetical protein
TGME49_084550	2	4	EF hand domain-containing protein
TGME49_065220	12	198	co-chaperone GrpE, putative
TGGT1_058000	7	155	co-chaperone GrpE, putative
TGGT1_026600	3	19	conserved hypothetical protein
TGME49_046520	2	9	hypothetical protein
TGGT1_104880	1	9	conserved hypothetical protein
TGME49_051870	15	869	histone H2B, putative
TGGT1_103250	3	56	histone H2B, putative
TGGT1_020580	7	74	programmed cell death, putative
TGME49_007690	6	67	double-stranded DNA-binding domain-containing protein
TGGT1_051650	2	7	conserved hypothetical protein
TGME49_039320	2	7	hypothetical protein, conserved
TGGT1_104480	2	30	conserved hypothetical protein
TGME49_078070	1	13	hypothetical protein
TGME49_086530	11	191	hypothetical protein
TGGT1_037950	3	12	conserved hypothetical protein
TGGT1_089170	2	8	conserved hypothetical protein

TGME49_112960	2	8	hypothetical protein
TGME49_120010	18	457	hypothetical protein
TGGT1_121380	15	441	conserved hypothetical protein
TGME49_042080	5	37	vesicle transport v-SNARE, putative
TGGT1_048340	3	28	vesicle transport v-SNARE domain-containing protein
TGME49_032150	38	884	hypothetical protein, conserved
TGGT1_115500	22	812	conserved hypothetical protein
TGME49_011410	20	590	hypothetical protein, conserved
TGGT1_123380	15	531	ligatin, putative
TGME49_046330	11	135	CRAL/TRIO domain-containing protein
TGGT1_026710	10	126	CRAL/TRIO domain-containing protein, putative
TGGT1_045440	16	971	conserved hypothetical protein
TGME49_044500	13	736	tubulin-tyrosine ligase family protein
TGME49_050330	6	52	hypothetical protein
TGGT1_101940	5	100	conserved hypothetical protein
TGGT1_078020	10	148	inositol monophosphatase domain-containing protein, putative
TGME49_023670	8	120	inositol monophosphatase domain-containing protein
TGGT1_124760	4	42	conserved hypothetical protein
TGME49_014340	4	42	hypothetical protein, conserved
TGME49_112270	21	2529	hypothetical protein
TGGT1_088530	10	291	conserved hypothetical protein
TGGT1_113760	12	248	nuclear movement domain-containing protein, putative
TGME49_033680	12	244	nuclear movement domain-containing protein
TGME49_040860	11	481	1-acyl-sn-glycerol-3-phosphate acyltransferase, putative
TGGT1_049900	8	142	1-acyl-sn-glycerol-3-phosphate acyltransferase, putative
TGME49_007440	34	5710	hypothetical protein
TGGT1_020420	15	1046	40S ribosomal protein S4, putative
TGGT1_114290	1	2	conserved hypothetical protein
TGME49_033200	1	2	hypothetical protein
TGME49_031630	52	101448	membrane skeletal protein IMC1, putative
TGGT1_116040	18	9799	conserved hypothetical protein
TGME49_032170	11	31	hypothetical protein
TGGT1_115470	1	1	conserved hypothetical protein
TGGT1_021140	1	4	oxidoreductase, putative
TGME49_008040	1	4	aldo/keto reductase family oxidoreductase, putative
TGGT1_032030	3	16	helicase, putative
TGME49_091090	3	16	SNF2 family N-terminal domain-containing protein
TGGT1_021410	38	6553	conserved hypothetical protein
TGME49_008400	31	4918	hypothetical protein
TGGT1_014920	22	2201	proline-rich protein, putative
TGME49_020240	13	1336	proline-rich protein
TGME49_062050	22	626	Rhoptry kinase family protein ROP39
TGGT1_008060	7	143	hypothetical protein

TGME49_068830	53	15688	hypothetical protein, conserved
TGGT1_107020	28	2576	conserved hypothetical protein
TGGT1_093910	2	16	conserved hypothetical protein
TGME49_116220	1	15	hypothetical protein, conserved
TGGT1_020400	1	10	hypothetical protein
TGGT1_058410	1	2	protein kinase domain-containing protein, putative
TGME49_037860	1	2	CAM Kinase, RAD family
TGME49_095650	10	269	hypothetical protein
TGGT1_054370	7	243	conserved hypothetical protein
TGME49_118760	3	6	hypothetical protein, conserved
TGGT1_122350	1	2	NBP2B protein, putative
TGME49_118410	45	3104	TCP-1/cpn60 family chaperonin, putative
TGGT1_122820	21	1548	chaperonin containing t-complex protein 1, zeta subunit, tcpz, putative
TGME49_119920	36	6107	dihydrolipoamide branched chain transacylase, E2 subunit, putative
TGGT1_121470	20	1920	lipoamide acyltransferase component of branched-chain alpha-keto dehyrogenase complex, putative
TGME49_020940	10	28	ribosomal RNA large subunit methyltransferase J, putative
TGGT1_066320	4	150	conserved hypothetical protein
TGME49_022710	4	150	hypothetical protein, conserved
TGGT1_038670	4	21	conserved hypothetical protein
TGME49_013830	4	21	hypothetical protein
TGGT1_071410	7	47	conserved hypothetical protein
TGME49_034480	5	36	hypothetical protein
TGGT1_085825	5	24	PX domain-containing protein
TGME49_109870	5	24	PX domain-containing protein
TGME49_100190	5	406	ribosomal protein L37a
TGGT1_095110	2	30	ribosomal protein L37A, putative
TGGT1_065090	2	9	conserved hypothetical protein
TGME49_021250	1	3	hypothetical protein
TGGT1_115070	1	2	DNA repair and recombination protein RAD54, putative
TGME49_032450	1	2	DNA repair protein RAD54, putative
TGGT1_035570	20	897	conserved hypothetical protein
TGME49_083550	17	823	hypothetical protein
TGME49_060300	7	29	zinc transporter ZIP domain-containing protein
TGGT1_009910	5	25	zinc transporter, putative
TGME49_031940	11	297	thiF family domain-containing protein
TGGT1_115820	10	365	conserved hypothetical protein
TGGT1_118160	1	2	conserved hypothetical protein
TGME49_036210	48	26684	mitochondrial-processing peptidase beta subunit, putative
TGGT1_069540	15	1073	mitochondrial processing peptidase beta subunit, putative
TGME49_087500	55	4897	TCP-1/cpn60 family chaperonin, putative
TGGT1_040270	28	3336	chaperonin containing t-complex protein 1, theta subunit, tcpq, putative

TGGT1_062450	2	2	conserved hypothetical protein
TGME49_006340	2	2	hypothetical protein
TGGT1_062040	13	2500	conserved hypothetical protein
TGME49_006690	12	1835	hypothetical protein, conserved
TGGT1_080550	3	13	conserved hypothetical protein
TGME49_025350	3	13	hypothetical protein, conserved
TGME49_070930	60	2418	hypothetical protein
TGGT1_109780	46	2626	conserved hypothetical protein
TGME49_119860	12	112	DNA polymerase epsilon catalytic subunit, putative
TGGT1_121530	9	88	DNA polymerase epsilon catalytic subunit, putative
TGGT1_016840	3	8	conserved hypothetical protein
TGME49_003880	2	5	hypothetical protein
TGME49_057370	8	82	hypothetical protein
TGGT1_013330	4	74	conserved hypothetical protein
TGGT1_032070	3	27	serine/threonine-protein kinase, putative
TGME49_091050	2	16	CAM kinase, SNF1 family
TGME49_114360	3	9	WW domain-containing protein
TGGT1_064320	6	218	pre-mRNA splicing factor, putative
TGME49_121290	6	218	pre-mRNA splicing factor 18, putative
TGGT1_118490	25	2137	conserved hypothetical protein
TGME49_029260	23	2016	hypothetical protein
TGME49_116340	27	4362	hypothetical protein
TGGT1_094150	18	4612	conserved hypothetical protein
TGGT1_038080	1	2	conserved hypothetical protein
TGME49_086680	1	2	hypothetical protein
TGGT1_053710	1	1	conserved hypothetical protein
TGME49_020880	1	1	hypothetical protein
TGME49_065370	7	91	hypothetical protein
TGGT1_057740	2	10	conserved hypothetical protein
TGGT1_008070	3	27	dolichyl glycosyltransferase, putative
TGME49_062030	3	27	ALG6, ALG8 glycosyltransferase family domain-containing protein
TGGT1_029840	2	7	TBC domain-containing protein, putative
TGME49_018870	2	7	TBC domain-containing protein
TGME49_052640	11	112	plasma-membrane H+-ATPase, putative
TGGT1_006350	6	90	conserved hypothetical protein
TGME49_099130	4	43	hypothetical protein
TGME49_002630	68	7404	AFG3 ATPase family protein
TGGT1_018340	13	210	peptidase M14 family protein
TGME49_114030	14	7615	hypothetical protein
TGGT1_090670	1	1	conserved hypothetical protein
TGGT1_080810	6	28	26S proteasome non-ATPase regulatory subunit, putative
TGME49_025580	6	28	26S Proteasome non-ATPase regulatory subunit 9, putative
TGME49_049630	20	904	glutaredoxin domain containing protein

TGGT1_101300	12	334	conserved hypothetical protein
TGME49_084560	20	4584	60S ribosomal protein L9, putative
TGGT1_036380	9	613	60S ribosomal protein L9, putative
TGGT1_113510	12	158	NBP2B protein, putative
TGME49_033820	10	120	DNA polymerase delta small subunit, putative
TGME49_024350	78	9412	aminopeptidase N, putative
TGGT1_078860	56	9062	aminopeptidase N, putative
TGGT1_026440	5	25	conserved hypothetical protein
TGME49_046670	5	25	hypothetical protein
TGGT1_046070	1	2	conserved hypothetical protein
TGME49_043790	1	2	hypothetical protein, conserved
TGGT1_022040	11	567	conserved hypothetical protein
TGME49_009100	8	249	hypothetical protein
TGME49_051470	1	27	hypothetical protein
TGME49_029990	38	4769	TCP-1/cpn60 family chaperonin, putative
TGGT1_117890	25	2843	chaperonin containing t-complex protein 1, alpha subunit, tcpa, putative
TGGT1_045960	2	4	DEAD/DEAH box helicase, putative
TGME49_041880	27	892	3', 5'-cyclic nucleotide phosphodiesterase, putative
TGGT1_048720	15	712	3', 5'-cyclic nucleotide phosphodiesterase, putative
TGGT1_046340	1	1	conserved hypothetical protein
TGME49_043520	1	1	hypothetical protein, conserved
TGGT1_016360	1	2	GTPase mss1/trme, putative
TGME49_004350	1	2	hypothetical protein
TGME49_110320	17	427	calnexin, putative
TGGT1_086550	3	29	calnexin, putative
TGGT1_039690	28	3132	conserved hypothetical protein
TGME49_013040	27	2126	hypothetical protein
TGME49_026070	34	656	hypothetical protein, conserved
TGGT1_081220	14	217	sphingomyelin phosphodiesterase, putative
TGME49_022170	11	400	dense-granule antigen DG32, putative
TGGT1_065990	6	68	dense-granule antigen DG32, putative
TGGT1_021440	9	98	transporter, putative
TGME49_008420	9	98	transporter, putative
TGGT1_020070	18	915	conserved hypothetical protein
TGME49_001250	17	835	chypothetical protein, conserved
TGGT1_095050	5	55	hypothetical protein
TGME49_100240	4	40	syntaxin family protein, putative
TGGT1_034340	4	14	hypothetical protein
TGME49_088940	4	14	hypothetical protein
TGGT1_065350	5	52	PUA domain-containing, cell cycle regulator protein, putative
TGME49_021490	5	52	PUA domain-containing, cell cycle regulator protein, putative
TGME49_097530	16	169	DNA-directed RNA polymerase I subunit RPA2, putative
TGGT1_096660	14	250	DNA-directed RNA polymerase I, beta subunit, putative

TGME49_063300	53	117739	porin, putative
TGGT1_043420	16	1567	porin, putative
TGME49_055350	18	388	hypothetical protein
TGGT1_076770	15	375	conserved hypothetical protein
TGME49_046560	13	187	vacuolar ATP synthase subunit G1, putative
TGGT1_026550	4	68	vacuolar ATP synthase subunit G1, putative
TGGT1_003190	14	571	conserved hypothetical protein
TGME49_053910	14	486	hypothetical protein
TGME49_031440	43	4082	hypothetical protein
TGGT1_116230	39	5013	conserved hypothetical protein
TGGT1_122510	2	8	DNA/pantothenate metabolism flavoprotein domain-containing protein, putative
TGME49_118600	2	8	DNA/pantothenate metabolism flavoprotein domain-containing protein
TGME49_101140	2	4	SRS19A
TGME49_014290	31	2429	intracellular protease, putative
TGGT1_124710	14	1509	intracellular protease, putative
TGME49_000400	16	285	protein phosphatase 2A, putative
TGGT1_060620	13	296	conserved hypothetical protein
TGME49_037260	10	43	hypothetical protein
TGGT1_068650	4	14	conserved hypothetical protein
TGME49_119710	13	356	kinesin motor domain-containing protein
TGGT1_121580	12	352	kinesin motor domain-containing protein, putative
TGME49_088720	22	3138	60S ribosomal protein L10, putative
TGGT1_034550	11	739	60S ribosomal protein L10, putative
TGME49_024890	37	3077	hypothetical protein, conserved
TGGT1_079800	35	3091	conserved hypothetical protein
TGGT1_037180	4	40	conserved hypothetical protein
TGME49_085780	2	15	hypothetical protein
TGGT1_101105	1	3	hypothetical protein
TGME49_049450	1	3	hypothetical protein
TGGT1_113710	1	4	conserved hypothetical protein
TGME49_033730	1	4	hypothetical protein
TGME49_105210	3	8	hypothetical protein
TGGT1_042290	2	4	conserved hypothetical protein
TGME49_021620	116	257799	tubulin beta chain, putative
TGGT1_065490	33	9750	tubulin beta chain, putative
TGME49_058090	41	4372	hypothetical protein
TGGT1_012120	17	389	conserved hypothetical protein
TGGT1_049500	3	10	hemolysin, putative
TGME49_041160	3	10	hemolysin, putative
TGME49_021470	61	6356	hypothetical protein
TGGT1_065330	24	2801	hypothetical protein
TGGT1_025030	1	7	conserved hypothetical protein
TGME49_047780	1	7	hypothetical protein

TGGT1_079310	4	17	conserved hypothetical protein
TGME49_024590	4	17	hypothetical protein
TGME49_066460	4	25	ubiquitin-like protein SMT3 precursor, putative
TGGT1_076410	2	16	hypothetical protein
TGGT1_108760	1	1	conserved hypothetical protein
TGME49_070230	1	1	hypothetical protein
TGME49_070360	6	82	hypothetical protein
TGGT1_108890	2	30	conserved hypothetical protein
TGME49_055690	5	22	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase domain-containing protein
TGGT1_077200	2	12	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, putative
TGGT1_115920	4	28	tola protein, putative
TGME49_031840	4	28	tola protein., putative
TGGT1_021090	2	21	hypothetical protein
TGME49_008000	2	21	hypothetical protein
TGME49_115180	8	58	hypothetical protein
TGGT1_092610	2	14	hypothetical protein
TGGT1_027530	3	113	conserved hypothetical protein
TGME49_045550	1	9	hypothetical protein
TGME49_019590	24	958	ruvB-like 1, putative
TGGT1_030480	16	616	conserved hypothetical protein
TGME49_106390	6	153	hypothetical protein
TGGT1_040950	2	32	conserved hypothetical protein
TGME49_106590	9	55	hypothetical protein
TGGT1_040750	5	25	hypothetical protein
TGME49_001400	27	4866	hypothetical protein
TGGT1_020020	25	4823	conserved hypothetical protein
TGGT1_107790	30	2172	serine/threonine protein phosphatase domain-containing protein, putative
TGME49_069460	22	1347	serine/threonine protein phosphatase domain-containing protein
TGME49_033890	14	486	hypothetical protein
TGGT1_113450	14	428	conserved hypothetical protein
TGME49_000320	32	2405	hypoxanthine-xanthine-guanine phosphoribosyl transferase
TGGT1_060540	13	724	hypoxanthine-xanthine-guanine phosphoribosyl transferase, putative
TGME49_056970	71	8447	vacuolar ATP synthase catalytic subunit A, putative
TGGT1_013780	36	4319	vacuolar ATP synthase, putative
TGME49_110790	32	1289	hypothetical protein
TGGT1_087110	6	362	conserved hypothetical protein
TGME49_031480	105	14370	translational activator, putative
TGGT1_116090	79	12730	translational activator, putative
TGME49_070010	7	94	hypothetical protein
TGGT1_108550	7	93	conserved hypothetical protein
TGGT1_079940	15	488	conserved hypothetical protein

TGME49_025000	13	338	hypothetical protein
TGME49_022960	6	26	SCY kinase-related protein (incomplete catalytic triad)
TGGT1_066500	5	22	protein kinase domain-containing protein, putative
TGME49_055900	3	50	hypothetical protein, conserved
TGME49_055440	7	362	hypothetical protein, conserved
TGGT1_076860	6	358	conserved hypothetical protein
TGGT1_096870	1	2	conserved hypothetical protein
TGME49_005180	9	51	U1 small nuclear ribonucleoprotein, putative
TGGT1_063850	6	31	U1 small nuclear ribonucleoprotein 70 kD, putative
TGME49_069180	173	34152	MIF4G domain-containing protein
TGGT1_107380	91	34601	MIF4G domain-containing protein, putative
TGGT1_046460	1	2	PX domain-containing protein, putative
TGME49_043400	1	2	PX domain-containing protein
TGGT1_051330	12	192	serine-pyruvate aminotransferase, putative
TGME49_039530	12	192	alanine--glyoxylate aminotransferase, putative
TGGT1_018010	6	18	para-aminobenzoate synthase, putative
TGME49_002920	6	18	para-aminobenzoate synthase, putative
TGGT1_019350	1	2	TPR domain-containing protein, putative
TGME49_001870	1	2	TPR domain-containing protein
TGGT1_044290	4	12	ubiquitin-conjugating enzyme domain-containing protein, putative
TGME49_064050	4	12	ubiquitin-conjugating enzyme domain-containing protein
TGGT1_113910	6	38	integral membrane transporter, putative
TGME49_033540	6	38	integral membrane transporter
TGME49_048330	1	1	zinc finger (C3HC4 RING finger) protein, putative
TGME49_016650	51	2876	hypothetical protein, conserved
TGGT1_098490	38	1833	conserved hypothetical protein
TGME49_031800	7	30	hypothetical protein
TGME49_086230	36	1877	hypothetical protein
TGGT1_037630	35	2698	conserved hypothetical protein
TGGT1_114740	1	1	conserved hypothetical protein
TGME49_032780	1	1	hypothetical protein
TGGT1_014350	1	6	WD-40 repeat-containing protein, putative
TGGT1_016810	29	3905	TBC domain-containing protein, putative
TGME49_003910	21	2650	TBC domain-containing protein
TGME49_054900	10	346	proteasome subunit beta type 2, putative
TGGT1_004470	9	339	proteasome subunit beta type, putative
TGGT1_037550	4	54	conserved hypothetical protein
TGME49_086140	4	54	hypothetical protein
TGME49_080370	33	4611	KH domain containing protein
TGGT1_072950	24	4143	KH domain containing protein, putative
TGGT1_089630	13	222	DEAD/DEAH box helicase, putative
TGME49_113240	11	187	DEAD/DEAH box helicase, putative
TGGT1_021080	16	322	ribosomal RNA methyltransferase, putative

TGME49_007990	11	180	ftsJ-like methyltransferase domain-containing protein
TGME49_090870	12	265	patched family domain-containing protein, conserved
TGGT1_032270	6	306	patched transmembrane domain-containing protein, putative
TGGT1_123510	2	3	conserved hypothetical protein
TGME49_011300	2	3	hypothetical protein
TGGT1_064060	1	1	conserved hypothetical protein
TGME49_004890	1	1	hypothetical protein
TGME49_094060	14	49	hypothetical protein
TGGT1_038560	4	26	conserved hypothetical protein
TGME49_013920	4	26	hypothetical protein
TGGT1_005560	12	523	HMG box domain-containing protein / high mobility group box domain-containing protein, putative
TGME49_081900	11	408	high mobility group (HMG) box domain-containing protein
TGME49_008340	10	84	hypothetical protein
TGGT1_021340	1	9	hypothetical protein
TGGT1_010120	1	1	conserved hypothetical protein
TGME49_060000	1	1	hypothetical protein, conserved
TGME49_018410	66	16468	60S acidic ribosomal protein P0
TGGT1_029230	19	2264	hypothetical protein
TGGT1_088320	1	2	conserved hypothetical protein
TGME49_112080	1	2	hypothetical protein
TGME49_058830	80	6078	hypothetical protein
TGGT1_011340	16	336	suppressor of actin mutations 2/vacuolar sorting protein, putative
TGGT1_112200	5	33	phosphoglycerate mutase protein, putative
TGME49_073030	4	22	phosphoglycerate mutase protein, putative
TGME49_116900	12	180	hypothetical protein
TGGT1_094640	7	92	something about silencing protein sas10, putative
TGME49_058210	29	882	DNA-directed RNA polymerase II subunit RPB2, putative
TGGT1_012000	24	857	DNA-directed RNA polymerase II subunit, putative
TGME49_091870	8	29	hypothetical protein
TGGT1_031670	5	23	conserved hypothetical protein
TGGT1_031850	7	64	MYND domain containing protein, putative
TGME49_091590	7	64	zinc finger MYND domain-containing protein
TGGT1_104910	16	992	conserved hypothetical protein
TGME49_077740	15	887	hypothetical protein
TGME49_085680	5	18	dihydrolipoamide acyltransferase, putative
TGGT1_037080	4	36	biotein requiring domain-containing protein, putative
TGGT1_083350	10	136	tRNA-dihydrouridine synthase, putative
TGME49_028080	10	136	hypothetical protein
TGME49_039340	2	4	hypothetical protein
TGGT1_043840	2	22	conserved hypothetical protein
TGME49_063600	2	22	hypothetical protein

TGME49_044100	11	299	putative snoRNA binding domain-containing protein
TGGT1_045850	9	286	pre-mRNA splicing factor prp31, putative
TGME49_085270	9	53	DnaJ domain-containing protein
TGGT1_036770	8	44	DnaJ domain-containing protein, putative
TGME49_080550	7	47	coatomer zeta-2 subunit, putative
TGGT1_072760	6	34	coatomer zeta-2 subunit, putative
TGME49_034550	1	24	60S ribosomal protein L29, putative
TGGT1_111610	4	21	1-O-acylceramide synthase, putative
TGME49_072420	4	21	1-O-acylceramide synthase, putative
TGGT1_112330	2	35	conserved hypothetical protein
TGME49_073160	2	35	hypothetical protein
TGME49_070560	5	18	cyclophilin, putative
TGGT1_109110	4	14	cyclophilin, putative
TGME49_069890	84	18201	M16 family peptidase, putative
TGGT1_108330	27	2681	insulysin, putative
TGME49_036020	1	1	BT1 transmembrane domain-containing protein
TGGT1_000460	3	12	conserved hypothetical protein
TGME49_075980	1	2	ubiquinone biosynthesis protein COQ4, putative
TGME49_066120	8	105	glutathione/thioredoxin peroxidase, putative
TGGT1_056630	6	57	glutathione/thioredoxin peroxidase, putative
TGGT1_034440	1	2	conserved hypothetical protein
TGME49_088820	1	2	hypothetical protein
TGME49_119510	13	345	hypothetical protein, conserved
TGGT1_121790	6	58	conserved hypothetical protein
TGGT1_109980	11	100	conserved hypothetical protein
TGME49_071100	5	44	hypothetical protein
TGGT1_070190	1	1	conserved hypothetical protein
TGME49_035590	1	1	hypothetical protein
TGGT1_040290	40	3510	conserved hypothetical protein
TGME49_087480	34	2894	hypothetical protein
TGME49_089490	14	98	hypothetical protein
TGGT1_033730	5	40	protein phosphatase 2C, putative
TGGT1_112900	1	6	conserved hypothetical protein
TGME49_121680	42	4025	nuclear RNA binding protein, putative
TGGT1_064710	29	2240	conserved hypothetical protein
TGGT1_044440	2	5	conserved hypothetical protein
TGME49_064200	2	3	hypothetical protein
TGGT1_087490	3	18	conserved hypothetical protein
TGME49_111180	3	18	hypothetical protein
TGGT1_094960	10	151	conserved hypothetical protein
TGME49_100330	8	71	hypothetical protein
TGGT1_011470	1	3	conserved hypothetical protein
TGME49_058720	1	3	ubiquitin-like protein 5, putative
TGME49_032000	13	173	hypothetical protein
TGGT1_115760	9	67	conserved hypothetical protein

TGME49_020360	25	2648	FAD-depdendent monooxygenase, putative
TGGT1_015060	22	2785	FAD-depdendent monooxygenase, putative
TGGT1_079900	10	169	nnp-1 protein, putative
TGME49_024970	8	135	Nucleolar protein,Nop52 domain containing protein
TGME49_007910	4	26	manganese resistance 1 protein, putative
TGGT1_020900	3	21	manganese resistance 1 protein, putative
TGGT1_044060	2	6	hypothetical protein
TGME49_051760	12	411	proteasome activator subunit, putative
TGGT1_103130	10	407	proteasome activator subunit, putative
TGGT1_013720	1	2	conserved hypothetical protein
TGME49_057020	1	2	hypothetical protein
TGGT1_115010	6	70	brca1-associated protein, putative
TGME49_032510	5	55	hypothetical protein
TGGT1_107340	1	7	DHHC domain-containing protein, putative
TGME49_069150	1	7	zinc finger DHHC domain-containing protein
TGGT1_029930	43	4695	conserved hypothetical protein
TGME49_018960	37	3330	hypothetical protein
TGGT1_039770	13	694	KH domain-containing protein, putative
TGME49_012980	12	612	KH domain-containing protein
TGGT1_079160	6	25	conserved hypothetical protein
TGME49_024540	5	15	hypothetical protein
TGME49_097770	4	8	hypothetical protein, conserved
TGGT1_096430	3	7	conserved hypothetical protein
TGME49_038930	2	5	hypothetical protein
TGGT1_052050	1	4	conserved hypothetical protein
TGGT1_066530	2	6	hypothetical protein
TGME49_022990	2	6	ribosomal protein S18 domain containing protein
TGME49_052070	3	5	hypothetical protein
TGME49_058930	9	88	peptidyl-prolyl cis-trans isomerase NIMA-interacting 1, putative
TGGT1_010940	5	57	peptidyl-prolyl cis-trans isomerase NIMA-interacting, putative
TGME49_108040	17	368	zinc-finger protein ZPR1, putative
TGGT1_042760	15	422	zinc finger protein ZPR1, putative
TGME49_008590	14	318	vacuolar ATP synthase subunit H, putative
TGGT1_021620	11	302	vacuolar ATP synthase subunit H, putative
TGME49_070810	1	1	hypothetical protein
TGME49_115250	12	148	melanocyte proliferating gene 1, putative
TGGT1_092680	9	176	conserved hypothetical protein
TGGT1_002040	3	11	conserved hypothetical protein
TGME49_053070	2	4	hydrolase, TatD family domain-containing protein
TGME49_059660	15	332	orotate phosphoribosyltransferase, putative
TGGT1_010360	11	258	orotate phosphoribosyltransferase, putative
TGME49_034410	33	349	hypothetical protein
TGGT1_071480	4	92	conserved hypothetical protein

TGME49_088990	4	10	hypothetical protein, conserved
TGME49_060310	22	1338	ATP-binding cassette protein subfamily B member 1
TGGT1_009900	6	25	multidrug resistance protein / ABC transporter, putative
TGGT1_082070	1	2	ABC transporter, putative
TGGT1_020110	1	2	conserved hypothetical protein
TGME49_055420	26	1876	hypothetical protein
TGGT1_076840	13	596	conserved hypothetical protein
TGGT1_116530	1	1	conserved hypothetical protein
TGME49_031180	1	1	hypothetical protein
TGGT1_032000	6	62	conserved hypothetical protein
TGME49_091120	6	55	hypothetical protein
TGME49_064180	2	4	hypothetical protein
TGME49_009730	3	6	hypothetical protein
TGGT1_022790	2	4	conserved hypothetical protein
TGGT1_070020	5	152	calcium/calmodulin-dependent protein kinase type I, putative
TGME49_035750	5	152	ULK kinase
TGGT1_000350	2	82	aminoalcoholphosphotransferase, putative
TGME49_076190	2	82	CDP-alcohol phosphatidyltransferase domain-containing protein
TGME49_026590	11	271	hypothetical protein
TGGT1_081740	2	24	conserved hypothetical protein
TGME49_066090	15	549	hypothetical protein, conserved
TGGT1_056670	7	201	conserved hypothetical protein
TGGT1_002140	2	8	conserved hypothetical protein
TGME49_053170	1	1	zinc carboxypeptidase, putative
TGME49_069920	18	126	phosphatidylserine decarboxylase proenzyme, putative
TGGT1_108460	5	30	phosphatidylserine decarboxylase proenzyme, putative
TGGT1_080750	3	18	zinc-iron transporter, putative
TGME49_025530	2	7	ZIP Zinc transporter domain-containing protein
TGME49_067330	40	1866	fumarase, putative
TGGT1_075710	24	1159	fumarase, putative
TGME49_115670	18	294	HEAT repeat containing protein
TGGT1_093350	12	268	HEAT repeat containing protein, putative
TGGT1_005380	4	83	conserved hypothetical protein
TGME49_082090	4	83	hypothetical protein
TGGT1_033680	7	68	WD-repeat protein, putative
TGME49_089550	6	66	hypothetical protein, conserved
TGGT1_084220	30	2126	protein transport protein sec7, putative
TGME49_028650	24	1584	sec7 domain-containing protein
TGME49_089950	10	49	hypothetical protein
TGGT1_033190	7	44	conserved hypothetical protein
TGME49_078260	5	46	hypothetical protein
TGGT1_104290	4	10	conserved hypothetical protein
TGGT1_111400	2	12	conserved hypothetical protein

TGME49_072190	1	4	hypothetical protein
TGME49_083510	34	7384	hypothetical protein
TGGT1_035530	25	7057	conserved hypothetical protein
TGGT1_035840	2	14	conserved hypothetical protein
TGME49_083810	1	8	hypothetical protein
TGME49_108870	28	2105	zinc finger (CCCH type) protein, putative
TGGT1_055900	24	3103	zinc finger (CCCH type) protein
TGME49_031400	2	3	tubulin alpha chain, putative
TGGT1_104550	16	839	conserved hypothetical protein
TGME49_077990	16	839	OTU-like cysteine protease domain-containing protein
TGGT1_104080	5	46	conserved hypothetical protein
TGME49_078470	2	11	hypothetical protein
TGME49_036160	3	29	hypothetical protein
TGGT1_069590	1	2	conserved hypothetical protein
TGGT1_093800	5	38	queuine tRNA-ribosyltransferase, putative
TGME49_116100	4	33	queuine tRNA-ribosyltransferase, putative
TGGT1_001150	1	2	conserved hypothetical protein
TGME49_052230	1	2	hypothetical protein
TGGT1_032110	2	6	elongin c, putative
TGME49_091010	2	6	elongin c, putative
TGGT1_078200	18	433	ATP-citrate synthase, putative
TGME49_023840	16	319	ATP-citrate synthase, putative
TGME49_019080	4	41	hypothetical protein
TGGT1_030050	3	31	conserved hypothetical protein
TGME49_035380	19	823	hypothetical protein
TGGT1_070610	15	772	conserved hypothetical protein
TGGT1_010770	1	1	conserved hypothetical protein
TGME49_023030	2	5	hypothetical protein
TGGT1_048850	31	2752	conserved hypothetical protein
TGME49_041850	30	2302	hypothetical protein
TGGT1_093870	12	205	FF domain-containing protein, putative
TGME49_116180	10	155	FF domain-containing protein
TGGT1_017680	14	556	conserved hypothetical protein
TGME49_003190	9	321	hypothetical protein
TGME49_062430	17	166	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, putative
TGGT1_007680	5	40	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, putative
TGGT1_057660	13	495	conserved hypothetical protein
TGME49_065440	13	495	hypothetical protein
TGME49_069960	13	218	14-3-3 protein, putative
TGGT1_108500	10	182	14-3-3 protein, putative
TGGT1_106970	4	30	hypothetical protein
TGME49_068790	3	21	hypothetical protein
TGGT1_027020	9	94	MIF4G domain-containing protein, putative

TGME49_046040	7	68	MIF4G domain-containing protein
TGGT1_047020	1	1	flavoprotein domain-containing protein, putative
TGME49_042880	1	1	flavoprotein domain-containing protein
TGGT1_098720	4	16	hypothetical protein
TGME49_016530	2	3	ribosome recycling factor domain-containing protein
TGGT1_088630	3	24	conserved hypothetical protein
TGME49_112350	3	24	RWD domain-containing protein
TGME49_039270	3	6	hypothetical protein
TGGT1_051700	1	4	conserved hypothetical protein
TGME49_070710	6	44	hypothetical protein
TGGT1_109470	6	40	conserved hypothetical protein
TGME49_012300	57	12825	hypothetical protein
TGGT1_072150	18	750	conserved hypothetical protein
TGGT1_002490	1	2	conserved hypothetical protein
TGME49_053410	1	2	hypothetical protein
TGGT1_075320	3	4	conserved hypothetical protein
TGME49_094970	3	4	hypothetical protein
TGGT1_036560	1	5	conserved hypothetical protein
TGME49_084640	1	5	hypothetical protein
TGGT1_073570	4	34	DHHC domain-containing protein, putative
TGME49_093220	4	34	zinc finger DHHC domain-containing protein
TGGT1_088800	8	50	conserved hypothetical protein
TGME49_112510	8	47	hypothetical protein
TGGT1_004510	2	8	conserved hypothetical protein
TGME49_054930	2	8	hypothetical protein, conserved
TGGT1_094410	3	28	conserved hypothetical protein
TGME49_116600	3	28	hypothetical protein
TGGT1_068200	8	202	conserved hypothetical protein
TGME49_075310	7	129	hypothetical protein
TGGT1_120510	2	2	hypothetical protein
TGME49_120760	2	2	hypothetical protein
TGME49_109290	7	55	HD domain-containing protein
TGGT1_056430	4	48	HD domain-containing protein, putative
TGGT1_081510	34	2653	conserved hypothetical protein
TGME49_026340	30	2166	hypothetical protein, conserved
TGGT1_023360	5	32	conserved hypothetical protein
TGME49_007120	5	17	hypothetical protein, conserved
TGGT1_013200	3	33	conserved hypothetical protein
TGME49_057500	2	18	hypothetical protein
TGGT1_118030	26	761	conserved hypothetical protein
TGME49_029750	20	327	hypothetical protein
TGGT1_031210	1	2	conserved hypothetical protein
TGME49_056840	21	342	hypothetical protein, conserved
TGGT1_013910	10	238	conserved hypothetical protein
TGGT1_087350	1	7	conserved hypothetical protein

TGME49_111040	1	7	hypothetical protein
TGGT1_003690	1	2	conserved hypothetical protein
TGME49_054270	1	2	hypothetical protein
TGME49_004340	43	2003	hypothetical protein
TGGT1_016370	26	1387	conserved hypothetical protein
TGGT1_050490	20	1144	conserved hypothetical protein
TGME49_040380	19	895	hypothetical protein
TGGT1_108780	49	5114	hypothetical protein
TGME49_070250	42	5128	dense granule protein 1 / major antigenp24
TGGT1_123250	2	2	conserved hypothetical protein
TGME49_011640	2	2	hypothetical protein
TGGT1_013520	1	2	hypothetical protein
TGME49_110210	9	152	hypothetical protein
TGGT1_086450	3	21	conserved hypothetical protein
TGME49_044250	20	1371	hypothetical protein
TGGT1_045690	17	1568	microneme protein-1, putative
TGME49_077760	49	1577	hypothetical protein, conserved
TGGT1_104890	24	1222	conserved hypothetical protein
TGME49_105880	1	1	PT repeat-containing protein
TGGT1_098400	26	1016	DNA replication licensing factor, putative
TGME49_016730	24	954	DNA replication licensing factor, putative
TGME49_089750	16	3551	ubiquitin / ribosomal protein CEP52 fusion protein, putative
TGME49_043960	50	9025	ras-GTPase-activating protein binding protein, putative
TGGT1_046000	39	8708	ras-GTPase-activating protein binding protein, putative
TGGT1_101420	2	10	hypothetical protein
TGME49_049740	2	10	hypothetical protein
TGME49_065840	28	1249	200 kDa antigen p200, putative
TGGT1_057040	23	1316	hypothetical protein
TGME49_009260	31	5031	cytochrome c oxidase, putative
TGGT1_022310	10	485	hypothetical protein
TGGT1_089650	34	3187	conserved hypothetical protein
TGME49_113260	31	3032	hypothetical protein
TGME49_009280	13	317	hypothetical protein
TGGT1_022320	12	286	conserved hypothetical protein
TGME49_070050	2	11	hypothetical protein
TGGT1_108570	1	10	conserved hypothetical protein
TGME49_005720	8	19	adenosine/AMP deaminase domain containing protein
TGGT1_063190	2	4	adenosine deaminase, putative
TGGT1_092960	1	5	hypothetical protein
TGME49_095330	26	1149	hypothetical protein
TGGT1_054730	8	100	conserved hypothetical protein
TGGT1_050590	14	1230	MACRO domain-containing protein
TGME49_040250	7	605	appr-1-p processing enzyme family domain-containing protein

TGME49_033010	15	1039	CMGC kinase, MAPK family TgMAPK2
TGGT1_114510	14	1035	serine/threonine-protein kinase / mitogen-activated protein kinase, putative
TGME49_027840	18	354	hypothetical protein
TGGT1_083140	11	259	conserved hypothetical protein
TGGT1_079730	1	1	conserved hypothetical protein
TGME49_024830	1	1	hypothetical protein, conserved
TGGT1_093610	3	7	DNA-directed RNA polymerase II 13.3 kDa polypeptide, putative
TGME49_115920	2	3	RNA polymerase II subunit RPB11, putative
TGGT1_004360	1	1	conserved hypothetical protein
TGME49_054810	1	1	hypothetical protein, conserved
TGGT1_075770	5	105	natural resistance-associated macrophage protein, putative
TGME49_067270	4	85	divalent metal transporter, putative
TGME49_055740	13	200	hypothetical protein
TGGT1_077250	11	196	conserved hypothetical protein
TGGT1_106920	1	3	CHCH domain-containing protein, putative
TGME49_068740	1	3	CHCH domain-containing protein
TGME49_047410	8	220	hypothetical protein
TGGT1_025490	4	78	conserved hypothetical protein
TGME49_091960	54	5298	Rhoptry kinase family protein ROP40 (incomplete catalytic triad)
TGGT1_031600	27	2909	conserved hypothetical protein
TGME49_035020	50	4482	coatomer protein complex subunit beta, putative
TGGT1_071090	42	4149	coatomer protein complex subunit beta, putative
TGME49_005060	3	6	hypothetical protein
TGGT1_063980	1	4	conserved hypothetical protein
TGGT1_010330	2	11	conserved hypothetical protein
TGME49_059700	2	11	hypothetical protein
TGGT1_106490	13	385	conserved hypothetical protein
TGME49_068430	12	330	hypothetical protein
TGGT1_121270	5	46	hypothetical protein
TGME49_120120	5	46	bis(5'-adenosyl)-triphosphatase, putative
TGME49_097130	1	4	hypothetical protein
TGGT1_097090	1	3	conserved hypothetical protein
TGME49_024210	14	536	pseudouridylate synthase 1, putative
TGGT1_078700	9	324	pseudouridylate synthase, putative
TGME49_093820	21	672	calpain family cysteine protease domain-containing protein
TGGT1_074200	15	798	calpain, putative
TGME49_094700	7	53	hypothetical protein
TGGT1_075020	5	48	conserved hypothetical protein
TGME49_062910	17	343	NADH-cytochrome B5 reductase, putative
TGGT1_007080	8	122	NADH-cytochrome B5 reductase, putative
TGGT1_032120	5	87	PHD-finger domain-containing protein, putative
TGME49_091000	4	75	PHD-finger domain-containing protein

TGGT1_060660	1	2	hypothetical protein
TGME49_000440	1	2	hypothetical protein
TGME49_031640	83	344533	membrane skeletal protein IMC1
TGGT1_116030	31	10515	membrane skeletal protein IMC1, putative
TGME49_033090	4	6	exonuclease, putative
TGGT1_106840	5	36	conserved hypothetical protein
TGME49_068660	5	36	hypothetical protein
TGME49_078630	8	54	hypothetical protein
TGGT1_103930	6	43	conserved hypothetical protein
TGGT1_103950	1	1	conserved hypothetical protein
TGME49_078610	1	1	hypothetical protein
TGME49_027810	42	3947	Rhoptry kinase family protein ROP11 (incomplete catalytic triad)
TGGT1_083110	21	1423	conserved hypothetical protein
TGME49_095730	14	206	TPR domain-containing protein
TGGT1_054290	9	186	TPR domain-containing protein, putative
TGGT1_011570	1	1	conserved hypothetical protein
TGME49_058630	1	1	hypothetical protein
TGME49_087160	29	2959	kinesin motor domain-containing protein, putative
TGGT1_038260	14	442	kinesin central motor, putative
TGGT1_011720	4	24	conserved hypothetical protein
TGME49_058490	2	10	hypothetical protein
TGGT1_002700	1	1	conserved hypothetical protein
TGGT1_011980	4	20	conserved hypothetical protein
TGME49_058230	2	6	Rhoptry kinase family protein ROP20
TGGT1_100510	6	62	conserved hypothetical protein
TGME49_048840	6	51	membrane transporter PFB0275w
TGGT1_025500	1	1	conserved hypothetical protein
TGME49_094790	15	1145	hypothetical protein
TGGT1_075110	11	394	conserved hypothetical protein
TGGT1_065880	11	258	conserved hypothetical protein
TGME49_022050	11	196	hypothetical protein
TGME49_025840	6	30	hypothetical protein
TGGT1_080980	5	33	conserved hypothetical protein
TGGT1_095070	2	6	conserved hypothetical protein
TGME49_100220	2	6	hypothetical protein
TGME49_043730	33	6342	p36 protein
TGGT1_046140	20	2194	p36 protein, putative
TGME49_111090	35	2855	ubiquitin carboxyl-terminal hydrolase, putative
TGGT1_087410	32	2824	ubiquitin carboxyl-terminal hydrolase, putative
TGGT1_017470	3	32	conserved hypothetical protein
TGME49_003360	3	32	hypothetical protein
TGGT1_073870	1	4	conserved hypothetical protein
TGME49_093490	1	4	hypothetical protein

TGGT1_082080	25	977	SET domain-containing protein / bromodomain-containing protein, putative
TGME49_026810	18	525	SET domain-containing protein / Bromodomain-containing protein
TGGT1_050890	14	478	conserved hypothetical protein
TGME49_039940	11	316	hypothetical protein
TGGT1_056880	3	18	conserved hypothetical protein
TGME49_065990	3	18	hypothetical protein
TGME49_094330	6	33	EGF-like domain-containing protein
TGGT1_074630	4	20	conserved hypothetical protein
TGME49_090260	6	69	hypothetical protein, conserved
TGGT1_032780	4	95	conserved hypothetical protein
TGME49_017600	13	170	CAM kinase, CDPK family CAM kinase TgTPK5
TGGT1_028460	10	176	calmodulin-domain protein kinase, putative
TGME49_069320	6	16	hypothetical protein
TGME49_112120	4	19	hypothetical protein
TGGT1_125970	11	292	conserved hypothetical protein
TGME49_015230	10	136	splicing factor U2AF-associated protein, related
TGME49_101410	25	2386	hypothetical protein
TGGT1_059840	24	2385	conserved hypothetical protein
TGGT1_043000	1	3	ectonucleoside triphosphate diphosphohydrolase, putative
TGME49_107800	1	3	ectonucleoside triphosphate diphosphohydrolase, putative
TGME49_054870	13	1555	hypothetical protein
TGGT1_004440	13	1546	conserved hypothetical protein
TGGT1_114770	1	1	conserved hypothetical protein
TGME49_032750	1	1	hypothetical protein
TGME49_070650	4	11	deoxyribose-phosphate aldolase, putative
TGGT1_109300	2	5	deoxyribose-phosphate aldolase, putative
TGME49_115160	1	1	hypothetical protein
TGME49_016460	10	258	hypothetical protein
TGGT1_098780	9	246	conserved hypothetical protein
TGGT1_114050	1	4	conserved hypothetical protein
TGME49_065250	68	10628	alpha-1 type II collagen, putative
TGGT1_057870	63	12583	conserved hypothetical protein
TGME49_033190	19	453	hypothetical protein, conserved
TGGT1_114300	12	417	nuclear protein skip, putative
TGGT1_005400	5	99	conserved hypothetical protein
TGME49_082070	4	64	hypothetical protein
TGME49_030980	7	17	myosin head motor domain-containing protein TgMyo I
TGGT1_116740	3	9	hypothetical protein
TGGT1_050190	9	147	S-adenosylmethionine synthetase, putative
TGME49_040690	9	147	S-adenosylmethionine synthetase, putative
TGGT1_099420	1	1	conserved hypothetical protein
TGME49_098620	1	1	hypothetical protein
TGME49_095370	8	24	hypothetical protein

TGGT1_054570	2	4	conserved hypothetical protein
TGGT1_050950	8	297	AMP-activated protein kinase subunit, putative
TGME49_039870	7	245	AMP-activated protein kinase subunit, putative
TGGT1_053740	16	524	conserved hypothetical protein
TGME49_020910	12	349	hypothetical protein
TGGT1_120700	8	60	conserved hypothetical protein
TGME49_120590	8	60	glycosyl hydrolases family 35 domain-containing protein
TGME49_057300	9	89	hypothetical protein
TGGT1_013510	6	99	conserved hypothetical protein
TGGT1_090340	13	277	ribosome biogenesis protein bms1, putative
TGME49_113830	12	245	ribosome biogenesis protein BMS1, putative
TGGT1_121710	1	1	conserved hypothetical protein
TGME49_119580	1	1	hypothetical protein
TGME49_068750	3	25	peptidyl-prolyl cis-trans isomerase E, putative
TGGT1_106930	3	24	dc50, putative
TGME49_037010	22	576	hypothetical protein, conserved
TGGT1_068900	15	459	conserved hypothetical protein
TGGT1_082300	1	2	conserved hypothetical protein
TGME49_026920	1	2	hypothetical protein
TGGT1_081190	4	36	conserved hypothetical protein
TGME49_026050	4	36	hypothetical protein
TGGT1_112420	2	8	molybdenum cofactor synthesis protein 2 large subunit, putative
TGGT1_065400	1	2	conserved hypothetical protein
TGME49_021540	1	2	hypothetical protein
TGGT1_124320	72	17050	conserved hypothetical protein
TGME49_014090	69	15371	hypothetical protein
TGGT1_114550	7	65	conserved hypothetical protein
TGME49_032970	7	65	hypothetical protein
TGGT1_073520	35	5907	conserved hypothetical protein
TGME49_093170	35	5907	hypothetical protein
TGME49_088350	23	443	nuclear NF-kB activating protein, putative
TGGT1_035050	7	139	conserved hypothetical protein
TGME49_054450	1	1	cell division protein pelota, putative
TGME49_114080	10	122	hypothetical protein
TGGT1_090720	8	174	conserved hypothetical protein
TGME49_031350	24	565	glucosamine--fructose-6-phosphate aminotransferase (isomerizing), putative
TGGT1_116350	7	78	glucosamine--fructose-6-phosphate aminotransferase
TGGT1_109490	5	27	conserved hypothetical protein
TGME49_070730	1	5	hypothetical protein
TGME49_013090	13	172	sec1 family domain-containing protein
TGGT1_039640	9	132	hypothetical protein
TGME49_086120	53	3963	prolyl endopeptidase, putative
TGGT1_037530	30	2773	prolyl endopeptidase, putative

TGME49_014150	3	14	mitochondrial import inner membrane translocase subunit, putative
TGGT1_124365	1	8	mitocondrial import inner membrane translocase subunit, putative
TGME49_115940	20	423	rhoptry protein, putative
TGGT1_093630	10	260	conserved hypothetical protein
TGGT1_087430	10	256	zinc finger (CCCH type) protein
TGME49_111100	6	120	zinc finger (CCCH type) protein, putative
TGME49_049760	1	4	vitamin K epoxide reductase complex subunit 1, putative
TGGT1_116730	7	54	conserved hypothetical protein
TGME49_030990	3	17	hypothetical protein
TGGT1_104800	12	421	trypsin, putative
TGME49_077850	11	360	trypsin, putative
TGGT1_124810	1	7	conserved hypothetical protein
TGME49_014400	1	7	hypothetical protein
TGME49_050770	69	17979	eukaryotic translation initiation factor 4A
TGGT1_102220	32	9036	conserved hypothetical protein
TGME49_119910	10	192	hypothetical protein
TGGT1_121480	8	189	conserved hypothetical protein
TGME49_020430	6	34	hypothetical protein
TGGT1_015140	5	51	conserved hypothetical protein
TGME49_078510	4	25	protein phosphatase 2C, putative
TGGT1_104050	4	24	protein phosphatase 2C, putative
TGME49_030930	16	1205	hypothetical protein
TGGT1_116890	16	1176	conserved hypothetical protein
TGME49_111710	16	107	hypothetical protein
TGGT1_039890	3	33	conserved hypothetical protein
TGME49_012870	2	20	hypothetical protein
TGME49_046340	11	181	DnaJ domain-containing protein
TGGT1_026690	10	232	DnaJ domain-containing protein, putative
TGGT1_036550	5	74	conserved hypothetical protein
TGME49_084620	5	74	hypothetical protein
TGGT1_023940	2	2	ubiquitin specific proteinase, putative
TGME49_048710	2	2	hypothetical protein
TGGT1_090210	7	105	conserved hypothetical protein
TGME49_113690	6	84	hypothetical protein, conserved
TGGT1_125240	1	1	conserved hypothetical protein
TGME49_014810	1	1	hypothetical protein
TGME49_104960	13	199	hypothetical protein
TGGT1_042530	9	191	protein phosphatase 2C, putative
TGGT1_042180	2	4	mRNA capping enzyme, putative
TGME49_105320	2	4	mRNA capping enzyme, putative
TGME49_043710	53	6648	TCP-1/cpn60 family chaperonin, putative
TGGT1_046170	36	4762	chaperonin containing t-complex protein 1, beta subunit, tcpb, putative
TGME49_031920	22	417	sepiapterin reductase, putative

TGGT1_115840	11	318	sepiapterin reductase, putative
TGME49_080440	2	21	ubiquitin-conjugating enzyme E2, putative
TGGT1_072870	2	12	ubiquitin-conjugating enzyme E2, putative
TGME49_086780	1	16	amiloride-sensitive amine oxidase, copper-containing, putative
TGME49_119300	23	322	glycosylphosphatidylinositol anchor attachment 1 protein, putative
TGGT1_121900	1	6	hypothetical protein
TGME49_111920	22	496	hypothetical protein
TGGT1_088170	19	749	conserved hypothetical protein
TGGT1_043510	2	4	conserved hypothetical protein
TGME49_063390	2	4	hypothetical protein
TGME49_053370	50	2609	hypothetical protein
TGGT1_002450	9	222	conserved hypothetical protein
TGGT1_087040	2	10	conserved hypothetical protein
TGME49_110730	2	10	non-muscle myosin heavy chain, putative
TGME49_109200	21	533	zinc finger (CCCH type) protein, putative
TGGT1_056350	3	86	conserved hypothetical protein
TGME49_043250	65	9249	myosin, putative
TGGT1_046630	24	745	myosin, putative
TGGT1_106440	4	16	RNA recognition motif-containing protein, putative
TGME49_068380	4	16	RNA recognition motif-containing protein
TGME49_109750	60	3831	succinyl-CoA ligase, putative
TGGT1_085710	18	1238	succinyl-CoA ligase, putative
TGME49_072350	14	341	poly(A)+ RNA export protein, putative
TGGT1_111540	13	292	poly(A)+ RNA export protein, putative
TGME49_109730	59	5937	thioredoxin reductase, putative
TGGT1_085690	23	2931	thioredoxin reductase, putative
TGME49_111900	5	10	hypothetical protein
TGGT1_088150	3	18	conserved hypothetical protein
TGME49_052630	9	175	hypothetical protein
TGGT1_001820	3	39	conserved hypothetical protein
TGME49_015400	7	40	RNA binding protein, putative
TGGT1_126250	6	48	RNA binding protein, putative
TGGT1_106240	13	512	hypothetical protein
TGME49_068160	13	464	hypothetical protein
TGGT1_110120	3	15	conserved hypothetical protein
TGME49_071250	3	15	KH domain-containing protein
TGME49_099250	8	141	hypothetical protein
TGGT1_006480	5	65	conserved hypothetical protein
TGGT1_047370	2	4	conserved hypothetical protein
TGME49_042450	2	4	hypothetical protein
TGME49_057470	31	1771	myosin IJ heavy chain, putative
TGGT1_013230	26	2109	myosin head motor domain-containing protein, putative
TGME49_094740	3	7	armadillo/beta-catenin-like repeat-containing protein

TGGT1_075060	1	3	conserved hypothetical protein
TGGT1_054400	14	908	zinc finger (CCCH type) protein
TGME49_095620	13	795	zinc finger (CCCH type) protein, putative
TGME49_118490	5	25	hypothetical protein
TGGT1_122630	4	24	conserved hypothetical protein
TGGT1_059830	4	16	conserved hypothetical protein
TGME49_101400	4	16	hypothetical protein
TGGT1_103160	1	2	conserved hypothetical protein
TGME49_051790	1	2	hypothetical protein
TGME49_118620	14	206	tRNA-splicing endonuclease positive effector protein, putative
TGGT1_122490	11	193	tRNA-splicing endonuclease positive effector, putative
TGGT1_001300	1	1	conserved hypothetical protein
TGME49_052390	1	1	hypothetical protein
TGME49_056770	28	1461	ATP-dependent helicase, putaive
TGGT1_013990	13	751	ATP-dependent helicase, putaive, putative
TGME49_101440	39	2723	CAM kinase, CDPK family TgCDPK1
TGGT1_059880	36	1896	calmodulin-domain protein kinase, putative
TGME49_087490	1	1	hypothetical protein, conserved
TGME49_109250	53	10021	AT hook motif-containing protein
TGGT1_056400	51	10579	AT hook motif-containing protein, putative
TGME49_032060	12	83	peptidase family T4, putative
TGGT1_115700	7	55	conserved hypothetical protein
TGME49_113490	8	44	hypothetical protein
TGGT1_089900	1	1	hypothetical protein
TGGT1_016630	3	24	splicing factor yt521-B, putative
TGME49_004070	2	6	YT521-B-like family domain-containing protein
TGGT1_029850	1	2	SF-assemblin, putative
TGME49_018880	1	2	SF-assemblin, putative
TGME49_005240	10	108	cleft lip and palate associated transmembrane protein 1, putative
TGGT1_063770	4	46	cleft lip and palate associated transmembrane protein, putative
TGGT1_019210	1	1	conserved hypothetical protein
TGME49_002010	1	1	hypothetical protein
TGME49_026000	16	738	ATP synthase, putative
TGGT1_081140	6	114	ATP synthase, putative
TGME49_065410	10	119	pre-mRNA splicing factor, putative
TGGT1_057690	9	113	pre-mRNA splicing factor, putative
TGGT1_066640	4	123	conserved hypothetical protein
TGME49_023100	3	99	hypothetical protein
TGME49_025880	11	230	UDP-glucose 4-epimerase, putative
TGGT1_081040	10	226	UDP-glucose 4-epimerase, putative
TGME49_030650	5	8	hypothetical protein
TGGT1_117200	4	7	conserved hypothetical protein
TGME49_048740	49	4105	hypothetical protein

TGGT1_100270	30	3343	conserved hypothetical protein
TGME49_109960	7	101	hypothetical protein
TGGT1_085900	6	72	conserved hypothetical protein
TGME49_048670	34	8119	H+-translocating inorganic pyrophosphatase TVP, putative
TGGT1_024080	5	42	H+-translocating inorganic pyrophosphatase TVP
TGGT1_069180	12	471	conserved hypothetical protein
TGME49_036660	11	422	hypothetical protein
TGGT1_121130	13	436	conserved hypothetical protein
TGME49_120160	13	436	hypothetical protein
TGME49_014940	40	6788	MIC2-associated protein M2AP
TGGT1_125360	8	770	MIC2-associated protein, putative
TGGT1_009360	4	13	conserved hypothetical protein
TGME49_060810	4	13	hypothetical protein, conserved
TGGT1_106250	6	164	conserved hypothetical protein
TGME49_068180	5	95	hypothetical protein
TGGT1_049850	3	13	C2 domain-containing protein, putative
TGME49_040910	3	13	C2 domain-containing protein
TGME49_099060	14	163	sodium/hydrogen exchanger, putative
TGGT1_006290	8	95	sodium/hydrogen exchanger, putative
TGME49_010820	12	366	hypothetical protein
TGGT1_124050	7	151	conserved hypothetical protein
TGME49_017890	42	3898	peroxiredoxin
TGGT1_028120	10	792	peroxiredoxin, putative
TGGT1_023000	71	7899	thioredoxin, putative
TGME49_009950	52	12656	thioredoxin, putative
TGME49_097490	3	9	hypothetical protein
TGME49_053080	10	78	replication protein A2, putative
TGGT1_002050	7	69	conserved hypothetical protein
TGGT1_111790	3	15	myo-inositol transprter, putative
TGME49_072500	2	7	sugar transporter, putative
TGGT1_019710	2	3	conserved hypothetical protein
TGME49_001720	2	3	hypothetical protein
TGGT1_011790	2	5	conserved hypothetical protein
TGME49_058420	1	1	hypothetical protein
TGGT1_107180	10	208	conserved hypothetical protein
TGME49_068990	8	160	hypothetical protein
TGME49_012930	4	33	NifU-like domain-containing protein
TGGT1_039820	4	32	conserved hypothetical protein
TGME49_008820	2	6	1-deoxy-D-xylulose 5-phosphate synthase, putative
TGGT1_021850	1	2	hypothetical protein
TGME49_069840	19	836	26S proteasome non-ATPase subunit, putative
TGGT1_108280	12	724	26S proteasome non-ATPase subunit, putative
TGME49_053030	14	205	alpha-glucosidase II, putative
TGGT1_002010	10	192	alpha-glucosidase II, putative
TGME49_069190	35	3128	glyceraldehyde-3-phosphate dehydrogenase

TGGT1_107390	25	1930	glyceraldehyde-3-phosphate dehydrogenase, putative
TGME49_071270	14	437	hypothetical protein
TGGT1_110140	7	36	conserved hypothetical protein
TGGT1_112760	1	2	conserved hypothetical protein
TGME49_073680	1	2	hypothetical protein
TGGT1_094170	1	8	conserved hypothetical protein
TGME49_116360	1	8	hypothetical protein
TGGT1_107740	1	6	conserved hypothetical protein
TGME49_069420	1	6	hypothetical protein, conserved
TGGT1_101790	3	15	conserved hypothetical protein
TGME49_050090	2	8	hypothetical protein
TGME49_012770	20	502	hypothetical protein
TGGT1_040010	19	547	conserved hypothetical protein
TGME49_015430	30	4029	hypothetical protein
TGGT1_126280	14	843	conserved hypothetical protein
TGME49_065450	35	1739	hexokinase
TGGT1_057650	25	1370	hexokinase, putative
TGGT1_090200	6	280	conserved hypothetical protein
TGME49_113680	2	70	hypothetical protein
TGGT1_074840	7	113	conserved hypothetical protein
TGME49_094440	7	113	hypothetical protein
TGGT1_040920	11	122	conserved hypothetical protein
TGME49_106430	9	92	hypothetical protein
TGME49_061080	29	1129	hypothetical protein, conserved
TGGT1_009190	15	436	conserved hypothetical protein
TGME49_033000	25	851	supt5h protein, putative
TGGT1_114520	22	806	supt5h protein, putative
TGME49_012250	29	997	XPG N-terminal domain containing protein
TGGT1_072000	7	186	conserved hypothetical protein
TGME49_041000	4	198	hypothetical protein
TGGT1_049670	3	170	conserved hypothetical protein
TGGT1_046300	9	301	hypothetical protein
TGME49_043590	9	209	endonuclease/exonuclease/phosphatase domain-containing protein
TGGT1_040710	3	5	tRNA (adenine-N(1)-)-methyltransferase catalytic subunit, putative
TGME49_106630	3	5	tRNA (adenine-N(1)-)-methyltransferase catalytic subunit, putative
TGGT1_070400	13	508	pantothenate kinase, putative
TGME49_035480	11	382	hypothetical protein, conserved
TGME49_057050	1	1	3-methyl-2-oxobutanoate hydroxymethyltransferase, putative
TGGT1_032500	32	2212	DNA mismatch repair protein, putative
TGME49_090640	29	1747	DNA mismatch repair protein, putative
TGGT1_081710	2	4	zinc finger (CCCH type) protein
TGME49_026560	2	4	zinc finger (CCCH type) protein, putative

TGME49_098610	110	33927	grb10 interacting GYF protein, putative
TGGT1_099410	91	33215	grb10 interacting GYF protein, putative
TGME49_109990	1	8	hypothetical protein, conserved
TGGT1_060650	9	314	deoxycytidylate deaminase, putative
TGME49_000430	9	314	deoxycytidylate deaminase, putative
TGME49_081610	5	62	hypothetical protein
TGGT1_005740	4	61	conserved hypothetical protein
TGME49_031430	3	24	oligosaccharyl transferase STT3, putative
TGGT1_114910	4	29	conserved hypothetical protein
TGME49_032610	4	29	hypothetical protein
TGGT1_068320	6	53	leucine carboxyl methyltransferase, putative
TGME49_037570	4	35	leucine carboxyl methyltransferase, putative
TGME49_048430	1	1	hypothetical protein
TGGT1_090780	1	8	general transcription factor IIIC, putative
TGME49_114230	1	3	general transcription factor IIIC, putative
TGME49_033470	3	6	hypothetical protein
TGGT1_114000	1	2	conserved hypothetical protein
TGME49_043930	50	4881	hypothetical protein
TGGT1_046030	4	238	hypothetical protein
TGGT1_004270	9	521	P35 surface protein, putative
TGME49_054720	8	5984	dense granule protein GRA8
TGGT1_071720	2	6	conserved hypothetical protein
TGME49_011850	2	6	hypothetical protein
TGME49_022670	3	7	hypothetical protein
TGGT1_066280	2	6	conserved hypothetical protein
TGGT1_117280	7	239	conserved hypothetical protein
TGME49_030570	6	180	hypothetical protein
TGME49_066440	7	31	hypothetical protein
TGGT1_076430	3	15	conserved hypothetical protein
TGME49_009210	19	596	KH domain-containing protein
TGGT1_022250	18	772	KH domain-containing protein, putative
TGME49_097720	45	4484	trehalose-6-phosphate synthase domain-containing protein
TGGT1_096460	35	5192	trehalose-6-phosphate synthase domain-containing protein, putative
TGGT1_029090	6	51	zinc finger (CCCH type) protein
TGME49_018360	6	51	zinc finger (CCCH type) protein, putative
TGGT1_125500	10	190	conserved hypothetical protein
TGME49_014980	10	175	hypothetical protein
TGGT1_110050	3	15	NBP2B protein, putative
TGME49_071180	3	15	hypothetical protein
TGME49_054820	2	6	hypothetical protein
TGME49_044850	1	1	GMC oxidoreductase, putative
TGME49_039280	2	2	hypothetical protein
TGME49_063730	37	4182	glycerol-3-phosphate dehydrogenase, putative
TGGT1_043990	12	318	glycerol-3-phosphate dehydrogenase, putative

TGGT1_101130	22	987	TPR domain-containing protein, putative
TGME49_049480	21	775	TPR domain-containing protein
TGGT1_051430	7	170	protein kinase, putative
TGME49_039440	6	152	protein kinase (incomplete catalytic triad)
TGGT1_081670	1	2	conserved hypothetical protein
TGME49_055980	16	157	hypothetical protein
TGGT1_014250	4	38	conserved hypothetical protein
TGGT1_010240	25	3674	conserved hypothetical protein
TGME49_059880	18	2392	hypothetical protein
TGGT1_100660	4	26	conserved hypothetical protein
TGME49_048990	2	8	hypothetical protein
TGME49_032130	30	1361	hypothetical protein, conserved
TGGT1_115520	19	1036	hypothetical protein
TGME49_023920	202	279223	hypothetical protein
TGGT1_078300	63	7709	conserved hypothetical protein
TGGT1_078620	16	582	conserved hypothetical protein
TGME49_024160	12	377	hypothetical protein
TGGT1_028150	8	242	DHHC domain-containing protein, putative
TGME49_017870	6	150	zinc finger DHHC domain-containing protein
TGGT1_081230	3	10	poly(A) polymerase, putative
TGME49_026080	3	10	poly(A) polymerase, putative
TGGT1_113920	3	8	conserved hypothetical protein
TGME49_033530	2	3	hypothetical protein
TGME49_120640	10	94	cyclophilin, putative
TGGT1_120650	7	85	cyclophilin, putative
TGME49_002650	1	1	hypothetical protein
TGGT1_008580	2	8	conserved hypothetical protein
TGME49_061650	2	8	hypothetical protein
TGGT1_058130	16	283	conserved hypothetical protein
TGME49_065090	7	82	hypothetical protein
TGME49_050820	63	21391	hypothetical protein
TGGT1_102370	50	23565	conserved hypothetical protein
TGME49_051840	20	1931	hypothetical protein
TGGT1_103210	19	1922	conserved hypothetical protein
TGGT1_105420	2	34	conserved hypothetical protein
TGME49_088340	7	62	UBX domain-containing protein
TGGT1_035060	5	52	UBX domain-containing protein, putative
TGME49_004380	13	398	hypothetical protein
TGGT1_016330	12	481	conserved hypothetical protein
TGGT1_003410	2	6	conserved hypothetical protein
TGME49_054130	2	6	hypothetical protein
TGGT1_083250	15	372	DNA-binding protein HU, putative
TGME49_027970	6	180	DNA-binding protein HU, putative
TGGT1_006360	2	4	conserved hypothetical protein
TGME49_099140	1	1	hypothetical protein

TGME49_005580	24	1135	hypothetical protein
TGGT1_063340	12	631	conserved hypothetical protein
TGME49_093690	31	2694	profilin family protein
TGGT1_074080	11	918	conserved hypothetical protein
TGME49_033220	27	2347	hypothetical protein
TGGT1_114270	25	3167	conserved hypothetical protein
TGME49_063190	17	98	adenylosuccinate lyase, putative
TGGT1_043290	7	83	adenylosuccinate lyase, putative
TGGT1_120840	8	167	conserved hypothetical protein
TGME49_120550	8	167	hypothetical protein
TGGT1_087900	11	258	leucine rich repeat protein, putative
TGME49_111650	8	168	leucine rich repeat protein, putative
TGME49_016720	5	52	hypothetical protein
TGGT1_098420	5	39	conserved hypothetical protein
TGME49_012840	1	1	hypothetical protein
TGME49_009190	21	904	ABC transporter, putative
TGGT1_022230	10	118	ABC transporter, putative
TGME49_009150	46	11991	mitochondrial alternative NADH dehydrogenase 1
TGGT1_022190	18	1005	pyridine nucleotide-disulphide oxidoreductase, putative
TGME49_114500	14	794	subtilisin-like protease TgSUB2
TGGT1_091290	3	11	subtilisin, putative
TGME49_047670	12	78	ribulose-phosphate 3-epimerase
TGGT1_025140	4	26	ribulose-phosphate 3-epimerase, putative
TGGT1_028560	1	2	conserved hypothetical protein
TGME49_017520	1	2	hypothetical protein
TGGT1_057030	12	467	conserved hypothetical protein
TGME49_065850	9	317	hypothetical protein
TGME49_070620	3	8	ATP-dependent RNA helicase, putative
TGGT1_109180	1	4	ATP-dependent RNA helicase, putative
TGGT1_054260	2	7	calcium-dependent protein kinase, putative
TGME49_095760	2	7	CAM kinase, CDPK family
TGGT1_007200	13	655	conserved hypothetical protein
TGME49_062780	12	564	hypothetical protein
TGME49_024860	3	10	exosome complex exonuclease, putative
TGGT1_079770	3	9	exosome complex exonuclease, putative
TGGT1_083160	3	18	conserved hypothetical protein
TGME49_027860	3	18	hypothetical protein
TGGT1_102430	5	384	DHHC domain-containing protein, putative
TGME49_050870	4	170	zinc finger DHHC domain-containing protein
TGGT1_066560	5	46	coproporphyrinogen oxidase, putative
TGME49_023020	5	35	coproporphyrinogen oxidase, putative
TGME49_099040	2	4	phosphatidylinositol N-acetylglucosaminyltransferase subunit, putative
TGME49_011400	6	51	helicase, putative
TGGT1_123390	5	47	helicase, putative

TGGT1_063410	2	3	integral membrane protein, putative
TGME49_005490	2	3	integral membrane protein, putative
TGME49_071470	1	1	hypothetical protein
TGGT1_098640	5	53	ap endonuclease, putative
TGME49_016600	4	33	DNA-(apurinic or apyrimidinic site) lyase, putative
TGME49_056980	30	1604	glycerol-3-phosphate acyltransferase, putative
TGGT1_013770	18	1354	glycerol-3-phosphate acyltransferase, putative
TGME49_081940	3	9	hypothetical protein, conserved
TGGT1_005530	2	8	conserved hypothetical protein
TGGT1_101400	2	4	conserved hypothetical protein
TGME49_049730	2	4	hypothetical protein
TGGT1_043810	1	2	conserved hypothetical protein
TGME49_063590	1	2	hypothetical protein
TGGT1_003770	1	6	conserved hypothetical protein
TGME49_054340	1	6	hypothetical protein, conserved
TGME49_101210	6	18	NAD(P) transhydrogenase, alpha subunit, putative
TGGT1_059630	3	9	NAD(P) transhydrogenase, alpha subunit, putative
TGGT1_008340	4	33	conserved hypothetical protein
TGME49_061790	4	33	hypothetical protein
TGME49_120600	8	958	glycine-rich protein 2, putative
TGGT1_120690	8	602	glycine-rich protein, putative
TGGT1_008570	4	52	conserved hypothetical protein
TGME49_061660	4	52	hypothetical protein
TGGT1_062050	2	6	conserved hypothetical protein
TGGT1_078670	16	561	pseudouridylate synthase, putative
TGME49_024200	15	316	hypothetical protein
TGME49_073560	76	6656	kinesin heavy chain, putative
TGGT1_112640	40	5216	kinesin motor domain-containing protein, putative
TGME49_110420	24	2741	hypothetical protein
TGGT1_086640	8	799	hypothetical protein
TGGT1_066250	4	365	hypothetical protein
TGME49_022640	2	175	leucine rich repeat containing protein
TGME49_109840	1	1	hypothetical protein
TGGT1_089070	2	4	conserved hypothetical protein
TGME49_112870	2	4	hypothetical protein
TGME49_112520	7	59	tRNA delta(2)-isopentenylpyrophosphate transferase, putative
TGGT1_088810	5	57	tRNA isopentenyl transferase, putative
TGGT1_105300	2	4	tRNA synthetase, putative
TGME49_077030	2	4	isoleucyl-tRNA synthetase, putative
TGGT1_083920	1	2	elongation factor Tu GTP-binding domain-containing protein
TGGT1_116550	13	484	conserved hypothetical protein
TGME49_031160	10	747	hypothetical protein
TGME49_045670	27	664	pyruvate dehydrogenase, putative

TGGT1_027400	12	347	pyruvate dehydrogenase, putative
TGME49_094270	6	73	arginine methyltransferase protein, putative
TGGT1_074560	4	69	protein arginine N-methyltransferase, putative
TGME49_072040	12	128	WD domain, G-beta repeat-containing protein
TGGT1_111160	10	208	WD domain, G-beta repeat-containing protein, putative
TGGT1_080730	1	1	conserved hypothetical protein
TGME49_025510	1	1	hypothetical protein, conserved
TGGT1_038600	1	4	conserved hypothetical protein
TGME49_013880	1	4	hypothetical protein, conserved
TGME49_015460	14	843	40S ribosomal protein S24, putative
TGGT1_126400	5	105	40S ribosomal protein S24, putative
TGME49_051620	19	270	flap endonuclease-1, putative
TGGT1_103010	15	290	flap endonuclease-1, putative
TGGT1_046080	4	32	conserved hypothetical protein
TGME49_043780	3	19	hypothetical protein
TGME49_072400	22	664	casein kinase II beta chain, putative
TGGT1_111590	15	566	casein kinase II beta chain, putative
TGME49_084190	42	1913	pyruvate carboxylase, putative
TGGT1_036120	23	981	hypothetical protein
TGME49_080610	27	1312	signal recognition particle protein, putative
TGGT1_072700	12	203	signal recognition particle protein, putative
TGME49_110100	1	4	GPI mannosyltransferase 1, putative
TGME49_034360	32	1010	DNA topoisomerase I, putative
TGGT1_071540	20	937	DNA topoisomerase I, putative
TGME49_071750	2	4	calmodulin, putative
TGME49_106970	13	450	thymidylate kinase, putative
TGGT1_055440	11	411	thymidylate kinase, putative
TGME49_116720	8	45	C2 domain-containing protein
TGGT1_094560	1	1	C2 domain-containing protein, putative
TGME49_021380	17	571	hypothetical protein, conserved
TGGT1_065230	8	182	conserved hypothetical protein
TGME49_030440	7	43	zinc finger (CHY type) protein, putative
TGGT1_117520	5	39	zinc finger (CHY type) protein
TGGT1_071310	4	26	DNA polymerase eta, putative
TGME49_034580	4	26	hypothetical protein, conserved
TGME49_016060	25	733	hypothetical protein
TGGT1_099190	11	292	conserved hypothetical protein
TGGT1_086830	1	2	nucleoporin FG repeat-containing protein, putative
TGGT1_017690	25	2302	leucine rich repeat containing protein / ankyrin repeat containing protein, putative
TGME49_003180	15	1145	leucine rich repeat containing protein / ankyrin repeat containing protein
TGGT1_066230	1	1	conserved hypothetical protein
TGME49_022420	1	1	hypothetical protein
TGGT1_112940	10	249	conserved hypothetical protein

TGME49_073860	7	132	hypothetical protein
TGME49_058860	18	632	hypothetical protein
TGGT1_011310	2	8	hypothetical protein
TGME49_014110	8	56	yip1 domain-containing protein
TGGT1_124340	7	46	YIP1 domain-containing protein, putative
TGGT1_037910	2	6	conserved hypothetical protein
TGME49_086500	1	4	hypothetical protein
TGME49_052500	3	20	hypothetical protein
TGGT1_001390	3	19	conserved hypothetical protein
TGME49_033760	2	4	hypothetical protein, conserved
TGME49_036510	5	21	hypothetical protein
TGGT1_069340	3	16	GAF domain containing protein, putative
TGGT1_031500	6	348	calmodulin-domain protein kinase 2 / protein kinase, putative
TGME49_092060	5	347	CAM kinase, CDPK family
TGGT1_121680	10	300	protein kinase domain-containing protein, putative
TGME49_119610	10	300	protein kinase (incomplete catalytic triad)
TGME49_022950	17	442	14 kDa phosphohistidine phosphatase, putative
TGGT1_066490	9	262	14 kDa phosphohistidine phosphatase, putative
TGME49_078950	3	14	U6 snRNA associated Sm-like protein LSm4, putative
TGME49_069450	9	119	hypothetical protein
TGGT1_107780	5	23	conserved hypothetical protein
TGGT1_016870	4	45	helicase, putative
TGME49_003840	2	14	helicase, putative
TGME49_053820	30	2323	hypothetical protein, conserved
TGGT1_002980	21	1991	conserved hypothetical protein
TGME49_014210	40	4139	centromere/microtubule binding protein, putative
TGGT1_124520	29	3582	centromere/microtubule binding protein, putative
TGME49_021220	2	4	hypothetical protein
TGME49_033480	20	3817	SRS29C (= SRS2, P35)
TGGT1_113990	7	186	SRS29C (= SRS2, P35)
TGGT1_023320	3	10	zinc finger (C2H2 type) protein
TGME49_007080	3	7	histone acetyltransferase
TGME49_018820	15	520	hypothetical protein, conserved
TGGT1_029790	8	266	conserved hypothetical protein
TGGT1_050070	5	44	RNA recognition motif-containing protein, putative
TGME49_040710	4	31	RRM domain-containing protein
TGGT1_056380	10	998	GTPase activating protein for Arf domain-containing protein, putative
TGME49_109220	10	998	ARF1-directed GTPase-activating protein, putative
TGME49_031130	14	258	hypothetical protein
TGGT1_116580	7	52	conserved hypothetical protein
TGME49_115270	64	8907	hypothetical protein
TGGT1_092700	39	8184	conserved hypothetical protein

TGGT1_024240	4	86	conserved hypothetical protein
TGME49_048610	4	86	hypothetical protein, conserved
TGME49_046140	6	72	hypothetical protein
TGGT1_026910	4	68	conserved hypothetical protein
TGME49_062640	17	552	hypothetical protein
TGGT1_007450	8	114	CG8, putative
TGME49_067310	4	7	myosin heavy chain, putative
TGGT1_077960	1	6	conserved hypothetical protein
TGME49_013430	1	1	hypothetical protein
TGGT1_077740	1	1	thrombospondin type 1 domain-containing protein, putative
TGGT1_037020	11	240	conserved hypothetical protein
TGME49_085520	6	124	hypothetical protein
TGGT1_007360	32	1442	conserved hypothetical protein
TGME49_062730	24	1022	Rhoptry kinase family protein ROP16
TGGT1_027030	5	30	conserved hypothetical protein
TGME49_046030	4	20	hypothetical protein, conserved
TGGT1_112690	3	3	conserved hypothetical protein
TGME49_073610	2	2	hypothetical protein
TGME49_091910	1	1	hypothetical protein
TGGT1_041880	4	32	conserved hypothetical protein
TGME49_105600	4	32	hypothetical protein
TGGT1_106190	11	84	conserved hypothetical protein
TGME49_068000	9	60	hypothetical protein
TGME49_028300	6	26	hypothetical protein
TGGT1_083870	4	20	conserved hypothetical protein
TGME49_111870	29	941	WD repeat protein, putative
TGGT1_088120	15	508	WD-repeat protein, putative
TGME49_030430	8	21	hypothetical protein
TGGT1_117540	2	5	vAMP-7, putative
TGME49_035190	1	1	hypothetical protein
TGGT1_041500	2	6	conserved hypothetical protein
TGME49_105970	2	6	hypothetical protein
TGME49_085980	60	6322	phosphoglucomutase/parafusin related protein 1, putative
TGGT1_037370	28	3227	phosphoglucomutase, putative
TGGT1_088180	11	287	GRAM domain-containing protein, putative
TGME49_111930	11	209	GRAM domain-containing protein
TGME49_045510	15	677	P-type ATPase, putative
TGGT1_027550	10	331	phospholipid-translocating P-type ATPase family protein
TGME49_052380	9	78	hypothetical protein
TGGT1_001290	8	93	conserved hypothetical protein
TGME49_068870	30	1445	TPR domain-containing protein
TGGT1_107060	27	1524	hypothetical protein
TGGT1_064050	2	8	conserved hypothetical protein
TGME49_005000	1	3	phosphoglycerate mutase domain-containing protein

TGGT1_028500	1	4	DNA-directed RNA polymerase II, 8.2 kDa polypeptide, putative
TGME49_017560	1	4	DNA-directed RNA polymerase II, 8.2 kDa polypeptide, putative
TGME49_031750	4	7	hypothetical protein
TGGT1_116020	1	1	conserved hypothetical protein
TGGT1_002000	6	30	conserved hypothetical protein
TGME49_053020	4	22	hypothetical protein
TGME49_104660	13	81	glutaredoxin domain containing protein
TGGT1_006710	5	38	prostaglandin-E synthase, putative
TGGT1_029400	6	40	RNA-binding protein nob1, putative
TGME49_018570	6	40	hypothetical protein
TGME49_049350	12	321	hypothetical protein, conserved
TGGT1_101020	10	317	conserved hypothetical protein
TGGT1_010500	3	9	conserved hypothetical protein
TGME49_059520	3	9	hypothetical protein
TGGT1_053200	16	871	DnaJ domain-containing protein, putative
TGME49_010430	14	799	DnaJ domain-containing protein
TGGT1_121320	2	11	CDK-activating kinase assembly factor, putative
TGME49_120070	2	11	CDK-activating kinase assembly factor, putative
TGME49_007930	4	30	phosphatidylethanolamine-binding protein, putative
TGGT1_020920	3	21	phosphatidylethanolamine-binding protein, putative
TGGT1_056140	1	2	hypothetical protein
TGME49_109100	1	2	hypothetical protein
TGME49_018300	28	3946	zinc finger (CCCH type) protein, putative
TGGT1_029030	27	4476	zinc finger (CCCH type) protein
TGME49_095790	25	714	hypothetical protein
TGGT1_054230	15	634	conserved hypothetical protein
TGME49_016610	3	5	hypothetical protein
TGME49_048870	2	9	hypothetical protein
TGME49_098030	2	8	hypothetical protein, conserved
TGME49_024470	1	10	hypothetical protein
TGGT1_126240	2	14	conserved hypothetical protein
TGME49_015390	2	14	Tim10/DDP family zinc finger containing protein
TGGT1_088710	14	560	conserved hypothetical protein
TGME49_112420	12	367	hypothetical protein
TGGT1_007150	2	3	conserved hypothetical protein
TGME49_062830	2	3	hypothetical protein
TGME49_046980	13	78	hypothetical protein
TGGT1_026030	4	19	glycosyl transferase, putative
TGME49_003390	17	448	CRAL/TRIO domain-containing protein
TGGT1_017440	9	207	CRAL/TRIO domain-containing protein, putative
TGME49_015940	9	32	acetyl-CoA transporter, putative
TGGT1_099320	1	2	acetyl-CoA transporter, putative
TGME49_099210	49	4049	CTP synthase, putative

TGGT1_006440	26	2464	CTP synthase, putative
TGGT1_031820	12	384	conserved hypothetical protein
TGME49_091620	10	280	hypothetical protein
TGME49_116400	86	204430	tubulin alpha chain
TGGT1_094210	26	8464	tubulin alpha chain, putative
TGGT1_090860	5	25	conserved hypothetical protein
TGME49_114280	4	16	hypothetical protein
TGME49_083730	7	21	endomembrane domain-containing protein
TGGT1_035770	4	10	endomembrane domain-containing protein, putative
TGME49_024460	27	603	aminopeptidase N, putative
TGGT1_078970	9	156	aminopeptidase N, putative
TGGT1_090650	1	6	selR domain-containing protein, putative
TGME49_114000	1	6	selR domain-containing protein
TGME49_093360	23	7046	hypothetical protein
TGGT1_073720	17	5634	conserved hypothetical protein
TGME49_068370	129	11301	hypothetical protein
TGGT1_106430	87	11157	hypothetical protein
TGGT1_005850	1	1	ribonuclease HI large subunit, putative
TGME49_081510	1	1	ribonuclease HI large subunit, putative
TGME49_090680	10	91	hypothetical protein
TGGT1_032350	6	34	conserved hypothetical protein
TGME49_093000	54	9214	3', 5'-cyclic nucleotide phosphodiesterase domain-containing protein
TGGT1_073240	51	9161	3', 5'-cyclic nucleotide phosphodiesterase, putative
TGME49_085760	11	138	peptidyl-prolyl cis-trans isomerase H, putative
TGGT1_037160	4	42	peptidyl-prolyl cis-trans isomerase H, putative
TGME49_110000	28	2427	lung seven transmembrane receptor domain-containing protein
TGGT1_086040	8	162	lung seven transmembrane receptor domain-containing protein, putative
TGME49_077570	11	344	GPI transamidase subunit PIG-U, putative
TGGT1_104990	6	64	GPI transamidase subunit PIG-U, putative
TGGT1_080700	3	61	calcium-dependent protein kinase, putative
TGME49_025490	3	61	CAM kinase, CDPK family TgPK6
TGGT1_068310	1	1	conserved hypothetical protein
TGME49_037590	1	1	hypothetical protein
TGGT1_068730	1	1	conserved hypothetical protein
TGME49_037190	1	1	hypothetical protein
TGME49_077820	4	18	hypothetical protein
TGGT1_104830	2	20	conserved hypothetical protein
TGGT1_007140	2	5	conserved hypothetical protein
TGME49_062850	2	5	hypothetical protein
TGGT1_079520	1	6	conserved hypothetical protein
TGME49_031070	49	3053	protein kinase
TGGT1_116650	26	1919	protein kinase, putative

TGGT1_071130	3	26	conserved hypothetical protein
TGME49_034990	3	26	hypothetical protein
TGME49_015680	4	118	SRS47B (= SAG2E)
TGGT1_126600	4	17	SRS47B (= SAG2E)
TGGT1_126530	1	2	conserved hypothetical protein
TGME49_114790	6	79	small nuclear ribonucleoprotein E/G, putative
TGGT1_092090	4	62	small nuclear ribonucleoprotein E/G, putative
TGME49_016020	41	367	hypothetical protein
TGGT1_099230	6	35	conserved hypothetical protein
TGME49_056990	64	4638	glycyl-tRNA synthetase, putative
TGGT1_013750	36	3298	glycyl-tRNA synthetase, putative
TGME49_071810	29	1405	inhibitor-1 of protein phosphatase type 2A
TGGT1_110610	14	918	inhibitor-1 of protein phosphatase type 2A, putative
TGGT1_032040	1	1	TATA-box binding protein, putative
TGME49_091080	1	1	TATA-box binding protein, putative
TGME49_105460	13	168	methionine aminopeptidase, type II, putative
TGGT1_042120	11	144	methionine aminopeptidase, type II, putative
TGME49_036540	26	1204	RRM domain-containing protein
TGGT1_069300	17	731	RNA recognition motif-containing protein, putative
TGGT1_071370	15	475	U2 snRNP auxiliary factor or splicing facotr, putative
TGME49_034520	15	475	U2 snRNP auxiliary factor or splicing facotr, putative
TGME49_018370	1	2	hypothetical protein
TGGT1_041870	6	91	conserved hypothetical protein
TGME49_105610	6	91	hypothetical protein
TGME49_048880	19	1795	Ras family domain-containing protein
TGGT1_100550	10	357	rab9 and, putative
TGGT1_064640	14	571	conserved hypothetical protein
TGME49_121600	14	387	hypothetical protein
TGGT1_110270	2	2	conserved hypothetical protein
TGME49_071400	2	2	tRNA intron endonuclease, catalytic C-terminal domain subfamily, putative
TGME49_106680	1	1	hypothetical protein
TGGT1_091240	1	2	conserved hypothetical protein
TGME49_114450	1	2	hypothetical protein
TGME49_024910	6	94	CutA1 divalent ion tolerance domain-containing protein
TGGT1_079820	5	93	CutA1 divalent ion tolerance domain-containing protein, putative
TGME49_045680	22	1744	ribosomal protein L21, putative
TGGT1_027390	8	212	ribosomal protein L21, putative
TGGT1_025020	3	29	conserved hypothetical protein
TGME49_047790	2	15	hypothetical protein
TGME49_015280	32	3741	iron-sulfur subunit of succinate dehydrogenase, putative
TGGT1_126010	9	212	iron-sulfur subunit of succinate dehydrogenase, putative
TGGT1_063830	1	3	conserved hypothetical protein

TGGT1_062180	43	5895	conserved hypothetical protein
TGME49_006580	36	4422	hypothetical protein, conserved
TGME49_003240	45	4606	hypothetical protein
TGGT1_017630	22	940	conserved hypothetical protein
TGME49_029650	25	1764	machado-joseph disease protein, putative
TGGT1_118120	24	1748	machado-joseph disease protein, putative
TGME49_039500	13	991	proteasome subunit alpha type 4, subunit
TGGT1_051360	13	928	proteasome subunit alpha type 4, subunit, putative
TGME49_099080	19	806	hypothetical protein
TGGT1_006310	14	817	xenotropic and polytropic murine leukemia virus receptor ids-4, putative
TGME49_048250	9	69	pre-mRNA-splicing factor, putative
TGGT1_024690	6	60	conserved hypothetical protein
TGME49_076210	5	48	hypothetical protein
TGGT1_000330	4	28	conserved hypothetical protein
TGME49_016370	2	16	hypothetical protein
TGGT1_045830	4	58	protamine P1 protein, putative
TGME49_044120	3	41	protamine P1 protein, putative
TGME49_039310	15	198	ribose 5-phosphate isomerase, putative
TGGT1_051660	11	208	hypothetical protein
TGME49_047650	3	69	sybindin-like family domain-containing protein
TGGT1_025160	3	67	synbindin, putative
TGME49_069340	7	101	hypothetical protein
TGGT1_107640	4	16	conserved hypothetical protein
TGGT1_065810	3	7	ww domain U1 zinc finger protein domain-containing protein, putative
TGGT1_012050	2	4	conserved hypothetical protein
TGME49_058160	1	2	hypothetical protein
TGGT1_018670	17	905	conserved hypothetical protein
TGME49_002410	12	605	hypothetical protein, conserved
TGME49_094890	2	9	hypothetical protein
TGME49_010810	5	47	hypothetical protein
TGGT1_124060	4	31	conserved hypothetical protein
TGGT1_006750	9	63	conserved hypothetical protein
TGME49_104700	9	63	hypothetical protein
TGGT1_034170	19	435	conserved hypothetical protein
TGME49_089080	19	435	adaptin N terminal region domain-containing protein
TGME49_048500	13	217	hypothetical protein
TGGT1_024460	7	125	conserved hypothetical protein
TGME49_113400	11	228	DnaJ domain-containing protein
TGGT1_089800	9	218	DnaJ domain-containing protein, putative
TGME49_032050	59	7602	DnaJ domain-containing protein
TGGT1_115710	23	1424	DnaJ domain-containing protein, putative
TGGT1_088840	8	136	NBP2B protein, putative
TGME49_112550	7	104	hypothetical protein

TGGT1_111440	4	36	heat shock protein HSLV, putative
TGME49_072230	4	36	heat shock protein HSLV, putative
TGME49_039800	5	77	hypothetical protein
TGGT1_051030	3	19	conserved hypothetical protein
TGGT1_013840	2	16	conserved hypothetical protein
TGME49_056910	1	14	hypothetical protein
TGME49_114540	12	184	hypothetical protein
TGGT1_091330	11	175	conserved hypothetical protein
TGGT1_037050	27	1452	conserved hypothetical protein
TGME49_085650	22	949	hypothetical protein
TGME49_116140	13	711	hypothetical protein
TGGT1_093850	5	65	conserved hypothetical protein
TGGT1_094340	4	39	conserved hypothetical protein
TGME49_116530	4	39	myosin heavy chain, putative
TGGT1_032950	2	9	conserved hypothetical protein
TGME49_090190	2	9	hypothetical protein
TGME49_089350	5	20	white protein, putative
TGGT1_033870	1	9	ABC transporter, putative
TGGT1_106620	5	70	conserved hypothetical protein
TGME49_068560	3	32	RAD2 endonuclease, putative
TGGT1_093790	1	1	conserved hypothetical protein
TGME49_066760	35	1403	isocitrate dehydrogenase, putative
TGGT1_076210	20	882	isocitrate dehydrogenase, putative
TGGT1_062320	60	11872	conserved hypothetical protein
TGME49_006450	43	7345	hypothetical protein
TGME49_066810	8	36	hypothetical protein, conserved
TGGT1_076150	3	15	conserved hypothetical protein
TGME49_045490	35	4058	microneme protein 8
TGGT1_027570	21	1008	microneme protein, putative
TGGT1_028520	2	3	conserved hypothetical protein
TGME49_017550	2	3	hypothetical protein
TGME49_029630	43	1642	protein kinase (incomplete catalytic triad)
TGGT1_118140	33	2233	hypothetical protein
TGME49_121500	32	3788	CELF family protein, putative
TGGT1_064550	18	2221	conserved hypothetical protein
TGME49_024340	1	8	hypothetical protein
TGGT1_122340	3	27	hypothetical protein
TGME49_118770	3	27	Aurora kinase(incomplete catalytic triad)
TGME49_014590	7	39	microfibrillar-associated protein 1, putative
TGGT1_124990	6	38	microfibrillar-associated protein, putative
TGME49_061400	10	58	hypothetical protein
TGGT1_008840	4	20	conserved hypothetical protein
TGME49_073360	16	288	ABC transporter, putative
TGGT1_112430	2	4	ABC transporter, putative
TGGT1_022210	8	108	conserved hypothetical protein

TGME49_009170	7	88	hypothetical protein
TGME49_090890	20	651	carbonyl reductase, putative
TGGT1_032250	14	548	oxidoreductase, putative
TGGT1_086610	1	2	conserved hypothetical protein
TGME49_110390	1	2	hypothetical protein
TGGT1_021540	1	3	conserved hypothetical protein
TGME49_008510	1	3	hypothetical protein
TGME49_029920	24	6604	hypothetical protein, conserved
TGGT1_117970	10	354	conserved hypothetical protein
TGME49_026250	58	8426	ATP-dependent RNA helicase, putative
TGGT1_081400	40	6898	ATP-dependent RNA helicase, putative
TGGT1_071360	1	1	hypothetical protein
TGME49_034530	1	1	hypothetical protein, conserved
TGME49_005230	17	540	hypothetical protein
TGGT1_063780	5	39	conserved hypothetical protein
TGGT1_113240	3	28	conserved hypothetical protein
TGME49_074050	3	28	hypothetical protein
TGGT1_079110	4	20	farnesyl pyrophosphate synthetase, putative
TGME49_024490	4	20	farnesyl pyrophosphate synthetase, putative
TGGT1_016140	6	79	conserved hypothetical protein
TGME49_004540	6	79	hypothetical protein
TGME49_014580	2	4	protein antigen, putative
TGGT1_104020	5	40	conserved hypothetical protein
TGME49_078540	5	40	hypothetical protein
TGME49_015560	21	245	protein arginine N-methyltransferase, putative
TGGT1_126490	11	143	protein arginine N-methyltransferase, putative
TGME49_057340	8	95	Ras family domain-containing protein
TGGT1_013460	3	36	RAB-2,4,14, putative
TGME49_012210	22	1588	hypothetical protein
TGGT1_071950	8	719	conserved hypothetical protein
TGME49_013000	13	148	activator 1 36 kDa, putative
TGGT1_039740	9	114	activator 1 36 kDa, putative
TGGT1_117170	25	1541	conserved hypothetical protein
TGME49_030680	21	1161	hypothetical protein
TGME49_090580	6	100	ABC transporter, putative
TGME49_006670	56	5612	hypothetical protein
TGGT1_062060	50	5706	conserved hypothetical protein
TGME49_061190	30	575	kringle domain-containing protein
TGGT1_009180	13	216	kringle domain-containing protein, putative
TGME49_081580	25	951	protein phosphatase 2C, putative
TGGT1_005780	15	989	protein phosphatase 2C, putative
TGGT1_098890	1	2	conserved hypothetical protein
TGME49_016360	1	2	hypothetical protein
TGGT1_118190	60	9559	membrane-associated calcum-binding protein, putative
TGME49_029480	40	14823	membrane-associated calcum-binding protein, related

TGGT1_071260	13	480	conserved hypothetical protein
TGME49_034640	10	320	hypothetical protein
TGGT1_004530	13	327	conserved hypothetical protein
TGME49_054950	11	251	hypothetical protein
TGGT1_044010	8	98	conserved hypothetical protein
TGME49_063750	5	32	hypothetical protein
TGGT1_111040	12	297	kelch motif domain-containing protein, putative
TGME49_072020	8	186	kelch motif domain-containing protein
TGME49_013520	11	240	amidohydrolase domain-containing protein
TGGT1_039100	10	206	amidohydrolase, putative
TGGT1_039250	6	53	conserved hypothetical protein
TGME49_013400	3	23	hypothetical protein
TGME49_024000	10	183	hypothetical protein
TGGT1_078480	9	203	conserved hypothetical protein
TGME49_053520	4	12	hypothetical protein
TGGT1_002690	3	15	conserved hypothetical protein
TGME49_094930	12	301	leucine rich repeat protein, putative
TGGT1_075280	9	285	leucine rich repeat protein, putative
TGME49_002550	45	2443	NLI interacting factor-like phosphatase domain-containing protein
TGGT1_018530	41	2751	conserved hypothetical protein
TGGT1_107110	1	4	inositol hexaphosphate kinase, putative
TGME49_068920	1	4	inositol polyphosphate kinase domain containing protein
TGME49_033400	8	55	UPF0101 protein CGI-137, putative
TGGT1_114070	5	35	conserved hypothetical protein
TGME49_029900	4	19	hypothetical protein
TGGT1_117990	3	18	conserved hypothetical protein
TGGT1_031720	22	787	regulator of nonsense transcripts 1 protein, putative
TGME49_091820	22	641	regulator of nonsense transcripts UPF1, putative
TGGT1_076730	14	414	zinc finger (CCCH type) protein
TGME49_055310	10	297	zinc finger (CCCH type) protein, putative
TGME49_117720	42	3652	eukaryotic translation initiation factor 3 subunit 7, putative
TGGT1_105580	25	2270	eukaryotic translation initiation factor 3 subunit, putative
TGME49_108920	35	1340	U2 small nuclear ribonucleoprotein auxiliary factor U2AF
TGGT1_055960	21	1110	U2 small nuclear ribonucleoprotein auxiliary factor U2AF, putative
TGME49_038400	27	1253	endonuclease/exonuclease/phosphatase domain-containing protein
TGGT1_052550	21	1536	synaptojanin, putative
TGME49_086210	4	22	serine/threonine protein phosphatase, putative
TGGT1_037610	3	18	serine/threonine protein phosphatase, putative
TGGT1_095830	12	201	conserved hypothetical protein
TGME49_099780	12	191	hypothetical protein

TGME49_073460	48	6437	eukaryotic translation initiation factor 3 subunit 6 interacting protein, putative
TGGT1_112540	24	4603	eukaryotic translation initiation factor 3 subunit 6 interacting protein, putative
TGME49_013910	13	81	hypothetical protein, conserved
TGGT1_038570	4	19	conserved hypothetical protein
TGME49_093390	7	73	cytochrome c heme lyase, putative
TGGT1_073750	6	72	cytochrome C heme lyase, putative
TGME49_016380	9	74	ATPase 2, putative
TGGT1_098870	4	30	ATPase, putative
TGGT1_018250	1	2	ribonuclease, putative
TGME49_002720	1	2	ribonuclease II RNB family protein
TGGT1_075270	2	6	conserved hypothetical protein
TGME49_094920	2	6	general transcription factor IIH polypeptide 2, putative
TGME49_110070	20	1061	hypothetical protein
TGGT1_086210	16	884	conserved hypothetical protein
TGME49_003450	16	312	hypothetical protein, conserved
TGGT1_017380	9	198	conserved hypothetical protein
TGGT1_098180	2	7	pre-mRNA splicing factor RNA helicase, putative
TGME49_016950	2	7	pre-mRNA splicing factor RNA helicase, putative
TGME49_007170	12	453	hypothetical protein
TGGT1_023410	5	70	conserved hypothetical protein
TGME49_054960	1	25	hypothetical protein
TGME49_002020	5	18	hypothetical protein
TGME49_097750	60	3024	hypothetical protein
TGGT1_096440	33	1985	conserved hypothetical protein
TGME49_116470	2	4	hypothetical protein
TGGT1_094280	1	1	conserved hypothetical protein
TGGT1_010460	16	860	endonuclease/exonuclease/phosphatase domain-containing protein, putative
TGME49_059560	16	860	endonuclease/exonuclease/phosphatase domain-containing protein
TGME49_011280	5	25	hypothetical protein
TGGT1_123540	1	1	conserved hypothetical protein
TGGT1_043850	5	97	RNA-binding protein, putative
TGME49_063610	5	97	RNA-binding protein, putative
TGME49_086750	25	1098	MA3 domain protein
TGGT1_038150	19	1018	MA3 domain protein, putative
TGGT1_025620	10	514	conserved hypothetical protein
TGME49_047380	9	413	hypothetical protein
TGGT1_074090	4	28	conserved hypothetical protein
TGME49_093700	4	28	hypothetical protein
TGGT1_009220	12	212	mitotic control protein dis3, putative
TGME49_061050	11	155	mitotic control protein dis3, putative
TGME49_049180	86	9419	bifunctional dihydrofolate reductase / thymidylate synthase

TGGT1_100740	33	3986	bifunctional dihydrofolate reductase-thymidylate synthase, putative
TGGT1_098210	2	7	conserved hypothetical protein
TGME49_016920	2	7	hypothetical protein, conserved
TGME49_094290	2	6	der1-like family domain-containing protein, conserved
TGGT1_074590	1	1	conserved hypothetical protein
TGGT1_071160	5	95	conserved hypothetical protein
TGME49_034960	4	58	hypothetical protein
TGME49_032340	42	2350	protein phosphatase 2C, putative
TGGT1_115290	30	2107	protein phosphatase 2C, putative
TGME49_086090	12	114	translation initiation factor SUI1, putative
TGGT1_037500	5	83	translation initiation factor SUI1, putative
TGME49_051490	5	25	hypothetical protein, conserved
TGGT1_102890	3	21	conserved hypothetical protein
TGME49_070900	6	30	ATPase, AAA family domain-containing protein
TGGT1_109750	5	29	C-Myc binding protein, putative
TGGT1_082810	3	27	conserved hypothetical protein
TGME49_027320	1	7	hypothetical protein
TGGT1_018220	1	2	exosome complex exonuclease, putative
TGME49_002750	1	2	exosome complex exonuclease, putative
TGME49_058130	9	238	Ras family domain-containing protein
TGGT1_012080	3	30	small GTP-binding protein domain-containing protein
TGGT1_083040	1	2	conserved hypothetical protein
TGME49_027630	1	2	hypothetical protein
TGME49_109800	12	119	small nuclear ribonucleoprotein U1A, putative
TGGT1_085760	6	48	small nuclear ribonucleoprotein U1A, putative
TGME49_017420	4	58	hypothetical protein, conserved
TGGT1_115400	1	5	conserved hypothetical protein
TGME49_032240	1	5	hypothetical protein, conserved
TGGT1_043560	13	381	conserved hypothetical protein
TGME49_063440	10	233	hypothetical protein
TGME49_078720	22	730	hypothetical protein
TGGT1_103850	11	207	conserved hypothetical protein
TGGT1_079780	33	3520	conserved hypothetical protein
TGME49_024870	19	1515	hypothetical protein
TGME49_077240	95	49523	nucleoside-triphosphatase I
TGGT1_105180	71	11749	nucleoside-triphosphatase I, putative
TGGT1_104490	2	13	double-strand break repair protein, putative
TGME49_078060	1	5	double-strand break repair protein, putative
TGME49_040880	3	9	hypothetical protein
TGME49_049650	8	93	apolipoprotein A-I binding protein, putative
TGGT1_101310	4	29	apolipoprotein A-I binding protein, putative
TGME49_077090	2	6	mitochondrial carrier domain-containing protein
TGGT1_105240	1	2	mitochondrial carrier domain-containing protein

TGME49_086450	14	8852	dense granule protein 5 precursor
TGGT1_037870	8	877	dense granule protein 5 precursor, putative
TGME49_007840	20	1918	40S ribosomal protein S17, putative
TGGT1_020820	11	602	40S ribosomal protein S17, putative
TGME49_016560	6	30	hypothetical protein
TGGT1_098690	3	12	conserved hypothetical protein
TGME49_046600	23	914	ABC1 family beta-lactamase, putative
TGGT1_026510	21	970	conserved hypothetical protein
TGME49_060200	11	120	ATP-dependent RNA helicase, putative
TGGT1_010030	10	136	ATP-dependent RNA helicase, putative
TGGT1_105380	3	61	ribosome associated membrane domain-containing protein, putative
TGME49_076940	3	61	Ribosome associated membrane domain-containing protein
TGME49_023390	87	8048	sec63 domain-containing DEAD/DEAH box helicase, putative
TGGT1_077470	56	6336	sec63 domain-containing DEAD/DEAH box helicase, putative
TGME49_067130	2	18	SRS38A
TGME49_051630	35	2930	cation efflux family protein, putative
TGGT1_103020	20	1979	cation efflux protein/ zinc transporter, putative
TGME49_094340	6	79	hypothetical protein
TGGT1_074640	6	78	conserved hypothetical protein
TGGT1_083510	4	38	conserved hypothetical protein
TGME49_028140	4	38	hypothetical protein
TGGT1_115420	1	2	SWIB/MDM2 domain-containing protein, putative
TGGT1_010410	1	2	conserved hypothetical protein
TGME49_059610	1	2	hypothetical protein
TGGT1_069160	11	160	zinc finger (C3HC4 RING finger) protein
TGME49_036780	9	75	zinc finger (C3HC4 RING finger) protein, putative
TGME49_010780	25	238	ubiquitin carboxyl-terminal hydrolase, putative
TGGT1_124100	13	137	ubiquitin carboxyl-terminal hydrolase, putative
TGME49_048810	48	2973	hypothetical protein
TGGT1_100460	18	1808	conserved hypothetical protein
TGME49_053380	10	55	hypothetical protein
TGGT1_002460	8	53	conserved hypothetical protein
TGGT1_011750	9	106	conserved hypothetical protein
TGME49_058460	5	30	hypothetical protein
TGGT1_013560	9	208	conserved hypothetical protein
TGME49_057150	7	143	NOT2 / NOT3 / NOT5 family domain-containing protein
TGME49_113350	13	56	WD-40 repeat-containing protein
TGGT1_089750	3	24	WD-40 repeat-containing protein, putative
TGGT1_013290	1	1	conserved hypothetical protein
TGME49_046540	28	4377	cytochrome C1, putative
TGGT1_026580	8	130	cytochrome C1, putative
TGGT1_013700	2	24	conserved hypothetical protein
TGME49_057030	2	24	hypothetical protein

TGGT1_088930	9	161	conserved hypothetical protein
TGME49_112640	6	73	hypothetical protein
TGME49_052860	1	1	hypothetical protein, conserved
TGME49_033450	28	2862	SRS29A (= SRS1)
TGGT1_114020	13	540	SRS29A (= SRS1)
TGME49_052220	34	2026	Hsc70/Hsp90-organizing protein, putative
TGGT1_001140	23	1608	heat shock protein 70 (HSP70)-interacting protein, putative
TGME49_022970	9	58	inositol monophosphatase, putative
TGGT1_066510	5	22	inositol monophosphatase, putative
TGGT1_024660	31	3263	WD domain-containing protein, putative
TGME49_048290	29	2906	WD domain-containing protein
TGGT1_068065	12	1286	granule antigen protein GRA6, putative
TGME49_075440	6	948	granule antigen protein GRA6
TGME49_116310	22	1224	superoxide dismutase
TGGT1_094100	9	430	superoxide dismutase, putative
TGME49_063990	8	156	hypothetical protein
TGGT1_044230	2	23	conserved hypothetical protein
TGME49_097840	6	12	DNA primase, large subunit, putative
TGGT1_096260	1	1	DNA primase, large subunit, putative
TGGT1_032340	10	309	conserved hypothetical protein
TGME49_090700	8	362	hypothetical protein
TGME49_062070	10	229	ribosomal RNA assembly protein, putative
TGGT1_008040	10	220	ribosomal RNA assembly protein, putative
TGME49_019070	3	9	cAMP-dependent protein kinase regulatory subunit, putative
TGME49_015730	8	72	hypothetical protein
TGGT1_126640	7	71	conserved hypothetical protein
TGME49_023060	13	141	MORN repeat-containing protein
TGME49_025850	34	1532	hypothetical protein
TGGT1_080990	16	316	conserved hypothetical protein
TGGT1_044260	5	46	conserved hypothetical protein
TGME49_064020	5	46	hypothetical protein
TGGT1_034370	4	27	conserved hypothetical protein
TGME49_088910	3	27	hypothetical protein, conserved
TGME49_052880	75	3133	hypothetical protein
TGGT1_001950	9	135	conserved hypothetical protein
TGGT1_084250	12	321	hypothetical protein
TGME49_028670	12	321	mitochondrial carrier domain-containing protein
TGME49_106600	6	109	RNA recognition motif domain-containing protein
TGGT1_040740	6	105	ribonucleic acid binding protein S1, putative
TGME49_097090	3	3	hypothetical protein, conserved
TGME49_048950	4	12	mitochondrial carrier domain-containing protein
TGGT1_100620	2	10	mitochondrial carrier domain-containing protein
TGME49_106610	4	7	3' exoribonuclease family, domain 1 containing protein

TGGT1_040730	1	2	conserved hypothetical protein
TGME49_047490	3	9	hypothetical protein
TGGT1_088780	10	343	conserved hypothetical protein
TGME49_112490	10	343	hypothetical protein
TGME49_002220	19	190	hypothetical protein
TGGT1_018900	9	92	conserved hypothetical protein
TGGT1_110360	1	4	conserved hypothetical protein
TGME49_071480	1	4	hypothetical protein
TGME49_010350	9	114	hypothetical protein
TGGT1_053280	7	110	conserved hypothetical protein
TGGT1_017940	11	133	conserved hypothetical protein
TGME49_002980	10	117	hypothetical protein
TGME49_112480	20	1070	uracil phosphoribosyltransferase
TGGT1_088770	12	387	uracil phosphoribosyltransferase, putative
TGME49_110120	4	11	hypothetical protein, conserved
TGGT1_086240	2	7	conserved hypothetical protein
TGME49_007460	10	223	GTP-binding protein, putative
TGGT1_020440	4	30	conserved hypothetical protein
TGGT1_031360	5	42	poly(A) binding protein
TGME49_092200	3	24	poly(A) binding protein, putative
TGGT1_070080	5	168	conserved hypothetical protein
TGME49_035690	5	168	hypothetical protein
TGGT1_024310	1	1	SPX domain-containing protein, putative
TGME49_048550	1	1	SPX domain-containing protein
TGGT1_058390	2	9	conserved hypothetical protein
TGME49_037870	2	9	hypothetical protein, conserved
TGGT1_011660	22	1948	conserved hypothetical protein
TGME49_058540	18	1531	hypothetical protein
TGGT1_107210	2	22	conserved hypothetical protein
TGME49_069020	1	10	hypothetical protein
TGME49_030920	4	10	clathrin coat assembly protein AP50, putative
TGGT1_116900	2	6	clathrin coat associated protein ap-50, putative
TGGT1_100840	19	1321	hypothetical protein
TGME49_049260	16	1072	CMGC kinase
TGGT1_016890	10	412	conserved hypothetical protein
TGME49_003820	9	396	hypothetical protein
TGGT1_071420	12	203	conserved hypothetical protein
TGME49_034470	9	153	hypothetical protein
TGME49_060670	4	46	centrin, putative
TGGT1_009500	4	42	centrin, putative
TGME49_091290	92	27770	hypothetical protein, conserved
TGGT1_031920	91	33369	conserved hypothetical protein
TGGT1_124480	5	24	RAB geranylgeranyl transferase alpha subunit, putative
TGME49_014170	5	24	hypothetical protein
TGME49_080480	3	16	calmodulin, putative

TGME49_023930	12	232	RNA binding protein, putative
TGGT1_078310	8	211	RNA binding protein, putative
TGGT1_038750	8	40	conserved hypothetical protein
TGME49_013740	7	30	methyltransferase domain containing protein
TGME49_004120	6	45	hypothetical protein
TGGT1_016590	5	44	conserved hypothetical protein
TGME49_019320	58	157391	acid phosphatase, putative
TGGT1_030300	16	1045	acid phosphatase, putative
TGME49_120460	15	134	ABC transporter protein, putative
TGGT1_120930	5	51	multidrug resistance protein / ABC transporter, putative
TGGT1_083910	13	220	elongation factor Tu GTP binding domain-containing protein, putative
TGME49_028340	13	120	elongation factor Tu GTP binding domain-containing protein
TGME49_070780	4	32	RNA polymerase II subunit RPB6, putative
TGGT1_109540	3	28	conserved hypothetical protein
TGME49_035570	1	11	hypothetical protein
TGME49_018200	6	30	UDP-N-acetylglucosamine pyrophosphorylase, putative
TGGT1_028910	5	29	UDP-N-acteylglucosamine pyrophosphorylase, putative
TGGT1_026050	3	12	conserved hypothetical protein
TGME49_046960	1	2	hydrolase-like protein
TGME49_097500	56	5523	TCP-1/cpn60 family chaperonin, putative
TGGT1_096700	27	3491	chaperonin containing t-complex protein 1, eta subunit, tcph, putative
TGGT1_112320	5	25	zinc finger (C3HC4 RING finger) protein
TGME49_073150	4	18	zinc finger (C3HC4 RING finger) protein, putative
TGME49_026430	8	457	reticulon domain-containing protein
TGGT1_081600	7	195	reticulon domain-containing protein, putative
TGGT1_109760	1	2	glycerol-3-phosphate acyltransferase, putative
TGME49_070910	1	2	glycerol-3-phosphate acyltransferase, putative
TGME49_078830	87	7846	glucose-6-phosphate dehydrogenase, putative
TGGT1_103530	39	4876	glucose-6-phosphate dehydrogenase, putative
TGGT1_000540	28	1874	microneme protein, putative
TGME49_075800	28	1492	microneme protein, putative
TGME49_037880	8	125	hypothetical protein
TGGT1_058380	6	115	conserved hypothetical protein
TGME49_116200	4	9	hypothetical protein
TGGT1_093890	2	3	conserved hypothetical protein
TGGT1_125520	51	8334	conserved hypothetical protein
TGME49_014990	43	5882	hypothetical protein
TGME49_032520	8	79	brix domain-containing protein
TGGT1_115000	6	75	brix domain-containing protein, putative
TGME49_115490	27	1423	hypothetical protein
TGGT1_092950	15	600	conserved hypothetical protein
TGME49_009160	32	1304	glycylpeptide N-tetradecanoyltransferase, putative

TGGT1_022200	26	1265	glycylpeptide N-tetRadecanoyltransferase, putative
TGGT1_111920	2	6	small nuclear ribonucleoprotein, putative
TGME49_072630	2	6	small nuclear ribonucleoprotein, putative
TGGT1_046590	7	68	met-10 domain-containing protein, putative
TGME49_043280	7	68	met-10 domain-containing protein
TGGT1_064460	23	1696	conserved hypothetical protein
TGME49_121410	22	1337	hypothetical protein
TGME49_002070	7	106	hypothetical protein, conserved
TGME49_061240	32	1663	histone H3
TGGT1_009010	3	56	histone H3, putative
TGME49_027990	8	32	hypothetical protein
TGGT1_083270	2	6	conserved hypothetical protein
TGME49_113190	9	198	Ras family domain-containing protein
TGGT1_089480	2	14	small GTP-binding protein domain-containing protein
TGME49_047300	18	506	hypothetical protein
TGGT1_025700	17	620	NBP2B protein, putative
TGGT1_086440	1	1	conserved hypothetical protein
TGME49_110200	1	1	hypothetical protein
TGME49_068310	2	4	oocyst wall protein, putative
TGGT1_018870	4	16	conserved hypothetical protein
TGME49_002250	2	4	hypothetical protein
TGME49_009420	13	474	hypothetical protein
TGGT1_022360	9	440	conserved hypothetical protein
TGME49_033300	12	143	rhoGAP protein, putative
TGGT1_114180	9	202	rhoGAP protein, putative
TGME49_072370	14	67	hypothetical protein
TGGT1_111560	4	16	conserved hypothetical protein
TGME49_022860	59	4262	eukaryotic translation initiation factor 3 subunit 9, putative
TGGT1_066370	32	3360	hypothetical protein
TGME49_053420	2	9	hypothetical protein, conserved
TGGT1_106350	8	179	pyroglutamyl-peptidase I, putative
TGME49_068280	6	120	chromatin organization modifier domain-containing protein
TGGT1_022170	8	136	regulator of chromosome condensation RCC1 domain-containing protein
TGME49_009130	8	123	regulator of chromosome condensation RCC1 domain-containing protein
TGGT1_064590	8	125	hypothetical protein
TGME49_121540	4	39	hypothetical protein
TGGT1_110000	14	742	endonuclease/exonuclease/phosphatase domain-containing protein, putative
TGME49_071120	13	521	endonuclease/exonuclease/phosphatase domain-containing protein
TGGT1_101200	18	683	conserved hypothetical protein
TGME49_049550	17	626	hypothetical protein
TGME49_013680	8	32	MmgE/PrpD family domain containing protein

TGGT1_038830	2	6	MmgE/PrpD family protein
TGME49_093350	9	80	mitochondrial carrier domain-containing protein
TGGT1_073710	3	5	mitochondrial carrier domain-containing protein
TGGT1_041260	14	542	conserved hypothetical protein
TGME49_106190	12	446	hypothetical protein
TGGT1_050510	1	4	conserved hypothetical protein
TGME49_040350	1	4	hypothetical protein
TGME49_072520	16	375	hypothetical protein, conserved
TGGT1_111810	14	316	conserved hypothetical protein
TGGT1_001320	12	269	protein arginine N-methyltransferase, putative
TGME49_052420	9	166	hypothetical protein, conserved
TGGT1_110930	1	2	conserved hypothetical protein
TGME49_071920	1	2	hypothetical protein
TGME49_025480	7	59	hypothetical protein
TGGT1_080690	5	45	conserved hypothetical protein
TGME49_099970	9	388	TPR domain-containing protein
TGGT1_095540	3	106	TPR domain-containing protein, putative
TGME49_094550	90	5084	dynein heavy chain, putative
TGGT1_074850	55	4201	dynein heavy chain, putative
TGGT1_021400	7	72	conserved hypothetical protein
TGME49_008390	7	72	capping protein alpha-like subunit, putative
TGME49_027950	106	9810	zinc metalloprotease 2, putative
TGGT1_083230	52	6122	zinc metalloprotease, putative
TGME49_061070	2	5	hypothetical protein
TGGT1_009200	1	1	conserved hypothetical protein
TGGT1_000450	11	210	conserved hypothetical protein
TGME49_075990	8	77	hypothetical protein
TGME49_030960	69	6943	splicing factor 3B subunit 3, putative
TGGT1_116860	38	4138	splicing factor 3B subunit, putative
TGGT1_016740	6	41	dolichyl-di-phosphooligosaccharide-protein glycotransferase, putative
TGME49_003970	5	26	dolichyl-di-phosphooligosaccharide-protein glycotransferase, putative
TGME49_063180	106	89486	myosin D, putative
TGGT1_043280	19	1581	myosin D, putative
TGGT1_088100	28	2202	conserved hypothetical protein
TGME49_111860	18	1372	hypothetical protein
TGGT1_066010	1	1	hypothetical protein
TGME49_022200	1	1	hypothetical protein
TGME49_030690	36	1757	lipin, putative
TGGT1_117060	19	1129	lipin, putative
TGME49_033710	20	708	sulfate transporter, putative
TGGT1_113730	15	688	sulfate transporter, putative
TGME49_093370	18	147	arginine N-methyltransferase, putative
TGGT1_073730	2	6	protein arginine N-methyltransferase, putative

TGME49_097110	39	832	kinesin motor domain containing protein
TGGT1_097110	5	63	kinesin motor domain-containing protein, putative
TGME49_028390	7	60	hypothetical protein
TGGT1_083960	5	56	conserved hypothetical protein
TGGT1_078250	7	82	conserved hypothetical protein
TGME49_023880	7	82	hypothetical protein
TGGT1_017370	1	4	conserved hypothetical protein
TGME49_114970	46	1850	hypothetical protein
TGGT1_092270	19	702	conserved hypothetical protein
TGME49_032730	6	131	diacylglycerol acyltransferase, putative
TGGT1_114790	4	57	diacylglycerol acyltransferase, putative
TGME49_109070	5	50	hypothetical protein
TGGT1_056100	4	38	conserved hypothetical protein
TGME49_044470	41	3208	viral A-type inclusion protein repeat family
TGGT1_045460	21	1702	conserved hypothetical protein
TGGT1_002660	7	115	conserved hypothetical protein
TGME49_053490	6	92	hypothetical protein
TGME49_089690	152	105183	glyceraldehyde-3-phosphate dehydrogenase
TGGT1_033520	24	8755	glyceraldehyde-3-phosphate dehydrogenase, putative
TGME49_093580	15	459	hypothetical protein
TGGT1_073970	11	399	conserved hypothetical protein
TGME49_011630	13	1075	hypothetical protein
TGGT1_123260	12	1047	conserved hypothetical protein
TGME49_009030	125	43630	actin
TGGT1_021970	35	12151	actin, putative
TGME49_036860	12	164	cation-transporting ATPase, putative
TGGT1_069080	9	217	cation-transporting ATPase, putative
TGGT1_087390	4	15	conserved hypothetical protein
TGME49_111070	3	13	hypothetical protein, conserved
TGME49_063740	1	9	ATP-binding cassette, putative
TGME49_119730	9	202	hypothetical protein
TGGT1_121560	8	156	conserved hypothetical protein
TGME49_055650	3	87	hypothetical protein
TGGT1_077170	2	3	zinc finger protein DHHC domain containing protein, putative
TGME49_044560	187	953513	heat shock protein 90, putative
TGGT1_045380	54	15593	heat shock protein, putative
TGGT1_046550	11	209	conserved hypothetical protein
TGME49_043310	6	98	hypothetical protein
TGGT1_112860	4	13	conserved hypothetical protein
TGME49_073780	3	7	hypothetical protein, conserved
TGME49_005330	4	11	hypothetical protein
TGGT1_063580	3	10	conserved hypothetical protein
TGGT1_041820	1	2	conserved hypothetical protein
TGME49_105760	1	2	hypothetical protein
TGME49_087210	37	1582	proteasome subunit alpha type 2, putative

TGGT1_038310	10	401	proteasome subunit alpha type, putative
TGGT1_060640	2	6	deoxycytidylate deaminase, putative
TGME49_000420	1	2	hypothetical protein
TGGT1_026890	37	1418	hypothetical protein
TGME49_046170	34	902	ARID/BRIGHT DNA-binding domain-containing protein
TGGT1_013970	1	2	conserved hypothetical protein
TGME49_056780	1	2	hypothetical protein, conserved
TGGT1_010080	16	1464	mannose-6-phosphate isomerase, putative
TGME49_060140	11	1089	mannose-6-phosphate isomerase, putative
TGME49_029670	12	3347	ribosomal protein S23, putative
TGGT1_118100	4	105	ribosomal protein S23, putative
TGGT1_084290	7	78	conserved hypothetical protein
TGME49_028710	3	15	hypothetical protein
TGGT1_047240	4	30	60S ribosomal protein L24, putative
TGME49_042680	3	13	60S ribosomal protein L24, putative
TGME49_030420	99	102350	calcium-transporting ATPase, putative
TGGT1_117550	34	3875	calcium-transporting ATPase, putative
TGGT1_041220	11	135	CDC73 domain protein, putative
TGME49_106210	10	131	hypothetical protein
TGME49_004480	28	1251	DnaJ domain-containing protein
TGGT1_016200	15	771	DnaJ domain-containing protein, putative
TGME49_007770	29	1071	proteasome PCI domain-containing protein
TGGT1_020650	17	855	proteasome PCI domain-containing protein, putative
TGME49_081920	6	31	vacuolar ATP synthase subunit D, putative
TGGT1_005550	4	30	vacuolar ATP synthase subunit D, putative
TGME49_007640	54	3952	isoleucine-tRNA synthetase, putative
TGGT1_020530	30	2880	isoleucine-tRNA synthetase, putative
TGME49_070640	5	43	RNA binding motif-containing protein
TGGT1_109290	5	32	hypothetical protein
TGME49_054710	8	68	hypothetical protein
TGGT1_004260	6	64	conserved hypothetical protein
TGME49_060260	22	1039	60S acidic ribosomal protein P1, putative
TGGT1_009950	7	423	60S acidic ribosomal protein P1, putative
TGME49_078820	2	7	F-box domain-containing protein
TGGT1_103540	1	6	F-box domain-containing protein, putative
TGME49_058870	56	5866	hypothetical protein
TGGT1_011200	25	3197	conserved hypothetical protein
TGME49_115320	28	5028	SRS52A
TGGT1_092760	16	1828	SRS52A
TGGT1_056570	1	2	conserved hypothetical protein
TGME49_066310	1	2	hypothetical protein
TGME49_072670	4	12	peptidase family M3 domain containing protein
TGGT1_111970	3	15	mitochondrial intermediate peptidase, putative
TGME49_048440	3	13	hypothetical protein
TGME49_069870	45	2068	hypothetical protein, conserved

TGGT1_108310	21	767	conserved hypothetical protein
TGGT1_115550	3	41	conserved hypothetical protein
TGME49_032110	3	22	hypothetical protein
TGME49_115780	13	297	myosin regulatory light chain, putative
TGGT1_093480	10	158	myosin regulatory light chain, putative
TGME49_028170	143	65212	serine/threonine protein phosphatase, putative
TGGT1_083640	53	5003	serine/threonine protein phosphatase, putative
TGGT1_117880	39	3342	conserved hypothetical protein
TGME49_030000	38	2614	hypothetical protein
TGGT1_123350	3	9	conserved hypothetical protein
TGME49_011440	3	9	hypothetical protein, conserved
TGGT1_083500	2	12	conserved hypothetical protein
TGME49_028130	2	12	hypothetical protein
TGGT1_097980	22	1235	transcriptional adaptor, putative
TGME49_017050	20	842	transcriptional co-activator ADA2-A
TGME49_110030	18	1336	adenylyl cyclase associated protein
TGGT1_086070	11	694	adenylyl cyclase associated protein, putative
TGGT1_017000	5	45	conserved hypothetical protein
TGME49_003730	5	45	hypothetical protein
TGME49_109570	13	199	nucleolar RNA helicase, putative
TGGT1_085530	3	43	conserved hypothetical protein
TGGT1_068360	2	6	conserved hypothetical protein
TGME49_037530	2	6	hypothetical protein
TGGT1_006050	1	1	conserved hypothetical protein
TGME49_081400	1	1	hypothetical protein, conserved
TGME49_085240	3	93	trans-2,3-enoyl-CoA reductase, putative
TGGT1_010800	1	1	ubiquitin-conjugating enzyme domain-containing protein, putative
TGME49_059090	1	1	ubiquitin-conjugating enzyme domain-containing protein
TGME49_006320	10	270	hypothetical protein
TGGT1_062470	9	257	conserved hypothetical protein
TGME49_058850	19	577	hypothetical protein
TGGT1_011320	1	1	conserved hypothetical protein
TGGT1_058200	9	94	glutamate 5-kinase, putative
TGME49_065010	5	51	glutamate 5-kinase, putative
TGME49_031960	36	2451	ppg3, putative
TGGT1_115800	34	3480	hypothetical protein
TGME49_033100	39	8401	prohibitin, putative
TGGT1_114410	12	544	prohibitin, putative
TGGT1_068070	7	49	conserved hypothetical protein
TGME49_075430	3	27	hypothetical protein
TGME49_015950	60	1074	hypothetical protein
TGGT1_099310	12	209	hypothetical protein
TGGT1_025370	8	162	ABC transporter, putative
TGME49_047540	7	138	ABC transporter, putative

TGME49_047550	84	77614	heat shock protein 60
TGGT1_025360	33	8079	heat shock protein, putative
TGME49_029210	16	179	U2 small nuclear ribonucleoprotein, putative
TGGT1_118630	8	101	U2 small nuclear ribonucleoprotein, putative
TGGT1_103410	7	127	ubiquitin--protein ligase, putative
TGME49_078940	7	127	ubiquitin--protein ligase, putative
TGME49_047220	16	620	hypothetical protein
TGGT1_025780	15	686	hypothetical protein
TGME49_039110	1	1	leucine rich repeat protein, putative
TGME49_044080	6	68	hypothetical protein
TGGT1_045870	1	4	conserved hypothetical protein
TGGT1_087740	3	18	RNA (guanine-9-) methyltransferase domain-containing protein
TGME49_111390	3	18	tRNA (guanine-N1)-methyltransferase, putative
TGME49_069440	26	2466	calmodulin, putative
TGGT1_107770	7	343	calmodulin, putative
TGME49_088530	18	518	NOL1/NOP2/sun family domain-containing protein
TGGT1_034860	16	467	conserved hypothetical protein
TGME49_119620	11	281	hypothetical protein
TGGT1_121670	9	273	conserved hypothetical protein
TGME49_048850	24	730	methionine aminopeptidase, putative
TGGT1_100520	15	615	methionine aminopeptidase, putative
TGME49_065310	28	1092	heat shock protein 40, putative
TGGT1_057800	14	592	DnaJ domain-containing protein, putative
TGME49_007240	5	58	hypothetical protein
TGGT1_023480	5	57	conserved hypothetical protein
TGME49_104680	18	1333	DNA-damage inducible protein, putative
TGGT1_006730	17	1310	conserved hypothetical protein
TGME49_100310	25	1158	26S protease regulatory subunit 6a, putative
TGGT1_094980	22	1377	hypothetical protein
TGME49_057380	15	544	hypothetical protein
TGGT1_013320	10	737	conserved hypothetical protein
TGGT1_028340	2	5	conserved hypothetical protein
TGME49_017700	1	3	hypothetical protein
TGGT1_060670	1	12	conserved hypothetical protein
TGME49_000450	1	12	hypothetical protein
TGME49_049870	13	120	C2 domain-containing protein
TGGT1_101560	5	39	C2 domain-containing protein, putative
TGME49_028770	18	406	hypothetical protein, conserved
TGGT1_084350	15	493	conserved hypothetical protein
TGME49_067370	24	423	kinesin motor domain-containing protein
TGGT1_075670	14	316	kinesin motor domain-containing protein, putative
TGGT1_006210	3	10	conserved hypothetical protein
TGME49_098980	3	10	RNA pseudouridylate synthase, putative
TGME49_067280	9	25	hypothetical protein

TGGT1_075760	2	3	conserved hypothetical protein
TGGT1_063910	46	6668	NBP2B protein, putative
TGME49_005130	42	5575	glyoxalase, putative
TGGT1_006080	1	2	poly(A) polymerase cid, putative
TGME49_081370	1	2	hypothetical protein, conserved
TGME49_025270	17	268	hypothetical protein
TGGT1_080470	11	208	hypothetical protein
TGME49_049540	3	11	hypothetical protein
TGGT1_101190	2	16	conserved hypothetical protein
TGME49_021520	20	546	coatomer delta subunit, putative
TGGT1_065380	15	481	coatomer delta subunit, putative
TGGT1_002730	11	210	conserved hypothetical protein
TGME49_053560	9	154	hypothetical protein
TGME49_063870	37	2372	glutamyl-tRNA synthetase, putative
TGGT1_044090	23	2862	glutamyl-tRNA synthetase, putative
TGME49_011270	9	67	sushi domain-containing protein / SCR repeat-containing protein
TGGT1_083290	3	12	conserved hypothetical protein
TGME49_028010	3	12	hypothetical protein
TGGT1_045150	6	43	conserved hypothetical protein
TGME49_044670	4	17	hypothetical protein, conserved
TGGT1_016880	2	5	zinc finger (C3HC4 type) protein
TGME49_003830	1	2	zinc finger (C3HC4 type) / FHA domain-containing protein
TGGT1_076650	14	292	conserved hypothetical protein
TGME49_055240	13	208	hypothetical protein
TGME49_091300	3	23	RER1 protein, putative
TGGT1_031915	1	15	RER1 protein, putative
TGGT1_071650	35	3994	conserved hypothetical protein
TGME49_034240	29	2942	hypothetical protein
TGGT1_046450	14	299	chaperone binding protein, putative
TGME49_043410	13	237	hypothetical protein
TGGT1_016320	3	18	conserved hypothetical protein
TGME49_004390	2	10	hypothetical protein
TGME49_038950	25	664	fatty acyl-CoA desaturase, putative
TGGT1_052030	21	647	fatty acyl-CoA desaturase, putative
TGGT1_105330	3	15	membrane associated progesterone receptor, putative
TGME49_076990	3	8	cytochrome b5-like Heme/Steroid binding domain containing protein
TGME49_114700	6	38	hypothetical protein
TGGT1_091980	3	33	conserved hypothetical protein
TGME49_057160	2	25	hypothetical protein
TGGT1_013550	1	4	hypothetical protein
TGGT1_041960	10	247	conserved hypothetical protein
TGME49_105510	8	183	hypothetical protein
TGME49_066020	2	3	hypothetical protein

TGGT1_056750	1	2	conserved hypothetical protein
TGGT1_017710	3	5	conserved hypothetical protein
TGME49_003160	3	5	hypothetical protein, conserved
TGME49_066400	5	25	hypothetical protein
TGGT1_095510	1	2	conserved hypothetical protein
TGGT1_008500	1	1	conserved hypothetical protein
TGME49_061730	1	1	hypothetical protein
TGME49_034430	21	131	ATPase, AAA family domain-containing protein
TGME49_076910	2	10	endoplasmic reticulum retention receptor, putative
TGME49_064790	8	177	hypothetical protein
TGME49_064440	20	1315	signal recognition particle receptor subunit beta, putative
TGGT1_081630	18	813	conserved hypothetical protein
TGME49_026470	15	558	hypothetical protein
TGME49_014140	37	3792	hypothetical protein
TGME49_064860	28	656	zinc finger (C3HC4 RING finger) protein, putative
TGME49_067480	19	314	N2,N2-dimethylguanosine tRNA methyltransferase, putative
TGME49_066390	1	1	DNA mismatch repair protein, putative
TGME49_118520	10	184	hypothetical protein
TGGT1_122600	10	178	conserved hypothetical protein
TGME49_120190	10	164	SRS16B (= SRS9)
TGGT1_121100	7	68	SRS16B (= SRS9)
TGME49_067670	4	32	hypothetical protein
TGGT1_065840	1	1	NLI interacting factor-like phosphatase domain-containing protein
TGME49_021200	26	1437	CW-type zinc finger domain-containing protein
TGGT1_065020	17	892	zinc finger (CW-type) protein
TGME49_067810	16	1315	Rab 5
TGME49_120020	31	11945	major facilitator superfamily domain-containing protein
TGGT1_121370	14	8274	conserved hypothetical protein
TGGT1_011580	1	1	hypothetical protein
TGGT1_083310	9	86	hypothetical protein
TGME49_028040	6	43	peptidyl-prolyl cis-trans isomerase, putative
TGME49_101420	12	341	hypothetical protein, conserved
TGGT1_059860	9	130	conserved hypothetical protein
TGGT1_003655	1	2	MSF-1 like conserved region domain-containing protein
TGME49_054250	1	2	MSF1-like conserved region domain-containing protein
TGME49_062790	4	56	hypothetical protein
TGME49_055630	1	25	hypothetical protein
TGME49_091610	1	10	hypothetical protein
TGGT1_031370	3	12	conserved hypothetical protein
TGME49_092190	3	8	hypothetical protein
TGME49_121360	4	10	RNA recognition motif domain-containing protein
TGGT1_064400	1	1	RNA recognition motif domain-containing protein, putative

TGME49_064450	23	708	DNA topoisomerase III beta-1, putative
TGME49_029350	1	2	hypothetical protein
TGME49_090860	3	9	amino acid transporter, putative
TGME49_065020	10	97	hypothetical protein
TGGT1_058190	5	68	conserved hypothetical protein
TGME49_064880	13	163	ubiquitin-activating enzyme, putative
TGGT1_001960	1	19	hypothetical protein
TGME49_052990	1	19	hypothetical protein
TGME49_027300	2	3	hypothetical protein
TGGT1_082690	1	1	hypothetical protein
TGME49_003600	9	253	hypothetical protein
TGME49_089710	23	2024	hypothetical protein
TGME49_079370	1	2	SNARE associated Golgi protein family domain-containing protein
TGME49_112710	3	16	hypothetical protein
TGME49_060600	44	4020	pumilio-family RNA binding repeat domain-containing protein
TGGT1_009580	19	1203	conserved hypothetical protein
TGME49_076220	15	303	hypothetical protein
TGME49_026780	4	20	zinc finger (C3HC4 RING finger) protein, putative
TGGT1_082060	1	6	conserved hypothetical protein
TGME49_051460	5	81	hypothetical protein
TGGT1_102860	2	22	hypothetical protein
TGME49_110310	13	91	WD domain, G-beta repeat-containing protein
TGME49_097790	5	10	hypothetical protein
TGME49_065820	3	4	hypothetical protein
TGGT1_057060	1	1	NBP2B protein, putative
TGME49_078790	6	106	hypothetical protein
TGME49_011710	6	197	TB2/DP1, HVA22 domain-containing protein
TGGT1_123170	5	175	TB2/DP1, HVA22 domain-containing protein, putative
TGGT1_096195	1	2	hypothetical protein
TGME49_097910	1	2	hypothetical protein
TGME49_064770	10	533	hypothetical protein
TGME49_108950	8	95	histidine acid phosphatase domain containing protein
TGME49_030120	1	1	hypothetical protein, conserved
TGME49_063660	8	443	hypothetical protein
TGGT1_043930	4	197	conserved hypothetical protein
TGME49_011100	1	2	hypothetical protein
TGME49_064690	1	3	cyclin, putative
TGME49_027030	13	344	hypothetical protein
TGGT1_082520	8	262	conserved hypothetical protein
TGGT1_109460	2	14	conserved hypothetical protein
TGME49_070700	2	14	hypothetical protein
TGME49_089800	15	461	transmembrane trafficking protein, transmembrane protein Tmp21 precursor, putative

TGME49_026510	24	1151	sec23/Sec24 helical domain-containing protein
TGGT1_081660	23	1465	sec23/Sec24 helical domain-containing protein, putative
TGGT1_017920	8	189	serine/threonine-protein kinase, putative
TGME49_003010	5	120	Aurora kinase
TGGT1_047680	1	2	hypothetical protein
TGME49_042240	1	2	Rhoptry kinase family protein ROP19
TGGT1_078740	22	1213	protein phosphatase 2C, putative
TGME49_024240	19	882	protein phosphatase 2C, putative
TGME49_027410	5	20	hypothetical protein
TGGT1_082910	5	19	hypothetical protein
TGGT1_108190	8	68	chromatin organization modifier domain-containing protein
TGME49_069760	7	41	chromatin organization modifier domain-containing protein
TGME49_079330	10	64	ATP-dependent RNA helicase, putative
TGME49_064840	1	3	DEAD/DEAH box helicase, putative
TGME49_005360	33	1108	hypothetical protein
TGGT1_063550	7	71	conserved hypothetical protein
TGME49_088040	19	344	hypothetical protein
TGME49_034230	22	503	hypothetical protein, conserved
TGGT1_034140	5	36	conserved hypothetical protein
TGME49_089110	5	36	hypothetical protein
TGGT1_081020	20	554	conserved hypothetical protein
TGME49_025870	19	433	hypothetical protein
TGME49_113820	8	96	dynactin p25, putative
TGGT1_090320	1	3	dynactin, putative
TGME49_065170	3	20	hypothetical protein
TGGT1_058060	2	10	hypothetical protein
TGME49_043240	2	4	G-protein beta-subunit, putative
TGME49_043910	18	702	haloacid dehalogenase-like hydrolase domain-containing protein
TGGT1_046050	13	623	conserved hypothetical protein
TGME49_109900	11	179	hypothetical protein
TGME49_077230	1	4	hypothetical protein
TGME49_049200	1	1	ctr copper transporter domain-containing protein
TGME49_005250	48	4818	Rhoptry kinase family protein ROP18
TGGT1_063760	32	2203	rhoptry protein / protein kinase domain-containing protein, putative
TGME49_063760	1	1	hypothetical protein
TGME49_079340	4	48	hypothetical protein
TGGT1_123290	4	14	conserved hypothetical protein
TGME49_011600	4	14	hypothetical protein
TGME49_064710	4	25	hypothetical protein
TGME49_070920	2	3	Rhoptry kinase family protein ROP32
TGGT1_109770	1	2	kinase domain-containing protein, putative
TGME49_047430	1	1	hypothetical protein
TGGT1_022700	4	38	conserved hypothetical protein

TGME49_009640	4	38	hypothetical protein
TGME49_120180	15	463	SRS16C (= BSR4, P36)
TGGT1_121110	9	144	SRS16C (= BSR4, P36)
TGME49_023460	2	5	hypothetical protein
TGME49_063160	1	1	hypothetical protein
TGME49_111010	2	17	hypothetical protein, conserved
TGGT1_063265	2	7	200 kDa antigen p200, putative
TGME49_005640	2	7	200 kDa antigen p200, putative
TGME49_079430	5	22	hypothetical protein
TGGT1_003680	3	15	hypothetical protein
TGME49_054260	2	8	hypothetical protein
TGME49_069210	1	1	hypothetical protein
TGME49_044840	1	4	zinc knuckle domain-containing protein
TGME49_020270	29	4719	hypothetical protein, conserved
TGGT1_014960	15	1848	conserved hypothetical protein
TGME49_067640	8	111	hypothetical protein
TGME49_020510	14	326	hypothetical protein
TGGT1_015240	3	30	conserved hypothetical protein
TGME49_009120	1	5	hypothetical protein
TGME49_113030	4	48	hypothetical protein
TGGT1_045980	1	1	conserved hypothetical protein
TGME49_043980	1	1	hypothetical protein, conserved
TGME49_067600	10	74	RNA recognition motif-containing protein
TGME49_009870	5	40	pyridoxal phosphate phosphatase, putative
TGME49_063700	25	3319	40s ribosomal protein S14, putative
TGGT1_043965	11	969	40S ribosomal protein S14, putative
TGME49_099820	6	57	hypothetical protein
TGGT1_095780	2	12	conserved hypothetical protein
TGME49_066030	10	250	PRP38 family domain-containing protein
TGGT1_056740	9	205	conserved hypothetical protein
TGME49_062040	23	1204	eukaryotic translation initiation factor 3 subunit 11, putative
TGGT1_008065	14	986	eukaryotic translation initiation factor 3 subunit 11, putative
TGGT1_070960	10	308	hypothetical protein
TGME49_035140	1	9	hypothetical protein
TGME49_029030	12	514	hypothetical protein, conserved
TGGT1_118710	10	267	conserved hypothetical protein
TGME49_035150	9	95	transporter, putative
TGGT1_070955	8	85	transporter, putative
TGME49_022370	3	7	SRS13
TGGT1_066180	2	8	SRS13
TGME49_028700	1	25	hypothetical protein
TGME49_067440	7	137	zinc finger (C3HC4 RING finger) protein, putative
TGME49_064750	12	140	hypothetical protein
TGME49_079450	13	99	adenylosuccinate synthetase, putative
TGME49_048520	5	36	hypothetical protein, conserved

TGME49_078390	4	33	hypothetical protein
TGGT1_104170	2	5	conserved hypothetical protein
TGGT1_043190	3	14	hypothetical protein
TGME49_078370	2	20	hypothetical protein
TGGT1_053770	25	2865	conserved hypothetical protein
TGME49_064740	2	8	hypothetical protein
TGME49_047800	1	1	hypothetical protein
TGGT1_039990	9	151	conserved hypothetical protein
TGME49_012790	8	73	hypothetical protein
TGME49_064890	24	1406	hypothetical protein
TGME49_007430	13	233	ATP-dependent RNA helicase DDX family protein, putative
TGGT1_020410	11	260	ATP-dependent helicase / DEAD box protein, putative
TGGT1_076620	1	2	hypothetical protein
TGGT1_043250	3	12	DNA excision repair protein ERCC6, putative
TGME49_063140	1	4	DNA excision repair protein, putative
TGME49_054000	7	75	hypothetical protein
TGGT1_003280	6	71	hypothetical protein
TGGT1_030820	1	2	hypothetical protein
TGGT1_008150	7	104	hypothetical protein
TGME49_061960	6	97	hypothetical protein
TGME49_114820	2	17	ATP synthase epsilon chain, putative
TGGT1_092110	2	12	ATP synthase epsilon chain, putative
TGME49_106380	9	129	U1 small nuclear ribonucleoprotein, putative
TGME49_044370	19	796	hypothetical protein
TGGT1_045570	13	468	conserved hypothetical protein
TGVEG_109480	6	24	glutaredoxin-1, grx1, putative
TGME49_079400	3	43	glutaredoxin, putative
TGME49_043640	1	2	hypothetical protein
TGGT1_076460	1	1	M3 protein, putative
TGME49_067730	4	63	GDP-fucose transporter 1, putative
TGGT1_031320	11	127	zinc finger (C2H2 type) protein, putative
TGME49_092240	9	82	hypothetical protein, conserved
TGME49_077270	34	3220	nucleoside-triphosphatase II, putative
TGME49_010300	17	1149	zinc finger (CCCH type) protein, putative
TGME49_105230	2	6	hypothetical protein
TGME49_060580	13	114	hypothetical protein
TGGT1_009600	8	66	conserved hypothetical protein
TGME49_064720	3	5	hypothetical protein
TGME49_110090	1	1	hypothetical protein
TGME49_038080	3	5	hypothetical protein, conserved
TGME49_009270	3	6	hypothetical protein
TGGT1_037380	1	1	hypothetical protein
TGGT1_051870	2	4	hypothetical protein
TGME49_054220	7	52	hypothetical protein
TGME49_087980	21	969	forkhead-aasociated domain-containing protein

TGGT1_081240	2	5	DEAD box ATP-dependent RNA helicase, putative
TGME49_026090	2	5	RNA helicase, putative
TGGT1_050620	36	3743	conserved hypothetical protein
TGME49_040220	30	3244	hypothetical protein
TGME49_082140	7	34	hypothetical protein
TGME49_073130	19	3594	SRS30A
TGGT1_112300	9	257	SRS30A
TGME49_088030	10	793	hypothetical protein
TGME49_067580	27	1119	cyclin, N-terminal domain-containing protein
TGME49_088000	36	805	hypothetical protein
TGGT1_068130	8	88	SRS47E
TGME49_075370	7	92	SRS47E
TGME49_046760	1	4	hypothetical protein
TGME49_006410	7	20	hypothetical protein
TGGT1_062380	2	14	conserved hypothetical protein
TGME49_060240	14	152	CCR4-NOT transcription complex subunit, putative
TGGT1_009970	10	125	CCR4-NOT transcription complex subunit, putative
TGGT1_050960	3	55	hypothetical protein
TGME49_020210	1	2	hypothetical protein
TGME49_024030	2	3	hypothetical protein
TGGT1_078510	1	2	conserved hypothetical protein
TGGT1_018030	5	219	hypothetical protein
TGME49_089700	1	2	hypothetical protein
TGME49_113080	43	5963	hypothetical protein
TGGT1_034050	13	202	hypothetical protein
TGME49_089190	11	150	TPR domain-containing protein
TGME49_020950	21	4146	hypothetical protein
TGME49_001880	7	65	hypothetical protein
TGGT1_019340	4	29	conserved hypothetical protein
TGME49_013280	13	1080	hypothetical protein
TGGT1_039470	10	265	conserved hypothetical protein
TGME49_067430	18	535	DnaJ domain-containing protein
TGGT1_098910	1	1	conserved hypothetical protein
TGME49_016340	1	1	hypothetical protein
TGME49_054540	5	18	CDP-alcohol phosphatidyltransferase domain-containing protein
TGGT1_004110	1	3	cdp-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase, putative
TGME49_078190	3	18	hypothetical protein, conserved
TGGT1_104370	1	2	hypothetical protein
TGGT1_123230	2	4	conserved hypothetical protein
TGME49_011660	2	4	hypothetical protein
TGGT1_065100	1	4	DAHP synthetase, putative
TGME49_021260	1	4	DAHP synthetase, putative
TGGT1_106370	1	2	conserved hypothetical protein

TGME49_068300	1	2	hypothetical protein
TGME49_092320	15	164	hypothetical protein
TGGT1_031250	1	3	hypothetical protein
TGME49_067850	2	2	hypothetical protein
TGGT1_005730	4	58	aldo-keto reductase, putative
TGME49_081620	4	58	aldo/keto reductase family oxidoreductase, putative
TGME49_070840	15	224	NAD(+) ADP-ribosyltransferase, putative
TGGT1_109595	3	26	hypothetical protein
TGME49_050850	16	955	CMGC kinase, putative
TGGT1_102410	5	224	hypothetical protein
TGGT1_031290	13	454	SRS36D (= SAG5C)
TGME49_092280	10	305	SRS36D (= SAG5C)
TGGT1_112060	1	1	conserved hypothetical protein
TGME49_072880	1	1	hypothetical protein, conserved
TGME49_064780	5	58	UDP-N-acetylhexosamine pyrophosphorylase, putative
TGME49_084180	5	96	hypothetical protein
TGME49_067750	12	102	hypothetical protein
TGME49_111150	2	9	hypothetical protein, conserved
TGME49_087970	18	1021	hypothetical protein
TGME49_109720	17	151	hypothetical protein
TGME49_059050	1	1	hypothetical protein
TGME49_071420	2	3	porphobilinogen deaminase, putative
TGME49_009800	15	4302	formate/nitrite transporter, putative
TGGT1_022840	7	1584	conserved hypothetical protein
TGME49_057600	1	3	hypothetical protein
TGME49_024790	1	1	SRS40A
TGGT1_083200	21	627	conserved hypothetical protein
TGME49_027900	17	376	hypothetical protein
TGME49_060650	2	6	hypothetical protein
TGGT1_009520	1	3	conserved hypothetical protein
TGME49_027890	2	16	hypothetical protein
TGME49_064830	38	1381	hypothetical protein
TGME49_114380	2	7	hypothetical protein
TGME49_100230	1	1	hypothetical protein
TGME49_048680	8	42	hypothetical protein
TGGT1_069070	6	145	hypothetical protein
TGME49_114880	4	139	hypothetical protein
TGME49_010320	1	2	SRS37A
TGME49_060150	8	49	hypothetical protein
TGME49_024850	44	2015	polyadenylate-binding protein, putative
TGME49_067840	7	29	TPR domain containing protein
TGGT1_000560	3	12	conserved hypothetical protein
TGME49_075770	1	2	hypothetical protein
TGGT1_028510	3	26	hypothetical protein
TGME49_067650	4	8	hypothetical protein

TGME49_118160	18	3264	major sperm protein domain-containing protein
TGGT1_123090	9	749	major sperm protein domain-containing protein, putative
TGME49_037520	30	1132	hypothetical protein
TGGT1_068370	7	235	hypothetical protein
TGME49_106300	5	33	hypothetical protein
TGME49_106280	3	12	hypothetical protein
TGGT1_037700	16	762	hypothetical protein
TGME49_086280	15	579	hypothetical protein
TGME49_088230	1	2	hypothetical protein
TGME49_067570	1	2	hypothetical protein
TGME49_002200	41	3003	hypothetical protein
TGGT1_018920	19	1014	conserved hypothetical protein
TGME49_032270	1	1	hypothetical protein
TGGT1_093570	2	3	conserved hypothetical protein
TGME49_115880	1	2	hypothetical protein
TGME49_033870	2	18	hypothetical protein
TGME49_008550	4	24	hypothetical protein
TGME49_014070	1	8	hypothetical protein
TGGT1_065360	13	514	dual specificity protein phosphatase, catalytic domain-containing protein, putative
TGME49_021500	6	198	dual specificity protein phosphatase, catalytic domain-containing protein
TGME49_064670	14	196	DNA polymerase zeta catalytic subunit, putative
TGME49_066370	3	19	BT1 transmembrane domain-containing protein
TGME49_067450	1	1	hypothetical protein
TGME49_063170	5	24	hypothetical protein, conserved
TGGT1_082000	10	272	translin, putative
TGME49_026720	10	272	translin, putative
TGME49_064760	15	282	oxysterol-binding protein, putative
TGME49_110720	2	6	hypothetical protein
TGME49_099100	10	252	hypothetical protein
TGGT1_084020	2	4	conserved hypothetical protein
TGME49_067550	18	457	hypothetical protein, conserved
TGME49_113780	3	13	hypothetical protein
TGME49_075790	2	2	microneme protein, putative
TGME49_067710	16	174	hypothetical protein
TGGT1_008775	2	7	snRNA-associated Sm-like protein, putative
TGME49_061470	1	1	U6 snRNA-associated Sm-like protein, putative
TGGT1_071300	1	2	conserved hypothetical protein
TGME49_067720	22	2043	hypothetical protein
TGGT1_110320	9	500	hypothetical protein
TGME49_071460	8	432	hypothetical protein
TGME49_078650	3	13	hypothetical protein
TGME49_067740	2	10	hypothetical protein
TGME49_067800	21	424	dynamin-like protein, putative

TGME49_090230	1	3	hypothetical protein
TGME49_070380	18	1641	40S ribosomal protein S13, putative
TGGT1_108910	8	209	40S ribosomal protein S13, putative
TGME49_079390	49	5028	proliferation-associated protein 2G4, putative
TGME49_023640	2	3	hypothetical protein
TGME49_067390	7	95	DNA-directed RNA polymerases I and III subunit RPAC1, putative
TGME49_012860	9	94	hypothetical protein
TGGT1_039900	3	27	conserved hypothetical protein
TGGT1_068890	6	46	hypothetical protein
TGME49_109080	3	5	hypothetical protein
TGGT1_068510	31	2790	hypothetical protein
TGME49_037290	28	2619	hypothetical protein
TGGT1_065150	1	4	hypothetical protein
TGME49_067510	4	36	hypothetical protein, conserved
TGME49_067560	6	32	protein phosphatases PP1 regulatory subunit, putative
TGGT1_018620	10	287	zinc finger (C2H2 type) protein
TGME49_002470	8	222	zinc finger (C2H2 type) domain-containing protein
TGGT1_118065	5	35	hypothetical protein
TGME49_029710	3	15	OTU-like cysteine protease domain-containing protein
TGME49_064640	25	1453	bromodomain domain-containing protein
TGME49_120230	5	18	SRS15C
TGGT1_121060	3	6	SRS15C
TGGT1_080120	3	4	programmed cell death protein, putative
TGME49_067760	1	1	hypothetical protein
TGME49_053960	1	1	short chain dehydrogenase domain-containing protein
TGME49_004160	65	5855	hypothetical protein
TGGT1_016450	51	7525	conserved hypothetical protein
TGME49_067700	7	109	hypothetical protein
TGME49_024650	7	41	hypothetical protein
TGGT1_079370	5	37	conserved hypothetical protein
TGME49_067770	2	6	hypothetical protein
TGME49_113630	2	5	hypothetical protein
TGGT1_080650	1	2	hypothetical protein
TGGT1_081115	1	3	SOH1 domain-containing protein
TGME49_025970	1	3	SOH1 domain-containing protein
TGGT1_043220	1	1	NBP2B protein, putative
TGME49_097740	42	1564	hypothetical protein
TGGT1_096445	10	230	hypothetical protein
TGGT1_081800	1	1	conserved hypothetical protein
TGME49_043800	45	3832	long-chain fatty acid CoA ligase, putative
TGGT1_046060	29	2741	long-chain fatty acid CoA ligase, putative
TGME49_046930	13	175	calmodulin, putative
TGGT1_096670	12	171	conserved hypothetical protein

TGME49_097520	10	102	hypothetical protein
TGME49_064820	6	87	WD-40 repeat protein, putative
TGME49_043600	19	504	acetyltransferase domain-containing protein
TGGT1_046310	18	558	acetyltransferase domain-containing protein, putative
TGME49_002170	1	3	hypothetical protein
TGME49_005570	31	1214	hypothetical protein, conserved
TGME49_086630	20	887	peroxiredoxin family protein/glutaredoxin, putative
TGGT1_038055	9	428	peroxiredoxin family protein, putative
TGME49_059900	5	36	hypothetical protein, conserved
TGME49_033030	21	429	myosin-A docking protein, putative
TGGT1_114490	4	28	myosin-A docking protein, putative
TGGT1_035920	1	4	jmjC domain-containing protein, putative
TGME49_083890	1	4	hypothetical protein, conserved
TGGT1_088370	1	3	conserved hypothetical protein
TGME49_112130	1	3	hypothetical protein
TGGT1_043760	3	31	conserved hypothetical protein
TGME49_063540	3	31	hypothetical protein
TGGT1_065280	1	1	conserved hypothetical protein
TGME49_021430	1	1	hypothetical protein
TGME49_019480	1	1	hypothetical protein
TGME49_112560	2	12	hypothetical protein, conserved
TGME49_067540	4	21	AGC kinase
TGGT1_066445	2	6	sedlin, putative
TGME49_022920	2	6	trafficking protein particle complex protein, putative
TGGT1_039280	11	258	conserved hypothetical protein
TGME49_013370	11	258	hypothetical protein
TGGT1_018880	1	1	conserved hypothetical protein
TGGT1_030740	9	129	conserved hypothetical protein
TGME49_019740	6	81	hypothetical protein
TGGT1_062230	6	82	conserved hypothetical protein
TGME49_006530	4	64	hypothetical protein
TGME49_070350	1	1	hypothetical protein
TGME49_085290	3	48	hypothetical protein
TGME49_064660	131	26192	SRS domain-containing, N-acetylglucosamine-phosphate mutase, putative
TGME49_016980	1	4	hypothetical protein
TGME49_020220	1	2	hypothetical protein
TGME49_067680	15	297	microneme protein, putative
TGME49_003210	1	12	hypothetical protein
TGGT1_004160	6	91	hypothetical protein
TGME49_054610	6	91	tim10/DDP zinc finger domain-containing protein
TGGT1_022010	1	2	conserved hypothetical protein
TGME49_009070	1	2	hypothetical protein
TGME49_018610	16	510	UPF0202 family protein

TGGT1_029460	13	286	UPF0202 family protein
TGME49_108010	15	451	hypothetical protein, conserved
TGGT1_042795	5	59	hypothetical protein
TGME49_095110	119	228140	Rhoptry kinase family protein ROP4 / ROP7 (incomplete catalytic triad) ***WARNING: GENE MODEL INACCURATE ***
TGGT1_075470	12	1526	rhoptry antigen ROP8, putative
TGME49_080490	35	1774	ubiquitin conjugation factor, putative
TGGT1_072820	27	1700	conserved hypothetical protein
TGME49_068630	1	4	uncharacterized ACR, YagE family COG1723 domain-containing protein
TGME49_067400	14	1885	ribosomal protein L32, putative
TGME49_119570	19	401	WD-40 repeat protein, putative
TGGT1_121720	13	337	WD-repeat protein, putative
TGME49_088020	1	1	hypothetical protein
TGGT1_034390	1	2	conserved hypothetical protein
TGME49_088880	1	2	hypothetical protein
TGME49_079420	36	2469	hypothetical protein
TGME49_106650	8	342	hypothetical protein
TGME49_014220	12	558	hypothetical protein
TGME49_088180	7	34	hypothetical protein
TGGT1_035230	1	4	conserved hypothetical protein
TGME49_037180	12	252	hypothetical protein
TGGT1_068740	6	60	conserved hypothetical protein
TGME49_026830	89	18563	DnaK family protein
TGGT1_082100	38	3695	heat shock protein 70kD, putative
TGME49_055290	1	25	hypothetical protein
TGME49_091900	1	18	hypothetical protein
TGME49_022160	21	1324	aldehyde dehydrogenase, putative
TGGT1_065980	18	1066	aldehyde dehydrogenase, putative
TGGT1_112680	3	27	conserved hypothetical protein
TGME49_073600	1	7	hypothetical protein
TGME49_064940	3	16	hypothetical protein
TGME49_067500	17	1153	hypothetical protein, conserved
TGME49_002840	18	3007	zinc finger (C3HC4 type) / FHA domain-containing protein
TGGT1_018100	12	2613	zinc finger (C3HC4 type) protein
TGGT1_083060	2	41	DNA methyltransferase 2, putative
TGME49_027660	1	38	DNA methyltransferase 2, putative
TGGT1_075840	5	71	hypothetical protein
TGME49_067090	1	14	hypothetical protein
TGME49_049640	1	1	hypothetical protein, conserved
TGME49_079410	1	5	hypothetical protein
TGME49_109130	1	1	hypothetical protein, conserved
TGME49_067820	4	8	electron transfer flavoprotein beta-subunit, putative
TGME49_035720	2	2	hypothetical protein
TGME49_110570	10	483	hypothetical protein

TGME49_064650	17	308	hypothetical protein
TGGT1_038870	1	1	hypothetical protein
TGME49_013650	1	1	conserved hypothetical ATP binding domain-containing protein
TGGT1_106405	2	23	glycosyltransferase family 28 C-terminal domain-containing protein
TGME49_068340	2	23	glycosyltransferase family 28 C-terminal domain-containing protein
TGME49_067660	14	918	hypothetical protein
TGME49_078760	2	6	hypothetical protein
TGME49_003780	41	1631	hypothetical protein
TGGT1_016950	35	1786	conserved hypothetical protein
TGME49_019260	5	13	cation-transporting ATPase, putative
TGGT1_030220	2	8	cation-transporting ATPase worm, putative
TGME49_060220	6	200	folate/methotrexate transporter, putative
TGME49_063450	7	36	hypothetical protein
TGGT1_043570	6	40	conserved hypothetical protein
TGGT1_069350	1	2	conserved hypothetical protein
TGME49_067690	8	82	hypothetical protein
TGME49_067590	12	144	hypothetical protein
TGGT1_081450	3	21	hypothetical protein
TGGT1_122430	8	61	RNA recognition motif-containing protein, putative
TGME49_118690	8	61	RNA recognition motif-containing protein
TGME49_011460	11	69	hypothetical protein
TGGT1_123330	7	45	conserved hypothetical protein
TGME49_067420	7	141	mago nashi protein, putative
TGME49_064470	2	4	hypothetical protein
TGGT1_105730	1	2	conserved hypothetical protein
TGME49_040740	1	2	syntaxin binding protein, putative
TGME49_101580	2	6	hypothetical protein, conserved
TGGT1_017890	2	6	hypothetical protein
TGME49_090300	6	159	hypothetical protein
TGGT1_032750	5	73	conserved hypothetical protein
TGME49_078050	27	1117	proteasome subunit alpha type 1, putative
TGGT1_104500	13	920	proteasome subunit alpha type, putative
TGME49_023490	6	71	hypothetical protein
TGGT1_077760	6	63	conserved hypothetical protein
TGME49_097820	3	5	armadillo/beta-catenin-like repeat protein
TGGT1_019760	2	14	conserved hypothetical protein
TGME49_001670	1	5	DnaJ domain-containing protein
TGGT1_022490	1	2	conserved hypothetical protein
TGME49_009530	1	2	hypothetical protein
TGME49_064900	14	261	hypothetical protein
TGGT1_063270	3	20	conserved hypothetical protein
TGME49_005630	3	20	hypothetical protein
TGGT1_064380	1	1	hypothetical protein

TGME49_057640	1	4	hypothetical protein
TGME49_079360	2	3	hypothetical protein
TGGT1_007430	1	4	hypothetical protein
TGME49_062660	1	1	hypothetical protein
TGME49_051690	58	6201	seryl-tRNA synthetase, putative
TGME49_013730	10	74	lanthionine synthetase C-like protein 1, putative
TGGT1_038760	5	54	conserved hypothetical protein
TGME49_074180	3	29	hypothetical protein
TGME49_079380	26	1466	hypothetical protein, conserved
TGGT1_106870	1	2	hypothetical protein
TGME49_073120	15	609	SRS30C
TGME49_028290	29	919	phospholipase/carboxylesterase, putative
TGGT1_083860	10	467	phospholipase/carboxylesterase, putative
TGME49_068390	2	3	hypothetical protein
TGGT1_043940	1	1	ribosomal RNA methyltransferase, putative
TGME49_063670	1	1	ribosomal RNA methyltransferase, putative
TGME49_023080	15	219	hypothetical protein
TGGT1_066620	11	211	conserved hypothetical protein
TGME49_052440	12	73	GPI-anchor transamidase, putative
TGGT1_001340	4	16	GPI transamidase, putative
TGME49_009700	4	7	hypothetical protein
TGGT1_022760	1	2	conserved hypothetical protein
TGME49_110650	1	1	hypothetical protein, conserved
TGGT1_122650	6	43	conserved hypothetical protein
TGME49_118470	6	43	hypothetical protein
TGME49_114020	25	358	hypothetical protein
TGME49_093480	13	188	molybdopterin biosynthesis MoeA protein, putative
TGGT1_073860	10	183	molybdopterin biosynthesis MoeA protein, putative
TGME49_102050	13	164	elongation factor Tu, putative
TGME49_072710	32	645	hypothetical protein
TGME49_089510	10	74	hypothetical protein, conserved
TGME49_084630	1	2	hypothetical protein
TGME49_033120	49	2626	hypothetical protein
TGME49_093740	5	76	hypothetical protein
TGME49_024780	1	3	SRS domain-containing protein
TGGT1_001830	2	10	translation initiation factor 2B, delta subunit, putative
TGGT1_028550	1	1	hypothetical protein
TGME49_113790	5	73	hypothetical protein
TGME49_010090	17	3367	Rhoptry kinase family protein ROP43 (incomplete catalytic triad)
TGME49_020920	2	23	hypothetical protein
TGGT1_053780	11	705	hypothetical protein
TGME49_019860	2	49	hypothetical protein
TGME49_068260	9	95	hypothetical protein

TGME49_055060	1	139	cytochrome oxidase I
TGME49_026620	11	74	hypothetical protein
TGVEG_030040	29	2534	Rhoptry kinase family protein ROP2A (incomplete catalytic triad)
TGME49_102000	1	4	chaperone clpB protein, putative
TGME49_115570	6	81	hypothetical protein
TGME49_071370	6	133	hypothetical protein
TGGT1_126670	10	313	rhoptry antigen ROP8, putative
TGME49_051060	2	4	hypothetical protein
TGME49_097760	2	8	hypothetical protein, conserved
TGME49_015270	1	1	hypothetical protein, conserved
TGME49_047990	1	10	hypothetical protein
TGME49_054600	3	18	ubiquitin, putative
TGME49_014250	8	211	hypothetical protein
TGME49_015770	23	55016	Rhoptry kinase family protein ROP8 (incomplete catalytic triad)
TGME49_104870	2	9	hypothetical protein
TGGT1_060480	19	1606	PAN domain-containing protein, putative
TGGT1_104190	2	6	conserved hypothetical protein
TGME49_000230	4	61	PAN domain-containing protein
TGME49_064930	1	1	hypothetical protein
TGME49_042120	6	41	hypothetical protein
TGME49_075460	3	30	hypothetical protein
TGME49_014740	13	104	retinitis pigmentosa 1-like 1 protein, putative
TGME49_071030	5	21	hypothetical protein, conserved
TGME49_003120	1	1	hypothetical protein
TGME49_053310	1	1	hypothetical protein
TGME49_019720	4	30	Ras family domain-containing protein
TGME49_026800	6	156	glutathione synthetase, putative
TGME49_086650	7	87	hypothetical protein
TGME49_105220	1	4	C-Myc binding protein, putative
TGME49_000250	14	765	PAN domain-containing protein
TGME49_086030	4	84	glutamic acid-rich protein, putative
TGME49_067160	7	65	SRS38D
TGGT1_080720	1	1	hypothetical protein
TGME49_005320	34	9760	hypothetical protein
TGGT1_099960	1	1	conserved hypothetical protein
TGME49_097190	1	1	hypothetical protein, conserved
TGME49_079100	34	10040	hypothetical protein
TGME49_085710	11	127	hypothetical protein
TGME49_000270	3	34	PAN domain-containing protein
TGME49_075300	27	53919	Rhoptry kinase family protein ROP2B (incomplete catalytic triad)
TGME49_053060	5	42	hypothetical protein
TGME49_059970	2	6	hypothetical protein
TGME49_064430	12	758	hypothetical protein

TGVEG_050040	51	4625	Rhoptry kinase family protein ROP4 / ROP7 (incomplete catalytic triad) ***WARNING: GENE MODEL INACCURATE ***
TGGT1_102650	1	2	conserved hypothetical protein
TGME49_121550	8	352	hypothetical protein
TGME49_015780	36	137272	Rhoptry kinase family protein ROP2A (incomplete catalytic triad)
TGME49_107860	1	2	hypothetical protein
TGME49_034310	3	29	hypothetical protein
TGME49_072300	1	1	hypothetical protein
TGGT1_041240	3	17	conserved hypothetical protein
TGME49_042230	11	226	Rhoptry kinase family protein ROP29
TGME49_079320	57	9750	hypothetical protein
TGME49_063800	2	10	hypothetical protein
TGME49_072380	4	34	hypothetical protein
TGME49_000240	5	77	PAN domain-containing protein
TGGT1_016910	3	28	hypothetical protein
TGME49_053450	16	2248	hypothetical protein
TGME49_090180	9	74	hypothetical protein
TGME49_121700	1	7	hypothetical protein
TGME49_067140	14	1473	SRS38B
TGME49_036060	3	9	hypothetical protein
TGME49_022240	19	1883	hypothetical protein
TGME49_103450	2	3	hypothetical protein, conserved
TGME49_067620	24	1243	hypothetical protein, conserved
TGME49_014840	41	3952	hypothetical protein
TGGT1_075410	3	11	hypothetical protein
TGGT1_049970	1	1	conserved hypothetical protein
TGME49_100650	1	1	sufB/sufD domain-containing protein
TGME49_061810	1	1	hypothetical protein, conserved
TGME49_043610	16	456	DNA methyltransferase, putative
TGME49_087740	1	2	hypothetical protein
TGME49_044420	1	1	hypothetical protein, conserved
TGME49_069910	2	7	hypothetical protein
TGGT1_007050	20	2059	conserved hypothetical protein
TGGT1_064740	3	13	protein kinase domain-containing protein, putative
TGME49_080170	6	162	hypothetical protein
TGME49_052870	3	5	hypothetical protein
TGME49_082210	3	45	hypothetical protein
TGME49_085910	2	3	hypothetical protein
TGME49_092260	3	15	SRS36B (= SAG5D)
TGME49_104150	1	1	protein kinase, putative
TGME49_114800	2	5	hypothetical protein, conserved
TGME49_009840	5	95	hypothetical protein
TGME49_061250	34	3870	histone H2A, putative
TGME49_064920	5	107	hypothetical protein

TGME49_010110	13	294	Rhoptry kinase family protein ROP44 (incomplete catalytic triad)
TGGT1_066680	1	1	hypothetical protein
TGME49_071390	2	6	hypothetical protein, conserved
TGME49_067530	3	6	hypothetical protein
TGGT1_051950	1	1	hypothetical protein
TGME49_094720	13	532	hypothetical protein
TGME49_036190	2	4	hypothetical protein
TGGT1_088990	1	6	hypothetical protein
TGGT1_082220	1	1	hypothetical protein
TGME49_106890	17	535	platelet binding protein GspB, putative
TGME49_068730	1	5	hypothetical protein, conserved
TGME49_027880	7	50	hypothetical protein
TGME49_015300	2	11	hypothetical protein, conserved
TGME49_081950	4	14	hypothetical protein
TGME49_081910	2	12	hypothetical protein
TGGT1_017510	1	2	hypothetical protein
TGME49_064870	9	192	sodium-dependent transporter, putative
TGME49_037120	1	2	hypothetical protein
TGME49_037130	2	5	cytochrome b, putative
TGME49_108970	8	92	hypothetical protein
TGGT1_047990	1	1	protein kinase domain-containing protein, putative
TGGT1_124090	1	4	alr/erv, putative
TGGT1_110260	1	1	conserved hypothetical protein
TGME49_097210	2	3	hypothetical protein
TGME49_055670	2	4	methyltransferase domain-containing protein
TGME49_109260	7	74	product unspecified
NCLIV_027290	5	64	Ribosomal protein S21-maize (ISS), related
NCLIV_060400	3	16	hypothetical protein

Table S3: Identification of 673 gene trascript up-regulated in *Toxoplasma gondii* oocysts in comparison to tachzoite/bradyzoite

[Gene ID]	[Fold Change]	[Product Description]
TGME49_037080	979.3	hypothetical protein
TGME49_002100	625.3	hypothetical protein
TGME49_002110	463.4	hypothetical protein
TGME49_120540	443.1	hypothetical protein
TGME49_080430	434.4	hypothetical protein
TGME49_002090	359.3	hypothetical protein
TGME49_055380	322	hypothetical protein
TGME49_087250	277.9	hypothetical protein
TGME49_120530	244.6	hypothetical protein
TGME49_062470	238.9	beta antigen, putative
TGME49_078180	233.4	hypothetical protein
TGME49_071490	208.6	hypothetical protein
TGME49_069110	201.3	ornithine aminotransferase, putative
TGME49_029320	175.3	haloacid dehalogenase-like hydrolase domain-containing protein
TGME49_013850	168.6	hypothetical protein
TGME49_066860	161.2	BTB/POZ domain-containing protein
TGME49_004520	152.6	hypothetical protein
TGME49_031390	146.9	hypothetical protein
TGME49_089660	142.5	hypothetical protein
TGME49_080500	142.3	sulfate transporter, putative
TGME49_049520	141.9	hypothetical protein
TGME49_027100	141.6	glutaredoxin, putative
TGME49_094600	139.3	hypothetical protein
TGME49_031110	136.7	hypothetical protein
TGME49_075660	136.2	hypothetical protein
TGME49_006660	135.1	hypothetical protein
TGME49_115260	132.5	alanine dehydrogenase, putative
TGME49_059900	131.5	hypothetical protein, conserved
TGME49_054760	129.4	hypothetical protein
TGME49_009470	118.6	hypothetical protein
TGME49_070950	118.5	hypothetical protein
TGME49_005090	117.5	hypothetical protein
TGME49_120580	113.5	hypothetical protein
TGME49_076880	111	late embryogenesis abundant domain-containing protein
TGME49_065320	110.2	hypothetical protein
TGME49_064630	109.6	hypothetical protein
TGME49_069120	107.7	oxidoreductase, putative
TGME49_071850	106.6	hypothetical protein
TGME49_009610	101.5	oocyst wall protein COWP, putative
TGME49_076850	93.5	late embryogenesis abundant domain-containing protein
TGME49_115480	93	acyl-CoA oxidase, putative
TGME49_042600	92	6-phosphogluconate dehydrogenase, putative

TGME49_006400	91.8	DNA photolyase, putative
TGME49_090610	89.1	hypothetical protein
TGME49_089860	88.6	hypothetical protein
TGME49_076870	88.4	LEA1 protein, putative
TGME49_058910	87.9	hypothetical protein
TGME49_026300	87.8	oxidoreductase, putative
TGME49_011320	86.3	C protein immunoglobulin-A-binding beta antigen
TGME49_063010	85.4	BTB/POZ domain-containing protein
TGME49_055360	83.8	hypothetical protein
TGME49_058550	80	SRS28 (= SporoSAG)
TGME49_073040	78.8	hypothetical protein, conserved
TGME49_039650	77.8	hypothetical protein
TGME49_078150	75.6	hypothetical protein
TGME49_089870	75.3	hypothetical protein
TGME49_066790	75.1	hypothetical protein
TGME49_116190	72.7	superoxide dismutase, putative
TGME49_037440	71.8	hypothetical protein, conserved
TGME49_005760	70.2	hypothetical protein, conserved
TGME49_076860	69	late embryogenesis abundant domain-containing protein
TGME49_089590	68.6	hypothetical protein
TGME49_023400	65.3	hypothetical protein
TGME49_039640	65.3	hypothetical protein
TGME49_097410	65.3	hypothetical protein
TGME49_006440	64.8	plasmodium falciparum CPW-WPC domain-containing protein
TGME49_023770	64.8	hypothetical protein
TGME49_088310	64.8	hypothetical protein
TGME49_069280	64.5	hypothetical protein
TGME49_003890	64.4	hypothetical protein
TGME49_104450	64.4	hypothetical protein
TGME49_005520	62.5	hypothetical protein
TGME49_056040	60.9	LCCL domain-containing protein
TGME49_030530	59.5	hypothetical protein
TGME49_105000	58.8	hypothetical protein
TGME49_073700	57.8	hypothetical protein
TGME49_078560	57.2	hypothetical protein
TGME49_061560	56.7	TBPIP domain-containing protein
TGME49_067410	55.8	scavenger receptor protein TgSR1, putative
TGME49_069170	54.9	hypothetical protein
TGME49_116290	54.4	hypothetical protein
TGME49_016150	54.2	oligoendopeptidase F, putative
TGME49_005070	53.8	hypothetical protein
TGME49_116550	53.4	hypothetical protein
TGME49_019160	52.3	hypothetical protein
TGME49_048440	51.9	hypothetical protein
TGME49_047500	51.5	acyl-coenzyme A oxidase, putative
TGME49_007780	51.4	hypothetical protein

TGME49_003070	50.9	hypothetical protein
TGME49_032200	49.4	hypothetical protein
TGME49_110690	47.8	hypothetical protein
TGME49_023700	46.9	LCCL domain-containing protein / F5/8 type C domain-containing protein
TGME49_029330	46.1	haloacid dehalogenase-like hydrolase domain containing protein
TGME49_018510	45.9	hypothetical protein, conserved
TGME49_053480	45.7	meiotic recombination protein SPO11, putative
TGME49_121580	45	hypothetical protein
TGME49_092350	44.8	hypothetical protein
TGME49_010770	44.7	hypothetical protein
TGME49_016400	44.4	meiotic recombination protein DMC1-like protein, putative
TGME49_059590	43.5	hypothetical protein
TGME49_010260	43.4	glycerol-3-phosphate dehydrogenase, putative
TGME49_048380	43	hypothetical protein
TGME49_014570	42.6	hypothetical protein
TGME49_065340	42.5	hypothetical protein
TGME49_061530	42.2	hypothetical protein
TGME49_025460	41.5	hypothetical protein
TGME49_093250	41.5	hypothetical protein
TGME49_072680	41.1	hypothetical protein
TGME49_091840	40.9	hypothetical protein
TGME49_007250	40.6	hypothetical protein
TGME49_023120	39.7	hypothetical protein
TGME49_032480	39.5	paramyosin, putative
TGME49_085300	39.5	hypothetical protein, conserved
TGME49_111710	38.7	hypothetical protein
TGME49_107640	38.5	CMGC kinase, CK2 family
TGME49_079440	38.3	hypothetical protein, conserved
TGME49_009920	38.2	PAN domain-containing protein
TGME49_019610	37.5	hypothetical protein
TGME49_016090	37.2	tropomyosin 1 alpha chain, putative
TGME49_066870	37.1	major facilitator family protein, putative
TGME49_111420	36.9	hypothetical protein
TGME49_066130	36.2	glutathione/thioredoxin peroxidase, putative
TGME49_081590	36.2	hypothetical protein
TGME49_114050	35.8	splicing factor U2AF 65 kDa subunit, putative
TGME49_057460	35.7	hypothetical protein
TGME49_077570	35.7	GPI transamidase subunit PIG-U, putative
TGME49_033280	34.9	hypothetical protein
TGME49_097350	34.2	hypothetical protein
TGME49_115540	33.6	microneme protein, putative
TGME49_007730	33.2	hypothetical protein
TGME49_028620	32.9	hypothetical protein
TGME49_040530	32.8	hypothetical protein
TGME49_017490	32.7	hypothetical protein

TGME49_019110	32.5	hypothetical protein, conserved
TGME49_085690	32.5	notch (DSL) domain-containing protein
TGME49_039670	32.4	hypothetical protein
TGME49_067470	31.8	hypothetical protein
TGME49_110830	31.7	delta3,5-delta2,4-dienoyl-CoA isomerase, putative
TGME49_025950	31.5	hypothetical protein
TGME49_085920	31.2	glutathione S-transferase, putative
TGME49_062940	30.8	eukaryotic aspartyl protease, putative
TGME49_099800	30.8	hypothetical protein
TGME49_048760	30.7	subtilase family serine protease, putative
TGME49_049750	30.6	hypothetical protein, conserved
TGME49_058400	30.6	LCCL domain-containing protein
TGME49_071220	30.6	hypothetical protein
TGME49_113940	30	hypothetical protein
TGME49_053670	29.7	hypothetical protein
TGME49_072240	29.2	hypothetical protein
TGME49_066770	28.9	plasmodium falciparum CPW-WPC domain-containing protein
TGME49_071210	28.8	4-alpha-glucanotransferase, putative
TGME49_106050	28.7	hypothetical protein
TGME49_065290	28.4	hypothetical protein
TGME49_028110	28	hypothetical protein
TGME49_088220	28	hypothetical protein
TGME49_062090	27.8	hypothetical protein
TGME49_078620	27.3	hypothetical protein
TGME49_031900	27.2	acyl-CoA dehydrogenase, putative
TGME49_054780	27	hypothetical protein
TGME49_064810	26.4	hypothetical protein
TGME49_032950	26.1	hypothetical protein
TGME49_048350	26.1	hypothetical protein
TGME49_008770	25.9	hypothetical protein
TGME49_038160	25.5	hypothetical protein
TGME49_032170	25.4	hypothetical protein
TGME49_078690	25.2	hypothetical protein
TGME49_088520	25.1	thyroid hormone receptor interactor, putative
TGME49_059620	24.5	hypothetical protein
TGME49_005400	24.2	DNA repair protein RAD54, putative
TGME49_056790	24.1	hypothetical protein
TGME49_040680	24	hypothetical protein
TGME49_115350	23	hypothetical protein, conserved
TGME49_088590	22.9	hypothetical protein
TGME49_059670	22.6	von Willebrand factor type A domain-containing protein
TGME49_031420	22.5	BT1 transmembrane domain-containing protein
TGME49_005660	22.3	F5/8 type C domain-containing protein
TGME49_066980	21.9	hypothetical protein
TGME49_119890	21.7	hypothetical protein
TGME49_073740	21.5	acetyl-CoA acyltransferase B, putative

TGME49_094300	21.5	hypothetical protein
TGME49_016640	21.2	homoserine kinase, putative
TGME49_034570	21.1	peroxisomal multifunctional enzyme type 2, putative
TGME49_008330	20.6	hypothetical protein
TGME49_061630	20.6	hypothetical protein
TGME49_012910	20.1	rhomboid family domain-containing protein
TGME49_005450	20	hypothetical protein
TGME49_028430	19.3	hypothetical protein
TGME49_086770	19.3	hypothetical protein, conserved
TGME49_092150	18.9	hypothetical protein
TGME49_068700	18.8	hypothetical protein
TGME49_039430	18.7	hypothetical protein, conserved
TGME49_099150	18.3	hypothetical protein, conserved
TGME49_033780	18.2	hypothetical protein
TGME49_063410	18.2	scavenger receptor protein SR2
TGME49_035000	18.1	hypothetical protein
TGME49_049930	18	hypothetical protein, conserved
TGME49_068930	18	hypothetical protein
TGME49_115550	17.8	microneme protein, putative
TGME49_113000	17.7	hypothetical protein
TGME49_072510	17.6	eukaryotic aspartyl protease, putative
TGME49_015290	17.5	saccharopine dehydrogenase, putative
TGME49_104730	17.5	mob1/phocein family domain-containing protein
TGME49_003230	17.4	hypothetical protein
TGME49_004460	17.4	hydroxymethylglutaryl-CoA lyase, putative
TGME49_078570	17.1	hypothetical protein
TGME49_086250	17	oocyst wall protein, putative
TGME49_026490	16.9	hypothetical protein
TGME49_090930	16.9	hypothetical protein
TGME49_114330	16.7	ABC transporter, putative
TGME49_091660	16.6	hypothetical protein
TGME49_066910	16.1	CMGC kinase
TGME49_048330	15.9	zinc finger (C3HC4 RING finger) protein, putative
TGME49_012970	15.8	protein kinase (incomplete catalytic triad)
TGME49_058810	15.8	SRS27B
TGME49_027070	15.6	hypothetical protein
TGME49_070280	15.6	hypothetical protein
TGME49_066740	15.3	DNA/RNA binding protein, putative
TGME49_086040	15.3	hypothetical protein
TGME49_073510	15.2	hypothetical protein
TGME49_109870	15.2	PX domain-containing protein
TGME49_043940	15.1	hypothetical protein
TGME49_085470	15	Patched family domain containing protein
TGME49_118540	15	hypothetical protein
TGME49_097160	14.9	hypothetical protein

TGME49_118650	14.7	transhydrogenase, putative
TGME49_005020	14.6	hypothetical protein
TGME49_101370	14.6	zinc finger DHHC domain-containing protein, conserved
TGME49_115560	14.6	ABC transporter, putative
TGME49_032420	14.4	hypothetical protein
TGME49_033360	14.4	hypothetical protein
TGME49_033370	14.3	hypothetical protein, conserved
TGME49_099790	14.3	hypothetical protein
TGME49_075760	14.2	hypothetical protein
TGME49_060830	14	hypothetical protein
TGME49_025140	13.9	folate/methotrexate transporter FT1, putative
TGME49_030540	13.7	hypothetical protein
TGME49_040480	13.6	hypothetical protein, conserved
TGME49_064070	13.6	antigenic protein, putative
TGME49_080680	13.5	hypothetical protein
TGME49_028440	13.4	hypothetical protein
TGME49_039360	13.4	hypothetical protein
TGME49_004010	13	hypothetical protein
TGME49_057010	12.9	insulysin, putative
TGME49_017040	12.8	hypothetical protein
TGME49_032470	12.8	hypothetical protein
TGME49_042590	12.8	hypothetical protein
TGME49_009590	12.6	hypothetical protein
TGME49_045600	12.6	hypothetical protein
TGME49_078120	12.6	SCP-like extracellular domain-containing protein
TGME49_055660	12.5	hypothetical protein
TGME49_012080	12.3	meiosis-specific nuclear structural protein 1, putative
TGME49_048730	12.3	oocyst wall protein, putative
TGME49_065060	12.3	hypothetical protein
TGME49_082230	12.3	sulfate adenylyltransferas-adenylylsulfate kinase, putative
TGME49_026530	12.2	hypothetical protein
TGME49_051550	12.2	acyl-CoA-binding protein, putative
TGME49_089610	12.1	hypothetical protein
TGME49_032580	11.8	acetyl-coenzyme A synthetase, putative
TGME49_040610	11.8	hypothetical protein, conserved
TGME49_020840	11.6	threonine synthase, putative
TGME49_099260	11.6	hypothetical protein
TGME49_075700	11.5	hypothetical protein
TGME49_068620	11.4	41-3 antigen, putative
TGME49_107790	11.2	hypothetical protein
TGME49_003280	10.9	hypothetical protein
TGME49_026540	10.9	protein kinase
TGME49_016100	10.8	hypothetical protein
TGME49_015360	10.7	hypothetical protein
TGME49_049980	10.5	leucine rich repeat protein, putative

TGME49_118450	10.5	zinc finger (C3HC4 RING finger) protein, putative
TGME49_002890	10.4	hypothetical protein
TGME49_042390	10.4	enoyl-CoA hydratase/isomerase family protein, putative
TGME49_116890	10.4	hypothetical protein
TGME49_024860	10.3	exosome complex exonuclease, putative
TGME49_000350	10.2	subtilase family serine protease, putative
TGME49_039570	10.2	hypothetical protein, conserved
TGME49_115090	10	hypothetical protein
TGME49_085790	9.9	hypothetical protein
TGME49_043600	9.6	acetyltransferase domain-containing protein
TGME49_049910	9.6	mitochondria-associated granulocyte macrophage CSF signaling molecule,
TGME49_116560	9.5	hypothetical protein
TGME49_105160	9.4	histone H2B, putative
TGME49_110620	9.4	hypothetical protein
TGME49_012740	9.1	phenylalanine-4-hydroxylase, putative
TGME49_069600	9	biotin carboxyl carrier protein, putative
TGME49_088490	9	hypothetical protein, conserved
TGME49_058160	8.9	hypothetical protein
TGME49_082140	8.9	hypothetical protein
TGME49_105280	8.9	hypothetical protein
TGME49_008010	8.8	hypothetical protein
TGME49_032290	8.8	hypothetical protein
TGME49_044060	8.8	ankyrin repeat-containing protein
TGME49_013730	8.7	lanthionine synthetase C-like protein 1, putative
TGME49_039370	8.7	hypothetical protein, conserved
TGME49_042890	8.7	hypothetical protein
TGME49_078140	8.5	hypothetical protein
TGME49_097290	8.5	hypothetical protein, conserved
TGME49_111650	8.4	leucine rich repeat protein, putative
TGME49_034400	8.3	hypothetical protein
TGME49_052330	8.3	hypothetical protein, conserved
TGME49_110460	8.3	RAB6 protein
TGME49_031770	8.2	tubulin alpha chain, putative
TGME49_003090	8.1	hypothetical protein
TGME49_086670	8.1	hypothetical protein
TGME49_110160	8.1	piwi-PAZ domain-containing protein
TGME49_016940	8	hypothetical protein
TGME49_094810	8	RRM domain-containing protein
TGME49_051820	7.9	hypothetical protein
TGME49_026230	7.8	hypothetical protein
TGME49_057480	7.8	succinate-semialdehyde dehydrogenase, putative
TGME49_066380	7.8	hypothetical protein
TGME49_016160	7.7	hypothetical protein
TGME49_027570	7.7	transmembrane amino acid transporter domain-containing protein
TGME49_114910	7.7	hypothetical protein, conserved

TGME49_094850	7.6	actin-like protein 7, putative
TGME49_052510	7.5	hypothetical protein, conserved
TGME49_099780	7.5	hypothetical protein
TGME49_120280	7.5	hypothetical protein
TGME49_021310	7.4	aminopeptidase N, putative
TGME49_007650	7.3	OTU-like cysteine protease domain-containing protein
TGME49_030180	7.3	hypothetical protein
TGME49_065120	7.3	hypothetical protein, conserved
TGME49_062890	7.2	hypothetical protein
TGME49_093560	7.2	hypothetical protein
TGME49_048430	7.1	hypothetical protein
TGME49_118550	7.1	hypothetical protein
TGME49_069590	6.9	14-3-3 protein, putative
TGME49_089930	6.9	phosphoenolpyruvate carboxykinase, putative
TGME49_105490	6.9	programmed cell death protein, putative
TGME49_046490	6.8	hypothetical protein
TGME49_095460	6.8	got1-like family domain-containing protein
TGME49_095640	6.8	endothelin-converting enzyme, putative
TGME49_101120	6.8	acetyil CoA acetyltransferase/thiolase, putative
TGME49_011360	6.7	hypothetical protein
TGME49_036510	6.7	hypothetical protein
TGME49_047620	6.7	hypothetical protein, conserved
TGME49_121430	6.7	DNA repair protein recA , putative
TGME49_029140	6.6	peroxisomal multifunctional enzyme, putative
TGME49_039510	6.6	hypothetical protein
TGME49_065280	6.6	hypothetical protein
TGME49_120270	6.6	hypothetical protein
TGME49_009620	6.5	eukaryotic aspartyl protease, putative
TGME49_105530	6.5	hypothetical protein
TGME49_105540	6.5	hypothetical protein
TGME49_031780	6.4	hypothetical protein
TGME49_051830	6.4	hypothetical protein, conserved
TGME49_053820	6.4	hypothetical protein, conserved
TGME49_011280	6.3	hypothetical protein
TGME49_028780	6.3	hypothetical protein
TGME49_053030	6.3	alpha-glucosidase II, putative
TGME49_053320	6.3	hypothetical protein, conserved
TGME49_115230	6.3	hypothetical protein
TGME49_018450	6.2	hypothetical protein
TGME49_057310	6.2	nuclear movement domain-containing protein
TGME49_071080	6.2	hypothetical protein
TGME49_071190	6.2	Bin3 domain-containing protein
TGME49_097280	6.2	hypothetical protein, conserved
TGME49_112870	6.2	hypothetical protein
TGME49_115840	6.2	hypothetical protein

TGME49_001240	6.1	hypothetical protein
TGME49_040960	6.1	hypothetical protein
TGME49_073710	6.1	hypothetical protein, conserved
TGME49_078680	6.1	hypothetical protein
TGME49_117710	6.1	enoyl-CoA hydratase, putative
TGME49_012070	6	meiosis-specific nuclear structural protein 1, putative
TGME49_106450	6	short chain dehydrogenase family protein, putative
TGME49_033950	5.9	hypothetical protein
TGME49_076990	5.9	cytochrome b5-like Heme/Steroid binding domain containing protein
TGME49_003110	5.8	citrate synthase, putative
TGME49_015140	5.8	hypothetical protein
TGME49_027370	5.8	X-Pro dipeptidyl-peptidase domain-containing protein
TGME49_043720	5.8	peroxisomal biogenesis factor 11 domain-containing protein
TGME49_053710	5.8	hypothetical protein
TGME49_081620	5.8	aldo/keto reductase family oxidoreductase, putative
TGME49_097910	5.8	hypothetical protein
TGME49_119350	5.8	SRS domain containing protein
TGME49_030650	5.7	hypothetical protein
TGME49_046750	5.7	hypothetical protein
TGME49_048410	5.7	hypothetical protein
TGME49_092100	5.7	2-oxoisovalerate dehydrogenase, putative
TGME49_003500	5.6	alanine dehydrogenase, putative
TGME49_028030	5.6	hypothetical protein
TGME49_054680	5.6	hypothetical protein
TGME49_086650	5.6	hypothetical protein
TGME49_089160	5.6	hypothetical protein
TGME49_092960	5.6	transmembrane domain-containing protein
TGME49_097820	5.6	armadillo/beta-catenin-like repeat protein
TGME49_007740	5.5	hypothetical protein
TGME49_054010	5.5	serine carboxypeptidase S28, putative
TGME49_012820	5.4	ubiquitin family domain-containing protein
TGME49_054950	5.4	hypothetical protein
TGME49_069610	5.4	hypothetical protein
TGME49_025920	5.3	hypothetical protein
TGME49_031040	5.3	exosome complex exonuclease, putative
TGME49_068670	5.3	hypothetical protein
TGME49_074090	5.3	tim10/DDP zinc finger domain-containing protein
TGME49_120020	5.3	major facilitator superfamily domain-containing protein
TGME49_052350	5.2	hypothetical protein
TGME49_054800	5.2	hypothetical protein
TGME49_057690	5.2	hypothetical protein
TGME49_069380	5.2	hypothetical protein
TGME49_072540	5.2	protein kinase (incomplete catalytic triad)
TGME49_088550	5.2	hypothetical protein
TGME49_089760	5.2	biotin-[acetyl-CoA-carboxylase] ligase, putative
TGME49_091170	5.2	hypothetical protein

TGME49_110340	5.2	hypothetical protein
TGME49_110510	5.2	hypothetical protein
TGME49_003720	5.1	vitamin K epoxide reductase complex subunit 1, putative
TGME49_015210	5.1	F-box domain-containing protein
TGME49_016610	5.1	hypothetical protein
TGME49_020540	5.1	hypothetical protein
TGME49_026600	5.1	syntaxin, putative
TGME49_037400	5.1	protein phosphatase 2C, putative
TGME49_058200	5.1	glutamic acid-rich protein, putative
TGME49_064110	5.1	hypothetical protein
TGME49_077090	5.1	mitochondrial carrier domain-containing protein
TGME49_078580	5.1	hypothetical protein
TGME49_097370	5.1	ruvB-like 1, putative
TGME49_012860	5	hypothetical protein
TGME49_030110	5	hypothetical protein
TGME49_030640	5	hypothetical protein
TGME49_046720	5	hypothetical protein
TGME49_074140	5	RNA recognition motif-containing protein
TGME49_104630	5	hypothetical protein
TGME49_001850	4.9	peroxisome biogenesis factor 7, putative
TGME49_001870	4.9	TPR domain-containing protein
TGME49_010400	4.9	hypothetical protein
TGME49_064640	4.9	bromodomain domain-containing protein
TGME49_087260	4.9	hypothetical protein
TGME49_090670	4.9	cytosol aminopeptidase
TGME49_105450	4.9	acetyltransferase domain-containing protein
TGME49_027580	4.8	hypothetical protein
TGME49_031350	4.8	glucosamine--fructose-6-phosphate aminotransferase (isomerizing), putative
TGME49_032400	4.8	PAN domain-containing protein
TGME49_039260	4.8	histone H4, putative
TGME49_047610	4.8	snRNP protein Lsm5, putative
TGME49_068020	4.8	transporter, major facilitator family domain containing protein
TGME49_078910	4.8	O-acetylserine (thiol) lyase, putative
TGME49_106590	4.8	hypothetical protein
TGME49_005130	4.7	glyoxalase, putative
TGME49_063540	4.7	hypothetical protein
TGME49_078040	4.7	G-patch domain-containing protein
TGME49_086460	4.7	hypothetical protein
TGME49_009070	4.6	hypothetical protein
TGME49_026010	4.6	pterin-4a-carbinolamine dehydratase
TGME49_040690	4.6	S-adenosylmethionine synthetase, putative
TGME49_043580	4.6	HIT domain-containing protein
TGME49_078090	4.6	hypothetical protein
TGME49_100380	4.6	endoplasmic reticulum oxidoreductin, putative
TGME49_109550	4.6	hypothetical protein

TGME49_114780	4.6	myosin G
TGME49_004130	4.5	membrane-attack complex / perforin domain-containing protein
TGME49_005050	4.5	hypothetical protein
TGME49_009930	4.5	hypothetical protein
TGME49_023430	4.5	hypothetical protein
TGME49_023520	4.5	hypothetical protein
TGME49_024610	4.5	IQ calmodulin-binding motif domain-containing protein
TGME49_028410	4.5	tubulin-tyrosine ligase family protein
TGME49_033390	4.5	hypothetical protein
TGME49_046660	4.5	hypothetical protein
TGME49_047310	4.5	plasmodium falciparum CPW-WPC domain-containing protein
TGME49_051710	4.5	hypothetical protein
TGME49_053300	4.5	hypothetical protein, conserved
TGME49_054770	4.5	serine/threonine protein phosphatase, putative
TGME49_111170	4.5	hypothetical protein
TGME49_112540	4.5	hypothetical protein
TGME49_120740	4.5	hypothetical protein
TGME49_002930	4.4	hypothetical protein
TGME49_004490	4.4	hypothetical protein
TGME49_055710	4.4	ULK kinase
TGME49_056890	4.4	AGC kinase
TGME49_072200	4.4	AGC kinase
TGME49_082030	4.4	hypothetical protein
TGME49_090890	4.4	carbonyl reductase, putative
TGME49_091030	4.4	hypothetical protein
TGME49_091600	4.4	interferon gamma-inducible protein 30
TGME49_114670	4.4	hypothetical protein
TGME49_002400	4.3	fibrillin-2 precursor, putative
TGME49_006450	4.3	hypothetical protein
TGME49_010950	4.3	oocyst wall protein, putative
TGME49_013340	4.3	glucose-methanol-choline oxidoreductase domain-containing protein
TGME49_020420	4.3	3'5'-cyclic nucleotide phosphodiesterase, putative
TGME49_035350	4.3	hypothetical protein
TGME49_039140	4.3	hypothetical protein
TGME49_050110	4.3	hypothetical protein, conserved
TGME49_057130	4.3	zinc finger (CCCH type) protein, putative
TGME49_115730	4.3	apical membrane antigen, putative
TGME49_119880	4.3	MORN repeat-containing protein
TGME49_011090	4.2	cysteine desulfurase, putative
TGME49_019670	4.2	zinc finger (CCCH type) protein, putative
TGME49_029280	4.2	hypothetical protein
TGME49_031090	4.2	hypothetical protein
TGME49_044540	4.2	mitochondrial carrier domain-containing protein
TGME49_055460	4.2	EGF-like domain-containing protein
TGME49_064030	4.2	aminotransferase, putative
TGME49_089290	4.2	hypothetical protein

TGME49_112930	4.2	cystathionine beta-lyase, putative
TGME49_112950	4.2	hypothetical protein
TGME49_006570	4.1	kinesin, putative
TGME49_032350	4.1	lactate dehydrogenase
TGME49_037860	4.1	CAM Kinase, RAD family
TGME49_049760	4.1	vitamin K epoxide reductase complex subunit 1, putative
TGME49_050750	4.1	hypothetical protein
TGME49_058370	4.1	Rhoptry kinase family protein ROP28
TGME49_059910	4.1	hypothetical protein
TGME49_073630	4.1	DnaJ domain-containing protein
TGME49_114000	4.1	selR domain-containing protein
TGME49_115770	4.1	cytochrome p450, putative
TGME49_002340	4	proline oxidase, putative
TGME49_052490	4	vacuolar protein sorting 29, putative
TGME49_053580	4	CMGC kinase, CDK family
TGME49_054180	4	radial spoke 3 protein
TGME49_064140	4	hypothetical protein
TGME49_070820	4	hypothetical protein, conserved
TGME49_089950	4	hypothetical protein
TGME49_002840	3.9	zinc finger (C3HC4 type) / FHA domain-containing protein
TGME49_014310	3.9	hypothetical protein
TGME49_023480	3.9	thrombospondin type 1 domain-containing protein
TGME49_040860	3.9	1-acyl-sn-glycerol-3-phosphate acyltransferase, putative
TGME49_044260	3.9	hypothetical protein
TGME49_072370	3.9	hypothetical protein
TGME49_075650	3.9	p97 protein
TGME49_108020	3.9	SRS57 (= SAG3, P43)
TGME49_112470	3.9	hypothetical protein
TGME49_012030	3.8	hypothetical protein
TGME49_015170	3.8	serine/threonine protein phosphatase, putative
TGME49_022430	3.8	ubiquitin-transferase domain containing protein
TGME49_027050	3.8	pyruvate dehydrogenase kinase, putative
TGME49_049030	3.8	endonuclease/exonuclease/phosphatase domain-containing protein
TGME49_053570	3.8	hypothetical protein
TGME49_063210	3.8	hypothetical protein
TGME49_064050	3.8	ubiquitin-conjugating enzyme domain-containing protein
TGME49_071330	3.8	hypothetical protein
TGME49_086440	3.8	hypothetical protein
TGME49_110270	3.8	hypothetical protein, conserved
TGME49_118570	3.8	SFT2-like domain-containing protein
TGME49_039120	3.7	hypothetical protein
TGME49_048820	3.7	hypothetical protein
TGME49_050030	3.7	hypothetical protein, conserved
TGME49_066700	3.7	hypothetical protein
TGME49_073860	3.7	hypothetical protein

TGME49_093190	3.7	endonuclease/exonuclease/phosphatase domain-containing protein
TGME49_106270	3.7	hypothetical protein
TGME49_006370	3.6	hypothetical protein
TGME49_032680	3.6	hypothetical protein
TGME49_038090	3.6	hypothetical protein
TGME49_043930	3.6	hypothetical protein
TGME49_044360	3.6	hypothetical protein, conserved
TGME49_046080	3.6	hypothetical protein
TGME49_060540	3.6	hypothetical protein
TGME49_066730	3.6	leucyl-tRNA synthetase, putative
TGME49_082020	3.6	hypothetical protein
TGME49_090960	3.6	phosphoglucan, water dikinase protein, putative
TGME49_095900	3.6	cyclic nucleotide-binding domain-containing protein
TGME49_116730	3.6	hypothetical protein
TGME49_006350	3.5	ER lumen protein retaining receptor 1, putative
TGME49_019230	3.5	acetoacetyl-CoA synthetase, putative
TGME49_031000	3.5	phosphatidylcholine transfer protein, putative
TGME49_031870	3.5	TPR domain-containing protein
TGME49_032250	3.5	peroxisomal catalase
TGME49_036120	3.5	hypothetical protein, conserved
TGME49_036620	3.5	protein kinase (incomplete catalytic triad)
TGME49_046640	3.5	hypothetical protein
TGME49_087490	3.5	hypothetical protein, conserved
TGME49_094640	3.5	ribonucleoside-diphosphate reductase, large subunit, putative
TGME49_096000	3.5	hypothetical protein
TGME49_097380	3.5	hypothetical protein
TGME49_104490	3.5	hypothetical protein
TGME49_108850	3.5	hypothetical protein
TGME49_110420	3.5	hypothetical protein
TGME49_110480	3.5	hypothetical protein
TGME49_113650	3.5	hypothetical protein
TGME49_115360	3.5	hypothetical protein
TGME49_001730	3.4	hypothetical protein, conserved
TGME49_001830	3.4	hypothetical protein, conserved
TGME49_004420	3.4	oocyst wall protein COWP, putative
TGME49_024030	3.4	hypothetical protein
TGME49_029200	3.4	hypothetical protein
TGME49_033010	3.4	CMGC kinase, MAPK family TgMAPK2
TGME49_042790	3.4	trichohyalin, putative
TGME49_043500	3.4	CAM kinase, SNF1/AMK1 family ToxPK1
TGME49_049380	3.4	zinc finger DHHC domain-containing protein
TGME49_049970	3.4	hypothetical protein
TGME49_064800	3.4	hypothetical protein
TGME49_068310	3.4	oocyst wall protein, putative
TGME49_069750	3.4	hypothetical protein

TGME49_070550	3.4	gamma-glutamyl phosphate reductase, putative
TGME49_073750	3.4	hypothetical protein, conserved
TGME49_110760	3.4	protein phosphatase 2C, putative
TGME49_110810	3.4	apyrase, putative
TGME49_001100	3.3	hypothetical protein
TGME49_004550	3.3	hypothetical protein
TGME49_016790	3.3	ABC transporter, putative
TGME49_026310	3.3	zinc finger (CCCH type) protein, putative
TGME49_032280	3.3	hypothetical protein
TGME49_033290	3.3	hypothetical protein
TGME49_036880	3.3	hypothetical protein, conserved
TGME49_048400	3.3	lactoylglutathione lyase, putative
TGME49_056330	3.3	hypothetical protein
TGME49_065230	3.3	electron transfer flavoprotein-ubiquinone oxidoreductase, putative
TGME49_070870	3.3	hypothetical protein
TGME49_095710	3.3	ubiquitin-transferase domain-containing protein
TGME49_098840	3.3	hypothetical protein
TGME49_106640	3.3	vesicle transport v-SNARE domain-containing protein
TGME49_110390	3.3	hypothetical protein
TGME49_113360	3.3	hypothetical protein
TGME49_115130	3.3	protein-L-isoaspartate O-methyltransferase, putative
TGME49_002310	3.2	glycoprotease family domain-containing protein
TGME49_002790	3.2	dihydrouridine synthase domain-containing protein
TGME49_002990	3.2	hypothetical protein
TGME49_003320	3.2	hypothetical protein
TGME49_005000	3.2	phosphoglycerate mutase domain-containing protein
TGME49_008380	3.2	hypothetical protein
TGME49_012240	3.2	tubulin beta chain, putative
TGME49_020380	3.2	hypothetical protein
TGME49_026910	3.2	glycogen debranching enzyme, putative
TGME49_034470	3.2	hypothetical protein
TGME49_038490	3.2	SRS22E
TGME49_040250	3.2	appr-1-p processing enzyme family domain-containing protein
TGME49_051460	3.2	hypothetical protein
TGME49_053150	3.2	oxidoreductase, putative
TGME49_057520	3.2	synaptobrevin-like protein, putative
TGME49_061660	3.2	hypothetical protein
TGME49_064150	3.2	hypothetical protein
TGME49_066690	3.2	hypothetical protein
TGME49_068630	3.2	uncharacterized ACR, YagE family COG1723 domain-containing protein
TGME49_078830	3.2	glucose-6-phosphate dehydrogenase, putative
TGME49_082000	3.2	hypothetical protein
TGME49_086750	3.2	MA3 domain protein
TGME49_089900	3.2	N-acetyltransferase 5, putative
TGME49_090850	3.2	WD-40 repeat protein, putative
TGME49_104480	3.2	3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 2, putative

TGME49_118210	3.2	hypothetical protein
TGME49_120710	3.2	hypothetical protein
TGME49_007460	3.1	GTP-binding protein, putative
TGME49_009480	3.1	hypothetical protein
TGME49_019570	3.1	hypothetical protein
TGME49_026380	3.1	hypothetical protein
TGME49_028360	3.1	FK506-binding protein 1, putative
TGME49_043350	3.1	gamma-glutamyl hydrolase, putative
TGME49_044010	3.1	hypothetical protein
TGME49_046570	3.1	hypothetical protein
TGME49_047940	3.1	hypothetical protein
TGME49_065050	3.1	hypothetical protein
TGME49_068730	3.1	hypothetical protein, conserved
TGME49_069680	3.1	acyl-CoA carboxyltransferase beta chain, putative
TGME49_070800	3.1	GAF domain-containing protein
TGME49_072420	3.1	1-O-acylceramide synthase, putative
TGME49_081650	3.1	hypothetical protein, conserved
TGME49_088230	3.1	hypothetical protein
TGME49_094330	3.1	EGF-like domain-containing protein
TGME49_094590	3.1	hypothetical protein
TGME49_100080	3.1	hypothetical protein, conserved
TGME49_100360	3.1	mitochondrial carrier domain-containing protein
TGME49_105140	3.1	patatin-like phospholipase domain-containing protein
TGME49_110080	3.1	long-chain-fatty-acid-CoA ligase, putative
TGME49_110230	3.1	hypothetical protein, conserved
TGME49_113030	3.1	hypothetical protein
TGME49_018590	3	protein phosphatase 2C, putative
TGME49_022110	3	UBA/TS-N domain-containing protein
TGME49_025820	3	hypothetical protein
TGME49_026730	3	aconitate hydratase, putative
TGME49_027450	3	mRNA decapping protein, putative
TGME49_028200	3	vacuolar (H+)-ATPase G subunit domain-containing protein
TGME49_031960	3	ppg3, putative
TGME49_051470	3	hypothetical protein
TGME49_067270	3	divalent metal transporter, putative
TGME49_077930	3	hypothetical protein
TGME49_091080	3	TATA-box binding protein, putative
TGME49_093550	3	hypothetical protein
TGME49_105930	3	hypothetical protein
TGME49_111620	3	WD repeat domain-containing protein
TGME49_113290	3	phosphatidylinositol-4-phosphate 5-kinase, putative