A microarray approach to understanding stress in a coral reef fish

Thesis submitted by
Karin Sonja KASSAHN B.Sc.(Hons), The University of Adelaide
in November 2006

for the degree of Doctor of Philosophy
in Zoology
within the School of Marine and Tropical Biology
James Cook University

Statement of Access

I, the undersigned, the author of this thesis, understand that James Cook University will make this thesis available for use within the University Library and, via the Australian Digital Theses network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and;

I wish this work to be embargoed until 22nd of February 2008.

22/02/2007

(Karin Sonja Kassahn) (Date)

Statement of Sources

DECLARATION

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

22/02/2007

(Karin Sonja Kassahn) (Date)

Electronic Copy

I, the undersigned, the author of this work, declare that the electronic copy of this thesis provided to James Cook University, is an accurate copy of the print thesis submitted, within the limits of the technology available.

22/02/2007

(Karin Sonja Kassahn) (Date)

Statement of the contribution of others

All chapters of this thesis include collaborative work with my supervisors Prof. Ross H. Crozier and Dr. M. Julian Caley. The thesis chapters two, three, and four also include collaborative work with Dr. Alister C. Ward, Dr. Glenn Stone and Dr. Ashley R. Connolly. While undertaking these collaborations I was responsible for the project concept and design, the collection of the majority of the samples, the laboratory work, the statistical analyses, synthesis, and the preparation of manuscripts and this thesis.

Funding for this project was provided by James Cook University and the Great Barrier Reef Marine Park Authority under the Science for Management Award program. My supervisors assisted with funds obtained from James Cook University (RHC), The Australian Research Council (RHC, MJC), The Estate of W. V. Scott (RHC), and the Hermon Slade Foundation (MJC). In-kind support was provided by Dr. Alister C. Ward (microarray slides), and the Adelaide Microarray Facility (use of laboratory facilities). I received stipend support from an International Postgraduate Research Scholarship (Australian Government) and School of Tropical Biology (James Cook University) as well as a completion scholarship from James Cook University.

Lewis Anderson and Dr. Line Bay provided assistance with the collection of *Pomacentrus moluccensis* and the performance of stress experiments. Zebrafish genomic DNA samples were kindly provided by the Zebrafish Facility, The University of Adelaide.

Dr. Glenn Stone provided support with the experimental design of the microarray experiments and the statistical analyses of the microarray data. Dr. Ashley R. Connolly provided support with the design and performance of the microarray experiments. Ching Crozier provided assistance with the laboratory work.

Dr. Andrew Perkins, Milena Gongora, Simon Wilkins and The ARC Special Research Centre in Function and Applied Genomics at The Institute for Molecular Bioscience, University of Queensland, provided assistance with the annotation of the zebrafish microarray features.

Lauryne Grant, Caroline Chong, Zoe Baker, and Dr. Matthew Symonds provided help with proof-reading and editing of this thesis.

This work conforms to ethical requirements at relevant authorities and collection of animals was authorised from the Great Barrier Reef Marine Park Authority, permit number G04/12331.1.

Acknowledgements

Many people have helped me conduct and complete the research in this thesis. I would like to thank:

• My supervisors and mentors, Prof. Ross Crozier and Dr. Julian Caley, for their continuous support and guidance. It has been great working with you. Thank you for all your support.

• My collaborators Dr. Ashley Connolly, Dr. Alister Ward and Dr. Glenn Stone for their many contributions to this work. It has been very enjoyable working with you.

• Ching Crozier for her help with the laboratory work and being such a kind and accommodating lab manager.

• Prof. David Miller for his advice with the microarray experiments and many interesting discussions.

• The organisations that provided financial assistance: James Cook University, GBRMPA and grants awarded to my supervisors.

• Staff at The Lizard Island Research Station, especially Lyle and Anne, for providing assistance during field collections and stress experiments.

• The Zebrafish Facility, The University of Adelaide for providing the zebrafish samples.

• Dr. Andrew Perkins, Milena Gongora, and Simon Wilkins at the Institute for Molecular Bioscience, University of Queensland, for help with the annotation of the zebrafish array features.

• The members of the Crozierlab and fellow postgraduate students and post-docs at JCU, in particular Lewis Anderson, Zoe Baker, Michelle Boyle, Sara Ceccarelli, Caroline Chong, Faye Christidis, Jane DeGabriel, Kirsten Evertz, Michelle Guzik, Kylie Gorringe, Ben Moore, Thomas Rayner, Familie Schlüns, Angela Shuetrim, Matthew Symonds for many discussions, encouragement throughout this project and, above all, your friendship.

• My officemates Lauryne Grant and Rodrigo Esparza-Salas for many discussions, hours of entertainment, and your fabulous officemateship.

General Abstract

Coral reef fishes are expected to experience a rise in sea surface temperatures due to climate change. How well tropical reef fishes will respond to these increased temperatures and which genes are important in the response to elevated temperatures is not known. Microarray technology provides a powerful tool for gene discovery studies, but the development of microarrays for individual species can be expensive and time-consuming. There often are, however, microarrays available for related species. I show that inter-species genomic hybridisation experiments can be used to assess which genes are conserved enough for microarray analysis across species, and thus introduce a novel application of microarray technology. I performed a series of tests to determine whether a microarray developed for the zebrafish *Danio rerio* is useful for measuring gene regulation in the coral reef fish *Pomacentrus moluccensis*. I hybridised genomic DNA from both taxa onto the *D. rerio* microarray, and based on significant cross-hybridisation, inferred that most genes share significant sequence similarity between the two taxa. I also sequenced eight nuclear genes. These genes showed an average sequence similarity of 81%. Finally, I used quantitative real-time PCR to validate the microarray data for differential expression. The results of the genomic hybridisation experiments, direct sequence comparisons, and quantitative real-time PCR indicate that the *D. rerio* microarray is useful for measuring gene regulation in *P. moluccensis*. I then used the *D. rerio* microarray to characterise the transcriptional responses of *P. moluccensis* to elevated temperatures over five days. Heat stress elicited differential expression of 324 genes. The functions of heat-responsive genes indicated that prolonged heat exposure leads to oxidative stress and protein damage, challenges the immune system, and causes re-allocation of energy sources. I have shown that a temperature increase of three degrees above normal can lead to significant gene regulation in a coral reef fish suggesting that climate change will have measurable impacts upon coral reef fish physiology. In order to identify upstream regulators of the transcriptional responses observed and test for the presence of a general stress response, I measured the early gene responses of *P. moluccensis* to hypoxic, hyposmotic, cold and heat shock. Early stress responses three hours after exposure were generally associated with a

suppression of transcription, but the responses of individual genes varied depending on the type of stressor applied. Only a few genes showed consistent regulation across stress treatments. However, a series of gene functions showed consistent responses across stress treatments, suggesting that there are common effects of stress on biological function. I present a conceptual model of the interactions between stress responses at different levels of biological organisation. I propose that stress commonly leads to a reduction in cellular oxygen levels and oxidative stress. Reduced cellular oxygen levels initiate endocrine, cardio-respiratory, and cellular responses many of which are aimed at restoring cellular oxygen balance. Oxidative stress in turn activates certain signal transduction pathways, immediate early genes, and transcription factors. The transcriptional stress profiles measured in environmental genomic studies are likely the result of the activation of these redox-sensitive signalling pathways and transcription factors. Further, I tested whether genes with stress-related functions evolve at accelerated rates. To do this, I competitively hybridised genomic DNA from *D. rerio* and *P. moluccensis* to a *D. rerio* microarray. 985 genes showed evidence of accelerated rates of sequence evolution between *D. rerio* and *P. moluccensis*. Rapidly diverging genes were over-represented for receptor, transcription co-activator, and cell signalling functions, but not for stress-related gene functions. I obtained orthologous sequences to *D. rerio* for the teleosts *Takifugu rubripes* and *Gasterosteus aculeatus*. A selection of rapidly diverging candidate genes showed accelerated rates of sequence evolution across multiple teleost lineages. I have shown that genomic hybridisation experiments on microarrays can be successfully used to identify rapidly diverging genes, in particular in species that currently lack genome sequence data and for which bioinformatic approaches are thus not applicable.

Table of contents

List of Tables

List of Figures

CHAPTER 1

General Introduction

Current models of climate change predict a rise in sea surface temperatures of between 2ºC and 5ºC by the year 2100 (IPCC 2001, Done *et al.* 2003). Such increased temperatures can have adverse effects on coral reef organisms. For example, coral bleaching can occur when temperatures rise only 1ºC to 2ºC above normal values (Hoegh-Guldberg 1999). The predicted climate change, therefore, has the potential to induce significant physiological stress and mortality, and may drive shifts in geographic ranges (Parmesan and Yohe 2003). The potential of climate change to induce stress and alter animal function has already been investigated in a variety of organisms (Hoegh-Guldberg 1999, Etterson and Shaw 2001, Pörtner *et al.* 2001, Hoffmann *et al.* 2003, Stillman 2003, Somero 2005). But we currently lack data to predict the likely effects of elevated sea surface temperatures on the physiology of coral reef fishes. We, therefore, need to understand the extent to which rising temperatures will induce stress in these species and the capacity of these species to evolve resistance to such stresses.

Stress may be regarded as a displacement from homeostasis (Johnson *et al.* 1992, Moberg 2000), causing a potentially injurious change in a biological system (Hoffmann and Parsons, 1991). The negative effects of stress manifest themselves at all levels of biological organisation. For example, the injurious effects of stress commonly include impairments of development, growth and reproduction (Hochachka and Somero 2002). In addition, stress can suppress immune functions, making the organism more prone to developing pathologies (Johnson *et al.* 1992, Moberg 2000). Stress further causes changes in cardio-respiratory, endocrine, cellular, and molecular functions (Johnson *et al.* 1992, Somero and Hofmann 1996, Bonga 1997, Feder and Hofmann 1999, Hochachka and Somero 2002, Pörtner 2002, Somero 2005). Finally, stress commonly leads to extensive changes in gene regulation (Gasch *et al.* 2000, Gracey *et al.* 2001, Buckley *et al.* 2003, Enjalbert *et*

al. 2003, Williams *et al*. 2003b, Gracey *et al*. 2004, Koskinen *et al*. 2004, Podrabsky and Somero 2004, Yoshida *et al*. 2004, Krasnov *et al*. 2005).

At present, however, understanding gene regulation during stress still poses many challenges. In particular, the significance of gene regulation for organismal function often remains unclear, because the interactions between physiological and transcriptional stress responses are not well understood. Thus far, reviews of stress responses have focused on individual aspects of the stress response in isolation (Johnson *et al*. 1992, Somero and Hofmann 1996, Bonga 1997, Feder and Hofmann 1999, Hochachka and Somero 2002, Pörtner 2002, Somero 2005). A conceptual model, which would synthesise data on different aspects of 'the stress response' (see, for example, Seyle 1936) and which would account for the key regulators of physiological and transcriptional stress responses and their main interactions, would add to our current understanding of stress responses. In this context, microarray technology and the assessment of gene regulation during stress have already provided new insights into the effects of stress on biological function.

DNA microarrays for gene expression analysis were first developed in the mid 1990s (Schena *et al*. 1995). DNA microarrays have since found a broad range of applications, including functional genomic studies in ecology and evolution (Feder and Mitchell-Olds 2003). In functional genomic studies, and because genes are transcribed into RNA which is translated into proteins that carry out biological functions, assessment of the transcriptome can reveal how genetic information is translated into biological function. At present, however, there are few examples of the use of microarrays to understand gene regulation in coral reef organisms and these are currently limited to the study of gene responses in corals (e.g. Edge *et al*. 2005). The limited application of microarray technology to study coral reef organisms is likely due to the challenges associated with transferring this technology to species for which genomic resources, such as EST libraries or genome sequence data, are currently not available.

In general, commercial microarrays are only available for widely studied species and species of interest to large research communities. Where there are no commercial microarrays available and for the study of non-model species, researchers wanting to apply microarray technology may pursue one of the following strategies. Firstly,

custom cDNA microarrays can be developed for any species of interest. The development of new cDNA microarrays requires the isolation and cloning of a cDNA library, which can then be PCR amplified and spotted onto microarray slides (Oleksiak *et al*. 2001). Clones that show interesting regulation during gene expression profiling are sequenced and the sequences are compared against genome databases to determine the identity of the clones. While the development of cDNA microarrays is, in principle, possible for all organisms, developing a new microarray is time-consuming and expensive. Given the broad range of species that are of interest to functional genomic studies, it is unlikely that microarrays will be developed for all these species any time soon.

An alternative strategy to using a species-specific microarray is to use a microarray developed for a closely related species. Microarray experiments that are performed using a species other than the one for which the microarray was originally designed, are commonly referred to as heterologous microarray experiments. Human cDNA microarrays, for example, have been used to study bovine, pig and even salmon gene regulation (Medhora *et al*. 2002, Tsoi *et al*. 2003, Adjaye *et al*. 2004). Oligonucleotide microarrays can also be successfully applied in heterologous microarray experiments (Ji *et al*. 2004). In general, cross-hybridisation on microarrays is positively correlated with sequence similarity (Wu *et al*. 2001, Hinchliffe *et al*. 2003, Brunelle *et al*. 2004). The success of heterologous microarray experiments would thus depend on the ability of the heterologous DNA species to bind to the microarray, which is in turn determined by the level of sequence similarity between the two taxa. In general, microarray analysis is robust to some level of sequence divergence and genes with less than 25 percent sequence divergence generally show significant cross-hybridisation (Kane *et al*. 2000). With increasing sequence divergence, the ability to accurately measure gene regulation decreases (Renn *et al*. 2004). However, the level of sequence divergence differs across genes with some genes diverging more rapidly than others (Makalowski *et al*. 1996). Therefore, heterologous microarrays may be more successfully employed to measure gene responses at conserved gene loci, while genes that have diverged significantly may fail to cross-hybridise.

For these reasons, heterologous microarray experiments are most suitable for gene discovery studies where the aim is to identify candidate genes for studies of gene regulation. In this context, loss of signal for some genes on the microarray would not compromise the aims of the study and genes with the most interesting gene regulation can be further validated using an alternative method, such as quantitative real-time PCR. In this thesis, I demonstrate that inter-species hybridisation experiments using genomic DNA can be used to identify which of the genes represented on a microarray are conserved enough to allow heterologous microarray analysis. For this purpose, genomic DNA (gDNA) is extracted from the species of interest and the species for which the microarray was designed and both gDNA samples are competitively hybridised to the microarray. Genes that share significant sequence similarity between the two species will produce even signal intensities, while genes with significant sequence divergence will show reduced signal intensities in the heterologous species. To my knowledge, this is the first study to use genomic hybridisation experiments to test the performance of a heterologous microarray platform.

Typically, comparative genomic hybridisation experiments are used to identify gene copy number changes and chromosomal deletion and duplication events (Pollack *et al.* 1999). More recently, comparative genomic microarray experiments have also been used to compare sequence similarity across genomes and to identify rapidly diverging genes between closely related taxa (Kim *et al.* 2002, Hinchliffe *et al.* 2003, Brunelle *et al.* 2004). Rapidly diverging genes interest evolutionary biologists because genes with high rates of evolutionary change have been implicated in the process of speciation and species diversification (Turner *et al.* 2005, Harr 2006). Furthermore, high rates of evolutionary change may indicate strong directional selection and molecular adaptation, especially where selection on gene function can be directly related to environmental pressures (Tautz and Schmid 1998, Lecompte *et al.* 2001, Matzkin 2005, Fairhead and Dujon 2006). Examples of rapidly diverging genes include genes with functions in reproduction, especially in sperm competition and species recognition (Swanson and Vacquier 1998, Hellberg and Vacquier 1999, Swanson *et al.* 2001), and genes with functions in environmental interactions (Jordan *et al.* 2001, Clark *et al.* 2003). Rapidly diverging genes thus appear to be good candidates to study the mechanisms underlying molecular adaptation.

Thus far, the application of microarray technology for the identification of rapidly diverging genes has been restricted to the study of microbial and fungal genomes (Hinchliffe *et al*. 2003, Brunelle *et al*. 2004, Le Quere *et al*. 2006). However, microarray approaches for the identification of rapidly diverging genes should be feasible for any organism and are of particular interest for application in species with few genomic resources. Such species do not lend themselves to bioinformatic approaches for the identification of rapidly diverging genes because of a lack of genome sequence data. The potential of comparative genomic hybridisation experiments on microarrays to identify rapidly diverging genes across a broader range of organisms thus warrants further investigation.

Because of their ability to screen thousands of gene transcripts and gene sequences at once, microarrays are an attractive tool for environmental genomic and comparative genomic studies. As outlined above, there are two main applications of microarray technology in ecology and evolution. Firstly, microarrays can be used to assess the transcriptome and to measure gene regulation, for example during conditions of environmental stress (Figure 1). Since the transcriptome is the mediator between the genome and the proteome, understanding of the transcriptome can contribute to our understanding of how genetic information is translated into biological function. Secondly, microarrays can be used to assess the genome itself, to understand sequence divergence, and to identify those genes, which can be measured in heterologous microarray experiments (Figure 1). The latter is a novel application of microarray technology developed as part of this PhD research.

The coral reef fish *Pomacentrus moluccensis* is an abundant damselfish species on the Great Barrier Reef (Figure 2). This species is particularly amenable to housing in aquaria and, therefore, is a suitable species on which to perform environmental stress experiments and to investigate the effects of heat on gene regulation in coral reef fishes. However, there are currently no genomic resources or microarrays available for *P. moluccensis* or any other species of coral reef fishes.

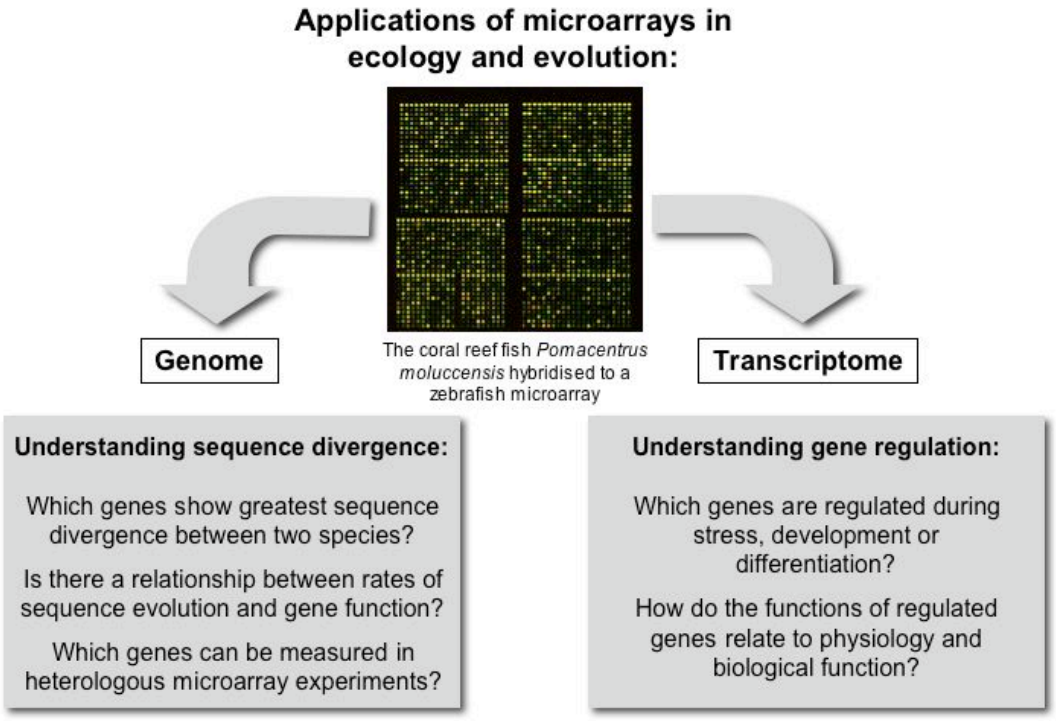


Figure 1 Applications of microarray technology in ecology and evolution.

Figure 2 The coral reef fish *Pomacentrus moluccensis* on the Great Barrier Reef, Australia.

In contrast, the zebrafish *Danio rerio* has been widely studied, its genome has been completely sequenced, and there is an olignucleotide (65mers) microarray available for this species consisting of nearly 16,000 gene products. This *D. rerio* microarray is one of the largest teleost microarrays available and would, therefore, offer the best chance of detecting large numbers of genes involved in environmental stress responses. Because the *D. rerio* microarray has been computationally designed and is widely used in the zebrafish research community, the identity of the clones represented on the microarray is known and there are continuing efforts into improving the annotation of this microarray. For these reasons, the *D. rerio* microarray was an obvious choice to elucidate environmental stress responses in coral reef fishes.

The specific aims of this PhD research were:

1. To examine strategies for how microarray technology can be applied to the study of non-model species and, specifically, to test the performance of the *D. reri*o microarray to study gene regulation in the coral reef fish *P. moluccensis*.
2. To apply heterologous microarray experiments to study environmental stress responses in *P. moluccensis*, in particular transcriptional responses to heat stress.
3. To apply comparative genomic hybridisation experiments and bioinformatic approaches to identify rapidly diverging genes in teleosts.
4. To develop a conceptual model of the key regulators of organismal, cellular, molecular, and transcriptional responses to stress and their main interactions.

This thesis is structured as a series of stand-alone but conceptually interconnected publications. In **Chapter 2**, I report on the utility of a *D. rerio* microarray for measuring gene regulation in *P. moluccensis*. I demonstrate that inter-species genomic hybridisation experiments can identify which genes can be measured in microarray analysis across species. In order to test the performance of the heterologous microarray to measure gene regulation in *P. moluccensis*, I further validate differential gene expression for a selection of candidate genes using quantitative real-time PCR.

In **Chapter 3**, I report heterologous microarray experiments designed to study environmental stress responses in *P. moluccensis*, in particular transcriptional responses to heat stress. This is the first study to investigate the effects of elevated temperatures on gene regulation in a coral reef fish. In order to understand the generality of transcriptional responses to heat, I examine how different types and severities of stress affect gene regulation. For this purpose, I designed a series of experiments, which allowed me to assess the effects of different dimensions of stress on gene regulation. Throughout this thesis I use the term 'stress' to refer to some form of disturbance from homeostasis or normal biological function. Hence, genes that change expression level following different treatments indicate a change in biological function; the treatment is thus considered a 'stressor'; and the change in gene expression is considered to form part of the 'stress response'.

In **Chapter 4**, I report on inter-species genomic DNA hybridisation experiments designed to identify rapidly diverging genes between *P. moluccensis* and *D. rerio*. I develop a cut-off threshold based on relative signal intensities to distinguish between relatively conserved and rapidly diverging genes. I examine the functions of rapidly diverging genes and test for an over-representation of gene functions amongst rapidly diverging genes. In addition, I test whether genes identified as rapidly diverging between *P. moluccensis* and *D. rerio* by means of comparative genomic microarray experiments generally show high rates of sequence evolution in teleosts. For this purpose, I used a bioinformatic approach and obtained orthologous sequences to *D. rerio* for the teleosts *Takifugu rubripes* and *Gasterosteus aculeatus.* I estimated rates of sequence evolution based on tree branch lengths and the number of nucleotide changes per site.

In **Chapter 5**, I review current literature on physiological, endocrine, cellular, molecular, and transcriptional responses to stress and develop a conceptual model of the interactions between these different levels of biological organisation. I identify key players of the stress response and their main interactions.

Lastly, in **Chapter 6**, I summarise and discuss the main findings and conclusions of this PhD thesis. I discuss questions that remain in our understanding of environmental stress responses and point out areas of potential future research.

CHAPTER 2

Heterologous microarray experiments used to identify the early gene response to heat stress in a coral reef fish

Publication: Kassahn KS, Caley MJ, Ward AC, Connolly AR, Stone G, Crozier RH. (*In press*) Heterologous microarray experiments used to identify the early gene response to heat stress in a coral reef fish. *Molecular Ecology*.

Abstract

Coral reef fishes are expected to experience a rise in sea surface temperatures due to climate change. How well tropical reef fishes will respond to increased temperatures and which genes are important in the response to elevated temperatures is not known. Microarray technology provides a powerful tool for gene discovery studies, but the development of microarrays for individual species can be expensive and time-consuming. In this study, I tested the suitability of a *Danio rerio* oligonucleotide microarray for application in a species with few genomic resources, the coral reef fish *Pomacentrus moluccensis*. Results from a comparative genomic hybridisation experiment and direct sequence comparisons indicate that for most genes there is considerable sequence similarity between the two species, suggesting that the *D. rerio* microarray is useful for genomic studies of *P. moluccensis*. I employed a heterologous microarray approach to characterise the early transcriptional responses to heat stress in *P. moluccensis*. A total of 111 gene loci, many of which are involved in protein processing, transcription, and cell growth, showed significant changes in transcript abundance following exposure to elevated temperatures. Changes in transcript abundance were validated for a selection of candidate genes using quantitative real-time PCR. This study demonstrates that heterologous microarrays can be successfully employed to study species for which specific microarrays have

not yet been developed, and so have the potential to greatly enhance the utility of microarray technology to the field of environmental and functional genomics.

Introduction

Current models of climate change predict a rise in sea surface temperatures of between 2ºC and 5ºC by the year 2100 (IPCC 2001, Done *et al.* 2003). Such increased temperatures can have adverse effects on coral reef organisms. For example, coral bleaching can occur when temperatures rise only 1ºC to 2ºC above normal values (Hoegh-Guldberg 1999). The predicted climate change, therefore, has the potential to induce significant physiological stress and mortality, and may drive shifts in geographic ranges (Parmesan and Yohe 2003). To predict the effects of rising sea surface temperatures on reef species, we need to understand the extent to which rising temperatures will induce stress in these species and the capacity of these species to evolve resistance to such stresses.

While the responses of coral reef fishes to temperature stress are poorly understood, temperature stress responses in temperate fishes include changes in gene expression, which vary depending on the type and duration of the stress. For example, constant high, constant low and fluctuating temperatures initiate different transcriptional responses in the eurythermal killifish (*Austrofundulus limnaeus*) (Podrabsky and Somero 2004). In channel catfish (*Ictalurus punctatus*), cold stress initiates rapid expression changes at a large number of loci (Ju *et al.* 2002), and the transcriptional response to cold stress in common carp (*Cyprinus carpio*) appears to resemble that observed during muscular atrophy (Gracey *et al.* 2004). Acclimation to seasonal temperature shifts in temperate fishes also entails changes in gene expression. The expression of warm temperature acclimation-related protein (Wap65) increases in summer-acclimated goldfish (*Carassius auratus*) (Kikuchi *et al.* 1995) and in warm-acclimated populations of the salt-water minnow (*Fundulus heteroclitus*) (Picard and Schulte 2004). Similarly, the expression of β-actin and 5.8 rRNA increases during seasonal acclimation to higher temperatures in common carp (Vera *et al.* 1997, Sarmiento *et al.* 2000). Gene expression changes in response to temperature stress are thought to help maintain cellular homeostasis and function during altered environmental conditions (Gabai and Sherman 2002, Sonna *et al.* 2002).

Microarray technology has become a powerful tool for studying stress-related gene expression responses in teleosts because it allows screening of a large proportion of the transcriptome (Gracey *et al*. 2001, Williams *et al*. 2003, Gracey *et al*. 2004, Koskinen *et al*. 2004, Picard and Schulte 2004, Podrabsky and Somero 2004, Krasnov *et al*. 2005). In some cases where a species-specific microarray has not been available, microarrays developed for a related species have been used effectively (Girke *et al*. 2000, Hittel and Storey 2001, Nowrousian *et al*. 2005). Currently, however, there are only a few examples of cross-species microarray experiments in teleosts and these are limited to cDNA arrays (Hogstrand *et al*. 2002, Tsoi *et al*. 2003, Renn *et al*. 2004, Rise *et al*. 2004, Aubin-Horth *et al*. 2005). Oligonucleotide arrays may provide increased specificity and sensitivity as the oligos deposited on the array can be especially designed for this purpose (Kane *et al*. 2000, Irizarry *et al*. 2005). There are currently no examples of cross-species microarray experiments using long oligonucleotide microarrays. Given the potential utility of microarrays for understanding stress responses in tropical reef species and the low probability of the development of species-specific microarrays for these species in the near future, the potential to successfully perform heterologous hybridisations on different microarray platforms warrants further investigation.

In this study, I test the suitability of an oligonucleotide microarray developed for the zebrafish *Danio rerio* to measure gene responses in the coral reef fish *Pomacentrus moluccensis*. The most recent common ancestor of *P. moluccensis* (Perciformes, Acanthopterygii, Teleostei) and *D. rerio* (Cypriniformes, Ostariophysi, Teleostei) dates from between 110 and 300 million years ago (Cantatore *et al*. 1994, Wittbrodt *et al*. 2002). Therefore, successful cross-hybridisation of *P. moluccensis* to the *D. rerio* microarray would depend on significant sequence conservation. Sequence similarity between *D. rerio* and *P. moluccensis* was estimated by direct sequence comparison and by performing a comparative genomic hybridisation experiment where genomic DNA from *D. rerio* and *P. moluccensis* was competitively hybridised to the *D. rerio* microarray. I then employed this heterologous microarray approach to compare the transcriptome of heat-stressed *P. moluccensis* to that of *P. moluccensis* kept at ambient temperature. The expression changes were further tested for a selection of candidate genes using quantitative real-time PCR. Based on these results

I evaluate the utility of heterologous hybridisations on long oligonucleotide microarrays and discuss the early gene response of *P. moluccensis* to heat stress.

Materials and Methods

Thermal stress experiments

Adult *P. moluccensis* were collected around Lizard Island, northern Great Barrier Reef, Australia (14º40'S, 145º28'E) by divers on SCUBA using barrier and scoop nets and transferred to the Lizard Island Research Station. Fish were housed in groups of up to 20 individuals in aquaria supplied with flow-through seawater at ambient temperature (approximately 26ºC). Sections of PVC pipes and dead branching coral were provided for shelter. After two days of acclimation to these conditions, animals were transferred for three hours to individual tanks containing water of either 34ºC (thermal stress) or 26ºC (ambient, control). Previous studies have shown that the transcriptional response to temperature stress in teleosts starts within hours after exposure (Ju *et al.* 2002, Podrabsky and Somero 2004), suggesting that a relatively short exposure to heat stress should lead to measurable changes in gene expression. It is possible that handling and capture stress amplified the effects of the heat stress treatment. However, fish that were kept at ambient temperature would have also suffered from the same handling and capture stress. It is likely, therefore, that the differences between treatments observed here are mainly associated with the effects of heat stress. Following these thermal treatments, fish were killed by placing them on ice. Livers were excised and stored in RNAlater™ (Qiagen) for microarray and quantitative real-time PCR analysis. Gonads were preserved in FAAC (formaldehyde 4%, acetic acid 5%, calcium chloride 1.3%) so the fish could be sexed. In teleost fishes, gene expression can vary considerably between the sexes with females showing larger inter-individual variation than males (Williams *et al.* 2003). Because such inter-individual variation could obscure the identification of those genes involved in the stress response, we restricted our analyses to adult male fish of similar standard length (46 ± 4 mm).

Microarray platform

Because there is currently no microarray available for any species of pomacentrid fish (family Pomacentridae), I tested the suitability of a *D. rerio* microarray for application in *P. moluccensis*. I used the Compugen 16K *D. rerio* oligonucleotide microarray, which contains 16,399 oligos (65-oligomers) representing 15,806 unique *D. rerio* gene clusters plus controls. This array platform represents one of the largest fish microarrays available and therefore offered the greatest chance of detecting large numbers of genes involved in temperature stress responses. The arrays were printed by the Adelaide Microarray Facility. The list of genes immobilized on the array is available at http://www.microarray.adelaide.edu.au/libraries/microarrays.html.

Comparative genomic DNA hybridisation experiment

In order to assess the hybridisation potential of *P. moluccensis* DNA to the *D. rerio* microarray, I performed four comparative genomic DNA hybridisations on *D. rerio* microarrays. Genomic DNA was extracted from fish caudal fins of both species using CTAB (Murray and Thompson 1980) and quantified using a spectrophotometer. Five μg of gDNA per sample was digested with *DpnII* and labelled using the BioPrime®Plus Array CGH Genomic Labeling System (Invitrogen). In order to account for potential dye bias, dye usage was swapped between *P. moluccensis* and *D. rerio* gDNA samples. Genomic DNA from three individuals of *D. rerio* were pooled. Four biological replicates of *P. moluccensis* gDNA were each mixed with an aliquot of fluorescently labelled *D. rerio* pooled gDNA. Each of the four gDNA mixtures was hybridised to an individual *D. rerio* microarray. Prior to hybridisation, each microarray slide was immersed in distilled water at 60ºC for 5min and dried by centrifugation at 650 x g for 5min. Fifty μg of human Cot-1 was added to each labelled gDNA sample, dried under reduced pressure, resuspended in 14μl formamide and 14μl of 6.25 X SSC, denatured by heating to 100ºC for 3min and transferred directly to ice. Finally, 0.6μl of 10% SDS was added to each sample. The probes were applied to the array and incubated at 42ºC overnight in a humidified chamber. The arrays were washed in 0.5 X SSC containing 0.01% SDS for 1min, 0.5 X SSC for 3min and 0.2 X SSC for 3min. The slides were scanned using an Axon 4000B microarray scanner. Single image .tif files were saved for data analysis.

Microarray analysis of heat-stressed P. moluccensis

Total RNA from liver tissue of heat-stressed *P. moluccensis* and *P. moluccensis* kept at ambient temperature was extracted using TRIzol® (Invitrogen) according to the manufacturer's instructions and purified using RNeasy™ columns (Qiagen), ethanol-precipitated and subsequently resuspended in nuclease-free water. The concentration and purity of RNA was determined by spectrophotometer readings at 260nm and 280nm. The integrity of the RNA was confirmed by agarose gel electrophoresis. Forty μg of total RNA was mixed with 4μg of anchored polyT(V)N and 0.5μg of random hexamers, and incubated at 70ºC for 10min. The samples were placed on ice and mixed with 6μl of 5X Superscript II buffer (Invitrogen), 2μl of 0.1M dithiothreitol (DTT), 2μl of Superscript II (200U/μl) (Invitrogen) and 0.6μl of aminoallyl (aa) dNTP mix (25mM dATP, 25mM dGTP, 25mM dCTP, 10mM dTTP and 15mM aa dUTP). After incubation at 42ºC for 2.5 hours, residual RNA was hydrolysed in 10μl of 0.25M NaOH and 10μl of 0.5M EDTA (pH 8.0) by incubating at 65ºC for 15min. The reactions were neutralised by adding 15μl of 0.2M acetic acid and purified using a QIAquick PCR purification kit (Qiagen). The purified cDNA was dried under reduced pressure, dissolved in 9μl of 0.1M NaHCO3 (pH 9.0), mixed with Cy3 or Cy5 and left in the dark to couple at room temperature for 60min. The labelled cDNA was mixed with 41μl of MilliQ water and purified using a QIAquick PCR purification kit (Qiagen). The purified fluorescently labelled cDNA samples were eluted into a clean tube with 90μl of MilliQ water and dried under reduced pressure. Equal amounts of RNA samples from four *P. moluccensis* kept at ambient temperature were pooled and used as a common reference. Labelled cDNA from four heat-stressed *P. moluccensis* was competitively hybridised against the pooled control in four microarray hybridisations using dye swaps. Microarray hybridisations of labelled cDNA were performed as described for the gDNA samples above except that 2.5μg human Cot-1 and 4μg poly A were added to each labelled cDNA sample.

Microarray statistical data analysis and data mining

The Cy5 and Cy3 fluorescent signal intensity of each gene on the array was extracted using SPOT software (CSIRO Mathematical and Information Sciences, Australia).

The background fluorescence was subtracted and the ratio of the resultant signal intensities (Cy5/Cy3) was $\log_2$-transformed. Statistical analyses were performed using the software package LIMMA implemented in the R statistical software environment (Smyth 2005), following Smyth (2004). The transformed signal intensities of each grid on the array and global signal intensity were print-tip Loess normalised (Smyth and Speed 2003), and scale normalised between arrays. Loess normalisation subtracts a Loess regression curve from the data in order to linearise the data, while scale normalisation between arrays ensures that signal intensities are comparable across arrays.

Pomacentrus moluccensis gDNA will not preferentially hybridise to a *D. rerio* oligonucleotide microarray in the presence of *D. rerio* gDNA. I was, therefore, able to measure the random noise in the comparative genomic hybridisation experiment by using spots with positive M-values (Figure 1). Due to the Loess normalisation discussed above potentially shifting the M-value corresponding to equal hybridisation away from M=0, I used the modal M-value, estimated from a density estimate of M-values, as the centre of the distribution. The 99th percentile of spots with M-values greater than this mode was taken as a threshold to decide which spots had poorly hybridised with the *P. moluccensis* gDNA and spots with M-values that exceeded this threshold below the mode were flagged as such (Figure 1). Genes with negative M-values outside this range are attributed to *D. rerio* gDNA hybridising in preference to *P. moluccensis* gDNA due to sequence divergence between the two taxa. In contrast, genes that lie within the 99th percentile show similar hybridisation signal in *P. moluccensis* and *D. rerio* and are thus considered to share significant sequence similarity in the two taxa. Although I take no account of variability amongst replicates and the 99th percentile is somewhat arbitrarily chosen, this approach identifies more spots as poorly hybridised than several other methods I have considered (see Chapter 4). Ultimately, the purpose of the comparative genomic hybridisation experiment was to identify any spots that show any evidence of reduced hybridisation when using *P. moluccensis* on the *D. rerio* microarray. Spots identified in this way were excluded from further analysis of the gene expression data. This approach provides a conservative estimate of the number of genes involved in the heat stress response of *P. moluccensis*.

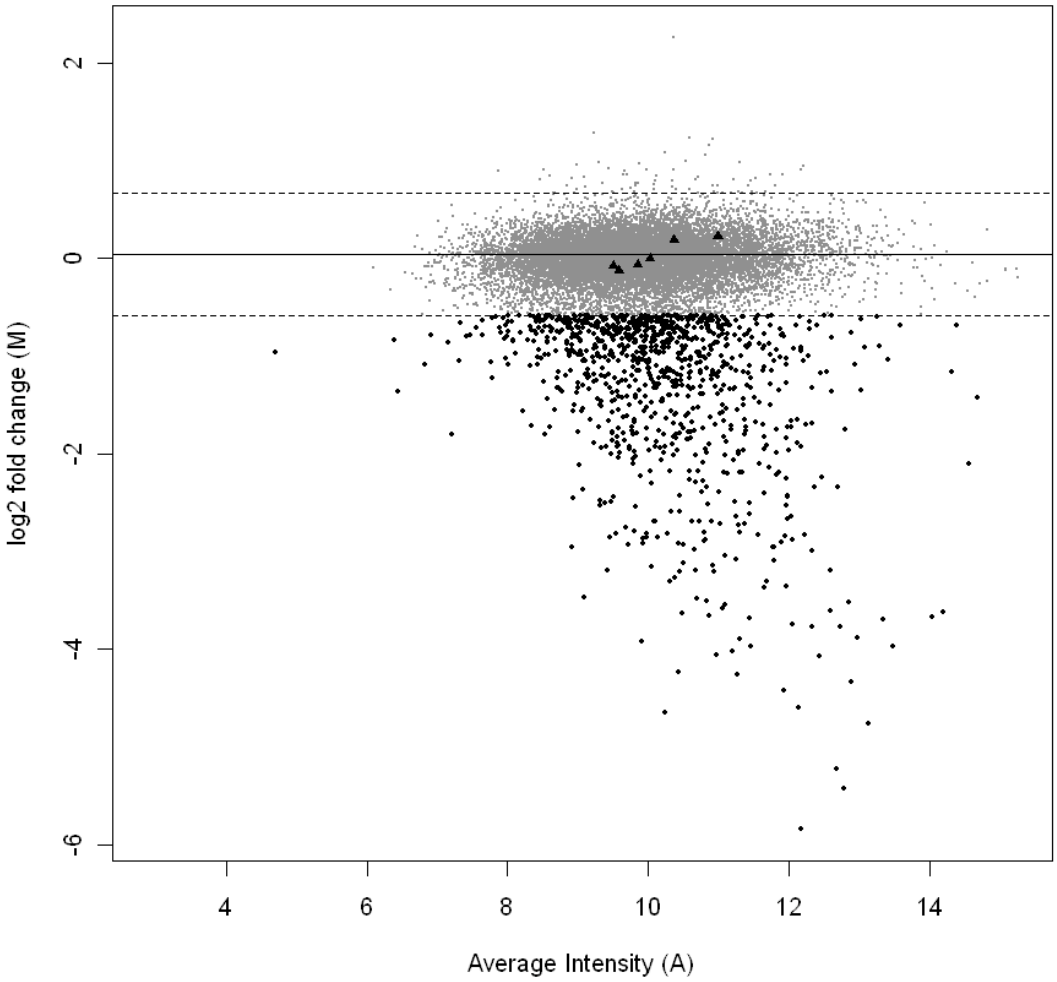


Figure 1 The plot illustrates the relative signal intensity difference between *Pomacentrus moluccensis* and *Danio rerio* against average signal intensity A. Each dot represents one of the 15,806 genes represented on the microarray. The solid line represents the modal M-value of the distribution of M-values. The dashed lines represent the M-value corresponding to the 99th percentile of positive M-values and its extrapolation towards negative M-values. Genes that show reduced hybridisation signal for *P. moluccensis* as compared to *D. rerio* have negative M-values that fall below this threshold (●). MA values for six nuclear gene loci that are known to have high sequence identity between *P. moluccensis* and *D. rerio* are also marked (▲).

For analysis of the gene expression data, a moderated t-statistic was calculated for each gene on the array employing an empirical Bayes method (Smyth 2004). This method uses the information from all genes represented on the array in order to moderate the standard errors of the estimated expression changes for each individual gene. This approach results in more stable inference and improved power and is particularly useful for experiments with small numbers of arrays (Smyth 2004). Benjamini and Hochberg's (1995) method for controlling the false discovery rate (FDR) was used to control experiment-wise error rates in the face of multiple testing. Genes with FDR-adjusted p-values < 0.1 were considered differentially expressed. The raw microarray data on which these analyses are based is deposited at the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/projects/geo/; Barrett et al. (2005)) under the following accessions: GPL3365 (microarray platform), GSE4047 (data series), GSM104737, GSM104739, GSM104741, GSM104742, and GSM92653 to GSM92656 (samples). The raw image .tif files can be downloaded at ftp://ftp.ncbi.nih.gov/pub/geo/DATA/ supplementary/samples/.

Gene annotation was performed using the program Resourcerer 12.0 (Tsai *et al.* 2001). Gene function of identified candidate genes was estimated by gene functions determined for *D. rerio* using a combination of AmiGo (The Gene Ontology Consortium 2000) (http://www.godatabase.org/cgi-bin/amigo/go.cgi), iHop (http://www.pdg.cnb.uam.es/UniPub/iHOP/) (Hoffmann and Valencia 2004), and databases at NCBI (http://www.ncbi.nlm.nih.gov/). Overrepresentation of gene ontologies in candidate genes compared to the total of genes represented on the microarray was tested using the program The Ontologizer 2.0 (Robinson *et al.* 2004) and data files downloaded from the Gene Ontology (gene_ontology.obo accessed September 27, 2005; http://www.geneontology.org/) and the Zebrafish Information Network (gene_association.zfin September 21, 2005; http://zfin.org/cgi-bin/webdriver?MIval=aa-ZDB_home.apg).

Quantitative real-time PCR verification of array data

In order to independently verify that candidate genes for heat stress identified in the heterologous microarray experiments were differentially expressed, I performed quantitative real-time PCR on six candidate genes. These genes were chosen because

their reported gene function made them interesting candidates, and because combined they represent a variety of functional classes. With the exception of β-actin, degenerate primers to amplify the selected candidate genes in *P. moluccensis* were designed using BlockMaker and CODEHOP (http://blocks.fhcrc.org/codehop.html) (Rose *et al.* 1998, Rose *et al.* 2003) and sequence information available on GenBank for other vertebrates. For β-actin I used the primers of Forlano *et al.* (2005) (Appendix 1). Where possible, PCR amplification targeted the same gene region that was represented by the oligo on the microarray. However, the choice of gene region for PCR amplification also depended on sequence information available in GenBank and the availability of suitable priming sites. Gene fragments of 217bp to 729bp per gene were amplified using cDNA synthesised with oligo dT primers. PCR products were cloned into pGEM®-T Easy Vector Systems (Promega) following the manufacturer's instructions. Clones were sequenced using DYEnamic™ ET Dye Terminator (GE Healthcare) on an ABI 377 automated sequencer. Sequence homology to genes of interest was confirmed using BLASTX search, translated query vs. protein database, on the NCBI website (all E values ≤ $3e^{-24}$, Appendix 1). Real-time PCR primers specific for *P. moluccensis* were designed using a combination of Primer Express® Software v.2.0 (Applied Biosystems) and Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (Rozen and Skaletsky 2000) following the recommendations of Bustin (2000) and Giulietti et al. (2001), in particular with respect to primer annealing temperature and amplicon length (Appendix 2).

Samples for quantitative real-time PCR data validation consisted of the same RNA samples employed in the microarray experiment, except for one control sample, which had been completely used during microarray analysis and hence was not available for the real-time PCR assay. In order to maximise sample sizes for quantitative real-time PCR, RNA from five non-stressed fish and two heat-stressed fish that had not been used in the microarray experiment were added to the sample pool, resulting in a total of eight and six samples from non-stressed and heat-stressed individuals, respectively. All RNA samples were treated with DNA-free™ DNase Treatment and Removal Reagents (Ambion) in order to remove any contamination with genomic DNA. The absence of genomic DNA contamination after DNase treatment was confirmed by the absence of bands following PCR with β-actin

genomic DNA primers. The quality and quantity of total RNA was assessed using spectrophotometry and formaldehyde-agarose gel electrophoresis. 1μg of total RNA was used for cDNA synthesis using oligo dT primers and the SuperScript™ III First-Strand Synthesis System for RT-PCR (Invitrogen) following the manufacturer's instructions. After treatment with RNase H, cDNA samples were cleaned of free nucleotides and enzyme using QIAquick PCR clean-up columns (Qiagen) and eluted in 30μl of nuclease free water. The amount of cDNA was quantified using spectrophotometry. PCR reactions were done in duplicate using a Corbett Robotics CAS-1200™ and run on a Corbett Research Rotor-Gene™ 3000. PCR reactions were carried out in a final volume of 15μl containing 7.5μl SYBR Green PCR Master Mix (Applied Biosystems), 5ng of cDNA, and between 50nM and 900nM forward and reverse primers. Optimum primer concentrations for each gene were determined using different combinations of 50nM, 300nM, and 900nM forward and reverse primers (Appendix 2). In order to reduce between-run variability, the same dilution of cDNA samples and the standard were used in each PCR run. PCR conditions were one cycle of 95ºC for 10min, followed by 40 cycles of 95ºC 15sec and 60ºC 1min. During each of the 40 cycles the fluorescence was acquired at the end of the 60ºC step. The PCR run was followed by a melt curve analysis in order to confirm amplification of specific product only. Negative controls for PCR reactions included RNA that had not been reverse-transcribed as well as no-template controls.

The amount of cDNA was quantified using the standard curve method. Each PCR run included serial dilutions of the cDNA sample designated as standard, ranging from 10ng cDNA to 0.05ng cDNA. PCR efficiencies (E) were calculated using the dilution series of the standard and the formula $E = 10^{(-1/slope)}$ implemented in the program REST (Pfaffl *et al*. 2002). PCR efficiencies ranged between 1.90 and 2.06 with R^2 values between 0.994 and 0.999 indicating good PCR performance for all loci. Two commonly used housekeeping genes, 18S rRNA and hypoxanthine-guanine phosphoribosyl transferase (HPRT), were also amplified for each sample for use as endogenous controls. The program BestKeeper by Pfaffl *et al*. (2004) and not-normalised results from the randomisation test in REST© were used to test the stability of the two housekeeping genes. Both, HPRT and 18S were suitable reference genes based on their invariant expression amongst control and treatment samples and the significant correlation between HPRT and 18S values. For relative

quantification, the real-time PCR data were normalised to the geometric mean of 18S and HPRT following the method of Vandesompele *et al*. (2002). The non-parametric Mann Whitney U test for two independent samples was used to test for significant differences in group means between heat-stressed and non-stressed fish.

Direct sequence comparison between D. rerio *and* P. moluccensis

As part of designing the real-time PCR assay, I sequenced segments of the coding region of eight nuclear gene loci in *P. moluccensis*. I was thus able to directly compare sequence identity between *P. moluccensis* and *D. rerio* at these loci. The *D. rerio* data consisted of the clone sequences represented on the microarray and sequence data downloaded from GenBank. Sequences were aligned using Se-Al v2.0a11 (Rambaut 1996) and sequence identity was calculated using MEGA 3.1 (Kumar *et al*. 2004).

Results

Estimation of sequence identity between D. rerio *and* P. moluccensis

Direct sequence comparison between *D. rerio* and *P. moluccensis* at eight nuclear gene loci revealed, on average, 81% DNA sequence identity (Table 1, see Appendix 3 for sequence alignments). The below average level of sequence identity at locus *cebpd* was due to 18 codon insertions and two codon deletions. Six of the eight nuclear gene loci sequenced were represented on the microarray. The M-values for these six loci did not exceed the threshold for M-values set in the comparative genomic hybridisation experiment, indicating that hybridisation to the *D. rerio* array at these loci was comparable for the two species (Figure 1). I have therefore demonstrated that genes with high sequence identity between *P. moluccensis* and *D. rerio* produce similar fluorescent signals in both species when competitively hybridised to the *D. rerio* array.

Table 1 Direct sequence comparison of *Pomacentrus moluccensis* and *Danio rerio* nuclear genes.

Gene name	GenBank Accession *P.moluc-censis* clone	GenBank Accession *D. rerio* clone	Length of sequence alignment (bp)	Percent nucleotide identity	Percentage of nucleotide changes that occur at third codon position	Percent amino acid identity
rhoC	DQ243817	AI959074	165	80.0	81.8	98.2
cdk5	DQ243818	AF203736	588	81.3	81.8	99.0
Loc402870	DQ243819	BM181148/ BC056714	441	79.6	83.3	94.6
cebpd	DQ243820	BE017827/ NM131887	525	63.0	44.3	62.3
cct6a	DQ243821	AI437239/ NM201290	963	79.1	81.1	91.9
β-actin	DQ243822	AF025305	618	90.0	75.8	98.1
HPRT	DQ243824	BC046003	171	79.5	85.7	94.7
18S	DQ243823	BX296557	618	95.3	non-coding	non-coding
Average				**81.0**	**76.3**	**91.3**

The comparative genomic hybridisation experiment identified loci where the hybridisation of *P. moluccensis* to the *D. rerio* array may have been compromised. Using the M-values greater than the mode (M_{mode}=0.038) I chose the 99th percentile as a threshold and classified all spots below the mode minus this threshold as showing evidence of reduced hybridisation in *P. moluccensis*. The negative threshold chosen by this method was -0.584 and this resulted in 985 spots, approximately 6% of genes assayed, being classified as showing evidence of reduced hybridisation (Figure 1, Appendix 4). These loci were excluded from analysis of the microarray gene expression data. Differences in hybridisation signal at these loci were as great as 50fold, but only 112 genes showed a 5fold change or larger. It is likely that hybridisation at these gene loci was compromised due to sequence divergence between *P. moluccensis* and *D. rerio*. However, for most gene loci, the fluorescent signal in the comparative genomic hybridisation experiment was comparable between the two taxa, suggesting that *P. moluccensis* hybridises well to the *D. rerio* array (Figure 1).

The early gene expression response to heat stress

Relative to *P. moluccensis* kept at ambient temperature, heat-stressed *P. moluccensis* showed expression changes in 111 genes, approximately 0.8% of assayed genes, as assessed by FDR (Table 2, Figure 2, Appendix 4). All but six of these genes were down-regulated in the heat-stressed samples with fold changes of up to 3.8 (Table 2). Information was available regarding the functions of 55 of the 111 differentially expressed genes (Table 2). While many functional classes of genes were affected by heat-stress, many of these candidate genes are involved in protein processing (15%), cell cycle and cell growth (15%), and transcription and translation (8%) (Table 2, Figure 3). Differentially expressed genes in this group included several ribosomal proteins (28S ribosomal protein S15, 40S ribosomal protein S17, 40S ribosomal protein S5), the transcription factors *cebpd*, *ercc8*, and *nkrf*, the growth factors *cspg5* and *FGF1*, and *ANAPC4*, a gene involved in cytokinesis (Table 2). Most of these genes were down-regulated following heat stress, except 40S ribosomal proteins S5 and S17 and carboxypeptidase A2, which were up-regulated. A further 11% of identified candidate genes are cytoskeletal proteins or are involved in their processing. Several gene ontologies were over-represented in the identified candidate genes as compared to the total of genes represented on the microarray (Table 3). In particular, 'macromolecule metabolism' and 'protein processing' were over-represented in the category Biological Process, 'structural molecule activity' and 'motor activity' in the category Molecular Function, and 'ribosome', 'actin cytoskeleton', and 'actin filament' in the category Cellular Component. The gene ontologies of five candidate genes, *rhoC*, α-cardiac actin, *cebpd*, *cct6a* and β-actin, included the response to stress.

Table 2 Changes in mRNA expression levels in heat-stressed *Pomacentrus moluccensis* (only genes with information regarding gene function are shown; where multiple functions are associated with a gene the gene function most relevant in the context of this study is represented). GenBank accession numbers refer to the *Danio rerio* clones represented on the microarray. Genes were ranked according to statistical significance as determined by Bayesian analysis of the expression response across four biological replicates. Negative values of fold change indicate down-regulation of gene in heat-stressed *P. moluccensis*, while positive values indicate up-regulation (p-values are FDR-corrected). Real-time PCR data for six candidate genes are provided for comparison to the array data (- indicates loci for which real-time PCR was not performed).

Rank	GenBank Accession *D.rerio* clone	UniGene ID	Gene Symbol	Putative Identification	Functional Classification	Fold Change Array	p	Fold Change qPCR
Protein Processing								
5	AW344170	Dr.7695	gp25L2	Glycoprotein 25L2 precursor	Protein Carrier	-2.55	0.027	-
7	AI957736	Dr.33713	wu:fd02h12	Aurora-like serine/threonine kinase	Cytokinesis, Protein Kinase	-2.62	0.027	-
8	BI880263	Dr.4218	ubiE	Ubiquinone methlytransferase	Ubiquinone Biosynthesis	-2.58	0.027	-
33	AF203736	Dr.10688	cdk5	Cyclin-dependent protein kinase 5	Protein Kinase	-2.19	0.044	-1.92
42	AW826222	Dr.23208	zgc:64014	zgc:64014	Protein Binding	-1.82	0.046	-
83	AW019487	Dr.31372	Ela2	Elastase 2	Proteolysis and Peptidolysis	-2.69	0.078	-
101	AI641408	Dr.1576	si:ch211-240l19.1	si:ch211-240l19.1	ATP Binding, Protein Kinase	-1.95	0.091	-
108	BI887176	Dr.20448	kpnb1	Importin beta-1 subunit	Protein Binding	-1.63	0.095	-
Cell Cycle and Cell Growth								
12	BM182231	Dr.9164	ANAPC4	Anaphase promoting complex subunit 4	Cytokinesis	-2.16	0.027	-
24	BI710602	Dr.41827	cspg5	Chondroitin sulfate proteoglycan 5	Growth Factor	-2.04	0.029	-
26	AW173992	Dr.669	sycp3	Cholinephosphotransferase 1	Cell Growth	-1.93	0.035	-
31	BG306448	Dr.27147	FGF1	Heparin-binding growth factor 1 precursor	Growth Factor	-1.80	0.044	-
89	AI496860	Dr.31546	CPA2	Carboxypeptidase A2	Cell Growth	+3.78	0.081	-
Cytoskeleton								
4	AF116824	Dr.10694	actcl	Alpha-cardiac actin	Cytoskeleton	-2.31	0.027	-
111	AF025305	Dr.1109	bactin2	Beta-actin	Cytoskeleton	-2.28	0.095	-3.59
Transcription and Translation								
15	AI545168	Dr.31567	hnrpab	Heterogeneous nuclear ribonucleoprotein A/B	DNA Binding	-2.93	0.027	-
22	BI709452	Dr.9814	MRPS15	28S ribosomal protein S15	Protein Biosynthesis	-2.68	0.027	-
27	AW116173	Dr.35537	ADA2B	Transcriptional adaptor 2: ADA2 beta	Transcription	-2.13	0.035	-
45	BE017827	Dr.1280	cebpd	cebpd protein/ transcription factor CC/EBP-2	Transcription Factor	-1.91	0.046	-1.5
52	BI888812	Dr.29118	RPS17	40S ribosomal protein S17	Protein Biosynthesis	+2.30	0.047	-
64	BI879671	Dr.32926	im:7151282	Retinoic acid-responsve protein	Transcription	-2.71	0.059	-
78	BM155576	Dr.36769	ercc8	Cockayne syndrome WD-repeat protein CSA	RNA Pol II Transcription Factor	-1.73	0.072	-
81	BI325685	Dr.5816	ING1	Inhibitor of growth family, member 1-like	Signal Transduction	-1.62	0.078	-
112	BE017895	Dr.1324	rps5	40S ribosomal protein S5	Protein Biosynthesis	+3.48	0.095	-
114	AI545725	Dr.4910	nkrf	NF-kappa-B-repressing factor	DNA Binding, Transcription	-1.54	0.095	-
118	BG306385	Dr.2818	zgc:91986	zgc:91986	Transcription	-1.83	0.097	-

Table 2 *cont.*

Rank	GenBank Accession *D.rerio* clone	UniGene ID	Gene Symbol	Putative Identification	Functional Classification	Fold Change Array	p	Fold Change qPCR
Cell Communication								
1	AI959074	Dr.18762	rhoC	GTP-binding gene RhoC	Signal Transduction	-3.05	0.027	-1.48
6	AF146429	Dr.8086	dlc	DeltaC	Cell Communication	-2.43	0.027	-
16	BM181148	Dr.28227	Loc402870	similar to KIAA0887 protein/ ETEA	Signal Transduction	-1.92	0.027	-1.56
20	AW174595	Dr.30368	FLJ22649	Microsomal signal peptidase 23 kDa subunit	Signal Peptidase	-1.92	0.027	-
28	AI793549	Dr.27534	Tc-mip	Truncated c-Maf-inducing protein	Signal Transduction	-1.91	0.035	-
69	BI843324	Dr.14009	GPR125	G protein-coupled receptor 125	Signal Transduction	-1.64	0.061	-
74	Y08426	Dr.114	ihhb	Indian hedgehog protein precursor (IHH)	Cell Communication	-1.91	0.071	-
100	L27585	Dr.36	shh	Sonic hedgehog protein precursor (SHH)	Cell Communication	-1.84	0.091	-
Carbohydrate Metabolism								
3	BI705511	Dr.9278	HK1	Hexokinase 1	Glycolysis	-2.45	0.027	-
35	AW115515	Dr.32807	PFKFB3	Inducible 6-phosphofructo-2-kinase	Fructose Metabolism	-2.53	0.046	-
110	BM183964	Dr.10029	glo1	Glyoxalase 1	Carbohydrate Metabolism	-1.96	0.095	-
Response to Stress								
62	AI437239	Dr.6928	cct6a	chaperonin containing TCP1, subunit 6A	Molecular Chaperone	+1.88	0.058	+1.33
Others								
18	AI957907	Dr.11635	COX15	Cytochrome c oxidase assembly protein	Electron Transport	-2.14	0.027	-
19	AW133635	Dr.1943	abce1	ATP-binding cassette, sub-family E, member 1	Electron Transport	-2.19	0.027	-
34	BG891929	Dr.1183	slco3a1	Solute carrier anion transporter, member 3A1	Ion Transport	-2.60	0.046	-
48	BI980125	Dr.30642	zgc:85626	similar to LOC394884 protein	Oxidoreductase	-1.87	0.046	-
53	AI477980	Dr.31384	vg1	Vitellogenin 1; 28kDa-1e apolipoprotein	Lipid Transport	-2.99	0.047	-
59	BE202218	Dr.23559	gst-13	Glutathione S-transferase subunit 13	Conjugation of Glutathione	-1.78	0.054	-
61	BI888327	Dr.28391	zdhhc24	zinc finger, DHHC-type containing 24	Zinc Ion Binding	-1.88	0.058	-
87	BM025184	Dr.16893	zgc:77412	Zgc:77412 protein	Phosphate Transport	-2.36	0.079	-
88	BI891136	Dr.322	wu:fa10g06	Zgc:77804 protein	Spliceosome Assembly	-1.64	0.079	-
95	BM026665	Dr.17063	ND6	NADH dehydrogenase I chain J	Electron Transport	-2.10	0.089	-
103	BG302556	Dr.26431	SCDR9	Short chain dehydrogenase reductase 9	Oxidoreductase	-1.75	0.091	-
106	BM071886	Dr.9766	cox5a	Cytochrome c oxidase subunit Va	Electron Transport	+1.98	0.094	-
107	BE200900	Dr.9792	zgc:66385	zgc:66385	Intracellular Transport	-1.96	0.095	-
109	AI958567	Dr.441	col9a2	Type IX collagen alpha 2	Phosphate Transport	-2.11	0.095	-
113	AI415849	Dr.2389	si:ch211-101n13.1	Asparaginase	Glycoprotein Catabolism	-1.59	0.095	-

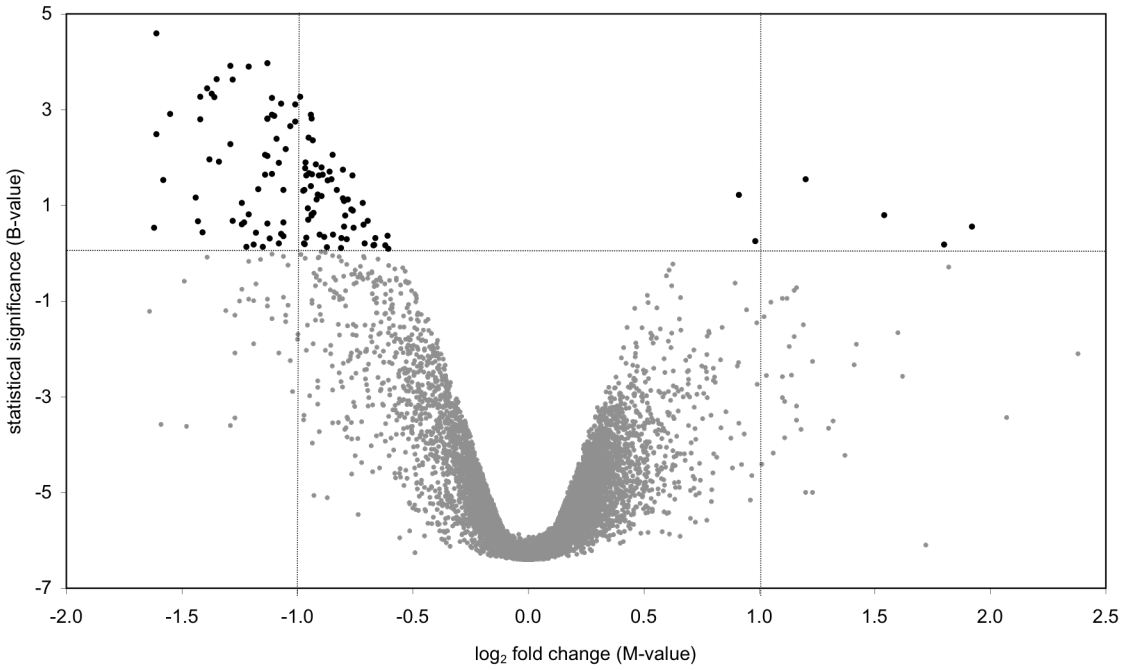


Figure 2 Volcano plot of Log odds (B-value, a measure of probability for differential expression) against relative expression differences between heat-stressed *Pomacentrus moluccensis* and *P. moluccensis* kept at ambient temperature. Each dot represents one of the 15,806 genes represented on the microarray after having excluded spots that showed poor hybridisation for *Pomacentrus moluccensis*. The X-axis displays log_2-transformed signal intensity differences between heat-stressed and non-stressed *P. moluccensis*; the Y-axis is the Log odds B for differential expression between heat-stressed and non-stressed *P. moluccensis*. The horizontal dashed line represents significance threshold corresponding to B_{min} with FDR-corrected p-value <0.1. The vertical dashed lines represent 2fold expression changes. All spots above the horizontal dashed line are genes that were identified as showing differential expression following heat stress in *P. moluccensis* (●).

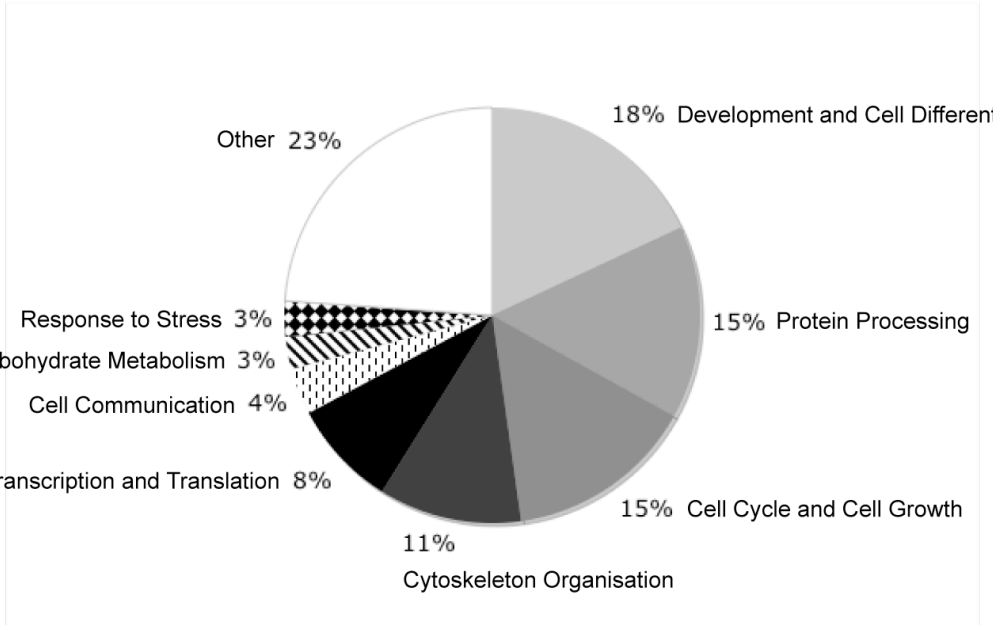


Figure 3 Gene ontologies of the 111 candidate genes identified in the microarray analysis of heat-stressed *Pomacentrus moluccensis*.

Table 3 Over-representation of gene ontologies amongst candidate genes for heat shock in *Pomacentrus moluccensis* as compared to the total of genes represented on the microarray.

Category	Gene ontology term	ID	Count	Percent candidate genes	Percent micro-array	E value
Biological Process	Macromolecule metabolism	GO:0043170	14	16.67	10.2	0.04
	Cellular protein catabolism	GO:0044257	4	4.76	1.56	0.04
	Protein processing	GO:0016485	3	3.57	0.07	0.00002
	Intein-mediated protein splicing	GO:0016539	2	2.38	0.02	0.00009
Molecular Function	Structural molecule activity	GO:0005198	5	5.95	2.06	0.03
	Peptidase activity	GO:0008233	4	4.76	1.57	0.04
	Structural constituent of ribosome	GO:0003735	3	3.57	0.91	0.04
	Structural constituent of cytoskeleton	GO:0005200	2	2.38	0.1	0.003
	Motor activity	GO:0003774	2	2.38	0.36	0.04
Cellular Component	Ribosome	GO:0005840	3	3.57	0.84	0.03
	Peroxisome	GO:0005777	2	2.38	0.09	0.003
	Actin cytoskeleton	GO:0015629	2	2.38	0.34	0.03
	Microbody	GO:0042579	2	2.38	0.09	0.003
	Actin filament	GO:0005884	2	2.38	0.1	0.003

Verification of microarray data using quantitative real-time PCR

The real-time PCR results supported the results from the microarray analyses. Four of the five genes that were down-regulated according to the microarray analyses were also down-regulated according to the real-time PCR results, and *cct6a*, which was up-regulated in the microarray analyses showed the same trend in the real-time PCR results (Table 2, Figure 4). While real-time PCR confirmed the trend for down-regulation at locus *cebpd*, the real-time PCR results were not significant at the α-level of 0.05, which was most likely due to difficulties in precisely estimating group means at this locus.

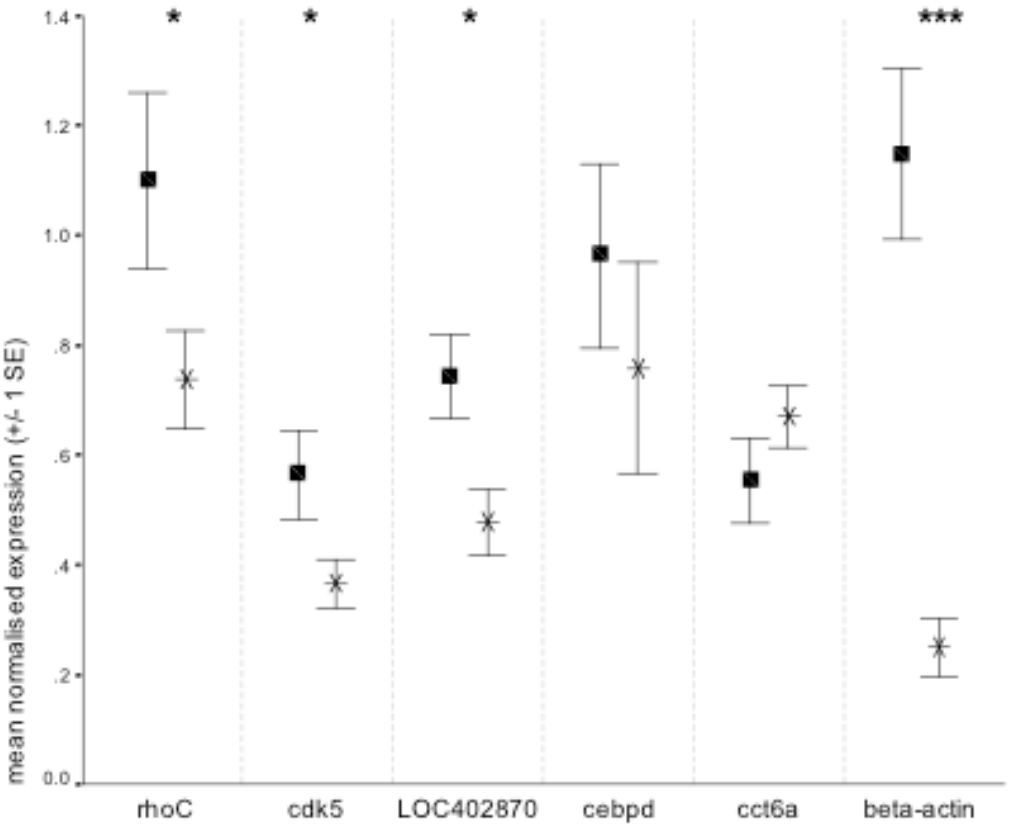


Figure 4 Normalised gene expression level of candidate genes in heat-stressed *Pomacentrus moluccensis* (x, n=6) and *P. moluccensis* kept at ambient temperature (■, n=8) as determined by quantitative real-time PCR. Expression levels were normalised to the geometric mean of the two housekeeping genes 18S and HPRT. Significance values between group means were determined using the non-parametric Mann Whitney U test (* $p < 0.05$; *** $p < 0.001$).

Discussion

The results of this study confirm the utility of using a heterologous *D. rerio* oligonucleotide array for gene discovery studies in the coral reef fish *P. moluccensis*. For a cross-species microarray platform to be useful, sequence identity needs to be great enough to allow cross-hybridisation of the heterologous sample to the array. I assessed sequence identity between *P. moluccensis* and *D. rerio* and hybridisation potential of *P. moluccensis* to the *D. rerio* array in a number of ways and also validated results from the microarray study using an alternative method for measuring differential expression. Firstly, direct sequence comparison between *P. moluccensis* and *D. rerio* at a selection of nuclear gene loci revealed an average sequence identity of 81% and, as expected, these selected gene loci showed similar fluorescent intensities for both taxa in the comparative genomic hybridisation experiment (Figure 1). These data suggest that the hybridisation of DNA to oligonucleotide microarrays is robust to some sequence variation between target and oligo. These results are in accordance with Kane *et al.* (2000) who have demonstrated that 75% sequence identity between target and oligo is sufficient to produce significant hybridisation signal on oligonucleotide microarrays. Secondly, the comparative genomic hybridisation experiment demonstrated that for most genes hybridisation of *P. moluccensis* gDNA to the *D. rerio* array was comparable to the hybridisation observed for *D. rerio* gDNA. Thus, the two fish taxa appear to share significant sequence identity at most gene loci suggesting that the heterologous microarray is useful for measuring gene regulation in *P. moluccensis*. Lastly, I employed quantitative real-time PCR to verify differential expression following heat stress for a selection of candidate genes. Four of the six candidate genes that were followed-up also showed significant expression changes in the real-time PCR assay. While the remaining two loci showed the same direction of expression change (down- or up-regulation) as in the microarray data, the results were not statistically significant at α-level 0.05, due to the relatively large variability observed between biological replicates and the small fold changes measured. The sequence alignments for eight nuclear gene loci further show that the majority of nucleotide changes (76.3%, Table 1) occur at the third codon position, and thus would not affect amino acid sequences. The high degree of amino acid sequence identity (average of 91.3% and up to 99%, Table 1) suggests that homologous genes in *P. moluccensis* and *D.*

rerio largely function in the same way, potentially allowing inference of gene function in *P. moluccensis* from known gene function in *D. rerio*. In summary, the results of this study indicate that a heterologous microarray approach is useful for measuring gene expression responses in a species with few genomic resources, the coral reef fish *P. moluccensis*.

Interpreting results from heterologous microarray experiments

The level of cross-hybridisation to a heterologous array depends on sequence similarity and therefore, on the phylogenetic distance between the probe and target species. With increasing sequence divergence the number of cross-hybridising spots and the power to detect small fold-changes in gene expression decreases (Renn *et al*. 2004). However, significant and biologically meaningful results can still be obtained using heterologous microarrays, even ones developed for distantly related species (Renn *et al*. 2004). This study is the first to use a 65mers-oligonucleotide microarray in heterologous microarray experiments and the results suggest that long oligonucleotide microarrays also have potential use in heterologous experiments.

Signal specificity may be affected in heterologous microarray experiments. In such experiments, the number of 'yellow spots' that indicate constant gene expression across treatment groups might be expected to be greater than in species-specific hybridisations due to non-specific binding. This will be true especially if the stringency of the hybridisation conditions is reduced in order to encourage binding of the sample to the array. Therefore, the biological significance of genes that are not differentially expressed is hard to gauge in heterologous microarray experiments. Here, I limit the discussion to those genes that were differentially expressed in the early response to heat stress. I cannot, however, dismiss the possibility that genes other than those discussed here may have been regulated. In contrast to the possibility of not detecting genes that are regulated, false positives in candidate gene identification due to cross-species hybridisations are less likely. False positives require considerable sequence similarity between a probe and an unrelated target. Such similarity is unlikely to occur, except for closely related members of a gene family. Cross-reactivity between gene family members increases as their sequence similarity increases and this also applies to species-specific microarray

hybridisations (Evertsz *et al*. 2001, Miller *et al*. 2002). The susceptibility to cross-hybridisation also depends on the design of oligonucleotides on the microarray. Oligonucleotide arrays of the type used here are thought to provide increased specificity compared to cDNA microarrays (Cossins and Crawford 2005). However, since the purpose of the heterologous microarray experiments presented here was to identify the participation of genes in the early response to heat stress, false negatives and cross-hybridisation between closely related gene family members constitute a minor problem. Therefore, the genes I report here are likely to be an underestimate of all the genes involved in the early response to heat stress.

The early gene response to heat stress

While quantitative PCR confirmed the results of the microarray gene expression analyses for a selection of identified candidate genes, the limited number of biological replicates (n = 4) employed in the microarray analyses in this study calls for cautious interpretation of the observed expression changes following heat shock. It is possible that the transcriptional responses observed here are not representative of the responses of the coral reef fish population as a whole. Nevertheless, genes that showed significant expression changes in this study constitute good candidates for heat stress and the transcriptional profiles discussed here may guide future work.

In heat-shocked *P. moluccensis,* the greatest expression changes occurred at *CPA2* (3.8-fold up-regulation), *rps5* (3.5-fold up-regulation), and *rhoC* (3.1-fold down-regulation). *CPA2* has been previously associated with cell growth, metal ion binding, and vacuolar protein catabolism (Pascual *et al*. 1989). Up-regulation of *CPA2* during heat stress may thus be related to increased rates of protein catabolism under conditions of heat stress. Increased protein catabolism appears to be a key response in the early response to heat shock in this study (Table 3). The ribosomal protein *rps5* is a component of the 40S subunit of ribosomes and thus associated with protein biosynthesis. Typically, heat shock causes the suppression of ribosomal protein synthesis (Bell *et al*. 1988). However, transcription of some ribosomal proteins, e.g. 40S ribosomal protein 8A, increases in response to heat stress (Podrabsky and Somero 2004). It is possible that some ribosomal proteins are

particularly sensitive to heat and that rates of transcription reflect the different heat sensitivities of ribosomal proteins.

The GTPase *rhoC* is one of the four differentially expressed genes that encode cytoskeletal proteins or are involved in their processing. These genes include α-cardiac actin, β-actin, *rhoC*, and *cct6a*. As a result, the gene ontologies 'structural molecule activity', 'structural constituent of cytoskeleton', 'motor activity', 'actin cytoskeleton', and 'actin filament' were over-represented in the group of candidate genes. The GTPase *rhoC* is involved in the organisation and biogenesis of the actin filament, while the molecular chaperone *cct6a* is important for the folding of actin. Cytoskeletal proteins have been previously shown to be regulated in response to temperature stress in teleosts (Sarmiento *et al*. 2000, Ju *et al*. 2002, Podrabsky and Somero 2004), possibly because the cytoskeleton needs to be stabilised during temperature stress (Podrabsky and Somero 2004). The expression pattern of β-actin during temperature stress and acclimation is complex. For example, summer-acclimatised carp have increased β-actin levels as compared to winter-acclimatised carp (Sarmiento *et al*. 2000). However, when channel catfish are transferred to cold, β-actin levels are similarly induced and return to baseline levels only six weeks after acclimation to the cold environment (Ju *et al*. 2002).

The molecular chaperone *cct6a*, which was up-regulated following heat shock in this study, belongs to the cytosolic chaperonin-containing t-complex polypeptide 1. It assists in the proper folding of many cellular proteins, including actin and tubulin, and consists of multiple subunits with each subunit exhibiting a unique expression pattern (Kubota 2002). Other subunits can also be affected by temperature stress in teleosts (Gracey *et al*. 2004, Podrabsky and Somero 2004); e.g. *cct7* is induced during heat stress and repressed during cold stress (Podrabsky and Somero 2004), while *cct5* is induced during cold stress (Gracey *et al*. 2004). Contrary to expectation, no other molecular chaperones, such as heat shock proteins, were amongst the differentially expressed genes. In this study and based on the values of the moderated t-statistic, the highest ranked heat shock proteins were *hsp47* and *hsp4* (FDR-corrected p-values of 0.164 and 0.278, respectively). The comparative genomic hybridisation experiment showed no evidence of reduced hybridisation potential at these loci suggesting that the array is useful to measure their regulation

during heat stress. Both loci showed small, 1.4- and 1.6-fold, changes in heat-stressed fish and there was comparatively large individual variation in gene expression amongst biological replicates (data not shown). It is possible that the small magnitude of expression change and the high variability amongst replicates prevented them from being classified as differential expressed or that these loci are not regulated in the early response to heat stress in this species. Real-time PCR may be used to follow-up on individual genes of interest.

New candidate genes identified in this study and which are not typically associated with heat shock responses include the hedgehog proteins *shh* and *ihhb*. Hedgehog proteins have known functions in cell communication and development, but their role in heat shock responses is yet to be determined. This study further identified a number of heat-responsive genes with unknown function, for example zgc:64014, si:ch211-240l19.1, zgc:91986 (Table 2). Their homology to genes from other organisms has not yet been determined. Fifty-six of heat-responsive genes in this study are unidentified (Appendix 4). It is likely that this group includes many novel genes and their identification and annotation will benefit from genome sequence data that is accumulating for other teleosts and vertebrates.

Stress responses in teleosts commonly result in gene expression changes at a large number of loci associated with protein processing, transcription and translation (Gracey *et al*. 2001, Williams *et al*. 2003, Gracey *et al*. 2004, Podrabsky and Somero 2004, Krasnov *et al*. 2005). Cold stress appears to induce transcription and protein translation (Gracey *et al*. 2004). This response may reflect compensation for decreased enzymatic rates resulting from low temperatures and increased protein synthesis in order to maintain biochemical function. In contrast, other stressors, including heat stress, commonly repress transcription and protein synthesis, probably reflecting the suppression of non-critical activities during stress. These responses, however, can vary among genes. Some translation elongation factors and ribosomal proteins are induced following chronic heat stress (Podrabsky and Somero 2004). In this study, the ontology of 15% of the differentially expressed genes included protein processing and another 8% translation and transcription. Over-representation of the gene ontology ‘cellular protein catabolism’ amongst candidate genes in this study suggests increased protein breakdown following the onset of heat stress. Most

differentially expressed genes were down-regulated in this study indicating a repression of transcriptional activity. Down-regulation was also observed for genes involved in cell cycle and cell growth, such as *CPA2*, *sycp3*, and the growth factors *cspg5* and *FGF1*. This gene expression profile is likely to reflect suppressed cell growth during stress. Such a response is consistent with gene expression responses reported for other species experiencing different types of stress. The yeast *Saccharomyces cerevisiae* responds to a range of environmental stressors with a stereotypical change in gene expression and the suppression of around 10% of assayed genes, many of which are involved in protein synthesis and cell growth (Gasch *et al*. 2000). In the goby *Gillichthys mirabilis*, hypoxia causes down-regulation of many cytoskeletal and ribosomal proteins and this transcription profile may reflect the reorganisation of metabolism and the suppression of major energy-requiring processes shortly after the onset of stress (Gracey *et al*. 2001). The gene expression response of *P. moluccensis* to heat stress observed in this study indicates suppressed cell growth, the repression of transcriptional activity and increased protein breakdown. These responses are consistent with an interpretation of metabolic reorganisation following the onset of stress, which would precede the induction of genes and *de novo* synthesis of proteins, which may ultimately allow the organism to cope with prolonged exposure to stress. Future studies using longer-term exposure to heat stress are needed to reveal those genes involved in stress responses beyond the initial genomic responses observed here.

Conclusion

Coral reef fishes are predicted to experience a significant rise in sea surface temperatures within the next decades. It is likely that different species of coral reef fishes differ at loci that will be important for thermal stress resistance in the future, and that this will result in varying success among species in coping with the predicted climatic change. In this study, I have shown that transcriptome profiling of the response to heat stress using a heterologous microarray approach is useful for identifying genes associated with heat stress resistance and their regulation in coral reef fishes.

CHAPTER 3

The transcriptional responses of a coral reef fish to stress: conservation of gene function responses but variable gene responses

Publication: Kassahn KS, Caley MJ, Ward AC, Stone G, Crozier RH (*In prep*) The transcriptional responses of a coral reef fish to stress: conservation of gene function responses but variable gene responses.

Abstract

Global warming is predicted to lead to an increase in sea surface temperatures in the near future. While recently considerable knowledge has been gained regarding the potential impacts of climate change on corals and the incidence of coral bleaching events, we currently know little about how increased temperatures may affect the physiology of coral reef fishes. In this study, I used microarrays to measure liver gene expression responses of the coral reef fish *Pomacentrus moluccensis* to heat over a period of five days. 324 gene loci were differentially expressed in response to heat, most of which were induced. Genes with functions in protein folding, lipid and carbohydrate metabolism, immune responses, and the response to oxidative stress were significantly over-represented amongst heat-responsive genes. The results of this study suggest that prolonged heat stress leads to oxidative stress and protein damage, a challenge of the immune system, and a re-allocation of energy sources, potentially reducing the available energy for other cellular and organismal functions, such as growth and reproduction. In order to elucidate early gene responses and upstream regulators of transcriptional stress responses, I further exposed *P. moluccensis* to hypoxic, hyposmotic, cold and heat shock over three hours. Early gene responses three hours after exposure were generally associated with a suppression of transcriptional activity, but the responses of individual genes varied depending on the type of stressor applied. Only a few genes showed consistent repression or induction across stress treatments. While individual genes showed

variable expression responses across stress treatments, a series of gene functions were consistently involved in all stress responses examined here, suggesting that different types of stress have common effects on biological function.

Introduction

The world's coral reefs are facing many anthropogenic threats, including over-fishing, pollution, run-off from agricultural activity, and increased water turbidity and sedimentation due to land clearing (Hughes 1994, Gardner *et al*. 2003, Pandolfi *et al*. 2003). Furthermore, current models of climate change predict a significant rise in sea surface temperatures within the next decades (IPCC 2001). The predicted rise in sea surface temperatures may lead to increased levels of physiological stress and mortality in coral reef organisms. For example, elevated sea surface temperatures a few degrees above normal can lead to coral bleaching (Hughes *et al*. 2003). While we have recently gained considerable knowledge about the relationship between sea surface temperatures and the incidence of coral bleaching, we presently have only limited data to estimate the potential effects of elevated temperatures on the physiology of coral reef fishes. We therefore need to gain a better understanding of how heat alters biological function in coral reef fishes.

In teleost fishes, thermal stress can lead to changes in ventilation and circulation rates (Farrell 2002), changes in mitochondrial densities and their properties (St-Pierre *et al*. 1998, Pörtner 2002), and a reduction in cellular oxygen levels (Mark *et al*. 2002, Lannig *et al*. 2004). Reduced cellular oxygen levels are associated with increased levels of oxidative stress and hence, the cellular response to thermal stress generally includes responses aimed at alleviating oxidative stress (Pörtner 2002, Heise *et al*. 2006a). Antioxidant enzymes, such as superoxide dismutase, catalase, and glutathione peroxidase are commonly activated during thermal stress (Parihar *et al*. 1997, Abele and Puntarulo 2004). Oxidative stress and the cellular damage associated with oxidative stress can further induce a heat shock response, which is primarily aimed at the molecular repair of protein damage (Dietz 1994, Airaksinen *et al*. 1998, Iwama *et al*. 1998, Airaksinen *et al*. 2003). In addition, thermal stress has been shown to lead to extensive changes in gene expression (Ju *et al*. 2002, Gracey *et al*. 2004, Podrabsky and Somero 2004, Cossins *et al*. 2006). These transcriptional

responses are likely the result of the stress-dependent activation of only a limited number of upstream regulators, namely the activation of immediate early genes (Morgan and Curran 1995, Gius *et al*. 2004) and mitogen-activated protein kinases (MAPK), in particular the JNK and p38 signalling pathways (Cohen 1997, Gabai and Sherman 2002, Cowan and Storey 2003). It is still not understood, however, how activation of these same pathways under different stress conditions results in the different transcriptional profiles that are observed downstream (Cowan and Storey 2003). These complex transcriptional responses to stress necessarily precede adjustments at the protein level, and thus, are expected to form an important component of the cellular response to stress. However, at present, we only have a limited understanding of the extent and generality of transcriptional responses to thermal stress in teleosts.

While there have been some studies on the transcriptional responses to thermal stress in temperate teleost fishes (Ju *et al*. 2002, Gracey *et al*. 2004, Podrabsky and Somero 2004, Cossins *et al*. 2006), these responses have not yet been investigated in coral reef fishes. Coral reef fishes inhabit relatively stable thermal environments with little diurnal and seasonal variations in temperature. Organisms inhabiting stable thermal environments have already been shown to have increased sensitivity towards thermal abnormalities (Hofmann *et al*. 2000, Pörtner 2002). For these reasons, one may expect that coral reef fishes have evolved different responses to thermal stress than fishes inhabiting environments that are naturally more variable. Transcriptional responses of temperate teleost fishes to thermal stress may, therefore, not be representative of the responses of coral reef fishes. Understanding thermal stress responses in coral reef fishes may, thereby, offer fundamental insights into mechanisms of environmental adaptations and shed light on the expected sensitivity of coral reef fishes to elevated sea surface temperatures in the future.

In this study, therefore, I measured short- and medium-term transcriptional responses of the coral reef fish *Pomacentrus moluccensis* to elevated temperatures. In Chapter 2 of this thesis, I have demonstrated the utility of a heterologous microarray for studying stress responses in *P. moluccensis*. Here, I use a heterologous microarray approach and gene expression profiles to infer the effects of elevated temperatures on biological function in coral reef fishes. In addition, I studied early gene responses to

stress and the upstream regulation of transcriptional stress responses. For this purpose, I exposed *P. moluccensis* over three hours to a variety of stress conditions including heat, cold, hyposmotic and hypoxic shock. I tested for the presence of a general stress response in this species. In the presence of a general stress response, one would expect strong similarities between the early gene responses to different stressors and induction of a common set of transcriptional regulators. In contrast, a specific stress response would be indicated by the induction of largely independent sets of genes depending on the type of stressor applied. I discuss components of a common stress response in *P. moluccensis* and also point out main differences in the responses to different types of stress. Last but not least, I discuss early gene responses in the light of common upstream regulators of transcriptional stress responses.

Materials and Methods

Stress experiments

Summer-acclimated adult *P. moluccensis* were collected around Lizard Island, northern Great Barrier Reef, Australia (14º40'S, 145º28'E) by divers on SCUBA using barrier and scoop nets and transferred to the Lizard Island Research Station. Fish were housed in groups of up to 20 individuals in aquaria supplied with flow-through seawater at ambient temperature (approximately 28ºC). Sections of PVC pipes and dead branching coral were provided for shelter. After two days of acclimation to these conditions, animals were transferred to aquaria of 31ºC for five days or remained at ambient temperature (28ºC) for five days. These treatments were aimed to target responses beyond the primary responses measured in Chapter 2. To elucidate the upstream regulators of transcriptional stress responses, I exposed additional *P. moluccensis* individuals to severe heat shock (34ºC), moderate heat shock (31ºC), cold shock (22ºC), hypoxia (23-36% air saturation), or hyposmotic conditions (20ppt salinity) for three hours. Hypoxic conditions were created by applying a constant flow of nitrogen to the tanks. Hyposmotic conditions were created by mixing seawater with distilled water, reducing the salinity of the seawater from 36ppt to 20ppt. This treatment was expected to constitute a significant osmotic challenge and be outside the normal range of naturally occurring salinity changes for

this species. On any one day, an equal number of fish were subjected to either one of the stress treatments or transferred to aquaria at ambient conditions, in order to provide time-matched controls. In Chapter 2 of this thesis, I have performed a severe heat shock treatment at 34ºC with winter-acclimated *P. moluccensis*. I was thus able to compare the summer responses to severe heat shock to those observed in winter where ambient conditions were two degree Celsius cooler (ambient conditions were 26ºC in winter and 28ºC in summer). Acclimation to different seasonal temperatures may be expected to affect gene responses to heat shock. Following these treatments, fish were killed by placing them on ice. Livers were excised and stored in RNAlater™ (Qiagen) for microarray analysis. Liver was chosen for the analyses of gene responses because liver is a metabolically important tissue and because a large number of genes are differentially expressed during heat shock in this tissue in *P. moluccensis* (see Chapter 2 of this thesis). Since age and gender can affect expression responses and due to the limited number of microarrays available, I restricted the analyses of gene responses to males of similar standard length assuming that this approach would target fish of largely similar age. Hence, only adult male fish of standard length 47 ± 4 mm were used in the microarray analyses.

Microarray platform

There is currently no microarray available for any species of pomacentrid fish. However, in Chapter 2 of this thesis, I have demonstrated that the Compugen 16K *D. rerio* oligonucleotide array is useful for studying gene responses in *P. moluccensis*. This array represents one of the largest teleost microarrays available containing 16,399 oligos (65-oligomers) representing 15,806 unique *D. rerio* gene clusters plus controls. This array offered, therefore, the greatest chance of detecting large numbers of genes involved in transcriptional responses to stress. The arrays were printed by the Adelaide Microarray Facility. The list of genes immobilised on the array is available at http://www.microarray.adelaide.edu.au/libraries/ microarrays.html.

Microarray analysis of stressed P. moluccensis

Total RNA from liver tissue of stressed *P. moluccensis* and *P. moluccensis* kept at ambient conditions was extracted using TRIzol® (Invitrogen) according to the

manufacturer's instructions and purified using RNeasy™ columns (Qiagen), ethanol-precipitated and subsequently resuspended in nuclease-free water. The concentration and purity of RNA was determined by spectrophotometer readings at 260 and 280nm. The integrity of the RNA was confirmed by agarose gel electrophoresis. Forty μg of total RNA were reverse transcribed and labelled using the SuperScript Plus Indirect cDNA Labeling System (Invitrogen) according to the manufacturer's instructions. Both, oligo dT and random hexamers were used for priming of the reverse transcription reaction. Prior to hybridisation, each microarray slide was immersed in distilled water at 60ºC for 5min and dried by centrifugation at 650 x g for 5min. The purified fluorescently labelled cDNA samples were mixed with 5μg human Cot-1 and 8μg poly A, dried under reduced pressure, resuspended in 14μl formamide and 14μl of 6.25 X SSC, denatured by heating to 100ºC for 3min and transferred directly to ice. Finally, 0.6μl of 10% SDS was added to each sample. The probes were applied to the arrays and incubated at 42ºC overnight in a humidified chamber. The arrays were washed in 0.5 X SSC containing 0.01% SDS for 1min, 0.5 X SSC for 3min and 0.2 X SSC for 3min. The slides were scanned using an Axon 4000B microarray scanner. Single image .tif files were saved for data analysis.

Microarray data analysis and data mining

In total, 59 microarray hybridisations representing 118 individual *P. moluccensis* were performed using dye swaps and a balanced design (Figure 1). While the aim of the study was to identify gene regulation in response to stress using expression levels of fish kept at ambient conditions as reference, direct comparisons involving two stress treatments were included to make the experimental design more robust to the potential failure of some microarray hybridisations and to facilitate comparison between stressors if later deemed important. In total, 985 of the 16,897 array spots were excluded from analysis of the microarray data because these spots showed poor cross-hybridisation between *P. moluccensis* and the *D. rerio* microarray in comparative genomic hybridisations (see Chapter 2 of this thesis). For the remainder of the spots on the array, the Cy5 and Cy3 fluorescent signal intensities were extracted using SPOT software (CSIRO Mathematical and Information Sciences, Australia). The data were background corrected using the Spot morphological close/open method. Spot weights were calculated on the basis of the number of

pixels in the spot, with spot areas between 30 and 300 pixels given full weight. The ratio of the resultant signal intensities (Cy5/Cy3) was $\log_2$-transformed. Statistical analyses were performed using the software package LIMMA (Smyth 2005), following Smyth (2004). The transformed signal intensities of each grid on the array and global signal intensity were print-tip Loess normalised (Smyth and Speed 2003), and scale normalised between arrays. A moderated t-statistic was calculated for each gene on the array using an empirical Bayes method (Smyth 2004). Benjamini and Hochberg's (1995) method for controlling the false discovery rate (FDR) was used to control experiment-wide Type I error rates in the face of multiple testing. Genes with FDR-adjusted p-values < 0.1 were categorised as differentially expressed.

Within-ambient comparisons were performed to test whether fish treated on different days showed significantly different expression patterns (Figure 1). Pooling of ambient controls is justified where expression levels are consistent across ambient controls. A linear model was, therefore, fitted to the expression data with the ambient controls from the 31°C three hours treatment group as reference. Moderated t-statistics and F-statistics measuring the overall significance of the between-control contrasts were calculated (Smyth 2004). Holm's method was used to adjust p-values for multiple testing (Holm 1979). The purpose here was to test whether there were significant differences between fish kept at ambient conditions on different days. Four genes had an adjusted p-value of less than 0.1. Three of these genes occupied neighbouring spots on the arrays. Two microarrays were outliers for these spots indicated by the residual M-value from the model. After removing these two arrays and refitting the linear model, only one spot remained significantly different amongst control groups. I decided to retain this array, but treat conclusions about this gene with caution. For the remaining analyses, ambient control samples were pooled and used as an ambient reference.

The microarray platform on which these analyses were based is deposited at the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/projects/geo/; Barrett *et al.* (2005)) under the accession GPL3365. The data series and samples including the raw image .tif files are in the process of being submitted to the GEO website, with the accession numbers available shortly.

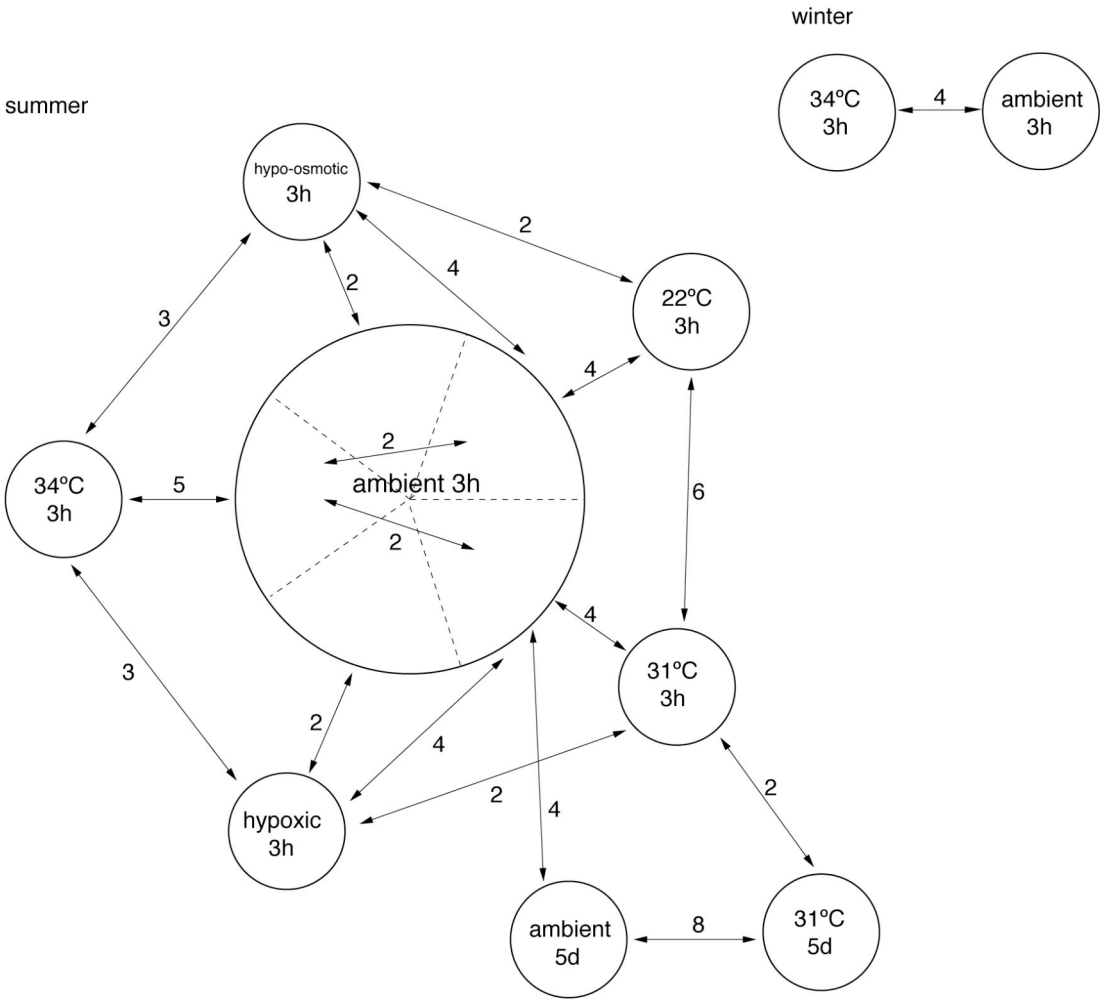


Figure 1 Experimental design employed in the microarray gene expression analysis of *Pomacentrus moluccensis* stress responses. Text within each circle names the stress and the duration it was applied, being three hours (3h) or five days (5d). Numbers above arrows indicate the number of biological replicates used. Arrowheads indicate use of Cy3 dye, which was swapped between samples to account for potential dye bias. The dotted lines within the ambient 3h group indicate five groups of fish exposed to ambient conditions on separate days. Ambient conditions were 28ºC in summer and 26ºC in winter.

Gene annotation was performed using the program Resourcerer 12.0 (Tsai *et al.* 2001). Gene function of identified candidate genes was estimated by gene functions determined for *D. rerio* using a combination of AmiGo (The Gene Ontology Consortium 2000) (http://www.godatabase.org/cgi-bin/amigo/go.cgi), iHop (http://www.pdg.cnb.uam.es/UniPub/iHOP/) (Hoffmann and Valencia 2004), and databases at NCBI (http://www.ncbi.nlm.nih.gov/).

Identification of gene functions responsive to stress

The gene expression responses to stress can be tested for the presence of common gene function responses. Knowledge of the type of gene functions commonly associated with the response to stress can shed light onto the significance of individual gene responses for maintaining biological function during stress. For this purpose, individual genes are grouped on the basis of their gene function. I then tested whether genes of a certain function were more likely to be differentially expressed in response to stress than expected if differential expression was independent of gene function. For this purpose, I used gene class testing and the gene set resampling (GSR) algorithm implemented in the software ermineJ (Lee *et al.* 2005). The aim of gene class testing is to identify sets of genes that are significantly over-represented amongst regulated genes. In contrast to other commonly used gene class testing algorithms, the gene set resampling algorithm does not require a threshold for gene selection. Instead, all genes belonging to a particular gene ontology class are used to compute a raw score $r = -\sum_i \log(p_i)$, where p_i is the p-value for differential expression for each gene in the gene ontology class. Thus, GSR employs the continuous evidence contained in the p-values for differential expression and is the method of choice for gene class testing where there is no *a priori* gene grouping information, such as in gene expression studies. Gene ontology classes of size k = 5–100 were examined. For genes that were represented multiple times on the array, the minimum p-value was used. In order to calculate the distribution of raw scores under the null hypothesis of random distribution of gene ontologies, a random set of genes of the same size as each of the gene classes of interest was drawn from the data and the raw score r was computed for the random set. I performed one million iterations of this procedure. The significance for a gene set class was calculated as the fraction of random trials resulting in a score higher than r and the resulting p-values for overall significance were FDR-corrected.

Hierarchical clustering and visualisation of gene responses across treatments

Once stress-responsive genes had been identified, I wanted to visualise their responses across stress treatments to test for the presence of a set of commonly

induced or suppressed genes. For this purpose, I extracted the expression data from differentially expressed genes across all stress treatments and organised the expression profiles using unsupervised hierarchical clustering and the software programs Cluster and TreeView (Eisen *et al*. 1998). In order to cluster genes, array- and gene-normalised values of the t-statistic and complete linkage uncentered correlation was used. The values of the t-statistic were chosen over average M-values, because the t-statistic not only reflects the magnitude of the average expression change, but also retains information on the variability in expression response across biological replicates. Visualisation of M-values representing individual arrays was not appropriate in this study because both direct and indirect comparisons were used for estimation of stress expression responses and indirect comparisons could not be unambiguously assigned to individual stress treatments.

In order to determine whether there was a relationship between cluster identity and gene function, i.e. whether co-regulated genes shared similar gene functions, I performed gene class testing on the clusters obtained by hierarchical clustering using over-representation analysis (ORA) implemented in ermineJ (Lee *et al*. 2005). The ORA algorithm is most appropriate here because the genes for this analysis naturally fell into two groups, being either in the cluster of interest or not. Gene functions of genes that were part of the cluster in question were compared to the gene functions of all genes represented on the microarray and I tested for a significant over-representation of gene functions amongst genes within a cluster.

Results

The transcriptional responses to prolonged heat

Prolonged exposure to elevated temperatures at 31°C over five days resulted in 324 differentially expressed genes (Figure 2, Table 1, Table 2). In general, the magnitude of expression changes was small, with only fifteen genes showing greater than two-fold expression changes (Table 2). Amongst the genes with the greatest expression changes were the protein kinase C substrate 80K-H *prkcsh*, rab escort protein 1, semaphorin 3ab, distal-less homeobox 3, and TTF-I interacting peptide 5, all of which were induced in response to prolonged heat (Table 1). Most regulated genes

were induced in response to prolonged heat (Figure 3). Differentially expressed genes belonged to a variety of gene functions. Based on gene ontology terms, only four percent of differentially expressed genes known stress-related gene functions, while 47 percent of identified candidate genes were of yet unknown gene function (Figure 4).

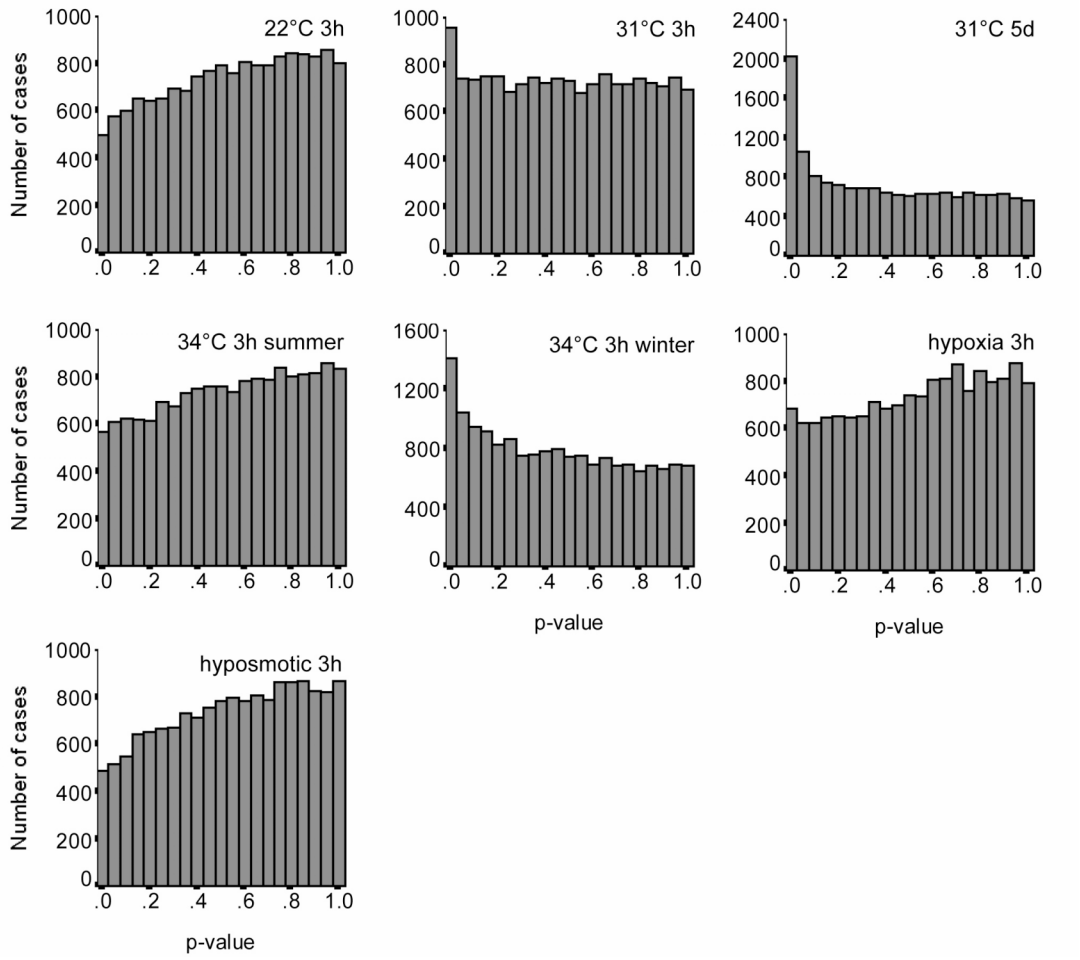


Figure 2 Summary of gene expression responses of *Pomacentrus moluccensis* to stress: Distribution of p-values for differential expression under different stress conditions.

The gene expression responses of *P. moluccensis* to prolonged heat were significantly associated with the gene functions 'regulation of translation', 'protein folding', 'carbohydrate metabolism', 'lipid metabolism', 'response to pest, pathogen or parasite', 'response to temperature stimulus', 'response to oxidative stress', and 'hemoglobin complex' (Table 3). Several gene functions related to cell growth and cytoskeleton were further associated with the expression responses to prolonged heat as well as the signal transduction pathways 'MAPKKK cascade', 'JNK cascade', and 'Rho protein signal transduction' (Table 3).

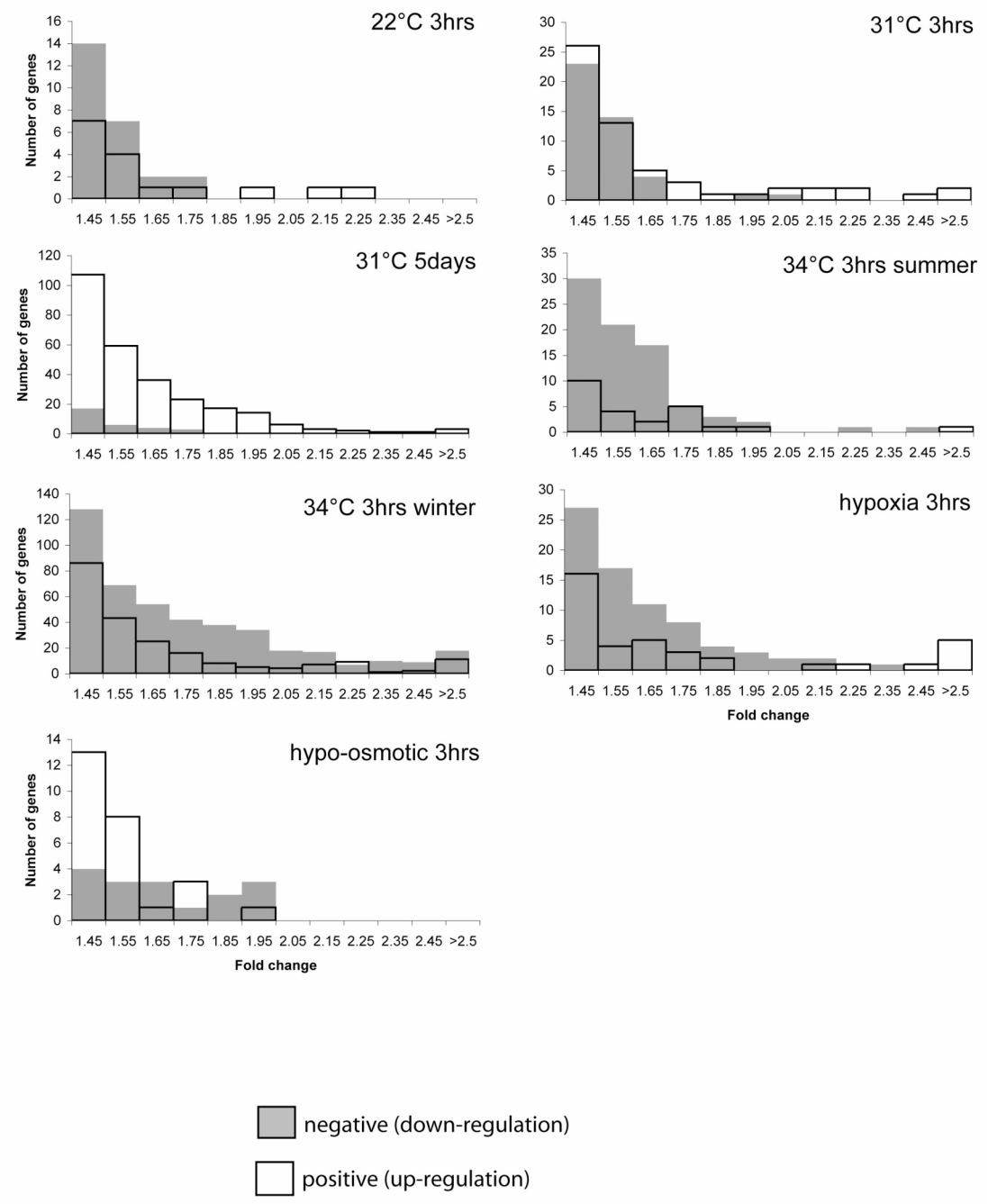


Figure 3 Summary of gene expression responses of *Pomacentrus moluccensis* to stress: Distribution of fold changes for differential expression under different conditions of stress. Grey boxes indicate down-regulation, white boxes indicate up-regulation of gene in stressed fish as compared to fish kept at ambient conditions.

Table 1 Changes in mRNA expression levels in heat-stressed *Pomacentrus moluccensis*, exposed to elevated temperatures (31ºC) for five days compared to *P. moluccensis* kept at ambient temperature (28ºC) for five days. Only genes for which information regarding gene function is currently available are reported here. Where multiple functions have been identified for a gene, the gene function most relevant in the context of this study is reported. GenBank accession numbers refer to the *Danio rerio* clones represented on the microarray. Genes were ranked according to statistical significance as determined by Bayesian analysis of the expression response across biological replicates. Negative values of fold change indicate down-regulation of gene in heat-stressed *P. moluccensis*, while positive values indicate up-regulation (p-values are FDR-corrected).

Rank	GenBank Accession *D.rerio* clone	UniGene ID	Gene Symbol	Putative Identification	Function	Fold Change	p
Cell Adhesion							
21	AF081128	Dr.28760	fn1	Fibronectin 1	Protein binding; Cell adhesion	+2.09	0.0227
45	AI641045	Dr.9212	DKEYP-18C4.2	Podocalyxin precursor	Negative regulation of cell adhesion	+1.93	0.0386
72	AF081128	Dr.28760	fn1	Fibronectin 1	Protein binding; Cell adhesion	+1.90	0.0414
81	BI882735		nrxn3	Neurexin 3-alpha	Cell adhesion	+1.28	0.0428
161	BI846718		papln	Papilin, proteoglycan-like sulfated glycoprotein	Cell adhesion; Protein binding	+1.61	0.0614
203	BI891338			Protocadherin 1 isoform 1 precursor	Cell adhesion; Calcium ion binding	+1.56	0.0737
211	AW019629	Dr.6507	eva1	Epithelial V-like antigen 1	Cell adhesion; Protein binding	+1.35	0.0752
255	AI384393	Dr.418	pcdh2g9	Protocadherin 2 gamma 9	Homophilic cell adhesion	-1.26	0.0907
Cell Cycle and Cell Growth							
26	BG308696	Dr.11045	GRIM19	Cell death-regulatory protein GRIM19	Apoptosis	+1.75	0.0227
33	BG308520	Dr.33755	TP53INP1	Tumor protein p53 inducible nuclear protein 1	Apoptosis	+1.54	0.0300
61	BI672829	Dr.13625	slah21	Seven in absentia homolog 2 (Drosophila)-like	Mitotic checkpoint	+1.61	0.0392
70	AF229449	Dr.8287	jag2	Jagged 2 isoform 1	Cell cycle; Cell communication	+1.36	0.0392
100	BI472637	Dr.15050	SEPT3	Septin 3	Cell cycle; Cytokinesis	+1.35	0.0504
144	AI959372	Dr.1212	ccng1	Cyclin G1	Cell cycle checkpoint	+1.34	0.0614
173	AW154091	Dr.33730	cdca3	Cell division cycle associated 3 isoform 1	Cell division	+1.35	0.0667
205	BI672089	Dr.14498	TXNL4B	Thioredoxin-like 4B	Mitosis	+1.99	0.0745
285	BG727086	Dr.41921	TGFB1	Transforming growth factor beta-1-binding protein	Cell Growth; Inflammatory response	+1.33	0.0960
288	BI672476	Dr.14472		Amyloid-like protein 2 precursor	Apoptosis; Cell adhesion; Heparin binding	+1.29	0.0965
307	BI672424	Dr.14459	bax	BCL2-associated X protein	Regulation of apoptosis; Cell cycle	+1.38	0.0975
Cytoskeleton							
46	BM172691	Dr.14192	DNAH11	Dynein, axonemal, heavy polypeptide 11	Microtubule motor activity	+1.53	0.0386
57	BM025904	Dr.6441	DNAH11	Dynein, axonemal, heavy polypeptide 11	Microtubule motor activity	+1.22	0.0392
126	AW116733	Dr.19450	SPTB	Spectrin	Structural constituent of cytoskeleton	+1.63	0.0586
147	BI979131	Dr.13282	DNAL4	Dynein, axonemal, light polypeptide 4	Microtubule-based movement	+1.54	0.0614
148	AI584969	Dr.28264	gas8	Growth arrest-specific 8	Cytoskeleton	+1.31	0.0614
197	BI864873	Dr.6173	tubb2	Tubulin beta-2	Structural constituent of cytoskeleton	+1.50	0.0710
231	BI889708	Dr.1964	PTPN13	Protein tyrosine phosphatase, non-receptor type 13	Cytoskeleton; Protein binding	+1.28	0.0804
239	AA658732	Dr.867	wdr1	WD repeat domain 1	Cytoskeleton	+1.36	0.0830
244	BI671622		filamin 2	Gamma Filamin	Cytoskeleton; Actin binding	+1.28	0.0887

Table 1 *cont.*

Rank	GenBank Accession *D.rerio* clone	UniGene ID	Gene Symbol	Putative Identification	Function	Fold Change	p
Metabolism							
35	BF156211	Dr.11446	ACAD	Acyl-CoA dehydrogenase	Electron transport; Metabolism	-1.25	0.0311
48	BI892323	Dr.30594	dgat2l	Diacylglycerol acyltransferase 2-like	Lipid metabolism	+1.41	0.0386
59	AA494790	Dr.396	dcxr	Dicarbonyl/L-xylulose reductase	Glucose metabolism	+1.39	0.0392
65	AW305632	Dr.36924	HADH2	3-hydroxyacyl-CoA dehydrogenase type II	Oxidoreductase activity; Lipid metabolism	+1.54	0.0392
78	BI710519	Dr.29442	crabp2	Cellular retinoic acid binding protein 2	Lipid binding	+1.22	0.0418
103	BM104068		wac	WW domain containing adaptor with coiled-coil	Glucan 1,4-alpha-glucosidase activity	+1.26	0.0527
107	AI545065	Dr.7466	ak3l1	Adenylate kinase 3-like 1	ATP binding; Adenylate kinase activity	+1.52	0.0558
118	AW281616	Dr.12080	Glb1	Beta-galactosidase	Carbohydrate metabolism	+1.30	0.0586
127	AW116000	Dr.7640	setd8	SET domain containing (lysine methyltransferase) 8	Methyltransferase activity	+1.44	0.0586
186	AI957831	Dr.15140	mettl6	Methyltransferase-like 6	Methyltransferase activity	+1.91	0.0698
188	BM103899	Dr.12654	elovl6l	ELOVL family member 6	Fatty acid elongation	+1.33	0.0698
219	BM024412	Dr.2359	nvl	Nuclear VCP-like	ATP binding	+1.53	0.0792
221	BI885898		dnmt	DNA (cytosine-5-)-methyltransferase 3	Methyltransferase activity; DNA methylation	+1.33	0.0792
234	AW280062	Dr.9165	ZDHH9	Zinc finger, DHHC-type containing 9	Metal ion binding; Transferase activity	-1.23	0.0821
250	BI882221	Dr.33925	rdh12l	Odd Oz/ten-m homolog 3; retinol dehydrogenase 12, like	Oxidoreductase activity	+1.46	0.0900
278	AW115782	Dr.6619	pgd	Phosphogluconate hydrogenase	Oxidoreductase; Pentose-phosphate shunt	-1.26	0.0959
281	AW232289	Dr.1041	fuca1	Fucosidase, alpha-L- 1, tissue	Carbohydrate metabolism	+1.47	0.0959
286	BM070575	Dr.11133	dpysl5b	Dihydropyrimidinase-like 5b	Hydrolase activity	+1.26	0.0960
300	AW232323	Dr.30175	AGPAT3	1-acylglycerol-3-phosphate O-acyltransferase 3	Acyltransferase activity	+1.48	0.0968
Protein Processing							
2	AW116649	Dr.1099	prkcsh	Protein kinase C substrate 80K-H	Protein kinase cascade	+3.87	0.0131
18	BM102551	Dr.17149	PAPPA	Pregnancy-associated plasma protein A	Proteolysis	+2.06	0.0195
36	BI672058	Dr.2642	ntf2	Nuclear transport factor 2	Protein transport	+1.68	0.0317
74	AF130460	Dr.8275	trpc4apb	Transient receptor potential cation channel	Protein transport	+1.49	0.0414
86	BG306270	Dr.15263	snrk	SNF1-related kinase	Protein kinase activity	+1.57	0.0441
91	AW170941	Dr.7993	USP29	Ubiquitin specific peptidase 29	Ubiquitin-dependent protein catabolism	-1.25	0.0464
102	BI704448	Dr.17040	chm	Rab escort protein 1; choroideremia	Intracellular protein transport	+2.31	0.0509
108	BM101665	Dr.9559	ctss	Cathepsin S	Proteolysis	-1.39	0.0563
113	BI886677	Dr.18008	RBM12	RNA binding motif protein 12	Protein binding	-1.17	0.0583
130	BI979883	Dr.31059	psmc3	Proteasome (prosome, macropain) 26S subunit, ATPase, 3	Ubiquitin-dependent protein catabolism	-1.34	0.0586
132	BI891871	Dr.3615	csnk1d	Casein kinase 1, delta	Protein serine/threonine kinase activity	+1.34	0.0587
140	AW134164	Dr.20940	fbxl12	F-box and leucine-rich repeat protein 12	Ubiquitin-dependent protein catabolism	+1.63	0.0587
145	AW019421	Dr.6291	tnpo2	Transportin 2 (importin 3, karyopherin beta 2b)	Protein transport	-1.27	0.0614
164	AI883718		ATG7	ATG7 autophagy related 7 homolog	Positive regulation of protein modification	+1.85	0.0625
175	AI588515	Dr.2974	Rabac1	Rab acceptor 1 (prenylated)	Protein binding; Golgi apparatus	+1.34	0.0667
178	BG737266		ru2	Ruby eye2-like protein	Protein binding	+1.35	0.0667
200	BG728947	Dr.12053	COG2	Component of oligomeric golgi complex 2	Protein transport	-1.18	0.0732
218	BG737844		mark1	MAP/microtubule affinity-regulating kinase 3	Protein kinase activity	+1.37	0.0789
259	BG302807	Dr.28615	snx12	Sorting nexin 12	Protein transport	-1.26	0.0914

Table 1 *cont.*

Rank	GenBank Accession *D.rerio* clone	UniGene ID	Gene Symbol	Putative Identification	Function	Fold Change	p
Protein Processing cont.							
267	BG304171	Dr.20155	ss18	Synovial sarcoma translocation, chromosome 18	Protein binding	+1.55	0.0936
279	AW058757	Dr.20362	cdc7	Cell division cycle 7-related protein kinase	Protein serine/threonine kinase activity	-1.17	0.0959
290	BI704278	Dr.26555	ctsla	Cathepsin L, a	Proteolysis and peptidolysis	+1.35	0.0965
301	BI980628	Dr.16147	osgepl1	Novel glycoprotease; O-sialoglycoprotein endopeptidase-like 1	Proteolysis and peptidolysis	+1.31	0.0968
323	BI882056	Dr.6513	cyhr1	Cysteine and histidine rich 1	Protein binding	+1.30	0.0998
Response to stress							
3	AF082662	Dr.28283	hbbe1	Hemoglobin beta embryonic-1	Oxygen transport	+1.86	0.0131
54	AW344134	Dr.9667	dnajb11	DnaJ (Hsp40) homolog, subfamily B, member 11	Heat shock protein binding; Protein folding	+1.82	0.0392
60	BI673277	Dr.2704	stambp	Associated molecule with the SH3 domain of STAM	Anti-apoptosis; Ubiquitin cycle	+1.27	0.0392
67	AA605696	Dr.13845	prdx6	Peroxiredoxin 6	Response to reactive oxygen species	+1.53	0.0392
75	AI884178	Dr.4306	clpx	Caseinolytic peptidase X homolog (E. coli)	Protein folding	+1.63	0.0414
128	AF082662	Dr.28283	hbbe1	Hemoglobin beta embryonic-1	Oxygen transport	+1.65	0.0586
224	AI964223	Dr.2970	apoea	Apolipoprotein Ea	Induction of apoptosis; Antioxidant activity	+1.31	0.0792
230	BM185394	Dr.14011	PFDN4	Prefoldin subunit 4	Protein folding	+1.29	0.0804
235	AI793830	Dr.31082	ptges3	Prostaglandin E synthase 3 (cytosolic)	Hsp90 binding; Protein folding	+1.34	0.0821
256	AI353083	Dr.1450	hbae3	Hemoglobin alpha embryonic-3	Oxygen transport	+1.90	0.0914
282	AF246176	Dr.30472		T-cell receptor alpha variable region	Immune Response; Receptor Activity	+1.37	0.0959
296	BG884044	Dr.4867	HP	Haptoglobin	Defense response; Hemoglobin binding	+1.27	0.0965
298	BI890693	Dr.13371	hif1an	Hypoxia-inducible factor 1, alpha subunit inhibitor	Oxidoreductase; Regulation of transcription	-1.23	0.0966
Signal Transduction							
17	BM183950	Dr.14422	Cdon	Cell adhesion molecule-related/down-regulated by oncogenes	Smoothened signaling pathway	+1.46	0.0195
40	BM184012	Dr.3319	CALM3	Calmodulin 3 (phosphorylase kinase, delta)	G-protein coupled receptor protein signaling	+1.23	0.0368
43	BG306148	Dr.16542	GRWD1	Glutamate-rich WD repeat containing 1	Signal transduction	+1.85	0.0386
55	AF083382	Dr.8112	sema3ab	Semaphorin 3ab	Cell-cell signaling	+2.14	0.0392
66	AF116853	Dr.8085	frzb	Frizzled-related protein	Wnt receptor signaling pathway	+1.54	0.0392
104	BF938356	Dr.11228	arl2bp	ADP-ribosylation factor-like protein 2	Small GTPase regulator activity	+1.58	0.0527
109	AW171604	Dr.8015	Cnksr2	Connector enhancer of kinase suppressor of Ras 2	Regulation of signal transduction	-1.24	0.0583
129	AW128372	Dr.7282	IQGAP2	IQ motif containing GTPase activating protein 2	Small GTPase mediated signal transduction	+1.54	0.0586
137	AW115682	Dr.7257	rasgef1b	RasGEF domain family, member 1B	Small GTPase mediated signal transduction	+1.29	0.0587
142	AW115765	Dr.9538	NCK2	NCK adaptor protein 2	Signal complex formation	-1.26	0.0609
163	AJ007742	Dr.8054	ptc2	Patched2	Hedgehog receptor activity	+1.36	0.0625
168	BG307536	Dr.9665	rhoG	Ras homolog gene family, member G	Signal Transduction	-1.33	0.0649
176	BI886464	Dr.25497	szl	Sizzled	BMP signaling pathway	+1.69	0.0667
181	BM183274	Dr.15125		Small inducible cytokine subfamily A	Signal Transduction	+1.27	0.0667
185	U49405		FZD4	Frizzled homolog 4	Frizzled signaling pathway	+1.29	0.0696
207	AW170891	Dr.35466	IGBP1	Immunoglobulin (CD79A) binding protein 1	Signal transduction	+1.43	0.0752
222	BI878477	Dr.10893	sara2	SAR1a gene homolog 2 (S. cerevisiae)	Small GTPase mediated signal transduction;	+1.31	0.0792
229	AF105152	Dr.354	rho	Rhodopsin	Rhodopsin mediated signaling	+1.33	0.0800
246	AI477343	Dr.2419	tiaf1	TGFB1-induced anti-apoptotic factor 1	I-kappaB kinase/NF-kappaB cascade	-1.18	0.0899

Table 1 *cont.*

Rank	GenBank Accession *D.rerio* clone	UniGene ID	Gene Symbol	Putative Identification	Function	Fold Change	p
Signal Transduction cont.							
251	AW078288	Dr.32732	invs	Inversin	Wnt receptor signaling pathway	+1.24	0.0900
252	AI965042	Dr.4874	rab1a	RAB1A, member RAS oncogene family	Two-component signal transduction system	+1.55	0.0900
268	BM183955	Dr.2109	rap1a	Ras-related protein RAP-1A	Signal transduction	-1.27	0.0952
297	AW116767	Dr.4451	chp	Calcium binding protein p22	Calcium-mediated signaling; Transcytosis	-1.23	0.0966
303	BE557009	Dr.28754	PRKAR1A	cAMP-dependent protein kinase type I-alpha regulatory chain	Intracellular Signaling Cascade	+1.20	0.0969
317	AW116392	Dr.6091	rhoQ	Ras homolog gene family, member Q	Signal transduction	-1.19	0.0997
Transcription							
7	BM156785	Dr.15390	FOXL1	Forkhead box protein L1	DNA-dependent regulation of transcription	+1.81	0.0164
8	BI897147	Dr.29850	runx1	Runt-related transcription factor 1	DNA-dependent regulation of transcription	-1.79	0.0164
10	BM072263	Dr.16724	znf238	BTB/POZ domain protein; Zinc finger protein 238	DNA-dependent regulation of transcription	+1.72	0.0166
13	U24225	Dr.625	Snail2	Snail2	Transcription factor activity	-1.37	0.0170
14	AW077156	Dr.27746	RBMS3	RNA binding motif, single stranded interacting protein	RNA binding	+1.64	0.0170
20	BM026839	Dr.9702	rmb19	RNA binding motif protein 19	Nucleic acid binding	+1.34	0.0218
22	BM024211	Dr.6364	ppp1r10	Protein phosphatase 1, regulatory subunit 10	Nucleic acid binding; Defense response	+1.71	0.0227
29	BI887656	Dr.356	gata2	GATA-binding protein 2	Transcription factor activity; Heme binding	-1.30	0.0238
39	BM182314	Dr.14269	tmf1	TATA element modulatory factor 1	DNA-dependent regulation of transcription	+1.31	0.0365
44	X65060		dlx3	Distal-less homeobox 3	DNA-dependent regulation of transcription	+2.53	0.0386
52	U14592	Dr.334	otx2	Orthodenticle homolog 2 (homeobox protein OTX2)	DNA-dependent regulation of transcription	+2.00	0.0392
62	AJ344448	Dr.12575	dcp1a	Decapping enzyme	Positive regulation of transcription	+1.80	0.0392
64	BI889395		CNOT6L	CCR4-NOT transcription complex, subunit 6-like	RNA processing and modification	+1.57	0.0392
77	AW018998	Dr.4008	msl-1	Male-specific lethal-1 protein	Chromatin binding; Dosage compensation	-1.28	0.0418
85	BG303824	Dr.14895	ZNF135	Zinc finger protein 135	DNA-dependent regulation of transcription	-1.27	0.0441
98	U84616	Dr.2328	elf2	E74-like factor 2 (ets domain transcription factor)	DNA-dependent regulation of transcription	+1.66	0.0475
111	BI878117	Dr.85	nucb2a	Nucleobindin 2a	DNA binding; Calcium ion binding	+1.39	0.0583
116	BI430221	Dr.37036	SLU7	Step II splicing factor SLU7	mRNA splice site selection	+1.33	0.0586
133	AI601470	Dr.35889	hcfc1	Host cell factor C1 (VP16-accessory protein)	Regulation of transcription; Cell cycle	+1.23	0.0587
141	AW170975	Dr.36935	ilf3	Interleukin enhancer binding factor 3	DNA-dependent regulation of transcription	+1.41	0.0594
155	BG306387	Dr.11261	A2bp1	Ataxin 2-binding protein 1	RNA binding	+1.74	0.0614
156	BI891001	Dr.29941	top2a	DNA topoisomerase II	DNA unwinding during replication	+1.35	0.0614
157	BI866527	Dr.13962	gins3	GINS complex subunit 3	DNA-dependent DNA replication	-1.18	0.0614
158	AW305388	Dr.10033	lsm7	LSM7 homolog, U6 small nuclear RNA associated	Nuclear mRNA splicing, via spliceosome	+1.39	0.0614
160	BM096095	Dr.361	seph	Selenoprotein H	DNA binding	+1.41	0.0614
162	BI891601	Dr.17679	taf12	TAF12 RNA polymerase II,	DNA-dependent regulation of transcription	+1.27	0.0618
169	AW116245	Dr.14278	znf131	Zinc finger protein 131	DNA-dependent regulation of transcription	+1.32	0.0655
172	BM095242	Dr.14888	PRPF8	U5 snRNP-specific protein (220kDa)	Nuclear mRNA splicing, via spliceosome	-1.35	0.0667
179	BI877633	Dr.19658	Tip5	TTF-I interacting peptide 5	DNA-dependent regulation of transcription	+2.48	0.0667
184	BI984001	Dr.24310	zcchc17	Zinc finger, CCHC domain containing 17	Nucleic acid binding; Zinc ion binding	+1.31	0.0696
190	AF071268		hoxd9a	Homeobox protein D9a	DNA-dependent regulation of transcription	+1.36	0.0698
191	AJ293862		sall1a	Danio rerio partial sall1a gene for putative spalt protein.	Nucleic acid binding; Zinc ion binding	+1.25	0.0698

Table 1 *cont.*

Rank	GenBank Accession *D.rerio* clone	UniGene ID	Gene Symbol	Putative Identification	Function	Fold Change	p
Transcription cont.							
196	BI892074	Dr.9145	H3F3B	H3 histone, family 3B	DNA binding	+1.26	0.0710
204	BI887415	Dr.4758		Proline-rich protein	Nucleic acid binding	+1.35	0.0745
208	BM156999	Dr.1992	lmx1b	LIM/homeobox protein LMX1B	Transcription factor activity	+1.29	0.0752
210	AI558282	Dr.21124	Sertad2	SERTA domain containing 2	DNA-dependent regulation of transcription	-1.23	0.0752
223	AW171228	Dr.115	tardbpl	TAR DNA binding protein, like	DNA-dependent regulation of transcription	-1.29	0.0792
237	BI890768		znf235	Zinc finger protein 93 homolog	DNA-dependent regulation of transcription	+1.34	0.0821
245	BI878611	Dr.8928	cnot8	CCR4-NOT transcription complex, subunit 8	DNA-dependent regulation of transcription	-1.31	0.0896
261	AI957820	Dr.27315	fen1	Flap structure-specific endonuclease 1	DNA repair	+1.31	0.0914
262	AF168008	Dr.618	pea3	ETS-domain transcription factor pea3	DNA-dependent regulation of transcription	+1.51	0.0914
266	AW420369	Dr.27962	scml2	Sex comb on midleg-like 2	DNA-dependent regulation of transcription	+1.47	0.0936
302	X65060		dlx3	Distal-less homeobox 3	DNA-dependent regulation of transcription	+1.26	0.0969
308	AF071240		hoxa11a	Homeobox protein A11a	DNA-dependent regulation of transcription	+1.24	0.0975
Translation							
4	BM181896	Dr.33915	rps6	40S ribosomal protein S6	Protein biosynthesis	+1.32	0.0131
32	BI889409	Dr.14812	mrpl36	Mitochondrial ribosomal protein L36	Structural constituent of ribosome	+1.43	0.0298
34	BM071666	Dr.16980	mrpl30	Mitochondrial ribosomal protein L30	Structural constituent of ribosome	+1.33	0.0302
97	AI667224	Dr.34694	EIF2C1	Eukaryotic translation initiation factor 2C 1	Protein biosynthesis	+1.60	0.0475
131	AI964245	Dr.7888	rrs1	Ribosome biogenesis regulatory protein	Ribosome biogenesis and assembly	+1.32	0.0587
276	AW232304	Dr.4626	eif2b2	Eukaryotic translation initiation factor 2B, subunit 2 beta	Protein biosynthesis	+1.43	0.0959
283	AI545168	Dr.48984	DDX3	DEAD-box protein 3	Nucleic acid binding	+1.26	0.0959
299	AW117050	Dr.2876	EIF4B	Eukaryotic translation initiation factor 4B	Protein biosynthesis	+1.35	0.0966
Transport							
79	AW076983	Dr.7965	clcn4	Chloride channel 4	Chloride transport	+1.39	0.0418
171	AI793405	Dr.4160		Collagen alpha 1	Phosphate transport; Skeletal development	+1.31	0.0667
233	AI667527	Dr.5169	abcb7	ATP-binding cassette, sub-family B, member 7	Transport; ATPase activity	+1.27	0.0815
Other							
30	BI350560	Dr.34580	CRTAC1	Cartilage acidic protein 1	Calcium ion binding	+1.51	0.0247
47	BM171804	Dr.23461	stc2	Stanniocalcin 2	Hormone activity	+1.52	0.0386
56	AW566602	Dr.36369	olfml3	Olfactomedin-like 3	Extracellular space	+1.62	0.0392
82	BG727594	Dr.12251	cmg2a	Capillary morphogenesis protein 2A	Receptor activity; Protein binding	+1.47	0.0428
96	AI877586	Dr.36948	NRN1	Neuritin	Axonogenesis	+2.19	0.0474
101	BE605771	Dr.9651	Ifitm1	Interferon induced transmembrane protein 1	Anterior/posterior pattern formation	+1.42	0.0507
106	BI882405	Dr.10079	crybb3	Crystallin, beta B3	Structural constituent of eye lens	+1.48	0.0558
206	BG985499	Dr.1940	Spna2	Spectrin alpha 2	Barbed-end actin filament capping	+1.28	0.0752
225	BI673745	Dr.14543	rtdr1	Rhabdoid tumor deletion region protein 1	Binding	+2.03	0.0792
291	AW281815	Dr.22797	LAPTM5	Lysosomal-associated protein transmembrane 5	Lysosome	+1.64	0.0965

Table 2 Summary of gene expression responses of *Pomacentrus moluccensis* to stress. Gene expression changes are compared to expression levels of *P. moluccensis* kept at ambient conditions (28°C summer, 26°C winter, 100% air saturation, 36ppt salinity). All expression responses were measured in summer, except the winter responses to severe heat shock, which have been reported in more detail in Chapter 2 of this thesis.

Treatment	Condition	Number of genes with >two-fold induction	Number of genes with >two-fold repression	Number of differentially expressed genes
Prolonged exposure to heat over five days				
Moderate heat shock	31°C	15	0	324
Short-term stress exposures over three hours				
Cold shock	22°C	2	0	0
Moderate heat shock	31°C	9	1	1
Severe heat shock (summer)	34°C	1	2	6
Severe heat shock (winter)	34°C	38	87	121
Hypoxia	23-36% air sat.	8	5	10
Hyposmotic	20ppt salinity	0	0	0

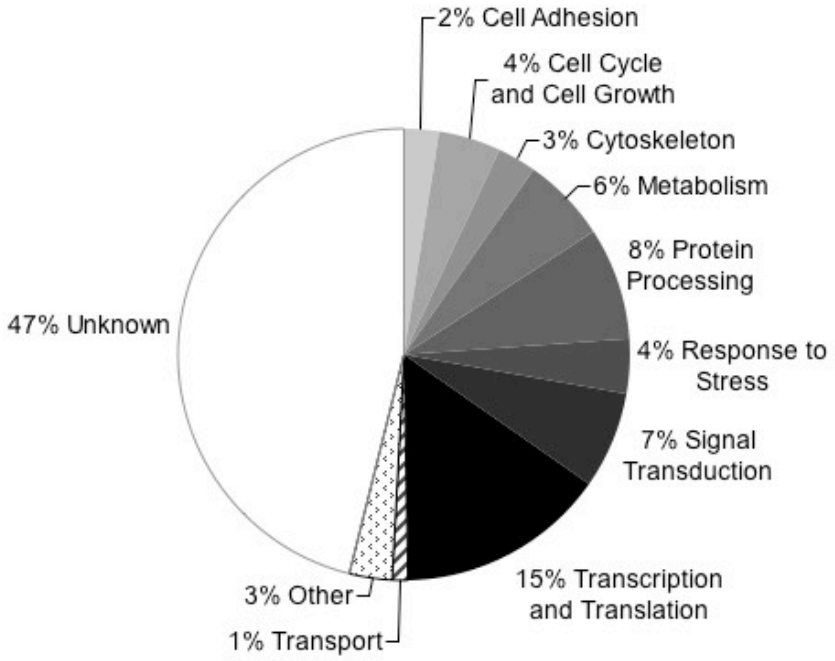


Figure 4 Gene function categories associated with the gene expression response of *Pomacentrus moluccensis* to heat exposure at 31ºC over five days.

Table 3 Results of gene set analysis of gene expression responses to stress in *Pomacentrus moluccensis*. For each gene ontology an aggregate score was computed based on the p-values for differential expression of the genes in that category. The significance of the resulting gene set score was determined by random sampling of the data using the software ermineJ. P-values are FDR-corrected with values smaller than 0.1 indicated in bold.

Gene Ontology	GO ID	Number of probes on array	Number of genes on array	p-value						
				31ºC 5d	22ºC 3h	31ºC 3h	34ºC 3h summer	34ºC 3h winter	hypoxia	hyp-osmotic
Cell growth and Cytoskeleton										
Actin cytoskeleton organization and biogenesis	GO:0030036	253	81	**0.019**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.013**
Anti-apoptosis	GO:0006916	76	72	**0.017**	**0.000**	**0.079**	**0.033**	**0.025**	**0.000**	**0.065**
Induction of apoptosis	GO:0006917	75	73	**0.001**	**0.001**	**0.034**	**0.007**	**0.002**	0.220	**0.000**
Cell growth	GO:0016049	65	63	**0.044**	**0.000**	0.454	**0.000**	**0.001**	**0.075**	0.167
Mitotic cell cycle	GO:0000278	30	30	0.106	**0.082**	0.572	0.127	**0.027**	**0.069**	0.417
Protein turnover										
Regulation of translation	GO:0006445	82	81	**0.000**	**0.000**	**0.002**	**0.005**	**0.000**	**0.001**	**0.012**
Ubiquitin-dependent protein catabolism	GO:0006511	98	96	0.200	**0.092**	**0.024**	**0.002**	**0.001**	**0.003**	**0.011**
Protein folding	GO:0006457	75	74	**0.000**	0.191	**0.005**	**0.000**	**0.002**	**0.000**	0.446
Metabolism										
Carbohydrate metabolism	GO:0005975	67	67	**0.057**	0.217	**0.076**	**0.011**	**0.000**	**0.001**	**0.088**
Lipid metabolism	GO:0006629	75	75	**0.001**	0.113	**0.007**	**0.000**	**0.000**	0.127	**0.018**
Glycolysis	GO:0006096	31	31	0.319	0.598	0.304	0.127	**0.009**	**0.086**	0.226
Gluconeogenesis	GO:0006094	21	21	0.644	0.408	0.353	0.154	**0.013**	**0.087**	0.430
Stress										
Response to pest, pathogen or parasite	GO:0009613	92	89	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.024**
Response to temperature stimulus	GO:0009266	57	52	**0.023**	0.148	**0.001**	**0.000**	**0.010**	0.163	0.225
Osmoregulation	GO:0018987	16	16	0.272	0.136	0.162	**0.035**	**0.067**	0.457	0.178
Response to oxidative stress	GO:0006979	22	22	**0.072**	0.306	0.274	0.270	0.865	0.273	0.239
Other										
DNA repair	GO:0006281	84	84	0.115	**0.040**	0.291	**0.081**	**0.000**	**0.068**	0.365
Hemoglobin complex	GO:0005833	16	11	**0.017**	0.651	0.227	**0.027**	**0.050**	0.151	0.446
Signal Transduction Pathways										
MAPKKK cascade	GO:0000165	98	97	**0.000**	**0.013**	0.267	**0.000**	**0.004**	**0.002**	**0.065**
JNK cascade	GO:0007254	37	36	**0.006**	**0.083**	0.250	**0.000**	**0.019**	0.402	0.105
Rho protein signal transduction	GO:0007266	27	27	**0.084**	0.334	0.301	**0.001**	**0.009**	0.510	**0.064**

The early gene responses to cold, heat, hypoxic, and hyposmotic stress

Only a limited number of genes showed significant expression changes in the early gene responses, three hours after exposure, to cold, heat, hypoxic, and hyposmotic shock (Figure 2, Table 2, Table 4). Most expression changes were small and regulated genes were generally suppressed in the early gene response to stress (Figure 3, Table 2). Some genes were not classified as differentially expressed despite average expression changes of greater than two-fold between control and treatment groups. This was due to these genes showing significant variation in expression response between biological replicates (data not shown). Since differential expression was assessed based on the values of the t-statistic, both, magnitude and consistency of expression change across biological replicates were taken into account.

Only two genes showed greater than two-fold expression changes during cold shock (Table 2, Table 4). Moderate heat shock at 31°C over three hours resulted in one gene being differentially expressed, *plec1*, while ten genes showed greater two-fold changes in expression level (Table 4). Severe heat shock at 34°C in summer led to significant expression changes at six gene loci, amongst which were *pdip5*, *rhoA* and *ZCCHC11*, a protein involved in DNA repair. Only three genes showed greater than two-fold expression changes in the summer response to severe heat shock at 34°C (Table 2, Table 4). The early gene response of *P. moluccensis* to heat shock at 34°C in winter has been reported in more detail in Chapter 2 of this thesis. In comparison to the summer response, the 34°C heat shock response in winter involved many more genes and the observed expression changes generally were of greater magnitude (Figure 2, Figure 3, Table 2). 121 genes were differentially expressed and 125 genes showed greater two-fold expression changes in the winter response to 34°C heat shock. The early transcriptional response to hypoxia involved ten differentially expressed genes, amongst which were *acin1a*, *β-actin*, *lbr*, and *ela2*, and a total of thirteen genes showed greater two-fold expression changes (Table 2). Hyposmotic shock resulted in no significant expression changes and none of the genes assayed showed greater than two-fold expression changes.

Table 4 Changes in mRNA levels in *Pomacentrus moluccensis* exposed to three-hour stress treatments compared to *P. moluccensis* kept at ambient conditions (28ºC, 100% air saturation). GenBank accession numbers refer to the *Danio rerio* clones represented on the microarray. Statistical significance was determined using Bayesian analysis of the expression response across biological replicates (p-values are FDR-corrected). Negative values of fold change indicate down-regulation of gene in stressed *P. moluccensis*, while positive values indicate up-regulation. Only genes with FDR-corrected p-values <0.1 and genes with greater than two-fold expression changes are reported here.

Rank	GenBank Accession *D.rerio* clone	UniGene ID	Gene Symbol	Putative Identification	Function	Fold Change	p
Cold shock (22ºC)							
1	BI533884	Dr.17570	mknk2	MAP kinase-interacting kinase 2	Response to stress; Protein kinase cascade	+2.20	0.1667
16	AI601458	Dr.17520	ube2h	Ubiquitin-conjugating enzyme E2H	Protein catabolism; Response to heat	+2.13	0.9994
Moderate heat shock (31ºC)							
1	AW134054	Dr.36134	plec1	Plectin 1	Structural constituent of cytoskeleton	+1.75	0.0015
9	AI601458	Dr.17520	ube2h	ubiquitin-conjugating enzyme E2H	Protein catabolism; Response to heat	+2.44	0.2658
72	BI710730	Dr.31497	ncl	Nucleolin	RNA binding	+2.78	0.5414
170	AW116649	Dr.1099	prkcsh	Protein kinase C substrate 80K-H	Kinase activity; Calcium ion binding	-2.05	0.5604
172	AI476925	Dr.1218	EXOSC2	Exosome component 2	Exonuclease activity; rRNA processing	+2.15	0.5604
186	AB026980	Dr.8092	odz4	odd Oz/ten-m homolog 4	MAPKKK cascade; immune response	+2.30	0.5604
315	BI673605	Dr.6875	TRIAD3	TRIAD3 protein	Protein binding; Ubiquitin cycle	+2.57	0.6326
353	BI865754	Dr.59783	hmg4	High mobility group protein 4	DNA-dependent regulation of transcription	+2.28	0.6608
464	AW826769	Dr.32001	ptk9l	Protein tyrosine kinase 9-like	Cytoskeleton; Actin binding	+2.05	0.7174
1039	AB057355	Dr.12576	ednra	Endothelin receptor type A	Signal Transduction	+2.14	0.7895
Severe heat shock (34ºC)							
1	AF387900	Dr.6509	pdip5	Protein disulfide isomerase-related protein P5	Protein folding	+4.89	0.0000
2	AI959074	Dr.18762	rhoA	Ras-like protein RhoA	Regulation of cell cycle	-1.60	0.0006
3	BG729495	Dr.31651	ZCCHC11	Zinc finger CCHC domain containing 11 isoform b	DNA replication, recombination, and repair	+1.99	0.0039
4	AW175080	Dr.8390	SC4MOL	Sterol-C4-methyl oxidase-like	Catalytic activity; Sterol biosynthesis	-1.66	0.0624
5	AW279775	Dr.3130	rgs4	Regulator of G-protein signalling 4	Signal transducer activity	-1.30	0.0624
6	BI876768	Dr.43918	znf395	Zinc finger protein 395	DNA-dependent regulation of transcription	-1.52	0.0624
197	AF246169	Dr.30471	tcra	T-cell receptor alpha variable region	T-cell receptor complex	-2.25	0.9998
845	BE557072	Dr.10066	birc6	Baculoviral IAP repeat-containing 6	Ubiquitin conjugating enzyme activity	-2.49	0.9998
Hypoxia (23-36% air saturation)							
1	AI558655	Dr.36283	acin1a	Apoptotic chromatin condensation inducer 1a	DNA binding; Apoptosis	+2.47	0.0008
2	AI522803	Dr.34133	Ela2	Elastase 2	Proteolysis and peptidolysis	+2.90	0.0011
3	AI558632	Dr.33926	ElaA	Elastase A	Proteolysis and peptidolysis	+2.99	0.0085
5	BI476356	Dr.16450	CHST2	Carbohydrate sulfotransferase 2	Inflammatory response; Sulfur metabolism	+1.32	0.0085
7	BM183382	Dr.16654	AVD	Avidin	Biotin binding	+2.25	0.0176
8	AI496860	Dr.31546	Cpa2	Carboxypeptidase A2	Proteolysis and peptidolysis	+2.84	0.0258
9	AF025305	Dr.1109	bactin2	Actin, cytoplasmic 1 (Beta-actin)	Structural constituent of cytoskeleton	+1.60	0.0647
10	BM103972	Dr.2437	amy2a	Amylase-3 protein	Carbohydrate metabolism; Hydrolase activity	+2.76	0.0651
16	AW777717	Dr.2317	NKX2-8	NK2 transcription factor related, locus 8	Transcription factor activity	-2.19	0.2928
112	AF052252	Dr.21004	fdk9	Fork head domain protein FKD9	Transcription factor activity	-2.38	0.6438
332	AB026980	Dr.8092	odz4	Odd Oz/ten-m homolog 4	MAPKKK cascade; immune response	-2.06	0.9476
578	BI710730	Dr.31497	ncl	Nucleolin	RNA binding	-2.01	0.9998

Several gene ontologies were consistently associated with early gene responses to stress, irrespective of the type of stress applied, e.g. 'actin cytoskeleton organization and biogenesis', 'anti-apoptosis', 'regulation of translation', 'ubiquitin-dependent protein catabolism', and 'response to pest, pathogen or parasite' (Table 3). 'Cell growth', 'mitotic cell cycle', and 'MAPKKK cascade' were also associated with many of the early gene responses examined here. The early gene responses to heat, hypoxic and hyposmotic shock were further associated with the gene function 'carbohydrate metabolism'. 'Lipid metabolism' was associated with expression responses to heat and hyposmotic shock, but not with those to cold and hypoxic shock, while 'gluconeogenesis' and 'glycolysis' were only associated with severe heat shock in winter and hypoxia. Genes with known function in the 'response to temperature stimulus' were associated with the responses to heat shock but were not associated with hypoxic or hyposmotic responses. The gene ontology 'osmoregulation' was significantly associated with gene responses to severe heat shock at 34°C, but not with those to hyposmotic shock. The ontology 'protein folding' was associated with all gene responses, but those observed during cold and hyposmotic shock and 'DNA repair' was associated with cold shock, severe heat shock and hypoxia. Lastly, genes that are part of the 'hemoglobin complex' were significantly associated with the responses to severe heat shock at 34°C (Table 3).

The expression profile of stress-responsive genes across stress treatments

Hierarchical clustering of genes that showed significant expression responses in at least one treatment identified four main clusters (Figure 5). Clustering appeared to have been driven primarily by the transcriptional responses of the 34°C winter and the five days at 31°C treatments. The observed expression responses varied depending on the type of stress applied (Figure 5). Candidate genes that showed differential expression in one treatment did not necessarily show a significant response in another treatment. Furthermore, some candidate genes showed opposite expression changes in response to different stressors, i.e. they were induced in response to one stressor but suppressed in response to another stressor.

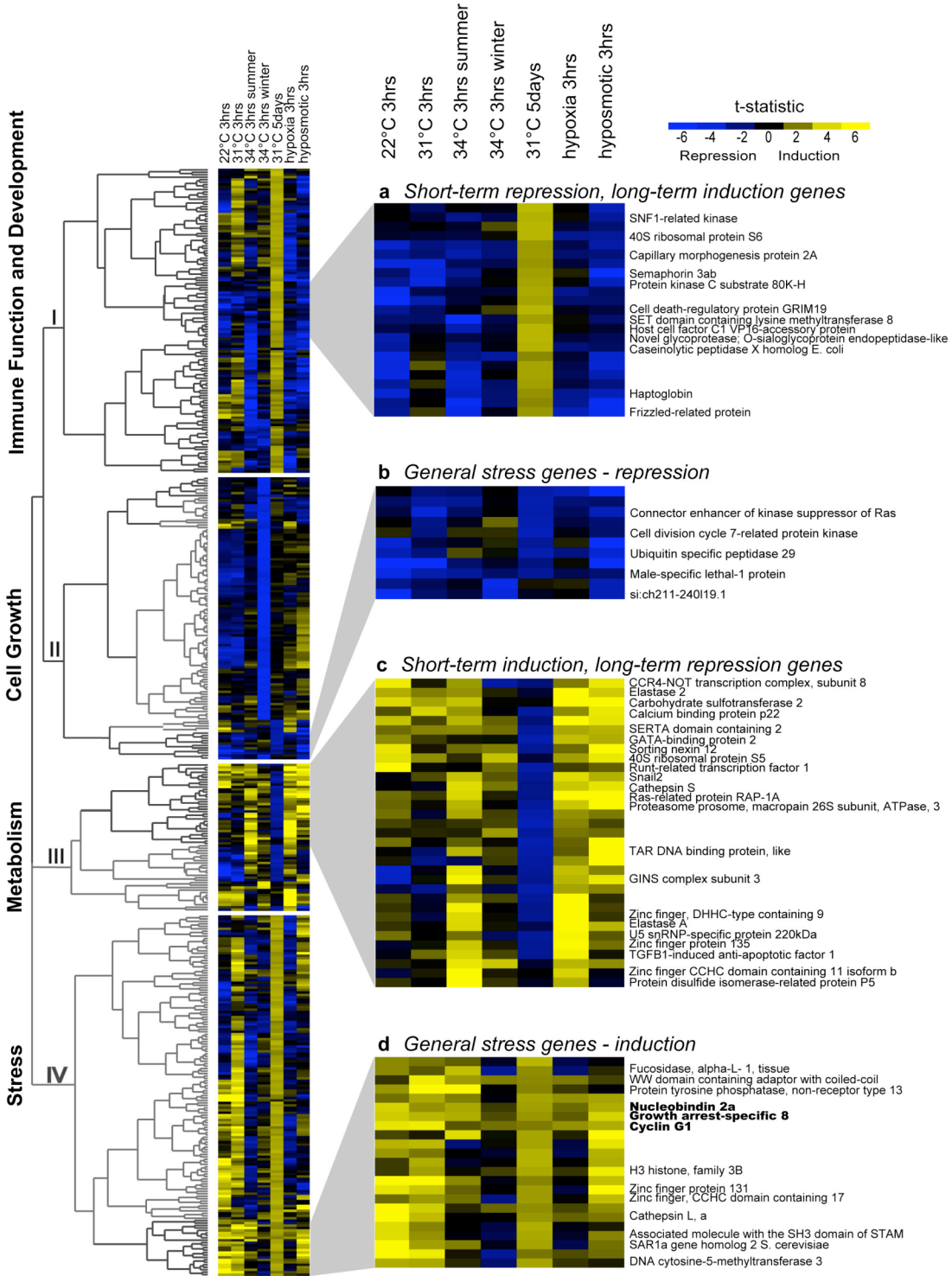


Figure 5 Results of hierarchical clustering of genes that showed differential expression in response to stress in *Pomacentrus moluccensis*. Represented values are the values of the t-statistic summarising the expression response of *P. moluccensis* across biological replicates.

Cluster I comprises genes which were strongly induced in the 31°C five days treatment, but showed variable responses in the short-term treatments. A subset of cluster I genes was generally down-regulated in the short-term treatments while being strongly induced in the long-term heat treatment (group a, Figure 5). Over-representation analysis of gene ontologies identified 'MHC class I protein binding' as significantly over-represented amongst cluster I genes when compared to all genes represented on the microarray (Table 5). Genes with developmental functions, such as 'neurite morphogenesis' and 'cell fate determination' showed a trend for over-representation amongst cluster I genes.

Table 5 Results of over-representation analysis of gene ontologies within gene clusters compared to the total compliment of genes represented on the microarray.

Gene Ontology	ID	Number of probes on array	Number of genes on array	Number of genes in gene cluster	Raw p-value	FDR-corrected p-value
Cluster I - Immune Function and Development						
MHC class I protein binding	GO:0042288	7	7	2	0.0000	0.0561
Neurite morphogenesis	GO:0048812	62	60	3	0.0032	0.1289
Cell fate determination	GO:0001709	59	45	2	0.0106	0.1807
Cluster II - Cell Growth						
Vesicle transport along actin filament	GO:0030050	181	9	2	0.0000	0.0251
SWR1 complex	GO:0000812	182	10	2	0.0000	0.0119
Cortical actin cytoskeleton	GO:0030864	183	11	2	0.0000	0.0122
Condensed nuclear chromosome	GO:0000794	13	13	2	0.0001	0.0168
Histone acetylation	GO:0016573	186	14	2	0.0001	0.0178
Actin filament	GO:0005884	187	15	2	0.0001	0.0190
Meiosis	GO:0007126	51	51	3	0.0002	0.0269
M phase of mitotic cell cycle	GO:0000087	97	96	3	0.0024	0.0779
Cluster III - Metabolism						
Sterol metabolism	GO:0016125	12	12	3	0.0000	0.0001
Trypsin activity	GO:0004295	14	14	3	0.0000	0.0001
Positive regulation of protein metab.	GO:0051247	18	18	3	0.0000	0.0002
Cholesterol metabolism	GO:0008203	18	18	3	0.0000	0.0002
Cluster IV - Stress						
Lytic vacuole	GO:0000323	42	42	3	0.0001	0.2378
Lysosome	GO:0005764	54	54	3	0.0004	0.3168
Response to UV	GO:0009411	6	5	1	0.0004	0.1620
Replicative cell aging	GO:0001302	5	5	1	0.0004	0.1296
Response to water deprivation	GO:0009414	5	5	1	0.0004	0.1080
Response to reactive oxygen species	GO:0000302	7	7	1	0.0008	0.1349

Cluster II contains a series of genes that were strongly down-regulated in the 34°C winter treatment and which showed variable expression responses in the other treatments. A subset of cluster II genes was down-regulated in most stress treatments

examined (group b, Figure 5). Cluster II genes showed significant over-representation of gene functions related to cell growth and cytoskeleton, e.g. 'cortical actin cytoskeleton', 'condensed nuclear chromosome', 'meiosis', 'M phase of mitotic cell cycle' (Table 5).

Cluster III was comprised of genes that were strongly suppressed during prolonged heat at 31°C. Most of these genes were induced in the short-term treatments. Group c includes those genes that showed the strongest suppression in the 31°C five days treatment and the strongest induction in the short-term treatments (Figure 5). Metabolic gene functions were significantly over-represented amongst cluster III genes, e.g. 'sterol metabolism', 'cholesterol metabolism', 'positive regulation of protein metabolism' (Table 5).

Lastly, cluster IV included genes that were strongly induced in the 31°C five days treatment, similar to cluster I genes. However, in contrast to cluster I genes many of the genes in cluster IV were also induced in the short-term treatments, and this was particularly so for genes in group d (Figure 5). Stress-related gene functions showed a trend for over-representation amongst cluster IV genes when compared to the total of genes represented on the microarray (Table 5). For example, 'response to UV', 'replicative cell aging', 'response to water deprivation', and 'response to reactive oxygen species' showed significant raw p-values, which, however, just missed the threshold for significance after FDR-correction.

Discussion

Coral reef fishes are predicted to experience a significant rise in sea surface temperatures due to global warming in the near future. At present, we have only limited data to estimate the potential physiological effects of elevated temperatures on coral reef fishes, since, thus far, temperature stress responses have been investigated in a small number of temperate fishes only. This is the first study to examine gene regulation in response to prolonged heat in a coral reef fish. Over 300 gene loci were regulated in response to a three-degree Celsius temperature rise over five days in the coral reef fish *P. moluccensis*, most of which were induced in response to heat. I have further identified a series of gene functions that were

significantly associated with the expression responses to prolonged heat. For example, expression responses to prolonged heat were significantly associated with the gene functions 'response to oxidative stress' and 'protein folding', indicating increased levels of oxidative stress and protein damage during heat stress. Protein and cellular damage arising from oxidative stress are likely to incur significant energetic costs. Furthermore, prolonged heat exposure appears to challenge the immune system of *P. moluccensis*, potentially increasing the risk for developing pathologies. I have also examined early gene responses three hours after exposure to a variety of environmental stressors, i.e. cold, heat, hypoxic, and hyposmotic shock. Early gene responses generally showed a suppression of transcriptional activity. The responses of individual gene loci varied depending on the type of stressor applied. However, stress-responsive genes could be clustered into four main groups of co-regulated genes, i.e. genes belonging to the same group showed similar expression responses across stress treatments. There was a significant relationship between gene cluster identity and gene function; co-regulated genes shared similar gene functions. In summary, early gene responses showed variability in the responses of individual gene loci, but there was conservation of gene function responses across stressors. Gene functions commonly associated with the response to stress can thus inform about the generality of the effects of different types of stress on biological function.

Common gene function responses elucidate how stress affects biological function in P. moluccensis

Gene functions commonly associated with the gene expression responses of *P. moluccensis* to a variety of environmental stressors included cell growth and cytoskeleton functions. Stress negatively affects cell growth (Carmeliet *et al.* 1998, Massague 2004). Since the actin cytoskeleton stabilises the cell during normal cellular function, cell growth and during stress, cytoskeletal elements commonly show differential expression in response to stress (Pedersen *et al.* 2001, Hall 2005). In the present study, the gene ontologies 'actin cytoskeleton organization and biogenesis', 'anti-apoptosis', 'induction of apoptosis', and 'cell cycle' were associated with most stress responses. Individual genes with function in cell growth and cytoskeleton and which were differentially expressed in response to stress in this study include *β-actin, plectin 1, dynein, spectrin, tubulin beta-2, filamin 2, acin1a,*

GRIM19, TP53INP1, septin 3, cyclin G1, cdca3, and *bax*. Genes that were generally suppressed during all stress responses examined here (cluster II genes) showed an over-representation of gene functions relating to cell growth and cytoskeleton, such as 'cortical actin cytoskeleton', 'actin filament' and 'M phase of mitotic cell cycle' (Table 5). Genes with functions in cell growth and cytoskeletal elements were generally suppressed in response to stress indicating a disruption and suppression of cell growth during stress in *P. moluccensis*.

Gene functions related to protein turnover were also commonly associated with the gene expression responses of *P. moluccensis* to environmental stress. 'Regulation of translation' was associated with all, 'protein folding' with most individual stress treatments examined. 'Ubiquitin-dependent protein catabolism' was significantly associated with all stressors but exposure to prolonged heat. Individual genes that were part of this component of the common stress response include *pdip5, USP29, psmc3, fbxl12*, the translation initiation factors *eif2b2, eif2c1, eif4b,* and the ribosomal proteins *rps6, mrpl30, mrpl36*. Stress increases rates of protein damage and ubiquitin-dependent protein catabolism (Hofmann and Somero 1995, 1996, Gabai and Sherman 2002). The gene expression responses of *P. moluccensis* measured in this study indicate that all stress treatments led to increased protein damage, an adjustment of the proteome, and *de novo* synthesis of proteins. Protein synthesis in ectotherms accounts for about 20 percent of the cellular energy budget (Houlihan 1991, Hofmann and Somero 1995). *De novo* protein synthesis and protein repair in response to stress are thus likely to incur substantial energetic costs in *P. moluccensis*.

A third group of gene functions commonly associated with stress responses in *P. moluccensi*s included metabolic gene functions, for example functions in carbohydrate and lipid metabolism. Genes that were generally induced in the initial responses to stress, but suppressed during prolonged heat (cluster III genes), showed significant over-representation of the gene functions 'cholesterol metabolism', 'sterol metabolism', and 'positive regulation of protein metabolism'. Genes with metabolic functions that were regulated in the present study include *SC4MOL, CHST2, amy2a, ACAD, dgat2l, dcxr,* and *fuca1*. Metabolic adjustments and the re-allocation of

energy resources during conditions of stress may reflect the need to allow for increased levels of protein and cellular repair during stress.

Genes with stress-related functions and functions in cellular repair were also commonly associated with the transcriptional responses to stress in *P. moluccensis*. For example, genes that were commonly induced in response to stress in *P. moluccenses* showed a trend for over-representation of the gene functions 'response to UV', 'replicative cell aging', 'response to reactive oxygen species', and 'response to water deprivation' (cluster IV genes, Figure 5, Table 5). Many other stress-related gene functions showed significant associations with the gene expression responses in individual treatments. The gene function 'response to temperature stimulus' was associated with all heat shock treatments, and 'osmoregulation' was associated with severe heat shock at 34ºC. Individual genes with stress-related gene functions and which were regulated in the present study include *dnajb11,* peroxiredoxin 6*, ptges3, apoea, hif1an,* and elastase 2. Elastase 2 hydrolyses collagen-IV and elastin and its activity has been associated with increased production of reactive oxygen species (Aoshiba *et al.* 2001). Regulation of elastase 2 during hypoxia suggests that *P. moluccensis* experienced oxidative stress as a result of reduced oxygen levels. Oxidative stress was also significantly associated with prolonged exposure to heat (Table 4). Three genes showed consistent induction across all stress treatments examined and which may constitute molecular markers of general stress in *P. moluccensis*: nucleobindin 2a, growth arrest-specific 8, and cyclin G1 (Figure 5). These genes are typically associated with cell cycle and cell growth and were consistently induced during stress in *P. moluccensis.*

Finally, genes with immune functions were also typically associated with the transcriptional responses to stress, potentially indicating a challenge to the immune system under conditions of stress. For example, the gene ontology 'response to pest, pathogen or parasite' was significantly associated with all individual stress responses examined here. The gene ontology 'MHC class I protein binding' was significantly over-represented amongst stress-responsive genes that were suppressed during short-term stress and induced during prolonged heat (cluster I genes, Figure 5). Differentially expressed genes with immune function identified in this study included *CHST2,* T-cell receptor alpha variable region, and haptoglobin. These data suggest

that different types of stress generally challenge the immune system of *P. moluccensis*. Stress has been previously associated with immunosuppression in teleosts (Bly and Clem 1991, Lupes *et al*. 2006). In *P. moluccensis*, the immediate response to stress appears to lead to a suppression of immune function genes, while prolonged exposure to heat appears to induce the transcription of immune function genes.

To summarise, a series of gene functions were consistently associated with the gene expression responses of *P. moluccensis* to cold, heat, hypoxic and hyposmotic shock. Conservation of these gene function responses under a variety of stress conditions suggests that there are common effects of stress on biological function. These common effects of stress include a) suppression of cell growth, b) increased protein damage and *de novo* synthesis of proteins, c) metabolic adjustments and a re-allocation of energy resources possibly related to increased protein and cellular repair, d) induction of stress genes and cellular repair systems, e) suppression of the immune system.

The gene expression response to prolonged heat

In addition to this common set of stress responses, there were also important differences in the responses to different types of stress. For example, 'ubiquitin-dependent protein catabolism' was significantly associated with gene expression responses in all short-term stress treatments, but not with gene expression responses to prolonged heat. In contrast, the gene ontology 'protein folding' was associated with most stress responses, including the expression response to prolonged heat. Association of the gene function 'protein folding' with gene expression responses to heat stress over five days may thus indicate increased levels of protein damage even after prolonged exposure to heat. Individual genes with function in protein folding and which were regulated in response to prolonged heat include *dnajb11, clpx, PFDN4, ptges3*, all of which were induced (Table 1). Hence, initial adjustments of the proteome during stress include the increased degradation of (stress-damaged) proteins and increased levels of ubiquitin-dependent protein catabolism. Over time, rates of protein catabolism may return to baseline levels, while protein chaperone systems remain induced. Given that protein synthesis in ectotherms accounts for

around 20 percent of the cellular energy budget (Houlihan 1991, Hofmann and Somero 1995), induction of protein chaperone systems over this prolonged period is likely to incur a considerable energetic cost to heat-stressed *P. moluccensis* and may potentially reduce the energy available for other cellular and organismal functions.

Prolonged exposure to heat was further significantly associated with the gene function 'response to oxidative stress'. Oxidative stress is caused by the formation of reactive oxygen species and has been associated with many forms of stress (Freeman *et al.* 1990, Rifkin *et al.* 1993, Abele *et al.* 1998, Flanagan *et al.* 1998, Abele *et al.* 2002, Abele and Puntarulo 2004, Gius *et al.* 2004, Heise *et al.* 2006a, Heise *et al.* 2006b, Lesser 2006). In this study, peroxiredoxin 6 and apolipoprotein Ea, two genes with antioxidant function, were both induced in response to prolonged heat. It is likely that oxidative stress and the associated cellular and molecular damage incur additional energetic costs during prolonged exposure to heat in *P. moluccensis*. Since 'carbohydrate metabolism' and 'lipid metabolism' remain altered even during prolonged exposure to heat, stress appears to lead to a continuous re-allocation of energy resources. For example, the genes *ACAD, dgat2l, dcxr, HADH2, crabp2, Glb1, elovl6l, fuca1, AGPAT3* all function in carbohydrate or lipid metabolism and were regulated in response to prolonged heat. This metabolic re-organisation is likely due to the energetic needs associated with protein and cellular repair arising from prolonged heat stress.

As discussed above, stress commonly suppresses immune function and chronic stress is typically associated with an increased risk of developing pathologies (Bly and Clem 1991, Lupes *et al.* 2006). The association of 'response to pest, pathogen or parasite' with the expression response to prolonged heat suggests a continuous challenge of the immune system and, potentially, an elevated risk for disease that persists over time. The ontology 'hemoglobin complex' was associated with the response to prolonged heat and heat shock at 34ºC. Hence, heat shock of severe magnitude or of moderate magnitude but prolonged duration appears to compromise oxygen supply and induce genes with function in oxygen delivery. In the present study, hemoglobin beta embryonic-1, hemoglobin alpha embryonic-3, and haptoglobin were induced during prolonged heat, reflecting this need to enhance oxygen delivery systems during heat stress.

This is the first study to examine the transcriptional responses of a coral reef fish to prolonged exposure to elevated temperatures. While I used liver for the analysis of gene expression responses, other tissues may have been affected differently. However, liver is a metabolically important tissue and would thus allow detection of a large number of physiological responses. This study identified significant gene regulation in liver, suggesting alterations of biological functions in response to a temperature rise of three degrees Celsius. Prolonged heat was associated with increased protein damage and induction of protein chaperone systems, challenge of the immune system, and increased levels of oxidative stress. These data suggest that prolonged heat leads to cellular and protein damage and would thus incur significant energetic costs to cells and the organism as a whole. Metabolic functions were altered in response to prolonged heat in *P. moluccensis* suggesting that energy resources are re-allocated during heat, potentially reducing the available energy for other cellular and organismal functions, such as growth and reproduction. These results provide the first evidence that the transcriptome and biological functions of coral reef fishes can be significantly altered in response to a three-degree Celsius increase in temperature. Such temperature rise is well within the range of predicted temperature increases of current climate change models. It is possible that the observed effects may improve or deteriorate following longer-term exposure to heat. Future work may thus investigate the longer-term effects of increased temperatures on gene regulation as well as on growth, reproduction, and immune function. Since the functions of many heat-responsive genes are still unknown, experimental characterisation of gene functions during stress would also enhance our understanding of the effects of elevated temperatures on coral reef fishes.

The early gene response to cold shock

Early gene responses three hours after exposure to a variety of stressors showed some important differences to gene responses to prolonged heat exposure. For example, cold stress is typically associated with an induction of transcriptional activity and increased rates of protein translation (Ju *et al.* 2002, Gracey *et al.* 2004, Cossins *et al.* 2006). This increase in rates of transcription and translation in response to cold may compensate for reduced metabolic and enzymatic rates at lower temperatures. In the present study, most genes showing greater than 1.5-fold

expression changes following cold shock were down-regulated, although the two genes with the greatest average expression changes, *mknk2* and *ube2h*, were induced. Both of these genes have previously identified functions in stress-related responses, being part of protein kinase cascades and protein catabolism. 'DNA repair' was further associated with the expression response to cold shock, possibly indicating that cold shock led to increased levels of DNA damage. Various types of stress can cause DNA and chromosome damage (Fischman *et al*. 1996). The gene function 'DNA repair' was associated with many stress responses in the present study. Stress may thus commonly exhibit genotoxic effects.

The early gene response to moderate heat shock

Short-term moderate heat shock at 31°C resulted in one gene, *plec1*, being significantly up-regulated. This gene functions as a structural constituent of the cytoskeleton and its induction in response to heat shock likely reflects the cell's need to stabilise the cytoskeleton and modify rates of cell growth during stress. The gene ontologies 'cell growth' and 'mitotic cell cycle' were not significantly associated with the response to the 31°C short-term heat shock. This heat shock treatment appeared, therefore, to only have a minor effect on cell growth functions. There was an approximately even distribution of down- and up-regulated genes in response to moderate heat shock at 31°C. However, most genes with greater than two-fold average expression changes were induced. For example, u*be2h* functions in protein catabolism and was also highly induced in response to cold shock. Regulation of the protein kinases and signal transduction genes *prkcsh*, *ptk9l*, *odz4*, and *ednra* suggests the involvement of protein kinase related signal transduction pathways in response to heat shock. The gene function 'response to temperature stimulus' was significantly associated with the observed expression responses to heat shock and included genes such as *ube2h*, *mapk8*, and *ppp1cb*.

The early gene response to severe heat shock

In summer, severe heat shock at 34°C over three hours resulted in the suppression of four genes, *rhoA*, *SC4MOL*, *rgs4*, and *znf395*, while two genes, *pdip5* and *ZCCHC11*, were induced. The expression of *pdip5* changed almost five-fold in this treatment. Since *pdip5* has known function in protein folding, its induction in

response to severe heat stress most likely reflects increased levels of protein damage in heat-shocked *P. moluccensis*. Transcription levels of most gene loci showing greater than 1.5-fold expression changes in this treatment were suppressed compared to expression levels at ambient temperature. Besides the common stress responses, the gene functions 'response to temperature stimulus', 'osmoregulation', and 'DNA repair' were significantly associated with the gene expression responses to severe heat shock. Their involvement suggests a substantial disruption of cellular homeostasis during severe heat shock. 'Hemoglobin complex' was further associated with the response to severe heat shock possibly indicating that these cells suffered from reduced oxygen supply as a result of the heat shock. During severe heat shock, oxygen demand may no longer be matched by oxygen supply and delivery systems, effectively leading to hypoxia in the cells (Pörtner 2002, Pörtner *et al.* 2004). The regulation of hemoglobin genes in this study is, therefore, likely to reflect the need for increased oxygen delivery capacities during severe heat shock.

Acclimation

There was a remarkable difference between the summer and winter gene expression responses to severe heat shock at 34°C over three hours. In winter, where ambient temperatures were 26°C, 121 gene loci showed significant expression changes in response to severe heat shock at 34°C (see Chapter 2). In summer, where ambient temperatures were 28°C, only six genes showed significantly altered expression levels. This difference between the summer and winter gene responses to severe heat shock may be either due to a real biological phenomenon, such as acclimation, or a technical artefact due to the different experimental design employed in the analyses. Winter heat-shocked *P. moluccensis* were compared to a common reference consisting of four pooled *P. moluccensis* kept at ambient temperature. It is thus possible that the winter response mainly reflected the transcriptional differences between individually heat-shocked *P. moluccensis* and this particular pooled control, rather than the population of *P. moluccensis* kept at ambient temperature as a whole. The direct comparisons between individual heat-shocked *P. moluccensis* and individual *P. moluccensis* kept at ambient temperature used in the summer study are more likely to assay gene responses that are representative of the reef fish population as a whole. However, seasonal acclimation has been shown to lead to extensive

transcriptional adjustments in teleosts (Kikuchi *et al.* 1995, Vera *et al.* 1997, Sarmiento *et al.* 2000, Picard and Schulte 2004). Seasonal acclimation may, therefore, account for at least some of the transcriptional differences observed. However, further data are required to understand the effects of seasonal acclimation on heat shock responses. In particular, comparisons of summer and winter responses across multiple years are required to conclusively test for seasonal differences in transcriptional responses. There were also important similarities between winter and summer responses to severe heat shock measured in this study. Winter and summer responses to severe heat shock resulted in a general suppression of transcriptional activity and mostly involved the same type of gene functions. For example, the gene functions 'response to temperature stimulus', 'osmoregulation', 'DNA repair' and 'hemoglobin complex' were significantly associated with both summer and winter expression responses.

The early gene response to hypoxia

Many known responses to hypoxia are aimed at compensating for reduced cellular oxygen levels by increasing capacities for oxygen delivery and by enhancing oxygen-independent ATP production by means of glycolysis (Semenza *et al.* 1994). The hypoxia-inducible transcription factor *HIF-1* has been shown to play an important role in coordinating transcriptional responses to hypoxia. For example, *HIF-1* commonly induces the transcription of glycolytic enzymes during hypoxia (Semenza *et al.* 1994, Schmedtje and Ji 1998, Semenza 1998, 2000). Activation of *HIF-1* during hypoxia has also been shown to increase glucose transporter levels as well as gene products which have angiogenic or vasodilatory effects (Schmedtje and Ji 1998, Semenza 2000, Hirota and Semenza 2006, Semenza 2006). In the present study, *HIF-1a* showed only very limited regulation. *HIF-1a* was 1.2-fold induced during severe heat shock, -1.1-fold suppressed during prolonged heat, and not regulated during hypoxia (data not shown). Hypoxia-inducible factor 1, alpha subunit inhibitor (*HIF1an*) was significantly down-regulated during prolonged heat (Table 1). *HIF1an* mediates repression of HIF-1 transcriptional activity. Hence, further data are required to understand the role of *HIF-1a* and *HIF1an* in the response to hypoxia and prolonged heat in *P. moluccensis.*

However, association of the gene ontologies 'glycolysis' and 'gluconeogenesis' with the expression response to hypoxia measured in this study is consistent with the need to increase oxygen-independent ATP production during hypoxia. Amongst the genes that showed the most significant expression responses to hypoxia was amylase-3 protein (*amy2a*), which functions in carbohydrate metabolism and which was induced in response to hypoxia. *Amy2a* endohydrolyses 1,4-α-D-glucosidic linkages in starch, glycogen, and related polysaccharides to make dextrin, which can then be reduced to α-D-glucose. Thus, during hypoxia, *amy2a* appears to free stored sugars for use in glycolysis and thus would facilitate oxygen-independent ATP production. Contrary to expectation, 'hemoglobin complex' was not significantly associated with the transcriptional responses to hypoxia in this study despite the fact that hemoglobin is an important oxygen transporter and oxygen delivery systems are commonly induced during hypoxia. In addition, there were important differences in metabolic regulation between hypoxia and heat shock. For example, 'lipid metabolism' was significantly associated with the responses to heat shock, but not with the expression responses to hypoxia. It thus appears that *P. moluccensis* utilises different energy resources during hypoxia and heat shock.

The early gene response to hyposmotic stress

Pomacentrus moluccensis appears to possess a relatively large tolerance towards low salinity. Exposure to hyposmotic conditions where the salinity of the seawater was reduced from 36ppt to 20ppt did not result in significant expression changes and none of the genes assayed showed greater than two-fold expression changes. Nevertheless, a number of gene functions that were discussed above as part of the common response to stress were significantly over-represented amongst top ranked genes (Table 3). These results suggest that cells did perceive the low salinity conditions as stress. However, the transcriptional responses to this stress appear to have been too moderate to allow detection of significant responses at individual gene loci. It may also be possible that the gene expression response to low salinity conditions involves very small expression changes, which may not be easily detected using a heterologous microarray approach and a limited number of biological replicates.

Strategies for managing and interpreting environmental genomic data

As demonstrated in this study, a large number of gene loci are typically regulated in response to environmental stress. The complexity of transcriptional responses to environmental stress can pose significant challenges to the interpretation of gene responses. In particular, the significance of gene regulation for biological function often remains unclear. In this study, I have demonstrated that gene class testing using the gene set resampling algorithm and the over-representation analysis can provide a useful means for comparing transcriptional responses across stress treatments and to relate gene regulation to biological function. The gene set resampling algorithm employs the information that is contained in the gene ranking in order to determine the type of genes that are ranked most highly and the type of functions that are most likely involved in the observed expression responses. Such information would be lost if only genes with individually significant expression responses were considered. Using this approach, I was able to identify a set of gene functions that was consistently associated with all stress responses examined here. Based on these shared gene function responses, I was able to infer the common effects of stress on biological function.

Another strategy for interpreting the complex gene responses to environmental stress involves the identification of stress-activated signal transduction pathways and transcription factors. For example, stress commonly activates the p38 and the *c-Jun* N-terminal kinase (JNK) signalling pathways (Gabai and Sherman 2002, Cowan and Storey 2003). These signalling pathways form part of the superfamily of mitogen-activated protein kinases (MAPK). JNK induces the transcription factors *ATF2, Elk-1*, *Smad3*, tumor suppressor p53, *c-myc*, *c-fos* and *c-Jun*, which in turn are responsible for coordinating transcriptional activities at a very large number of gene loci (Cowan and Storey 2003). Many of the transcription factors under the control of the JNK pathway are immediate early genes. The p38 pathway interacts with small heat shock proteins, e.g. *hsp27*, the eukaryotic translation initiation factor *elF-4E*, the *CREB* transcription factors, and histone H3 to coordinate translation, transcription, chromatin remodelling and cytoskeleton stabilisation in response to stress (Cowan and Storey 2003). Hence, identification of signal transduction pathways, immediate

early genes, and transcription factors that are activated in response to stress should allow better interpretation of the vast number of transcriptional changes that occur downstream of these primary responses. Ultimately, we would like to understand how different environmental stressors elicit different transcriptional responses despite induction of the same primary response genes and the same signalling pathways.

In the present study, the gene ontology 'MAPKKK cascade' was significantly associated with the expression responses in all treatments except the three-hour heat shock at 31ºC. It is possible that the MAPKKK cascade is involved in all stress responses, but that the three-hour heat shock at 31ºC was too moderate in order to measure its involvement in the response. The ontology 'JNK cascade', which is a daughter term of the 'MAPKKK cascade', was significantly associated with all temperature stresses examined except the three-hour 31ºC heat shock, and was not associated with hypoxic or hyposmotic stress. I have further identified a number of genes that are directly related to these signal transduction pathways and which were regulated in the stress responses measured here. Tumor protein p53 inducible nuclear protein 1, H3 histone, family 3B, the eukaryotic translation initiation factors 2B, 2C, and 4B were regulated during prolonged heat at 31ºC over five days, while *CEBPD* and *nkrf* were both down-regulated in the winter 34ºC treatment. In addition, a series of genes of the *ras* family of small GTPases were regulated in this study, e.g. *rhoG*, *rhoQ*, and *rap1a* were suppressed, while *rasgef1b* and *rab1a* were induced in response to prolonged heat. The GTPase r*hoA* was down-regulated in the severe heat shock treatments at 34ºC in summer and winter, while *rhoC* was down-regulated in the 34ºC winter treatment. The gene ontology 'rho protein signal transduction' was significantly associated with the prolonged heat stress at 31ºC, severe heat stress at 34ºC, and hyposmotic stress. In summary, these results support an important involvement of the MAPKKK signalling cascade in all stress responses examined here, an involvement of the JNK cascade in temperature stress responses, and an important role for *rho* signal transduction in severe and prolonged heat stress.

I also examined the expression of known immediate early genes, such as *c-myc*, *c-Jun*, *c-fos*, *JunB*, *egr-1*, and *SRF* in the present study. There was no evidence for significant regulation of these genes in the present study (data not shown).

Microarray gene expression profiling, and in particular the use of a heterologous array platform, may not allow detection of the small fold changes we expect for these genes. Quantitative real-time PCR may be better suited to detect expression changes for these immediate early genes under conditions of environmental stress. As we accumulate more information regarding the induction of MAPK signalling pathways under different conditions of environmental stress and the cross-talk between these pathways, we will be better able to interpret the complex gene expression changes that occur in response to environmental stress.

Conclusion

In this chapter, I have presented evidence that the transcriptome and biological functions of coral reef fishes can be significantly altered in response to a three-degree Celsius increase in temperature. Such temperature rise is well within the range of predicted temperature increases of current climate change models. Future work may address the longer-term effects of heat on gene regulation and biological function.

By examining transcriptional profiles under a variety of stress conditions I have identified conservation of gene function responses, but variability in the responses of individual genes. Conservation of gene function responses suggests that there are common effects of stress on biological function, for example, the suppression of cell growth and increased protein damage. Analysis of gene function responses can thus help relate gene regulation to biological function. Several signal transduction pathways were significantly associated with the transcriptional responses observed, while immediate early genes did not show significant gene regulation. In Chapter 5, I expand on the relationships between transcriptional stress responses and responses at other levels of biological organisation and identify common effects of stress on biological function. I also discuss the role of signal transduction pathways and immediate early genes in coordinating transcriptional stress responses in more detail.

CHAPTER 4

Using inter-species genomic hybridisations to identify rapidly diverging genes

Publication: Kassahn KS, Crozier RH, Connolly AR, Ward AC, Stone G, Caley MJ (*In prep*) Using inter-species genomic hybridisations to identify rapidly diverging genes.

Abstract

Rapidly evolving genes interest evolutionary biologists because they can be indicative of selection and adaptive evolution. While the increasing amount of whole genome sequence data is facilitating *in silico* studies of molecular evolution including the identification of rapidly diverging genes, many species of interest to comparative genomic studies have not yet been sequenced. Comparative genomic hybridisation techniques on microarrays can provide an alternative to *in silico* approaches for the identification of rapidly diverging genes. In this study, I used inter-species genomic hybridisations on microarrays to identify rapidly diverging genes between *Danio rerio* and the coral reef fish *Pomacentrus moluccensis*. 985 genes, approximately six percent of all genes assayed, showed accelerated rates of sequence evolution between *D. rerio* and *P. moluccensis*. Rapidly diverging genes were over-represented in receptor, transcription coactivator and cell signalling functions. In order to test whether accelerated rates of sequence evolution could also be confirmed in other teleost lineages, I obtained orthologous DNA sequences to *D. rerio* in the teleosts *Takifugu rubripes* and *Gasterosteus aculeatus*. I estimated rates of sequence evolution based on average tree branch lengths and the number of nucleotide changes per site. *In silico* studies confirmed longer average branch lengths for rapidly diverging genes as compared to genes that showed evidence of significant sequence similarity in the genomic hybridisation experiment. The receptor gene *reverbb2*, the cell cycle gene *geminin*, the homeobox gene *msh-D* and the signal transduction gene *Tradd* showed some of the greatest rates of sequence evolution

across multiple teleost lineages. I have here demonstrated the utility of inter-species genomic hybridisation experiments to identify rapidly diverging genes. Such an approach is particularly useful for species for which there are few genomic resources.

Introduction

Rapidly diverging genes are of interest to evolutionary biologists because genes with high evolutionary rates have been implicated in speciation and species diversification (Turner *et al.* 2005, Harr 2006). In addition, high rates of evolutionary change may indicate strong directional selection and/or adaptation, especially where selection on gene function can be directly related to environmental pressures (Tautz and Schmid 1998, Lecompte *et al.* 2001, Matzkin 2005, Fairhead and Dujon 2006). A number of rapidly diverging genes have already been identified. For example, genes involved in olfaction (Clark *et al.* 2003), receptor genes (Jordan *et al.* 2001), the breast cancer BRCAI gene and the testis-determining factor SRY (Makalowski *et al.* 1996), the fertilization gene lysin (Swanson and Vacquier 1998, Hellberg and Vacquier 1999), and other reproductive genes (Swanson *et al.* 2001) have been shown to diverge rapidly between closely related species. Gene regions encoding extra-cellular protein families and extra-cellular portions of proteins also show accelerated rates of sequence evolution when compared to their cytoplasmic counterparts (Luz and Vingron 2006). Rapidly diverging genes, therefore, appear to commonly have functions in reproduction, especially in sperm competition, species recognition, and in environmental interactions. Rapid evolution may also occur in cases of loss of gene function resulting in pseudogenes. However, in this thesis chapter, I only consider instances of rapid adaptive evolution where the gene remains functional.

In silico approaches to compare whole genome sequence data have proven useful for identifying rapidly diverging genes. However, the general applicability of *in silico* approaches is limited by the number of species for which there are genome sequence data. Comparative genomic hybridisation experiments on microarrays can provide an alternative to *in silico* approaches for the identification of rapidly diverging genes. Comparative genomic hybridisation experiments are typically used to assess copy number changes and gene deletion or duplication events (Pollack *et al.* 1999). The

same technique can also be adapted to estimate sequence divergence by competitively hybridising genomic DNA of two taxa onto microarrays (Kim *et al.* 2002, Hinchliffe *et al.* 2003, Brunelle *et al.* 2004). Genes diverging at accelerated rates can be identified on the basis of significant differences in cross-hybridisation in such inter-species genomic hybridisations. It is possible that other genes also have rapid rates of sequence evolution but that saturation of those sites which can be varied prevents these sequences from diverging (for an example, see DeSalle *et al.* 1987). Here, I refer to rapidly diverging genes as those genes where fast rates of sequence evolution have led to significant sequence divergence.

The coral reef fish *Pomacentrus moluccensis* has been the subject of functional genomic studies investigating the physiological effects of elevated sea surface temperatures on coral reef fishes (chapters two and three of this thesis). However, there are few genomic resources available for *P. moluccensis* and its genetic potential for adaptation to elevated sea surface temperatures is not well understood. In previous studies, I have therefore used a heterologous microarray developed for *Danio rerio* to measure gene responses of *P. moluccensis* to heat stress (chapters two and three of this thesis). However, genes with stress-related functions may be expected to evolve at accelerated rates because such genes can be under strong directional selection during the process of adaptation to new environments. Indeed, some genes related to environmental interactions and stress responses have been shown to evolve at accelerated rates in mammals and fungi (Jordan *et al*. 2001, Le Quere *et al*. 2006). Since cross-hybridisation to a heterologous microarray largely depends on sequence similarity, I sought to determine the level of sequence divergence between these two species and test whether genes with stress-related functions are rapidly diverging since this would compromise the ability to measure stress responses using a heterologous microarray approach.

Here I report inter-species genomic hybridisation experiments designed to identify genes with the greatest sequence divergence between the coral reef fish *P. moluccensis* and *D. rerio*. Both taxa were competitively hybridised to an oligonucleotide microarray designed for *D. rerio*. Sequence divergence was estimated on the basis of relative hybridisation signals in comparative genomic hybridisations. I tested for a relationship between rates of evolution and gene

function and for an over-representation of stress-related gene functions amongst the genes that were identified as rapidly diverging. I also examined whether genes identified as rapidly diverging between *P. moluccensis* and *D. rerio* were also rapidly diverging in other teleost lineages using a bioinformatic approach. Orthologous sequences to *D. rerio* were obtained for *Takifugu rubripes* and *Gasterosteus aculeatus* and rates of sequence evolution were estimated based on tree branch lengths. Branch lengths of genes identified as rapidly diverging in the genomic hybridisation experiments were compared, assuming that if genes were also rapidly diverging in other teleost lineages they would show comparatively longer branch lengths. Genes that were identified as showing significant sequence similarity between *P. moluccensis* and *D. rerio* and genes commonly used in phylogenetic studies were used as a baseline for comparing relative branch lengths and rates of sequence evolution.

Material and Methods

Inter-species genomic DNA hybridisation experiments

I performed four individual inter-species genomic DNA hybridisations (IGH) in which I competitively hybridised *P. moluccensis* and *D. rerio* genomic DNA (gDNA) to the Compugen 16K *D. rerio* oligonucleotide microarray. This microarray contains 16,399 oligos (65-oligomers) representing 15,806 unique *D. rerio* gene clusters plus controls. The list of genes immobilized on the array is available at http://www.microarray.adelaide.edu.au/libraries/microarrays.html. Genomic DNA was extracted from fish caudal fins of both species using CTAB (Murray and Thompson 1980) and quantified using spectrophotometry. Five μg of gDNA per sample was digested with *DpnII* and labelled using the BioPrime®Plus Array CGH Genomic Labeling System (Invitrogen). Dye usage was swapped between *P. moluccensis* and *D. rerio* samples in order to account for potential dye bias. Genomic DNA from three individuals of *D. rerio* was pooled. Four biological replicates of *P. moluccensis* gDNA were each mixed with an aliquot of fluorescently labelled *D. rerio* pooled gDNA. Each of the four gDNA mixtures was hybridised to an individual *D. rerio* microarray. Prior to hybridisation, each microarray slide was immersed in distilled water at 60ºC for 5min and dried by centrifugation at 650 x g

for 5min. Fifty μg of human Cot-1 was added to each labelled gDNA sample, dried under reduced pressure, resuspended in 14μl formamide and 14μl of 6.25 X SSC, denatured by heating to 100ºC for 3min and transferred directly to ice. Finally, 0.6μl of 10% SDS was added to each sample. The probes were applied to the array and incubated at 42ºC overnight in a humidified chamber. The arrays were washed in 0.5 X SSC containing 0.01% SDS for 1min, 0.5 X SSC for 3min and 0.2 X SSC for 3min. The slides were scanned using an Axon 4000B microarray scanner. Single image .tif files were saved for data analysis.

Microarray statistical data analysis and data mining

The Cy5 and Cy3 fluorescent signal intensity of each gene on the array was extracted using SPOT software (CSIRO Mathematical and Information Sciences, Australia). The background fluorescence was subtracted and the ratio of the resultant signal intensities (Cy5/Cy3) was $\log_2$-transformed. Statistical analyses were performed using the software package LIMMA (Smyth 2005). The transformed signal intensities of each grid on the array and global signal intensity were print-tip Loess normalised (Smyth and Speed 2003), and scale normalised between arrays.

These IGH experiments required a new statistical approach for analysis. Because *P. moluccensis* gDNA will not preferentially hybridise to a *D. rerio* oligonucleotide microarray in the presence of *D. rerio* gDNA, I was able to measure the random noise in the IGH experiments by using spots with positive M-values. Due to the Loess normalisation discussed above potentially shifting the M-value away from M=0 (corresponding to equal hybridisation), I used the modal M-value, estimated from a density estimate of M-values, as the centre of the distribution. The 99th percentile of spots with M-values greater than this mode was taken as a threshold to decide which spots had poorly hybridised with the *P. moluccensis* gDNA and spots with M-values that exceeded this threshold below the mode were flagged as showing significantly reduced hybridisation signal in *P. moluccensis*. Genes with negative M-values outside this range are attributed to *D. rerio* gDNA hybridising in preference to *P. moluccensis* gDNA due to sequence divergence between the two taxa. In contrast, genes that lie within the 99th percentile show similar hybridisation signal in *P.*

moluccensis and *D. rerio* and are thus considered to share significant sequence similarity in the two taxa.

I compared this method for identifying rapidly diverging genes to other approaches that were based on the values of the t-statistic. Approaches based on the values of the t-statistic may be expected to provide more reliable identification of divergent genes because variability between replicates in the IGH experiments can be taken into account. Similar to the M-value approach, positive t-statistics were used to estimate the random noise in the IGH experiments and the t-value corresponding to the 99th percentile of t-values above the mode was extrapolated towards negative values to determine which genes showed significantly reduced hybridisation signals in *P. moluccensis*. I further performed individual t-tests for significant differences in hybridisation signal between *P. moluccensis* and *D. rerio* on a gene-by-gene basis and Holm-corrected the resulting p-values for multiple testing. I compared the number of genes identified as rapidly diverging using these different approaches and also examined the M-A and t-A plots relating to these analyses. M-A and t-A plots provide information about the distribution of M- and t-values and can thus help determine how different analytical approaches affect the identification of rapidly diverging genes.

The microarray platform on which these analyses are based is deposited at the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/projects/geo/; Barrett *et al*. (2005)) under the accession GPL3365 (microarray platform), GSE4047 (data series), GSM104737, GSM104739, GSM104741, and GSM104742 (samples). The raw image .tif files can be downloaded at ftp://ftp.ncbi.nih.gov/pub/geo/DATA/supplementary/samples/.

Gene annotation was performed using the program Resourcerer 12.0 (Tsai *et al.* 2001). Gene function of identified candidate genes was estimated by gene functions determined for *D. rerio* using a combination of AmiGo (The Gene Ontology Consortium 2000) (http://www.godatabase.org/cgi-bin/amigo/go.cgi), iHop (http://www.pdg.cnb.uam.es/UniPub/iHOP/) (Hoffmann and Valencia 2004), and databases at NCBI (http://www.ncbi.nlm.nih.gov/).

Since the Compugen *D. rerio* microarray employed in this study was designed to minimise the distance of the oligos from the 3' end of genes, some oligos are positioned outside of coding gene regions. These oligonucleotides thus inform about the level of sequence divergence in untranslated gene regions. Because rates of evolution are expected to differ between coding and untranslated gene regions, I determined which of the rapidly diverging genes were identified by oligos complementary to coding gene regions and which were complementary to untranslated gene regions. Therefore, all the 65mer oligonucleotide sequences representing genes identified as being rapidly diverging were blasted against the *D. rerio* protein database using netblast 2.2.14 and BLASTX. Genes that produced a significant hit (E-value < 0.0001) were considered to match the coding region of a gene. Oligos that did not produce hits against the *D. rerio* protein database were in untranslated gene regions. I then examined the functions of rapidly diverging genes. I distinguished between rapidly diverging genes identified by oligos in either coding or untranslated gene regions. Gene function analyses were performed using gene-class testing and the over-representation analysis algorithm implemented in the software ermineJ (Lee *et al*. 2005). The ORA algorithm is most appropriate here because the genes for this analysis naturally fell into two groups, having been identified as rapidly diverging or not. Gene functions of rapidly diverging genes, those identified by oligos in coding gene regions and those identified by oligos in untranslated gene regions, were thus contrasted against the gene functions of all genes represented on the microarray. The resulting p-values were FDR-corrected and values < 0.1 were considered statistically significant.

Estimation of evolutionary rates across teleost lineages using tree branch lengths

I tested whether genes identified as rapidly diverging in the IGH experiment of *P. moluccensis* and *D. rerio* were also rapidly diverging in other teleost lineages. For this purpose, I calculated tree branch lengths for rapidly diverging candidate genes using orthologous sequences for *D. rerio* and the teleosts *T. rubripes* and *G. aculeatus*. Only rapidly diverging genes that had been identified by oligos in coding gene regions in the IGH experiments were used in the *in silico* study. Orthologs of type one2one, that is orthologs which do not have known paralogs, were obtained

using the Ensembl release 40, August 2006. Sequences were aligned using a combination of CLUSTALX v1.8 (Thompson *et al.* 1997) and Se-Al v2.0a11 (Rambaut 1996). The tree topology describing the phylogenetic relationships between these teleost species is firmly established and the relative position of *P. moluccensis* was inferred from the phylogenetic position of other members of the Labroidei (Metscher and Ahlberg 1999, Chen *et al.* 2004, Roest Crollius and Weissenbach 2005) (Figure 1).

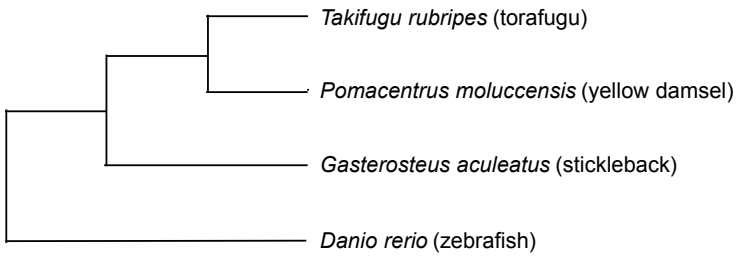


Figure 1 Phylogenetic relationships of the teleost species used for the identification of rapidly diverging genes. *Danio rerio* and *Pomacentrus moluccensis* were used in the comparative genomic hybridisation experiments, while *D. rerio, Gasterosteus aculeatus* and *Takifugu rubripes* were used in the calculation of tree branch lengths based on publicly available sequence data. Relative phylogenetic positions were inferred using NCBI taxonomy (http://www.ncbi.nlm.nih.gov/).

The number of nucleotide changes per site for each branch of this phylogeny was determined by calculating a neighbour joining tree based on Maximum Likelihood, a transition/ transversion ratio of 2.0, and the F84 model (Felsenstein and Churchill 1996) using the DNADIST and Neighbour programs from the PHYLIP v3.573c package (Felsenstein 1989). Total tree branch lengths were determined by adding the total number of nucleotide changes per site across all branches for each gene. Tree branch lengths of ten genes that were identified as rapidly diverging in the IGH experiments were compared to those of twelve genes that showed good cross-hybridisation (M values not significantly different from 0) in the IGH experiment and were thus thought to share high sequence identity between *P. moluccensis* and *D. rerio*. The choice of genes for *in silico* sequence comparisons was limited to those genes for which orthologs could be unambiguously identified in the other teleost species. Genes that satisfied this criterion were selected at random for the estimation of tree branch lengths. I further estimated tree branch lengths for a selection of genes that are commonly used in teleost phylogenetic studies and which may, therefore, provide a useful baseline for judging relative rates of sequence evolution: cytochrome oxidase I, cytochrome b, elongation factor 1α, and β-actin.

Results

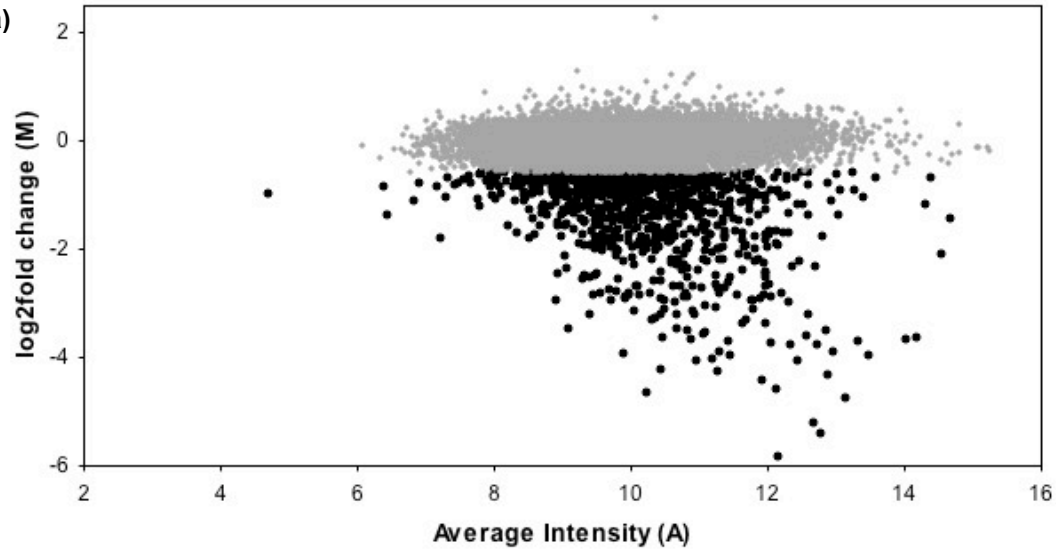
Identification of rapidly diverging genes using inter-species genomic hybridisations

The mode of the distribution of M-values in the IGH experiment was 0.038. The 99th percentile of positive M-values above the mode corresponded to an M-value of 0.622. Extrapolating this M-value below the mode resulted in -0.584 as a threshold for assessing cross-hybridisation success in the IGH experiment (Figure 2). Genes with negative M-values smaller than -0.584 were deemed to show significantly reduced cross-hybridisation in *P. moluccensis* (Figure 2). These genes were, therefore, considered to be rapidly diverging between *P. moluccensis* and *D. rerio*. This approach identified 985 genes as rapidly diverging, approximately 6% of all genes assayed (Figure 3). Signal intensities for *P. moluccensis* were up to 58-fold lower than for *D. rerio*. 484 genes showed greater than two-fold signal intensity losses in *P. moluccensis* compared to *D. rerio* (Appendix 6).

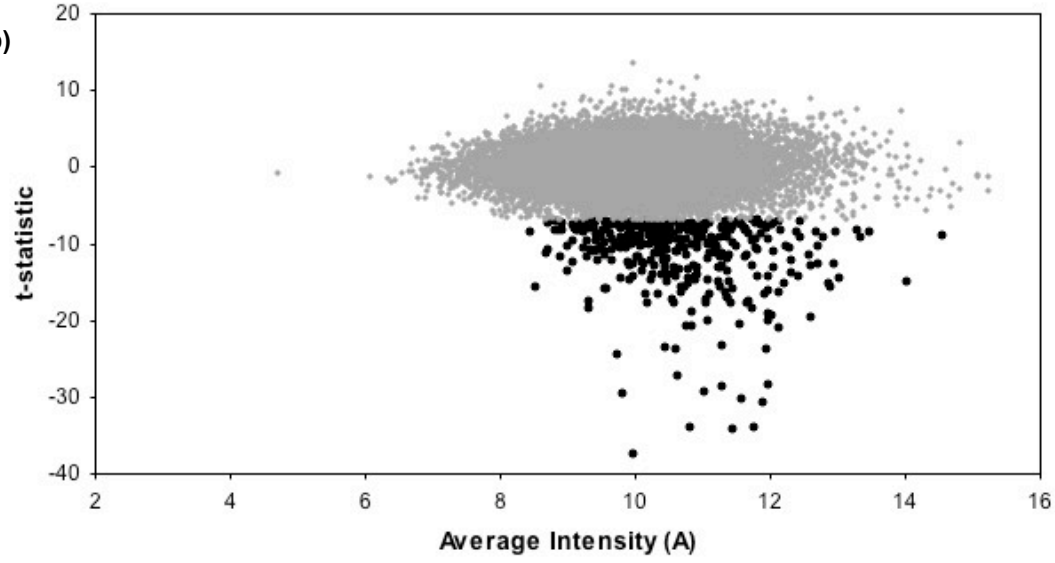
I have also tested the applicability of other approaches for the identification of rapidly diverging genes. Using the 99th percentile of t-values, 390 spots were identified as rapidly diverging between *P. moluccensis* and *D. rerio* (Figure 2 and Figure 3). Individual t-tests for significant differences in hybridisation signal between *P. moluccensis* and *D. rerio* on a gene-by-gene basis identified 48 spots as rapidly diverging between the two species (Figure 3). Comparison of the M-A and the t-A plots relating to these analyses showed a difference in the spread of these plots with the t-A plot showing a considerably reduced spread compared to the M-A plot (Figure 2). Rapidly diverging genes showed a trend for increased variability compared to genes that had shown significant cross-hybridisation in the inter-species genomic hybridisation experiment (Figure 4).

Figure 2 Log_2-fold change (M) versus signal intensity (A) of the inter-species genomic hybridisation experiment competitively hybridising *Danio rerio* and *Pomacentrus moluccensis* gDNA to a *D. rerio* microarray (a). Values of the t-statistic versus signal intensity (b). Genes identified as rapidly diverging (●) and genes showing evidence of significant sequence similarity (●) are indicated. For comparison, the M-A plot is shown with the genes marked that were not identified as rapidly diverging on the basis of t-statistics (●) and with the genes marked that were only identified as rapidly diverging by t-statistics (▲) (c).

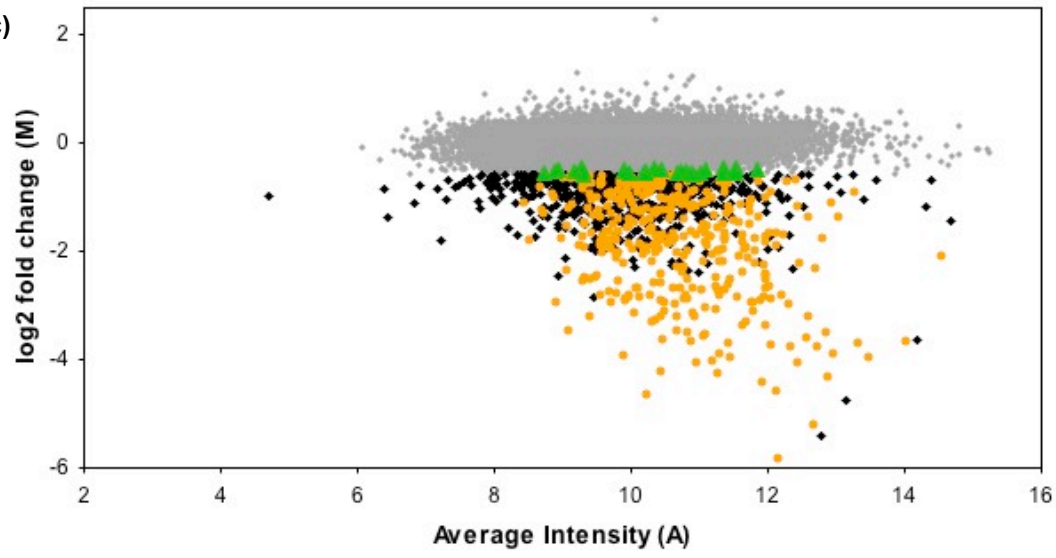
a)


log2fold change (M)
2
0
-2
-4
-6
2
4
6
8
10
12
14
16
Average Intensity (A)

b)


t-statistic
20
10
0
-10
-20
-30
-40
2
4
6
8
10
12
14
16
Average Intensity (A)

c)


log2 fold change (M)
2
0
-2
-4
-6
2
4
6
8
10
12
14
16
Average Intensity (A)

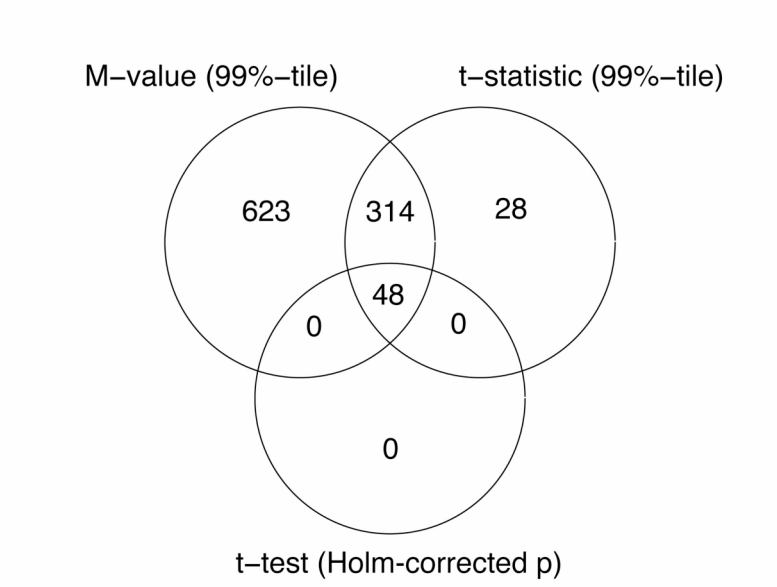


Figure 3 Venn diagram of the number of genes identified as rapidly diverging by three alternative methods of analysis: the 99th percentile of M-values, the 99th percentile of t-values, and individual t-tests with Holm-correction for multiple testing (please refer to text for further explanation).

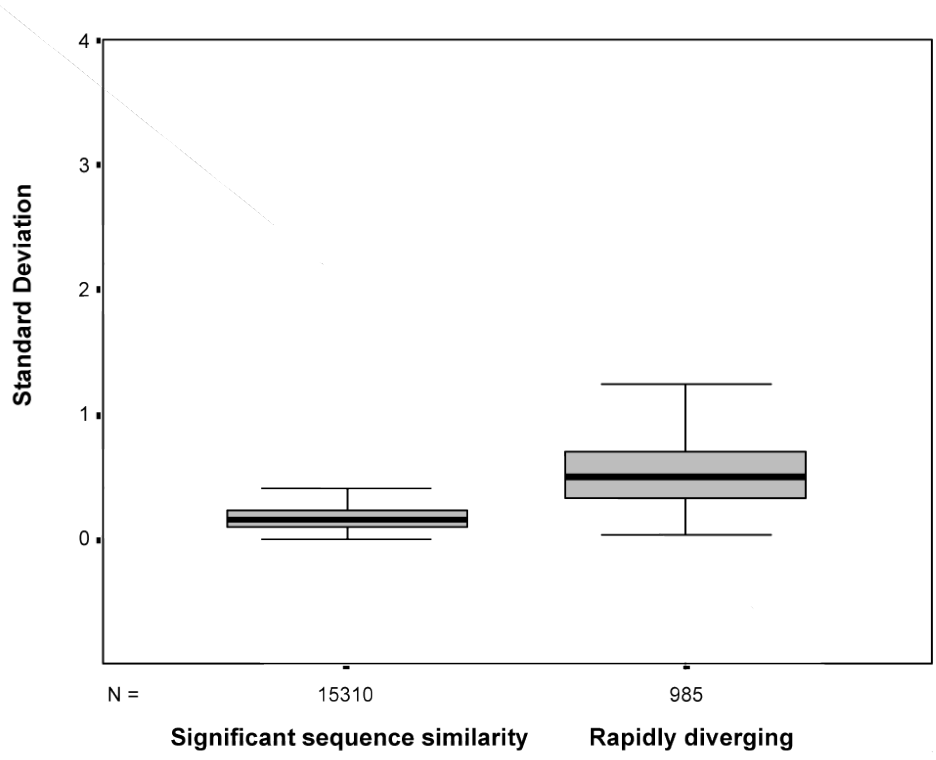


Figure 4 Boxplot of the standard deviation of M-values calculated for each gene across four biological replicates in the inter-genomic hybridisation experiment of *Pomacentrus moluccensis* and *Danio rerio*. Rapidly diverging genes showed significant differences in signal intensity between *P. moluccensis* and *D. rerio* when competitively hybridising genomic DNA to a *D. rerio* microarray. Genes that showed significant cross-hybridisation in these experiments were considered to share significant sequence similarity.

Functions of rapidly diverging genes

BLASTX of oligonucleotide sequences against the *D. rerio* protein database produced significant hits for 74 of the 985 rapidly diverging genes (Table 1, Appendix 6). Rapidly diverging genes belonged to a variety of gene functions (Figure 5a,b). Eighty-six percent of rapidly diverging genes identified by oligos in untranslated gene regions were of unknown gene function, while only fifty-one percent of rapidly diverging genes identified by oligos in coding gene regions were of unknown gene function (Figure 5c,d).

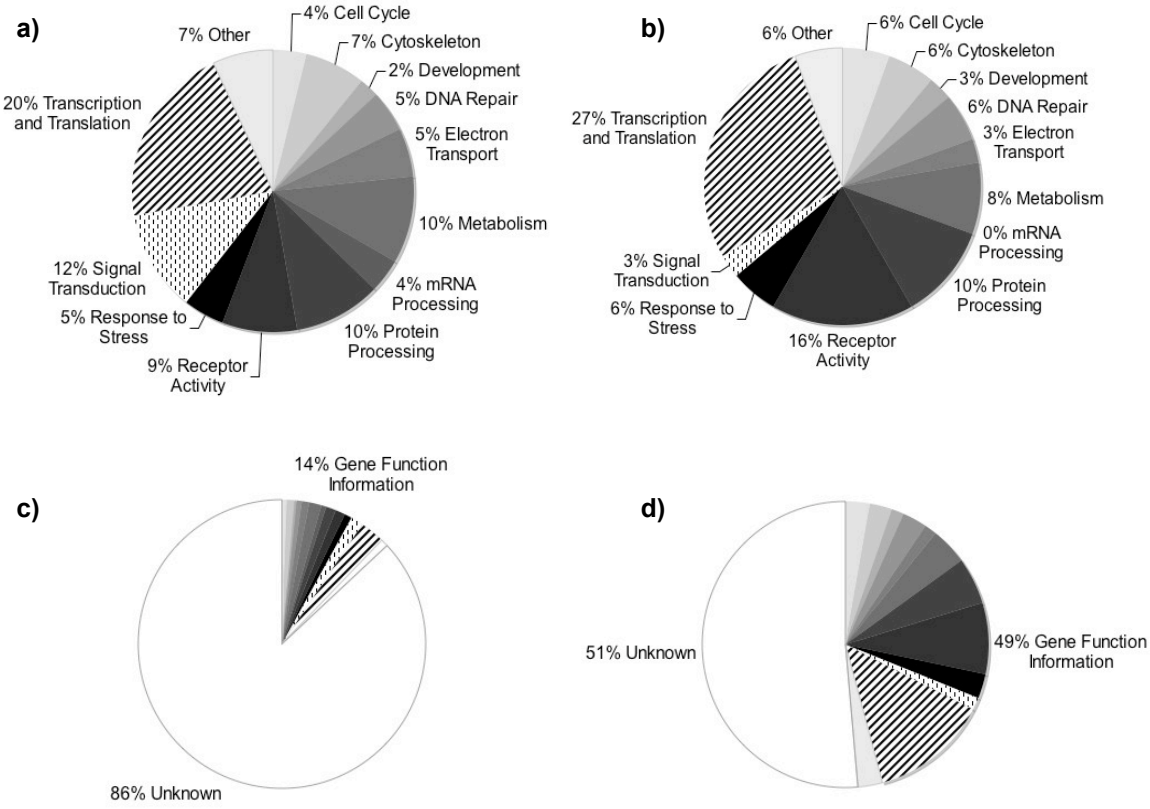


Figure 5 Gene ontologies of rapidly diverging genes with known gene function only (a,b). Rapidly diverging genes identified by oligos in untranslated gene regions (a), and coding gene regions (b). Gene ontologies of all rapidly diverging genes, identified by oligos in untranslated gene regions (c), and coding gene regions (d).

Table 1 Rapidly diverging candidate genes identified by inter-species genomic hybridisation experiments. These genes showed significant differences in hybridisation signal when *Danio rerio* and *Pomacentrus moluccensis* genomic DNA were competitively hybridised to a *D. rerio* microarray. Fold change refers to the difference in signal intensity obtained for *P. moluccensis* and *D. rerio* gDNA. Only genes for which information about putative gene function was available are shown.

GenBank ID	UniGene ID	Gene symbol	Putative Gene Identification	Gene Ontology	Fold Change	Oligo in coding region
Cell Cycle and Cell Adhesion						
AI601467	Dr.4363	cask	Calmodulin-dependent serine protein kinase	Cell Adhesion; Protein Kinase Activity	-1.53	
BM184447	Dr.14739	gmnn	Geminin, DNA replication inhibitor	Cell Cycle; Neg. Reg. of DNA Replication	-2.07	+
BI844156	Dr.2291	mcm2	Minichromosome maintenance deficient 2, mitotin	Cell Cycle; Transcription	-1.57	
AF265345	Dr.30271	TLK1	PKU-beta protein kinase	Cell Cycle; Response to DNA Damage Stimulus	-1.62	
Cytoskeleton						
AI942574	Dr.15424	arpc5a	Actin related protein 2/3 complex, subunit 5A	Regulation of Actin Filament Polymerization	-1.86	
BG728366	Dr.9511	FGL2	Fibrinogen-like protein 2	Fibrinogen Complex	-4.02	
BI878485	Dr.26692	ktn1	Kinectin 1	Microtubule-Based Movement	-1.56	
BI880252	Dr.16886		MGC53005 protein	Structural Constituent of Cytoskeleton	-1.55	
BI878304	Dr.16118		Microtubule-associated protein tau	Microtubule Stabilization	-1.66	+
BG728409	Dr.30602	myl9	Myosin, light polypeptide 9, regulatory	Regulation of Muscle Contraction	-6.49	
BI890773	Dr.20552	rdx	Radixin	Structural Constituent of Cytoskeleton	-1.65	
AW154375	Dr.2363	tagln2	Transgelin 2	Actin Binding	-1.60	+
AW128192	Dr.23476		zgc:66125	Microtubule Motor Activity	-1.77	
Development						
BI981137	Dr.17172	vangl1	Vang-like 1 (van gogh, Drosophila)	Development	-2.26	
AL590147		CTNND2	Catenin delta 2	Development	-6.04	+
AB006084	Dr.562	spon2a	Spondin 2a, extracellular matrix protein	Axon Guidance	-2.15	
DNA Repair						
BM101602	Dr.15882		Checkpoint 1 protein	DNA Damage Checkpoint	-1.51	+
BM182396	Dr.24937	cspg6	Chondroitin sulfate proteoglycan 6 (bamacan)	DNA Repair, Cell Cycle, Signal Transduction	-1.80	
BM024762	Dr.17357	h2afx	H2A histone family, member X	DNA Damage Checkpoint	-1.72	
AW154320	Dr.38362	HELLS	Helicase, lymphoid-specific	DNA Methylation; Anti-apoptosis	-1.54	+
BI842492	Dr.13952		Protein-tyrosine phosphatase	DNA Damage Checkpoint	-1.51	
AI667329		RFC5	Replication factor C subunit 5	DNA Replication; DNA Repair	-1.68	

Table 1 *cont.*

GenBank ID	UniGene ID	Gene symbol	Putative Gene Identification	Gene Ontology	Fold Change	Oligo in coding region
Electron Transport						
AI558444		Cyb5b	Cytochrome b5	Electron Transport	-1.83	
BI839328	Dr.24827	jup	Junction plakoglobin	Mitoch. Electron Transport, NADH to ubiquinone	-2.06	
BI865067	Dr.15779		NADH-ubiquinone oxidoreductase chain 1	Mitoch. Electron Transport, NADH to ubiquinone	-1.76	
BI880056	Dr.26950	pdc1	Phosducin 1	Electron Transport	-1.78	
BI885320	Dr.15025		Sb:cb825 protein	Electron Transport; Protein-Nucleus Import	-1.53	
AY054971	Dr.30224	smox	Spermine oxidase	Oxidoreductase; Electron Transport	-1.55	+
BI865578	Dr.30224	smox	Spermine oxidase	Oxidoreductase; Electron Transport	-2.45	
Metabolism						
AW232030		ACBP	Acyl-CoA-binding protein	Lipid Binding	-1.58	+
BE201971	Dr.25168	asns	Asparagine synthetase	Glutamine Metabolism	-6.60	
AI397087		DDC	Dopa decarboxylase	Carboxylic Acid Metabolism	-12.34	+
AI545421	Dr.3019	dspg3	Dermatan sulfate proteoglycan 3	Glycosaminoglycan Binding	-1.95	
BG303052	Dr.26560	elovl5	ELOVL family member 5	Very-Long-Chain Fatty Acid Metabolism	-4.16	
AI461355	Dr.37150		Galk2-prov protein	Galactose Metabolism	-1.56	
AW420718	Dr.28437	lgals9l1	Lectin, galactoside-binding, soluble, 9-like 1	Galactoside Binding	-1.64	
BG728832	Dr.43727		LOC398483 protein	Very-Long-Chain Fatty Acid Metabolism	-2.52	
AW420694	Dr.848		LOC398631 protein	L-serine Biosynthesis	-1.80	
BM095990	Dr.13990		Slc25a1-prov protein	Citrate Transporter Activity	-1.56	
AI958233		URKL1	Uridine kinase-like 1	Biosynthesis	-2.11	+
BG303586	Dr.25009	vg1	Vitellogenin 1	Lipid Transport	-3.02	
AW421066	Dr.17850		zgc:56592	L-Glutamate Transport	-2.19	
mRNA Processing						
AW175474	Dr.27187	rnpc1	RNA-binding region containing 1	RNA Binding	-1.65	
BG727249	Dr.28420	ddx5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	mRNA Catabolism, nonsense-mediated	-1.66	
BI877823	Dr.31567	pl10	Pl10	mRNA Processing	-1.79	
U62018	Dr.25770	elavl3	Embryonic lethal, abnormal vision, Drosophila-like 3	mRNA 3'-UTR binding; RNA Catabolism	-1.52	
BM005426	Dr.2043	auh	AU RNA binding protein	mRNA Catabolism, deadenylylation-dependent	-1.55	
Protein Processing						
AW280558	Dr.32494	PSMD4	26S proteasome non-ATPase regulatory subunit 4	Ubiquitin-Dependent Protein Catabolism	-1.61	+

Table 1 *cont.*

GenBank ID	UniGene ID	Gene symbol	Putative Gene Identification	Gene Ontology	Fold Change	Oligo in coding region
Protein Processing *cont.*						
AW019526		Ctrb1	Chymotrypsinogen B1	Proteolysis and Peptidolysis	-1.60	
AI793437	Dr.32382		Dual-spec. tyrosine-phosphorylation reg. kinase 2	Protein Kinase Activity	-1.72	
AW077976	Dr.10525	hectd1	HECT domain containing 1	Ubiquitin-dependent Protein Catabolism	-1.92	+
AI666982	Dr.25629	ivns1abpa	influenza virus NS1A binding protein a	Protein Binding	-7.12	
AI794113	Dr.25629	ivns1abpa	influenza virus NS1A binding protein a	Protein Binding	-8.55	
AW128332	Dr.18106	melk	Maternal embryonic leucine zipper kinase	Protein Kinase Activity	-1.72	+
BI980448	Dr.9029		MIF2 suppressor	Protein-Nucleus Import	-1.50	
BI888687	Dr.4179	ptp4a1	Protein tyrosine phosphatase type IVA, member 1	Protein Amino Acid Dephosphorylation	-1.77	
BI708483	Dr.4091	rpl8	Ribosomal protein L8	Protein Biosynthesis	-1.51	
AW154568	Dr.12666	rnf146	Ring finger protein 146	Protein Binding, Ion Binding	-2.35	
BI840104	Dr.30110	slc7a3	Solute carrier family 7, member 3	Amino Acid Transport	-1.69	+
BM036448	Dr.14973	stx5al	Syntaxin 5A, like	Intracellular Protein Transport	-1.61	
Receptor Activity						
AY048971		adra2a	Alpha2A-adrenergic receptor	MAPKKK Cascade; Receptor Activity	-3.27	+
U93460		COUP-TF	Chicken ovalbumin upstr. promoter trans. factor 3	Steroid Hormone Receptor Activity	-1.51	+
AF359430	Dr.8680	cldnj	Claudin j	Transmembrane Receptor Activity	-3.10	
AW232855	Dr.5765	rtk6	Eph-like receptor tyrosine kinase 6	Ephrin Receptor Activity; Signaling Pathway	-2.34	
U93466		ERR2	Estrogen-related receptor gamma	Receptor activity; Transcription Factor Activity	-1.69	+
U49409		fzd8a	Frizzled 8a protein	Receptor Activity; Cell Motility	-2.05	+
BI891150	Dr.6631	kdelr2	KDEL endoplasmic retic. prot. retention receptor 2	Receptor Activity	-1.55	
AL591442			Major histocompatibility class II protein	G-protein Coupled Receptor Activity	-2.37	
BG306469	Dr.36420	nrp2b	Neuropilin 2b	Cell Adhesion; Receptor Activity	-2.83	+
U93481		reverbb2	Rev-Erb beta 2	Thyroid Hormone Receptor Activity	-1.64	+
BI883230	Dr.21056		zgc:77734	Receptor Binding	-1.80	
Response to Stress and Chaperone Activity						
AI721437	Dr.15608	bag2	BCL2-associated athanogene 2	Chaperone Activity; Protein Folding	-1.58	
AF210640		hsp70	Hsp70	Protein Folding	-1.90	+
AW019171	Dr.1995		Wu:fb60f06 protein	Response to Pest, Pathogen or Parasite	-1.78	+
BI672376	Dr.2813		zgc:55461	Chaperone Activity	-4.93	
BI850032	Dr.6391		zgc:63524	Chaperone Activity	-1.78	
BI704996			zgc:63896	Response to Stress	-1.87	

Table 1 *cont.*

GenBank ID	UniGene ID	Gene symbol	Putative Gene Identification	Gene Ontology	Fold Change	Oligo in coding region
Signal Transduction						
BI891199	Dr.19510	adam10	A disintegrin and metalloprotease domain 10	Cell-Cell Signaling	-1.60	
BI428482	Dr.20261	anp32a	Acidic nuclear phosphoprotein 32 family, member A	Intracellular Signaling Cascade	-1.61	
BI979187	Dr.32157	ar	Androgen receptor	Cell-Cell Signaling	-1.63	
AB032262	Dr.8294	axin1	Axin 1	Apoptosis; Frizzled Signaling Pathway	-1.61	
AF359430	Dr.8680	cldnj	Claudin j	Transmembrane Receptor Activity	-3.10	
BI429881	Dr.20115	cfl1	Cofilin 1 (non-muscle)	Rho-Protein Signaling Transduction	-2.29	
BI877949	Dr.27120	guca1e	Guanylate cyclase activator 1e	Signal Transduction	-1.84	
AI584734	Dr.4736	mpp1	Membrane protein, palmitoylated 1	Signal Transduction	-25.07	
BF158097	Dr.4736	mpp1	Membrane protein, palmitoylated 1	Signal Transduction	-7.79	
BG727421	Dr.4736	mpp1	Membrane protein, palmitoylated 1	Signal Transduction	-19.13	
BG307787	Dr.354	rho	Rhodopsin	Rhodopsin Mediated Signaling	-4.32	
AW018967	Dr.36	shh	Sonic hedgehog	Cell-Cell Signaling	-2.79	
BG305930	Dr.36	shh	Sonic hedgehog	Cell-Cell Signaling	-15.25	
BI845415	Dr.28615	snx12	Sorting nexin 12	Signaling Cascade; Protein Binding	-2.25	
AF231014		Tradd	Tnfrsf1a-associated via death domain	Signal Transduction	-1.74	+
Transcription and Translation						
AI558513		CNOT-3	CCR4-NOT transcription complex subunit 3	Regulation of Transcription, DNA-dependent	-1.55	+
U03876	Dr.429	dlx5a	Distal-less homeobox gene 5a	Transcription Factor Activity	-1.78	+
AI884034		ERCC8	Excision repair cross-complementing rodent repair deficiency, complementation group 8	Transcription-Coupled Nucleotide-Excision Repair	-1.63	+
AF052249	Dr.590	foxd3	Forkhead box D3	Transcription Factor Activity	-1.52	
AW344193	Dr.9123	foxk1	Forkhead box K1	Transcription Factor Activity	-1.56	
U18312	Dr.356	gata2	GATA-binding protein 2	Transcription Factor Activity	-1.55	
BM095922	Dr.11866	gtf2b	General transcription factor IIB	Regulation of Transcription, DNA-dependent	-1.68	
AW826795	Dr.40048	HIST4H4	Histone H4	DNA Binding	-1.54	
Y14533	Dr.5726	hoxb7a	Homeo box B7a	Transcription Factor Activity	-1.55	+
AF071258		hoxc11a	Homeobox protein hoxc11a	Regulation of Transcription, DNA-dependent	-2.01	+
BI889237	Dr.10302		HTATIP protein	Transcription Coactivator Activity	-1.53	+
AB022286	Dr.23470	krml2.2	Kreisler maf-related leucine zipper homolog 2.2	Regulation of Transcription	-1.52	
X65062		msh-D	Muscle segment homeobox D	Regulation of Transcription, DNA-dependent	-2.09	+
U50563	Dr.41	msxe	Muscle segment homeobox E	Regulation of Transcription, DNA-dependent	-1.79	

Table 1 *cont.*

GenBank ID	UniGene ID	Gene symbol	Putative Gene Identification	Gene Ontology	Fold Change	Oligo in coding region
Transcription and Translation *cont.*						
BI880399	Dr.266	mef2a	Myocyte enhancer factor 2a	Transcription	-1.63	
AI793736	Dr.2883	polr2gl	Polymerase (RNA) II polypeptide G-like	Transcription-Coupled Nucleotide-Excision Repair	-6.99	
BM082684	Dr.25643	pias2	Protein inhibitor of activated STAT, 2	Transcription Factor Activity	-1.52	+
AW343823	Dr.2562	prkcbp1l	Protein kinase C binding protein 1, like	Protein Binding; Regulation of Transcription	-1.50	
U93479		RAR-gamma	Retinoic acid receptor gamma	Regulation of Transcription, DNA-dependent	-1.53	+
BG985543	Dr.33499	snai2	Slug protein	Neg. Reg. of Transcription from Pol II Promoter	-1.53	
U24225	Dr.625	snai1b	Snail homolog 1b (Drosophila)	Transcription Factor Activity	-1.53	
AF253325	Dr.8231	tbx20	T-box 20	Transcription Factor Activity	-1.63	
AI965244	Dr.921	timeless	Timeless homolog (Drosophila)	Transcription; Development	-1.64	+
AI957437	Dr.36535	EEF1D	Translational elongation factor 1 delta	Translation Elongation Factor Activity	-1.80	
AW233578	Dr.10946	ubtf	Upstream binding transcr. factor, RNA polymerase I	Transcription	-1.73	
AI723126	Dr.3216		zgc:77008	Phosphorylase Activity; DNA Modification	-1.62	
AW019173	Dr.33575		zgc:77008	Phosphorylase Activity; DNA Modification	-1.74	
Other						
BI866388	Dr.10209	apaf1	Apoptotic protease activating factor	Regulation of Apoptosis	-1.67	
BI879509	Dr.30206	cacna1db	Calcium channel, L type, alpha 1D subunit, b	Ion Transport	-3.16	
BI981073	Dr.26356	hbae1	Hemoglobin alpha embryonic-1	Oxygen Transport; Heme Binding	-1.60	
AW115757	Dr.31376	hpx	Hemopexin	Heme Transport	-1.64	+
AI617291	Dr.1330	sepp1a	Selenoprotein P, plasma, 1a	Selenium Binding	-2.83	
AW232016	Dr.34427	slc17a6l	Solute carrier family 17, member 6 like	Transport	-2.14	
AW422249	Dr.4765	txndc9	Thioredoxin domain containing 9	ATP Binding	-2.22	+
AI723219	Dr.345	unc45r	Unc-45 (C. elegans) related	Binding	-3.51	
AW305605	Dr.27155	vamp2	Vesicle-associated membrane protein 2	Regulation of Exocytosis	-1.51	

Of the total of 985 rapidly diverging genes, only 129 could be identified and annotated with putative gene function (Table 1). The remainder are currently unidentified gene loci and expressed sequence tags (ESTs). Only rapidly diverging genes identified by oligos in the untranslated gene region included genes with functions in mRNA processing, associated with four percent of genes in this group (Figure 5a). Amongst these genes were *rnpc1*, *ddx5*, *pl10*, *elavl3*, and *auh* (Table 1). Over-representation analysis of gene functions produced no significant results. However, several gene ontologies showed a trend for over-representation based on raw p-values before FDR-correction (Table 2). Amongst these were several gene functions related to mRNA and DNA processing, e.g. 'DNA modification', 'mRNA 3' UTR binding', 'mRNA catabolism', and 'DNA damage checkpoint' (Table 2). Rapidly diverging genes identified by oligos in the untranslated regions showed a higher proportion of genes with function in signal transduction than rapidly diverging genes identified by oligos in coding regions (Figure 5a,b). Among rapidly diverging genes with function in signal transduction were *adam10*, *anp32a*, *axin1*, *cldhnj, cfl1, guca1e, mpp1, rho, shh,* and *snx12* (Table 1). The genes *mpp1* and *shh* were identified as rapidly diverging by multiple oligos on the array, representing different parts of the untranslated gene region of these genes. Other genes identified in this way had function in DNA repair, e.g. *cspg6*, *h2afx*, protein-tyrosine phosphatase, and *RFC5* (Table 1). The major histocompatibility class II gene was also identified as rapidly diverging in this study.

Rapidly diverging candidate genes identified by oligos in coding regions showed a higher proportion of genes with receptor activity and function in transcription and translation (Figure 5b). Genes in this group showed significant over-representation of the functions 'thyroid hormone receptor activity', 'steroid hormone receptor activity', 'transcription coactivator activity', 'skeletal development', 'negative regulation of apoptosis', and 'cell-cell signalling' (Table 2). Among these are genes with receptor activity were *adra2a, COUP-TF, ERR2, fzd8a, nrp2b*, *RAR-gamma* and *reverbb2* (Table 1). Rapidly diverging genes with function in transcription and translation included *CNOT-3, ERCC8*, HTATIP protein, and the homeobox genes *dlx5a, hoxb7a, hoxc11a*, and *msh-D*. Other rapidly diverging candidate genes identified in this way included the molecular chaperone *hsp70*, the signal transduction gene *Tradd*, the transcription factors *snai1b* and *timeless*, the heme

transport protein *hemopexin*, and *PSMD4* and *hectd1*, two genes with function in ubiquitin-dependent protein catabolism (Table 1).

Table 2 Gene ontologies with significant over-representation amongst rapidly diverging genes as compared to the total of genes represented on the microarray. Rapidly diverging candidate genes were split into those for which the oligo was in the untranslated region of the gene and those for which the oligo was in the coding region. Only gene functions with two or more genes amongst identified candidates are shown. Raw p-values are p-values before correction for multiple testing.

Gene ontology	ID	Number of probes on array	Number of genes on array	Number of genes amongst can-didates	Raw p-value	FDR-corr. p-value
Gene loci with oligos in the untranslated region of the gene						
mRNA and DNA processing						
DNA modification	GO:0006304	5	5	2	0.002	1.000
mRNA 3'-UTR binding	GO:0003730	7	7	2	0.006	1.000
mRNA catabolism	GO:0006402	16	16	3	0.012	1.000
DNA damage checkpoint	GO:0000077	24	24	3	0.048	1.000
Other gene functions						
Coenzyme binding	GO:0050662	7	7	2	0.006	1.000
Prenylated protein tyrosine phosphatase	GO:0004727	8	8	2	0.009	1.000
Phosphorylase activity	GO:0004645	9	9	2	0.023	1.000
Wnt receptor activity	GO:0042813	16	14	2	0.044	1.000
Fatty acid biosynthesis	GO:0006633	14	14	2	0.044	1.000
Gene loci with oligos in the coding region of the gene						
Receptor activity						
Thyroid hormone receptor activity	GO:0004887	18	18	2	0.000	0.082
Steroid hormone receptor activity	GO:0003707	51	51	3	0.000	0.045
Other gene functions						
Transcription coactivator activity	GO:0003713	93	92	3	0.001	0.055
Skeletal development	GO:0001501	102	92	3	0.001	0.052
Negative regulation of apoptosis	GO:0043066	84	79	2	0.004	0.093
Cell-cell signaling	GO:0007267	94	87	2	0.005	0.099

Evolutionary rates of these rapidly diverging genes in other teleosts

Rapidly diverging genes identified in the IGH experiments had on average, a greater number of paralogs in *D. rerio* than genes that showed significant cross-hybridisation in the IGH experiments (Table 3). However, the number of paralogs in *D. rerio* differed greatly depending on the type of gene. For example, the receptor gene *reverbb2* had 15 paralogs in *D. rerio*, while the transcription factor *Tradd* had none. *Danio rerio* consistently showed the longest branch lengths in the unrooted tree of *D. rerio*, *T. rubripes* and *G. aculeatus* in line with known phylogenetic relationships and

relative ages of these taxa. Average branch lengths in all three taxa and average total branch lengths summed across taxa were greater for rapidly diverging candidate genes than genes that showed evidence of significant sequence similarity in the IGH experiment (Table 3). However, as for the number of paralogs, tree branch lengths for individual genes differed greatly depending on the type of gene (Figure 6). For example, *msh-D*, *reverbb2*, *gmnn*, and *Tradd* showed the longest total branch lengths and thus greatest rates of sequence evolution amongst rapidly diverging genes. In contrast, *fzd8a* and *hoxc11a*, which were also identified as rapidly diverging in the IGH experiment, showed shorter branch lengths than the average total branch lengths of genes that showed evidence of significant sequence similarity in the IGH experiment. Similarly, two genes that were identified as sharing significant sequence similarity in *D. rerio* and *P. moluccensis*, *irak4* and *alg1*, showed unusually long branch lengths in the *in silico* study. If we removed these two genes, the differences in average branch lengths for *D. rerio* and *G. aculeatus* and the differences in total branch lengths between rapidly diverging genes and genes with significant sequence similarity are significant at $\alpha < 0.05$.

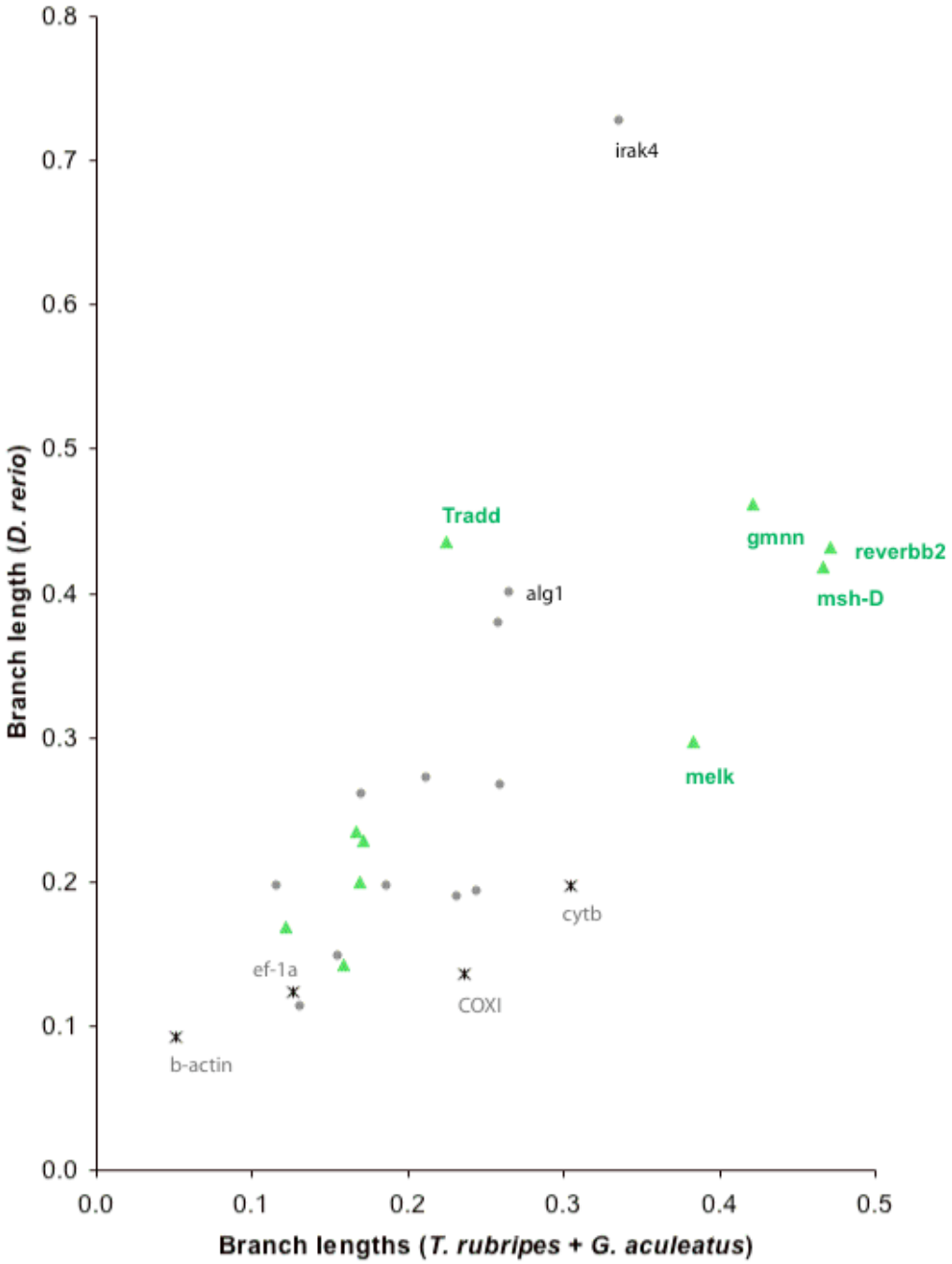


Figure 6 Tree branch lengths in *Danio rerio* versus the sum of branch lengths in *Takifugu rubripes* and *Gasterosteus aculeatus*. Rapidly diverging genes (▲) were identified by inter-species genomic hybridisation experiments of *D. rerio* and *Pomacentrus moluccensis*. For comparison, genes showing evidence of significant sequence similarity between *D. rerio* and *P. moluccensis* (●), and commonly used phylogenetic markers (x) are also given: cytb: cytochrome b; COXI: cytochrome oxidase I; ef-1a: elongation factor 1α; b-actin: β-actin. Please note that for simplicity only some genes names are marked. See Table 1 for gene name abbreviations used.

Table 3 Branch lengths in the unrooted tree of orthologous DNA sequences of *Danio rerio*, *Takifugu rubripes* and *Gasterosteus aculeatus*. The Ensembl Gene IDs of orthologous DNA sequences used for calculation of tree branch lengths are given. Rapidly diverging candidate genes were identified by inter-species genomic hybridisation experiments, competitively hybridising gDNA from the coral reef fish *Pomacentrus moluccensis* and gDNA from *D. rerio* to a *D. rerio* microarray. These genes are contrasted against genes that showed equal hybridisation signals of *P. moluccensis* and *D. rerio* in the IGH experiment and are thus expected to share significant sequence similarity in the two taxa. The difference in hybridisation signals between *P. moluccensis* and *D. rerio* in the IGH experiments and on the basis of which rapidly diverging candidate genes were identified is provided as IGH fold change.

GenBank ID	Gene symbol	Putative Gene Identification	Ensembl Gene IDs of orthologs			Number of paralogs in *D.rerio*	IGH fold change
			D. rerio	*T. rubripes*	*G. aculeatus*		
Rapidly diverging genes							
AI958233	URKL1	Uridine kinase	ENSDARG00000001686	NEWSINFRUG00000121541	ENSGACG00000012022	8	-2.11
X65062	msh-D	Msh-related homeobox	ENSDARG00000006982	NEWSINFRUG00000154539	ENSGACG00000020715	5	-2.09
U93481	reverbb2	ReverbB gene	ENSDARG00000009594	NEWSINFRUG00000121364	ENSGACG00000007986	15	-1.64
AI793761	thbs4	Thrombospondin 4	ENSDARG00000020072	NEWSINFRUG00000163207	ENSGACG00000016162	6	-1.67
AF071258	hoxc11a	Homeobox protein hoxc11a	ENSDARG00000028655	NEWSINFRUG00000149367	ENSGACG00000009392	4	-2.01
AW128332	melk	Maternal embryonic leucine zipper kinase	ENSDARG00000030759	NEWSINFRUG00000140679	ENSGACG00000016110	0	-1.72
BM184447	gmnn	Geminin, DNA replication inhibitor	ENSDARG00000035957	NEWSINFRUG00000154631	ENSGACG00000004689	0	-2.07
AF231014	Tradd	Death domain-containing adaptor molecule	ENSDARG00000036057	NEWSINFRUG00000126483	ENSGACG00000015872	0	-1.74
AY054971	smox	Smox protein	ENSDARG00000036967	NEWSINFRUG00000154633	ENSGACG00000017615	2	-1.55
U49409	fzd8a	Frizzled 8a protein	ENSDARG00000045444	NEWSINFRUG00000139617	ENSGACG00000001776	10	-2.05
Genes with significant sequence similarity							
AF375872	meis2.2	Myeloid ecotropic viral integration site 2.2	ENSDARG00000042361	NEWSINFRUG00000137931	ENSGACG00000010527	7	-1.04
AB041734	dvl3	Dishevelled, dsh homolog 3	ENSDARG00000015707	NEWSINFRUG00000146449	ENSGACG00000015451	5	-1.03
AW423082	irak4	Interleukin-1 receptor-associated kinase 4	ENSDARG00000010657	NEWSINFRUG00000122284	ENSGACG00000004050	0	-1.03
AW170913	alg1	Beta-1,4-mannosyltransferase	ENSDARG00000054963	NEWSINFRUG00000122557	ENSGACG00000014943	0	-1.03
AW133635	abce1	ATP-binding cassette, sub-family E (OABP), member 1	ENSDARG00000007216	NEWSINFRUG00000151650	ENSGACG00000018611	0	-1.04
BI891332	vil2	Villin 2	ENSDARG00000025091	NEWSINFRUG00000148580	ENSGACG00000010010	7	-1.03
BG302837	rhpn2	Rhophilin-2 (GTP-Rho-binding protein 2)	ENSDARG00000014577	NEWSINFRUG00000143237	ENSGACG00000014868	1	-1.03
BM182278	sars2	Seryl-tRNA synthetase 2	ENSDARG00000020115	NEWSINFRUG00000148538	ENSGACG00000005820	0	-1.03
AI667665	fts	Fused toes homolog	ENSDARG00000026862	NEWSINFRUG00000137681	ENSGACG00000016685	0	-1.04
BG985785	glrx5	Glutaredoxin 5	ENSDARG00000043665	NEWSINFRUG00000145950	ENSGACG00000004913	2	-1.04
AW594981	uhrf1	Ubiquitin-like, contain. PHD and RING finger domains, 1	ENSDARG00000009946	NEWSINFRUG00000152297	ENSGACG00000008745	0	-1.03
BI878078	carm1	Coactivator-associated arginine methyltransferase 1	ENSDARG00000018698	NEWSINFRUG00000147097	ENSGACG00000005531	1	-1.03

Table 3 *cont.*

GenBank ID	Gene symbol	Putative Gene Identification	Branch length *D. rerio*	Branch length *T. rubripes*	Branch length *G. aculeatus*	Total branch length
Rapidly diverging genes						
AI958233	URKL1	Uridine kinase	0.200	0.076	0.092	0.369
X65062	msh-D	Msh-related homeobox	0.419	0.249	0.217	0.885
U93481	reverbb2	ReverbB gene	0.432	0.262	0.209	0.904
AI793761	thbs4	Thrombospondin 4	0.235	0.100	0.067	0.401
AF071258	hoxc11a	Homeobox protein hoxc11a	0.169	0.071	0.051	0.291
AW128332	melk	Maternal embryonic leucine zipper kinase	0.298	0.186	0.197	0.680
BM184447	gmnn	Geminin, DNA replication inhibitor	0.462	0.179	0.242	0.882
AF231014	Tradd	Death domain-containing adaptor molecule	0.435	0.096	0.129	0.660
AY054971	smox	Smox protein	0.229	0.108	0.063	0.401
U49409	fzd8a	Frizzled 8a protein	0.142	0.074	0.086	0.301
		Mean	**0.302**	**0.140**	**0.135**	**0.577**
Genes with significant sequence similarity						
AF375872	meis2.2	Myeloid ecotropic viral integration site 2.2	0.113	0.102	0.029	0.244
AB041734	dvl3	Dishevelled, dsh homolog 3	0.261	0.105	0.065	0.431
AW423082	irak4	Interleukin-1 receptor-associated kinase 4	0.728	0.182	0.154	1.064
AW170913	alg1	Beta-1,4-mannosyltransferase	0.401	0.145	0.119	0.666
AW133635	abce1	ATP-binding cassette, sub-family E (OABP), member 1	0.148	0.087	0.068	0.303
BI891332	vil2	Villin 2	0.190	0.115	0.116	0.421
BG302837	rhpn2	Rhophilin-2 (GTP-Rho-binding protein 2)	0.272	0.113	0.099	0.484
BM182278	sars2	Seryl-tRNA synthetase 2	0.380	0.132	0.126	0.638
AI667665	fts	Fused toes homolog	0.197	0.109	0.077	0.383
BG985785	glrx5	Glutaredoxin 5	0.268	0.106	0.153	0.527
AW594981	uhrf1	Ubiquitin-like, contain. PHD and RING finger domains, 1	0.194	0.142	0.103	0.438
BI878078	carm1	Coactivator-associated arginine methyltransferase 1	0.197	0.064	0.052	0.313
		Mean	**0.279**	**0.117**	**0.097**	**0.493**

Discussion

This is the first study to use inter-species genomic hybridisation (IGH) experiments to identify rapidly diverging genes in teleosts and I have presented a new approach for their analysis. IGH experiments on a *D. rerio* 65mers-oligonucleotide micorarray identified 985 rapidly diverging candidate genes between the coral reef fish *P. moluccensis* and *D. rerio*. Rapidly diverging genes showed a significant over-representation of gene functions related to receptor activity, transcription coactivator activity, and cell-cell signalling. Evolutionary rates at rapidly diverging gene loci were also estimated in other teleost lineages using orthologous sequences to *D. rerio* from *T. rubripes* and *G. aculeatus*. In line with expectations, average branch lengths were longer for rapidly diverging candidate genes than for a randomly chosen set of genes, which showed evidence of significant sequence similarity in the IGH experiment. These comparative genomic data indicate that the IGH experiment successfully identified rapidly diverging genes and that these genes showed accelerated rates of sequence evolution across multiple teleost lineages. Rapidly diverging genes identified in this study include the homeobox gene *msh-D*, the receptor gene *reverbb2*, the cell cycle gene *gmnn*, and the signal transduction gene *Tradd*.

Analysis of inter-species genomic hybridisation experiments to identify rapidly diverging genes

I have presented here a new approach for the analysis of inter-species genomic hybridisation experiments. This approach takes advantage of the skewed distribution of M-values in inter-species genomic hybridisation experiments to estimate the random noise of the experiment and to determine which genes show unusually reduced cross-hybridisation. The approach presented here identifies those genes that show more extreme M-values than 99 percent of genes, which represent random noise, and hence genes identified in this way were considered rapidly diverging.

I also compared this method for identifying rapidly diverging genes to other approaches based on evaluating the values of the t-statistic. When calculating t-statistics the variability between biological replicates is taken into account: the

greater the variability between replicates the smaller the value of the t-statistic. In cases where variability increases as M-values become more extreme, this can lead to a reduction in the spread of the t-A plot as compared to the M-A plot. In this study, genes with very negative M-values displayed a tendency for greater variability between replicates (Figure 4), potentially explaining the reduction in spread of the t-A plot. In microarray experiments, variability is generally greater for spots with low signal intensity and spots with extreme M-values (Kane *et al.* 2000, Moody *et al.* 2002). In comparative genomic hybridisation experiments, signal intensity and M-values largely depend on sequence similarity between taxa (Wu *et al.* 2001, Hinchliffe *et al.* 2003, Brunelle *et al.* 2004, Le Quere *et al.* 2006). Genes with large sequence differences between taxa will produce lower signal intensities and more extreme M-values. However, as discussed above, genes with extreme M-values are expected to show increased variability between replicates. An analytical approach based on t-statistics would, therefore, penalise genes with the greatest sequence differences. Such genes may potentially be declared as sharing significant sequence similarity because the values of the t-statistic are down-weighed by the large variances relating to the measurement of the most divergent genes. For these reasons, approaches based on t-statistics do not appear suitable for the analysis of inter-species genomic hybridisation experiments.

The approach presented in this study benefits from the use of a relatively large microarray so that the random noise in the inter-species genomic hybridisation experiment can be reliably estimated. I used here a 65mers-oligonucleotide microarray representing approximately 16K genes, one of the largest teleost microarrays presently available. The approach presented here somewhat resembles that of the GACK algorithm, which compares the distribution of observed hybridisation signals to a normal distribution and determines absence/ presence calls from the ratio of expected to observed number of presence calls, the estimated probability of presence (EPP) (Kim *et al.* 2002). However, the analyses presented here do not rely on comparison of the actual signal distribution to a normal distribution, but instead uses positive M-values to calculate the random noise in the data. Other approaches require the availability of whole-genome sequence data for closely related taxa (Porwollik *et al.* 2002, Le Quere *et al.* 2006) and are thus only applicable to selected species.

Relationship between rates of evolution and gene function

There was a significant relationship between gene function and rates of sequence evolution for rapidly diverging genes identified by oligos in coding regions. These genes showed significant over-representation of the gene functions 'thyroid receptor activity', 'steroid receptor activity', and 'cell-cell signalling' and included the receptor genes *adra2a, COUP-TF, ERR2, fzd8a, nrp2b*, *RAR-gamma*, and *reverbb2*. Rapidly diverging genes in closely related species of bacteria, archaea and mammals have also been shown to be enriched for membrane proteins and receptor genes (Jordan *et al*. 2001, Hinchliffe *et al.* 2003). Since receptor proteins function in signal transduction and cellular responses to external stimuli, accelerated rates of sequence evolution at these gene loci may reflect interaction with and adaptation to different environments.

Many of the other rapidly diverging genes identified in this study were transcription factors and homeobox genes, e.g. *CNOT-3, dlx5a, ERCC8, hoxb7a, hoxc11a*, HTATIP protein, and *msh-D*. As a result, the gene functions 'transcription coactivator function' and 'skeletal development' were significantly over-represented amongst rapidly diverging genes. Hox gene clusters have undergone independent duplication events throughout animal and vertebrate evolution (Holland *et al*. 1994, Amores *et al*. 1998, Pollard and Holland 2000, Martinez and Amemiya 2002, Chourrout *et al.* 2006). There is also evidence that homeobox genes show accelerated rates of sequence evolution in different lineages (Malaga-Trillo and Meyer 2001, Wang and Zhang 2004, Edvardsen *et al*. 2005). Teleosts have experienced an additional Hox cluster duplication event early in their evolution (Crow *et al*. 2006). This additional Hox cluster duplication event may have resulted in a need to reduce genetic redundancies and as a consequence, duplicate homeobox genes may have been under selective pressure to diverge rapidly (Malaga-Trillo and Meyer 2001). The data presented here show accelerated rates of sequence evolution in teleost homeobox genes, possibly supporting the hypothesis of genomic instability after the Hox cluster duplication event.

I did not detect a significant over-representation of stress-related gene functions amongst rapidly diverging genes. However, four of the identified rapidly diverging genes have chaperone activity and function in protein folding, *bag2, hsp70*,

zgc:55461, and zgc:63524. Two further genes, wu:fb60f06 and zgc:63896, also have stress-related gene functions. In addition, *PSMD4*, a protein involved in ubiquitin-dependent protein catabolism, and *txndc9*, a gene containing the thioredoxin domain, were identified as rapidly diverging. Adaptation to new environments may be expected to involve the evolution of different environmental tolerances and stress resistances. As a consequence, positive selection may act on genes with stress-related functions during the process of adaptation to new environments, resulting in these genes diverging at accelerated rates. Despite the potential role of genes with stress-related functions in adaptive evolution, there have been no reports of an enrichment of stress-related gene functions amongst rapidly diverging genes. There are reports, however, of individual genes with stress-related functions evolving at accelerated rates. For example, in the fungus *Paxillus involutus* genes with function in redox balance, e.g. thioredoxin, peroxiredoxin, n-alkane inducible cytochrome P450, and glutathione S-transferase III homolog, were identified as evolving at accelerated rates (Le Quere *et al*. 2006). Similarly, duplicated chaperone genes in *Mycobacterium* evolve at accelerated rates (Hughes 1993). The molecular chaperone *hsp70* was identified as rapidly diverging in this study; however, *hsp70* has at least twelve paralogs in *D. rerio*. It is possible that the results of the inter-species genomic hybridisation experiment were confounded by the presence of different numbers of paralogs of *hsp70* in *D. rerio* and *P. moluccensis*. Irrespective of the source of these differences, these data may indicate differences in the heat shock response system of *D. rerio,* a cooler water species, and the warm-adapted coral reef fish, *P. moluccensis*. Finally, genes with stress-related functions typically exhibit a range of functions. These genes may, therefore, be under evolutionary constraints due to multiple selective pressures operating upon them.

Rapidly diverging genes identified by oligos in untranslated gene regions showed a trend for over-representation of the gene functions ‘DNA modification’, ‘mRNA 3’ UTR binding’, ‘mRNA catabolism’, and ‘DNA damage checkpoint’. This group included the checkpoint 1 protein, *cspg6, h2afx, HELLS*, protein-tyrosine phosphatase, *RFC5*, *rnpc1, ddx5, pl10, elavl3*, and *auh*. As stress has been associated with increased levels of DNA damage and mRNA catabolism (Fischman *et al*. 1996, Lesser 2006), these data may indicate that stress and adaptation to different environments exert indirect selective pressures on DNA sequence evolution by

adjusting DNA repair systems and mRNA metabolism. However, further work is required to test such relationships.

Comparative genomics of teleosts

In this study I took advantage of currently available genome sequence data to test whether accelerated rates of sequence evolution identified by means of inter-species genomic hybridisation experiments could be confirmed using bioinformatic approaches. Orthologous sequences in *D. rerio, T. rubripes*, and *G. aculeatus* showed longer average branch lengths for genes identified as rapidly diverging in the IGH experiments compared to genes that showed evidence of significant sequence similarity. These results suggest that IGH experiments can be used to guide the identification of rapidly diverging genes. Some of the rapidly diverging genes identified in this study showed evidence for accelerated rates of sequence evolution in multiple teleost lineages. However, there was substantial variation in branch lengths between different genes. Two genes, *irak4* and *alg1,* showed unusually long branch lengths and thus high rates of sequence evolution, despite the fact that the IGH experiment indicated them to share significant sequence similarity between taxa. It is possible that the gene region represented on the microarray for these genes was not representative of the level of sequence divergence across the complete gene sequence. The analyses of this study may have benefited from estimation of the branch length distribution for thousands of genes showing evidence of significant sequence similarity in the IGH experiment. Branch lengths of individual rapidly diverging genes could then be contrasted against this distribution. However, such analyses were outside of the scope of the present study, and were unnecessary to demonstrate that accelerated rates of sequence evolution could be confirmed for a selection of rapidly diverging genes in different teleost lineages.

Limitations of IGH experiments for the identification of rapidly diverging genes

The nature of genomic hybridisation experiments on microarrays imposes some limitations to the identification of rapidly diverging genes. For example, signal intensity in inter-species genomic hybridisation experiments not only depends on

sequence similarity but also on gene copy number (Wu *et al.* 2001, Hinchliffe *et al.* 2003). However, multiple copies of the same gene would have to retain significant sequence similarity in order to affect signal intensities. Furthermore, differences in gene copy number generally affect only a small number of gene loci. Thus, the greater the phylogenetic distance between two taxa, the more likely that differences in signal intensity are due to sequence divergence rather than differences in gene copy number (Brunelle *et al.* 2004). For these reasons, and because of the large phylogenetic distance between *P. moluccensis* and *D. rerio*, I assumed that most differences in signal intensity observed here can be attributed to sequence differences rather than gene copy number changes. The sensitivity of microarray technology imposes a further limitation on the identification of rapidly diverging genes. Because cross-hybridisation on microarrays is robust to some degree of sequence divergence, sequences need to differ by at least 12% to 21% before a loss in signal intensity can be detected (Kim *et al.* 2002; chapter one of this thesis). Thus, rapidly diverging genes reported here are likely to represent only genes with the greatest sequence divergence between *D. rerio* and *P. moluccensis*. Lastly, because the oligonucleotides on the microarray only represent a localised gene region, sequence differences detected in this way may not be representative of the level of sequence divergence across the complete gene sequence. Despite these limitations, IGH experiments can provide a useful tool for guiding the identification of rapidly diverging genes, in particular in species for which a lack of genome sequence data currently precludes bioinformatic approaches.

Outlook

In this study, I have demonstrated the use of IGH experiments and bioinformatic approaches for the identification of rapidly diverging genes in teleosts. Several teleost genomes have been, or are currently being sequenced, e.g. *D. rerio*, *T. rubripes*, *G. aculeatus*, *Tetraodon nigroviridis*, and *Oryzias latipes*, There are also significant sequencing efforts underway for *Salmo salar*, *Oncorhynchus mykiss*, and various cichlid species including *Oreochromis spp*. (Roest Crollius and Weissenbach 2005). Teleosts have thus become important models for comparative and functional genomic studies in vertebrates (Metscher and Ahlberg 1999, Mulley and Holland 2004, Cossins and Crawford 2005, Roest Crollius and Weissenbach 2005,

Froschauer *et al*. 2006). In addition, teleosts have experienced a whole genome duplication event after their divergence from holostei, possibly some 350 MYA (Amores *et al*. 1998, Christoffels *et al*. 2004, Jaillon *et al*. 2004, Crow *et al*. 2006). Gene duplicates may have been free to evolve new functions, and may in fact have needed to do so in order to avoid genomic instability due to genetic redundancies mentioned above (Malaga-Trillo and Meyer 2001). This whole genome duplication event may explain the overall accelerated rates of sequence evolution observed in teleosts compared to mammals (Jaillon *et al*. 2004). The great diversification of modern teleosts, which today are the most speciose vertebrate group with some 24,000 species, has also been causally related to this whole genome duplication event (Ruddle *et al*. 1994, Sidow 1996, Blomme *et al*. 2006, Froschauer *et al*. 2006), although there are debates about whether or not these events are indeed causally linked (Donoghue and Purnell 2005, Crow and Wagner 2006). For the above reasons, teleosts appear interesting models to study comparative genome evolution in vertebrates, gene duplication and gene loss events, and to study the relationship between gene duplication events, functional diversification, and species radiation.

Conclusion

Inter-species genomic hybridisation experiments hold great potential for the identification of rapidly diverging genes without the need for extensive genomic resources. Such experiments may be used to screen multiple closely related species and to identify genes, which show consistently accelerated rates of evolution. Application of IGH experiments for the identification of rapidly diverging genes can thus provide new insights into molecular sequence evolution and adaptation. In the present study, I have used IGH experiments to identify rapidly diverging genes between *D. rerio* and the coral reef fish *P. moluccensis* and I present a new approach for the analysis of IGH data. Rapidly diverging genes showed an over-representation of gene functions with roles in cellular responses to external stimuli. For example, the receptor genes *adra2a, reverbb2*, and *ERR2*, and the signal transduction gene *Tradd*. showed accelerated rates of sequence evolution. Homeobox genes were also significantly over-represented among rapidly diverging genes. I did not, however, detect a significant enrichment of stress-related gene functions among rapidly

diverging genes. I further obtained sequences orthologous to *D. rerio* in the teleosts *T. rubripes* and *G. aculeatus* and confirmed accelerated rates of sequence evolution for a selection of rapidly diverging genes across multiple teleost lineages. The results of this study suggest that IGH experiments can generate new hypotheses about rapidly diverging genes in the absence of genome sequence data. *In silico* comparative genomic approaches and direct sequencing may then be used to test hypotheses generated by means of IGH experiments.

CHAPTER 5

Modelling the stress response: molecular and physiological correlates and their evolutionary significance

Publication: Kassahn KS, Crozier RH, Caley MJ (*In prep*) Modelling the stress response: molecular and physiological correlates and their evolutionary significance.

Abstract

Recent advances in molecular technology and the use of DNA microarrays for gene expression profiling are providing new insights into the effects of stress by elucidating the transcriptional response to stress. Understanding gene regulation during stress can add to our understanding of how the genetic make-up of an organism is translated into biological function. In this endeavour to link the genome and phenome, microarray gene expression studies have brought together physiologists, population geneticists, and evolutionary biologists. However, interpretation of the complex transcriptional changes that occur during stress still poses many challenges. A conceptual model is needed to link physiological and transcriptional responses to stress. Here, I provide the basis for one such model by synthesising data from organismal, endocrine, cellular, molecular, and genomic studies. I propose that reduced cellular oxygen levels and oxidative stress are common to most stress conditions and that oxygen imbalance is a key regulator of physiological and transcriptional responses to stress. Furthermore, many transcriptional stress responses appear coordinated by only a limited number of transcription factors and signal transduction pathways. These upstream regulators of transcriptional stress responses have been largely neglected in environmental genomic studies thus far. There is considerable interaction between different levels of biological organisation. My proposed model may help the interpretation of environmental genomic data in the context of organismal function. I also provide an

evolutionary perspective on the implications of stress for adaptive evolution, particularly in the light of predicted climate change.

Introduction

Stress may be defined as any displacement from homeostasis (Johnson *et al.* 1992, Moberg 2000), and is usually considered to cause an injurious change in a biological system (Hoffmann and Parsons 1991). The deleterious effects of stress include, for example, impairments of development, growth and reproduction (Hochachka and Somero 2002). In addition, stress commonly suppresses immune functions, making the organism more prone to developing pathologies (Johnson *et al.* 1992, Moberg 2000). Hence, stress has been widely investigated in the context of human health, animal welfare, agriculture and aquaculture, and ecosystem conservation and management. The study of animal stress has become particularly topical in recent years due to the potentially negative effects of predicted climate change on animal function (Hoegh-Guldberg 1999, Etterson and Shaw 2001, Pörtner *et al.* 2001, Hoffmann *et al.* 2003, Stillman 2003, Somero 2005). Since the onset of stress delineates the limits of normal physiological function, the study of stress can provide important insights into organismal adaptation and adaptive evolution.

Early stress studies distinguished between an alarm reaction stage, a resistance stage, and an exhaustion stage, which were described as the 'General Adaptation Syndrome' (Seyle 1936). During the alarm reaction stage, the primary or 'flight or fight' response to stress involves changes in heart rate, blood pressure, and gastrointestinal activity (Cannon 1929). Induction of the hypothalamic-pituitary-adrenal (HPA) axis and secretion of glucocorticoids can lead to secondary responses, such as adrenal hypertrophy, gastrointestinal ulceration, and thymic and lymphoid shrinkage. During this stage, biological functions are adjusted to cope with the stressor. Impairments of development, growth, reproduction, and immune function are considered the tertiary response to stress. Persistence of the stressor may finally lead to exhaustion and death.

Since these early stress studies, the effects of stress have been investigated at all levels of biological organisation, including systemic, cardio-respiratory, endocrine,

cellular, and molecular responses to stress. These different aspects of the stress response have been reviewed in more detail elsewhere (Johnson *et al*. 1992, Somero and Hofmann 1996, Bonga 1997, Feder and Hofmann 1999, Hochachka and Somero 2002, Pörtner 2002, Somero 2005). More recently, microarray technology has identified extensive gene regulation in response to environmental stressors (Gasch *et al*. 2000, Gracey *et al*. 2001, Buckley *et al*. 2003, Enjalbert *et al*. 2003, Williams *et al*. 2003b, Gracey *et al*. 2004, Koskinen *et al*. 2004, Podrabsky and Somero 2004, Yoshida *et al*. 2004, Krasnov *et al*. 2005). However, interpretation of transcriptional stress profiles still poses many challenges and the significance of gene regulation for organismal function often remains unclear, largely because the relationship between transcriptional stress responses and responses at other levels of biological organisation are not well understood.

Transcriptional stress responses are not commonly considered in physiological stress studies despite the fact that the definition of physiology in the era of functional genomics may well include gene regulation. For the purpose of this review, I maintain this distinction between physiological and transcriptional stress responses. I use the term physiological stress response to refer to neuroendocrine, biochemical, metabolic, and reproductive adjustments. In contrast, I use the term transcriptional stress responses to refer to any changes in the transcriptome that occur as a result of stress. Both physiological and transcriptional responses to stress are aimed at recovering homeostasis and many of these responses are shared across a broad range of organisms. For example, catecholamines and cortisol coordinate behavioural and physiological stress responses in a range of organisms. Similarly, heat shock proteins are present in many organisms and play important roles in protein repair during stress. While the response to stress can vary depending on the type and severity of the stressor, there are a series of responses that are common to most stress conditions. These have been referred to as the generalised stress response (reviewed in Iwama *et al*. 2006).

In this chapter, I present a conceptual model which accounts for the key components of physiological and transcriptional stress responses and their main interactions (Figure 1).

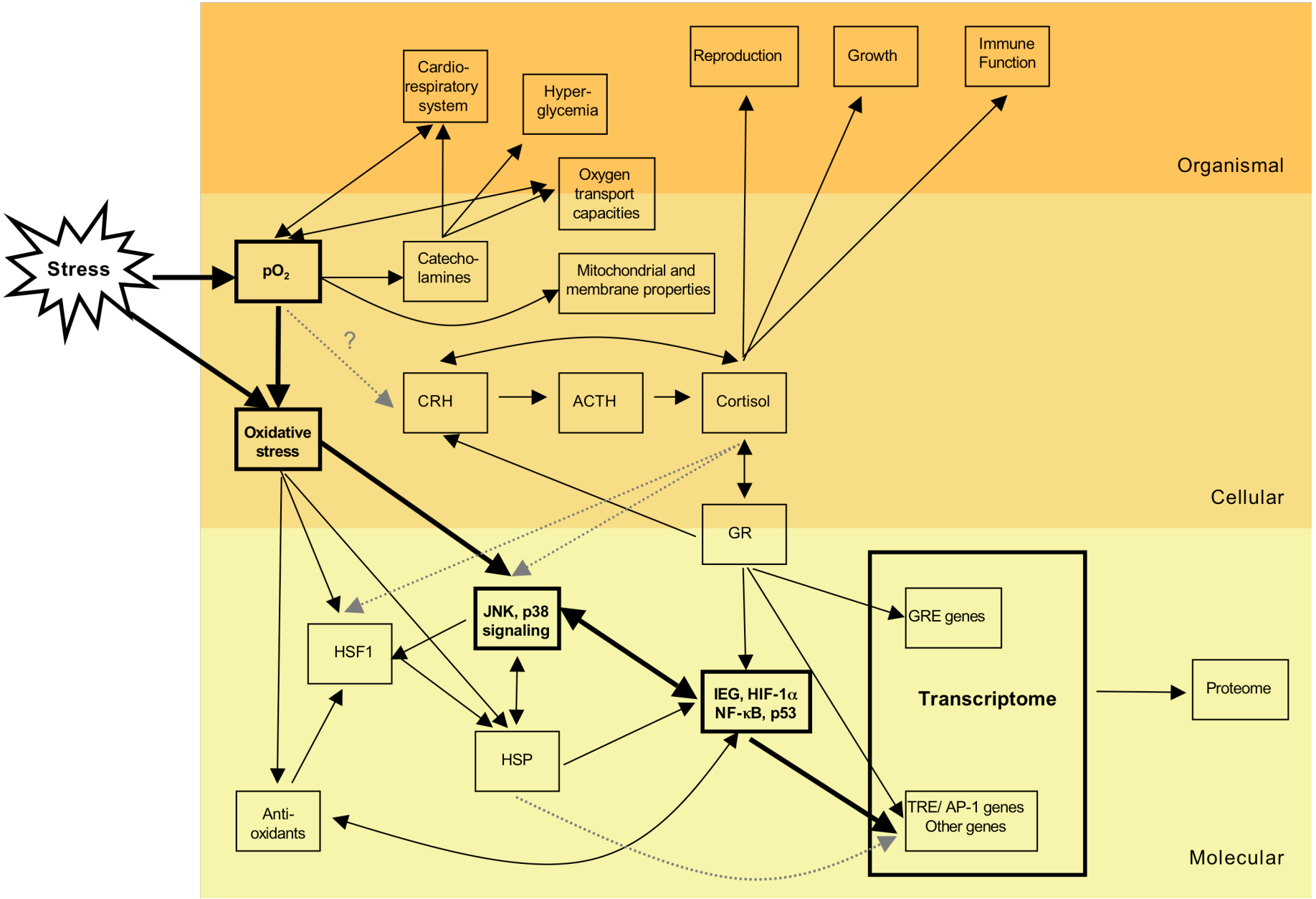
I propose that:

1. The common effects of stress include the disturbance of oxygen balance and the production of reactive oxygen species. Oxidative imbalance and oxidative stress play central roles in regulating physiological and transcriptional responses to stress.

2. Oxidative stress activates signal transduction pathways, immediate early genes, and other transcription factors. The transcriptional stress profiles measured in environmental genomic studies are the result of the redox-sensitive activation of these signalling pathways and transcription factors.

Throughout this review, I will highlight the central role of oxygen imbalance and reactive oxygen species in regulating physiological and transcriptional stress responses. I focus on the main links between stress responses at different levels of biological organisation. The reader is referred to Figure 1 for visualisation of these relationships. However, it is impossible to present a model inclusive of all interactions. I have focused instead on the features that are common to many stress conditions, key regulators of stress responses, and the main interactions between responses at different levels of biological organisation. While many examples are drawn from teleost studies, this model should apply in a similar way to a wider range of organisms. This conceptual model may contribute to our understanding of transcriptional stress responses and aid the interpretation of environmental genomic data in the context of organismal function. Finally, I discuss the implications of stress for adaptation to new environments, and the ability of organisms to cope with predicted climate change.

Figure 1 Conceptual model of the stress response relating organismal, cellular and molecular responses. Key regulators and pathways and their main interactions are indicated in bold. Weaker interactions and interactions that are based on the modulation of activity are indicated by dashed lines. Arrowheads indicate the direction of interaction. Transcriptional changes observed during stress are likely a result of activation of these key regulators and pathways. Knowledge of their interactions may aid interpretation of genomic data in ecological and functional genomic studies. Abbreviations: *ACTH* Adrenocorticotropic hormone; *CRH* Corticotropin-releasing hormone; *GR* Glucocorticoid receptor; *GRE genes* Genes with the glucocorticoid response element; *HIF-1*α Hypoxia-inducible factor 1 alpha; *HSF1* Heat shock factor 1; *HSP* Heat shock proteins; *IEG* Immediate early genes; *JNK* c-*Jun* NH_2-terminal kinase; *NF-*κ*B* Nuclear Factor kappa B; *TRE/ AP-1 genes* Genes with the TPA-responsive element/ Activation protein-1 site; *p38* The p38 signal transduction pathway; *p53* Tumor protein p53.


Organismal
Cellular
Molecular
Stress
pO_2
Cardio-respiratory system
Hyper-glycemia
Oxygen transport capacities
Catecho-lamines
Mitochondrial and membrane properties
Reproduction
Growth
Immune Function
?
CRH
ACTH
Cortisol
Oxidative stress
GR
JNK, p38 signaling
HSF1
HSP
IEG, HIF-1α NF-κB, p53
Anti-oxidants
GRE genes
Transcriptome
TRE/ AP-1 genes Other genes
Proteome

Stress commonly leads to reduced cellular oxygen levels

Stress appears to commonly cause a reduction in cellular oxygen levels (for examples see Sisson and Sidell 1987, Guderley and Blier 1988, St-Pierre *et al*. 1998). Changes in respiration and circulation (Crockett and Sidell 1990, Guderley 1990), adjustments of mitochondrial densities and properties (Miranda and Hazel 1996, Logue *et al*. 2000), alterations of enzyme capacities (Sommer *et al*. 1997, Pörtner *et al*. 2004), and alterations in membrane composition (Pörtner 2002) that have been observed during stress are, therefore, compensatory mechanisms targeted at re-establishing normal cellular oxygen levels. In some circumstances, these compensatory mechanisms appear insufficient to maintain aerobic function. As a result, some stress studies have reported a switch to anaerobic function and a reduction in cellular energy levels (Abele *et al*. 1998, Flanagan *et al*. 1998, Abele *et al*. 2002, Heise *et al*. 2003, Keller *et al*. 2004, Heise *et al*. 2006a). It has been suggested that an organism's ability to endure stress is largely a function of its ability to maintain aerobic function under stressful conditions (Abele and Puntarulo 2004, Morales *et al*. 2004, Lesser 2006). In this review, I suggest that reduced cellular oxygen levels and the associated oxidative stress play central roles in regulating physiological and transcriptional stress responses.

The endocrine system coordinates stress responses and affects rates of transcription

Stress commonly induces a neuroendocrine response, which is mainly coordinated by catecholamines and glucocorticoids. In teleosts, these hormones are part of the hypothalamic-sympathetic-chromaffin cell axis and the hypothalamic-pituitary-interrenal axis (Chrousos and Gold 1992). The release of catecholamines under stress has been shown to stimulate cardiac output and to increase ventilation rate, branchial blood flow, and oxygen transport capacities (Nikinmaa 1992a, 1992b, Soldatov 1996, Bonga 1997). Stress-related hyperglycemia is also mediated by catecholamines and the freeing of stored carbohydrates by glycogenolysis (Janssens and Waterman 1988, Mommsen *et al*. 1988). Elevated blood glucose levels and increased oxygen consumption are likely to reflect the increased metabolic rates observed during many

types of stress (Barton and Schreck 1987), although hypoxia and cold shock can cause metabolic depression (Donaldson 1981, van Ginneken *et al*. 2001).

Cortisol is the main glucocorticoid in teleosts and is commonly released during stress (Schreck *et al*. 1989, Young 1993). During stress, cortisol stimulates gluconeogenesis in the liver (Sheridan 1994) and increases plasma free fatty acid levels by promoting lipolysis (Hyllner *et al*. 1989, Gedamu and Zafarullah 1993, Ryan *et al*. 1995). High cortisol levels suppress growth, reproduction, and immune function (Vijayan *et al*. 1991). Please refer to Figure 1 for an illustration of these relationships. The main regulator of cortisol secretion is adrenocorticotropic hormone (ACTH) (Fryer 1989, Olivereau and Olivereau 1990), while ACTH levels are regulated by corticotropin-releasing hormone (CRH) (Bonga 1997, Pankhurst and Van Der Kraak 1997, Pottinger 1999). The regulation of cortisol and CRH levels appears complex. A variety of factors, including 12-O-tetradecanoylphorbol 13-acetate (TPA), the protein kinase A pathway, cyclic-AMP, growth hormone, thyroxin, cytokines, catecholamines, and glucocorticoids themselves have been reported to modulate the secretion of cortisol and CRH (Sumpter *et al*. 1994, Vamvakopoulos and Chrousos 1994, Mark *et al*. 2002, Pörtner 2002, Lannig *et al*. 2004, Pörtner *et al*. 2004, Romero 2004). The complex regulation of cortisol secretion suggests an interaction of neural, endocrine, cellular and immune functions. Neural sensing of reduced cellular oxygen levels may also modulate the secretion of CRH. However, at present, this relationship is still poorly understood.

Cortisol also directly interacts with transcriptional regulation, indicating a two-way link between endocrine and transcriptional stress responses. For example, cortisol can stimulate the expression of metallothionein, ubiquitin, and possibly heat shock proteins by interacting with heat shock factors (Vamvakopoulos and Chrousos 1994, Iwama *et al*. 2006). Cortisol further binds to glucocorticoid receptors and in this association interacts with the c-*Jun* component of the activation protein-1 (AP-1) transcription factor (Iwama *et al*. 2006). Downstream, AP-1 regulates the expression of the transcription factor NFκB, genes with the TPA-response element (TRE)/ AP-1 DNA motif, and genes with the glucocorticoid response element (GRE, Figure 1). Stress-dependent activation of glucocorticoid receptors by cortisol can, therefore, lead to transcriptional changes at a large number of gene loci. The transcriptional

effects of cortisol binding to glucocorticoid receptors has been shown to depend on the tissue type and the amount of heat shock protein 90 present (Farrell 2002). This observation suggests a link between cortisol-dependent gene induction and the heat shock response.

Production of reactive oxygen species during cellular stress

Another feature that appears to be common to many stress conditions is an increased production of reactive oxygen species (ROS) leading to oxidative stress (for examples see Rifkin *et al*. 1993, Brand *et al*. 2004, Brookes 2005, Lushchak and Bagnyukova 2006). The inner membrane of mitochondria is the primary site of ROS production and reduced cellular oxygen levels during stress appear to enhance mitochondrial production of ROS (Ames *et al*. 1993, Sohal and Weindruch 1996, Beckman and Ames 1998). Antioxidant defenses are commonly activated during stress (Sohal *et al*. 1993, Sohal and Weindruch 1996); their activation is likely a response to increased levels of oxidative stress. Oxidative stress also plays a central role in aging (Finkel and Holbrook 2000), leading to an accumulation of protein and DNA damage (Shigenaga *et al*. 1994). However, oxidative stress may itself increase with age because of age-related defects in antioxidant defenses and mitochondrial function (Lau and Nathans 1985, Rollins and Stiles 1988, Herschman 1991, McMahon and Monroe 1992, Melov *et al*. 2000). In systems where antioxidant defenses have been artificially enhanced, increased life span has been observed (Herschman 1991, Mohn *et al*. 1991, McMahon and Monroe 1992, Melov *et al*. 2000). Because there are many similarities between the response to stress and the responses to aging and disease it is possible to argue that aging and disease are in fact just another form of stress, arising from defects in biological function.

Oxidative stress causes protein damage and induces a heat shock response

Increased levels of oxidative stress and the formation of ROS during stress are likely to cause increased levels of protein damage. Ubiquitin labels denatured and damaged proteins for proteolysis. The amount of ubiquitin-labelled protein can be used as an indication of the level of irreversible protein damage in the cell (Parsell and

Lindquist 1993, Iwama *et al.* 1998, Feder and Hofmann 1999). The presence of damaged and ubiquitinated protein in the cell induces a heat shock response (Wu 1995), suggesting that oxidative stress is an indirect activator of the heat shock response (Figure 1). Most heat shock proteins are molecular chaperones and function in the folding, repair and catabolism of proteins (Buckley and Hofmann 2002). For example, *hsp70* prevents the aggregation of denatured proteins, while *hsp140* dissolves aggregates and rescues protein conformation (Moseley 1997). The magnitude of the heat shock response depends not only on the magnitude and duration of the stress (Sorensen *et al.* 2003, Iwama *et al.* 2004), but also on acclimation and previous exposure to stress (Hawkins 1991, Somero and Hofmann 1996). Because expression levels of heat shock proteins are largely context-dependant, the use of heat shock proteins as stress markers requires careful establishment of baseline expression levels prior to experimentation (Parsell and Lindquist 1993, Fader *et al.* 1994, Tomanek and Somero 1999, 2000). Heat shock proteins have further roles in buffering the expression of hidden genetic variation (see below).

The expression of heat shock proteins is under the regulation of the heat shock factor 1 (HSF1, Figure 1). Under normal conditions, HSF1 is bound to heat shock proteins and is therefore rendered inactive. During stress, protein damage leads to the dissociation of heat shock proteins from HSF1. HSF1 then migrates to the cell nucleus and induces the transcription of genes with the heat shock element (HSE), such as heat shock proteins (Kim *et al.* 1997, Kline and Morimoto 1997). The activation of HSF1 is acclimation-dependent (Parsell *et al.* 1994). HSF1 activity further depends on its phosphorylation state, which is under the control of mitogen-activated protein kinases (Somero and Hofmann 1996). While the heat shock response is activated in response to protein damage, increased production of ROS during stress may ultimately be the cause of increased protein damage. Oxidative stress thus has an indirect role in inducing the heat shock response. Oxidative stress further interacts with signal transduction pathways, some of which control the activity of HSF1. There are hence multiple interactions between cellular oxygen balance, the expression of heat shock proteins, and signal transduction pathways.

While the heat shock response system is aimed at repairing protein damage during stress, the costs associated with replacing and repairing damaged proteins are likely to be high. Protein synthesis accounts for approximately 20% to 30% of cellular ATP turnover (Feder and Hofmann 1999). In some organisms, heat stress results in the immediate reduction or elimination of synthesis of proteins other than heat shock proteins (Nakano *et al.* 2004). Elevated *hsp70* levels have also been associated with reduced somatic growth (Sonna *et al.* 2002). Furthermore, high concentrations of heat shock proteins can be toxic and interfere with normal cellular function (Fader *et al.* 1994). Cellular heat shock protein levels are thus likely to reflect a trade-off between the costs and benefits associated with maintaining a molecular chaperone system. While heat shock proteins play crucial roles in the repair of protein damage during stress, constraints in biochemical adaptation which would prevent protein damage have been discussed elsewhere (Hochachka and Somero 2002).

Insights from cell biology – activation of signal transduction pathways and immediate early genes during stress

Studies in cell biology have provided important insights into the mechanisms by which a cell perceives and responds to external stimuli including stress. This literature, however, is rarely referred to in physiological and environmental genomic studies. Here, I briefly summarise this literature as it relates to physiological and transcriptional stress responses. The c-*Jun* NH_2-terminal kinase (JNK) and the p38 kinase pathway are commonly activated during stress (Cowan and Storey 2003). JNKs are the primary kinases responsible for phosphorylation of *c-Jun* and activation of the AP-1 transcription factor (Whitmarsh *et al.* 1995), but JNKs also activate other transcription factors, including *myc* and *p53* (Cowan and Storey 2003). These JNK-regulated transcription factors initiate many of the transcriptional responses observed during stress (Gabai and Sherman 2002). There is cross-talk between the JNK signalling pathway and the heat shock response (Figure 1). The JNK pathway can modulate the expression of heat shock proteins, while heat shock proteins, in particular *hsp27* and *hsp72*, can modulate JNK signalling (Schenk *et al.* 1994). The downstream targets of the stress-induced p38 kinase pathway include heat shock factors and the heat shock proteins *hsp25* and *hsp27* (Mercurio and Manning 1999). Understanding the signal transduction pathways that are activated during stress can

therefore help interpret the gene expression changes that occur downstream of these stress-signalling cascades.

Studies in cell biology have further identified a series of genes, called immediate early genes (IEGs), that are induced rapidly and transiently within minutes of the cell receiving an external stimulus or experiencing stress (Herschman 1991, Hughes and Dragunow 1995). Again, this literature is rarely referred to in physiological and environmental genomic studies. Here, I suggest that understanding the functions of IEGs can help interpret transcriptional stress responses and I briefly summarise the current literature on IEG function.

Induction of IEGs occurs in a variety of ways, for example by serum response factor, increases in cellular Ca^{2+} levels, or activation of MAPKK or JNK signal transduction pathways (Cohen 1997, Edwards *et al*. 2004, Hughes and Dragunow 1995, Figure 1). IEGs include the transcription factors *c-fos, fosB, c-jun, junB*, *c-myc, egr-1*, but also genes with other functions, e.g. mtf, the cytokines KC and JE, actin, and fibronectin (Inuzuka *et al*. 1999, Cai *et al*. 2000, Dunn *et al*. 2005). As stress-inducible transcription factors, some IEGs are likely to play important roles in gene regulation during stress. For example, the protein products of *Fos* and *c-jun* can form heterodimers, such as the AP-1 complex. The AP-1 complex activates the expression of a large number of genes containing the TRE/ AP-1 DNA motif including some antioxidant defense genes (Schenk *et al*. 1994). The DNA-binding ability of *Fos* in AP-1 is further subject to redox regulation via the Ref-1 enzyme and thioredoxin (Kyriakis and Avruch 1996, Verheij *et al*. 1996, Shaulian and Karin 2002, Cowan and Storey 2003). Thus, immediate early genes are commonly activated in response to stress; some IEGs are transcription factors that regulate the expression of a large number of genes. Interpreting transcriptional stress profiles in the light of activated immediate early genes and signal transduction pathways can identify common features among transcriptional stress profiles. There is cross-talk between the induction of immediate early genes, JNK and p38 signal transduction pathways, heat shock proteins, and antioxidants (Figure 1). Since levels of oxidative stress regulate the expression of antioxidants, this suggests an indirect role of oxidative stress in the activation of immediate early genes and signal transduction pathways.

The activation of NFκB, HIF-1α and other transcription factors during stress

Oxidative stress also activates other transcription factors. The nuclear factor-κB (NF-κB), for example, regulates the transcription of a large number of genes usually involved in apoptosis, immunity and inflammation (Mercurio and Manning 1999). However, NF-κB is a redox-sensitive transcription factor and as such also involved in cellular responses to oxidative stress. NF-κB regulates the expression of many antioxidants, while antioxidants, e.g. thioredoxin, can suppress NF-κB activity (Semenza *et al*. 1994, Carmeliet *et al*. 1998, Lando *et al*. 2002, Treinin *et al*. 2003, Semenza 2004a, 2004b, Heise *et al*. 2006b). NF-κB expression is regulated by inhibitor protein IκB, which in turn is regulated by several MAPKs and interleukin-1 receptor-associated kinases, suggesting cross-talk between cellular redox state, NF-κB activity, and protein kinase signalling (Figure 1).

The hypoxia-inducible factor 1 (HIF-1) is another transcription factor commonly activated during stress (Semenza 1998). Under normoxic conditions, HIF-1α is rapidly ubiquitinated and degraded by 26S proteasome (Katschinski *et al*. 2002). However, during hypoxia and other stress conditions, HIF-1α is stabilised and regulates the expression of many genes involved in oxygen transport, angiogenesis, and glucose metabolism, e.g. erythropoietin, vascular endothelial growth factor, heme oxygenase-1, glyceraldehyde-3-P, phosphoglucomutase, aldolase A, and lactate dehydrogenase A (Soitamo *et al*. 2001). Heat shock protein 90 regulates the accumulation and activity of HIF-1α during heat and hypoxia (An *et al*. 1998, Carmeliet *et al*. 1998). These interactions suggest cross-talk between cellular oxygen levels, HIF-1α induction, and the heat shock response (Figure 1).

Other transcription factors activated by reactive oxygen species include the tumor suppressor p53 and heat shock factor (HSF). The activities of both factors are modulated by thioredoxin and other antioxidants (Rechsteiner 1987). The activity of p53 is further modulated by its interaction with HIF-1α. During hypoxia, binding of HIF-1α to p53 leads to the accumulation of p53 (Hofmann and Somero 1995). These interactions suggest cross-talk between oxidative stress, p53 and HSF1 induction, antioxidants and HIF-1α (Figure 1).

Immediate early genes, the transcription factors NF-κB, HIF-1α, p53, and HSF, and JNK and p38 signalling pathways are likely to be the regulators of the gene expression responses measured in environmental genomic studies. Many of these transcription factors and signalling pathways are activated by oxidative stress suggesting a pivotal role for cellular redox state in regulating transcriptional responses to stress. Interpretation of transcriptional stress responses in light of these upstream regulators can, thus, aid our understanding of how gene regulation relates to responses at other levels of biological organisation.

An evolutionary perspective on the relationship between stress and adaptive evolution

In this review, I have identified key components of the physiological and transcriptional responses to stress. However, stress is metabolically expensive, owing to the cellular and molecular damage it causes and the costs associated with subsequently re-establishing homeostasis (Koehn and Bayne 1989, Calow 1991, Feder and Hofmann 1999). Acclimation allows survival under conditions that otherwise would be lethal, but comes at the cost of reduced competitive ability and fecundity (Hoffmann 1995). For these reasons, it may be beneficial to avoid stress by evolving greater or different tolerances. So why can organisms not avoid stress in the first place? And what are the evolutionary constraints limiting greater environmental tolerances? In order to answer these questions it is necessary to examine the costs associated with evolving greater environmental tolerances and the genetic mechanisms underlying the evolution of stress resistance.

One mechanism for altering tolerances is modifying gene expression levels and such transcriptional variation can be of adaptive advantage. For example, expression levels of lactate dehydrogenase-B (*Ldh-B*) in the killifish *Fundulus heteroclitus* vary between populations and this variation is related to environmental temperatures (Schulte 2000), suggesting that transcriptional variation can lead to phenotypic changes and adaptive advantages. In addition, tolerances may be enhanced through biochemical adaptation. For example, the kinetic properties and allele frequencies of *Ldh-B* vary between warm- and cold-adapted populations of *F. heteroclitus* (Powers and Schulte 1998). However, there are many trade-offs in biochemical adaptation.

Enzymes need to retain the ability for conformational changes, while being rigid enough to allow binding of substrates (Hochachka and Somero 2002). The costs and benefits of a molecular chaperone system have been discussed above. Furthermore, adjustments of the transcriptome may temporarily ameliorate the negative effects of stress (Somero 2005). However, most variation in gene expression appears to occur between individuals, not populations (Oleksiak *et al*. 2002, Oleksiak *et al*. 2005), suggesting that the extent of adaptive variation in gene expression is limited. Trade-offs thus appear to occur at all levels of physiological and biochemical adaptation and would limit an organism's ability to avoid stress.

Stress is thus often unavoidable and delineates the limits of biochemical adaptation. Population genetic studies have identified further limitations to organismal adaptation. For example, the distribution of the fruit fly *Drosophila birchii* is limited by its inability to tolerate drought. Yet, while desiccation resistance has a high heritability in many *Drosophila* species, intense laboratory selection experiments have failed to evolve greater desiccation resistance in *D. birchii* (Hoffmann *et al*. 2003). *Drosophila birchii* appears to lack genetic variation which would allow evolution of increased desiccation resistance. Adaptation to new environments may be further limited by among-trait genetic correlations that are antagonistic to the direction of selection (Etterson and Shaw 2001), gene swamping from populations adapted to different environmental conditions (Kirkpatrick and Barton 1997; Lenormand 2002), and trade-offs in the adaptation to different environments (Somero 2005). Stress has thus been implicated in explaining geographic limits of species ranges (Hoffmann and Blows 1994). Populations are thought to experience increasing levels of stress towards the edge of their species' range, while lacking the ability to evolve resistance to conditions at the range margin.

While adaptation to new environments can strike limits due to such factors as limited stores of additive variation, stress can facilitate adaptation by releasing genetic variation previously concealed. Such release of previously hidden variation by stress was first proposed in the context of the canalisation model of Waddington (1942). More recently, several molecular mechanisms underlying the stress-dependant release of genetic variation have been identified. Heat shock protein 90 interacts with many signal transducers in development and cell cycle and also allows accumulation

of genetic variation by buffering the potentially negative effects of such variation. During stress, *hsp90* buffering is compromised resulting in the sudden appearance of new phenotypes (Rutherford and Lindquist 1998, Queitsch *et al*. 2002). These new phenotypes can then be selected for. In laboratory selection experiments, new phenotypes rapidly become independent of the *hsp90* context, allowing for permanent phenotypic change (Rutherford and Lindquist 1998, Cowen and Lindquist 2005). In bacteria, heat shock protein 70 has also been implicated in stress-dependent generation of genetic variation. During stress, over-expression of *hsp70* stimulates transposition events (Chow 2000). Whether *hsp70* also plays a role in the generation of new genetic variation in other organisms is yet to be determined. Another facilitator of adaptive evolution is the prion conformation of the Sup53 protein [PSI^+] in yeast, which allows read-through of stop codons during translation. This read-through results in proteins of extra length, which can be of selective advantage during altered environmental conditions. The activity of [PSI^+] appears to be modulated in a stress-dependent manner (Partridge and Barton 2000, True and Lindquist 2000). Under certain circumstances population bottlenecks can reduce genetic buffering leading to increased epistatic genetic variance and a 'release' of genetic variation (A.A. Hoffmann, *pers. comm.*). Since stress may be the cause of the population decline and bottleneck, this suggests another mechanism by which stress can facilitate phenotypic change. In summary, several mechanisms can release previously hidden genetic variation under conditions of environmental stress. Release of such genetic variation is likely to play an important role during adaptation to new environments.

There are continuing controversies about whether there is selection for increased evolvability (Wagner *et al*. 1999, Partridge and Barton 2000, Masel 2005). Mathematical modelling of the costs and benefits associated with an evolutionary capacitor system has demonstrated that, in principle, natural selection can favour the evolution of evolutionary capacitors (Masel 2005). In contrast, evolutionary capacitor mechanisms may be solely a by-product of other functions carried out by these systems (Meiklejohn and Hartl 2002). Partial loss of *hsp90* function during stress may in fact be predominantly detrimental, rather than adaptive, due to fact that the phenotype is least buffered against environmental disturbances when these are most extreme (Meiklejohn and Hartl 2002). The genetic variation that can be

released by evolutionary capacitor systems during stress is normally concealed by canalisation, i.e. the buffering of the phenotype against changes in genotype and environment (Waddington 1942). Similar to the controversy regarding the adaptiveness of evolutionary capacitor systems, there is contention about whether or not canalisation is adaptive. Stabilising selection may be expected to favour any mechanism, including canalisation, that buffers the effects of genetic and environmental variation on the phenotype. However, computer simulation studies have demonstrated that canalisation can occur in the absence of stabilising selection (Siegal and Bergman 2002). For example, the complexity of gene regulatory networks underlying development can make these networks insensitive to mutations at individual gene loci (Siegal and Bergman 2002). Selection experiments investigating the relationship between genetic variation and phenotypic plasticity may prove useful to resolve some of these controversies regarding the adaptiveness of canalisation (Scheiner 2002).

Stress and the effects of climate change on population viability

While stress may lead to accelerated rates of evolution, the current unprecedented rates of climate change may exceed the ability of many species to adapt to these changing conditions. Current models of climate change not only forecast an increase in temperatures, but also changes in the frequency and severity of El Niño Southern Oscillation events, tropical cyclones, and extreme temperature events, changes in rainfall and possibly ocean currents, and sea-level rise (Pittock 1999). Some populations and species may thus be lost due to climate change; indeed, some losses have already been reported (Glynn and Deweerdt 1991, Roemmich and McGowan 1995) and further losses have been predicted (Williams *et al*. 2003a). Some extinctions of marine species have been directly or indirectly linked to climatic change (Dulvy *et al*. 2003). For example, two species of coral, *Siderastrea glynni* and *Millipora boschmai*, are thought to have become globally extinct due to an extreme El Niño event in 1982/83 (Glynn and Deweerdt 1991, Dulvy *et al*. 2003). Some coral reef fishes also seem to have disappeared after the 1982/83 El Niño event: the damselfish *Azurina eupalama* appears to be globally extinct, while the harlequin leatherjacket *Oxymonacanthus longirostris* appears to have disappeared from the Maldives and Southern Japan (Dulvy *et al*. 2003).

In response to climatic change, species may also exhibit range shifts. Such range shifts have been documented for recent and historical times (Parker and Dixon 1998, Davis and Shaw 2001, Hellberg *et al.* 2001, Roy *et al.* 2001). For example, species composition in a reef fish community off North Carolina has changed significantly in response to a two-degree Celsius rise in sea surface temperatures over 15 years (Parker and Dixon 1998). Two tropical families and 29 tropical species are new to the area, while no new temperate species have been recorded and the previously most abundant temperate species have decreased drastically in abundance (Parker and Dixon 1998). Even though it is difficult to isolate the immediate causes for the shift in species composition, changes in the relative abundance of tropical and temperate species suggest a relationship with sea surface temperatures. Similarly, Californian rocky intertidal communities have undergone a change in invertebrate species composition between the 1930s and 1990s (Barry *et al.* 1995). The relative abundances of southern species have increased, while those of northern species have decreased. This shift in species composition was associated with and possibly caused by an increase in local sea surface and mean summer maximum temperatures between 1930s and 1990s (Barry *et al.* 1995). There are also reports of historical range shifts. For example, species ranges of marine bivalves in the central Pacific changed dramatically in response to sea level fluctuations and historical climate change (Paulay 1996).

Climate change generally adds to other pressures already confronting marine and terrestrial ecosystems. For example, current threats to coral reefs include overfishing, human damage due to mining, dredging, tourism, and boating activity, pollution through nutrient run-off, pesticides, and herbicides from agriculture, coral disease, mass mortality of the herbivorous urchin *Diadema antillarum* in the Caribbean, and outbreaks of crown-of-thorns starfish on the Great Barrier Reef (Hoegh-Guldberg 1999, Pittock 1999, Wilkinson 1999, Done *et al.* 2003, Gardner *et al.* 2003). As a result, coral cover has severely declined over the last three decades (Jackson *et al.* 2001, Wilkinson 2002, Gardner *et al.* 2003, Hughes *et al.* 2003, Pandolfi *et al.* 2003). Climate change is likely to add additional threats to coral reef ecosystem.

Corals seem to live on the edge of their thermal tolerance range. Increases in sea surface temperatures by only one degree Celsius above mean summer maximum

temperatures can cause coral bleaching (Hoegh-Guldberg 1999, Hughes *et al*. 2003) The increased frequencies of coral bleaching events that have been observed over the last 30 years are most likely related to the increases in sea surface temperatures and the increased frequencies of extreme temperature events that occurred over the same period of time (Hoegh-Guldberg 1999, Hughes *et al*. 2003). It is expected that this trend will continue and that coral reef ecosystems will be severely affected by climate change in coming decades (Hughes *et al*. 2003). In this context, the study of stress will contribute to our understanding of whether species will be able to adapt to climate change.

CHAPTER 6

General Discussion

General Summary

In this thesis, I have presented strategies for how microarray technology can be applied to study species with few genomic resources, but which are of interest in functional genomic studies. For many such species there are currently no microarrays available. However, there often are microarrays for related species. Here I demonstrate that inter-species genomic hybridisation experiments can inform us about which genes are conserved enough to allow microarray analysis across species.

I have performed heterologous microarray experiments to measure gene regulation in response to environmental stress in the coral reef fish *P. moluccensis*. The transcriptional responses to stress varied depending on the type of stressor, but there was conservation of gene function responses. Prolonged exposure to heat resulted in significant gene regulation, indicating increased protein damage, oxidative stress, and alterations in metabolism. Cellular and protein damage suggest that there are energetic costs associated with the heat response, potentially reducing energy available for growth and reproduction. The longer-term effects of these responses on reproductive output and fitness are yet to be determined. This is the first study to investigate transcriptional responses to elevated temperatures in a coral reef fish and the results of my PhD research indicate that a three-degree Celsius rise in temperature can cause significant alterations in gene expression in a coral reef fish.

I have further applied microarray technology to identify rapidly diverging genes between the coral reef fish *P. moluccensis* and the zebrafish *D. rerio*. This is a novel application of microarray technology and had, previously, only been applied to the study of microbial and fungal genomes. Rapidly diverging genes identified in this study showed an over-representation of receptor and transcription factor functions. Comparative genomic hybridisations on microarrays can thus provide a suitable

alternative to bioinformatic approaches for the identification of rapidly diverging genes, in particular for species that currently lack genome sequence data.

Finally, I have presented a conceptual model that relates transcriptional stress responses to stress responses at other levels of biological organisation, including systemic, endocrine, cellular, and molecular responses. Stress commonly leads to reduced cellular oxygen levels and oxidative stress, and as a result, many responses to stress are aimed at re-establishing oxygen balance and repairing cellular damage arising from oxidative stress. I have discussed the implications of stress for adaptation and have pointed out some mechanisms that can release genetic variation during conditions of stress and which can thus provide the raw material for selection to operate upon.

Summary and Discussion of Key Findings

The performance of heterologous microarray experiments

In Chapter 2, I report on inter-species genomic hybridisation experiments which demonstrated that most genes show high cross-hybridisation between the coral reef fish *P. moluccensis* and the zebrafish *D. rerio*. I also sequenced parts of the coding region of a selection of genes in *P. moluccensis*. These genes showed an average sequence similarity of 81% between *P. moluccensis* and *D. rerio* producing significant cross-hybridisation in the comparative genomic hybridisation experiment. I, therefore, infer that most genes share significant sequence similarity between *P. moluccensis* and *D. rerio*, sufficient to allow cross-hybridisation. I then used quantitative real-time PCR to validate expression changes identified in these heterologous microarray experiments. Quantitative real-time PCR confirmed significant expression changes for most gene loci examined. In summary, I have demonstrated in this first part of my thesis that the *D. rerio* microarray can be successfully used to measure gene regulation in *P. moluccensis*.

I also introduced a novel application of comparative genomic hybridisation experiments. Microarray hybridisation experiments using genomic DNA can be used to identify which genes are conserved enough to allow microarray analysis across

species. Using inter-species genomic hybridisation experiments, I have identified which of the genes represented on the *D. rerio* microarray allow measurement of gene regulation in *P. moluccensis*. Differences in gene copy number can potentially confound the results of inter-species genomic hybridisation experiments (Pollack *et al*. 1999, Hinchliffe *et al*. 2003), but with increasing phylogenetic distance differences in gene copy number should play a comparatively minor role (Brunelle *et al*. 2004). Researchers using cDNA microarrays may also need to account for differences arising from the different structure of genomic DNA and cDNA. Because genomic DNA includes introns and non-coding regions, probes on the microarray whose sequences span an exon-intron boundary in the genomic DNA sample may not bind genomic DNA, but would successfully bind cDNA during gene expression analysis. In these instances, cDNA derived from the species of interest and the species for which the microarray was designed may be competitively hybridised to the cDNA microarray and inform about the potential for cross-hybridisation (Renn *et al*. 2004). However, true biological differences in gene expression as well as gene copy number differences can confound the results of competitive cDNA hybridisation experiments across species. When using cDNA microarrays, the most thorough strategy for determining the expression of which genes can be successfully measured across species would thus consist of a combination of comparative genomic DNA and cDNA hybridisations.

Transcriptional responses of a coral reef fish to environmental stress

In Chapter 3 of this thesis, I report on heterologous microarray experiments designed to elucidate the transcriptional responses to environmental stress in *P. moluccensis*. This series of experiments allowed the assessment of different dimensions of the stress response, in particular how the type, intensity, and duration of stress affect gene regulation. A particular focus of this component of my PhD research was the identification of gene responses to heat stress. Exposure to heat over five days resulted in significant expression changes at 324 gene loci. Most regulated genes were up-regulated in response to prolonged heat. In contrast, most early gene responses to stress measured three hours after exposure resulted in significant down-regulation and suppression of transcription. I have thus identified a qualitative

difference between transcriptional responses to prolonged heat and early gene responses to stress.

Transcriptional responses to stress varied significantly depending on the type of stressor applied. Only few genes showed consistent regulation across stress treatments. In particular, three genes that had previously been associated with cell cycle functions were consistently induced in response to stress in this study, i.e. nucleobindin 2a, cyclin G1, and growth-arrest specific 8. Grouping of individual gene responses into gene function responses identified a series of gene functions that were commonly associated with the response to stress. These gene functions indicated that stress commonly led to a suppression of cell growth, alterations in metabolism, increased protein damage and protein catabolism, a challenge of the immune system, and the induction of genes to combat cellular damage. The response to prolonged heat was further associated with increased levels of oxidative stress.

I have demonstrated that a three-degree Celsius temperature rise can cause significant gene regulation in a coral reef fish. Such temperature rise is well within the range of temperature increases predicted by current models of climate change. Not only are sea surface temperatures predicted to rise, but also the frequency of El Niño and extreme temperature events is predicted to increase. Since the gene expression response to prolonged heat indicates increased levels of oxidative stress and protein damage, the response to heat appears associated with significant energetic costs. The observed regulation of metabolic genes may reflect this need to re-organise metabolic functions to account for increased protein and cellular repair. As a consequence, energy expended on protein and cellular repair in response to heat may reduce the energy available for other biological functions, such as growth and reproduction. The longer-term effects of these responses on the fitness and performance of these fishes are yet to be determined.

In this study I have focused on the analysis of gene expression responses in liver. While liver is a metabolically important tissue and shows extensive gene regulation in response to stress, it is likely that other tissue would show different gene responses. In addition, gene responses in liver may affect gene responses in other tissues and vice versa. However, analysis of gene responses in other tissues was beyond the scope of this research. It is possible that hyposmotic and hypoxic stress

may have caused more significant gene expression responses in other tissues, for example gill or muscle. Future research may address tissue-specific differences in expression changes.

The application of microarrays to identify rapidly diverging genes in nonmodel teleost species

In Chapter 4 of this thesis, I introduced the application of inter-species genomic hybridisation experiments for identifying rapidly diverging genes in teleosts. Previously, this application of microarray technology had been restricted to the study of microbial and fungal genomes. Rapidly diverging genes identified in this study of *P. moluccensis* and *D. rerio* included the homeobox genes *dlx5a, hoxb7a, hoxc11a,* and *msh-D*, and the receptor genes *adra2a, COUP-TF, ERR2, fzd8a,* and *reverbb2*. Accelerated rates of sequence evolution of receptor genes may indicate interaction with, and adaptation to, different environments. Homeobox gene clusters, on the other hand, have undergone independent duplication events throughout animal and vertebrate evolution (Holland *et al*. 1994, Amores *et al*. 1998, Pollard and Holland 2000, Martinez and Amemiya 2002, Chourrout *et al*. 2006). Gene duplication events may have resulted in a need to reduce genetic redundancies, and as a result, duplicate homeobox genes may have been under selective pressure to diverge rapidly (Malaga-Trillo and Meyer 2001). My data indicate rapid divergence of teleost homeobox genes and are consistent with the hypothesis of genomic instability after Hox cluster duplication events.

Genes with stress-related gene functions did not show accelerated rates of sequence evolution despite the fact that such genes may be expected to be under selective pressures during the process of adaptation to new environments. However, because genes with stress-related functions commonly have multiple functions associated with them, it is possible that these genes are under evolutionary constraints arising from there being multiple selective pressures operating on their evolution.

Direct sequence comparisons using publicly available genome sequence data confirmed accelerated rates of sequence evolution for a selection of rapidly diverging candidate genes in other teleost lineages. For example, the homeobox gene *msh-D*, the receptor gene *reverbb2,* the signal transduction gene *Tradd*, and the cell cycle

gene *geminin* showed some of the greatest rates of sequence evolution across multiple teleost lineages. While bioinformatic approaches can compare sequence similarity across the complete gene region, microarray approaches for measuring sequence similarity are restricted to the particular gene regions represented on the microarray. These partial gene regions represented on the microarray may not be representative of the level of sequence similarity across the complete gene. However, with further advances in microarray technology and the design of tiling arrays that interrogate the genome at very small intervals, the resolution of microarrays for measuring sequence similarity will only improve.

In summary of this fourth aim of my study, inter-species genomic hybridisation experiments can be used to identify rapidly diverging candidate genes in non-model species for which there are few genomic resources. Rapidly diverging genes identified in this study had functions in development and the interaction with the environment.

A conceptual model of the stress response

Lastly, in Chapter 5 of this thesis, I have synthesised data from the physiology, cell biology, and environmental genomics literature in order to identify key regulators of physiological and transcriptional stress responses. I have also identified a series of features common to most stress conditions. First, stress generally leads to a reduction in cellular oxygen levels. Reduced oxygen levels initiate endocrine, cardio-respiratory, and cellular responses that are aimed at improving oxygen delivery capacities. Second, stress is commonly associated with an increased production of reactive oxygen species and oxidative stress. Oxidative stress induces the expression of heat shock proteins and antioxidants and activates JNK and p38 signal transduction pathways, immediate early genes, and the transcription factors HIF-1α, p53, HSF, and NF-κB. Their activation ultimately produces the changes in the transcriptome and proteome that are observed in environmental genomic studies. There is cross-talk between many of the components of the stress response. I have pointed out some of these interactions. However, it is impossible to present a model inclusive of all interactions. I have focused instead on the features that are common to many stress conditions, key regulators of the stress response and the main

interactions between components at different level of biological organisation. The proposed model may help relate gene expression responses to stress responses at other levels of biological organisation. Finally, I have discussed the relationship between stress and adaptive evolution. I have discussed several mechanisms that can release genetic variation during conditions of stress and which can thus provide the raw material for selection to operate upon. I have discussed the implications of stress in the context of predicted climate change.

Overall Conclusions and Future Directions

Microarrays are powerful tools to investigate how the genetic make-up of an organism translates into biological function. There are many species and physiological responses that are of interest to functional genomic studies in ecology and evolution. It is hence unlikely that microarrays will be developed for all these species. In this PhD thesis I have discussed strategies that may be applied to species with few genomic resources. Inter-species genomic hybridisation experiments can be used to test the performance of a heterologous microarray. Heterologous microarray experiments can thus enhance the utility of microarray technology to the fields of environmental and functional genomics. Researchers at CSIRO, Australia are currently designing an oligonucleotide microarray for use across all vertebrate species. Future research may develop further microarrays specifically designed for use across species. In addition, future research may investigate in more detail the performance of heterologous microarray experiments to measure gene regulation across species.

While my research has identified extensive gene regulation in response to heat, the longer-term effects of heat exposure are still not understood. Future work may thus address how gene regulation is modulated over the longer term. Chronic heat may lead to transcript levels returning to pre-stress conditions. Alternatively, transcript levels may remain altered indicating a continuous strain on biological function. In addition, future research may investigate gene responses in different tissues and relate these to responses at other levels of biological organisation. The results of my PhD research indicated that exposure to heat over five days is associated with increased levels of protein and cellular damage. The repair of protein and cellular

damage would cause significant energetic costs and potentially reduce energy available for other biological functions, such as reproduction and growth. Future work may thus address these potential injurious effects of prolonged heat exposure on reproductive output and fitness.

In general, the relationship between gene regulation and physiological function are still poorly understood. I have proposed a conceptual model, outlining some of the key players and main interactions between transcriptional, molecular, cellular, and organismal stress responses. However, our understanding of stress responses would benefit from more research into the significance of gene regulation for biological function, including the experimental characterisation of gene functions during stress. Such research may measure gene regulation and relate gene expression profiles to other parameters of biological function, such as hormone secretion, cardio-respiratory function, reproductive output, swimming performance, etc. At present, much of this research has investigated these different aspects of the stress response in isolation. Future work may thus co-investigate responses at different levels of biological organisation.

Lastly, I have demonstrated that inter-species genomic hybridisation experiments can provide an alternative to bioinformatic approaches for the identification of rapidly diverging genes. Microarray approaches for the identification of rapidly diverging genes are particularly useful for species in which a lack of genome sequence data precludes bioinformatic approaches. In this context, future work may address the performance of different microarray platforms, e.g. cDNA, short and long oligonucleotide microarrays, for the identification of rapidly diverging genes. Cross-species hybridisation success to microarrays would differ depending on the design of the microarray. Since sequence similarity is inferred from relative hybridisation signals, future work may investigate the relationship between sequence similarity and cross-hybridisation on different microarray platforms.

Bibliography

Abele D, Burlando B, Viarengo A *et al.* (1998) Exposure to elevated temperatures and hydrogen peroxide elicits oxidative stress and antioxidant response in the Antarctic intertidal limpet *Nacella concinna*. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology,* **120,** 425-435.

Abele D, Heise K, Portner HO *et al.* (2002) Temperature-dependence of mitochondrial function and production of reactive oxygen species in the intertidal mud clam *Mya arenaria*. *Journal of Experimental Biology,* **205,** 1831-1841.

Abele D, Puntarulo S (2004) Formation of reactive species and induction of antioxidant defence systems in polar and temperate marine invertebrates and fish. *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology,* **138,** 405-415.

Adjaye J, Herwig R, Herrmann D *et al.* (2004) Cross-species hybridisation of human and bovine orthologous genes on high density cDNA microarrays. *BMC Genomics,* **5,** 83.

Airaksinen S, Rabergh CMI, Lahti A *et al.* (2003) Stressor-dependent regulation of the heat shock response in zebrafish, *Danio rerio*. *Comparative Biochemistry and Physiology A - Molecular & Integrative Physiology,* **134,** 839-846.

Airaksinen S, Rabergh CMI, Sistonen L *et al.* (1998) Effects of heat shock and hypoxia on protein synthesis in rainbow trout (*Oncorhynchus mykiss*) cells. *Journal of Experimental Biology,* **201,** 2543-2551.

Ames BN, Shigenaga MK, Hagen TM (1993) Oxidants, antioxidants, and the degenerative diseases of aging. *Proceedings of the National Academy of Sciences of the United States of America,* **90,** 7915-7922.

Amores A, Force A, Yan YL *et al.* (1998) Zebrafish hox clusters and vertebrate genome evolution. *Science,* **282,** 1711-1714.

An WG, Kanekal M, Simon MC *et al.* (1998) Stabilization of wild-type p53 by hypoxia-inducible factor 1[alpha]. *Nature,* **392,** 405.

Aoshiba K, Yasuda K, Yasui S *et al.* (2001) Serine proteases increase oxidative stress in lung cells. *American Journal Of Physiology-Lung Cellular And Molecular Physiology,* **281,** L556-L564.

Aubin-Horth N, Letcher BH, Hofmann HA (2005) Interaction of rearing environment and reproductive tactic on gene expression profiles in Atlantic salmon. *Journal of Heredity,* **96,** 261-278.

Barrett T, Suzek TO, Troup DB *et al.* (2005) NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Research,* **33,** D562-566.

Barry JP, Baxter CH, Sagarin RD *et al.* (1995) Climate-Related, Long-Term Faunal Changes in a California Rocky Intertidal Community. *Science,* **267,** 672-675.

Barton BA, Schreck CB (1987) Metabolic cost of acute physical stress in juvenile steelhead. *Transactions of the American Fisheries Society,* **116,** 257-263.

Beckman KB, Ames BN (1998) The free radical theory of aging matures. *Physiological Reviews,* **78,** 547-581.

Bell J, Neilson L, Pellegrini M (1988) Effect of heat shock on ribosome synthesis in *Drosophila melanogaster*. *Molecular And Cell Biology,* **8,** 91-95.

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate - A practical and powerful approach to multiple testing. *Journal Of The Royal Statistical Society Series B-Methodological,* **57,** 289-300.

Blomme T, Vandepoele K, De Bodt S *et al*. (2006) The gain and loss of genes during 600 million years of vertebrate evolution. *Genome Biology,* **7**.

Bly JE, Clem LW (1991) Temperature-mediated processes in teleost immunity - in vitro immunosuppression induced by in vivo low temperature in channel catfish. *Veterinary Immunology And Immunopathology,* **28,** 365-377.

Bonga SEW (1997) The stress response in fish. *Physiological Reviews,* **77,** 591-625.

Brand MD, Affourtit C, Esteves TC *et al*. (2004) Mitochondrial superoxide: Production, biological effects, and activation of uncoupling proteins. *Free Radical Biology And Medicine,* **37,** 755-767.

Brookes PS (2005) Mitochondrial H+ leak and ROS generation: An odd couple. *Free Radical Biology And Medicine,* **38,** 12-23.

Brunelle B, Nicholson T, Stephens R (2004) Microarray-based genomic surveying of gene polymorphisms in *Chlamydia trachomatis*. *Genome Biology,* **5,** R42.

Buckley BA, Gracey AY, Somero GN (2003) DNA microarray analysis of gene expression during heat stress in the eurythermal goby *Gillichthys mirabilis*. *Integrative and Comparative Biology,* **43,** 928-928.

Buckley BA, Hofmann GE (2002) Thermal acclimation changes DNA-binding activity of heat shock factor 1 (HSF1) in the goby *Gillichthys mirabilis*: Implications for plasticity in the heat-shock response in natural populations. *Journal of Experimental Biology,* **205,** 3231-3240.

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology,* **25,** 169-193.

Cai Y, Zhang C, Nawa T *et al*. (2000) Homocysteine-responsive ATF3 gene expression in human vascular endothelial cells: activation of c-Jun NH2-terminal kinase and promoter response element. *Blood,* **96,** 2140-2148.

Calow P (1991) Physiological costs of combating chemical toxicants: ecological implications. *Comparative Biochemistry And Physiology C-Pharmacology Toxicology & Endocrinology,* **100,** 3-6.

Cannon W (1929) *Bodily changes in pain, hunger, fear and rage: an account of recent researches into the function of emotional excitement,* New York, Appleton.

Cantatore P, Roberti M, Pesole G *et al*. (1994) Evolutionary analysis of cytochrome b sequences in some Perciformes - evidence for a slower rate of evolution than in mammals. *Journal of Molecular Evolution,* **39,** 589-597.

Carmeliet P, Dor Y, Herbert JM *et al*. (1998) Role of HIF-1 alpha in hypoxia-mediated apoptosis, cell proliferation and tumour angiogenesis. *Nature,* **394,** 485-490.

Chen WJ, Orti G, Meyer A (2004) Novel evolutionary relationship among four fish model systems. *Trends in Genetics,* **20,** 424-431.

Chourrout D, Delsuc F, Chourrout P *et al*. (2006) Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature,* **442,** 684-687.

Chow KC (2000) Hsp70 (DnaK) - an evolution facilitator? *Trends in Genetics,* **16,** 484-485.

Christoffels A, Koh EGL, Chia JM *et al*. (2004) *Fugu* genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. *Molecular Biology and Evolution,* **21,** 1146-1151.

Chrousos GP, Gold PW (1992) The concepts of stress and stress system disorders - overview of physical and behavioral homeostasis. *Journal Of The American Medical Association,* **267,** 1244-1252.

Clark AG, Glanowski S, Nielsen R *et al*. (2003) Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science,* **302,** 1960-1963.

Cohen DM (1997) Mitogen-activated protein kinase cascades and the signaling of hyperosmotic stress to immediate early genes. *Comparative Biochemistry And Physiology A-Physiology,* **117,** 291-299.

Cossins A, Fraser J, Hughes M *et al*. (2006) Post-genomic approaches to understanding the mechanisms of environmentally induced phenotypic plasticity. *Journal of Experimental Biology,* **209,** 2328-2336.

Cossins AR, Crawford DL (2005) Fish as models for environmental genomics. *Nature Reviews Genetics,* **6,** 324-333.

Cowan KJ, Storey KB (2003) Mitogen-activated protein kinases: new signaling pathways functioning in cellular responses to environmental stress. *Journal of Experimental Biology,* **206,** 1107-1115.

Cowen LE, Lindquist S (2005) Hsp90 potentiates the rapid evotution of new traits: Drug resistance in diverse fungi. *Science,* **309,** 2185-2189.

Crockett EL, Sidell BD (1990) Some pathways of energy metabolism are cold adapted in Antarctic fishes. *Physiological Zoology,* **63,** 472-488.

Crow KD, Stadler PF, Lynch VJ *et al*. (2006) The "fish-specific" Hox cluster duplication is coincident with the origin of teleosts. *Molecular Biology and Evolution,* **23,** 121-136.

Crow KD, Wagner GP (2006) What is the role of genome duplication in the evolution of complexity and diversity? *Molecular Biology and Evolution,* **23,** 887-892.

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change. *Science,* **292,** 673-679.

DeSalle R, Freedman T, Prager EM *et al*. (1987) Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *Journal of Molecular Evolution,* **26,** 157-164.

Dietz TJ (1994) Acclimation of the threshold induction temperatures for 70-kDa and 90-kDa heat-shock proteins in the fish *Gillichthys mirabilis*. *Journal of Experimental Biology,* **188,** 333-338.

Donaldson EM (1981) The pituary-interrenal axis as an indicator of stress in fish. In: *Stress in Fish* (eds. Pickering AD), London Academic.

Done T, Whetton P, Jones R *et al*. (2003) Global climate change and coral bleaching on the Great Barrier Reef. Final report to the State of Queensland Greenhouse Taskforce through the Department of Natural Resources and Mines.

Donoghue PCJ, Purnell MA (2005) Genome duplication, extinction and vertebrate evolution. *Trends in Ecology & Evolution,* **20,** 312-319.

Dulvy NK, Sadovy Y, Reynolds JD (2003) Extinction vulnerability in marine populations. *Fish and Fisheries,* **4,** 25-64.

Dunn KL, Espino PS, Drobic B *et al*. (2005) The Ras-MAPK signal transduction pathway, cancer and chromatin remodeling. *Biochemistry And Cell Biology-Biochimie Et Biologie Cellulaire,* **83,** 1-14.

Edge SE, Morgan MB, Gleason DF *et al*. (2005) Development of a coral cDNA array to examine gene expression profiles in *Montastraea faveolata* exposed to environmental stress. *Marine Pollution Bulletin,* **51,** 507-523.

Edvardsen RB, Seo HC, Jensen MF *et al*. (2005) Remodelling of the homeobox gene complement in the tunicate *Oikopleura dioica*. *Current Biology,* **15,** R12-R13.

Edwards MG, Sarkar D, Klopp R *et al*. (2004) Impairment of the transcriptional responses to oxidative stress in the heart of aged C57BL/6 mice. In: *Strategies For Engineered Negligible Senescence: Why Genuine Control Of Aging May Be Foreseeable*.

Eisen MB, Spellman PT, Brown PO *et al*. (1998) Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America,* **95,** 14863-14868.

Enjalbert B, Nantel A, Whiteway M (2003) Stress-induced gene expression in *Candida albicans*: Absence of a general stress response. *Molecular Biology of the Cell,* **14,** 1460-1467.

Etterson JR, Shaw RG (2001) Constraint to adaptive evolution in response to global warming. *Science,* **294,** 151-154.

Evertsz EM, Au-Young J, Ruvolo MV *et al*. (2001) Hybridization cross-reactivity within homologous gene families on glass cDNA microarrays. *Biotechniques,* **31,** 1182-1192.

Fader SC, Yu ZM, Spotila JR (1994) Seasonal variation in heat shock proteins (Hsp70) in stream fish under natural conditions. *Journal of Thermal Biology,* **19,** 335-341.

Fairhead C, Dujon B (2006) Structure of *Kluyveromyces lactis* subtelomeres: duplications and gene content. *FEMS Yeast Research,* **6,** 428-441.

Farrell AR (2002) Cardiorespiratory performance in salmonids during exercise at high temperature: insights into cardiovascular design limitations in fishes. *Comparative Biochemistry and Physiology A - Molecular and Integrative Physiology,* **132,** 797-810.

Feder ME, Hofmann GE (1999) Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. *Annual Review of Physiology,* **61,** 243-282.

Feder ME, Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Nature Reviews Genetics,* **4,** 651-657.

Felsenstein J (1989) PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics,* **5,** 164-166.

Felsenstein J, Churchill GA (1996) A hidden Markov Model approach to variation among sites in rate of evolution. *Molecular Biology and Evolution,* **13,** 93-104.

Finkel T, Holbrook NJ (2000) Oxidants, oxidative stress and the biology of ageing. *Nature,* **408,** 239-247.

Fischman HK, Pero RW, Kelly DD (1996) Psychogenic stress induces chromosomal and DNA damage. *International Journal Of Neuroscience,* **84,** 219-227.

Flanagan SW, Moseley PL, Buettner GR (1998) Increased flux of free radicals in cells subjected to hyperthermia: detection by electron paramagnetic resonance spin trapping. *FEBS Letters,* **431,** 285-286.

Forlano PM, Deitcher DL, Bass ARH (2005) Distribution of estrogen receptor alpha mRNA in the brain and inner ear of a vocal fish with comparisons to sites of aromatase expression. *Journal of Comparative Neurology,* **483,** 91-113.

Freeman ML, Spitz DR, Meredith MJ (1990) Does Heat-Shock Enhance Oxidative Stress - Studies With Ferrous And Ferric Iron. *Radiation Research,* **124,** 288-293.

Froschauer A, Braasch I, Volff JN (2006) Fish genomes, comparative genomics and vertebrate evolution. *Current Genomics,* **7,** 43-57.

Fryer JN (1989) Neuropeptides regulating the activity of goldfish corticotropes and melanotropes. *Fish Physiology And Biochemistry,* **7,** 21-27.

Gabai VL, Sherman MY (2002) Interplay between molecular chaperones and signaling pathways in survival of heat shock. *Journal of Applied Physiology,* **92,** 1743-1748.

Gardner TA, Cote IM, Gill JA *et al.* (2003) Long-term region-wide declines in Caribbean corals. *Science,* **301,** 958-960.

Gasch AP, Spellman PT, Kao CM *et al.* (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell,* **11,** 4241-4257.

Gedamu L, Zafarullah M (1993) Molecular analysis of rainbow trout metallothionein and stress protein genes: structure, expression and regulation. In: *Biochemistry and Molecular Biology of Fishes* (eds. Hochachka PW, Mommsen TP) Amsterdam, Elsevier.

Girke T, Todd J, Ruuska S *et al*. (2000) Microarray analysis of developing *Arabidopsis* seeds. *Plant Physiology,* **124,** 1570-1581.

Giulietti A, Overbergh L, Valckx D *et al.* (2001) An overview of real-time quantitative PCR: Applications to quantify cytokine gene expression. *Methods,* **25,** 386-401.

Gius D, Mattson D, Bradbury CM *et al*. (2004) Thermal stress and the disruption of redox-sensitive signalling and transcription factor activation: possible role in radiosensitization. *International Journal Of Hyperthermia,* **20,** 213-223.

Glynn PW, Deweerdt WH (1991) Elimination of two reef-building hydrocorals following the 1982-83 El-Nino warming event. *Science,* **253,** 69-71.

Gracey AY, Fraser EJ, Li W *et al*. (2004) Coping with cold: An integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 16970-16975.

Gracey AY, Troll JV, Somero GN (2001) Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*. *Proceedings of the National Academy of Sciences of the United States of America,* **98,** 1993-1998.

Guderley H (1990) Functional significance of metabolic responses to thermal acclimation in fish muscle. *American Journal of Physiology,* **259,** R245-R252.

Guderley H, Blier P (1988) Thermal acclimation in fish: conservative and labile properties of swimming muscle. *Canadian Journal Of Zoology,* **66,** 1105-1115.

Hall A (2005) Rho GTPases and the control of cell behaviour. *Biochemical Society Transactions,* **33,** 891-895.

Harr B (2006) Genomic islands of differentiation between house mouse subspecies. *Genome Research,* **16,** 730-737.

Hawkins AJS (1991) Protein turnover: a functional appraisal. *Functional Ecology,* **5,** 222-233.

Heise K, Puntarulo S, Nikinmaa M *et al*. (2006a) Oxidative stress during stressful heat exposure and recovery in the North Sea eelpout *Zoarces viviparus* L. *Journal of Experimental Biology,* **209,** 353-363.

Heise K, Puntarulo S, Nikinmaa M *et al.* (2006b) Oxidative stress and HIF-1 DNA binding during stressful cold exposure and recovery in the North Sea eelpout (*Zoarces viviparus*). *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology,* **143,** 494-503.

Heise K, Puntarulo S, Portner HO *et al*. (2003) Production of reactive oxygen species by isolated mitochondria of the Antarctic bivalve *Laternula elliptica* (King and Broderip) under heat stress. *Comparative Biochemistry And Physiology C-Toxicology & Pharmacology,* **134,** 79-90.

Hellberg ME, Balch DP, Roy K (2001) Climate-driven range expansion and morphological evolution in a marine gastropod. *Science,* **292,** 1707-1710.

Hellberg ME, Vacquier VD (1999) Rapid evolution of fertilization selectivity and lysin cDNA sequences in teguline gastropods. *Molecular Biology and Evolution,* **16,** 839-848.

Herschman HR (1991) Primary Response Genes Induced By Growth Factors And Tumor Promoters. *Annual Review Of Biochemistry,* **60,** 281-319.

Hinchliffe SJ, Isherwood KE, Stabler RA *et al.* (2003) Application of DNA microarrays to study the evolutionary genomics of *Yersinia pestis* and *Yersinia pseudotuberculosis*. *Genome Research,* **13,** 2018-2029.

Hirota K, Semenza GL (2006) Regulation of angiogenesis by hypoxia-inducible factor 1. *Critical Reviews In Oncology Hematology,* **59,** 15-26.

Hittel D, Storey KB (2001) Differential expression of adipose- and heart-type fatty acid binding proteins in hibernating ground squirrels. *Biochimica et Biophysica Acta,* **1522,** 238-243.

Hochachka PW, Somero GN (2002) *Biochemical Adaptation. Mechanisms and process in physiological evolution,* Oxford, Oxford University Press.

Hoegh-Guldberg O (1999) Climate change, coral bleaching and the future of the world's coral reefs. *Marine & Freshwater Research,* **50,** 839-866.

Hoffmann AA (1995) Acclimation: increasing survival at a cost. *Trends in Ecology & Evolution,* **10,** 1-2.

Hoffmann AA, Blows MW (1994) Species borders: ecological and evolutionary perspectives. *Trends in Ecology & Evolution,* **9,** 223-227.

Hoffmann AA, Hallas RJ, Dean JA *et al.* (2003) Low potential for climatic stress adaptation in a rainforest *Drosophila* species. *Science,* **301,** 100-102.

Hoffmann AA, Parsons PA (1991) Stress, species margins, and conservation. In: *Evolutionary Genetics and Environmental Stress* Oxford, Oxford University Press.

Hoffmann R, Valencia A (2004) A gene network for navigating the literature. *Nature Genetics,* **36,** 664.

Hofmann GE, Buckley BA, Airaksinen S *et al.* (2000) Heat-shock protein expression is absent in the Antarctic fish *Trematomus bernacchii* (family Nototheniidae). *Journal of Experimental Biology,* **203,** 2331-2339.

Hofmann GE, Somero GN (1995) Evidence for Protein Damage at Environmental Temperatures - Seasonal-Changes in Levels of Ubiquitin Conjugates and Hsp70 in the Intertidal Mussel *Mytilus trossulus*. *Journal of Experimental Biology,* **198,** 1509-1518.

Hofmann GE, Somero GN (1996) Interspecific variation in thermal denaturation of proteins in the congeneric mussels *Mytilus trossulus* and *M. galloprovincialis*: Evidence from the heat shock response and protein ubiquitination. *Marine Biology,* **126,** 65-75.

Hogstrand C, Balesaria S, Glover CN (2002) Application of genomics and proteomics for study of the integrated response to zinc exposure in a non-model fish species, the rainbow trout. *Comparative Biochemistry and Physiology B - Biochemistry and Molecular Biology,* **133,** 523-535.

Holland PWH, Garcia-Fernandez J, Williams NA *et al.* (1994) Gene Duplications And The Origins Of Vertebrate Development. *Development***,** Suppl., 125-133.

Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65-70.

Houlihan DF (1991) Protein turnover in ectotherms and its relationships to energetics. *Advances in comparative and environmental Physiology,* **7,** 1-43.

Hughes AL (1993) Contrasting evolutionary rates in the duplicate chaperonin genes of *Mycobacterium tuberculosis* and *Mycobacterium leprae*. *Molecular Biology and Evolution,* **10,** 1343-1359.

Hughes P, Dragunow M (1995) Induction Of Immediate-Early Genes And The Control Of Neurotransmitter-Regulated Gene-Expression Within The Nervous-System. *Pharmacological Reviews,* **47,** 133-178.

Hughes TP (1994) Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. *Science,* **265,** 1547-1550.

Hughes TP, Baird AH, Bellwood DR *et al.* (2003) Climate change, human impacts, and the resilience of coral reefs. *Science,* **301,** 929-933.

Hyllner SJ, Andersson T, Haux C *et al.* (1989) Cortisol induction of metallothionein in primary culture of rainbow trout hepatocytes. *Journal Of Cellular Physiology,* **139,** 24-28.

Inuzuka H, Nanbu-Wakao R, Masuho Y *et al.* (1999) Differential regulation of immediate early gene expression in preadipocyte cells through multiple signaling pathways. *Biochemical and Biophysical Research Communications,* **265,** 664-668.

IPCC (2001) Climate change 2001. IPCC Third Assessment Report. Port Chester, NY, Cambridge University Press.

Irizarry RA, Warren D, Spencer F *et al.* (2005) Multiple-laboratory comparison of microarray platforms. *Nature Methods,* **2,** 345-350.

Iwama GK, Afonso LOB, Todgham A *et al.* (2004) Are hsps suitable for indicating stressed states in fish? *Journal of Experimental Biology,* **207,** 15-19.

Iwama GK, Afonso LOB, Vijayan MM (2006) Stress in Fishes. In: *The Physiology of Fishes* (eds. Evans DH, Claiborne JB) 3rd edition ed. Boca Raton, FL, Taylor & Francis.

Iwama GK, Thomas PT, Forsyth RHB *et al.* (1998) Heat shock protein expression in fish. *Reviews In Fish Biology And Fisheries,* **8,** 35-56.

Jackson JBC, Kirby MX, Berger WH *et al.* (2001) Historical overfishing and the recent collapse of coastal ecosystems. *Science,* **293,** 629-638.

Jaillon O, Aury JM, Brunet F *et al.* (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature,* **431,** 946-957.

Janssens PA, Waterman J (1988) Hormonal regulation of gluconeogenesis and glycogenolysis in carp (*Cyprinus carpio*) liver pieces cultured *in vitro*. *Comparative Biochemistry And Physiology A-Physiology,* **91,** 451-455.

Ji W, Zhou W, Gregg K *et al.* (2004) A method for cross-species gene expression analysis with high-density oligonucleotide arrays. *Nucleic Acids Research,* **32,** e93.

Johnson EO, Kamilaris TC, Chrousos GP *et al.* (1992) Mechanisms of stress: A dynamic overview of hormonal and behavioral homeostasis. *Neuroscience & Biobehavioral Reviews,* **16,** 115-130.

Jordan IK, Kondrashov F, Rogozin I *et al.* (2001) Constant relative rate of protein evolution and detection of functional diversification among bacterial, archaeal and eukaryotic proteins. *Genome Biology,* **2,** 53.51-53.59.

Ju Z, Dunham RA, Liu Z (2002) Differential gene expression in the brain of channel catfish (Ictalurus punctatus) in response to cold acclimation. *Molecular Genetics And Genomics,* **268,** 87-95.

Kane M, Jatkoe T, Stumpf C *et al.* (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Research,* **28,** 4552-4557.

Katschinski DM, Le L, Heinrich D *et al.* (2002) Heat induction of the unphosphorylated form of hypoxia-inducible factor-1 alpha is dependent on heat shock protein-90 activity. *Journal of Biological Chemistry,* **277,** 9262-9267.

Keller M, Sommer AM, Pörtner HO *et al.* (2004) Seasonality of energetic functioning and production of reactive oxygen species by lugworm (*Arenicola marina*) mitochondria exposed to acute temperature changes. *Journal of Experimental Biology,* **207,** 2529-2538.

Kikuchi K, Yamashita M, Watabe S *et al.* (1995) The Warm Temperature Acclimation-Related 65-Kda Protein, Wap65, In Goldfish And Its Gene-Expression. *Journal of Biological Chemistry,* **270,** 17087-17092.

Kim C, Joyce E, Chan K *et al.* (2002) Improved analytical methods for microarray-based genome-composition analysis. *Genome Biology,* **3,** 65.61-65.17.

Kim Y, Nueda A, Meng YH *et al.* (1997) Analysis of the phosphorylation of human heat shock transcription factor-1 by MAP kinase family members. *Journal of Cellular Biochemistry,* **67,** 43-54.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *American Naturalist,* **150,** 1-23.

Kline MP, Morimoto RI (1997) Repression of the heat shock factor 1 transcriptional activation domain is modulated by constitutive phosphorylation. *Molecular And Cellular Biology,* **17,** 2107-2115.

Koehn RK, Bayne BL (1989) Towards a physiological and genetic understanding of the energetics of the stress response. *Biological Journal of the Linnean Society,* **37,** 157-171.

Koskinen H, Pehkonen P, Vehniainen E *et al.* (2004) Response of rainbow trout transcriptome to model chemical contaminants. *Biochemical and Biophysical Research Communications,* **320,** 745-753.

Krasnov A, Koskinen H, Pehkonen P *et al.* (2005) Gene expression in the brain and kidney of rainbow trout in response to handling stress. *BMC Genomics,* **6**.

Kubota H (2002) Function and regulation of cytosolic molecular chaperone CCT. *Vitamins and Hormones,* **65,** 313-331.

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. *Briefings in Bioinformatics,* **5,** 150-163.

Kyriakis JM, Avruch J (1996) Protein kinase cascades activated by stress and inflammatory cytokines. *Bioessays,* **18,** 567-577.

Lando D, Peet DJ, Whelan DA *et al.* (2002) Asparagine hydroxylation of the HIF transactivation domain: A hypoxic switch. *Science,* **295,** 858-861.

Lannig G, Bock C, Sartoris FJ *et al.* (2004) Oxygen limitation of thermal tolerance in cod, *Gadus morhua* L., studied by magnetic resonance imaging and on-line venous oxygen monitoring. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology,* **287,** R902-R910.

Lau LF, Nathans D (1985) Identification of a set of genes expressed during the G0/G1 transition of cultured mouse cells. *Embo Journal,* **4,** 3145-3151.

Le Quere A, Eriksen KA, Rajashekar B *et al.* (2006) Screening for rapidly evolving genes in the ectomycorrhizal fungus *Paxillus involutus* using cDNA microarrays. *Molecular Ecology,* **15,** 535-550.

Lecompte O, Ripp R, Puzos-Barbe V *et al.* (2001) Genome evolution at the genus level: Comparison of three complete genomes of hyperthermophilic Archaea. *Genome Research,* **11,** 981-993.

Lee HK, Braynen W, Keshav K *et al.* (2005) ErmineJ: Tool for functional analysis of gene expression data sets. *BMC Bioinformatics,* **6**.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution,* **17,** 183-189.

Lesser MP (2006) Oxidative stress in marine environments: Biochemistry and physiological ecology. *Annual Review of Physiology,* **68,** 253-278.

Logue JA, De Vries AL, Fodor E *et al.* (2000) Lipid compositional correlates of temperature-adaptive interspecific differences in membrane physical structure. *Journal of Experimental Biology,* **203,** 2105-2115.

Lupes SC, Davis MW, Olla BL *et al.* (2006) Capture-related stressors impair immune system function in sableflsh. *Transactions of the American Fisheries Society,* **135,** 129-138.

Lushchak VI, Bagnyukova TV (2006) Effects of different environmental oxygen levels on free radical processes in fish. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology,* **144,** 283-289.

Luz H, Vingron M (2006) Family specific rates of protein evolution. *Bioinformatics,* **22,** 1166-1171.

Makalowski W, Zhang JH, Boguski MS (1996) Comparative analysis of 1196 orthologous mouse and human full-length mRNA and protein sequences. *Genome Research,* **6,** 846-857.

Malaga-Trillo E, Meyer A (2001) Genome duplications and accelerated evolution of Hox genes and cluster architecture in teleost fishes. *American Zoologist,* **41,** 676-686.

Mark FC, Bock C, Portner HO (2002) Oxygen-limited thermal tolerance in Antarctic fish investigated by MRI and P-31-MRS. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology,* **283,** R1254-R1262.

Martinez P, Amemiya CT (2002) Genomics of the HOX gene cluster. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology,* **133,** 571-580.

Masel J (2005) Evolutionary capacitance may be favored by natural selection. *Genetics,* **170,** 1359-1371.

Massague J (2004) G1 cell-cycle control and cancer. *Nature,* **432,** 298-306.

Matzkin LM (2005) Activity variation in alcohol dehydrogenase paralogs is associated with adaptation to cactus host use in cactophilic *Drosophila*. *Molecular Ecology,* **14,** 2223-2231.

McMahon SB, Monroe JG (1992) Role Of Primary Response Genes In Generating Cellular-Responses To Growth-Factors. *Faseb Journal,* **6,** 2707-2715.

Medhora M, Bousamra M, Zhu DL *et al.* (2002) Upregulation of collagens detected by gene array in a model of flow-induced pulmonary vascular remodeling. *American Journal of Physiology - Heart and Circulatory Physiology,* **282,** H414-H422.

Meiklejohn CD, Hartl DL (2002) A single mode of canalization. *Trends in Ecology & Evolution,* **17,** 468-473.

Melov S, Ravenscroft J, Malik S *et al.* (2000) Extension of life-span with superoxide dismutase/catalase mimetics. *Science,* **289,** 1567-1569.

Mercurio F, Manning AM (1999) NF-kappa B as a primary regulator of the stress response. *Oncogene,* **18,** 6163-6171.

Metscher BD, Ahlberg PE (1999) Zebrafish in context: Uses of a laboratory model in comparative studies. *Developmental Biology,* **210,** 1-14.

Miller NA, Gong Q, Bryan R *et al.* (2002) Cross-hybridization of closely related genes on high-density macroarrays. *Biotechniques,* **32,** 620-625.

Miranda EJ, Hazel JR (1996) Temperature-induced changes in the transbilayer distribution of phosphatidylethanolamine in mitoplasts of rainbow trout (*Oncorhynchus mykiss*) liver. *Journal of Experimental Zoology,* **274,** 23-32.

Moberg G (2000) Biological response to stress: implications for animal welfare. In: *The biology of animal stress: basic principles and implications for animal welfare* (eds. Moberg G, Mench J) New York, CABI Publishing.

Mohn KL, Laz TM, Hsu JC *et al.* (1991) The immediate-early growth response in regenerating liver and insulin-stimulated H-35 cells: comparison with serum-stimulated 3T3 cells and identification of 41 novel immediate-early genes. *Molecular And Cellular Biology,* **11,** 381-390.

Mommsen TP, Walsh PJ, Perry SF *et al*. (1988) Interactive effects of catecholamines and hypercapnia on glucose production in isolated trout hepatocytes. *General And Comparative Endocrinology,* **70,** 63-73.

Moody D, Zou Z, McIntyre L (2002) Cross-species hybridisation of pig RNA to human nylon microarrays. *BMC Genomics,* **3,** 27.

Morales AE, Perez-Jimenez A, Hidalgo MC *et al*. (2004) Oxidative stress and antioxidant defenses after prolonged starvation in *Dentex dentex* liver. *Comparative Biochemistry And Physiology C-Toxicology & Pharmacology,* **139,** 153-161.

Morgan JI, Curran T (1995) Immediate-early genes: ten years on. *Trends in Neurosciences,* **18,** 66-67.

Moseley PL (1997) Heat shock proteins and heat adaptation of the whole organism. *Journal of Applied Physiology,* **83,** 1413-1417.

Mulley J, Holland P (2004) Comparative genomics - Small genome, big insights. *Nature,* **431,** 916-917.

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research,* **19,** 4321-4325.

Nakano K, Takemura A, Nakamura S *et al*. (2004) Changes in the cellular and organismal stress responses of the subtropical fish, the Indo-Pacific sergeant, *Abudefduf vaigiensis*, due to the 1997-1998 El Nino Southern Oscillation. *Environmental Biology of Fishes,* **70,** 321-329.

Nikinmaa M (1992a) How does environmental pollution affect red cell function in fish. *Aquatic Toxicology,* **22,** 227-238.

Nikinmaa M (1992b) Membrane-transport and control of hemoglobin-oxygen affinity in nucleated erythrocytes. *Physiological Reviews,* **72,** 301-321.

Nowrousian M, Ringelberg C, Dunlap JC *et al*. (2005) Cross-species microarray hybridization to identify developmentally regulated genes in the filamentous fungus *Sordaria macrospora*. *Molecular Genetics And Genomics,* **273,** 137-149.

Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. *Nature Genetics,* **32,** 261-266.

Oleksiak MF, Kolell KJ, Crawford DL (2001) Utility of natural populations for microarray analyses: Isolation of genes necessary for functional genomic studies. *Marine Biotechnology,* **3,** S203-S211.

Oleksiak MF, Roach JL, Crawford DL (2005) Natural variation in cardiac metabolism and gene expression in *Fundulus heteroclitus*. *Nature Genetics,* **37,** 67-72.

Olivereau M, Olivereau JM (1990) Corticotropin-like immunoreactivity in the brain and pituitary of three teleost species (goldfish, trout and eel). *Cell And Tissue Research,* **262,** 115-123.

Pandolfi JM, Bradbury RH, Sala E *et al*. (2003) Global trajectories of the long-term decline of coral reef ecosystems. *Science,* **301,** 955-958.

Pankhurst NW, Van Der Kraak G (1997) Effects of stress on reproduction and growth of fish. In: *Fish Stress and Health in Aquaculture* (eds. Iwama GK, Pickering AD *et al*.) Cambridge, Cambridge University Press.

Parihar MS, Javeri T, Hemnani T *et al*. (1997) Responses of superoxide dismutase, glutathione peroxidase and reduced glutathione antioxidant defenses in gills of the freshwater catfish (*Heteropneustes fossilis*) to short-term elevated temperature. *Journal of Thermal Biology,* **22,** 151-156.

Parker RO, Dixon RL (1998) Changes in a North Carolina reef fish community after 15 years of intense fishing - Global warming implications. *Transactions of the American Fisheries Society,* **127,** 908-920.

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature,* **421,** 37-42.

Parsell DA, Kowal AS, Singer MA *et al*. (1994) Protein Disaggregation Mediated by Heat-Shock Protein Hsp104. *Nature,* **372,** 475-478.

Parsell DA, Lindquist S (1993) The Function of Heat-Shock Proteins in Stress Tolerance - Degradation and Reactivation of Damaged Proteins. *Annual Review of Genetics,* **27,** 437-496.

Partridge L, Barton NH (2000) Evolving evolvability. *Nature,* **407,** 457-458.

Pascual R, Burgos FJ, Salva M *et al*. (1989) Purification and properties of five different forms of human procarboxypeptidases. *European Journal of Biochemistry,* **179,** 609-616.

Paulay Gb (1996) Dynamic clams: Changes in the bivalve fauna of Pacific islands as a result of sea-level fluctuations. *American Malacological Bulletin,* **12,** 45-57.

Pedersen SF, Hoffmann EK, Mills JW (2001) The cytoskeleton and cell volume regulation. *Comparative Biochemistry and Physiology a-Molecular and Integrative Physiology,* **130,** 385-399.

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research,* **30,** E36-E36.

Pfaffl MW, Tichopad A, Prgomet C *et al*. (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper - Excel-based tool using pair-wise correlations. *Biotechnology Letters,* **26,** 509-515.

Picard DJ, Schulte PM (2004) Variation in gene expression in response to stress in two populations of *Fundulus heteroclitus*. *Comparative Biochemistry and Physiology A - Molecular and Integrative Physiology,* **137,** 205-216.

Pittock AB (1999) Coral reefs and environmental change: Adaptation to what? *American Zoologist,* **39,** 10-29.

Podrabsky JE, Somero GN (2004) Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. *Journal of Experimental Biology,* **207,** 2237-2254.

Pollack JR, Perou CM, Alizadeh AA *et al.* (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nature Genetics,* **23,** 41-46.

Pollard SL, Holland PWH (2000) Evidence for 14 homeobox gene clusters in human genome ancestry. *Current Biology,* **10,** 1059-1062.

Porwollik S, Wong RMY, McClelland M (2002) Evolutionary genomics of *Salmonella*: Gene acquisitions revealed by microarray analysis. *Proceedings of the National Academy of Sciences of the United States of America,* **99,** 8956-8961.

Pottinger TG (1999) The impact of stress on animal reproductive activities. In: *Stress Physiology in Animals* (eds. Balm PHM) Sheffield, England, Sheffield Academic Press.

Powers DA, Schulte PM (1998) Evolutionary adaptations of gene structure and expression in natural populations in relation to a changing environment: A multidisciplinary approach to address to million-year saga of a small fish. *Journal of Experimental Zoology,* **282,** 71-94.

Pörtner HO (2002) Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. *Comparative Biochemistry and Physiology A - Molecular and Integrative Physiology,* **132,** 739-761.

Pörtner HO, Berdal B, Blust R *et al.* (2001) Climate induced temperature effects on growth performance, fecundity and recruitment in marine fish: developing a hypothesis for cause and effect relationships in Atlantic cod (*Gadus morhua*) and common eelpout (*Zoarces viviparus*). *Continental Shelf Research,* **21,** 1975-1997.

Pörtner HO, Mark FC, Bock C (2004) Oxygen limited thermal tolerance in fish? Answers obtained by nuclear magnetic resonance techniques. *Respiratory Physiology & Neurobiology,* **141,** 243-260.

Queitsch C, Sangster TA, Lindquist S (2002) Hsp90 as a capacitor of phenotypic variation. *Nature,* **417,** 618-624.

Rambaut A (1996) *Se-Al: Sequence Alignment Editor. Version 1.0 alpha 1.,* Oxford, University of Oxford.

Rechsteiner M (1987) Ubiquitin-Mediated Pathways for Intracellular Proteolysis. *Annual Review of Cell Biology,* **3,** 1-30.

Renn SCP, Aubin-Horth N, Hofmann HA (2004) Biologically meaningful expression profiling across species using heterologous hybridization to a cDNA microarray. *BMC Genomics,* **5**.

Rifkin JM, Abugo O, Levy A *et al.* (1993) Formation of free radicals under hypoxia. In: *Surviving Hypoxia: Mechanisms of Control and Adaptation.* (eds. Hochachka PW, Lutz PL *et al.*) Boca Raton, FL, CRC Press.

Rise ML, von Schalburg KR, Brown GD *et al.* (2004) Development and application of a salmonid EST database and cDNA microarray: Data mining and interspecific hybridization characteristics. *Genome Research,* **14,** 478-490.

Robinson P, Wollstein A, Bohme U *et al.* (2004) Ontologizing gene-expression microarray data: characterizing clusters with Gene Ontology. *Bioinformatics,* **20,** 979-981.

Roemmich D, McGowan J (1995) Climatic Warming and the Decline of Zooplankton in the California Current. *Science,* **267,** 1324-1326.

Roest Crollius H, Weissenbach J (2005) Fish genomics and biology. *Genome Research,* **15,** 1675-1682.

Rollins BJ, Stiles CD (1988) Regulation of *c-myc* and *c-fos* proto-oncogene expression by animal cell growth factors. *In Vitro Cellular & Developmental Biology,* **24,** 81-84.

Romero LM (2004) Physiological stress in ecology: lessons from biomedical research. *Trends in Ecology & Evolution,* **19,** 249-255.

Rose TM, Henikoff JG, Henikoff S (2003) CODEHOP (COnsensus-DEgenerate hybrid oligonucleotide primer) PCR primer design. *Nucleic Acids Research,* **31,** 3763-3766.

Rose TM, Schultz ER, Henikoff JG *et al.* (1998) Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Research,* **26,** 1628-1635.

Roy K, Jablonski D, Valentine JW (2001) Climate change, species range limits and body size in marine bivalves. *Ecology Letters,* **4,** 366-370.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. Krawetz S, Misener S) Totowa, NJ, Humana Press.

Ruddle FH, Bartels JL, Bentley KL *et al.* (1994) Evolution of Hox genes. *Annu Rev Genet,* **28,** 423-442.

Rutherford SL, Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. *Nature,* **396,** 336-342.

Ryan SN, Pankhurst NW, Wells RMG (1995) A possible role for ubiquitin in the stress response of the teleost fish blue mao mao (*Scorpis violaceus*). *Physiological Zoology,* **68,** 1077-1092.

Sarmiento J, Leal S, Quezada C *et al.* (2000) Environmental acclimatization of the carp modulates the transcription of beta-actin. *Journal of Cellular Biochemistry,* **80,** 223-228.

Scheiner SM (2002) Selection experiments and the study of phenotypic plasticity. *Journal of Evolutionary Biology,* **15,** 889-898.

Schena M, Shalon D, Davis RW *et al.* (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science,* **270,** 467-470.

Schenk H, Klein M, Erdbrugger W *et al.* (1994) Distinct effects of thioredoxin and antioxidants on the activation of transcription factors Nf-Kappa-B and AP-1. *Proceedings of the National Academy of Sciences of the United States of America,* **91,** 1672-1676.

Schmedtje JF, Ji YS (1998) Hypoxia and molecular cardiovascular medicine. *Trends In Cardiovascular Medicine,* **8,** 24-33.

Schreck CB, Bradford CS, Fitzpatrick MS *et al*. (1989) Regulation of the interrenal of fishes - non-classical control mechanisms. *Fish Physiology And Biochemistry,* **7,** 259-265.

Semenza GL (1998) Hypoxia-inducible factor 1: master regulator of O_2 homeostasis. *Current Opinion In Genetics & Development,* **8,** 588-594.

Semenza GL (2000) HIF-1: mediator of physiological and pathophysiological responses to hypoxia. *Journal of Applied Physiology,* **88,** 1474-1480.

Semenza GL (2004a) Hydroxylation of HIF-1: Oxygen sensing at the molecular level. *Physiology,* **19,** 176-182.

Semenza GL (2004b) O_2-regulated gene expression: transcriptional control of cardiorespiratory physiology by HIF-1. *Journal of Applied Physiology,* **96,** 1173-1177.

Semenza GL (2006) Regulation of physiological responses to continuous and intermittent hypoxia by hypoxia-inducible factor 1. *Experimental Physiology,* **91,** 803-806.

Semenza GL, Roth PH, Fang HM *et al*. (1994) Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor-1. *Journal of Biological Chemistry,* **269,** 23757-23763.

Seyle H (1936) A syndrome produced by diverse nocuous agents. *Nature,* **138,** 32.

Shaulian E, Karin M (2002) AP-1 as a regulator of cell life and death. *Nature Cell Biology,* **4,** E131-E136.

Sheridan MA (1994) Regulation of lipid metabolism in poikilothermic vertebrates. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology,* **107,** 495-508.

Shigenaga MK, Hagen TM, Ames BN (1994) Oxidative damage and mitochondrial decay in aging. *Proceedings of the National Academy of Sciences of the United States of America,* **91,** 10771-10778.

Sidow A (1996) Gen(om)e duplications in the evolution of early vertebrates. *Curr Opin Genet Dev,* **6,** 715-722.

Siegal ML, Bergman A (2002) Waddington's canalization revisited: Developmental stability and evolution. *Proceedings of the National Academy of Sciences of the United States of America,* **99,** 10528-10532.

Sisson JE, Sidell BD (1987) Effect of thermal acclimation on muscle fiber recruitment of swimming striped bass (*Morone saxatilis*). *Physiological Zoology,* **60,** 310-320.

Smyth G (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology,* **3,** Article 3.

Smyth G, Speed T (2003) Normalization of cDNA microarray data. *Methods,* **31,** 265-273.

Smyth GK (2005) Limma: linear models for microarray data. In: *Bioinformatics and Computational Biology Solutions using R and Bioconductor* (eds. Gentleman R, Carey V *et al.*) New York, Springer.

Sohal RS, Agarwal S, Dubey A *et al.* (1993) Protein oxidative damage is associated with life expectancy of houseflies. *Proceedings of the National Academy of Sciences of the United States of America,* **90,** 7255-7259.

Sohal RS, Weindruch R (1996) Oxidative stress, caloric restriction, and aging. *Science,* **273,** 59-63.

Soitamo AJ, Rabergh CMI, Gassmann M *et al.* (2001) Characterization of a Hypoxia-inducible Factor (HIF-1alpha) from Rainbow Trout. *Journal of Biological Chemistry,* **276,** 19699-19705.

Soldatov AA (1996) The effect of hypoxia on red blood cells of flounder: a morphologic and autoradiographic study. *Journal of Fish Biology,* **48,** 321-328.

Somero G (2005) Linking biogeography to physiology: Evolutionary and acclimatory adjustments of thermal limits. *Frontiers in Zoology,* **2,** 1.

Somero GN, Hofmann GE (1996) Temperature thresholds for protein adaptation: when does temperature change start to 'hurt'? In: *Global Warming: Implications for freshwater and marine fish* (eds. Wood CM, McDonald DG), Cambridge University Press.

Sommer A, Klein B, Portner HO (1997) Temperature induced anaerobiosis in two populations of the polychaete worm *Arenicola marina* (L). *Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology,* **167,** 25-35.

Sonna LA, Fujita J, Gaffin SL *et al.* (2002) Effects of heat and cold stress on mammalian gene expression. *Journal of Applied Physiology,* **92,** 1725-1742.

Sorensen JG, Kristensen TN, Loeschcke V (2003) The evolutionary and ecological role of heat shock proteins. *Ecology Letters,* **6,** 1025-1037.

St-Pierre J, Charest PM, Guderley H (1998) Relative contribution of quantitative and qualitative changes in mitochondria to metabolic compensation during seasonal acclimatisation of rainbow trout *Oncorhynchus mykiss*. *Journal of Experimental Biology,* **201,** 2961-2970.

Stillman JH (2003) Acclimation capacity underlies susceptibility to climate change. *Science,* **301,** 65-65.

Sumpter JP, Pottinger TG, Rand-Weaver M *et al.* (1994) The wide-ranging effects of stress on fish. In: *Perspectives in Comparative Endocrinology* (eds. Davey KG, Peter RE *et al.*) Ottawa, National Research Council of Canada.

Swanson WJ, Clark AG, Waldrip-Dail HM *et al.* (2001) Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America,* **98,** 7375-7379.

Swanson WJ, Vacquier VD (1998) Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein. *Science,* **281,** 710-712.

Tautz D, Schmid KJ (1998) From genes to individuals: developmental genes and the generation of the phenotype. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* **353,** 231-240.

The Gene Ontology Consortium (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics,* **25,** 25-29.

Thompson JD, Gibson TJ, Plewniak F *et al.* (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research,* **25,** 4876-4882.

Tomanek L, Somero GN (1999) Evolutionary and acclimation-induced variation in the heat-shock responses of congeneric marine snails (genus *Tegula*) from different thermal habitats: Implications for limits of thermotolerance and biogeography. *Journal of Experimental Biology,* **202,** 2925-2936.

Tomanek L, Somero GN (2000) Time course and magnitude of synthesis of heat-shock proteins in congeneric marine snails (genus *Tegula*) from different tidal heights. *Physiological and Biochemical Zoology,* **73,** 249-256.

Treinin M, Shliar J, Jiang HQ *et al.* (2003) HIF-1 is required for heat acclimation in the nematode *Caenorhabditis elegans*. *Physiological Genomics,* **14,** 17-24.

True HL, Lindquist SL (2000) A yeast prion provides a mechanism for genetic variation and phenotypic diversity. *Nature,* **407,** 477.

Tsai J, Sultana R, Lee Y *et al.* (2001) RESOURCERER: a database for annotating and linking microarray resources within and across species. *Genome Biology,* **2,** software0002.0001-0002.0004.

Tsoi SCM, Cale JM, Bird IM *et al.* (2003) Use of human cDNA microarrays for identification of differentially expressed genes in Atlantic salmon liver during *Aeromonas salmonicida* infection. *Marine Biotechnology,* **5,** 545-554.

Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *Plos Biology,* **3,** 1572-1578.

Vamvakopoulos NC, Chrousos GP (1994) Hormonal regulation of human corticotropin-releasing hormone gene expression: implications for the stress response and immune/inflammatory reaction. *Endocrine Reviews,* **15,** 409-420.

van Ginneken VJT, Onderwater M, Olivar OL *et al.* (2001) Metabolic depression and investigation of glucose/ethanol conversion in the European eel (*Anguilla anguilla* Linnaeus 1758) during anaerobiosis. *Thermochimica Acta,* **373,** 23-30.

Vandesompele J, De Preter K, Pattyn F *et al.* (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology,* **3,** research0034.0031-0034.0011.

Vera MI, Ríos HM, de la Fuente E *et al.* (1997) Seasonal acclimatization of the carp involves differential expression of 5.8S ribosomal RNA in pituitary cells. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology,* **118,** 777-781.

Verheij M, Bose R, Lin XH *et al.* (1996) Requirement for ceramide-initiated SAPK/JNK signalling in stress-induced apoptosis. *Nature,* **380,** 75-79.

Vijayan MM, Ballantyne JS, Leatherland JF (1991) Cortisol-induced changes in some aspects of the intermediary metabolism of *Salvelinus fontinalis*. *General And Comparative Endocrinology,* **82,** 476-486.

Waddington CH (1942) Canalization of development and the inheritance of acquired characters. *Nature,* **150,** 563-565.

Wagner GP, Chiu CH, Hansen TF (1999) Is Hsp90 a regulator of evolvability? *Journal of Experimental Zoology,* **285,** 116-118.

Wang XX, Zhang JZ (2004) Rapid evolution of mammalian X-linked testis-expressed homeobox genes. *Genetics,* **167,** 879-888.

Whitmarsh AJ, Shore P, Sharrocks AD *et al*. (1995) Integration of MAP kinase signal transduction pathways at the serum response element. *Science,* **269,** 403-407.

Wilkinson C, Ed. (2002) Status of Coral Reefs of the World. Townsville, Australia, Australian Institute of Marine Science.

Wilkinson CR (1999) Global and local threats to coral reef functioning and existence: review and predictions. *Marine and Freshwater Research,* **50,** 867-878.

Williams SE, Bolitho EE, Fox S (2003a) Climate change in Australian tropical rainforests: an impending environmental catastrophe. *Proceedings of the Royal Society B: Biological Sciences,* **270,** 1887.

Williams TD, Gensberg K, Minchin SD *et al.* (2003b) A DNA expression array to detect toxic stress response in European flounder (*Platichthys flesus*). *Aquatic Toxicology,* **65,** 141-157.

Wittbrodt J, Shima A, Schartl M (2002) Medaka - A model organism from the Far East. *Nature Reviews Genetics,* **3,** 53-64.

Wu C (1995) Heat shock transcription factors: Structure and regulation. *Annual Review Of Cell And Developmental Biology,* **11,** 441-469.

Wu LY, Thompson DK, Li GS *et al*. (2001) Development and evaluation of functional gene arrays for detection of selected genes in the environment. *Applied And Environmental Microbiology,* **67,** 5780-5790.

Yoshida T, Muller E, Stears R *et al.* (2004) Osmoadaptation-related genes in inner medulla of mouse kidney using microarray. *Biochemical and Biophysical Research Communications,* **322,** 250-257.

Young G (1993) Effects of hypophysectomy on coho salmon interrenal maintenance of steroidogenic pathway and restoration of *in vitro* responsiveness to adrenocorticotropin after handling. *General And Comparative Endocrinology,* **92,** 428-438.

Appendix 1

Degenerate primers used for PCR amplification of candidate and housekeeping genes in *Pomacentrus moluccensis*, GenBank Accession number of *P. moluccensis* clones and results of BLASTX search for *P. moluccensis* clones on 21st July 2005.

Gene	GenBank # of *Danio rerio* clone	Species for protein alignment†	Forward primer (5' → 3')	Reverse Primer (5' → 3')	Product size (bp)	GenBank # of *P. mol* clone	BLASTx match	E-value
rhoC	AI959074	Dr, Do, Pf, Tn	5'GGTGTACGTGCCTACCGTGTTYGARAAYTA-3'	5'GCTTCACCTCAGGGGTCCAYTTYTC-3'	217	DQ243817	*Danio rerio* rhoC (AAH53200)	3e-24
cdk5	AF203736	Dr, Hs, Xl	5'CGAGTGCGACTGGACGAYAYGAYGA-3'	5'CGTGGTGGCAGGGTACATNGGRTA-3'	639	DQ243818	*Danio rerio* cdk5 (NP571794)	9e-112
LOC402870	BM181148	Dr, Hs, Tn	5'GACGATCCCACCCTGTGTTYTAYCARGG-3' 5'CGACCACCAGGACACCGAYGARTTYTG-3'	5'CCTCTCGCTTCTTTCGGTCYTTYTCYTG-3'	496 399	DQ243819	*Homo sapiens* ETEA protein (AAH14001)	2e-73
cebpd	BE017827	Dr, Hs, Pt, Rn, Xl	5'CCGCCCCTGCCATGTAYGAYGAYGA-3'	5'GCACAGCTCCACCATCTTCTGYTGCATNTC-3'	558	DQ243820	*Danio rerio* cebpd (AAH62522)	3e-26
cct6a	AI437239	Dr, Hs, Tn, Xl	5'TGTACATGGTCGAGATCATGGANATGAA-3' 5'GCGACGAGCCAAGCGAMGNAAYATGGA-3'	5'CCGTCTCGGATGGCGTCYTTDATYTG-3' 5'GACATTCCGGCTCGCATDATYTCRTC-3'	617 644	DQ243821	*Danio rerio* cct6a (AAH44393)	3e-166
β-actin‡	AF025305	-	5'TTGAGACCTTCAACACCCC-3'	5'ACTCCTGCTTGCTGATCCAC-3'	648	DQ243822	*Monopterus albus* β-actin (AAQ21403)	2e-80
18S	-	Dr, Fh, Oe, Ol, St, Tn,	5'AGTTGGTGGAGCGATTTGTCTG-3' 5'CGGACGAAAGCGAAAGCATT-3'	5'GTAGCGACGGGCGGTGTG-3'	369 729	DQ243823	*Auxis rochei* 18S (AB193747)	0
HPRT	-	Dr, Ec, Gg, Xl	5'GGACCTGGAGTGCGTGTACATHCCNCAYGG-3'	5'TGCAGTAGGACTTCAGTCGGATRAARTCNAC-3'	229	DQ243824	*Danio rerio* HPRT (AAH71311)	1e-32

† Dr: *Danio rerio*; Do: *Discopyge ommata*; Ec: *Equus caballus*; Fh: *Fundulus heteroclitus*; Gg: *Gallus gallus*; Hs: *Homo sapiens*; Oe: *Oreochromis esculentus*; Ol: *Oryzias latipes*; Rn: *Rattus norvegicus*; St: *Salmo trutta*; Tn: *Tetraodon nigroviridis*; Pf: *Platichthys flesus*; Pt: *Pan troglodytes*; Xl: *Xenopus laevis*;

‡ Primers taken from Forlano *et al.* (2005).

Appendix 2

Primers for quantitative real-time PCR, optimised primer concentrations and expected size of PCR product.

Gene	Forward primer	Reverse Primer	Forward primer (nM)	Reverse primer (nM)	Product size (bp)
rhoC	5'TGGGACACAGCAGGTCAAGA-3'	5'TGTCGATGGAGAAGCACATGA-3'	300	300	91
cdk5	5'TCCTCATCAACCGAAATGGA-3'	5'GCACCAAACAGCACATCTGG-3'	300	50	133
LOC402870	5'GTTTTGCCGCTCCACGTTAT-3'	5'GCCTGGGACACTCTGTAGCC-3'	300	300	111
cebpd	5'GCCATCGACTTCAGCCAGTA-3'	5'AAATCCGCCTTCTCCTGCTT-3'	300	300	116
cct6a	5'CCGACACCCAGACATGAAGA-3'	5'CCTCTCCTCGGCACTCTTGT-3'	50	300	121
β-actin	5'AGCGTGGCTACTCCTTCACC-3'	5'CCCATCTCCTGCTCGAAGTC-3'	300	300	100
18S	5'GGCCGTTCTTAGTTGGTGGA-3'	5'GCTCAATCTCGTGTGGCTGA-3'	300	300	145
HPRT	5'GAGAGGCTGGCGAGAGAGAT-3'	5'TCGGTCACTGTTCCTGTTCA-3'	900	300	135

Appendix 3

Sequence alignments of nuclear gene loci sequenced in *Pomacentrus moluccensis* as part of this study and gene sequence data downloaded from GenBank for *Danio rerio*. GenBank accession numbers of gene sequences used in the alignments are given after taxon name.

rhoC

```
#D.rerio_AI959074       --C ATT GCT GAC ATT GAA GTC GAC AGC AAA CAG GTG GAG  [ 39]
#P.moluccensis_DQ243817 --T G.G ... ... ... ... ..T ..T ... ... ... ..T ...  [ 39]

#D.rerio_AI959074       CTG GCA TTG TGG GAC ACA GCA GGA CAG GAG GAC TAT GAC  [ 78]
#P.moluccensis_DQ243817 T.A ..T C.C ... ... ... ... ..T ..A ..A ... ..C ...  [ 78]

#D.rerio_AI959074       CGT CTC AGA CCT CTG TCT TAC CCA GAC ACA GAT GTC ATC  [117]
#P.moluccensis_DQ243817 A.A ..G C.C ..G ..C ... ..T ... ..T ..T ... ... ..T  [117]

#D.rerio_AI959074       CTC ATG TGC TTC TCC ATA GAC AGT CCC GAC AGT TTA GAG  [156]
#P.moluccensis_DQ243817 ... ... ... ... ... ..C ... ..C ..T ... ... C.T ...  [156]

#D.rerio_AI959074       AAT ATC CCA  [165]
#P.moluccensis_DQ243817 ..C ... ..T  [165]
```

cdk5

```
#D.rerio_AF203736       GGC GTA CCA AGT TCA GCT TTA CGA GAA ATT TGC CTC CTG  [ 39]
#P.moluccensis_DQ243818 ..A ... ... ... ... ... C.G A.G ... ..C ..T ..T ...  [ 39]

#D.rerio_AF203736       AAA GAA CTG AAG CAT AAA AAC ATT GTA AGG TTA CAT GAC  [ 78]
#P.moluccensis_DQ243818 ..G ... ..A ... ... ... ... ..A ..C ..A C.. ... ..T  [ 78]

#D.rerio_AF203736       GTC CTT CAC AGT GAT AAG AAA CTC ACA TTA GTG TTT GAA  [117]
#P.moluccensis_DQ243818 ..T T.G ... ... ..C ... ..G T.A ..C ..G ..T ... ...  [117]

#D.rerio_AF203736       TAC TGT GAT CAG GAT TTG AAG AAA TAC TTT GAC AGC TGT  [156]
#P.moluccensis_DQ243818 ... ... ... ... ... ... ... ..G ..T ... ... ... ..C  [156]

#D.rerio_AF203736       AAT GGT GAC TTG GAT CCA GAG ATT GTC AAG TCT TTT ATG  [195]
#P.moluccensis_DQ243818 ... ..G ..T C.A ... ..T ..A .C. ..G ... ..G ..C ...  [195]

#D.rerio_AF203736       TAC CAG TTG TTG AAG GGT CTT GCC TTC TGC CAC AGC AGA  [234]
#P.moluccensis_DQ243818 ... ... C.. ... ..A ..C ... ..T ... ..T ... ..T C..  [234]
```

```
#D.rerio_AF203736       AAT GTT CTC CAC CGT GAC CTT AAA CCA CAA AAC CTG CTC  [273]
#P.moluccensis_DQ243818 ..C ... ..T ..T A.A ..T ..G ..G ..G ..G ..T ..C ...  [273]

#D.rerio_AF203736       ATC AAC AGA AAC GGG GAA CTG AAG TTG GCT GAT TTC GGC  [312]
#P.moluccensis_DQ243818 ... ... C.. ..T ..A ..G T.A ... C.. ... ..C ..T ..G  [312]

#D.rerio_AF203736       TTG GCC AGA GCA TTT GGG ATC CCT GTA CGC TGT TAT TCT  [351]
#P.moluccensis_DQ243818 ... ..T C.. ..T ... ..C ..T ... ..G A.G ... ..C ..A  [351]

#D.rerio_AF203736       GCT GAG GTG GTG ACG TTG TGG TAC AGG CCT CCG GAT GTA  [390]
#P.moluccensis_DQ243818 ..A ... ... ... ..A ... ... ... C.. ... ..A ... ..G  [390]

#D.rerio_AF203736       CTG TTT GGT GCT AAA CTC TAC TCT ACC TCT ATC GAC ATG  [429]
#P.moluccensis_DQ243818 ... ... ... ... ... ..T ..T ... ... ... ... ... ...  [429]

#D.rerio_AF203736       TGG TCT GCA GGC TGT ATA TTT GCA GAA CTG GCA AAT GCA  [468]
#P.moluccensis_DQ243818 ... ... ... ... ..C ... ... ... ..G ... ..T ... ..G  [468]

#D.rerio_AF203736       GGA CGG CCT TTA TTT CCT GGT AAC GAT GTG GAC GAT CAG  [507]
#P.moluccensis_DQ243818 ..G ... ... ... ..C ..C ..C ... ... ... ..T ..C ...  [507]

#D.rerio_AF203736       CTA AAG AGG ATC TTT CGA TTG CTG GGA ACG CCA ACA GAA  [546]
#P.moluccensis_DQ243818 T.G ..A ..A ... ..C A.. ... T.. ..T ..A ... ..T ...  [546]

#D.rerio_AF203736       GAA CAG TGG CAG ACT ATG AAT AAA CTT CCA GAT TAT AAG  [585]
#P.moluccensis_DQ243818 ... ... ... ... ..A ... .CA ... ..G ... ... ..C ...  [585]

#D.rerio_AF203736       CCG  [588]
#P.moluccensis_DQ243818 ..A  [588]
```

Loc402870

```
#D.rerio_BC056714       ACA TAT AGC CAG GCT CTG AAT GAT GCC AAG CGA GAA CTT  [ 39]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [ 39]
#P.moluccensis_DQ243819 ..G ..C ... ... ..A ... ... ... ... ..A ... ..G ..C  [ 39]

#D.rerio_BC056714       CGC TAT TTG TTA GTT TAT CTT CAT GGA GAA GAT CAC CAG  [ 78]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [ 78]
#P.moluccensis_DQ243819 ..T ..C C.A ... ..G ..C ... ..C ..G ..C ..C ... ...  [ 78]
```

```
#D.rerio_BC056714       GAC ACT GAT GAG TTC TGC CGA TCC ACG TTA TGT TCA GAA  [117]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [117]
#P.moluccensis_DQ243819 ... ..C ... ... ..T ... ..C ... ... ... ... A.G ...  [117]

#D.rerio_BC056714       GAG GCG CTC ACC TTC ATC AAC ACG AGG ATG TTG TTC TGG  [156]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [156]
#P.moluccensis_DQ243819 ... .TC A.A ..T ... C.. ... G.. C.A ... C.C ..T ...  [156]

#D.rerio_BC056714       GCA TGT TCC ACC AGC AAA CCA GAG GGT TAC AGA GTC TCT  [195]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [195]
#P.moluccensis_DQ243819 ... ... ..A ..G ... ... ..T ... ..C ... ... ..G ..C  [195]

#D.rerio_BC056714       CAG GCC CTG CGT GAA AAC ACC TAT CCA TTC CTG GCC ATG  [234]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [234]
#P.moluccensis_DQ243819 ... ..A T.. ..G ..G ... ... ..C ... ... ... ... ...  [234]

#D.rerio_BC056714       ATC ATG CTG AAG GAC CGG AAG ATG ACG GTG GTT GGG AGA  [273]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [273]
#P.moluccensis_DQ243819 ..A ... ... ... ... ..C ... ... ..C ... ..A ..C ..G  [273]

#D.rerio_BC056714       CTG GAG GGT TTA ATT CAG CCG GAG GAT CTG ATC AAC CAG  [312]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [312]
#P.moluccensis_DQ243819 ... ... ... C.C ..C ... ..A ..A ..C ..C ... ..T ...  [312]

#D.rerio_BC056714       CTG ACC TTC ATC ATG GAG GCA AAC CAG ACA TAT CTC ATG  [351]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- --- --- --- --- ---  [351]
#P.moluccensis_DQ243819 ..C ..T ... ... ... ..T ..C ..T ..A ... ... ..G ...  [351]

#D.rerio_BC056714       TCA GAA CGA CTA GAG AGG GAA GAG AGG AAT CAG ACG CAG  [390]
#D.rerio_BM181148       --- --- --- --- --- --- --- --- ... ... ... ... ...  [390]
#P.moluccensis_DQ243819 ... ..G ..C ..T ..A C.C ..G ... ... ..C ... ..C ..A  [390]

#D.rerio_BC056714       GTA TTG CGA CAG CAG CAG GAC GAG GCT TAC GAG GCA TCA  [429]
#D.rerio_BM181148       ... ... ... ... ... ... ... ... ... ... ... ... ...  [429]
#P.moluccensis_DQ243819 ..G C.A A.G ... ... ... ... ... ..C ..T CTC ... ..C  [429]

#D.rerio_BC056714       CTG CGA GCA GAT  [441]
#D.rerio_BM181148       ... ... ... ...  [441]
#P.moluccensis_DQ243819 ..C ..T ..C ..C  [441]
```

cebpd

```
#D.rerio_NM131887       AGC GCC ATC GAC TTC AGC GCA TAC ATT GAG TCC ATG TCA  [ 39]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- --- --- --- ---  [ 39]
#P.moluccensis_DQ243820 ... ... ... ... ... ... CAG ... ..C ... ... ... A..  [ 39]

#D.rerio_NM131887       ACG GTC CCG --- TTG GAG ATC TGC AAC GAC GAG CTC TTC  [ 78]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- --- --- --- ---  [ 78]
#P.moluccensis_DQ243820 G.C ..G ..C AAC C.. ... C.G ... ... ... ... ... ...  [ 78]

#D.rerio_NM131887       GCA GAC CTG TTC AAC AAT ACT GTA AAG CAA GAA AAG CCT  [117]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- --- --- --- ---  [117]
#P.moluccensis_DQ243820 CTC ... ... ... ... --- ... ..G ... ..G ..G ... G.G  [117]

#D.rerio_NM131887       GAC TTT TAC ATG TCC AAC ACG TTT GCG CAC AAG AGC GCG  [156]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- --- --- --- ---  [156]
#P.moluccensis_DQ243820 ..T ..C ... .AC CTG C.G .GC .CC .T. .TG CCC G.. .GC  [156]

#D.rerio_NM131887       GAG AGG CAC TTG GAG --- GGT TTC GGA --- --- --- ---  [195]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- --- --- --- ---  [195]
#P.moluccensis_DQ243820 AT. CA. ..G CA. CT. TCC .CC .A. ACC GCA GAC AGG AGG  [195]

#D.rerio_NM131887       --- --- --- --- --- --- AAG GGC TCG TTT TGC GCA CCG  [234]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- --- --- --- ---  [234]
#P.moluccensis_DQ243820 GCG GAC GGC GGG CTG AAC ... ..G C.. ..C AA. ..T ..C  [234]

#D.rerio_NM131887       ATT AAG AAG GAG GCG GAC TGG AGC GAC AGC GAG CAC TCC  [273]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- --- --- --- ---  [273]
#P.moluccensis_DQ243820 ..C ... C.. ... T.C ... ... ..T ... ... ..C GTG ..T  [273]

#D.rerio_NM131887       TCG TCT TTA CCG TCG CAG ATC GAG GCG TGC GCG CAG ACC  [312]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- --- --- --- ---  [312]
#P.moluccensis_DQ243820 ..A ... C.G ... ..C ... ... ... A.C ... ..C ... ...  [312]

#D.rerio_NM131887       TCC GTC AAC TTC ATG CAC ACG GGA CAG CCG ACG CCC CCC  [351]
#D.rerio_BE017827       --- --- --- --- --- --- --- --- --- G.N ... ... ...  [351]
#P.moluccensis_DQ243820 ... ... .G. --- C.C .C. ..A ..G ... ..C ..C ..T ...  [351]

#D.rerio_NM131887       ACC ACA CCT GAG CCC --- GAG CCC GTG GCA CAC AGG AGG  [390]
#D.rerio_BE017827       ... ... ... ... ... --- ... ... ... ... ... ..N ...  [390]
#P.moluccensis_DQ243820 ... ..C ..G ... ... GTC TCC G.G ... TGC TCG GCC .A.  [390]
```

```
#D.rerio_NM131887       CCA --- --- --- --- --- GGC AAG GAG AAG GGC AAA AAG  [429]
#D.rerio_BE017827       ... --- --- --- --- --- ... ... ... ... ... ... ...  [429]
#P.moluccensis_DQ243820 T.C TCC CCG AGG AAG ATG ... .G. ... ... ... ..G ...  [429]

#D.rerio_NM131887       AAC GTG GAC AGG CAC AGT CCG GAG TAC CGG CAG CGG CGC  [468]
#D.rerio_BE017827       ... ... AC. ... ... ... ... ... ... ... ..N ... ...  [468]
#P.moluccensis_DQ243820 GCG ... ... C.. .TG ..C GT. ... ... ..A ... A.. ..A  [468]

#D.rerio_NM131887       GAG AGG AAC AAC ATC GCC GTG CGT AAA AGC AGA GAC AAG  [507]
#D.rerio_BE017827       ... ... ... ... ... C.. ... ... ... ... ... ... ...  [507]
#P.moluccensis_DQ243820 ... ... ..T ... ..T ..A .W. A.G ... ... ..G ... ..A  [507]

#D.rerio_NM131887       GCG AAG CAG CGC AAC TTG  [525]
#D.rerio_BE017827       ... ... ... ... ... ...  [525]
#P.moluccensis_DQ243820 ..C ... AG. ... ... ...  [525]
```

cct6a

```
#D.rerio_NM201290       CAC AAA ACC GAC AGT GAT ACT CAA CTG ATC AGA GGG TTG  [ 39]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [ 39]
#P.moluccensis_DQ243821 ... ..G ... ... T.C ..C ..A ... ... ... ... ..T ...  [ 39]

#D.rerio_NM201290       GTG TTG GAC CAC GGA GCC AGA CAT CCT GAC ATG AAG AAG  [ 78]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [ 78]
#P.moluccensis_DQ243821 ... C.. ... ... ..C ... C.. ..C ..A ... ... ... ...  [ 78]

#D.rerio_NM201290       AGA ATA GAA GAT GCT TTC ATT CTC ACG TGC AAC GTT TCC  [117]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [117]
#P.moluccensis_DQ243821 ..G G.G ..G ..C ..C .A. G.G ..G ... ... ... ..C ..T  [117]

#D.rerio_NM201290       TTG GAA TAC GAG AAA ACC GAA GTG AAC TCC GGG TTT TTC  [156]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [156]
#P.moluccensis_DQ243821 ... ..G ... ..A ..G ..G ..G ..C ... ..T ..C ... ...  [156]

#D.rerio_NM201290       TAC AAG AGC GCA GAC GAG AGA GAC AAG CTG GTG AAA GCT  [195]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [195]
#P.moluccensis_DQ243821 ... ... ..T ..C ..G ... ..G ..G ... ..T ... GCT ..G  [195]
```

```
#D.rerio_NM201290       GAG AGG AAG TTC ATT GAA GAT CGT GTG ATG AAA ATA ATC  [234]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [234]
#P.moluccensis_DQ243821 ... ... ... ... ..C ..G ..C ... ... CAA ..G ..C ...  [234]

#D.rerio_NM201290       GAC CTG AAG AAT AAA GTG TGC GCT GAT AAT AAA AAG GGC  [273]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [273]
#P.moluccensis_DQ243821 .C. ... ... ..C ... ..T ..T C.C A.. GGG G.G ... ..A  [273]

#D.rerio_NM201290       TTT GTG GTC ATC AAT CAG AAG GGT ATT GAT CCA TTT TCT  [312]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [312]
#P.moluccensis_DQ243821 ... ..C ... ... ..C ... ... ... ... ..C ... ... ..C  [312]

#D.rerio_NM201290       CTG GAT GCG CTG GCC AAA GAG GGC ATT GTG GCT CTG CGC  [351]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [351]
#P.moluccensis_DQ243821 ... ..C ..C ..C ... ..G ..A ... ... ..C ... ... ...  [351]

#D.rerio_NM201290       CGT GCA AAG AGA CGG AAT ATG GAA AGG CTG ACT CTG GCA  [390]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [390]
#P.moluccensis_DQ243821 A.G ..G ... ..G A.. ..C ... ..G ... ... ... ..C ..T  [390]

#D.rerio_NM201290       TGT GGT GGT GTA GCA ATG AAC TCT GTG GAG GAT CTC ACG  [429]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [429]
#P.moluccensis_DQ243821 ..C ... ..C A.C ..C ... ..T ..A ..T ..T ..C ... ...  [429]

#D.rerio_NM201290       CCA GAG TGT TTG GGA CAC GCC GGT CTT GTG TAC GAG TAC  [468]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [468]
#P.moluccensis_DQ243821 .TC ... ..C ..A ... ..G ..T ..G ..G ..T ..T ..A C..  [468]

#D.rerio_NM201290       ACA CTG GGT GAG GAG AAA TTC ACG TTT ATT GAG AAC TGT  [507]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [507]
#P.moluccensis_DQ243821 ... ... ..A ..A ... ... .A. ..A ..C ... ... ..G ...  [507]

#D.rerio_NM201290       AGC AAC CCT CGT TCT GTG ACC CTG CTG GTG AAA GGC CCA  [546]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [546]
#P.moluccensis_DQ243821 G.A ... ... ..Y ... ..T ... ... ... ... ..G ..A ..C  [546]

#D.rerio_NM201290       AAC AAG CAC ACC CTG ACA CAG ATC AAA GAC GCA GTG AGA  [585]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [585]
#P.moluccensis_DQ243821 ... ..A ... ... ..C ... ... ... ... ... ..C A.C C..  [585]
```

```
#D.rerio_NM201290       GAC GGA CTC AGA GCT GTC AAA AAC GCC ATT GAA GAT GGC  [624]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [624]
#P.moluccensis_DQ243821 ... ..T ..G C.G ..A ... ..G ... ... ... ..G ... ..A  [624]

#D.rerio_NM201290       TCT GTT GTA GCT GGG GCT GGT GCA TTT GAG GTA GCC GTG  [663]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- --- --- ---  [663]
#P.moluccensis_DQ243821 AGC ..C ..G T.. ..T ..A ... ... ..C ... ..C ..T ...  [663]

#D.rerio_NM201290       GCT GAT GCA TTG GTC AAA CAC AAG CCC AAA GTA AAG GGA  [702]
#D.rerio_AI437239       --- --- --- --- --- --- --- --- --- --- -.C ..C TA.  [702]
#P.moluccensis_DQ243821 ..A ..C ..T C.. ..A ... ... ... ... ..C ..G ..A ..C  [702]

#D.rerio_NM201290       CGA GCC CAG CTG GGT GTG CAG GCG TTT GCT GAT GCT CTT  [741]
#D.rerio_AI437239       GCC TAG GTT TAC ATA TAT T.C CTT ACA .A. CCC CTA TAA  [741]
#P.moluccensis_DQ243821 A.. ... ... ... ..A ..C ... ... ... ..A ... ... ..C  [741]

#D.rerio_NM201290       CTT GTC ATC CCT AAA GTT CTG GCT CAG AAT TCA GGC TAC  [780]
#D.rerio_AI437239       AGA AAA TAT ATG T.. ACC TA. ... T.. TTG AGT AAG ATA  [780]
#P.moluccensis_DQ243821 ..C ... ... ..C ..G ... T.. ..C ... ..C ..T ..T ..T  [780]

#D.rerio_NM201290       GAC CCA CAA GAG ACT CTT GTG AAG CTG CAG AGT GAA TTC  [819]
#D.rerio_AI437239       TCA GA. T.T .GA .A. GGC CAT .CT GC. AGT CTC AG. CA.  [819]
#P.moluccensis_DQ243821 ..T ... ..G ... ..C ..G C.. ... ... ... .CA ..G .A.  [819]

#D.rerio_NM201290       AAA GAG GCT GGA CAG TTG GTG GGA GTT GAC CTA AGC ACA  [858]
#D.rerio_AI437239       G.. TGT TTC AT. T.T ..T C.T TAT AGG .GA TCT GTA .GG  [858]
#P.moluccensis_DQ243821 ... ... Y.. ..T ... C.A ..C ... ..C ... ..C ... ...  [858]

#D.rerio_NM201290       GGC GAG CCT ATG GTT GCT GGA GAG GCT GGT GTA TGG GAC  [897]
#D.rerio_AI437239       TAA T.A .T. TA. T.. TTG TTG T.T AAA CTC TCT .TC AGT  [897]
#P.moluccensis_DQ243821 ..G ..A ..A ... ..G ..A ... ..A ... ... ..T ... ...  [897]

#D.rerio_NM201290       AAC TAC AGT GTG AAA AAG CAG CTG CTT CAC TCC TGC ACT  [936]
#D.rerio_AI437239       TT. ..A TT. TCT .CT C.T --- --- --- TT. CA. A.. ...  [936]
#P.moluccensis_DQ243821 ..T ..T ..C ..C ..G ... ... ..T ..C ..T ..A ... ..G  [936]

#D.rerio_NM201290       GTA ATC GCC AGT AAT ATC CTG CTG GTG  [963]
#D.rerio_AI437239       ... ... ... ... ... ... ... ... ...  [963]
#P.moluccensis_DQ243821 ..G ... ... ..C ..C ... ... T.. ...  [963]
```

β-actin

```
#D.rerio_AF025305       GTG CTG TCC CTG TAT GCC TCT GGT CGT ACC ACT GGT ATC  [ 39]
#P.moluccensis_DQ243822 ... ... ... ... ... ... ... ... ... ... ... ... ...  [ 39]

#D.rerio_AF025305       GTG ATG GAC TCT GGT GAT GGT GTC ACC CAC ACT GTG CCC  [ 78]
#P.moluccensis_DQ243822 ..C ... ... ..C ... ... ... ..G ... ... ..A ... ...  [ 78]

#D.rerio_AF025305       ATC TAC GAG GGT TAC GCC CTG CCC CAT GCC ATC CTC CGT  [117]
#P.moluccensis_DQ243822 ... ... ... ..C ... ..T ... ... ..C ... ... ..G ...  [117]

#D.rerio_AF025305       CTG GAC TTG GCT GGC CGT GAC CTG ACT GAC TAC CTC ATG  [156]
#P.moluccensis_DQ243822 ... ..T C.. ..C ... ..C ... ..C ..A ... ... ... ...  [156]

#D.rerio_AF025305       AAG ATC CTG ACC GAG AGA GGC TAC AGC TTC ACC ACC ACA  [195]
#P.moluccensis_DQ243822 ... ... ... ..A ... C.T ... ... TC. ... ... ... ...  [195]

#D.rerio_AF025305       GCT GAG AGG GAA ATT GTC CGT GAC ATC AAG GAG AAG CTC  [234]
#P.moluccensis_DQ243822 ..C ... ... ... ..C ..G ... ... ... ... ... ... ...  [234]

#D.rerio_AF025305       TGC TAT GTG GCC CTT GAC TTT GAG CAG GAG ATG GGC ACC  [273]
#P.moluccensis_DQ243822 ... ... ..T .G. ..G ... ..C ... ... ... ... ... ...  [273]

#D.rerio_AF025305       GCT GCT TCC TCC TCC TCC CTG GAG AAG AGC TAC GAG CTG  [312]
#P.moluccensis_DQ243822 ... ..C ... ... ... ... ... ... ... ... ..T ... ...  [312]

#D.rerio_AF025305       CCT GAC GGA CAG GTC ATC ACC ATT GGC AAT GAG AGG TTC  [351]
#P.moluccensis_DQ243822 ..C ... ... ... ... ... ... ..C ... ... ... ... ...  [351]

#D.rerio_AF025305       AGG TGC CCA GAG GCC CTG TTC CAG CCA TCC TTC TTG GGT  [390]
#P.moluccensis_DQ243822 C.T ... ... ... ... ..C ... ... ..T ... ... C.T ...  [390]

#D.rerio_AF025305       ATG GAA TCT TGC GGT ATC CAC GAG ACC ACC TTC AAC TCC  [429]
#P.moluccensis_DQ243822 ... ..G ..C ... ..A ... ... ... ..T ... ... ... AG.  [429]

#D.rerio_AF025305       ATC ATG AAG TGT GAC GTC GAC ATC CGT AAG GAC CTG TAT  [468]
#P.moluccensis_DQ243822 ... ... ... ... ..T ... ... ... ... ... ... ... ..C  [468]

#D.rerio_AF025305       GCC AAC ACT GTA TTG TCT GGT GGT ACC ACC ATG TAC CCT  [507]
#P.moluccensis_DQ243822 ... ... ..C ..G C.. ... ..A ... ... ... ... ... ..C  [507]
```

```
#D.rerio_AF025305       GGC ATT GCT GAC AGG ATG CAG AAG GAG ATC ACA TCC CTG  [546]
#P.moluccensis_DQ243822 ... ... ... ... ..T ... ... ... ... ... ... G-. ...  [546]

#D.rerio_AF025305       GCC CCT AGC ACA ATG AAG ATC AAG ATC ATT GCC CCA CCG  [585]
#P.moluccensis_DQ243822 ... ..A TC. ..C ... ... ... ..A ... ..C ... ... ..A  [585]

#D.rerio_AF025305       GAG CGT AAA TAC TCT GTC TGG ATC GGA GGC TCC  [618]
#P.moluccensis_DQ243822 ... ... ... ... ... ... ... ..T C.G A.. ...  [618]
```

HPRT

```
#D.rerio_BC046003       --G CTC ATC ATG GAC CGA ACT GAA CGT CTG GCC AGA GAT  [ 39]
#P.moluccensis_DQ243824 --A ... ... T.. ... A.G ..G ..G A.G ... ..G ... ..G  [ 39]

#D.rerio_BC046003       ATC ATG AAG GAC ATG GGT GGA CAC CAT ATA GTG GCT CTA  [ 78]
#P.moluccensis_DQ243824 ... ... ... ..A ... ... ... ..T ..C ..C ..C ..C ..C  [ 78]

#D.rerio_BC046003       TGT GTG CTC AAA GGA GGC TAC AAG TTT TTT GCT GAC TTA  [117]
#P.moluccensis_DQ243824 ... ... ..Y ... ..G ..Y ... ... ... ..C ..A ... C.G  [117]

#D.rerio_BC046003       CTA GAT TAC ATC AAA GCC CTT AAT CGC AAC AGT GAT CGC  [156]
#P.moluccensis_DQ243824 ..G ..C ... ... ..G ... ..G ..C A.G ... ... ..C ..A  [156]

#D.rerio_BC046003       TCC ATT CCC ATG ACA  [171]
#P.moluccensis_DQ243824 ... ..C ... ... ..G  [171]
```

18S

```
#D.rerio_BX296557       CGG AGG TTC GAA GAC GAT CAG ATA CCG TCG TAG TTC CGA  [ 39]
#P.moluccensis_DQ243823 ... ... ... ... ... ... ... ... ... ... ... ... ...  [ 39]

#D.rerio_BX296557       CCG TAA ACG ATG CCG ACC CGC GAT CCG GCG GCG TTA TTC  [ 78]
#P.moluccensis_DQ243823 ..A ... ... ... ..A ..T A.. ... ... ... ... ... ...  [ 78]

#D.rerio_BX296557       CCA TGA CCC GCC GGG CAG CGT GCG GGA AAC CAC GAG TCT  [117]
#P.moluccensis_DQ243823 ... ... ... ... ... ... ... C.. ... ... ..A -.. ...  [117]

#D.rerio_BX296557       TTG GGT TCC GGG GGG AGT ATG GTT GCA AAG CTG AAA CTT  [156]
#P.moluccensis_DQ243823 ... ... ... ... ... ... ... ... ... ... ... ... ...  [156]
```

```
#D.rerio_BX296557       AAA GGA ATT GAC GGA AGG GCA CCA CCA GGA GTG GAG CCT  [195]
#P.moluccensis_DQ243823 ... ... ... ... ... ... ... ... ... ... ... ... ...  [195]

#D.rerio_BX296557       GCG GCT TAA TTT GAC TCA ACA CGG GAA ACC TCA CCC GGC  [234]
#P.moluccensis_DQ243823 ... ... ... ... ... Y.. ... ... ... ... ... ... ...  [234]

#D.rerio_BX296557       CCG GAC ACG GAA AGG ATT GAC AGA TTG ATA GCT CTT TCT  [273]
#P.moluccensis_DQ243823 ... ... ... ... ... ... ... ... ... ... ... ... ...  [273]

#D.rerio_BX296557       CGA TTC TGT GGG TGG TGG TGC ATG GCC GTT CTT AGT TGG  [312]
#P.moluccensis_DQ243823 ... ... ... ... ... ... ... ... ... ... ... ... ...  [312]

#D.rerio_BX296557       TGG AGC GAT TTG TCT GGT TCA TTC CGA TAA CGA ACG AGA  [351]
#P.moluccensis_DQ243823 ... ... ... ... ... ... .A. ... ... ... ... ... ...  [351]

#D.rerio_BX296557       CTC CGG CAT GCT AAA TAG TTA CGC GGC CCC GCG CGG TCG  [390]
#P.moluccensis_DQ243823 ... ... ... ... ..C ... ... ... ... ... .T. ... ...  [390]

#D.rerio_BX296557       GCG TCC CAA CTT CTT AGA GGG ACA AGT GGC GTT CAG CCA  [429]
#P.moluccensis_DQ243823 ... ... -.. ... ... ... ... ... ... ... ... ... ...  [429]

#D.rerio_BX296557       CGC GAG ATG GAG CAA TAA CAG GTC TGT GAT GCC CTT AGA  [468]
#P.moluccensis_DQ243823 .A. ... ..T ... ... ... ... ... ... ... ... ... ...  [468]

#D.rerio_BX296557       TGT CCG GGG CTG CAC GCG CGC CAC AAT GGG CGG ATC AAC  [507]
#P.moluccensis_DQ243823 ... ... ... ... ... ... ... ... .C. .A. T.. ... .G.  [507]

#D.rerio_BX296557       GTG TGC CTA CCC TGC GCC GAG AGG CGC GGG TAA CCC GTT  [546]
#P.moluccensis_DQ243823 ... ..T ... ... .T. ... ... ... ..T ... ... ... .C.  [546]

#D.rerio_BX296557       GAA CCC CGC TCG TGA TTG GGA CTG GGG CTT GAA ACT GTT  [585]
#P.moluccensis_DQ243823 ... ... .A. ... ... .A. ... T.. ... A.. .C. .T. A..  [585]

#D.rerio_BX296557       TCC CAT CAA CGA GGA ATT CCC AGT AAG CGC AGG  [618]
#P.moluccensis_DQ243823 ... ... G.. ... ... ... ... ... ... ... G..  [618]
```

Appendix 4 - 6

Please find Appendix 4 – 6 on the CD attached.