ResearchOnline@JCU

JAMES COOK UNIVERSITY AUSTRALIA

This file is part of the following work:

Esparza Salas, Rodrigo (2008) *Molecular ecology of the endangered Gouldian Finch Erythrura gouldiae.* PhD Thesis, James Cook University.

Access to this file is available from:

https://doi.org/10.25903/c00w%2Dwh73

Copyright © 2008 Rodrigo Esparza Salas

The author has certified to JCU that they have made a reasonable effort to gain permission and acknowledge the owners of any third party copyright material included in this document. If you believe that this is not the case, please email researchonline@jcu.edu.au

Molecular ecology of the endangered Gouldian Finch *Erythrura gouldiae*

Thesis submitted by:

Rodrigo Esparza-Salas BSc

For the degree of Doctor of Philosophy
In Conservation Biology
Within the School of Marine and Tropical Biology
James Cook University
Townsville, Queensland, Australia
July 2008

A mis padres:

Graciela y Víctor

STATEMENT OF ACCESS

I, the undersigned author of this work, 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I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and;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 do not wish to place any further restriction on access to this work

____________________________________ ______________

Signature Date

STATEMENT ON 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 Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

……………………………………………..   ………………………
(Signature)   (Date)

STATEMENT OF CONTRIBUTION OF OTHERS

All data chapters of this thesis include collaborative work with my supervisors Prof. Ross H. Crozier and Prof. Chris N. Johnson.

Collection of blood samples from the Yinberrie Hills locality, used in chapters 2 to 4, was done in collaboration with Milton Lewis.

Collection of blood samples from Mornington Wildlife Sanctuary, used in chapters 2 to 4, was done in collaboration with Sarah Legge and Steve Murphy.

Blood samples from the captive population in Mareeba, QLD, used in chapters 2 and 3 were provided by the Mareeba Wetlands Foundation.

Tissue samples from museum specimens used in Chapter 2 were provided by Museum Victoria (Melbourne, Australia), the Natural History Museum (Tring, UK), the Western Australian Museum (Perth, Australia) and Queensland Museum (Brisbane, Australia).

Financial support during my PhD candidature was granted through a scholarship by the Mexican Council of Science and Technology (CONACYT- México, reference number 150502).

Financial support to cover costs associated with my PhD research was provided by the School of Tropical Biology (currently School of Marine and Tropical Biology) at James Cook University (JCU), Birds Australia through a Stuart Leslie Bird Research Award (SLBRA), and the Australian Bird Environment Foundation (ABEF). My supervisors assisted with funds obtained from Australian Research Council (ARC), the Australian Academy of Science and the School of Tropical Biology at JCU.

Logistic support during sample collection at the Yinberrie Hills location was provided by Parks and Wildlife Commission of the Northern Territory (PWCNT).

Logistic support during sample collection at Mornington Wildlife Sanctuary was provided by Australian Wildlife Conservancy (AWC).

Analyses for the reconstruction of phylogenies used in Chapter 4 were carried out using the High Performance Computing Facility at JCU (JCU-HPC).

DECLARATION OF ETHICS

The research presented and reported in this thesis was conducted within the guidelines for research ethics outlined in the *National Statement on Ethics Conduct in Research Involving Human (*1999*),* the *Joint NHMRC/AVCC Statement and Guidelines on Research Practice* (1997)*,* the *James Cook University Policy on Experimentation Ethics*. *Standard Practices and Guidelines* (2001), and the *James Cook University Statement and Guidelines on Research Practice* (2001). The proposed research methodology received clearance from the James Cook University Experimentation Ethics Review Committee (approval numbers A810 and A969).

....

Rodrigo Esparza-Salas (Date)

ACKNOWLEDGEMENTS

I wish to express my gratitude to the following people and organisations that provided assistance during different stages of the development of this thesis and the research associated with it:

- My supervisors Ross H. Crozier and Chris N. Johnson for their continuous guidance, advice and encouragement. Working with such good scientists has been an inspiring experience.
- CONACYT-México for financial support to cover tuition and living expenses during my PhD candidature.
- The School of Tropical Biology at James Cook University, Birds Australia and Australia Bird Environment Foundation for providing financial assistance to cover costs associated with the research presented in this thesis.
- Mareeba Wetlands foundation for providing samples from the captive population used in chapters 2 and 3.
- The natural History Museum (UK), Museum Victoria, the Western Australian Museum and Queensland Museum for providing tissue material from Gouldian Finch specimens in their collections.
- Parks and Wildlife Commission of the Northern Territory, especially the staff at Katherine Gorge National Park for logistic help during sample collection at the Yinberrie Hills location.
- Australian Wildlife Conservancy and the staff at Mornington Wildlife Sanctuary for logistic help during sample collection.
- Milton J. Lewis, Sarah Legge and Steve Murphy for assistance during sample collection, for providing additional samples, and for feedback that was essential for my research.
- All the volunteers who assisted during sample collection at the different locations.
- All members of the Crozier Lab during the past 5 years and other fellow PhD students at the School of Marine and Tropical Biology, especially David Abrego, Line Bay, Mark Bulmer, Sara Ceccarelli, Faye Christidis, Michelle Guzik, Michael Henshaw, Lauryne Grant, Karin Kassahn, Ellen Schlüns, Helge Schlüns, Angela Shuetrim and Itzel Zamora-Vilchis for continuous technical input, and for creating an exceptional working environment.

- A special mention to Ching Crozier who has been my mentor in laboratory techniques, and continuously and patiently provided important feedback throughout my PhD.
- The staff at the High Performance Computing Facility at James Cook University (JCU-HPC) for their help in the use of their facility for the reconstruction of phylogenies in Chapter 4.
- Line Bay, Sara Ceccarelli, Bernhard Schmierer, Angela Shuetrim, Itzel Zamora-Vilchis and Daniel Wilson for their input on different statistical methodologies used in this thesis.
- Line Bay, Karin Kassahn, Angela Shuetrim, Bernhard Schmierer and Itzel Zamora-Vilchis for their comments on previous drafts of some of the chapters.

GENERAL ABSTRACT

The Gouldian Finch (*Erythrura gouldiae*) is a species endemic to the Australian tropical savannas that has experienced drastic population declines in recent times. Population declines have been largely attributed to uncontrolled fire regimes, cattle grazing and disease. The decline of *E. gouldiae* populations has raised concerns about the state of conservation of tropical savannah habitats across northern Australia. There are a number of current conservation efforts aimed at understanding the biology of this species, and at using such knowledge to implement effective management for the recovery of its populations. An in depth understanding of the molecular ecology of *E. gouldiae* is important to complement the current knowledge of the species biology and will contribute to the implementation of management practices for its recovery.

I used molecular genetic techniques to resolve aspects of the ecology of *E. gouldiae* that are relevant to the management and recovery of their populations. I contrasted some of the patterns of genetic variation with those of the Long-tailed Finch (*Poephila acuticauda*); another tropical savannah species that is comparatively abundant and has not experienced the kind of population declines that have affected Gouldian finch populations. The research in my thesis has three main components:

1. Phylogeography of *E. gouldiae* and its relation to historical habitat changes in the Australian tropical savannas.

I investigated the geographical distribution of *E. gouldiae* lineages using mitochondrial DNA (mtDNA) control region sequences from extant populations and museum specimens. Results show a lack of lineage isolation among different regions. Patterns of mismatch distribution between mtDNA haplotypes are consistent with range expansions, dating from the Holocene period, approximately 7000 to 3600 years before present. Analyses of samples from different geographical regions suggest that population expansion events took place at different times: earlier at the easternmost localities and more recently at the westernmost parts of the species distribution. This proposed scenario of gradual colonisation of northern Australia can be explained by well documented climatic and socio-cultural changes during the Holocene. Change of

the landscape into the present drier habitats with grassy open vegetation would have favoured the expansion of *E. gouldiae* populations.

2. Genetic variability, population structure and migration in *E. gouldiae* and *P. acuticauda* populations.

I used microsatellite genotypes to investigate genetic variation, population structure and gene flow in *E. gouldiae* and *P. acuticauda* populations. Heterozygosity seems slightly higher in *P. acuticauda* than in *E. gouldiae*. I found no evidence of recent genetic bottlenecks or severe inbreeding in wild populations, although there is significantly lower heterozygosity and evidence for a recent population bottleneck in a captive stock that is part of a reintroduction programme for *E. gouldiae*. The indicators of intra-population inbreeding (F_{IS}) were relatively low in both species and there was no evidence for severe inbreeding. Inter-population differentiation (F_{ST}) is moderate for both species, which is consistent with their dispersal capabilities. Assignment tests show evidence of structuring between *Poephila acuticauda* populations, whereas such structuring is absent between individuals of *E. gouldiae* from the same locations. Migration rates appear to be higher for *E. gouldiae* than for *P. acuticauda.* These differences can partly explain the absence of population structure in *E. gouldiae*. However, mitochondrial DNA analyses have suggested the possibility of incomplete lineage sorting, which might potentially compromise this interpretation.

3. Variability, gene conversion and natural selection at Major Histocompatibility Complex (MHC) genes of *E. gouldiae* and *P. acuticauda*.

I amplified exon 2 of the MHC class IIβ gene in *E. gouldiae* and *P. acuticauda* to test for the presence of gene conversion and to search for evidence of positive selection. As previously found in other passerine bird species, the number of alleles per individual was high (up to 10 in *E. gouldiae* and up to 12 in *P. acuticauda*). The total number of haplotypes and nucleotide diversity are higher in *P. acuticauda*, although this difference is not significant. The observed individual allele diversity may be driven partly by gene conversion, which appears to occur more frequently in *E. gouldiae* than in *P.*

acuticauda. Analyses of MHC codon substitution show evidence of positive selection at 9 codons in each species, most of which are located at or adjacent to the predicted peptide binding amino acids in human MHC. There are differences in the positively selected amino-acid sites between both species It is likely that the observed higher rates of gene conversion and positive selection have been comparably more important in the maintenance of MHC diversity in *E. gouldiae* as a result of lower genetic variability in this species. This suggests *E. gouldiae* may be subject to higher selective pressure from pathogens, consistent with documented evidence on differences in parasite loads between the two species studied.

In summary, the present distribution of *E. gouldiae* populations appears to be relatively recent in origin. This may have resulted in the reduced levels of genetic variability, as is expected under a scenario of population expansion following a genetic bottleneck. Results from this research indicate *E. gouldiae* have a high dispersal capability, which is reflected by the lack of genetic structure between populations and high levels of gene flow. Genetic variability does not appear to be of immediate concern for this species, as there is no evidence of severe inbreeding or recent population bottlenecks in wild populations. However, monitoring of genetic diversity can be critical in the success of future captive breeding and reintroduction programs. Although MHC gene diversity seems to have been maintained (in terms of allele richness) in spite of past and present population reductions in *E. gouldiae*, it is still not known if the present diversity confers resistance to pathogens similar to that for other species whose populations have been more stable, particularly given that a high proportion of the observed MHC variants in this species appear to be of recombinant origin. Future research in this topic will help to determine the importance of gene conversion and selection on the maintenance of MHC diversity and parasite resistance in small and declining populations.

TABLE OF CONTENTS

STATEMENT OF ACCESS i

STATEMENT OF SOURCES ii

STATEMENT OF CONTRIBUTION OF OTHERS iii

DECLARATION OF ETHICS iv

ACKNOWLEDGEMENTS v

GENERAL ABSTRACT vii

TABLE OF CONTENTS x

LIST OF TABLES xv

LIST OF FIGURES xvii

CHAPTER ONE: General Introduction 1

Genetic diversity and population viability 1

Molecular genetics in ecological research 2

Resolution of taxonomic uncertainties 2

Calculation of genetic diversity 2

Phylogeography 3

Resolution of population structure and gene flow 4

Resolution of mating systems 4

Adaptive variation 5

Conservation genetics 5

Study species 8

Population trends and current threats 8

Conservation status and current conservation measures *9*

Research aims **10**

CHAPTER TWO: Historical phylogeography of the endangered Gouldian Finch *Erythrura gouldiae*: evidence of gradual population expansion during the Holocene **12**

Abstract **12**

Introduction **13**

Methods **17**

Sampling *17*
DNA extractions *17*
Amplification of mtDNA control region *18*
Genetic diversity *19*
Analysis of Molecular Variance and Median-joining network *20*
Mismatch distributions *20*
Tests of neutrality *21*
Relative dating of population expansions *21*

Results **22**

Genetic Diversity *22*
Genetic structure *24*
Mismatch distributions and tests of neutrality *26*
Estimated times of population expansion *28*

Discussion **28**

Population structure *29*
Population expansion and gradual colonisation of northern Australia *30*
Conservation implications *34*

CHAPTER THREE: Microsatellite variation, genetic structure and migration in Gouldian Finch *Erythrura gouldiae* and Long-tailed Finch *Poephila acuticauda* populations. **36**

Abstract **36**

Introduction **37**

Methods **40**

Sample collection and DNA extraction *40*
Cross-species amplification of microsatellite loci *41*
Evaluation of microsatellite loci *43*
Analyses of recent population bottlenecks *43*
Genetic diversity and genetic structure *44*
Assignment of individuals to populations *45*
Migration rates *45*

Results **46**

Cross-species microsatellite loci amplification *46*
Bottleneck analyses *47*
Genetic variability, genetic differentiation and inbreeding *47*
Assignment tests *49*
Migration rates *50*

Discussion **50**

Patterns of genetic diversity *50*
Recent population bottlenecks *51*
Population structure and migration *52*
Management implications *52*

CHAPTER FOUR: Gene conversion and natural selection in MHC class IIβ genes of Gouldian Finch *Erythrura gouldiae* and Long-tailed finch *Poephila acuticauda* **55**

Abstract **55**

Introduction **56**

Methods ... **59**

Sampling ... *59*

PCR + 1 cloning of MHC class IIβ genes ... *59*

SSCP gel electrophoresis ... *61*

Sequence analyses ... *62*

Phylogeny reconstruction ... *62*

Detection of gene conversion ... *63*

Detection of positive selection ... *63*

Results ... **65**

Patterns of allele variation ... *65*

Gene conversion ... *67*

Positive selection at MHC class IIβ genes ... *73*

Discussion ... **76**

Maintenance of MHC class II diversity ... *76*

Considerations on statistical methodologies ... *78*

CHAPTER FIVE: General discussion ... **80**

General summary ... **80**

Summary of key findings ... **81**

Conservation implications of results ... **82**

Overall conclusions ... **84**

Future research directions ... **85**

Comparative phylogeography in Australian Tropical Savannas ... *85*

Population structure ... *86*

Evolution and variation of MHC genes ... *86*

REFERENCES ... **89**

APPENDICES 103

Appendix I Posterior means and 95% credible intervals (CI) of panmictic heterozygosity (Hs), *f* and θ^{II} for *E. gouldiae* and *P. acuticauda* under three different models 104

Appendix II Assignment of E. gouldiae and P. acuticauda individuals to populations under hypothetical scenarios of two to six populations (K) 105

Appendix III Amino-acid sequences of MHC class IIβ haplotypes 106

LIST OF TABLES

Table 2.1 Localities of *E. gouldiae* samples used in this study 18

Table 2.2 Variable sites and geographic distribution of the 20 mtDNA control region haplotypes (HVRI, 208bp) in *E. gouldiae*. 23

Table 2.3 Haplotype and Nucleotide diversity of mtDNA control region HVRI for different *E. gouldiae* sample groupings 24

Table 2.4 Analyses of molecular variance, fixation indexes and significance for mtDNA HVRI of *E. gouldiae*. 24

Table 2.5 Mismatch distribution parameters and tests of neutrality for mtDNA control region HVRI sequences of *E. gouldiae* 27

Table 2.6 Pairwise comparisons between mismatch distributions of different geographic regions. 28

Table 3.1 Primer sequences, number of alleles and optimal annealing temperatures for microsatellite loci used in this study 42

Table 3.2 Evaluation of microsatellite loci used in this study 46

Table 3.3 Wilcoxon test and *M* ratio tests for recent population bottlenecks in *E. gouldiae* and *P. acuticauda* populations 47

Table 3.4 Summary of population parameters Hs (panmictic heterozygosity), *f* and θ^{I}, and 95% credible intervals (CI) calculated for *E. gouldiae* and *P. acuticauda* populations 48

Table 3.5 Model fit parameters calculated using the program Hickory 49

Table 3.6 Migration rates and 95% confidence intervals (CI) for *E. gouldiae* and *P. acuticauda* populations, obtained with the program BayesAss 50

Table 4.1 Diversity of MHC class IIβ alleles in *E. gouldiae* and *P. acuticauda* 65

Table 4.2 Hierarchical Analyses of Molecular Variance (AMOVA) for MHC class IIβ alleles in *E. gouldiae* and *P. acuticauda* 66

Table 4.3 Tests of neutrality for MHC class IIβ alleles in *E. gouldiae* and *P. acuticauda* 67

Table 4.4 Detected gene conversion events in MHC class IIβ haplotypes of *Erythrura gouldiae*.. 68

Table 4.5 Detected gene conversion events in MHC class IIβ haplotypes of *Poephila acuticauda* 70

Table 4.6 Amino-acid positions detected under positive selection 74

LIST OF FIGURES

Figure 2.1 Map of Northern Australia showing the distribution of *E. gouldiae* 15

Figure 2.2 Median-joining network of *E. gouldiae* mtDNA control region HVRI haplotypes. 25

Figure 2.3 Mismatch distributions of *E. gouldiae* mtDNA control region HVRI sequences.......... 26

Figure 2.4 Timeline of some climatic and socio-cultural events during the Holocene (above box), and approximate predicted times of *E. gouldiae* population expansions (below box) 33

Figure 3.1 Map of northern Australia showing the distribution of *E. gouldiae* and *P. acuticauda* 39

Figure 3.2 Assignment of *E. gouldiae* and *P. acuticauda* individuals to two hypothetical populations 49

Figure 4.1 Bayesian phylogenetic trees of MHC class IIβ alleles 71

Figure 4.2 Frequency of gene conversions detected between MHC class IIβ haplotypes of *E. gouldiae* and *P. acuticauda*, relative to nucleotide positions......... 73

Figure 4.3 Ratios of d_N/d_S per amino-acid site in MHC class IIβ genes 75

CHAPTER ONE: General Introduction

Genetic diversity and population viability

Genetic diversity can be defined as a level of variation of inheritable traits within a species. Along with ecosystem diversity and species diversity it is one of the three levels of diversity recognised by the world conservation union (IUCN). The reasons for conserving genetic diversity are based on the fact that genetic variability is necessary for the occurrence of evolution, and the effects of genetic variability on population fitness (Reed & Frankham 2003).

Small populations are susceptible to the fixation of random alleles and loss of alleles by drift which reduces the quantitative variation necessary for evolution. In addition, since selection is less effective in small populations, deleterious alleles may accumulate. Small population sizes also reduce the opportunities for random mating, which may result in increased mating among relatives, which promotes inbreeding (Keller & Waller 2002).

It has been revealed that inbreeding can affect individual and population performance across different taxa, although its effects vary across taxa, populations and environments (Keller & Waller 2002). Such loss of fitness due to reduced genetic variability is termed inbreeding depression. Populations of endangered species tend to be small, and therefore inbreeding depression has become a major concern in conservation (Hedrick & Kalinowski 2000).

The role of genetic factors in extinction risk has been previously debated due to the fact that other ecological factors such as high predation, disease outbreaks and metapopulation dynamics appear to drive small populations to extinction before genetic effects take place (Lande 1988; Caro & Laurenson 1994). However, strong evidence from experimental studies and case studies of some wild populations, combined with the fact that inbreeding depression interacts with other ecological factors (e.g. reduction of fertility, increased susceptibility to parasites) and the fact that genetic diversity tends to be lower in small and endangered populations (contrary to expectations if ecological factors were more important than genetic factors), points to the conclusion that genetic factors are involved in extinction risk. Failure to consider genetic factors may result on

the underestimation of extinction risks, and the implementation of inappropriate recovery strategies (Frankham & Ralls 1998; Frankham 2005).

Molecular genetics in ecological research

Molecular genetics has transformed research in ecology and evolution by allowing the genetic characterisation of individuals, populations and species. Different molecular methods are available and are increasingly becoming routine in ecological research (Freeland 2005). There are a number of laboratory techniques available for the study of nucleic acids and proteins. The applications of these techniques in ecological research are equally varied. Some of the more common applications are outlined below.

Resolution of taxonomic uncertainties

Knowledge of systematics is fundamental to ecology (Beebee & Rowe 2004). Organisms often evolve as separate evolutionary entities as a result of isolation of their populations. The degree of differentiation between populations is not always detectable using traditional taxonomic methodologies such as external morphology and anatomy. Since taxonomic classification is central for the determination of management priorities, it can have direct implications in conservation. DNA-based taxonomy can overcome many of the uncertainties of more traditional methodologies, although its role is still far from central in taxonomy (Tautz *et al.* 2003). One classical example of the implications of molecular-based taxonomy in species delimitation is that of the tuataras (*Sphenodon spp.*), which were split into two species after combined analyses of allozyme variation and morphology (Daugherty *et al.* 1990). The value of molecular systematics is not restricted to species delimitation. Phylogenies also contain information about the taxonomic distinctness of species, which can be used as a quantitative measure alternative to simple species count for the determination of management priorities (May 1990).

Calculation of genetic diversity

Knowledge of genetic diversity is essential in population ecology, as it is necessary for populations to adapt to environmental change (Beebee & Rowe 2004; Freeland 2005). Genetic diversity can be calculated using different molecular markers.

The interpretation of genetic diversity, either as genotype frequencies or allele frequencies is important in conservation given the deleterious effects that reduced genetic diversity can have on population fitness (Freeland 2005).

Diversity can be measured at the level of allele richness, allele frequencies at one or several loci, haplotype diversity, nucleotide diversity (for DNA sequences) or quantitative genetic variation. Genetic variation at neutral markers (such as microsatellite loci or non-coding DNA sequences) is normally used as a surrogate of overall genetic diversity. It has been found however that the correlation between neutral variation and quantitative variation is very weak or absent (Reed & Frankham 2001; Bekessy *et al.* 2003). This may be due to the fact that selection at neutral markers is effectively neutral, while quantitative variation is often affected by mutation at several loci (Reed & Frankham 2001). Nonetheless, the calculation of neutral variation is important, as it is related to fitness at the population level, and is useful for the detection of inbreeding depression (Reed & Frankham 2003). When information about the short-term ability of a population to adapt is required, it is necessary to measure directly quantitative variation instead of using neutral genetic variation as a surrogate (Reed & Frankham 2001).

Mitochondrial DNA has been commonly used in population genetics, given that it is often considered to be effectively neutral (Avise *et al.* 1987). However, Bazin *et al.* (2006) recently pointed out that mtDNA diversity may not reflect species abundance or ecology as expected for genuinely neutral markers; reasons for regarding mtDNA as a useful marker were noted by others (Berry 2006; Wares *et al.* 2006), and a strong correlation between mtDNA and allozyme diversity for eutherian mammals pointed out by (Mulligan *et al.* 2006).

Phylogeography

The geographic distribution of lineages can reveal aspects of the history of populations such as evidence of population bottlenecks (i.e. drastic reductions in population size), gene flow and population expansions that can be the consequences of past adaptations of organisms to changing environments (Avise 2000; Excoffier 2004). The development of coalescent theory has resulted in a growing interest in reconstructing intra-specific phylogenies, in order to resolve the genetic relationships of different lineages, back to a common ancestry in the past (Crandall & Templeton 1993). The signature that demographic events such as bottlenecks and population expansions

make can be detected in the distribution of haplotypes in populations (Rogers & Harpending 1992; Harpending 1994; Excoffier 2004). Such information can be used to calculate the times since coalescence of populations or since expansion or historical bottleneck events, making possible to relate the inferred population history of species to past habitat changes, as inferred by palaeoclimatic and palaeontologic evidence.

Resolution of population structure and gene flow

Although the detection of genetic variation in single populations is important, the genetic analyses of multiple populations can reveal other important aspects of species biology (Frankham *et al.* 2002; Freeland 2005). Populations can be structured at different levels as a result of mutation, drift due to finite population size, and selection acting as adaptation to local environments (Slatkin 1987). Gene flow (the movement of gametes, individuals or populations) can counteract the effects of isolation and local adaptation by spreading new alleles or preventing adaptation to local conditions (Slatkin 1987).

DNA analyses of multiple populations can help reveal the degree to which subpopulations have evolved into separate evolutionary entities and to detect movements of individuals between subpopulations, making possible the differentiation between completely isolated populations and different types of fragmented population structuring (e.g. island-mainland, stepping stone, metapopulations) (Frankham *et al.* 2002). The parameters F_{ST}, F_{IS} and F_{IT} (referred to as fixation indices or F-statistics) have been introduced as an effective measure of population structure (Weir & Cockerham 1984). They describe the amounts of heterozygosity at different levels of population structure (e.g. within individuals or among populations).

Resolution of mating systems

Sexual selection has led to a number of mechanisms enabling organisms to compete for reproduction. Mating systems vary in organisms from complete monogamy to several degrees of polygamy, to promiscuity (Beebee & Rowe 2004). Mating systems have often been inferred from field observations, but these observed mating systems often differ from real ones due to behavioural adaptations such as extra-pair paternity. Often the true mating system can only be resolved with the use of molecular markers (Beebee & Rowe 2004). The identity of parents and the heritability to offspring can be established with the use of polymorphic markers such as allozymes and microsatellites.

Adaptive variation

Neutral variation is important in the knowledge of several aspects of species biology such as dispersal, phylogeography and mating systems, but it can be a poor indicator of adaptive variation of populations and species. Unlike neutral variation, adaptive variation is concerned with traits that are of adaptive importance and often under selection, and that are important for the viability of populations (Beebee & Rowe 2004). Quantitative trait loci (regions of DNA that are associated with particular continuous phenotypic trait) can be used in combination with neutral genetic markers to investigate the relative importance of random genetic drift and selection on population differentiation (Merilä & Crnokrak 2001).

The study of genes that are known to be functionally important can be a good indicator of adaptive variation. Genes of the Major Histocompatibility Complex (MHC) of vertebrates are ideal candidates for the study of adaptive variation in host-parasite systems (Edwards & Hedrick 1998; Beebee & Rowe 2004).

Some novel genomic techniques such as suppression subtractive hybridisation and cDNA macroarrays have been used to identify genes that are differentially expressed after parasite infections (Wang *et al.* 2006). Such genes have the potential to be used as markers to investigate adaptive variation in short timescales, which may be relevant to determine the capacity of populations to respond to disease outbreaks or other stochastic events.

Conservation genetics

There is an alarmingly increasing number of ecological communities and species that are becoming threatened or endangered. The causes of the decline of species and populations are largely linked to the direct or indirect effect of human activities. Some of the major threats that have been identified include the loss of habitat, the introduction of exotic species, illegal trade and pollution (Frankham *et al.* 2002; CBD 2006).

Habitat loss leads to the fragmentation of suitable habitat for species. Habitat becomes fragmented in patches, which can lead to the fragmentation of large populations into population patches of smaller sizes (CBD 2006). There are several consequences of the reduction of populations on their genetic variability. The loss of

genetic diversity can be detrimental in small populations, since genetic variability is required for populations to adapt to environmental change (Frankham *et al.* 2002; Beebee & Rowe 2004; Freeland 2005). When a population is drastically reduced, the gene pool is also reduced; inbreeding increases and the risk of losing alleles of adaptive importance becomes larger. The risk of genetic diseases also becomes larger due to inbreeding and the fixation of deleterious alleles. Reduction of genetic variation can then result in inbreeding depression, the loss of fitness due to loss of genetic heterozygosity (Frankham *et al.* 2002). The relatively new discipline of conservation genetics deals with the genetic factors that affect extinction risk and the genetic management required to maintain genetic diversity in populations.

Knowledge of genetics of populations and species is useful in conservation in different ways. Some important applications include the resolution of taxonomic uncertainties; detection of inbreeding and loss of genetic diversity; identification of populations of conservation concern; resolution of genetic population structure; identification of management units within species; detection of hybridisation; identification of populations and sites for reintroduction; forensic analyses of products of animal and plant origin; and the knowledge of other aspects of species biology relevant to conservation such as mating systems, dispersal, migration, sex identification and parasite detection (Frankham *et al.* 2002).

The use of non-destructive tissue material as a source of nucleic acids for genetic studies has allowed sampling of individuals regardless of the conservation status of a species. Sampling of museum and herbarium specimens has also allowed genetic studies in some extinct species whose tissues have been preserved in museums. In recent times, natural history collections have been an invaluable source of genetic material for conservation genetic research (Wandeler *et al.* 2007). The use of biological material from zoological collections and herbaria has allowed researchers to investigate aspects of the biology of species that would otherwise be difficult to determine, such as loss of genetic diversity over time, detection of introgression and changes in connectivity, as well as investigating genetic diversity in past populations of critically endangered or extinct species (Wandeler *et al.* 2007).

The dramatic loss of natural habitats, species and populations observed at present suggests that conservation ex-situ (i.e. conservation outside the places where a species is normally found) will be a necessary practice to save species that will be incapable of surviving in their habitats in the near future (Frankham *et al.* 2002). Ex-situ

conservation generally involves founding a captive population, growing the population to a target size, managing the population for a number of generations, reintroducing individuals and managing the reintroduced population, relocating individuals to areas where the threat for their survival is absent, or to breeding individuals in captivity for later release (Frankham *et al.* 2002). There are important considerations on the genetic management of captive populations that relate to the reduced population sizes. One immediate problem that arises from these practices is the limitations in space for breeding, which consequently means the number of founder individuals is limited as well. In cases where the number of founder individuals is dramatically reduced, the success of captive breeding relies on more active genetic management aimed to maximise effective population sizes (Frankham *et al.* 2002). Zoos and botanical gardens around the world are involved in collaborative captive breeding, in which animals and plants are exchanged among zoos in order to interbreed them and to maximise the genetic variability in critically endangered species such as large felines (Frankham *et al.* 2002).

There have been important efforts to recover threatened populations of different taxa. Effective management and recovery of endangered species relies on a good understanding of ecological factors that can drive important changes in population trends, as well as on knowledge of genetic factors. The study of genetics is nowadays essential to the conservation of threatened species. There are a number of adverse effects of ignoring genetic factors in conservation planning, including: underestimation of extinction risk, inappropriate implementation of recovery strategies, the inappropriate use of individuals for reintroduction, failure to characterise fragmented populations and the detrimental effects of outbreeding depression (Frankham 2005). Considering these risks, assessments of susceptibility to extinction and recovery plans for threatened taxa that do not take account of genetic factors are no longer regarded as scientifically realistic (Frankham 2005).

Although much of the research in conservation genetics has focused on single species, genetic information can also be applied at the community level. For instance, phylogenetic diversity can be used to set priorities for conservation at the community level (Crozier 1997; Moritz & Faith 1998; Diniz-Filho 2004). Phylogenetic diversity can be a better indicator of conservation worth than other measures more traditionally used, such as species richness as it takes into account genetic content regardless of the taxonomic status of organisms (Crozier 1997; Moritz & Faith 1998).

Study species

The Gouldian Finch (*Erythrura gouldiae*, Aves: Passeriformes: Estrildidae) is endemic to the tropical savannas in Northern Australia. Its original distribution includes most of the tropical savannas in northern Australia, north of 20° latitude south (Barret *et al.* 2003).

Habitat for *E. gouldiae* in the Australian tropical savannas is characterised by a striking seasonality and a lack of prominent topography (Woinarski *et al.* 2005). The marked seasonality of Australian tropical savannas is reflected in the production of seeds on which *E. gouldiae* feed. Availability of seeds is higher during the early dry season, after which seeds become scarcer though removal by ants, rodents and birds, destruction by fire, burial in the soil, and being washed away by early rains (Dostine *et al.* 2001). Seed-eating birds may experience food shortages at the beginning of the wet season, when the seed supply from the previous season is exhausted (Dostine *et al.* 2001). In response to these changes in food availability, *E. gouldiae* undertake seasonal movements from breeding areas on hill woodlands to adjacent lowlands during the wet season (Dostine *et al.* 2001). Breeding in *E. gouldiae* occurs from January to August and depends on precipitation. Annual productivity calculated as nesting success has been calculated to be approximately 1.5 fledglings per adult per season (Tidemann *et al.* 1999).

Population trends and current threats

Erythrura gouldiae is a charismatic species that has become popular in aviculture. During the 1970s it was subject to intensive trapping for the pet market, which continued until it was declared a protected species in 1981 (Franklin *et al.* 1999). The declining trend in populations has been continuous even after the commercial trapping ceased. The decline in *E. gouldiae* populations has resulted in a marked reduction in the distribution of the species. Its present distribution consists of a reduced and fragmented portion of their original distribution (O'Malley 2006). The impacts have been the most severe in Cape York Peninsula, where recent sightings of have become increasingly uncommon and unpredictable.

The decline of *E. gouldiae* populations has been attributed to several causes; among which trapping for the aviculture market, cattle grazing activities, uncontrolled fire regimes and disease have drawn attention (Tidemann *et al.* 1992a; Franklin *et al.* 1999; Woinarski *et al.* 2005; O'Malley 2006). Cattle grazing affects tropical savannas by altering the soil quality and the availability of waterholes on which wildlife rely for drinking. The intensity and history of cattle grazing in northern Australia is consistent with the patterns of population declines of *E. gouldiae*. At the easternmost part of *E. gouldiae* distribution in Cape York Peninsula, cattle grazing activities have been established earlier than at the rest of the tropical savannas, and it is in this region where *E. gouldiae* decline has been most severe (Franklin 1999; Franklin *et al.* 2005).

At the onset of the wet season in the tropical savannas, there is a period of shortage of seeds available for *E. gouldiae* and other granivorous species. During this period, *E. gouldiae* rely on ripening seed from a number of perennial grass species (Dostine *et al.* 2001). Uncontrolled fires, particularly at the end of the dry season can destroy patches of perennial grasses, consequently depleting habitat of seeds during the resource bottleneck period (O'Malley 2006).

It has been suggested that disease may be a potentially driving factor for the decline of *Erythrura gouldiae* in the wild (O'Malley 2006). That theory is supported mainly by a study in which a higher incidence of the parasitic air-sac mite *Sternostoma tracheacolum* (Tidemann *et al.* 1992a) was found in *E. gouldiae*, compared to other tropical savannah bird species. This bird species also has a reputation of being delicate and susceptible to diseases in aviculture (Gelis 2003) which is consistent with the apparent higher susceptibility to pathogens in wild populations.

Information on genetic variation previous to the research described in this thesis is limited to study using allelic variation at an intron of the myoglobin gene (Heslewood *et al.* 1998) in which the authors found limited evidence of a lack of genetic differentiation between three geographically distant populations.

Conservation status and current conservation measures

Erythrura gouldiae is considered an endangered species under IUCN- World Conservation Union red list category EN C2a(ii) - a total population smaller than 2500 mature individuals that has had a continuous decline trend, with at least 95% of mature

individuals in a subpopulation (BirdLife International 2006) and is also considered endangered under Australian legislation (DEWR Australia 2007).

This categorisation, however, might be subject to revision in the near future, as recent data suggests the total population might exceed 2500 mature individuals, populations might not be declining any longer and *E. gouldiae* might not exist as a single population (O'Malley 2006). Although these considerations might result in the modification of the conservation status of this species, it would be classified as a vulnerable species of conservation concern (O'Malley 2006).

There is currently a recovery program for *E. gouldiae*. The objective of the recovery plan is to improve the conservation status of *E. gouldiae* through population increases. The specific actions to achieve recovery include: the implementation of adequate fire regimes at key sites, fire and cattle management at off-reserve key sites, the monitoring of individual and population health parameters, a reintroduction program and the encouragement of community participation (O'Malley 2006).

The decline of *E. gouldiae* rather than being an isolated phenomenon, correspond to human-induced habitat changes of the tropical savannas that has affected a number of different taxa, including 12 of the 49 native species of granivorous birds indigenous to the tropical and subtropical savannas (Franklin 1999; Franklin *et al.* 2005). Because *E. gouldiae* is a charismatic species, the noticeable decline in numbers during recent times has helped raising awareness about the conservation status of the tropical savannas. Habitat management aimed to restore suitable habitat for this species is likely to have a wider positive effect on the conservation of other species and ecological communities in the Australian monsoonal tropics, including other threatened and declining bird and mammal species (O'Malley 2006).

Given the lack of genetic studies in *E. gouldiae*, at present the recovery plan for this species does not take into account genetic factors on the recovery of the species that can be critical for the appropriate implementation of management strategies.

Research aims

The general aim of the research in this thesis is to use molecular genetic techniques to resolve aspects of the ecology of *E. gouldiae* that are relevant for their

conservation. This thesis is organised as individual research papers (Chapters 2, 3 and 4) that are interconnected.

In Chapter 2, I investigate the phylogeography of *E. gouldiae* using mitochondrial DNA (mtDNA) control region sequences. In this study I included samples from wild contemporary populations, specimens from museum collections that together represent 18 locations within their original geographic distribution. I look for evidence of heterozygosity loss over time; and of past demographic changes. I discuss such changes in relation to past climatic and socio-cultural events that produced habitat changes in tropical savannah habitats of Northern Australia.

In Chapter 3 I assess the usefulness of some previously developed microsatellite PCR primers in *E. gouldiae* and the sympatric and comparably more abundant long-tailed finch *Poephila acuticauda*. I used the obtained microsatellite genotypes of individuals from both species, sampled at two field locations in Northern Australia, as well as a captive population of *E. gouldiae* to calculate genetic diversity (Heterozygosity), population structure, to look for the evidence of recent population bottlenecks and to calculate reciprocal migration rates between the two sampled locations.

In Chapter 4, I investigate variability of genes of the Major Histocompatibility Complex (MHC) in *E. gouldiae* and *P. acuticauda* individuals. I also look for evidence of gene conversion events between alleles, and look for evidence of positive selection, in order to determine the relative importance that these evolutionary processes have had in the maintenance of variability of these genes, that has an important function in pathogen recognition and the initiation of the immune response.

In Chapter 5 I summarize the results of chapters 2, 3 and 4, and discuss their implications on current and future conservation activities for the recovery of *E. gouldiae* populations.

CHAPTER TWO: Historical phylogeography of the endangered Gouldian Finch *Erythrura gouldiae*: evidence of gradual population expansion during the Holocene.

Abstract

The present distribution of extant species has been influenced by historical events of habitat change that can have an effect on cause population sizes and distribution. These types of fluctuations often have had an effect on the genetic makeup of populations. Although the effects of past climatic changes are well documented for temperate environments, these are relatively less clear for the tropics. In this study I analyse mitochondrial DNA control region variation in the Gouldian Finch (*Erythrura gouldiae*), a species endemic to the Australian tropical savannas that has experienced drastic population declines in recent times. Results show a lack of lineage differentiation between individuals across the present distribution of this species, and a total lack of genetic diversity for a population that is part of a captive breeding and reintroduction program. Analyses of mismatch distributions fit predictions of population growth consistent with range expansions, dating during the Holocene period from approximately 7000 to 3600 years before present. Differences in mismatch distributions between separate geographical regions suggest population expansion events happened at different times in different regions. This gradual colonisation of northern Australia can be explained in terms of important climatic and socio-cultural changes during the Holocene that were favourable for the growth of *E. gouldiae* populations. The apparent lack of genetic differentiation between populations from different geographical regions, and the relatively low sequence divergence suggest a recent colonisation of the tropical Savannas by *E. gouldiae*. Implications of these results for current management practices aimed to recover *E. gouldiae* populations are discussed.

Introduction

Historical population dynamics, as well as current ecological and climatic factors, have had an important role in shaping the present geographic distribution of organisms (Beebee & Rowe 2004). Phylogeographic patterns (the distribution of phylogenetic lineages in geographic areas) can reveal traces of the population history of a species, caused by gene flow, coalescent events, population bottlenecks, and population expansions (Avise 2000; Excoffier 2004). Historical population trends can have important implications in the conservation and management of endangered species. For instance, geographically separated populations may consist of different evolutionary units that require separate management (Moritz 1994a; Frankham *et al.* 2002). Knowledge of phylogeography can be helpful in identifying the most suitable populations as sources of founders for captive breeding or reintroduction programs for endangered species (e.g. Matthee & Robinson 1999; Negro & Torres 1999; Godoy *et al.* 2004).

Past climatic events have had a large affection in the present patterns of distribution of extant species. Effects of Pleistocene glaciations driving past population contractions and expansions have been well documented for species inhabiting high latitudes and temperate climates, with common extinction and re-colonisation events during the Quaternary Ice Ages (Hewitt 2004). In contrast, the effects of past climatic events on populations of tropical and subtropical species have been comparably less studied, despite the high biodiversity of tropical regions (Hewitt 2004; Wüster *et al.* 2005; Davison & Chiba 2006). Evidence of altitudinal shifts in mountain forest species within the tropics, driven by Pleistocene climate fluctuations, has been found (Hewitt 2004). However, the effects of Pleistocene glaciations in other tropical environments are less well understood and are often subject to debate (Wüster *et al.* 2005). There is evidence that climate fluctuations during the Holocene have lo had an effect on the stability of habitats, and consequently on species diversity in temperate regions (Araújo *et al.* 2008). It is likely that climate fluctuations during the Holocene may have had also important effects on tropical environments.

Landscapes in tropical savannas of northern Australia are characterised by a relative uniformity in plant species composition, with local variation associated with factors such as fire history, rockiness and geology (Woinarski *et al.* 2005). Rainfall in Australian tropical savannas is highly seasonal compared to other monsoonal tropical

areas in the world, with a short wet season lasting for three to four months, followed by a long dry season. The seasonal variation in rainfall translates into marked fluctuations in resources throughout the year (Dostine *et al.* 2001; Woinarski *et al.* 2005).

For vertebrate fauna, there are three areas of endemism that have been identified within the tropical savannas of northern Australia: The Kimberley Plateau (north west), Arnhem Land (north central) and Cape York Peninsula (north east) (Cracraft 1991). There are delimited boundaries between these areas of endemism, which break the continuum of savannah habitats between them. The Kimberley Plateau is separated from Arnhem land on the east by the lowlands of the Victoria and Daly River valleys, whereas Arnhem land is separated from savannas in Cape York Peninsula by drier habitats around the Gulf of Carpentaria (Cracraft 1991). Several plant and animal species have subspecies or local races whose distribution limits are concordant with such habitat boundaries. For instance, distributions of species and subspecies of savannah grass finches (*Poephila spp.*) are delimited by the Kimberley Plateau – Arnhem Land and Carpentarian barriers (Cracraft 1986). Amongst them, the Long-tailed Finch *Poephila acuticauda* has two morphologically different races restricted to the Kimberley plateau and Top End regions respectively. The time of divergence of these two forms has been calculated to be approximately 0.3 million years before present (Jennings & Edwards 2005).

The population history of other species inhabiting the Australian tropical savannas and other dry habitats varies considerably. Populations of the Ghost bat *Macroderma gigas* show an extreme population structure (Worthington-Wilmer *et al.* 1994). In contrast, it has been found that two related species of woodswallows (*Artamus superciliosus* and *A. personatus.*) of the Australian dry environments show a polyphyletic origin with respect to each other, in spite of external morphological differences (Joseph *et al.* 2006). These contrasting differences have been attributed to a high phylopatry in the former, and an incomplete sorting of the common ancestor combined with dispersal abilities and possible introgression in the case of the latter.

The Gouldian Finch (*Erythrura gouldiae*) is a granivorous bird, endemic to the tropical savannas in Northern Australia (Dostine *et al.* 2001). These birds were formerly distributed across northern Australia, from approximately 17º 57’ S; 122º 14’ E on the westernmost limit, across the Kimberley plateau and Arnhem Land in northwest and north-central Australia, respectively, through a narrow corridor along the south-western

shore of the Gulf of Carpentaria, and in Cape York Peninsula in the northwest, as far east as approximately 20º S; 146º E (Fig. 2.1).

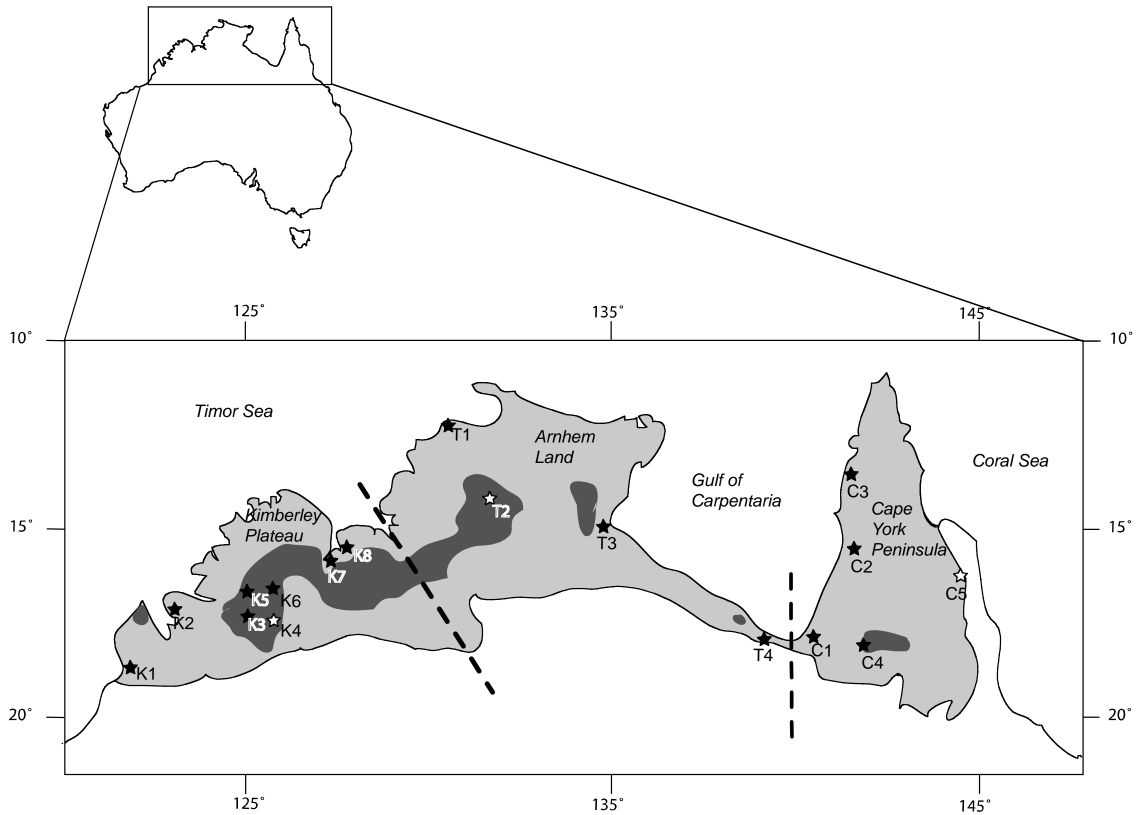


Figure 2.1 Map of Northern Australia showing the distribution of *E. gouldiae*. Light grey shaded areas represent the past distribution and dark grey shaded areas represent the current distribution. White stars represent blood sample collection localities. Black stars represent localities of origin of museum specimens used in this study. Location codes correspond to those used in Table 1. Dashed lines represent the approximate locations of natural barriers for tropical savannas. Habitat within the shown present distribution may be discontinuous and fragmented. Adapted from O'Malley (2006), and records in the Atlas of Australian Birds (Barret *et al.* 2003).

During the past four decades, suitable habitat for *E. gouldiae* has been degraded through human-related activities, such as intensification of cattle grazing and changes in fire regimes. Habitat degradation has resulted in the continuous decline of *E. gouldiae* populations over the last four decades. The decline in *E. gouldiae* numbers has been more evident in the eastern parts of the original distribution, particularly in Cape York Peninsula, from where most populations have disappeared. Current distribution of *E.*

gouldiae consists of a small proportion of the area in which it was once abundant (Fig. 2.1). *E. gouldiae* has been categorized as an endangered species on a global scale, on the basis of population size estimates of less than 2500 breeding adults in the wild, and a declining trend in population sizes (BirdLife International 2006). Key threats to the survival of this species include uncontrolled fire regimes, extensive cattle grazing, disease, and potential effects of predicted future climate change (O'Malley 2006).

There are currently a number of efforts to restore *E. gouldiae* populations. These efforts include long term population monitoring, improved fire management and removal of cattle in protected areas, and a captive breeding program at the Mareeba Wetlands at the eastern limit of the past distribution of *E. gouldiae*. The captive breeding program aims to re-establish a new population in that region, where *E. gouldiae* has been absent for several decades. The captive population of Mareeba Wetlands consists of birds descendant from individuals that were captured in Queensland (northeast Australia) during the 1980s (Nevard *et al.* 2002).

Suitable habitats for *E. gouldiae* may be divided by the above-mentioned natural barriers between areas of endemism of Australian tropical savannas, and this is expected to be reflected in the genetic population structure for this species. However, at present, the genetic structure of *E. gouldiae* populations is largely unknown. A previous study showed that, based on allelic variation in intron 2 of the myoglobin gene, populations of *E. gouldiae* in the Kimberley region, the Top End and Cape York Peninsula appear to lack in significant genetic differentiation (Heslewood *et al.* 1998). However, ability to identify population sub-structuring may have been compromised by the use of allelic variation at a single locus.

In this study I use mtDNA control region sequences obtained from wild and captive extant *E. gouldiae* populations as well as from specimens in museum collections, in order to determine the geographic distribution of lineages in present and past populations. Levels of haplotype diversity are compared between museum specimens and extant populations in order to determine wether there has been reductions in genetic diversity following the reported population size fluctuations in recent times. I calculate mismatch distributions of mtDNA control region to detect the signature of past expansion events in *E. gouldiae*, and to calculate the approximate time since coalescence in past populations. Calculated expansion times are discussed in relation to historical changes in habitat caused by climatic and socio-cultural events in

northern Australia. I further discuss the implications of the results obtained with respect to current management efforts for the recovery of *E. gouldiae* populations.

Methods

Sampling

A total of 157 individuals were used in this study. These individuals are representative of 18 localities from the past and present distribution of *E. gouldiae* (Table 2.1; Fig. 2.1). Blood samples were obtained from two of the remaining locations where *E. gouldiae* can be found reliably (Figure 1): Yinberrie Hills in the Top End (NT) (n=50) and Mornington Wildlife Sanctuary in the Kimberley region (n=30). Further blood samples from the captive population of *E. gouldiae* at the Mareeba Wetlands in Queensland were used (n=34). These samples are part of the founding population for the ongoing reintroduction program for north-eastern Australia. Blood samples were stored in either Queens lysis buffer (Seutin et al., 1991) or 70% ethanol prior to DNA extraction.

Foot skin or feather samples from specimens stored in museum collections in Australia and the United Kingdom provided genetic material of individuals from 15 additional locations across Northern Australia, including some at which *E. gouldiae* are now extinct (Table 2.1; Fig. 2.1).

DNA extractions

DNA was extracted from blood samples using a phenol-chloroform protocol (Friesen *et al.* 1997). DNA was extracted from museum specimens using a DNeasy kit (Qiagen) following recommendations of the manufacturer. In order to avoid cross-contamination with DNA obtained from blood samples, I carried out all extractions of DNA from museum specimens under a UV laminar flow unit, using a separate set of reagents, pipettes and other equipment from those used for DNA extraction from blood samples. All equipment and materials were sterilized before each round of DNA extractions of museum material. In order to detect possible cross-contamination, I included negative extraction controls (with the same amount of reagents but without starting tissue) in a 1:5 ratio for museum samples and 1:12 ratio for blood samples.

Table 2.1 Localities of *E. gouldiae* samples used in this study

Region	Locality name	Locality code	No. of samples	Latitude S	Longitude E	Year Collected
Kimberley	Broome	K1	2	17° 57'	122° 14'	N/A
	Kimbolton	K2	1	16° 41'	123° 50'	1975
	Mount Bell	K3	2	17° 09'	125° 17'	1969
	Mornington*	K4	30	17° 31'	126° 07'	2004-2005
	Beverley Springs	K5	1	16° 34'	125° 29'	1960
	Mount Elizabeth	K6	1	16° 23'	126° 16'	1969
	Wyndham	K7	3	15° 29'	128° 07'	1940
	Ord River	K8	2	15° 03'	128° 21'	1945/1971
	"Kimberley" [a]	K9	2	-	-	1891
Top End	Darwin	T1	10	12° 27'	130° 50'	1905
	Yinberrie Hills*	T2	50	14° 06'	132° 04'	2003
	Roper River	T3	1	14° 44'	135° 23'	1897
	Burketown	T4	9	17° 44'	139° 32'	1896 (n=8) 1911(n=1)
Cape York Peninsula	Normanton	C1	6	17° 40'	141° 04'	1897(n=1) 1905(n=5)
	Mitchell River	C2	1	15° 11'	141° 35'	1965
	Watson River	C3	1	13° 21'	141° 43'	1914
	Croydon	C4	1	12° 27'	142° 14'	1893
	Mareeba *[b]	C5	34	16° 59'	145° 25'	2002-2005
	Total		**157**			

* Contemporary populations
[a] Museum specimens lacked detailed locality information for these samples
[b] Captive population for reintroduction program

Amplification of mtDNA control region

I designed primers to amplify the mtDNA control region based on the alignment of a partial sequence of Gouldian Finch control region, GenBank accession AF407116 (Sorenson & Payne 2001) with control region sequences of other estrildid finch species: the Indigobird *Vidua chalybeata,* GenBank accession AF090341 (Mindell *et al.* 1998) and Zebra Finch *Taenyiopygia guttata*, GenBank accession DQ422742 (Mossman *et al.* 2006). Primer GouldCR5PF (5'-GCC CGC CGT TAT GAA TTT AAC-3' [located in the tRNA-Glu gene]) was paired with Primer FinchCR3PR (5'GGC CGT CTT GAC ATC TTC AG-3' [located in the tRNA-Phe gene]) to amplify the control region of contemporary samples (Mornington, Yinberrie and Mareeba populations). Polymerase Chain Reactions (PCR) were performed in a 25 µl final volume containing 1x GoTaq Green Master Mix (Promega), 10 pmoles of each primer and ~40 ng of template DNA.

Given that DNA obtained from museum specimens is highly degraded; a smaller section of the control region was amplified to improve the yield of PCR products from such samples. A total of 20 haplotypes (GenBank accession numbers EF094893 to EF094912) were identified from 48 sequences 1204 bp long, which were obtained with primers GouldCR5PF and FinchCR3PR. After aligning such haplotypes, I identified a section that contained the hypervariable region I (HVRI) of the control region. Primers GouldmtF (5'-CGT TGA GTA GTT CGG TTC TCG-3') and GouldmtR (5'-ACA GCC CAA GTG ATC CTA CC-3') were designed internally to amplify a 208 bp section (excluding primers) of the HVRI.

PCR reactions for DNA extracted from museum specimens were performed in a total volume of 50 µl and included 10 mM Tris-HCl (pH 8.4), 50 mM KCl, 2.5 mM $MgCl_2$, 0.8 µg/µl of bovine serum albumin, 0.2 mM of each dNTP, 20 pmoles of each primer, 1 unit of platinum *Taq* DNA polymerase (Invitrogen) and 2-5 µl of the initial extraction solution. In order to avoid cross-contamination of PCR reactions for museum specimens, the same precautions mentioned above for DNA extractions were applied for the respective PCR amplifications. All PCR reactions were run on a GeneAmp PCR System 9700 Thermal Cycler (Perkin-Elmer) with the following cycle conditions: an initial five minute denaturation at 94° C; followed by 35 cycles of denaturing at 94° C for 30 s, annealing at 52° C for 30 s and extension at 72° C for 40 s (90 s for amplification of the whole control region); and one final extension at 72 °C for ten minutes. For amplification of museum specimens I increased the number of cycles to 45 and added two seconds of extension at each cycle to improve yield (Ellegren 1994). Negative controls (with all the reagents but replacing DNA template with distilled water) were included in each set of PCR reactions in the same ratios as for negative DNA extraction controls. PCR products were purified using a Qiaquick kit (Qiagen), and were sequenced in forward and reverse directions. Sequences were aligned and sought for sequencing errors using Sequencher 4.5 (Gene Codes Corporation).

Genetic diversity

Haplotype diversity (*h*), nucleotide diversity (*d*) and their respective standard deviations were obtained using the program DnaSP 4.10 (Rozas *et al.* 2003) for different sample groupings: all samples, contemporary versus museum samples, and samples grouped in different geographical regions: Kimberley (KI), Top End (TE) (Which includes the Arnhem Land area of endemism mentioned above, plus the narrow

strip of savannas along the South-western shore of the Gulf of Carpentaria), Cape York Peninsula (CYP), and each of the sampled contemporary populations (Yinberrie Hills, Mornington Wildlife Sanctuary and the captive population in Mareeba Wetlands). Differences in haplotype and nucleotide diversities were tested using a modified *t* test (Nei 1987).

Analysis of Molecular Variance and Median-joining network

Hierarchical genetic population structure among regions and populations was evaluated using an Analysis of Molecular Variance (AMOVA, Weir & Cockerham 1984; Excoffier *et al.* 1992). As implemented in ARLEQUIN 3.01 (Excoffier *et al.* 2005). A distance matrix for the analysis was calculated using pairwise nucleotide differences. Significant deviations of fixation indices from zero were calculated by bootstrapping 10,000 times. Only populations with more than 5 individuals were included in the AMOVA analysis. Populations T1, T2 and T4 were grouped together, whereas populations K4 and C1 were considered a different group each.

In order to reconstruct evolutionary associations between HVRI haplotypes, a median-joining network (Bandelt *et al.* 1999) was constructed using the program Network 4.112 (www.fluxus-engineering.com). I used a tolerance level value (epsilon) of 0 to obtain the smallest number of alternative nodes between haplotypes. Samples from the captive population at Mareeba were treated as a separate group for this analysis.

Mismatch distributions

Populations that have experienced stable demographic history will show a multi modal distribution of pairwise haplotype differences. On the other hand, populations that have experienced a population expansion will show a unimodal Poisson-like distribution. These mismatch distributions were calculated using ARLEQUIN 3.01 (Excoffier *et al.* 2005). The program simulates mismatch distributions under a model of spatial expansion, which is suited to describe the expected pattern of genetic diversity under the scenario of a range expansion in species distributed over a large number of demes (Excoffier 2004). The program additionally calculates a raggedness index (Harpending 1994) for observed and simulated distributions. Goodness of fit tests between simulated and observed distributions, and their respective raggedness indices were also performed. Mismatch distributions were calculated for all samples combined

(excluding the captive population), as well as for the different geographic regions (KI, TE and CYP), and datasets containing contemporary populations only (Yinberrie Hills and Mornington Wildlife Sanctuary) and museum specimens only.

To determine whether different populations experienced the same expansion event, I carried out pairwise comparisons between simulated mismatch distributions of different geographic region groupings. Significance of the differences between mismatch distributions was tested with a goodness of fit test between the simulated mismatch distribution of each of the groupings to be tested, and their respective "intermatch distribution" (calculated from the combined datasets of the two datasets being tested) as recommended by Excoffier (2004).

Tests of neutrality

Although mtDNA control region is expected to be neutral, it is known that mitochondrial genes can be influenced by selection (Bazin *et al.* 2006) that can affect demographic estimations. Different tests of neutrality were performed on the same datasets used for calculating mismatch distributions, in order to determine deviations from neutrality and the causes of such deviations, if present. Tajima's D test of neutrality (Tajima 1989) was calculated to detect deviations from neutrality that can be caused by selection, genetic hitchhiking or population growth (Fu 1997). Fu's F_S test (Fu 1997) was calculated as an indication of population growth. Fu and Li's $D*$ and $F*$ tests (Fu & Li 1993) were carried out to detect background selection. All neutrality tests were carried out using DnaSp 4.01 (Rozas *et al.* 2003).

Relative dating of population expansions

The mismatch distribution analyses further allow estimation of the age of expansion in evolutionary units (τ). Values of τ were used to estimate the relative time in years since genetic coalescence (t) using the equation $t = \tau / 2u$ (where $u = \mu L$; μ = mutation rate and L = the sequence length in base pairs), and multiplying by the generation time. Dating expansion times requires an accurate estimate of μ. Since the mutation rate of control region has not been estimated for *E. gouldiae*, I used a μ value of 0.96 substitutions per site per million years (s/s/Myr), which has been calibrated for the HVRI of Adélie Penguins (*Pygoscelis adeliae*) using ancient DNA technology. (Lambert *et al.* 2002; Ritchie *et al.* 2004). The 95% highest posterior density (HPD) values of 0.53 to 1.43 s/s/Myr, obtained for the calculation of the mutation rate by

Lambert *et al.* (2002), were used to calculate a confidence interval for the estimates. Generation time was considered to be one year, which is the age at first breeding of *E. gouldiae* in the wild.

Results

Genetic Diversity

The 208 bp section of HVRI of the control region from a total of 157 samples analysed revealed 13 polymorphic nucleotide sites, and a total of 20 unique haplotypes (Table 2.2). Nine haplotypes were found in single birds, while 10 others were found in at least two different regions. Haplotype diversity values ranged from 0.781 for contemporary samples to 0.839 in museum samples. Nucleotide diversity values ranged from 0.007 to 0.01 in regional groupings. All individuals from the Mareeba population share a single haplotype, and therefore have nil diversities (Table 2.3).

Differences in haplotype diversities were non significant between contemporary and historical samples ($t_{(121)}$=0.95, P=0.344), or between the three regional groupings (KI vs. TE $t_{(112)}$=0.48, P=0.629; KI vs. CYP $t_{(51)}$=0.15, P=0.88; TE vs. CYP $t_{(77)}$=0.10, P=0.924).

Table 2.2 Variable sites and geographic distribution of the 20 mtDNA control region haplotypes (HVRI, 208bp) in *E. gouldiae*. Sequences have been entered in GenBank under accession numbers EF094913 to EF094932. Nucleotide positions are relative to the beginning of the sequence.

	Nucleotide position													Region				
Haplotype	10	32	40	49	52	54	76	85	114	144	156	163	196	Kimberley	Top End	Cape York Peninsula	Mareeba*	Total
HS01	G	G	C	T	A	C	A	A	T	T	T	T	A		1			1
HS02	•	•	•	•	•	T	•	•	•	•	•	•	•	4	5	1		10
HS03	•	•	•	•	•	T	G	•	•	•	•	•	G	2	1			3
HS04	A	•	•	•	•	T	G	•	•	•	•	•	•	5	3			8
HS05	•	•	•	•	•	T	G	•	•	•	•	•	•	19	27	4	34	84
HS06	•	•	•	•	G	T	•	•	•	C	•	•	•	4	11	2		17
HS07	•	•	•	C	•	T	G	•	•	•	•	•	•	2	2			4
HS08	•	•	•	•	G	T	G	•	•	•	•	•	•	4	5			9
HS09	•	•	•	•	•	T	G	•	C	•	•	•	•		1			1
HS10	A	•	•	•	•	T	G	•	C	•	•	•	•	2	3			5
HS11	A	•	•	•	•	T	•	•	•	•	•	•	•		4			4
HS12	•	•	T	•	G	T	•	•	•	C	•	•	•		1			1
HS13	•	•	•	•	•	T	•	•	•	•	•	G	•	1				1
HS14	•	•	•	•	G	T	•	•	•	•	•	•	•	1	1			2
HS15	A	A	•	•	•	T	G	•	•	•	•	•	•			1		1
HS16	A	A	•	•	G	T	G	•	•	•	•	•	•		1			1
HS17	•	•	•	•	•	T	G	•	•	C	•	•	G		1	1		2
HS18	•	•	•	•	•	T	G	•	•	•	C	•	•		1			1
HS19	•	•	T	•	•	T	•	•	•	•	C	•	•		1			1
HS20	•	•	•	•	•	T	G	G	•	•	•	•	•		1			1
													Total	44	70	9	34	157

* Captive population for reintroduction program.

Table 2.3 Haplotype and Nucleotide diversity of mtDNA control region HVRI for different *E. gouldiae* sample groupings.

Region	*N*	Haplotypes present	Haplotype diversity (s. d.)	Nucleotide diversity (s. d.)
All samples *	123	20	0.800 (± 0.031)	0.008 (± 0.001)
Museum samples only	43	15	0.839 (± 0.047)	0.009 (± 0.001)
Contemporary samples only	80	14	0.781 (± 0.039)	0.008 (± 0.001)
Mareeba (captive)	34	1	-	-
Regions:				
Kimberley	44	10	0.786 (± 0.054)	0.007 (± 0.001)
Kimberley historical				
Mornington				
Top End	70	18	0.818 (± 0.038)	0.009 (± 0.001)
Top End historical				
Yinberrie Hills				
Cape York Peninsula	9	5	0.806 (± 0.120)	0.010 (± 0.002)

* Excludes captive population (Mareeba)

Genetic structure

AMOVA analyses showed almost all of the variation in the dataset occurred among individuals within populations, with only 2% explained by variation among regions (Table 2.4). The variance component among populations within regions was negative, which indicates a lack of genetic structure. Such negative values arise in cases where genes from different regions are more related to each other than genes within populations. Likewise, the respective negative fixation index obtained for variation among populations within regions should be interpreted as zero (Long 1986). All three fixation indices were found to be not significantly different from zero (P> 0.05), which indicates all the variance components, including within population variation, are not significant (Table 2.4).

Table 2.4 Analyses of molecular variance, fixation indexes and significance for mtDNA HVRI of *E. gouldiae*.

	Observed partition			
	Variance Component	% of variation	Fixation index	P
Among regions	0.0171	2.05	Φ_{CT}=-0.05236	0.9959
Among populations within regions	-0.0427	-5.13	Φ_{SC}=-0.0308	0.9853
Within populations	0.8587	103.08	Φ_{ST}= 0.0205	0.3495
Total	0.8330			

The median-joining network shows limited reticulation between haplotypes, which are separated by up to six or nucleotide substitutions. Some homoplasy may be present, as is suggested by the presence of loops within the network, which indicate alternative mutation paths between haplotypes. Haplotypes do not show an association with geographic region, given that several haplotypes are present in samples from different regions, and such regions are represented through the whole network (Figure 2.2). All samples from the captive population at Mareeba are represented by the most common haplotype.

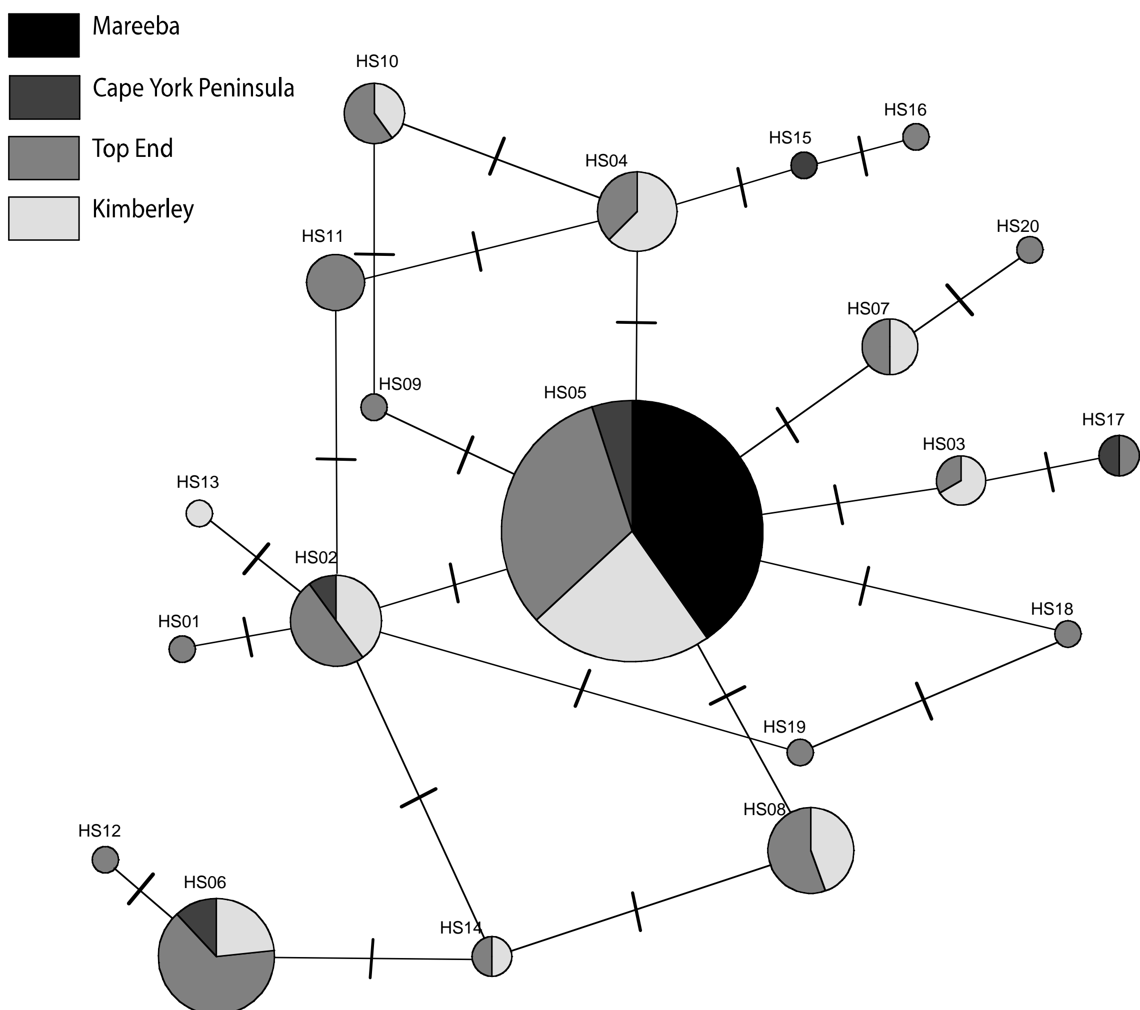


Figure 2.2 Median-joining network of *E. gouldiae* mtDNA control region HVRI haplotypes. Each node represents a different haplotype. Sizes of the nodes are proportional to relative haplotype frequencies. Bars between nodes represent one base pair change each. Pie slices in the nodes represent the relative frequencies of individuals from different geographical regions. Labels correspond to haplotypes listed in Table 2.2

Mismatch distributions and tests of neutrality

Mismatch analyses show a unimodal distribution in all datasets analysed, which is a characteristic signature of population expansion (Fig. 2.3). There were no significant deviations from either distribution simulated under a spatial expansion model (Table 2.5). Mismatch analysis could not be performed for the Mareeba samples, as this population has no genetic variation.

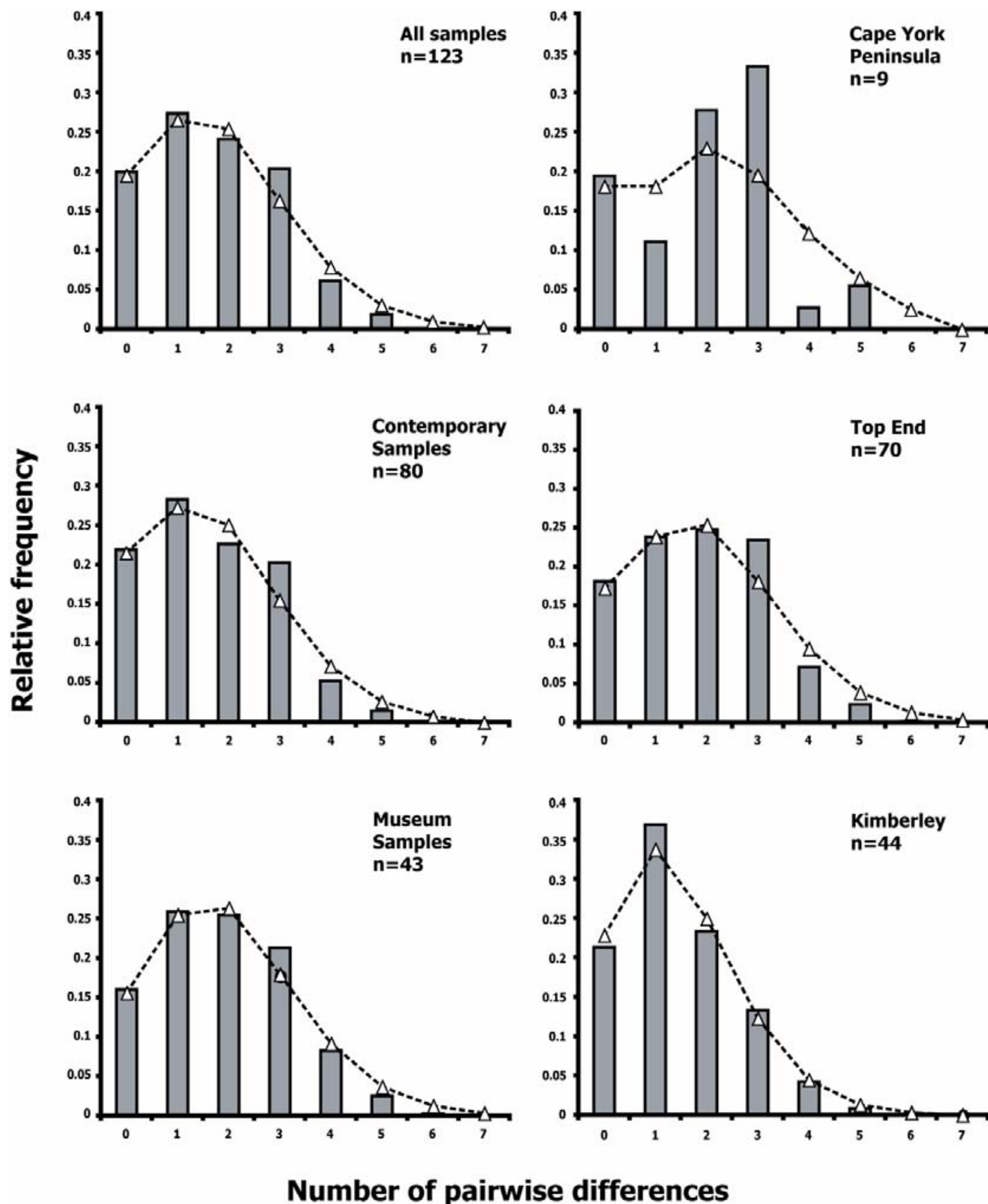


Figure 2.3 Mismatch distributions of *E. gouldiae* mtDNA control region HVRI sequences**.** Bars = observed pairwise differences; dotted line and white triangles = simulated under a spatial expansion model.

Table 2.5 Mismatch distribution parameters and tests of neutrality for mtDNA control region HVRI sequences of *E. gouldiae***.**

	All samples	Museum samples	Contemporary samples	Regions		
				Kimberley	Top End	Cape York Peninsula
Spatial expansion model						
θ	0.046	0.008	0.014	0.001	0.003	0.001
τ	1.865	2.047	1.836	1.472	2.124	2.580
M	16.991	30.340	13.502	99999.0	14.895	8.220
Point estimate of *t* (x1000 ybp)	4.67	5.13	4.66	3.69	5.32	6.46
(CI 95%)[a]	8.459- 3.135	9.284- 3.441	8.327- 3.086	6.676- 2.474	9.634- 3.570	11.702- 4.337
SSD	0.0024 n.s.	0.0016 n.s.	0.0036 n.s.	0.0016 n.s.	0.0042 n.s.	0.036 n.s.
P^{b}	0.606	0.706	0.5475	0.500	0.520	0.377
Raggedness	0.0304 n.s.	0.0323 n.s.	0.0320 n.s.	0.0620 n.s.	0.0329 n.s.	0.135 n.s.
P^{c}	0.802	0.750	0.827	0.368	0.457	0.475
Tests of Neutrality						
Tajima's *D*	-0.7606 n.s.	-0.7688 n.s.	-0.5062 n.s.	-0.5962 n.s.	-0.7158 n.s.	-0.3005 n.s.
Fu and Li's *D**	-0.9677 n.s.	-0.2198 n.s.	-0.6462 n.s.	0.6123 n.s.	-0.8608 n.s.	-0.2853 n.s.
Fu and Li's *F**	-1.0646 n.s.	-0.4688 n.s.	-0.7086 n.s.	0.2672 n.s.	-0.9609 n.s.	-0.3219 n.s.
Fu's F_S	-11.181 *	-8.353 *	-5.760 *	-3.716 *	-10.288 *	-0.787 n.s.

θ = Population size parameter (2*Nu*; where *N* = deme size and *u* = mutation rate per generation)
M = Migration parameter (M= 2*Nm*; where *N* = deme size and *m* = rate of exchange of migrants)
**P*<0.05
[a] Calculated using 95% HPD of 0.53 and 1.43 s/s/Myr (Lambert *et al.* 2002)
[b] *P* (Simulated SSD ≥ Observed SSD)
[c]*P* (Simulated Raggedness ≥ Observed Raggedness)

Calculated average τ values were approximately similar for groupings including all samples, museum specimens and contemporary samples. However, there are differences in τ values when analysing samples from different geographic regions separately, indicating different expansion events for each region (Table 2.5). Mismatch distributions of Kimberley and Top End samples were significantly different to their respective intermatch distribution (Table 2.6), but such difference is not significant between Cape York Peninsula and either of the other two regions (Table 2.6).

Tests of neutrality showed negative Tajima's *D*, and Fu and Li's *D** and *F** values for all sample groupings. Negative values of Fu's F_S were obtained for all samples, which were significant (*P*<0.05) for all groups, with the exception of Cape York Peninsula (Table 2.5).

Table 2.6 Pairwise comparisons between mismatch distributions of different geographic regions.

Mismatch distribution From:	To intermatch distribution with: Kimberley	Top End	Cape York Peninsula
Kimberley	-	<0.0001	0.1872
Top End	<0.0001	-	0.9393
Cape York Peninsula	0.0557	0.9654	-

Estimated times of population expansion

Estimated times indicate that *E. gouldiae* populations in northern Australia expanded across northern Australia approximately between 8400 and 3100 years before present (ybp) (Table 2.5). Estimated times were roughly similar for the whole dataset, compared to museum and contemporary samples analysed independently (t≈4600 to 5200 ybp). On the other hand, differences were greater when different regions were analysed separately. Estimate times since expansion were earliest for CYP (t≈6500 ybp), and more recent for TE (t≈5500 ybp) and KI regions (t≈3700 ybp) (Table 2.5).

Discussion

Results from this study show a lack of geographic subdivision among *E. gouldiae* individuals sampled at different localities, and unimodal mismatch distributions of pairwise nucleotide differences, which suggest a very recent history of expansion in a species that had a wide geographical distribution until recently.

Genetic diversity

Results from this study indicate that present populations of *E. gouldiae* are of relatively recent origin, and have not evolved into genetically different geographic lineages since the putative expansions into some of the areas within the present distribution. The low genetic diversity found in *E. gouldiae*, as well as results from mismatch distribution analyses, suggest that this species may have experienced a severe bottleneck in the past, with more recent population expansions, possibly at times when

habitat changes became more favourable for this species. This relatively recent growth of *E. gouldiae* populations may explain the low divergence between existing haplotypes, despite the spatial separation between the sampling locations in this study.

Nucleotide diversity in all sample groupings is low, reflecting the similarities in sequence composition between all the haplotypes detected. Haplotype diversity is by contrast higher in all cases (except the captive Mareeba population), which suggest evenness in the relative distribution of existing haplotypes within each of the sample groupings analysed. Although there appear to be some differences in haplotype diversity between contemporary and museum samples, as well as between regional sample groupings, they are non-significant in all comparisons. Therefore a loss of haplotypes following the observed demographic changes in past decades can not be proved.

By contrast, the lack of detectable genetic variation in the captive Mareeba population may be resulting from a small number of founders combined with the effects of genetic drift during the following generations.

Population structure

Populations of *E. gouldiae* lack a geographical structure, with identical haplotypes found across their present and former distribution range. Such a pattern of sympatric lineages combined with the low divergence between haplotypes, is typical of species with high gene flow, whose populations have often experienced rapid growth in relatively recent times, and have not been separated by long-term geographical barriers (Avise 2000). Results from the hierarchical AMOVA analyses support the lack of structure, as most of the variation is explained by differences between individuals within populations, with only 2% of the variation explained by differences between regional groupings.

In cases like the present study, for which populations have experienced a recent expansion, structure caused by more recent events (e.g. habitat fragmentation) is not likely to be reflected by patterns of mtDNA variation alone (Moritz 1994b; Spaulding *et al.* 2006). Studies combining variation at mtDNA and at nuclear multilocus data (e.g. microsatellite loci), as well as other more rapidly evolving loci such as genes under selection are often more adequate to address local variation at populations with incomplete mtDNA lineage sorting. Analyses of genetic variation between and within

E. gouldiae populations using such loci in chapters three and four of this thesis also fail to show genetic structuring between sampling areas, further supporting these results.

These results contrast with those found for the sympatric Long-tailed Finch *Poephila acuticauda*, for which divergence times between the two recognised lineages *P. a. acuticauda* and *P. a. hecki*, which are restricted to the Kimberley and Top End regions respectively, has been estimated to be within the Pleistocene, approximately 0.3 million ybp (Jennings & Edwards 2005).

Population expansion and gradual colonisation of northern Australia

Recent expansion of *E. gouldiae* populations is supported by mismatch distribution analyses and tests of neutrality (Table 2.4). I did not find significant differences between observed mismatch distributions and those simulated under the model of spatial expansion. Comparisons of observed distributions with those simulated under a pure demographic expansion were virtually identical (data not shown). Both scenarios of expansion produce a unimodal mismatch distribution and are often difficult to distinguish from each other, in particular in those cases where migration is relatively high (Excoffier 2004). The relatively higher values of the migration parameter ($M>10$) obtained for most of the sample groupings (Table 2.5) explain the similarities between the two models. However, a scenario of pure demographic expansion would not be appropriate to this study, as predictions can only be made about intra-deme diversity with such model. It is possible that *E. gouldiae* populations experienced both kinds of events, with range expansions following favourable habitat changes, and subsequent demographic expansions in local demes.

Results from the different tests of neutrality also support the scenario of population expansion. Results of Tajima's D test suggest a deviation from neutrality. Such deviations can be the product of selection, genetic hitchhiking or population growth. In order to differentiate between such mechanisms, other tests are necessary. Fu and Li's D^* and F^* tests are powerful for detecting background selection, while Fu's F_S is used to detect population growth. Non-significant values of D^* and F^* found in all datasets reject the possibility of selection. Analyses of most datasets also showed significant values of F_S, which support the scenario of population expansion. The exception was Cape York Peninsula, for which the F_S value was not significant (P=0.787).

Relative times since expansion for *E. gouldiae* populations are estimated to be approximately between 9000 and 3100 ybp. Significant differences in the mode of the mismatch distributions calculated between TE and KIM groups (Fig. 2.3; Table 2.6) suggest those populations experienced different expansion events (Excoffier 2004).

Under this scenario, expansion for *E. gouldiae* would have happened gradually, possibly starting in CYP at approximately 6500 ybp, then at 5300 ybp for TE, and at 3700 ybp for KI. The non-significant value of Fu's F_S for CYP can be indicative of the relative stability of populations in that area, which supports the scenario of a relatively older and more stable population that could have potentially been a source of founders for the rest of the regions.

Nonetheless, given the reduced sample size (n=9), results for CYP may not be conclusive. Analyses of a greater number of samples would be desirable to obtain more accurate estimates of expansion times. Unfortunately populations on CYP have been affected most adversely since European colonisation of Australia. Most *E. gouldiae* populations have been extirpated from most of the region, and recent sightings in this area have become uncommon and irregular (O'Malley 2006), making sampling of contemporary specimens virtually impossible. Therefore, museum specimens, although also rare for this region, are the only available source of tissue for molecular genetic analyses. Although the sample size for CYP is limited in this study, a significant difference between expansion times in TE and KI regions exists, which supports the hypothesis of a gradual westward expansion.

The times since expansion reported in this study should be considered relative rather than absolute, as there are important considerations on their calculation using mismatch distribution analyses. The conversion of time in evolutionary units (τ) into time in years (t) depends strongly on parameters that may lack some precision, such as the mutation rate of mtDNA HVRI and the generation time for the species. Traditionally this type of time inference has been calculated using a mutation rate of 0.208 s/s/Myr, which was initially calculated for the whole mitochondria using phylogenetic methods (reviewed in Lambert *et al.* 2002). For this study, I chose to use a rate of 0.96 s/s/Myr, as it has been calculated for the same DNA region used in this study. The methods used for calculating this rate of evolution combine palaeontologic records with phylogenetic information from sub-fossil and extant populations, therefore it should be more robust that a rate calculated by phylogeny of extant species alone.

The generation time in this study was considered to be one year, which is typically the age at first breeding for *E. gouldiae*. However, it is known that other estrildid birds reach maturity at an early age of only a few months. For instance Zebra Finches (*Taeniopygia guttata*) have a generation time in captivity of approximately 90 days (Clayton 1990). During a year with favourable rainfall and abundance of seeds, *E. gouldiae* individuals born at the beginning of the breeding season may reach maturity within months and may be able to reproduce within the season they were born, thus reducing the average generation time. A lower generation time would shift the calculation of expansion to a more recent time. Moreover, since the length of the breeding season is also dependent on rainfall (Tidemann *et al.* 1999), the generation time may also vary between localities, as there is a strong rainfall variation in northern Australia, from coastal to inland areas (Schulze *et al.* 1998). For illustration purposes I have used a generation time of 1 year, which assumes individuals would start breeding only the year after they are born. While time since expansion in this study has been calculated to be within the Holocene period (>10000 ybp) regardless of the sample groupings analysed, the use of a the phylogenetic mutation rate of 0.208 s/s/Myr would have shifted the estimation of expansion time to more than 20000 ybp, which corresponds to the last glacial maximum during the Pleistocene, and therefore would provide a more unrealistic scenario.

Irrespective of the errors on the calculation of absolute time since expansion (*t*), significant differences found in this study between the expansions in evolutionary time (τ) between some of the regional groupings analysed allow hypothesising a gradual expansion of this species in Northern Australia. Such gradual expansion of *E. gouldiae* populations can be hypothesised as consequences of habitat changes caused by climatic events and/or human-related activities during the Holocene. Some climatic, hydrological and socio-cultural events that might have influenced population expansions of *E. gouldiae* are summarised in Figure 2.4.

At approximately 6000 ybp, climate was stable in the lowland tropics in Australia (Shulmeister & Lees 1995; Shulmeister 1999). There is archaeological evidence that humans on Cape York Peninsula experienced population increases during that time. Subsequently, this was a period of intensification in the use of the landscape, through the use of new technologies, and an increase in the use of fire as a tool for landscape management, that continued until European colonisation (Haberle & David 2004). Intensification in the use of fire was aimed at increasing environmental productivity, by

creating a more open vegetation, that would be used by people to harvest large mammals such as kangaroos (Johnson 2006). These changes in vegetation, dominated by eucalypts and grasses would have been favourable for *E. gouldiae*, and may have promoted population expansion of this species.

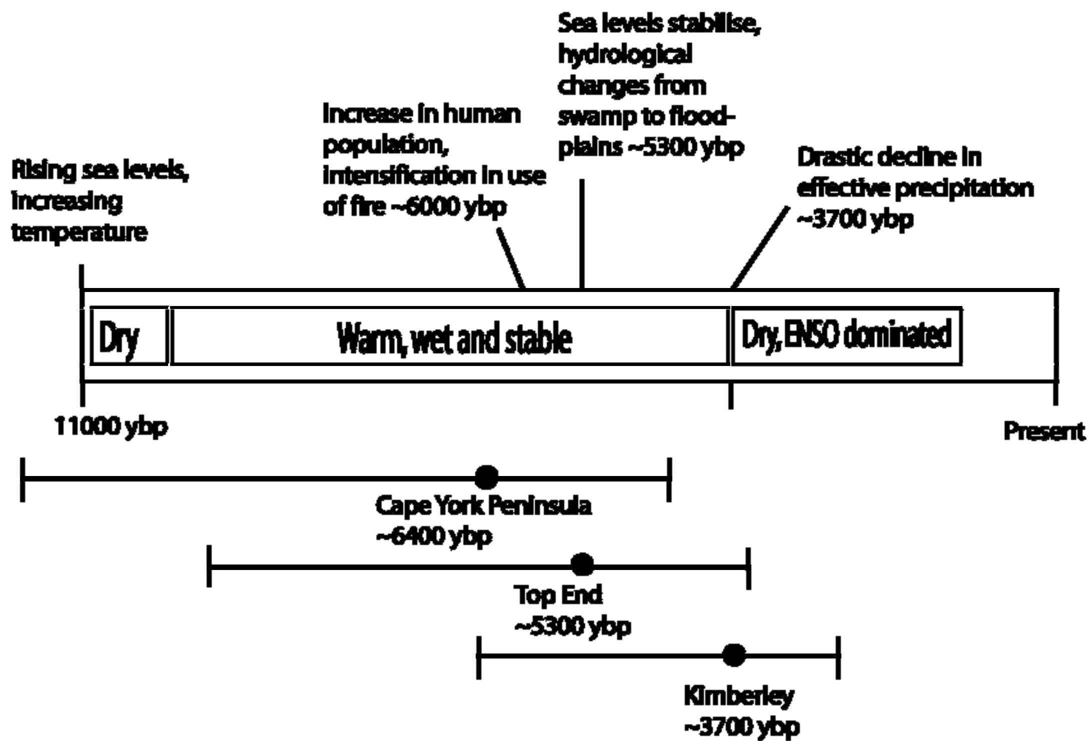


Figure 2.4 Timeline of some climatic and socio-cultural events during the Holocene (above box), and approximate predicted times of *E. gouldiae* population expansions (below box). Black dots represent the time since expansion calculated using a mutation rate for HVRI of 0.96 s/s/Myr, and whiskers represent the range covered when calculating time since expansion using mutation rates of 0.53 to 1.43 s/s/Myr (Lambert *et al.* 2002).

The later colonisation of the Top End, and especially of the Kimberley region, could have been the product of further habitat changes that disrupted the barriers between savannah habitats at different times. Studies of sediment cores in rivers of northern Australia indicate that, as sea levels stabilised, there was a gradual change in the estuarine habitats, from a “big swamp” phase from 6800 to 5300 ybp when mangroves were abundant, to a floodplain phase after 5300 ybp, where mangroves were replaced by freshwater wetland vegetation dominated by grasses and sedges (Woodroffe 1993; Woodroffe *et al.* 1993). A change from mangroves to more open vegetation in estuarine and deltaic plains at the southern tip of the Gulf of Carpentaria could have

posed an easier habitat barrier to overcome for *E. gouldiae* that may have allowed the colonisation of the Top End.

Precipitation in Northern Australia was at a maximum from approximately 5000 to 3700 ybp. After 3700 ybp there was a sharp decline in effective precipitation that created a drier, more variable climate, dominated by increasingly frequent El Niño-Southern Oscillation (ENSO) events that lasted until about 1200 ybp (Shulmeister & Lees 1995; Shulmeister 1999). There are pollen records from Vanderlin Island in the Gulf of Carpentaria that suggest that during that time, vegetation in the periphery of wetlands became more open, with decline of woodland species (Prebble *et al.* 2005). Such change in vegetation, could have facilitated the expansion of *E. gouldiae* into the Kimberley region from the Top End, as riparian forests that separated these regions were replaced my more open woodlands.

Conservation implications

A lack of significant differences between museum and contemporary samples suggests there is no evidence of loss of genetic diversity after the reported recent population declines for *E. gouldiae* since the 1970s. However it is important to consider that diversity in this species is limited, as indicated by the low divergence between existing haplotypes. Low genetic diversity can have consequences for fitness, particularly in small populations (Reed & Frankham 2003).

Captive breeding is a commonly used tool for the conservation of endangered species. It is desirable that the founder population consists of outbred individuals that will represent a good portion of genetic diversity found in the wild (Frankham *et al.* 2002). The lack of diversity found in the captive population in Mareeba may be the result of a small number of founders and the effects of genetic drift. This captive population may likely benefit from the addition of wild individuals to improve genetic diversity, combined with an intense management of the population aimed at preserving and maximising genetic diversity. The fact that wild *E. gouldiae* populations have not diverged into different lineages means that individuals from other regions can potentially be used to repopulate areas of Cape York Peninsula, without the risk of mixing lineages which are too dissimilar from each other. However, care should be taken when translocating individuals from other regions, to minimise the risk of introducing diseases from different areas, to which local birds may be maladapted. This may be important given some evidence of relatively high prevalence of parasites in wild

E. gouldiae compared to other species of savannah estrildid finches. These parasites include air sac mites *Sternostoma tracheacolum* (Tidemann *et al.* 1992b) and Trypanosomes in blood (R. E-S unpublished data).

This study shows an example of the population history and the effects of past habitat changes in a species typical of the monsoonal tropics. Future phylogeographic studies of tropical savannah taxa, with different life histories and habitat requirements will help to illustrate better the effects that past climate fluctuations and human activities have had in the distribution and genetic diversity of extant wildlife species in the dry tropics. Phylogeographic studies are particularly important to reconstruct the population history of species from tropical savannas and other tropical habitats for which there is little or no fossil record, such as most species of birds and invertebrates.

CHAPTER THREE: Microsatellite variation, genetic structure and migration in Gouldian Finch *Erythrura gouldiae* and Long-tailed Finch *Poephila acuticauda* populations.

Abstract

Species that have experienced population reductions are expected to be at higher extinction risk because of loss of genetic diversity. This can have deleterious consequences for population survival and reduce the potential of such populations to respond to environmental change. Such deleterious effects can be reduced by high levels of migration that contribute to the maintenance of genetic diversity. The Gouldian Finch (*Erythrura gouldiae*) has experienced recent population reductions that have resulted in its endangerment. In order to provide effective management aimed to achieve the recovery of an endangered species, it is necessary to investigate the genetic variation in this species.

In this study I investigate genetic variation, population structure and gene flow in Gouldian Finch populations, and contrast such patterns with those of the sympatric and abundant Long-tailed finch (*Poephila acuticauda*). Heterozygosity is slightly higher in *P. acuticauda* than in *E. gouldiae*. There was no evidence of recent severe genetic bottlenecks and the indicators of intra-population inbreeding (f) were relatively low in both species. Inter-population differentiation (θ^{II}) was moderate for both species.

Significantly lower heterozygosity and evidence of a recent population bottleneck was found for a captive population of *E. gouldiae* that is part of a reintroduction program in north-east Australia.

There was evidence of population structuring between *P. acuticauda* populations, while *E. gouldiae* shows no sign of population structure. The lack of genetic differentiation in *E. gouldiae* can be partly explained by relatively higher migration rates compared to *P. acuticauda*.

Introduction

Currently there is an alarming number of species that are classified as threatened. Including approximately 23% of vertebrate species evaluated by the International Union for Conservation of Nature IUCN (IUCN 2007). Habitat loss and fragmentation, often caused by human-related activities, have been identified as major threats for the survival of populations (CBD 2006).

Fragmentation is associated with the reduction of habitat patch sizes and isolation between remaining patches (CBD 2006), which results in reduction and isolation of plant and animal populations. One outcome of population reductions is the reduction of genetic diversity, which is necessary to adapt to environmental change (Freeland 2005). The detrimental effects of population fragmentation can be potentially reduced by migration between patches (Frankham *et al.* 2002). When a population has experienced a bottleneck (i.e. severe reduction in population sizes), it is more susceptible to demographic stochasticity, inbreeding, loss of genetic variation, fixation of deleterious alleles, is less able to adapt to environmental change and is therefore at a greater risk of population extinction (Luikart & Cornuet 1998).

Tropical savannah habitats in northern Australia have been modified, largely as a consequence of different economic activities such as pastoralism, as well as an alteration of the traditional fire regimes (O'Malley 2006). These changes in land management have resulted in the decline of a number of taxa (Franklin *et al.* 2005). Among the groups that have been most affected by these habitat changes are the native granivorous birds. Of the 49 species of granivorous birds of Australian tropical savannas, 12 have experienced population declines, including two taxa considered critically endangered, one endangered species and one presumably extinct (Franklin 1999).

One of the granivorous bird species that has experienced population declines is the Gouldian Finch, *Erythrura gouldiae*, which is endemic to the tropical savannas of northern Australia. Its original distribution included most of the tropical savannah environments north of 20º latitude south, from approximately 17º 57’ S; 122º 14’ E, across the north-west and north-central part of the continent, to Cape York Peninsula, as far east as approximately 20º S; 146º E (Fig. 3.1). The present distribution of the species consists of a greatly reduced and patchy fraction of the original distribution. The decline of Gouldian finch populations has been more noticeable in Cape York Peninsula, where

recent sightings have been infrequent and unpredictable (O'Malley 2006). The reduction in population numbers during the last three decades has been attributed to drastic habitat changes induced by altered fire regimes and cattle grazing activities. Both these activities reduce the availability of seeds, particularly of perennial grass species, which are of critical importance during a "resource bottleneck" period of the year when the more common annual species do not produce seeds (O'Malley 2006). The reduction of *E. gouldiae* populations has resulted in their classification as an endangered species on the global scale (BirdLife International 2006).

E. gouldiae has been a popular species in aviculture. It is a charismatic bird that has attracted attention towards the conservation of Australian tropical savannas. There are a number of efforts taking place for its conservation and recovery, which include efforts to reduce threats in key areas and long-term monitoring programs at some of the largest populations, as well as a captive breeding and reintroduction program at the Mareeba Wetlands (16° 59' S; 145° 25'E) which aims to restore *E. gouldiae* populations in an area where it has been extirpated.

In order to provide effective management for threatened species, it is necessary to obtain information on aspects of population biology, such as genetic diversity, the recent history of population trends and the degree of population differentiation. The use of molecular genetic techniques allows resolution of some of these aspects.

Analyses of mitochondrial DNA (mtDNA) control region showed that *E. gouldiae* lineages are not geographically subdivided, which is a potential consequence of relatively recent population expansion events during the Holocene (Chapter 2). Although evidence of population structure was not found using mtDNA, populations may have evolved separately as isolated populations in relatively more recent times. Evidence of population structure resulting from recent isolation events is more likely to be found when examining other kinds of genetic data such as allele frequencies at variable microsatellite loci (Johnson *et al.* 2003), which marker has been a popular tool for obtaining such fine scale ecological information (Selkoe & Toonen 2006).

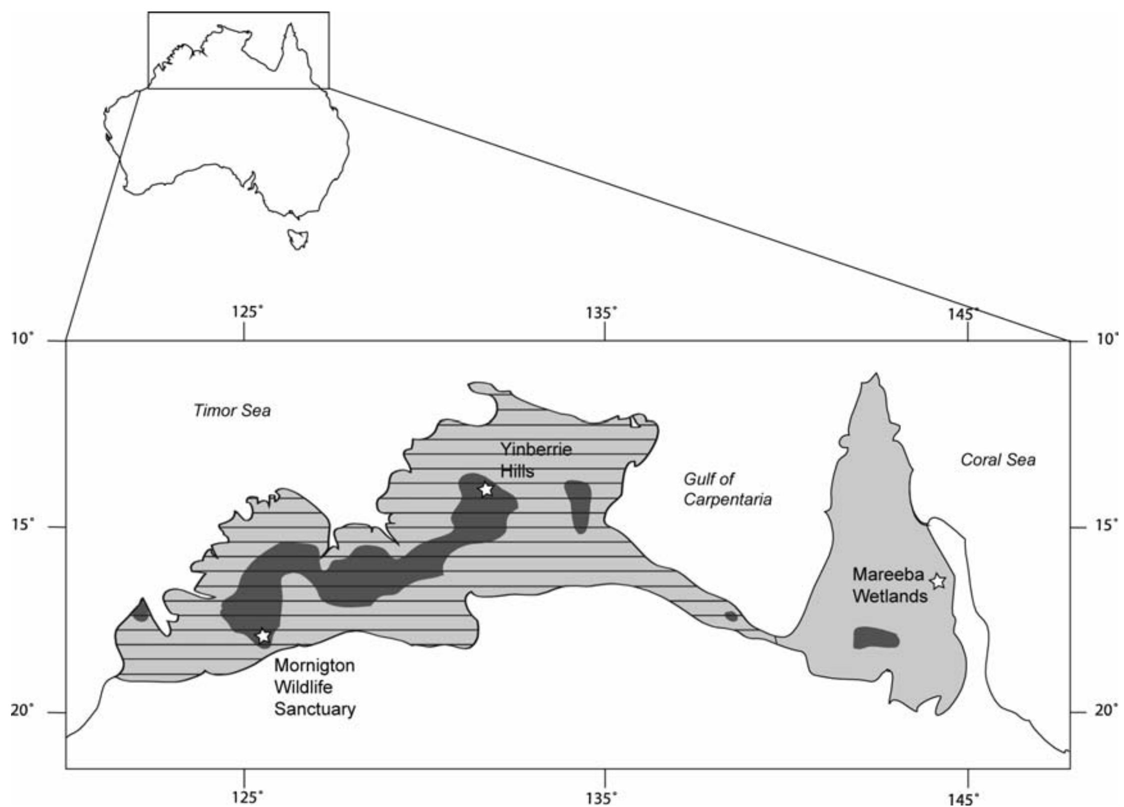


Figure 3.1 Map of northern Australia showing the distribution of *E. gouldiae* and *P. acuticauda***.** Areas shaded in light grey represent the original distribution of *E. gouldiae* while areas shaded in dark grey represent its present distribution. Areas shaded with horizontal lines represent the distribution of *P. acuticauda*. Sampling locations for this study are indicated with stars on the map. Adapted from O'Malley (2006), and records in the Atlas of Australian Birds (Barret *et al.* 2003).

Some other granivorous finch species in the Australian tropical savannas have not been adversely affected by population reductions. One of these species is the Long-Tailed Finch (*Poephila acuticauda*), which remains comparatively abundant. Unlike other granivorous finch species, its populations have not been affected by changes in tropical savannah habitat. *P. acuticauda* has a similar distribution to *E. gouldiae* in north-west and north-central Australia, but its distribution does not include the tropical savannas in Cape York Peninsula in the north-east (Fig. 3.1). Both species have similar habitat requirements, and are frequently found feeding and roosting together in areas where they are sympatric. There are however important life history differences between them. Unlike *E. gouldiae*, *P. acuticauda* is not strictly granivorous. During times when seeds are scarce, *P. acuticauda* have been observed feeding on other abundant resources, such as invertebrates, particularly termites (Dostine & Franklin 2002).

E. gouldiae show three different colour morphs that are not associated with geographical distribution (Southern 1945; Franklin & Dostine 2000), whereas *P. acuticauda* shows dimorphism in different areas of their distribution: *P. acuticauda* individuals from the Kimberley Region in north-west Australia have yellow bills, while birds east of the Kimberley have bright red bills. Some authors have previously classified these colour morphs as separate taxa (Cracraft 1986).

Comparing aspects of population genetics of a threatened species such as *E. gouldiae* with those of an abundant species such as *P. acuticauda* can provide a useful contrast that may show how different population histories of species with different conservation status have shaped the genetic makeup presently observed.

The aim of this study is to use allele frequency data at microsatellite loci to assess genetic variation and look for evidence of recent population bottlenecks in *E. gouldiae* and *P. acuticauda* populations. Moreover, I use allele frequency data to investigate the degree of differentiation and migration rates between two wild populations of both species, respectively. Conservation implications of the results obtained are discussed.

Methods

Sample collection and DNA extraction

Blood samples of *E. gouldiae* and *P. acuticauda* were collected at two different locations in northern Australia that are considered key conservation areas for *E. gouldiae* (O'Malley 2006). Samples from the Yinberrie Hills (14º 06’ S; 132º 04’ E, hereafter “Yinberrie” [*E. gouldiae* n=57; *P. acuticauda* n=30]) were collected from April to July 2003, and samples from Mornington Wildlife Sanctuary (17º 31’ S; 126º 07’ E, hereafter “Mornington” [*E. gouldiae* n=59; *P. acuticauda* n=47] were collected from August to October 2005 (Fig. 3.1). Birds were captured at waterholes using mist nets. A 20 to 75 µl blood sample was taken from the brachial vein of each bird before release. Additional blood samples were obtained from the captive *E. gouldiae* population at the Mareeba Wetlands (hereafter “Mareeba”) collected in 2002 (n=22). Blood samples were stored in either Queens lysis buffer (Seutin *et al.* 1991) or 70% ethanol prior to DNA extraction. DNA was extracted from all samples using a phenol-chloroform protocol (Friesen *et al.* 1997).

Cross-species amplification of microsatellite loci

Cross-species amplification of microsatellite loci (i.e. using PCR primer pairs on a species that have been developed for a different one, often related) is a commonly used option for microsatellite amplification in many species, particularly in taxa such as some marine invertebrates, lepidopterans and birds, for which microsatellite isolation is still comparably difficult (Galbusera *et al.* 2000; Selkoe & Toonen 2006). Cross-species amplification of microsatellite loci tends to be more successful between closely related species (Primmer *et al.* 1996; Galbusera *et al.* 2000; Selkoe & Toonen 2006). Given that microsatellite loci have not yet been developed specifically for *E. gouldiae* or *P. acuticauda*, I attempted cross-species amplification using primers that have been previously developed for related species: 11 loci from the Indigobird (*Vidua chalybeata*) (Sefc *et al.* 2001) and eight loci from the Bengalese Finch (*Lonchura striata* var. *domestica*) (Yodogawa *et al.* 2003). Additionally, I used a set of ten microsatellite loci from other passerine bird species that showed successful cross-species amplification in Zebra Finch (*Taenyiopygia guttata*) (Dawson *et al.* 2005). Loci that have successfully amplified products in *T. guttata* that were used in this study are: *Ase*9, *Ase*12, *Ase*18, *Ase*26, *Ase*32 for Seychelles Warbler *Acrocephalus seychellensis* (Richardson *et al.* 2000) ; *Esc*μ4 for Reed Bunting *Emberizia schoeniclus* (Hanotte *et al.* 1994) ; *LOX*4 for Red Crosbill *Loxia curvirostra* (Piertney *et al.* 1998) ; *Pdo*μ5, *Pdo*μ6 for House Sparrow *Passer domesticus* (Griffith *et al.* 1999) ; and *Ppi*2 for Magpie *Pica pica* (Martinez *et al.* 1999).

PCR Reactions were carried out in volumes of 10 μl containing 1x GoTaq PCR buffer (Promega), 2.5 mM of MgCl2, 0.2 μM of each dNTP, 0.25 μM of each primer, ~40ng of DNA template and 0.25units of GoTaq Flexi DNA polymerase (Promega). Cycling conditions consisted of an initial denaturation of 94ºC for 2 minutes, followed by 25 repeats of 94ºC for 15s, annealing temperature (optimised for each locus, Table 3.1) for 15s and 72º C for 30s. A final extension of 72ºC for 5 minutes was applied at the end of the cycles. All reactions were run on either GenAmp PCR System 9700 (Perkin Elmer) or Palm Cycler (Corbett Research) PCR thermal cyclers.

In order to establish the optimal annealing temperatures for each primer pair, I performed an initial PCR for three samples of each species at each locus, applying a temperature gradient of 45ºC to 67ºC at 2ºC intervals. These products were run in 2% agarose gels, stained with ethidium bromide and visualised under UV light. Optimal

annealing temperatures were determined by comparing the yields of PCR reactions obtained for each sample run with different annealing temperatures.

Polymorphism in loci that successfully amplified PCR products was detected by carrying out PCR reactions with a subset of 16 samples per species and locus using the optimised conditions. PCR products from these samples were run on non-denaturing gels containing 5% acrylamide: bis-acrylamide 19:1 and 0.6x TBE; using a GelScan2000 DNA fragment analysis system (Corbett Research) and stained with ethidium bromide for visualisation.

Primers that amplified polymorphic products were used to genotype all samples, with the addition of fluorescence markers (HEX, FAM or TET) at the 5' end of the forward primer. PCR products were cleaned through a column of Sephadex G-50 (GE Healthcare Life Sciences) prior to running on a MegaBace 1500 capillary electrophoresis system (GE Healthcare Life Sciences), using a MegaBace ET550-R standard (GE Healthcare Life Sciences) as size reference.

Table 3.1 Primer sequences, number of alleles and optimal annealing temperatures for microsatellite loci used in this study**.**

		Erythrura gouldiae n=(116)		*Poephila acuticauda* n=(77)	
Locus	Primer sequence	N° alleles	Ta	N° alleles	Ta
Ase26	F: 5'-GCTGGCCTTGCAAAAAACTTC-3' R: 5'-AACACCTCCCTGTCCCTGC-3'	5	64ºC	8	45ºC
BF05	F: 5'-CTCTTCCCTAGTGCTTTGTC-3' R: 5'-GCAATAAACAACCCCTCTCC-3'	4	49ºC	--	--
BF18	F: 5'-GGTGGTGCGTGGTGAGAGTA-3' R: 5'-TCACCCCGGATTCTAGCACG-3'	9	54ºC	10	49ºC
Escµ4	F: 5'-TTCCCTCACAATTTTCCGAC-3' R: 5'-TATGTGCTGAAGTGAACCATTC-3'	6	47ºC	5	49ºC
Indigo28	F: 5'-CCCAGGAAGTATCCCAGAA-3' R: 5'-CCTCCAATGCTTTAGTGACC-3'	6	50ºC	--	--
Indigo29	F: 5'-TCAGGGAGCAAATCTCTACG-3' R: 5'-GGAAGAAGGCTGGGTAAAAT-3'	4	52ºC	8	51ºC
Indigo30	F: 5'-TGCTTCCTGAGTTCCTATTCT-3' R: 5'-CTGGGTGTTCATAAAGACCT-3'	--	--	4	45ºC
Indigo37	F: 5'-CTGCAATGACACAGACATGG-3' R: 5'-CAAAATCCCCTCGGAGAG-3'	15	48ºC	--	--
Indigo41	F: 5'-GACAGTGTTCAGGAGAAGATAC-3' R: 5'-TTTCCTACAGGATTCCCTAC-3'	23	59ºC	12	59ºC
Pdoµ5	F: 5'-GATGTTGCAGTGACCTCTCTTG-3' R: 5'-GCTGTGTTAATGCTATGAAAATGG-3'	29	55ºC	--	--
Ppi2	F: 5'-CACAGACCATTCGAAGCAGA-3' R: 5'-GCTCCGATGGTGAATGAAGT-3'	29	47ºC	--	--

Ta= Optimal annealing temperature used

Evaluation of microsatellite loci

Alleles for each locus were scored using the program Fragment Profiler (GE Healthcare Life Sciences). Allele scoring errors can cause inaccurate results. When PCR conditions are not ideal or when mutations in the microsatellite flanking region inhibit primer binding, certain alleles in one locus may fail to amplify ("null alleles") or shorter alleles can amplify more efficiently than large ones ("large allele dropout") (Selkoe & Toonen 2006). Other errors include the misinterpretation of artefact bands, such as stutter patterns. I used the program Micro-checker (Van Oosterhout *et al.* 2004) to test the datasets for scoring errors due to stutter, large allele dropout and the presence of null alleles. The algorithms used in the program have the advantage of being able to distinguish between inbreeding and Wahlund effects and Hardy-Weinberg deviations caused by null alleles, as well as being able to detect other genotyping errors such as large allele dominance and some typographic errors. The program GenePop (Raymond & Rousset 1995) was used to detect deviations from Hardy-Weinberg equilibrium and genotypic disequilibrium in loci that passed the null alleles tests.

Analyses of recent population bottlenecks

When a population has experienced a bottleneck event, alleles are lost faster than heterozygosity, resulting in a heterozygosity excess that can be detected with the analysis of allele frequency data (Luikart & Cornuet 1998). Evidence for recent bottleneck events at each population in both species was assessed with a one-tailed Wilcoxon signed-rank test for heterozygosity excess, as implemented in the program BOTTLENECK (Cornuet & Luikart 1996; Piry *et al.* 1999) using 20,000 iterations to calculate probability. The Wilcoxon test was chosen as it is more powerful and robust in the detection of bottlenecks when the number of loci is low (<20) (Piry *et al.* 1999). All analyses were performed using three different models of mutation: the infinite allele model (IAM) in which every mutation results in a new allele, and number of mutations that produce an existing allele is negligible (Kimura & Crow 1964); the stepwise mutation model (SMM) in which mutations occur as steps forwards or backwards with the same probability and therefore mutations towards an existing allele are possible (Ohta & Kimura 1973); and a "Two-phase model" (TPM) (Di Rienzo *et al.* 1994) which is intermediate to the IAM and the SMM, in that alleles can be produced by single step mutations, but some infrequent large jumps in repeat numbers are possible. The IAM and SMM are extreme scenarios within a range of possible mutation models, and most

loci are likely to follow a mutation model in between those extremes (Piry *et al.* 1999). For the TPM, the program was set to values of 70% of SMM.

The heterozygosity excess caused by a bottleneck is temporary, and is expected to be detectable only after 0.2 – 4 N_e generations, where N_e is the bottleneck effective size (Luikart & Cornuet 1998) therefore limiting the detection of a past bottleneck. The ratio of the number of microsatellite alleles to allele size (hereafter *M* ratio) decreases when a population is reduced in size. The effects of population reductions in *M* ratios can be detected up to approximately 125 generations after a severe population reductions and up to approximately 500 generations when population sizes remains small (Garza & Williamson 2001). The program M (Garza & Williamson 2001) was used to calculate *M* ratios in the same sample groupings as those used with BOTLENECK. For this analyses I considered 90% of mutations to be one-step, a mean size of 3.5 for mutations larger than one step and an average mutation rate μ of 5 x 10^{-4}, as recommended by the authors (Garza & Williamson 2001). Ratios were calculated assuming scenarios of pre-bottleneck population sizes of 2,000; 5,000 and 12,500 individuals respectively. Probability for *M* ratios was calculated by simulating 10,000 replicates.

Genetic diversity and genetic structure

Inbreeding coefficients and panmictic (expected) heterozygosities were calculated using a Bayesian approach as implemented in the program HICKORY 1.0.5 (Holsinger & Lewis 2003). The program calculates Hs (panmictic heterozygosity), as well as the parameters f and θ^{II}, which are the Bayesian estimates of Wright's F_{IS} (the inbreeding coefficient due to differences in allele frequencies within populations) and F_{ST} (the inbreeding coefficient due to differences in allele frequencies between populations) respectively. Values of $f > 0$ indicate inbreeding within subpopulations, while values of $\theta^{II} > 0$ indicate population structure (differentiation between populations). Analyses were carried out using a burn-in of 50,000 iterations followed by 250,000 sampling iterations retaining values every 50th iteration for a final sample size of 5000.

The program analyses the data using three models: a "full" model that calculates f and θ^{II} simultaneously, a model that calculates θ^{II} without calculating f (f = 0 model) and a model that calculates f without calculating θ (θ = 0 model). Wild populations (Yinberrie and Mornington) were analysed respectively for *E. gouldiae* and *P. acuticauda* to calculate Hs, f and θ^{II} using the three different models. The θ = 0 model

was used to calculate panmictic heterozygosity and *f* separately for the captive *E. gouldiae* population (Mareeba).

The calculation of *f* and θ^{II} may be affected by the choice of model (Holsinger & Wallace 2004). The Deviance Information Criterion (DIC) (Spiegelhalter *et al.* 2002) calculated by HICKORY was used to choose the model that best fits to the respective dataset. The DIC combines the model fit to the data $\bar{D}$ and the model complexity *pD*. A model that produces a lower DIC value is preferred, as it indicates a better fit to the dataset (Holsinger & Wallace 2004).

Differences between Hs, *f* and θ^{II} values were considered statistically significant in cases where the 95% credible intervals (CI) from the posterior distributions of the groups being compared did not overlap.

Assignment of individuals to populations

Assignment of individuals to different populations was done using the program STRUCTURE 2.1 (Pritchard *et al.* 2000). The program uses a Bayesian clustering approach to assign individuals probabilistically to one or more population based on allele frequency data. The program was run for each species using a population admixture model. I tested prior assumptions of one to six populations (*K*). Each run was set with an initial burn-in of 1,000,000 iterations, followed by 100,000 Markov-chain Monte Carlo (MCMC) samples. Birds from the captive *E. gouldiae* population (Mareeba) were excluded from this analysis.

Migration rates

Recent migration rates between Yinberrie and Mornington populations of E. gouldiae and P. acuticauda were estimated with the Bayesian method implemented in the program BayesAss 1.3 (Wilson & Rannala 2003).The program uses a Bayesian approach to calculate simultaneously asymmetric migration rates between populations. The model does not assume sampling of all subpopulations or Hardy-Weinberg equilibrium within subpopulations. The program was run using a MCMC chain of 1,000,000 burn-in iterations, followed by 2,000,000 steps with sampling every 2,000 iterations. Initial input parameter values of Δp (allele frequency), Δm (migration) and ΔF (inbreeding coefficient) were set at 0.15 for each respectively. Two independent runs were run for the dataset of each species to confirm consistency of results.

Differences between migration rates were considered significant when the 95% confidence intervals from the posterior distribution did not overlap.

Results

Cross-species microsatellite loci amplification

Cross-species amplification of polymorphic microsatellites was successful for ten loci in *E. gouldiae* and for six loci in *P. acuticauda*. Four loci in *E. gouldiae* and one in *P. acuticauda* were eliminated from the datasets, as there was evidence for the presence of null alleles (Table 3.2). The final dataset therefore consisted of six loci for *E. gouldiae* and five for *P. acuticauda*. There were deviations from Hardy-Weinberg equilibrium at some loci in some populations, although deviations were never consistent across all populations of the respective species (Table 3.2).

Table 3.2 Evaluation of microsatellite loci used in this study

	Erythrura gouldiae				*Poephila acuticauda*		
	Hardy-Weinberg exact test (S.E.)			Null allele test	Hardy-Weinberg exact test (S.E.)		Null allele test
Locus	Yinberrie	Mornington	Mareeba		Yinberrie	Mornington	
Ase26	--	--	--	Null	0.0228 (0.0041)	0.2056 (0.0173)	Passed
BF05	--	--	--	Null	--	--	--
BF18	0.9573 (0.0061)	0.4012 (0.0280)	0.9199 (0.0013)	Passed	0.2601 (0.0234)	0.4158 (0.0267)	Passed
Escµ4	--	--	--	Null	--	--	Null
Indigo28	0.4545 (0.0193)	0.0860 (0.0104)	0.9445 (0.0038)	Passed	--	--	--
Indigo29	0.5527 (0.0067)	0.4045 (0.0025)	0.7066 (0.0076)	Passed	0.0845 (0.0178)	0.5082 (0.0343)	Passed
Indigo30	--	--	--	--	0.2337 (0.0091)	0.1892 (0.0087)	Passed
Indigo37	0.7203 (0.0307)	0.7805 (0.0231)	0.2709 (0.0051)	Passed	--	--	--
Indigo41	0.5429 (0.0414)	0.0353 (0.0119)	0.0244 (0.0047)	Passed	0.0459 (0.0107)	0.1923 (0.0306)	Passed
Pdoµ5	--	--	--	Null	--	--	--
Ppi2	0.2360 (0.0359)	0.3134 (0.0424)	0.4859 (0.0178)	Passed	--	--	--
Fisher's test All loci							
χ^2	7.6158	18.0495	12.4581		24.2675	12.8999	
P	0.8144	0.1142	0.4096		0.0069	0.2293	

Bottleneck analyses

Significant values from the Wilcoxon test for heterozygosity excess ($P<0.05$) under the IAM were found for all *E. gouldiae* populations, and the Mornington population of *P. acuticauda* (Table 3.3). Values were non-significant for all populations of both species under the TPM and the SMM. Values for the captive population (Mareeba) calculated using the TPM were nearly significant (P=0.0781).

On the other hand, the *M* ratio tests were significant for population reduction in the Mareeba sample when assuming pre-bottleneck sizes of 2,000 (P=0.0089) and 5,000 (P= 0.0243). The same tests were non significant for the rest of the sample groups analysed.

Table 3.3 Wilcoxon test and *M* ratio tests for recent population bottlenecks in *E. gouldiae* and *P. acuticauda* populations. Significant values ($P<0.05$) are highlighted in bold**.**

		Wilcoxon test *P* values			*M* ratio (all loci)	*M* ratio *P* values		
Species	Population	Model				Pre-bottleneck population size		
		IAM	TPM	SMM		2,000	5,000	12,500
E. gouldiae	Yinberrie	**0.0391**	0.9219	1.0000	0.8013	0.3747	0.5223	0.6774
	Mornington	**0.0391**	0.9219	0.9922	0.7604	0.1983	0.3079	0.4235
	Mareeba	**0.0234**	0.0781	0.4219	0.6034	**0.0089**	**0.0243**	0.0821
P. acuticauda	Yinberrie	0.0781	0.6875	0.9688	0.6954	0.0867	0.1552	0.2714
	Mornington	**0.0156**	0.6875	1.000	0.7161	0.1153	0.1978	0.3227

Genetic variability, genetic differentiation and inbreeding

In *E gouldiae*, values of panmictic heterozygosity (Hs) are not significantly different between Yinberrie and Mornington populations. They are however significantly lower for Mareeba than for either Yinberrie or Mornington (Table 3.4). Heterozygosity was not significantly different between both populations of *P. acuticauda*. Heterozygosity values for the combined wild populations (i.e. Yinberrie + Mornington) are significantly higher for *P. acuticauda* than for *E. gouldiae* (Table 3.4).

Mean estimates of population differentiation (θ^{II}) are moderate for both species, but significantly different from zero (95% CI >0) which indicates moderate population

differentiation in both species. Value of θ^{II} is significantly higher for *P. acuticauda* than for *E. gouldiae*, indicating a stronger population differentiation in the former. Inbreeding coefficients (*f*) are also moderate and significantly different from zero in both species (Table 3.4).

Table 3.4 Summary of population parameters Hs (panmictic heterozygosity), *f* and θ^{II}, and 95% credible intervals (CI) calculated for *E. gouldiae* and *P. acuticauda* populations. Values of Hs correspond to the mean value calculated using the "full model" and values of *f* and θ^{B} are means obtained using the model with the best fit for the respective dataset. Parameters Hs and *f* for Mareeba were calculated using the $\theta = 0$ model. For complete results and 95% credible intervals obtained for all models refer to Appendix I

Species	Population	Hs (95% CI)	*f* (95% CI)	θ^{II} (95% CI)
E. gouldiae	Yinberrie + Mornington	0.7541 (0.7394 – 0.7693)	0.0535 (0.0271 – 0.0853)	0.0800 (0.0663 – 0.0959)
	Yinberrie	0.7608 (0.7405 – 0.7817)		
	Mornington	0.7475 (0.7264 – 0.7686)		
	Mareeba	0.6627 (0.6254 – 0.6965)	0.0847 (0.0308 – 0.1578)	
P. acuticauda	Yinberrie + Mornington	0.7987 (0.7778 – 0.8173)	0.0838 (0.0437 – 0.1297)	0.1244 (0.0983 – 0.1547)
	Yinberrie	0.7948 (0.7626 – 0.8221)		
	Mornington	0.8026 (0.7777 – 0.8250)		

For *E. gouldiae*, the $f = 0$ model had a better fit to the data than the full model. The full and $f = 0$ models produced identical rDIC values for the *P. acuticauda* dataset, however the full model was preferred as it appears to have a slightly better fit to the data (lower $\overline{D}$) the $f = 0$ model (Table 3.5).

Table 3.5 Model fit parameters calculated using the program Hickory. Lowest DIC values obtained for each dataset are highlighted in bold.

Species	Model	$\overline{D}$	*pD*	DIC
E. gouldiae	Full	1756.04	154.30	1907.34
	$f = 0$	1736.74	146.53	**1883.26**
	$\theta^B = 0$	3481.36	0.9364	3482.30
P. acuticauda	Full	1042.82	87.05	**1129.87**
	$f = 0$	1045.48	84.40	**1129.87**
	$\theta^B = 0$	2086.44	0.9102	2087.35

Assignment tests

Assignment tests using STRUCTURE failed to assign *E. gouldiae* individuals to more than one population. A clear pattern was found for *P. acuticauda* with the assumption of two populations (K=2) (Fig. 3.2). The program failed to assign individuals to more than two populations (Appendix II). Figure 3.2 shows some *P. acuticauda* individuals from the Yinberrie population have been assigned largely to the population to which Mornington individuals were mostly assigned (Fig. 3.2 red bars). Separate analyses including each sample location separately failed to assign individuals to more than one population in either species (data not shown).

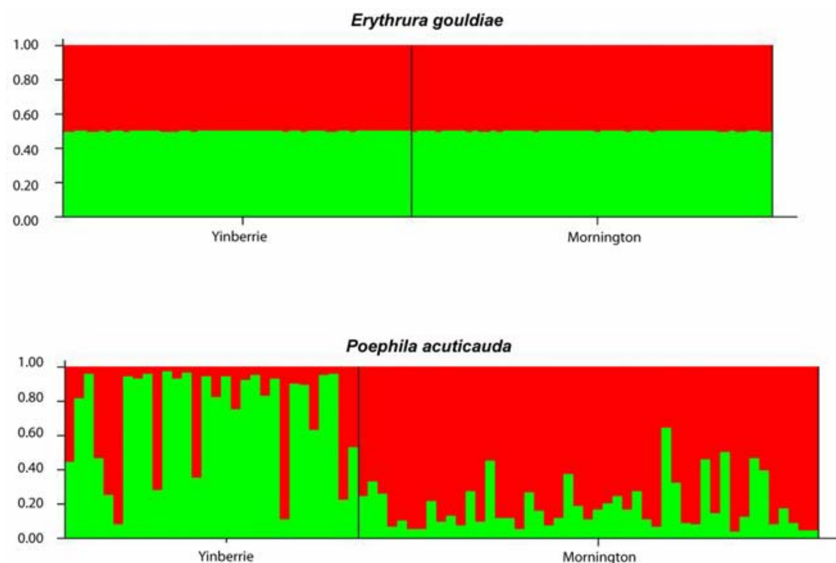


Figure 3.2 Assignment of *E. gouldiae* and *P. acuticauda* individuals to two hypothetical populations**.** Bars represent individual birds. Bar colours represent two hypothetical populations, and the proportion colours in the bars indicate the probability of individuals to be assigned to each population. A black vertical line separates birds from different localities.

Migration rates

Mean posterior values of non-migration rates indicate there is relatively little migration between the two studied populations. However, inter-population migration rates appear to be relatively higher for E. gouldiae (m >0.2) than for P. acuticauda (m <0.1) (Table 3.6). Nevertheless, these differences between species are not statistically significant, as the 95% confidence interval values overlap slightly (Table 3.6). Results from the two independent runs on each dataset produced similar results, which indicate convergence of the MCMC chains. Differences between the direction of migration (Yinberrie to Mornington and vice-versa) are also non significant within each species.

Table 3.6 Migration rates and 95% confidence intervals (CI) for *E. gouldiae* and *P. acuticauda* populations, obtained with the program BayesAss**.** Values correspond to the proportion of individuals in a population that originate from the same population (non-migration rates, and the corresponding proportion of individuals of immigrant origin (inter-population migration).

	Species			
	E. gouldiae	(95% CI)	*P. acuticauda*	(95% CI)
Non-migration				
Yinberrie	0.7453	(0.6685 – 0.8666)	0.9079	(0.8195 – 0.9777)
Mornington	0.7777	(0.6713 – 0.9479)	0.9801	(0.9355 – 0.9994)
Inter-population migration				
Mornington to Yinberrie	0.2547	(0.1334 – 0.3314)	0.0921	(0.0223 – 0.1805)
Yinberrie to Mornington	0.2223	(0.0521 – 0.3287)	0.0199	(0.0005 – 0.0645)

Discussion

Patterns of genetic diversity

Higher heterozygosity estimates for *E. gouldiae* wild populations than in Mareeba and are consistent with results obtained using mitochondrial DNA (Chapter 2). Within-population inbreeding coefficients appear to be moderate for wild populations of both species, as well as for the captive Mareeba population of *E. gouldiae*, which indicates a lack of severe effects of inbreeding due to small population sizes. Heterozygosity

differences between *P. acuticauda* and *E. gouldiae* might reflect the differences in recent population histories of both species.

Recent population bottlenecks

Results of tests for heterozygosity excess showed evidence of recent population bottlenecks when assuming all loci fit the IAM. However, most microsatellites are thought to evolve in a stepwise manner, closer to the assumption of the SMM. The more relaxed TPM is thought to be more appropriate for microsatellites than the SMM (Di Rienzo *et al.* 1994), therefore being more realistic than either the IAM or the SMM. Although there is no significant evidence of recent population bottlenecks under the SMM or the TPM, values for the captive population in Mareeba are close to significant under the TPM.

The results from the heterozygosity excess analyses contrast with those of the *M* ratio differences, under which evidence of a recent population bottleneck in the Mareeba population is clearer. These differences are indicative of the relative timing since the population bottleneck occurred for this population. Heterozygosity excess after a population bottleneck can be detected only after a reduced number of generations (less than $4N_e$ generations) after which populations reach a new equilibrium between mutation and drift that can mask the evidence of a severe bottleneck (Luikart & Cornuet 1998). On the other hand, under some circumstances heterozygosity excess in alleles evolving under the SMM might not be detectable (Cornuet *et al.* 1999; Piry *et al.* 1999). In contrast, reductions of the *M* ratio caused by a population bottleneck are detectable after up to approximately 125generations after the time of population reduction (Garza & Williamson 2001). The significant results under scenarios of pre-bottleneck population sizes of 2,000 and 5,000 can be considered realistic given the current total population estimates of more than 2,500 but less than 10,000 adult *E. gouldiae* (O'Malley 2006).

The Mareeba population of *E. gouldiae* was established around 1982 from a stock of wild individuals (Nevard *et al.* 2002). If the number of founder individuals was low, the number of generations before sampling has been potentially large enough to mask the signature of a population bottleneck in allele frequencies. The nearly significant trend in the heterozygosity excess bottleneck tests for Mareeba obtained with the IAM and the TPM, and the significant results from the *M* ratio test might be worth of

consideration for the genetic management of present and future captive populations to be used for reintroduction.

Immigration can also have an effect on the duration and severity of a population bottleneck (Keller *et al.* 2001). If bottleneck events have recently happened in wild *E. gouldiae* populations, their negative effects could be reduced in a few generations as a consequence of high migration rates between populations.

Population structure and migration

Estimated values of θ^{II} show a moderate population structuring in both species. Such structure appears to be more significantly higher for *P. acuticauda*. These differences between both species are also supported by the results obtained with the assignment method implemented in STRUCTURE.

Although *P. acuticauda* populations in different regions present different colour phenotypes, past admixture between populations can not be excluded, as shown by the assignment of Yinberrie individuals to the Mornington population. In contrast, there is no obvious pattern in the assignment of *E. gouldiae* to different populations (Fig. 3.2). This lack of population structuring is consistent with results obtained using mtDNA (Chapter 2).

The differences in population structuring can be explained by differences in migration rates between both species. The apparent higher migration rates between *E. gouldiae* populations may reduce the effects of population isolation, while in *P. acuticauda* migration may be small enough to maintain genetic identity between populations.

Management implications

Results from this study indicate that the two wild *E. gouldiae* populations studied do not suffer from severe inbreeding or recent population bottlenecks, despite the recorded population reductions. However the interpretation of these results might not be applicable to other *E. gouldiae* populations. Both Mornington and Yinberrie are relatively large populations (estimate sizes of 100-250 adults in each), in which conservation measures are currently taking place (O'Malley 2006). Other locations with smaller number of individuals are potentially at a higher risk of inbreeding depression if their populations are not stable.

Results also show that migration rates of *E. gouldiae* are higher compared to other finch species. This higher migration capability of *E. gouldiae* is a possible adaptation to the unpredictable nature of their food resources in tropical savannas. When the particular seeds on which they feed are not available, they may be capable of flying longer distances than other species in order to find new areas with more resources. The evidence of higher migration rates is consistent with radio-tracking evidence. *E. gouldiae* individuals appear to move daily within an approximately 10km radius between nesting and feeding areas, both at Yinberrie Hills (Milton Lewis, pers. com.) and Mornington Wildlife Sanctuary (Sarah Legge & Steve Murphy, pers. com) locations. Movements are more pronounced at the "resource bottleneck" period during the wet season, when seeds become more difficult to obtain and individuals are forced to fly longer distances to feed.

Yinberrie and Mornington are approximately 750 km apart; therefore migration rates between these localities may not be a consequence of first generation migration, but rather of several generations of migrants between neighbouring demes in a stepping-stone manner. Although *E. gouldiae* seem to be capable of relatively long-distance movements, this capability may be impeded by the lack of suitable habitat with available food and water resources between suitable locations. A network of suitable habitat areas between key locations for *E. gouldiae* would be ideal. Current fire management practices that are being implemented in some of the protected areas could be extended to other locations to promote a wider availability of resources, and would permit the natural migration pattern of this species.

Yinberrie is the only population of *E. gouldiae* for which there are available counts for the last ten years. *E. gouldiae* populations in that locality have been fluctuating markedly since 1997 (O'Malley 2006). Population fluctuations in this locality can be the effect of different habitat conditions throughout the years. In the other hand, if Yinberrie is a destination locality for *E. gouldiae* migrants, rather than a source, the observed fluctuations might be the effect of fluctuations in the number of migrants from surrounding localities with more stable habitat conditions.

The inclusion of more of other *E. gouldiae* populations in future population genetic analyses will provide a more complete picture of patterns of genetic diversity and migration. Knowledge of between Yinberrie and other locations in north-central Australia would help identifying the relative importance of different localities as

sources or destination of migrants, which would allow taking more efficient management decisions at the regional and landscape levels.

The captive population at Mareeba has lower genetic diversity than the two studied wild populations, which is consistent with results obtained from mitochondrial DNA (Chapter 2). Moreover there is some evidence of a potential population bottleneck for this population. The introduction of wild individuals in this population would potentially increase its genetic variability, and therefore would increase their survival potential. The lack of population structure in *E. gouldiae* means that, in the absence of suitable populations as a source of individuals from Cape York Peninsula, introduction of wild individuals from other localities wild populations is potentially feasible, as the risk of outbreeding depression by mixing populations that are genetically too different is less likely.

CHAPTER FOUR: Gene conversion and natural selection in MHC class IIβ genes of Gouldian Finch *Erythrura gouldiae* and Long-tailed finch *Poephila acuticauda*

Abstract

Genes of the Major Histocompatibility Complex (MHC) have an important role in pathogen recognition and the initiation of the immune response. MHC genes are among the most variable genes in vertebrates. Their high diversity is thought to be maintained by gene duplication, balancing selection and gene conversion. I investigated the variability of MHC class IIβ genes of the endangered Gouldian Finch (*Erythrura gouldiae*) and the sympatric and relatively abundant Long-tailed Finch (*Poephila acuticauda*). I looked for evidence for the presence of gene conversion events between haplotypes, in order to determine its relative importance on the maintenance of genetic variability in these two species, and calculated d_N/d_S (the ratio of non-synonymous to synonymous substitution rates) as an indication of the relative strength of selection at this gene.

Erythrura gouldiae had an apparently lower variability at MHC class II genes than *P. acuticauda*. Gene conversion events between haplotypes were detected in both species, being more prevalent in *E. gouldiae* than in *P. acuticauda*. This indicates that gene conversion has been relatively more important in the maintenance of genetic variability of MHC genes in *E. gouldiae*. Evidence of positive selection at MHC class II genes was found for both species. The patterns of positive selection at the amino-acid level appear roughly similar between both species, although in some cases I found positions that are under purging selection in *E. gouldiae*, while in *P. acuticauda* are under strong positive selection.

Differences found between both species in variability of MHC genes, the prevalence of gene conversion events and the rates of positive selection may reflect possible differences in evolutionary pressures caused by pathogens in each species.

Introduction

Genetic variability represents the evolutionary potential of populations. Its maintenance is important to allow populations to adapt to environmental changes. Genetic variability increases with mutation or migration and is lost by directional selection and drift. Selection tends to be more effective in large populations, while the effects of drift are stronger in small populations (Frankham *et al.* 2002).

Reduced genetic variability in populations can reduce their ability to adapt to habitat changes. Knowledge of genetic variability is therefore important in the management and recovery of small and threatened populations (Frankham *et al.* 2002; Freeland 2005). While genetic variability is often studied using neutral genetic markers such as non-coding DNA sequences, allele frequencies at neutral loci and single nucleotide polymorphisms, knowledge of variation at molecular markers of adaptive importance is necessary to understand adaptive processes within and between populations (Sommer 2005).

The study of variation at genes of the Major Histocompatibility Complex (MHC) has been increasingly used to answer questions in the evolutionary ecology of vertebrates (Bernatchez & Landry 2003; Sommer 2005). MHC genes are ideal candidates for the study of adaptations in natural populations, as they are highly polymorphic, the molecular processes on which they are involved have been intensely studied, and details about the structure and function of their products are available (Bernatchez & Landry 2003).

MHC genes are members of the immunoglobulin gene superfamily. They have a key role in the adaptive immune response of vertebrates (Edwards & Hedrick 1998). MHC molecules bind foreign peptides and present them to the cell surface. Cells of the immune system recognise the peptides presented by MHC molecules, and are then programmed to kill the identified pathogens, or to aid in the proliferation of cells that will produce the antibodies to neutralise them (Edwards & Hedrick 1998). Class I and Class II MHC genes are known to be highly polymorphic. MHC Class I genes are expressed in all nucleated cells, and bind endogenous peptides, such as those of viral origin. MHC Class II genes are expressed only on antigen-presenting cells such as macrophages and B-cells, and bind exogenous peptides such as those originated from extracellular parasites (Edwards & Hedrick 1998). The α and β domains of the MHC

class II molecule are encoded by separate genes and together form the peptide-binding region (PBR) in which foreign peptides are bound, and later recognised by T-cells.

The unusual polymorphism in MHC genes is thought to be maintained by gene duplication, gene conversion, and balancing selection (Ohta 1991). MHC genes are typically under strong positive selection, particularly around the PBR (Hughes & Nei 1988). There are two theories of balancing selection that have attempted to explain the maintenance of high polymorphism of MHC genes. Under the "heterozygote advantage" hypothesis, individuals with a higher number of different alleles can recognise a wider range of pathogens than homozygous individuals, and under the "frequency-dependent selection" theory, a balanced polymorphism results from an arms race between the host and pathogen (Takahata & Nei 1990). There are a number of studies supporting both hypotheses, most of which are derived from experimental studies in laboratory conditions (Sommer 2005).

MHC variation is important for long-term survival and fitness of small and fragmented populations (Sommer 2005). There is evidence for the maintenance of MHC diversity despite low levels of variation at neutral loci in different taxa (Aguilar *et al.* 2004; Jarvi *et al.* 2004; Schad *et al.* 2004). These contrasting differences between MHC and neutral loci have been attributed to balancing selection. Although selection can explain the high diversity of MHC genes (Bernatchez & Landry 2003), gene conversion (a non-reciprocal form of recombination, between alleles and between loci (Slatkin 1986)) can also have a key role in the maintenance of the observed diversity in MHC genes. Gene conversion and recombination allow the creation of new alleles by shuffling genetic material between existing alleles (Ohta 1991; Miller & Lambert 2004; Schaschl *et al.* 2005; Bos & Waldman 2006; Schaschl *et al.* 2006).

The Gouldian Finch *Erythrura gouldiae* is an endangered species endemic to the tropical savannas in northern Australia. This species has experienced continuous population declines in the last three decades, presumably as a consequence of changes in the landscape of tropical savannas, caused to a great extent by a gradual intensification of cattle grazing and altered fire regimes across northern Australia (O'Malley 2006). The distribution of *E. gouldiae* formerly included the tropical savannas of northern Australia north of 20° latitude south, from approximately 17º 57' S; 122º 14' E, across the Kimberley plateau and Arnhem land, to Cape York Peninsula as far east as approximately 20º S; 146º E. Declines have been most severe in Cape York Peninsula, where recent records of this species have been particularly uncommon

and unpredictable (O'Malley 2006). Consequently, the present distribution of *E. gouldiae* consists of a reduced and fragmented part of the original distribution.

E. gouldiae populations retain relatively low levels of genetic variability (Chapter 2; Chapter 3), potentially resulting from a historical population bottleneck and relatively recent population expansion events during the Holocene (Chapter 2).

The decline of *E. gouldiae* has raised concerns about the health of Australian tropical savannas. This species is subject to a recovery plan, and several conservation measures, including monitoring of representative populations and a captive breeding program for reintroduction, are taking place (O'Malley 2006).

While *E. gouldiae* populations have experienced recent declines, there are other related finch species in the Australian tropical savannas whose numbers have remained stable. One of these species is the comparatively abundant Long-tailed Finch *Poephila acuticauda*. *Poephila acuticauda* has a similar distribution to *E. gouldiae* across northern Australia, with the exception of Cape York Peninsula, where it is absent (Fig. 4.1). Despite having similar habitat requirements, *E. gouldiae* and *P. acuticauda* differ in their recent population histories, and in their apparent susceptibility to parasites.

Empirical evidence shows marked differences in parasite load between *E. gouldiae* and *P. acuticauda*. A study found that 62% of a sample of wild *E. gouldiae* individuals from two wild populations were infected with the air-sac mite *Sternostoma tracheacolum*, whereas the parasite was not detected in *P. acuticauda* (Tidemann *et al.* 1992b). There is also evidence of differences in the infestation of blood parasites from the genus *Trypanosoma* from a PCR survey. Seventeen percent of *E. gouldiae* individuals screened (n=131) tested positive for *Trypanosoma spp.* infection, while only 0.8% of *P. acuticauda* (n=122) were infected (R. Esparza-Salas, unpublished data). MHC variability may explain these apparent differences, as there is some evidence that variability in MHC genes can confer resistance to infections in other vertebrate species, including ectoparasites (Hedrick *et al.* 2001; Olsson *et al.* 2005; Šimková *et al.* 2006).

In this study I characterise the variation, rates of gene conversion and positive selection of MHC class IIβ genes in *E. gouldiae* and *P. acuticauda*. This comparison aims to determine the degree to which selection and gene flow contribute to maintain adaptive genetic variation in two related species with similar habitat requirements and with important differences in population history and susceptibility to pathogens. Such differences are expected to be reflected in the patterns of diversity and selection at MHC genes.

Methods

Sampling

Blood samples from *E. gouldiae* and *P. acuticauda* were collected at the Yinberrie Hills (14° 06’ S; 132° 04’ E) during April to July 2003, and from Mornington Wildlife Sanctuary (17° 31’ S; 126° 07’ E) during August to October 2005. A volume of 30 to 75 µl of blood was obtained from the brachial vein of each individual before release. Blood samples were stored either in Queens lysis buffer (Seutin *et al.* 1991) or 70% alcohol prior to DNA extraction. For this study I used a total of 14 *E. gouldiae* samples (6 from Yinberrie Hills and 8 from Mornington Wildlife Sanctuary) and 14 *P. acuticauda* samples (7 from Yinberrie Hills and 7 from Mornington Wildlife Sanctuary). DNA was extracted from all samples using a phenol-chloroform protocol (Friesen *et al.* 1997).

PCR + 1 cloning of MHC class IIβ genes

Polymerase Chain Reactions (PCR) involving multi-gene systems such as MHC can produce artefacts when the sequence of one partially amplified allele is repaired with a different template on subsequent PCR cycles (Borriello & Krauter 1990). This phenomenon can result in an artificial “in vitro” recombination between alleles, which can potentially lead to overestimations of allele numbers, and may produce misleading results when studying positive selection. In order to avoid in vitro recombination and sequencing artefactually recombinant haplotypes, I used the PCR + 1 method (Borriello & Krauter 1990) to clone and identify non-recombinant MHC class II haplotypes. The PCR + 1 technique consists of a reaction in two steps. For the first step, one of the PCR primers is added to the reaction mix in excess with respect to its complementary primer, resulting in an excess of single stranded non-recombinant fragments that lack the complementary strand. On the second step, the PCR products from the first reaction are re-amplified in fresh PCR master mix that includes only the complementary primer to the one that has added in excess in the first step, and to which a restriction site has been added. The primers with the restriction site anneal to the non-recombinant single strands and their respective complementary strands are formed. Identification of non-recombinants in the final reaction is achieved by cloning and subsequent digestion of the amplified cloned sequences with the respective restriction endonuclease.

As an additional measure to avoid the in vitro formation of recombinant haplotypes, the times of the thermal cycling profiles used for PCR reactions were modified by increasing the time in the extension steps, which has been demonstrated to suppress PCR-mediated recombination (Judo *et al.* 1998).

Primers HOPE1 (5'-GAA AGC TCG AGT GTC ACT TCA CGA ACG GC-3) and HOPE2 (5'-GGG TGA CAA TCC GGT AGT TGT GCC GGC AG-3') (Vincek *et al.* 1997) were paired to amplify a 173bp section of the second exon of MHC class IIβ gene of *E. gouldiae* and *P. acuticauda* individuals. The amplified exon includes part of the PBR of the gene.

The initial PCR reaction was carried out in 50 µl total volume including 1x GoTaq Flexi PCR buffer (Promega), 1.5 mM $MgCl_2$, 5% Dimethyl sulfoxide (DMSO) 0.2 mM of each dNTP, 4 pmoles of primer HOPE1, 40 pmoles of primer HOPE2, 1.25 units of GoTaq Flexi DNA polymerase (Promega) and approximately 40ng of template DNA.

Cycling conditions included an initial denaturation at 94ºC for 3 minutes; followed by 14 touchdown cycles of 92 °C for 15 seconds, annealing at 65 °C to 58 °C for 20 seconds, with 0.5 °C decrement at each cycle and 74 °C for 3 minutes. Twenty cycles of 92 °C for 15 seconds; 50 °C for 30 seconds and 74 °C for 3 minutes followed the touchdown cycle. A final extension at 72 °C for 60 minutes was applied at the end of the cycles.

The second reaction was performed in 50 µl containing 1x GoTaq Flexi PCR buffer (Promega), 1.5 mM $MgCl_2$, 5% Dimethyl sulfoxide (DMSO) 0.2 mM of each dNTP, 40 pmoles of primer *Hind*III-HOPE1 that included a *Hind*III restriction sequence at the 5' end (5'-<u>AAG CTT</u> GAA AGC TCG AGT GTC ACT TCA CGA ACG GC-3'), 1.25 units of GoTaq Flexi DNA polymerase (Promega) and 20 µl of PCR product from the initial PCR reaction.

The temperature profile for the second reaction consisted of a denaturation step at 94 °C for 4 minutes, followed by 50 °C for 2 minutes and 74 °C for 13 minutes.

Products obtained with the second PCR reaction were cloned on a pGem-T easy vector (Promega) following instructions from the manufacturer. For each individual, bacterial colonies that included the MHC insert were identified with blue-white screening, collected and resuspended in 500 µl of nuclease-free water.

Clones were re-amplified in total volumes of 10 µl containing 1x GoTaq Green Master Mix (Promega); 5 pmoles of each of the vector primers SP6 (5'-GAT TTA GGT GAC ACT ATA G-3'), and T7 (5'-GTA ATA CGA CTC ACT ATA GGG C-3'); and

2 μl of the solution containing the bacteria colonies. Cycling for PCR reactions of cloned alleles consisted of an initial denaturation at 94 °C for three minutes, followed by 25 cycles of 94 °C for 15 seconds; 55 °C for 30 seconds and 72 °C for one minute; and a final extension of 72 °C for 10 minutes.

1 μl of the PCR reaction of each clone was digested overnight with 2 units of *Hind*III restriction endonuclease (New England Biolabs), in total volumes of 10 μl and in the presence of 1X NE buffer 2 (New England Biolabs). The digestion reactions (10 μl) were run on 2% agarose gels. Clones that presented bands of approximately 85bp or 105bp after digestion (i.e. the length of the vector sequence from the primers to the restriction site at the insert, depending on the direction in which the PCR product was inserted in the vector) were selected for subsequent mutation detection analyses.

SSCP gel electrophoresis

Analysis of single stranded conformation polymorphism (SSCP) is a widely used technique for the detection of mutation in fragments of DNA of the same length (Sunnucks *et al.* 2000). Denatured fragments of DNA are run through non-denaturing polyacrylamide gels. As single strands travel through the gels, they fold differently into structures according to their nucleotide sequences. Different strands assume different shapes and travel through the gels with different mobility.

A total of 48 clones per individual, previously identified as non-recombinant with the PCR + 1 technique, were screened using SSCP in order to identify representative alleles in each individual for sequencing. Cloned PCR products were re-amplified using forward and reverse primers labelled with fluorescence, in order to facilitate visualisation in SSCP gels. Total volumes of 10 μl were used, containing 1x GoTaq Flexi buffer (Promega), 1.5 mM $MgCl_2$, 2% Dimethyl sulfoxide (DMSO) 0.2 mM of each dNTP, 5 pmoles of each primer HOPE1 and HOPE2 (labelled with the fluorescent marker HEX at the 5' end), 0.25 units of GoTaq Flexi DNA polymerase (Promega) and 2 μl of PCR product. These reactions were run with an initial denaturation of 94 °C for 3 minutes, followed by five cycles of 94 °C for 30 seconds; 58 °C for 1 minute and 72 °C for 1 minute.

The resulting PCR products were diluted in a 1:3 ratio with formamide containing bromophenol blue for visualisation. This mixture was heated for three minutes at 95 °C, snap-cooled on ice slurry for ten minutes and run through denaturing gels containing 4% Acrylamide: bis-acrylamide 49:1; 2% Glycerol and 0.6x TBE. Gels were run at

1200V and 22 °C for 40 minutes on a GelScan2000 DNA fragment analysis system (Corbett Research).

Representative alleles were identified by the unique pattern of two bands produced by clones with different conformation on SSCP gels. Clones from each representative allele were sequenced. In order to test the accuracy of the technique, for each individual I sequenced two randomly chosen pairs of clones with identical band patterns.

Sequence analyses

All sequences were searched for running errors and aligned using the program SEQUENCHER 4.2 (Gene Codes Corporation). Separate alignments were made for *E. gouldiae* and *P. acuticauda* haplotypes. All sequences were cut to the nearest codon position at 171bp for subsequent codon-based analyses. Nucleotide and haplotype diversity were calculated using the program DnaSp 4.10.9 (Rozas *et al.* 2003).

For each species, a hierarchical Analysis of Molecular Variance (AMOVA) was performed to determine the relative contribution of variance components: among populations, among individuals within populations, and among alleles within individuals. Distance matrices for these analyses were calculated using pairwise differences between sequences. Significant deviations from zero of the fixation indices were calculated by bootstrapping 10,000 times.

Different tests of neutrality were performed on the datasets of each species using DnaSp 4.01 (Rozas *et al.* 2003), to determine the effects of demography and selection on deviations from neutrality. Tajima's D test (Tajima 1989) was calculated to detect deviations from neutrality. Fu's F_S test (Fu 1997) was calculated as an indication of population growth. Fu and Li's D^* and F^* tests (Fu & Li 1993) were carried out to detect background selection. All neutrality tests were performed using the total number of segregating sites in each dataset.

Phylogeny reconstruction

A phylogenetic tree was constructed with the identified MHC class IIβ alleles found in each species, in order to illustrate their genealogies. Alignments were divided into first, second and third codon positions using the program CODONSPLIT (Ingrid Jakobsen, University of Queensland). A χ^2 stationarity test was performed on every codon partition using the program TREEPUZZLE 5.2 (Schmidt *et al.* 2002) to detect

potential variation in sequence compositions among each datasets. After passing stationarity tests, the best model of nucleotide substitution for each partition was found using the program PAUP* version 4.0b10 (Swofford 1998) with the add-on MrMODELTEST 2.2 (Nylander 2004). Phylogenies were reconstructed with a Bayesian method using the program MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003) running two parallel analyses for 4,000,000 Markov Chain Monte Carlo (MCMC) generations, and sampling every 100 generations. One cold chain and three heated chains were used for each analysis. Consensus trees were constructed using the final 2,000 trees of both parallel runs. Phylogenetic trees were drawn using the program TREEVIEW 1.6.6 (Page 1996).

Detection of gene conversion

The program GENECONV 1.81 (Sawyer 1989; 1999) was used to detect gene conversion events between MHC class IIβ haplotypes in *E. gouldiae* and *P. acuticauda*. The program searches for stretches of nucleotides that are more similar to each other than expected by chance. The option that includes only silent (i.e. synonymous) sites of coding sequences was used in the analyses, as the effects of strong selection on amino-acid polymorphisms may cause the significance of detected fragments to be overestimated (Sawyer 1999). Sequences from each species were analysed separately in order to detect differences in gene conversion patterns. Gene conversion events were considered significant when simulated *P* values were equal or less than 0.05 after 10,000 iterations.

Detection of positive selection

The degree of natural selection can be calculated as d_N/d_S ratio, which is the ratio of non-synonymous substitutions (i.e. those that produce a change in the amino-acid sequence of the protein), to synonymous substitutions (i.e. those that do not produce a change in the amino-acid sequence of the protein) between gene-coding nucleotide sequences. Non-synonymous substitutions are expected to be under the effect of selection, while synonymous substitutions are essentially neutral (Bernatchez & Landry 2003). Thus, ratios of $d_N/d_S > 1$ indicate positive selection, $d_N/d_S = 1$ indicates neutral evolution and $d_N/d_S < 1$ indicates purging selection.

I used the program OmegaMap 0.5 (Wilson & McVean 2006) to calculate d_N/d_S ratios in MHC class II genes of *E. gouldiae* and *P. acuticauda*. The program uses a

population genetics approximation to the coalescent, and Bayesian inference to calculate selection as d_N/d_S ratios (ω), in the presence of recombination. Prior parameters included exponential distributions of μ (synonymous transversion), κ (synonymous transition), ω (selection), ρ (recombination rate) and φ (rate of insertions/deletions) set at 3.6, 3.0, 1.0, 0.2 and 0.1 respectively. Codon usage to be used as input for the selection analyses was calculated for each dataset using the program GCUA (McInerney 1998). After calculation, some codons may have a usage of zero in the datasets. Setting the frequency of some codons to be zero in omegaMap implies making an assumption that such codons have never been present in the history of the sequences and can never been present. In order to overcome this assumption, I corrected codon usage values by adopting a semi-Bayesian approach with the equation (1+(n/S))/(61+L), where n is the number of times the codon is observed, S is the number of sequences and L the number of codons per sequence. After this correction the minimum value is 1/(61+L) instead of zero (Daniel Wilson, pers. com.).

Global d_N/d_S ratios (across the whole sequence) were calculated using a model that assumes a constant ω for the complete sequence. Amino-acid site specific d_N/d_S ratios were calculated using a model that assumes independent ω for each codon. Amino-acid sites were considered to be under positive selection when d_N/d_S values were greater than 1, and their respective posterior probabilities for selection were equal or greater than 0.95. Sampling parameters for all omegaMap analyses included a total of 500,000 Markov Chain Monte Carlo (MCMC) steps, with a burn-in of 5,000 steps and thinning of 10 for a final sample size of 4500 iterations.

Differences between d_N/d_S ratios of *E. gouldiae* and *P. acuticauda* were tested, using the posterior distribution comparison tool implemented in the software HICKORY 1.0.5 (Holsinger & Lewis 2003). The program compares pairwise differences between the two independent runs at every sampling step, and calculates a median and 95% confidence interval (CI) the differences. If the 95% CI falls above or below zero, a significant difference exists between the compared results. On the contrary, if the 95% CI includes zero, differences are not significant. This strategy has been employed to test differences between results obtained with other types of Bayesian approaches (Holsinger & Wallace 2004). For these comparisons, the first 5,000 steps of each run were discarded as burn-in.

Results

Patterns of allele variation

A total of 171 different MHC class IIβ sequences were found in this study: 78 in *E. gouldiae* and 94 in *P. acuticauda*. The identification of alleles using SSCP proved to be reliable, as alleles with identical bands in gels produced identical sequences. None of the sequences obtained contained stop codons. Most of the sequences were found in single individuals. Only six alleles were found in more than one *E. gouldiae* individual and seven in more than one *P. acuticauda*. One allele (Ergo04.05) was found in individuals of both species. All MHC sequences in this study have been deposited in GenBank under accession numbers EF535335-EF535505. Amino-acid sequences of each allele are listed in Appendix III. The numbers of haplotypes varied from four to nine in *E. gouldiae* individuals, and from five to 12 in *P. acuticauda* individuals (Table 4.1), suggesting there are five or more MHC class II loci in *E. gouldiae* and six or more loci in *P. acuticauda*. Haplotype diversity values were high in both species and nucleotide diversity was slightly lower in *E. gouldiae* compared to *P. acuticauda*.

Table 4.1 Diversity of MHC class IIβ alleles in *E. gouldiae* and *P. acuticauda*.

	Erythrura gouldiae	*Poephila acuticauda*
Sample size	14	14
Total number of alleles in sample	78	94
Range of number of alleles per individual	4-10	5-12
Average number of alleles per individual	6.43	7.36
Number of polymorphic sites	93	110
Haplotype diversity in sample (Hd)	0.994	0.997
Nucleotide diversity in sample (π)	0.129	0.158

Hierarchical analysis of molecular variance (AMOVA) show most of the variation in the datasets of either species is explained by differences of alleles within individuals ;and to a lesser extent by variation among individuals within population; and among populations (Table 4.2). In both species variance components were significant among alleles within individuals and a mong individuals within populations, but variation among populations is not significant (P> 0.05; Table 4.2)

Table 4.2 Hierarchical Analyses of Molecular Variance (AMOVA) for MHC class IIβ alleles in *E. gouldiae* and *P. acuticauda*, fixation indexes for the variance components and their respective significance.

	Source of variation	Observed partition		Fixation index	*P*
		Variance component	% of variation		
Erythrura gouldiae	Among Populations	0.2081	1.65	Φ_{CT}= 0.0927	0.1743
	Among individuals within populations	1.1507	9.12	Φ_{SC}= 0.1077	0.0029
	Among alleles within individuals	11.2559	89.23	Φ_{ST}= 0.0165	0.0012
	Total	12.6147			
Poephila acuticauda	Among Populations	-0.1204	0.83	Φ_{CT}=-0.00829	0.7351
	Among individuals within populations	1.5217	10.48	Φ_{SC}= 0.10396	0
	Among alleles within individuals	13.1154	90.35	Φ_{ST}= 0.09653	0
	Total	14.5167			

Table 4.3 shows results of the different neutrality tests applied to the dataset of each species. Values of Tajima's *D* test were positive and non-significant for both species, indicating a deviation from neutrality. Fu and Li's *D** and *F** tests were negative, and significant for *D** in both species. Fu's *Fs* were negative and significant for both species. Positive Tajima's *D* values are indicative of a deviation from neutrality caused by a decrease in population size and/or balancing selection. Negative and significant values of Fu and Li's *D** can be interpreted as evidence for selection. Significant values of Fu's *Fs* are often indicative of population growth. However, *Fs* is sensitive to the presence of recombination, as it normally inflates the total number of alleles (Fu 1997).Therefore, Fu's *Fs* results should be interpreted carefully in this study, given that gene conversion is present.

Table 4.3 Tests of neutrality for MHC class IIβ alleles in *E. gouldiae* and *P. acuticauda*. Significant values are marked with an asterisk (*).

	E. gouldiae n = 90	*P. acuticauda* n = 103
Tajima's *D*	0.6801 n.s.	0.9266 n.s.
Fu and Li's *D**	-2.7192 *	-2.3780 *
Fu and Li's *F**	-1.5718 n.s.	-1.1605 n.s.
Fu's F_S	-49.841 *	-66.250 *

Gene conversion

The program GENECONV detected a total of 87 significant gene conversion events between *E. gouldiae* haplotypes (Table 4.4), and 30 events between *P. acuticauda* sequences (Table 4.5). The length of the converted fragments varied from 45 to 138 base pairs in *E. gouldiae* and from 57 to 127 base pairs in *P. acuticauda*. Gene conversion events appear to involve only a portion of all the tested sequences. There were 38 haplotypes involved in gene conversion events for *E gouldiae*, and 30 for *P. acuticauda*, corresponding to 48.7% and 31.9% of the sample total for each species respectively.

Table 4.4 Detected gene conversion events in MHC class IIβ haplotypes of *Erythrura gouldiae*. Simulated *P* values correspond to those obtained after 10,000 iterations. Fragment start and fragment end are relative to the start of the sequence.

Haplotype pairs	Simulated *P* value	Fragment start	Fragment End	Fragment length (bp)	Number of fragment polymorphic sites	Number of Fragment mismatches
Ergo12.03 Ergo16.01	0.0000	1	98	98	47	26
Ergo12.03 Ergo16.02	0.0000	1	98	98	47	25
Ergo12.03 Ergo16.03	0.0000	1	98	98	47	25
Ergo05.01 Ergo16.01	0.0000	1	106	106	50	23
Ergo04.02 Ergo12.03	0.0000	1	79	79	37	30
Ergo16.01 Ergo03.03	0.0000	1	98	98	47	24
Ergo16.02 Ergo03.03	0.0000	1	98	98	47	24
Ergo03.03 Ergo16.03	0.0000	1	98	98	47	24
Ergo12.02 Ergo15.02	0.0000	18	98	81	43	26
Ergo15.02 Ergo20.01	0.0000	18	98	81	43	26
Ergo04.02 Ergo03.03	0.0000	1	79	79	37	29
Ergo04.02 Ergo03.06	0.0000	1	79	79	37	29
Ergo12.03 Ergo04.06	0.0000	1	79	79	37	29
Ergo02.01 Ergo15.02	0.0000	21	98	78	40	27
Ergo05.01 Ergo16.02	0.0000	1	106	106	50	22
Ergo05.01 Ergo16.03	0.0000	1	106	106	50	22
Ergo15.02 Ergo14.02	0.0000	18	98	81	43	25
Ergo16.01 Ergo09.03	0.0000	1	98	98	47	23
Ergo16.02 Ergo09.03	0.0000	1	98	98	47	23
Ergo09.03 Ergo16.03	0.0000	1	98	98	47	23
Ergo04.02 Ergo09.03	0.0003	1	79	79	37	28
Ergo04.02 Ergo05.04	0.0003	1	79	79	37	28
Ergo03.03 Ergo04.06	0.0003	1	79	79	37	28
Ergo03.06 Ergo04.06	0.0003	1	79	79	37	28
Ergo16.01 Ergo19.03	0.0005	1	98	98	47	22
Ergo16.02 Ergo19.03	0.0005	1	98	98	47	22
Ergo16.03 Ergo19.03	0.0005	1	98	98	47	22
Ergo04.02 Ergo03.05	0.0006	18	79	62	33	30
Ergo04.02 Ergo05.01	0.0006	1	79	79	37	27
Ergo04.02 Ergo19.03	0.0006	1	79	79	37	27
Ergo04.06 Ergo09.03	0.0006	1	79	79	37	27
Ergo04.06 Ergo05.04	0.0006	1	79	79	37	27
Ergo16.02 Ergo18.03	0.0006	54	138	85	42	24
Ergo11.07 Ergo16.05	0.0006	54	133	80	40	25
Ergo04.02 Ergo05.02	0.0010	1	77	77	35	28
Ergo02.02 Ergo12.05	0.0013	1	106	106	50	20
Ergo16.01 Ergo03.06	0.0014	1	87	87	41	24
Ergo16.02 Ergo03.06	0.0014	1	87	87	41	24
Ergo03.06 Ergo16.03	0.0014	1	87	87	41	24
Ergo02.01 Ergo15.05	0.0018	29	98	70	33	29
Ergo03.05 Ergo04.06	0.0018	18	79	62	33	29
Ergo05.01 Ergo04.06	0.0018	1	79	79	37	26
Ergo04.06 Ergo19.03	0.0018	1	79	79	37	26
Ergo18.03 Ergo16.06	0.0019	1	138	138	70	14

Table 4.4 (Continued)

Haplotype pairs	Simulated P value	Fragment start	Fragment End	Fragment length (bp)	Number of fragment polymorphic sites	Number of fragment mismatches
Ergo16.01 Ergo05.04	0.0021	1	86	86	40	24
Ergo04.06 Ergo05.02	0.0024	1	77	77	35	27
Ergo16.01 Ergo03.05	0.0024	18	90	73	38	25
Ergo16.02 Ergo03.05	0.0024	18	90	73	38	25
Ergo03.05 Ergo16.03	0.0024	18	90	73	38	25
Ergo12.02 Ergo15.05	0.0032	29	98	70	33	28
Ergo15.05 Ergo20.01	0.0032	29	98	70	33	28
Ergo16.01 Ergo09.04	0.0032	22	98	77	39	24
Ergo16.02 Ergo09.04	0.0032	22	98	77	39	24
Ergo16.03 Ergo09.04	0.0032	22	98	77	39	24
Ergo16.01 Ergo03.04	0.0033	24	98	75	37	25
Ergo16.02 Ergo03.04	0.0033	24	98	75	37	25
Ergo03.04 Ergo16.03	0.0033	24	98	75	37	25
Ergo16.02 Ergo05.04	0.0054	1	86	86	40	23
Ergo16.03 Ergo05.04	0.0054	1	86	86	40	23
Ergo15.05 Ergo14.02	0.0071	29	98	70	33	27
Ergo16.01 Ergo08.05	0.0085	24	98	75	37	24
Ergo16.02 Ergo08.05	0.0085	24	98	75	37	24
Ergo16.03 Ergo08.05	0.0085	24	98	75	37	24
Ergo15.05 Ergo14.03	0.0089	33	98	66	30	29
Ergo04.02 Ergo18.07	0.0109	1	65	65	32	27
Ergo04.02 Ergo16.05	0.0115	1	97	97	46	19
Ergo04.02 Ergo09.04	0.0151	22	79	58	29	29
Ergo16.01 Ergo05.02	0.0197	1	77	77	35	24
Ergo04.02 Ergo18.05	0.0202	1	65	65	32	26
Ergo04.06 Ergo18.07	0.0202	1	65	65	32	26
Ergo05.01 Ergo05.03	0.0212	54	106	53	22	36
Ergo16.01 Ergo18.03	0.0234	54	120	67	33	25
Ergo04.06 Ergo09.04	0.0261	22	79	58	29	28
Ergo15.01 Ergo19.02	0.0281	1	97	97	46	18
Ergo04.06 Ergo16.05	0.0281	1	97	97	46	18
Ergo15.02 Ergo14.03	0.0293	33	98	66	30	27
Ergo16.02 Ergo05.02	0.0346	1	77	77	35	23
Ergo05.02 Ergo16.03	0.0346	1	77	77	35	23
Ergo16.01 Ergo02.04	0.0359	30	98	69	32	25
Ergo16.01 Ergo18.07	0.0359	1	65	65	32	25
Ergo04.06 Ergo18.05	0.0359	1	65	65	32	25
Ergo04.02 Ergo08.05	0.0395	24	79	56	27	29
Ergo02.02 Ergo16.01	0.0443	29	98	70	33	24
Ergo16.02 Ergo19.02	0.0443	81	139	59	33	24
Ergo02.02 Ergo12.04	0.0443	1	106	106	50	16
Ergo12.03 Ergo05.03	0.0453	54	98	45	19	39
Ergo03.04 Ergo05.03	0.0453	54	98	45	19	39

Table 4.5 Detected gene conversion events in MHC class IIβ haplotypes of *Poephila acuticauda*. Simulated *P* values correspond to those obtained after 10,000 iterations. Fragment start and fragment end are relative to the start of the sequence.

Haplotype pairs	Simulated *P* value	Fragment start	Fragment End	Fragment length (bp)	Number of fragment polymorphic sites	Number of Fragment mismatches
Poac12.05 Poac15.03	0.0000	1	76	76	47	31
Poac05.06 Poac13.05	0.0002	28	122	95	62	22
Poac05.06 Poac13.03	0.0006	28	122	95	62	21
Poac03.08 Poac09.07	0.0065	28	96	69	43	27
Poac06.03 Poac14.04	0.0077	44	106	63	38	30
Poac19.02 Poac03.08	0.0109	45	103	59	35	32
Poac03.08 Poac09.09	0.0140	28	96	69	43	26
Poac13.02 Poac06.03	0.0162	44	106	63	38	29
Poac13.02 Poac10.05	0.0162	44	106	63	38	29
Poac06.07 Poac14.04	0.0162	44	106	63	38	29
Poac19.02 Poac19.06	0.0175	1	87	87	53	21
Poac19.03 Poac19.06	0.0175	1	87	87	53	21
Poac19.01 Poac03.08	0.0185	45	103	59	35	31
Poac03.01 Poac03.09	0.0266	23	96	74	47	23
Poac13.02 Poac06.07	0.0291	44	106	63	38	28
Poac13.02 Poac10.02	0.0291	44	106	63	38	28
Poac17.01 Poac14.04	0.0291	44	106	63	38	28
Poac17.02 Poac14.04	0.0291	44	106	63	38	28
Poac14.04 Poac10.05	0.0291	44	106	63	38	28
Poac16.03 Poac17.04	0.0316	50	129	80	51	21
Poac07.03 Poac10.03	0.0316	45	171	127	82	13
Poac19.01 Poac19.06	0.0356	1	87	87	53	20
Poac13.02 Poac09.07	0.0379	53	105	53	30	34
Poac17.03 Poac19.07	0.0413	26	87	62	37	28
Poac13.02 Poac17.01	0.0494	44	106	63	38	27
Poac13.02 Poac17.02	0.0494	44	106	63	38	27
Poac14.04 Poac10.02	0.0494	44	106	63	38	27
Poac03.01 Poac03.07	0.0500	23	96	74	47	22
Poac03.01 Poac13.04	0.0500	23	96	74	47	22
Poac12.03 Poac03.09	0.0500	23	96	74	47	22

Sequences that were involved in gene conversion events appear clustered in some branches in phylogenetic trees. In the case of *E. gouldiae*, branches that were clustered as outliers with respect of the rest of the tree contain mostly sequences detected in gene conversion, and therefore they might potentially be of recombinant origin (Fig 4.1). The presence of gene conversion in the datasets can affect the reconstruction of the phylogeny, as the recombinant sequences should not reflect the evolutionary relationships of the MHC alleles under the models of substitution used to reconstruct the phylogenetic trees in Fig. 4.1. Hence, these phylogenies should be interpreted carefully.

A)

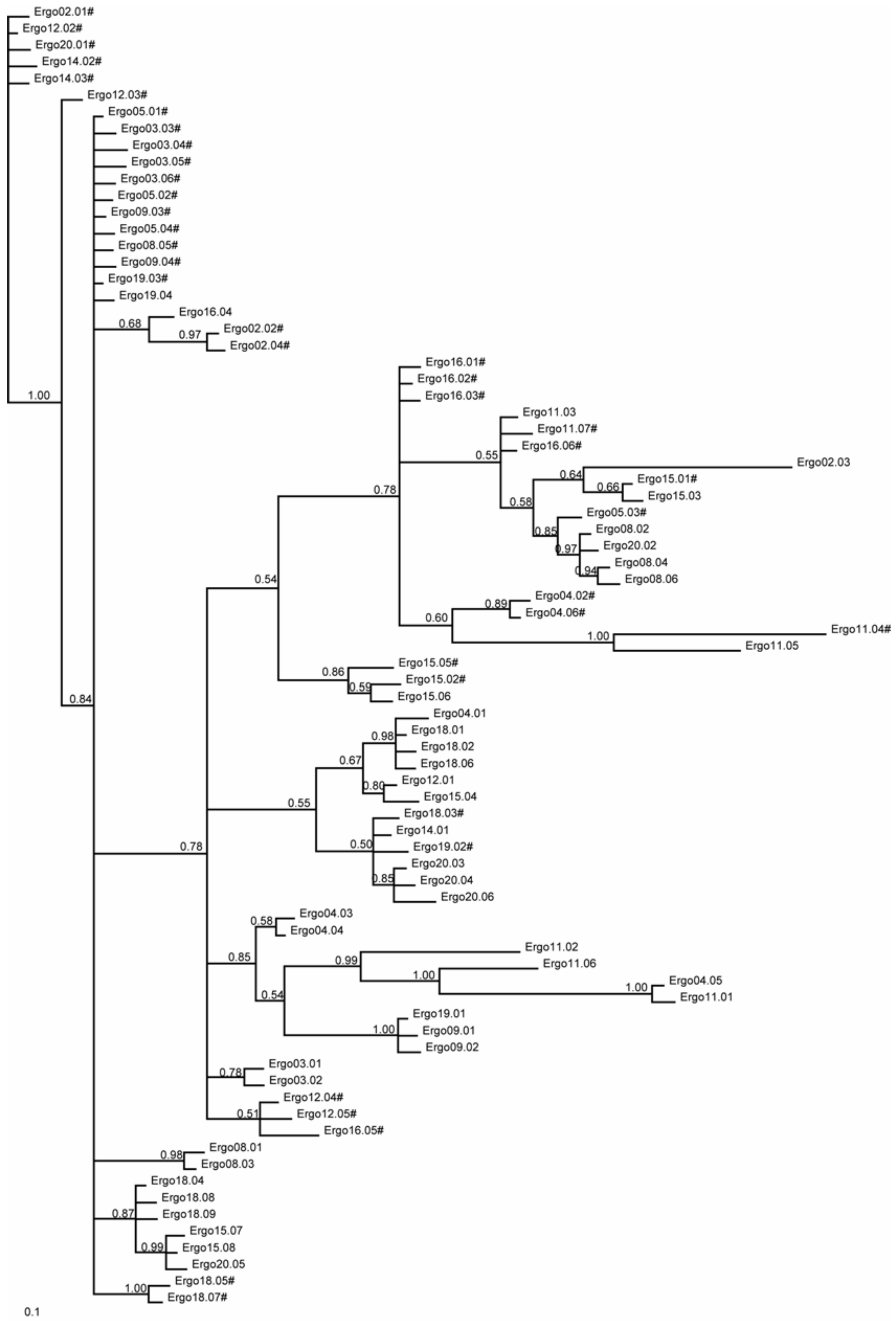


Figure 4.1 Bayesian phylogenetic trees of MHC class IIβ alleles A) *Erythrura gouldiae*; B) *Poephila acuticauda*. Haplotypes marked with a # sign correspond to those involved in putative gene conversion events. Values above branches correspond to posterior probabilities for the respective lineages.

B)

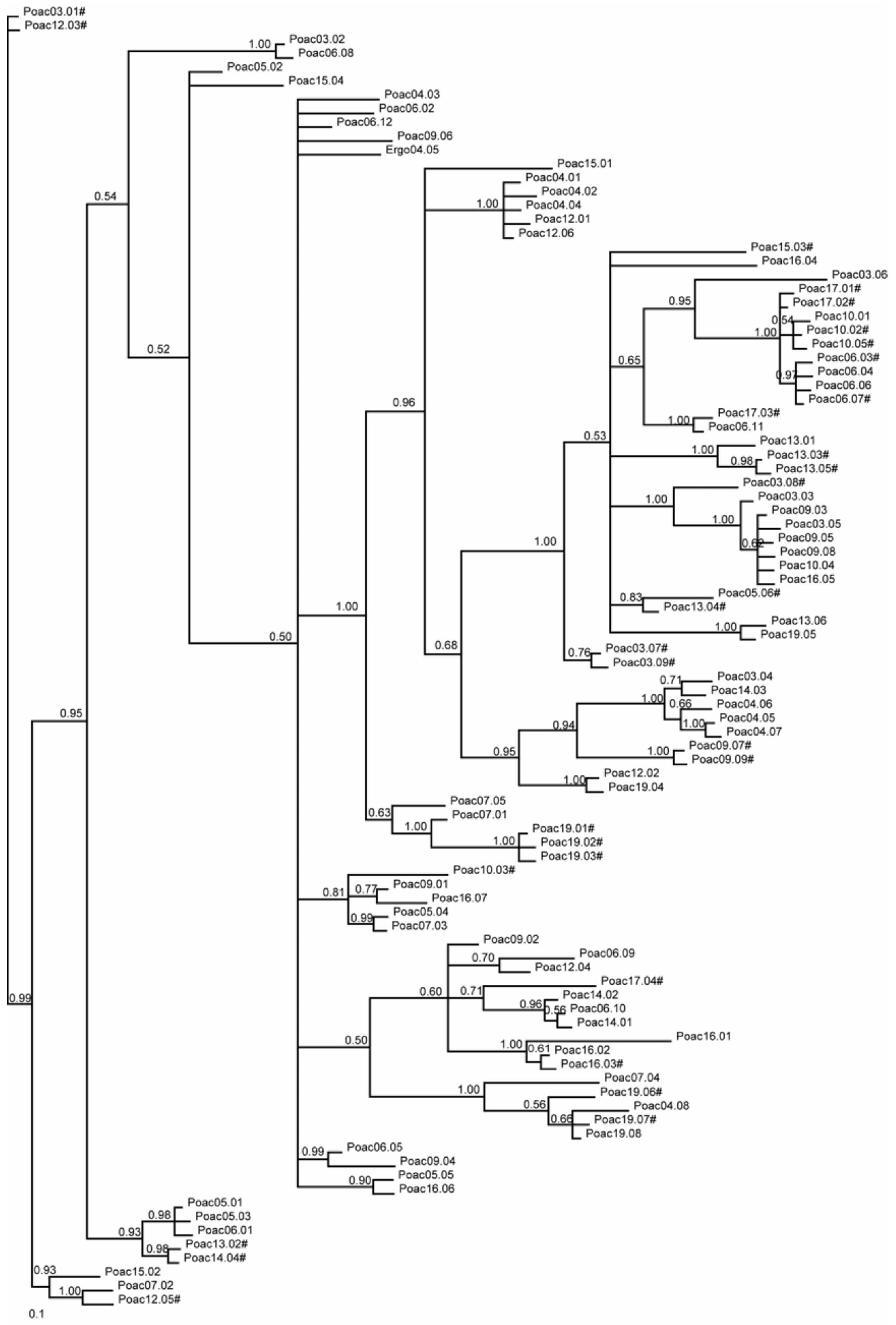


Figure 4.1 (Continued).

Figure 4.2 shows gene conversions detected in *E. gouldiae* involved mostly the first 98 nucleotides of the sequence, while in *P. acuticauda* most gene conversion events occurred from nucleotide positions 28 to 106.

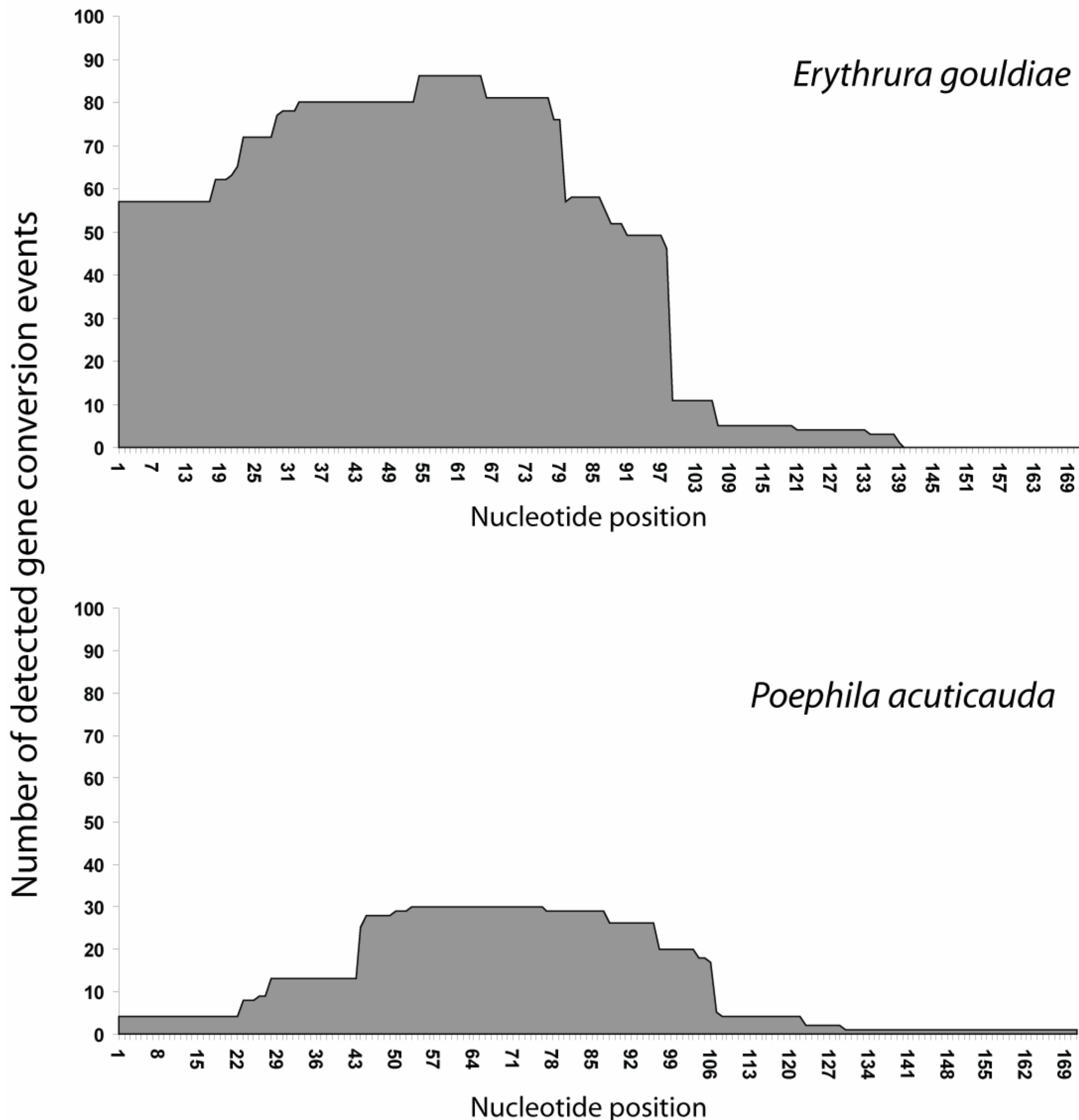


Figure 4.2 Frequency of gene conversions detected between MHC class IIβ haplotypes of *E. gouldiae* and *P. acuticauda*, relative to nucleotide positions.

Positive selection at MHC class IIβ genes

Evidence for positive selection was found in both species. Global d_N/d_S ratio is greater for *E. gouldiae* than for *P. acuticauda*, although such difference is not statistically significant (Table 4.6). A total of 9 amino-acid positions were identified as

positively selected in *E. gouldiae* and 10 in *P. acuticauda* (Table 4.6). There are many similarities between the positively selected amino-acid sites in both species, the majority of which either correspond to or are adjacent to putative PBCs in the human DRB1 gene (Brown *et al.* 1993) (Fig. 4.3). Although there appears to be a common pattern in amino-acid specific d_N/d_S ratios, only five amino-acid positions were found to be positively selected in both species (Table 4.6; Fig. 4.3).Comparisons between the results of both species showed there are significant differences between amino-acid position 37 (mean difference = -2.393; 95% CI -4.771,-0.447) and 44 (mean difference = -1.5; 95% CI -3.368,-0.1123). In both of these cases mean d_N/d_S ratios are >1 for *P. acuticauda* and <1 for *E. gouldiae* (Fig. 4.3).

Table 4.6 Amino-acid positions detected under positive selection, indicating estimate d_N/d_S ratios, 95% highest posterior density (HPD) intervals and posterior distributions for positive selection. Global d_N/d_S ratios below were obtained from all 57 amino-acid positions using a constant selection model. Amino-acid positions highlighted in bold correspond to PBC positions in the human DRB1 gene (Brown *et al.* 1993).

Species	Amino-acid position	d_N/d_S ratio (point estimate)	95% HPD interval	Posterior probability for positive selection
Erythrura gouldiae	**10**	2.9041	1.4502 – 6.3438	0.9928
	18	2.7926	1.2476 – 6.0874	0.9870
	27	2.2923	0.9430 – 4.8936	0.9586
	33	3.5257	1.5579 – 8.1063	0.9913
	39	2.5874	1.2390 – 5.2247	0.9928
	40	3.1525	1.3542 – 6.8282	0.9929
	41	3.1562	1.5355 – 6.5798	0.9938
	51	3.3124	1.5878– 6.9077	0.9994
	57	2.3530	0.9254 – 5.1986	0.9574
	Global	1.2623	0.8486 – 1.9440	0.8701
Poephila acuticauda	**10**	2.6667	0.9980 – 4.4438	0.9677
	33	4.0121	1.9994 – 8.5180	0.9994
	36	2.9337	1.5830 – 5.5444	0.9949
	37	2.0959	0.9460 – 4.1994	0.9619
	40	2.1669	1.0891 – 4.3338	0.9749
	41	2.9941	1.6074 – 5.9646	0.9942
	47	2.5651	1.2460 – 5.4611	0.9880
	50	3.1333	1.5824 – 6.2089	0.9998
	51	3.3963	1.7409 – 6.7071	1.0000
	54	2.1925	1.0639 – 4.5559	0.9674
	Global	0.9152	0.6758 – 1.2406	0.2832

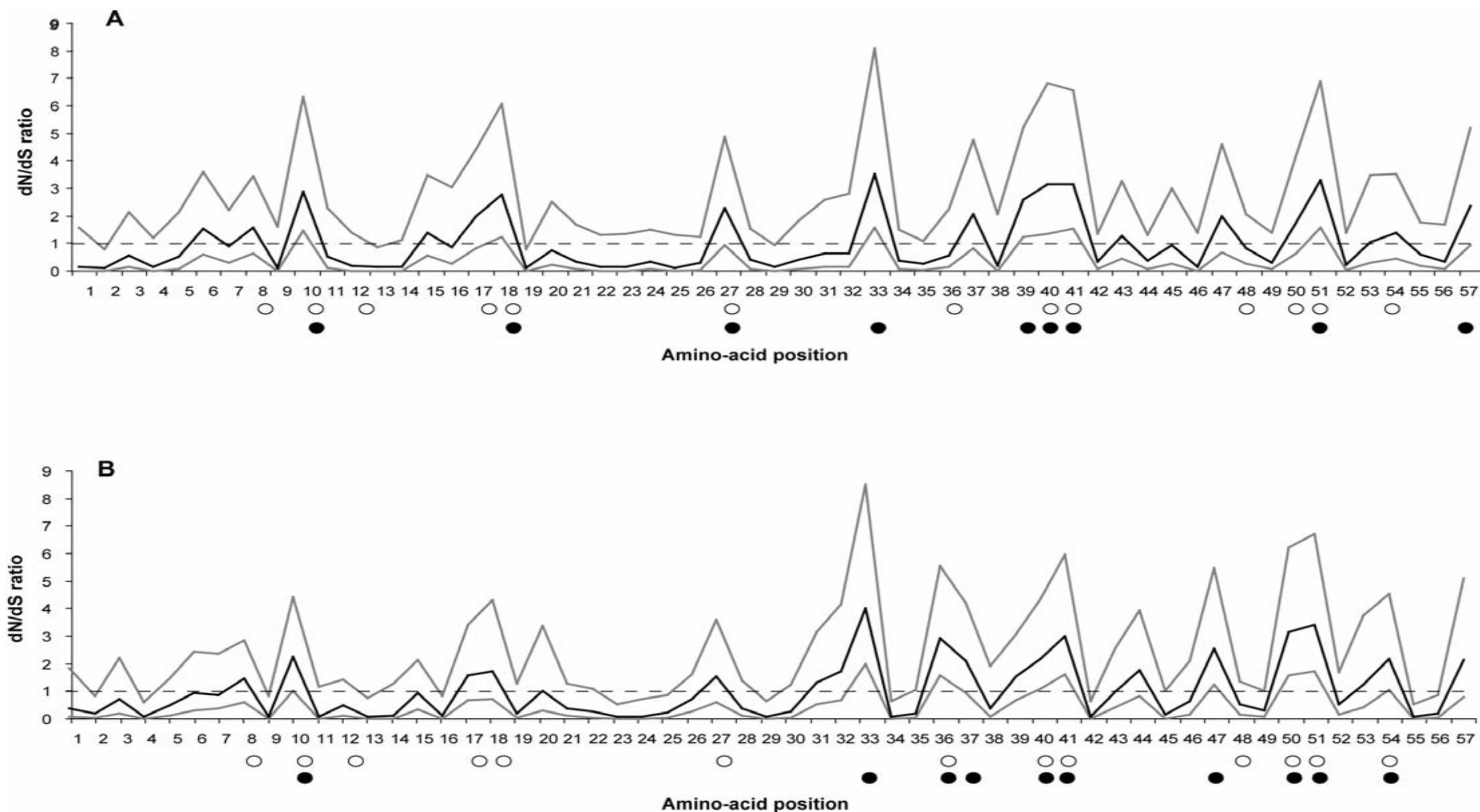


Figure 4.3 Ratios of d_N/d_S per amino-acid site in MHC class IIβ genes: A) *Erythrura gouldiae*; B) *Poephila acuticauda*. Solid black lines represent calculated d_N/d_S values and solid grey lines represent upper and lower bounds of the 95% higher posterior densities. Broken lines indicate d_N/d_S ratio level of 1. Amino-acid sites under significant positive selection are indicated with filled ovals. Predicted peptide binding amino-acid positions in human DRB1 gene (Brown *et al.* 1993) are indicated with open ovals.

Discussion

Maintenance of MHC class II diversity

Both studied species possess a relatively high number of alleles relative to the number of individuals screened. The fact that most alleles were found in single individuals suggests the actual number of alleles in a population may be much higher than those that were detected in this study. Although there are some differences in the number of haplotypes, polymorphic nucleotide sites, and nucleotide diversity in both species, these differences are not significant. AMOVA results suggest selection has favoured allele variation within individuals, and to a lesser extent individual variation within populations. Contrastingly, variation between populations is not significant in spite of the distance between the sample locations, and the fact that populations of *P. acuticauda* are structured, and often considered separate taxa in each of such locations.

Allele richness seems to have been maintained in both species; although there is evidence that some of the alleles detected in this study may be of recombinant origin, particularly in *E. gouldiae*, for which gene conversion appears to be a more common phenomenon than for *P. acuticauda*. Many of its alleles are products of gene conversion between existing alleles and may have been maintained through selection.

Gene conversion can act to restore allelic diversity for MHC genes following a bottleneck more quickly than could mutation alone (Schaschl *et al.* 2006). The higher prevalence of gene conversion events in *E. gouldiae* may reflect a strategy for the maintenance of allelic diversity that compensates for a lower allelic diversity resulting from a historical population bottleneck, for which there is evidence from analysis of mitochondrial DNA (Chapter 2).

These differences in MHC variability and the way it is maintained in both species may be directly or indirectly related to the striking differences between parasite loads between them.

Selection appears to be strong in both species, which is sound for genes involved in pathogen resistance such as MHC. Selection is also supported by neutrality tests in both species. The negative significant values of Fu's *Fs* test are often indicative of population growth. However, given that the test is sensitive to recombination, these results are likely to be inaccurate. The positive results from Tajima's *D* and the significant results from Fu and Li's *D** also favour the hypothesis of selection over demographic changes to explain deviations from neutrality.

Despite the fact that these two species are related, live in similar environments and presumably are subject to similar evolutionary pressures, there are important differences in the rates of positive selection at the amino-acid level. At least there are two clear cases in which amino-acids are positively selected in one species (*P. acuticauda*) and subject to purging selection in the other (*E. gouldiae*).

These differences may be indicative of different evolutionary pressures acting on MHC variation in each species. Variation in the relative position of positively selected amino-acid sites may also reflect differences in the secondary structure of the MHC protein. MHC are important proteins in all vertebrates, and their function in pathogen recognition is similar amongst different vertebrate groups. However, it is not likely that the secondary structure is highly conserved. Figure 4.3 shows there are strong similarities in the rates of positive selection and the relative position of selected amino-acids between the two studied species. However, although the selected amino-acids in this study are largely adjacent or at the sites identified for human MHC class II genes, there are other important differences.

In the case of some MHC amino-acid positions in this study, mean d_N/d_S ratios are >1, which is indicative of positive selection. However, in many cases, the positive selection is non-significant. This uncertainty might be explained by the fact that the datasets analysed in this study contain alleles from several loci. If selective pressures act differently in different loci, there might be cases in which an amino-acid position is positively selected in one locus and under purging selection in another. Unfortunately, the peculiar nature of passerine MHC means it is difficult to distinguish between loci, and moreover, recombination can happen between different loci as well as between alleles in the same locus. Another possibility is the presence of pseudogenes in the datasets. Although I did not find evidence of such, given the absence of stop codons in the sequence, their presence is still a possibility. In such case, selection acting in expressed alleles potentially combined with neutral evolution of pseudogenes may be the cause of uncertainty for the positive selection at some codons.

Some of the differences found between the two studied species may be explained by the fact that apparently there were more loci amplified in *P. acuticauda* than in *E. gouldiae*. These differences might be due to the fact that there are more copies of the gene in the former, although I can not rule out the possibility that some loci may have failed to amplify during the PCR. Analyses of individual loci may be useful to resolve

some of these uncertainties. This may be possible using laboratory approaches such as motif-specific amplification of MHC alleles (Westerdahl *et al.* 2004b).

Considerations on statistical methodologies

Other studies that investigate selection at MHC genes in non-model organisms (e.g. Jarvi *et al.* 2004; Miller & Lambert 2004; Schad *et al.* 2004; Meyer-Lucht & Sommer 2005) have used methods such as that of Nei and Gojobori (1986) to calculate d_N/d_S ratios in peptide-binding codons (PBC) and non-PBC between aligned sequences, often using the predicted human PBC (Brown *et al.* 1988; Brown *et al.* 1993) as a reference. However, results from this study show that amino-acid positions detected under positive selection do not always match those identified as PBC in humans. Other studies in species of ungulates similarly show an incomplete match of amino-acids under selection with PBCs in human DRB1 genes (Schaschl *et al.* 2005; Schaschl *et al.* 2006). The assumption that MHC acts similarly amongst different vertebrate groups is likely to produce misleading results, as some positively selected sited might often be unnoticed under such assumption. Specific codon-based analyses of selection are a better approach for the detection of selection at particular sites.

The presence of recombination or gene conversion events may increase the rate of false positive results for recombination (Kosakovsky Pond *et al.* 2006; Wilson & McVean 2006). There are a number of mathematical models available for the detection of positive selection in genes. Some of the more widely used such as M8 (Yang *et al.* 2000), FEL and REL (Kosakovsky Pond *et al.* 2005) require the inference of a phylogeny for the detection of selection. When recombination occurs in the gene in question, reconstructed phylogenies do not reflect the real genealogies between taxa, and therefore the rate of false positives in the detection of selection can be potentially high (Wilson & McVean 2006). Some methods address this issue by detecting recombination breakpoints, reconstructing partial phylogenies from fragments of the gene separated by breakpoints, and use them for the detection of selection (Kosakovsky Pond *et al.* 2006; Scheffler *et al.* 2006). However, in gene conversion events, alleles do not necessarily recombine at defined breakpoints, as is suggested by results from this study (Fig. 4.3; tables 4.4 and 4.5). Therefore, detection of recombination breakpoints and reconstruction of partial phylogenies can also produce misleading results in the presence of gene conversion. A method for the calculation of d_N/d_S with a population genetics approach that does not rely on phylogeny inference, such as that implemented

in omegaMap (Wilson & McVean 2006), is therefore a better approximation for the calculation of selection in the presence of gene conversion.

CHAPTER FIVE: General discussion

General summary

The use of molecular genetic techniques for the research described in this thesis reveals several previously unknown aspects of population history, ecology and evolution of the endangered Gouldian Finch, *Erythrura gouldiae*. Lineages of *E. gouldiae* are not geographically isolated in spite of the extension of their distribution. This may result from a past population bottleneck and subsequent population growth and geographic expansion events during the Holocene period throughout northern Australia. There is no evidence of recent population bottlenecks or severe inbreeding in two of the largest known wild populations of *E. gouldiae*, although lower genetic diversity and evidence of a recent population bottleneck was found in an *E. gouldiae* population which is part of a reintroduction program. Results show population structure is absent in *E. gouldiae*, while those of *P. acuticauda* show a clearer structure between different regions. These differences in structure can be explained by the higher migration rates in *E. gouldiae* compared to those of *P. acuticauda*. Allelic richness at MHC genes is high in *E. gouldiae* and *P. acuticauda*, although the total number of alleles, average number of alleles per individual, number of polymorphic sites and nucleotide diversity are higher in the latter. The high allelic richness at MHC class IIβ is likely to be maintained in both species through gene conversion and balancing selection. The number of gene conversion events is higher in *E. gouldiae*, suggesting this evolutionary mechanisms is relatively more important in maintaining allelic richness at MHC class IIβ genes in *E. gouldiae* than *P. acuticauda*. Because *E. gouldiae* is a species of conservation concern, there are important conservation implications of the results obtained from this thesis, and results outlined from the research described in this thesis are therefore expected to be considered in current and future implementation of management practices for the recovery of the species.

Summary of key findings

Historical phylgeography

The historically observed distribution of *Erythrura gouldiae* appears to be of relatively recent origin. The apparently low divergence between present mtDNA haplotypes combined with the unimodal mismatch distribution suggests this species suffered a population reduction in historical times, followed by population expansion events throughout northern Australia (Chapter 2). The differences between mismatch distributions between individuals of different regional groups indicate *E. gouldiae* population expansions have happened gradually in Northern Australia rather than uniformly. Evidence suggests more recent expansion events in the western part of the distribution than in the east. The approximate calculated times since expansion of *E. gouldiae* population using mismatch distributions of mtDNA haplotypes corresponds to events during the Holocene, at approximately 7000 to 3600 years before present, when there were important habitat changes in the tropical savannas of northern Australia that might have been favourable to the population growth of *E. gouldiae*.

Current genetic variability and population structure

Results from Chapter 3 indicate microsatellite heterozygosity is higher in *P. acuticauda* than in *E. gouldiae*. Evidence of recent population bottlenecks was not found in wild populations of either species, but a recent population bottleneck appears to have happened in the captive population for reintroduction in north-eastern Australia. Intra-population inbreeding coefficients (*f*) were relatively low in both species, indicating there is no severe inbreeding within populations. Inter-population genetic differentiation (θ^{II}) was moderate in both, and significantly larger for *P. acuticauda*. Assignment tests support evidence of population structuring between *P. acuticauda* populations, which is absent in *E. gouldiae* populations. Migration rates appear higher in *E. gouldiae* compared to *P. acuticauda*.

Genetic diversity in a captive population for reintroduction

The current population for the reintroduction program of *E. gouldiae* in the Mareeba wetlands appears to lack in gene diversity at mtDNA control region (Chapter 2). Analyses of microsatellite loci reveal some evidence of a recent population bottleneck and significantly lower heterozygosity than either of two wild populations

(Chapter 3). This low genetic diversity can have adverse consequences in the potential of this population to adapt to stochastic events and the outbreak of diseases.

Variability, gene conversion and selection in MHC class IIβ genes

Analyses of the variability of MHC class IIβ genes of *E. gouldiae* and *P. acuticauda* (Chapter 4) shows the former has a lower number of alleles in the sample analysed than the latter. Gene conversion events between haplotypes were more prevalent in *E. gouldiae* than in *P. acuticauda*, indicating gene conversion has been relatively more important in the maintenance of genetic variability of MHC genes in *E. gouldiae*. Positive selection at the amino-acid level appears to have a similar in pattern between the two species, but there are nevertheless examples of amino-acid positions that are under purging selection in *E. gouldiae* and positively selected in *P. acuticauda*. These differences in variability of MHC genes, the prevalence of gene conversion events and differences in patterns selection may reflect differences in selective pressures posed by pathogens.

Conservation implications of results

The results described in this thesis have several implications for the efficient implementation of conservation strategies.

Erythrura gouldiae lineages are not geographically segregated (Chapter 2). A lack of population structure as indicated by the analyses of mitochondrial DNA (Chapter 2) and microsatellite loci (Chapter 3) suggests *E. gouldiae* populations have not evolved into separate evolutionary entities, and therefore the whole distribution of the species can be considered to be a single evolutionarily significant unit following the criteria of Moritz (1995). An important implication of this finding in current conservation is the possibility of allowing translocation of individuals between subpopulations, without the risk of mixing populations that are genetically different. The genetic variability of small populations may therefore be increased by the translocation of individuals from a different area, in cases where natural migration may be somehow impeded.

There is no evidence of severe inbreeding or recent population bottlenecks in the two wild populations studied in this thesis (Yinberrie Hills and Mornington Wildlife Sanctuary) and heterozygosities in either of these two populations are not significantly lower than those of the relatively more abundant *P. acuticauda*. It is however important

to consider that the populations mentioned are amongst the largest recorded for this species (O'Malley 2006) and may not reflect the genetic variation at other locations within the current distribution of *E. gouldiae*. In some of such locations, lower habitat quality may result in important population fluctuations that may result in the loss of genetic diversity due to drift.

Migration rates between *E. gouldiae* populations appear to be higher than those of *P. acuticauda* (Chapter 3), although they do not differ statistically.

Some potential consequences of captive breeding include the rapid loss of genetic diversity in populations with a small number of founders that can result in rapid inbreeding. There is also a risk of adaptation of individuals to captivity, resulting in maladaptation of released individuals to habitat conditions in the wild (Frankham *et al.* 2002). The captive *E. gouldiae* population at Mareeba Wetlands that is part of a reintroduction program lacks genetic diversity at the mtDNA control region (Chapter 2) and has lower heterozygosity at microsatellite loci compared to wild populations (Chapter 3). The observed low levels of genetic variability in that population may be the consequence of a small number of founder individuals in the initial breeding stock, and the subsequent loss of diversity through drift after more than twenty years of breeding in captivity. This scenario of founder effect is supported by the results of bottleneck analyses in Chapter 3. The genetic variability observed in that population can be further reduced through genetic drift in subsequent generations, which would reduce overall fitness in the long term and consequently its capability to respond to stochastic events. Moreover, if variation at adaptively important genes such as MHC is also reduced, the potential of this captive population to resist disease outbreaks can also be reduced, increasing risk of extinction.

It may be necessary to manage this population actively in order to increase genetic variability. Given that *E. gouldiae* populations are not genetically separated (Chapters 2 and 3), the introduction of individuals from wild populations into this captive population is a possibility, as there would be no risk of outbreeding depression, which may happen when mixing populations that are genetically dissimilar (Frankham *et al.* 2002). The introduction of wild individuals would increase the genetic variability and consequently the potential for better survival of released individuals.

It is necessary however to consider the presence of pathogens before the introduction of wild individuals into this captive population, including, as *E. gouldiae* appear to be more susceptible to parasitism than other estrildid finch species (Tidemann

et al. 1992b). This consideration is important to avoid the introduction of new parasite strains to which a population might be locally maladapted.

In addition, it will be necessary to monitor any released populations over several generations to ensure genetic diversity is maintained, and if necessary maximise the effective population sizes and avoid inbreeding by equalising family sizes and through circular mating between individuals (Frankham *et al.* 2002)

The number of detected MHC alleles detected in *E. gouldiae* is lower than in *P. acuticauda* (Chapter 4), which is consistent with low genetic variation found in mtDNA (Chapter 2). This apparently lower richness in MHC genes may explain the higher susceptibility to parasites in this species (Tidemann *et al.* 1992b). Although management aimed to eradicate diseases in wildlife may not be viable, indirect ways to avoid physiological stress in individuals, such as habitat restoration would promote immunocompetence in individuals. Maximising genetic variability through natural gene flow or translocation of individuals would increase the gene pool in small and bottlenecked populations, and would promote the creation of novel MHC alleles in following generations through gene conversion and selection.

Overall conclusions

This thesis includes different pieces of research that use different approaches to answer different questions about the population history, ecology and evolution in *E. gouldiae* contrasting some aspects with *P. acuticauda*. The present geographic distribution of *E. gouldiae* populations is of relatively recent origin (Chapter2). Lineages are more or less homogeneously distributed across Northern Australia. This evidence indicates the species suffered a population bottleneck, possibly at the end of the Pleistocene or during the early Holocene. Later habitat changes in the monsoonal tropics of Northern Australia, that changed the landscape to the more open and grassy habitat we see now, might have favoured the expansion of *E. gouldiae* across northern Australia. There is also some evidence that this expansion occurred first at the easternmost part of the distribution, and more recently in the Kimberley region, at the westernmost part of the distribution, consistent with the relatively higher nucleotide diversity at mtDNA control in Cape York Peninsula and lowest in the Kimberley, which suggests the present *E. gouldiae* populations might have originated in the savannas of

Cape York Peninsula and expanded westwards as habitat conditions became more favourable. This population expansion after a historical bottleneck can explain the reduced genetic variability in the species, as is evident from mtDNA control region variation (Chapter 2), and lower variability than *P. acuticauda* at microsatellites (Chapter 3) and MHC class II genes (Chapter 4).

Analyses of microsatellite allele distribution reveals wild *E. gouldiae* populations have not experienced recent population bottlenecks, and that *E. gouldiae* populations do not show signs of population structure while *P. acuticauda* show moderate structuring. Such results are consistent with differences in migration rates in both species, as revealed by analyses of microsatellite allele frequencies. The fact that *E. gouldiae* are capable of longer-distance migration created an apparently panmictic structure of their populations. Individuals of this species are able to reach neighbouring populations throughout their lives and maintain genetic admixture in successive generations.

Analyses of variation at MHC class II genes in *E. gouldiae* and *P. acuticauda* show allelic richness is relatively high in both species, although nucleotide variation is lower in *E. gouldiae*. There are also important differences in the number of gene conversion events detected between detected haplotypes of each species, and in the patterns of positive selection at the amino-acid level. This shows selection and gene conversion have had a more important role in the maintenance of MHC variability in *E. gouldiae*, compared to *P. acuticauda*, which is consistent with the differences in genetic variability and the conservation status of these two species.

Future research directions

Comparative phylogeography in Australian Tropical Savannas

In contrast with temperate ecosystems, studies on the phylogeography of tropical and subtropical organisms are uncommon, particularly in the tropical lowlands (Hewitt 2004; Wüster *et al.* 2005; Davison & Chiba 2006). Phylogeographic studies on taxa from the different habitats of the monsoonal tropics of Australia, including tropical savannas and other key habitats such as riparian habitats, mangroves and monsoonal rainforests would help elucidate the effects that climate changes and the intensification of habitat use by humans have on the observed geographic distribution patterns of

different ecological communities and their organisms. For example, the past habitat changes that were favourable for the growth and geographical expansion of *E. gouldiae* populations might have affected other tropical savannah species, including grass species, granivorous and grazing animals, as well as some predators. On the other hand, species of riparian habitats and monsoonal rainforests may have experienced habitat contractions following such habitat changes.

Population structure

The high capability of *E. gouldiae* for long-distance migration, as is evident by the lack of genetic structure (chapters 2 and 3) and the comparably higher rates of migration than those of *P. acuticauda* (Chapter 3), suggests that the population structure at the regional and landscape levels might be complex. A study on the population structure of *E. gouldiae* which includes a larger number of populations is desirable in order to detect population structure at different geographical scales. This information, in combination with the available evidence on local movements (Dostine *et al.* 2001), reproductive biology (Tidemann *et al.* 1999) and current long-term population monitoring would permit to identify populations that effectively act as sources of migrants or as receptors of immigrants and would allow making more accurate management decisions at regional and landscape levels.

Evolution and variation of MHC genes

Variation at MHC genes has a potential role as a tool for monitoring the temporal effect of selection in populations, since variation at these genes is known to fluctuate through generations (Westerdahl *et al.* 2004a). Results from Chapter 4 of this thesis show there are important differences in the role of gene conversion and selection in the evolution of MHC genes between species. Ideally, this kind of study would include a large number of individuals and populations. However, the methods outlined in Chapter 4 for the isolation of MHC alleles and the measures to avoid PCR artefacts can be demanding and costly, making them less practical as a regular long-term monitoring tool. The use of novel methodologies for indirect detection of mutation such as capillary electrophoresis SSCP (Lento *et al.* 2003; Endo *et al.* 2005) combined with motif-specific screening of MHC (Westerdahl *et al.* 2004b) are good alternatives for the less costly screening of higher numbers of samples simultaneously, particularly in

species with unusually high number of alleles at these genes such as *E. gouldiae* and *P. acuticauda*.

There are a growing number of studies addressing the relationships between resistance to parasites with number of MHC alleles and local adaptation to certain alleles to endemic strains of the parasite (e.g. Meyer-Lucht & Sommer 2005; Westerdahl *et al.* 2005; Bonneaud *et al.* 2006). Such kind of research would be important for *E. gouldiae*, considering their relatively higher susceptibility of to pathogens. There are methods available for the non-destructive detection of parasites (Sehgal *et al.* 2001; Richard *et al.* 2002; Hellgren *et al.* 2004; Hellgren 2005; Meyer-Lucht & Sommer 2005; Sehgal *et al.* 2005) which can be used together with MHC variation to address this type of questions. Differences in the intensity of gene conversion in the maintenance of MHC genes in each species should also be taken into account in host-parasite interactions, as the heritability of alleles is affected by that evolutionary mechanism.

In addition to MHC class II, the study of variation and adaptation at MHC class I genes would be important, given that these MHC genes recognise different arrays of pathogens, for instance intracellular vs. Extracellular parasites (Edwards & Hedrick 1998). The comparison of evolution in both types of genes would elucidate the importance of different evolutionary strategies for maintaining variation in both types of MHC genes, and can be an indirect indicator of the types of parasites that have driven the evolution of these genes in different species.

Extending studies of MHC evolution to a larger number of related species would allow the comparison of different evolutionary mechanisms for the maintenance of allele richness in species with different population histories and habitat requirements. The Australian estrildid finches can be an ideal model for studying trans-species evolution of MHC genes. The different species represent a wide range in terms of habitat requirements, nesting requirements, conservation status, and geographic distribution.

The process of development of specific molecular markers can be expensive and time consuming. The molecular genetic methods used in this thesis can be used in other species of the family Estrildidae, sometimes with minor modifications to the protocols in some cases. This would be important in conservation, since 18 species of this family worldwide are considered to be under some threatened species category (IUCN 2006). The methods described in Chapter 3 for the assessment, optimisation and cross-species

amplification of variable microsatellite loci can be repeated in other species of the family Estrildidae. There is potential success in cross-species amplification as is indicated by the successful amplification in the two species of different genera used in this thesis. Hence, the demanding and costly process of developing specific primers for other estrildid species can be avoided. Primers used in this study for the amplification of MHC class IIβ genes have successfully amplified PCR products in a wide range of passerine species, including members of the families Emberizidae and Tyrannidae (Sato *et al.* 2001), as well as Estrildidae (this thesis), Melliphagidae and Acanthizidae (Itzel Zamora-Vilchis, pers. com.) and Ptilonorhynchidae (Pers. obs.).

Aguilar A, Roemer G, Debenham S*, et al.* (2004) High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *PNAS* **101**, 3490-3494.

Araújo M, Nogués-Bravo D, Diniz-Filho N*, et al.* (2008) Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography* **31**, 8-15.

Avise JC (2000) *Phylogeography: the history and formation of species* Harvard University Press, Cambridge, USA.

Avise JC, Arnold J, Ball RM*, et al.* (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489-522.

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37-48.

Barret G, Silcocks A, Barry S, Cunningham R, Poulter R (2003) *The New Atlas of Australian Birds* Royal Australasian Ornithologists Union, Hawthorn, Victoria.

Bazin E, Glemin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. *Science* **312**, 570-572.

Beebee TJC, Rowe G (2004) *An Introduction to Molecular Ecology* Oxford University Press, New York.

Bekessy SA, Ennos RA, Burgman MA, Newton AC, Ades PK (2003) Neutral DNA markers fail to detect genetic divergence in an ecologically important trait. *Biological Conservation* **110**, 267-275.

Bernatchez L, Landry C (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology* **16**, 363-377.

Berry O (2006) Mitochondrial DNA and Population Size. *Science* **314**, 1388-1390.

BirdLife International (2006) *Erythrura gouldiae*. In: *2006 IUCN Red List of Threatened Species* http://www.iucnredlist.org*. Downloaded on 30 October 2006.* (ed. 2006 I).

Bonneaud C, Perez-Tris J, Federici P, Chastel O, Sorci G (2006) Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution* **60**, 383-389.

Borriello F, Krauter KS (1990) Reactive site polymorphism in the murine protease inhibitor gene family is delineated using a modification of the PCR reaction (PCR+1). *Nucleic Acids Research* **18**, 5481-5487.

Bos DH, Waldman B (2006) Evolution by recombination and transspecies polymorphism in the MHC class I gene of Xenopus laevis. *Molecular Biology and Evolution* **23**, 137-143.

Brown JH, Jardetzky T, Saper MA*, et al.* (1988) A hypothetical model of the foreign antigen-binding site of class-II histocompatibility molecules. *Nature* **332**, 845-850.

Brown JH, Jardetzky TS, Gorga JC*, et al.* (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33-39.

Caro TM, Laurenson MK (1994) Ecological and genetic factors in conservation: a cautionary tale. *Science* **263**, 485-486.

CBD (2006) *Global Biodiversity Outlook 2* Secretariat of the Convention on Biological Diversity, Montreal.

Clayton NS (1990) Assortative Mating in Zebra Finch Subspecies, Taeniopygia guttata guttata and T. g. castanotis. *Philosophical Transactions: Biological Sciences* **330**, 351-370.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989-2000.

Cracraft J (1986) Origin and evolution of continental biotas - speciation and historical congruence within the Australian avifauna. *Evolution* **40**, 977-996.

Cracraft J (1991) Patterns of diversification within continental biotas: Hierarchical congruence among the areas of endemism of Australian vertebrates. *Australian Systematic Botany* **4**, 211-227.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* **134**, 959-969.

Crozier RH (1997) Preserving the information content of species: genetic diversity, phylogeny and conservation worth. *Annual Review of Ecology and Systematics* **28**, 243-268.

Daugherty CH, Cree A, Hay JM, Thompson MB (1990) Neglected taxonomy and continuing extinctions of tuatara (*Sphenodon*). *Nature* **347**, 177-179.

Davison A, Chiba S (2006) The recent history and population structure of five Mandarina snail species from subtropical Ogasawara (Bonin Islands, Japan). *Molecular Ecology* **15**, 2905-2919.

Dawson DA, Chittock JC, Jehle R*, et al.* (2005) Identification of 13 polymorphic microsatellite loci in the zebra finch, *Taeniopygia guttata* (Passeridae, Aves). *Molecular Ecology Notes* **5**, 298-301.

DEWR Australia (2007) *Erythrura gouldiae in Species Profile and Threats Database, Department of the Environment and Water Resources (DEWR), Canberra. Available from: http://www.environment.gov.au/sprat. Accessed 2007-04-28.*

Di Rienzo A, Peterson AC, Garza JC*, et al.* (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 3166-3170.

Diniz-Filho J (2004) Phylogenetic diversity and conservation priorities under distinct models of phenotypic evolution. *Conservation Biology* **18**, 698-704.

Dostine PL, Franklin DC (2002) A comparison of the diet of three finch species in the Yinberrie Hills area, Northern Territory. *EMU* **102**, 159-164.

Dostine PL, Johnson GC, Franklin DC, Zhang Y, Hempel C (2001) Seasonal use of savanna landscapes by the Gouldian finch *Erythrura gouldiae*, in the Yinberrie Hills area, Northern Territory. *Wildlife Research* **28**, 445-458.

Edwards SV, Hedrick PW (1998) Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends in Ecology and Evolution* **13**, 305-311.

Ellegren H (1994) Genomic DNA from museum bird feathers. In: *Ancient DNA: Recovery and Analysis of Genetic Material from Paleontological, Archaeological, Museum, Medical and Forensic Specimens* (eds. B. H, Hummel S), pp. 211-217. Springer-Verlag, New York.

Endo Y, Zhang L, Katashima R*, et al.* (2005) Effect of polymer matrix and glycerol on rapid single-strand conformation polymorphism analysis by capillary and microchip electrophoresis for detection of mutations in K-ras gene. *Electrophoresis* **26**, 3380-3386.

Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology* **13**, 853-864.

Excoffier L, Laval LG, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Informatics Online* **1**, 47-50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Frankham R (2005) Genetics and extinction. *Biological Conservation* **126**, 131-140.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics* Cambridge University Press, Cambridge, UK.

Frankham R, Ralls K (1998) Conservation biology: Inbreeding leads to extinction. *Nature* **392**, 441-442.

Franklin DC (1999) Evidence of disarray amongst granivorous bird assemblages in the savannas of northern Australia, a region of sparse human settlement. *Biological Conservation* **90**, 53-68.

Franklin DC, Burbidge AH, Dostine PL (1999) The harvest of wild birds for aviculture: an historical perspective on finch trapping in the Kimberley with special emphasis on the Gouldian Finch. *Australian Zoologist* **31**, 92-109.

Franklin DC, Dostine PL (2000) A note on the frequency and genetics of head colour morphs in the Gouldian finch. *Emu* **100**, 236-239.

Franklin DC, Whitehead PJ, Pardon G*, et al.* (2005) Geographic patterns and correlates of the decline of granivorous birds in northern Australia. *Wildlife Research* **32**, 399-408.

Freeland JR (2005) *Molecular Ecology* John Wiley & Sons, Chichester, UK.

Friesen VL, Congdon BC, Walsh HE, Birt TP (1997) Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. *Molecular Ecology* **6**, 1047-1058.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915-925.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* **133**, 693-709.

Galbusera P, Van Dongen S, Matthysen E (2000) Cross-species amplification of microsatellite primers in passerine birds. *Conservation Genetics* **1**, 163-168.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10**, 305-318.

Gelis S (2003) The Gouldian Finch (Erythrura gouldiae) in health and disease. *Seminars in Avian and Exotic Pet Medicine* **12**, 215-227.

Godoy JA, Negro JJ, Hiraldo F, Donazar JA (2004) Phylogeography, genetic structure and diversity in the endangered bearded vulture (*Gypaetus barbatus*, L.) as revealed by mitochondrial DNA. *Mol Ecol* **13**, 371-390.

Griffith SC, Stewart IRK, Dawson DA, Owens IPF, Burke T (1999) Contrasting levels of extra-pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an 'island effect'? *Biological Journal of the Linnean Society* **68**, 303-316.

Haberle SG, David B (2004) Climates of change: human dimensions of Holocene environmental change in low latitudes of the PEPII transect. *Quaternary International* **118-19**, 165-179.

Hanotte O, Zanon C, Pugh A*, et al.* (1994) Isolation and characterization of microsatellite loci in a passerine bird - the Reed Bunting *Emberiza schoeniclus*. *Molecular Ecology* **3**, 529-530.

Harpending HC (1994) Signature of ancient population-growth in a low-resolution mitochondrial-DNA mismatch distribution. *Human Biology* **66**, 591-600.

Hedrick P, Kalinowski S (2000) Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics* **31**, 139-162.

Hedrick PW, Kim TJ, Parker KM (2001) Parasite resistance and genetic variation in the endangered Gila topminow. *Animal Conservation*.

Hellgren O (2005) The occurrence of haemosporidian parasites in the Fennoscandian bluethroat (Luscinia svecica) population. *Journal of Ornithology* **146**, 55-60.

Hellgren O, Waldenstr, m J, Bensch S (2004) A new PCR assay for simultaneous studies of *Leucocytozoon*, *Plasmodium*, and *Haemoproteus* from avian blood. *Journal of Parasitology* **90**, 797-802.

Heslewood MM, Elphinstone MS, Tidemann SC, Baverstock PR (1998) Myoglobin intron variation in the Gouldian Finch *Erythrura gouldiae* assessed by temperature gradient gel electrophoresis. *Electrophoresis* **19**, 142-151.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **359**, 183-195.

Holsinger KE, Lewis PO (2003) HICKORY v.1.0 http://darwin.eeb.uconn.edu/hickory/hickory.html.

Holsinger KE, Wallace LE (2004) Bayesian approaches for the analysis of population genetic structure: an example from *Platanthera leucophaea* (Orchidaceae). *Molecular Ecology* **13**, 887-894.

Hughes AL, Nei M (1988) Nucleotide substitution at Major Histocompatibility Complex Class I loci evidence for overdominant selection. *Nature* **335**, 167-170.

IUCN (2006) *IUCN Red List of Threatened Species.* *www.iucnredlist.org*. *Downloaded on 27 April 2007.*

IUCN (2007) 2007 IUCN Red List of Threatened Species. <http://www.iucnredlist.org>. Downloaded on 21 January 2008.

Jarvi SI, Tarr CL, McIntosh CE, Atkinson CT, Fleischer RC (2004) Natural selection of the major histocompatibility complex (Mhc) in Hawaiian honeycreepers (Drepanidinae). *Molecular Ecology* **13**, 2157-2168.

Jennings WB, Edwards SV (2005) Speciational history of Australian grass finches (Poephila) inferred from thirty gene trees. *Evolution* **59**, 2033-2047.

Johnson CN (2006) *Australia's Mammal Extinctions: A 50 000 year history* Cambridge University Press, Melbourne, VIC.

Johnson JA, Toepfer JE, Dunn PO (2003) Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. *Mol Ecol* **12**, 3335-3347.

Joseph L, Wilke T, Ten Have J, Chesser RT (2006) Implications of mitochondrial DNA polyphyly in two ecologically undifferentiated but morphologically distinct migratory birds, the masked and white-browed woodswallows Artamus spp. of inland Australia. *Journal of Avian Biology* **37**, 625-636.

Judo MSB, Wedel AB, Wilson C (1998) Stimulation and suppression of PCR-mediated recombination. *Nucleic Acids Research* **26**, 1819-1825.

Keller LF, Jeffery KJ, Arcese P*, et al.* (2001) Immigration and the ephemerality of a natural population bottleneck: evidence from molecular markers. *Proceedings of the Royal Society of London Series B : Biological sciences* **268**, 1387-1394.

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. *Trends in Ecology and Evolution* **17**, 230-241.

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725-738.

Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676-679.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW (2006) Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution* **23**, 1891-1901.

Lambert DM, Ritchie PA, Millar CD*, et al.* (2002) Rates of evolution in ancient DNA from Adelie penguins. *Science* **295**, 2270-2273.

Lande R (1988) Genetics and demography in biological conservation. *Science* **241**, 1455-1460.

Lento GM, Baker CS, David V*, et al.* (2003) Automated single-strand conformation polymorphism reveals low diversity of a Major Histocompatibility Complex Class II gene in the threatened New Zealand sea lion. *Molecular Ecology Notes* **3**, 346-349.

Long JC (1986) The allelic correlation structure of Gainj-speaking and Kalam-speaking people. I. The estimation and interpretation of Wright's F-statistics. *Genetics* **112**, 629-647.

Luikart G, Cornuet JM (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology* **12**, 228-237.

Martinez JG, Soler JJ, Soler M, Moller AP, Burke T (1999) Comparative population structure and gene flow of a brood parasite, the great spotted cuckoo (*Clamator glandarius*), and its primary host, the magpie (*Pica pica*). *Evolution* **53**, 269-278.

Matthee CA, Robinson TJ (1999) Mitochondrial DNA population structure of roan and sable antelope: implications for the translocation and conservation of the species. *Molecular Ecology* **8**, 227-238.

May RM (1990) Taxonomy as destiny. *Nature* **347**, 129-130.

McInerney JO (1998) GCUA: general codon usage analysis. *Bioinformatics* **14**, 372-373.

Merilä J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology* **14**, 892-903.

Meyer-Lucht Y, Sommer S (2005) MHC diversity and the association to nematode parasitism in the yellow-necked mouse (*Apodemus flavicollis*). *Molecular Ecology* **14**, 2233-2243.

Miller HC, Lambert DM (2004) Gene duplication and gene conversion in class II MHC genes of New Zealand robins (Petroicidae). *Immunogenetics* **56**, 178-191.

Mindell DP, Sorenson MD, Dimcheff DE (1998) Multiple independent origins of mitochondrial gene order in birds. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 10693-10697.

Moritz C (1994a) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* **3**, 401-411.

Moritz C (1994b) Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology and Evolution* **9**, 373-375.

Moritz C (1995) Uses of molecular phylogenies for conservation. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **349**, 113-118.

Moritz C, Faith DP (1998) Comparative phylogeography and the identification of genetically divergent areas for conservation. *Molecular Ecology* **7**, 419-429.

Mossman JA, Birkhead TR, Slate JON (2006) The whole mitochondrial genome sequence of the zebra finch (Taeniopygia guttata). *Molecular Ecology Notes* **6**, 1222-1227.

Mulligan CJ, Kitchen A, Miyamoto MM (2006) Comment on "Population Size Does Not Influence Mitochondrial Genetic Diversity in Animals". *Science* **314**, 1390a-.

Negro JJ, Torres MJ (1999) Genetic variability and differentiation of two bearded vulture *Gypaetus barbatus* populations and implications for reintroduction projects. *Biological Conservation* **87**, 249-254.

Nei M (1987) *Molecular Evolutionary Genetics* Columbia University Press, New York, USA.

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**, 418-426.

Nevard T, Goulding W, Kearney F*, et al.* (2002) Poposal to Reintroduce Gouldian Finches to the Wild in North-east Queensland, Cairns, QLD.

Nylander JAA (2004) MrModeltest v2. *Program distributed by the author. Evolutionary Biology Centre, Uppsala University.*

O'Malley C (2006) *Recovery Plan for the Gouldian Finch (Erythrura gouldiae)* WWF-Australia and Parks and Wildlife NT, Sydney and Palmerston.

Ohta T (1991) Role of diversifying selection and gene conversion in evolution of major histocompatibility complex loci. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 6716-6720.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research Cambridge* **22**, 201-204.

Olsson M, Madsen T, Wapstra E*, et al.* (2005) MHC, health, color, and reproductive success in sand lizards. *Behavioral Ecology and Sociobiology* **58**, 289-294.

Page RDM (1996) TreeView: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357-358.

Piertney SB, Marquiss M, Summers R (1998) Characterization of tetranucleotide microsatellite markers in the Scottish crossbill (*Loxia scotica*). *Molecular Ecology* **7**, 1261-1263.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**, 502-503.

Prebble M, Sim R, Finn J, Fink D (2005) A Holocene pollen and diatom record from Vanderlin Island, Gulf of Carpentaria, lowland tropical Australia. *Quaternary Research* **64**, 357-371.

Primmer CR, Moller AP, Ellegren H (1996) A wide-range survey of cross-species microsatellite amplification in birds. *Molecular Ecology* **5**, 365-378.

Pritchard JK, Stephens M, Donelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Raymond M, Rousset F (1995) Genepop (Version-1.2) - Population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* **55**, 1095-1103.

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. *Conservation Biology* **17**, 230-237.

Richard FA, Sehgal RNM, Jones HI, Smith TB (2002) A comparative analysis of PCR-based detection methods for avian malaria. *Journal of Parasitology* **88**, 819-822.

Richardson DS, Jury FL, Dawson DA*, et al.* (2000) Fifty Seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. *Molecular Ecology* **9**, 2226-2231.

Ritchie PA, Millar CD, Gibb GC, Baroni C, Lambert DM (2004) Ancient DNA enables timing of the Pleistocene origin and Holocene expansion of two Adelie penguin lineages in Antarctica. *Molecular Biology and Evolution* **21**, 240-248.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic-differences. *Molecular Biology and Evolution* **9**, 552-569.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496-2497.

Sato A, Mayer WE, Tichy H*, et al.* (2001) Evolution of Mhc class IIB genes in Darwin's finches and their closest relatives: birth of a new gene. *Immunogenetics* **53**, 792-801.

Sawyer SA (1989) Statistical tests for detecting gene conversion. *Molecular Biology and Evolution* **6**, 526-538.

Sawyer SA (1999) GENECONV: A computer package for the statistical detection of gene conversion. *Distributed by the author, Department of Mathematics, Washington University in St. Louis, available at* http://www.math.wustl.edu/~sawyer.

Schad J, Sommer S, Ganzhorn JU (2004) MHC variability of a small lemur in the littoral forest fragments of southeastern Madagascar. *Conservation Genetics* **5**, 299-309.

Schaschl H, Suchentrunk F, Hammer S, Goodman SJ (2005) Recombination and the origin of sequence diversity in the DRB MHC class II locus in chamois (*Rupicapra spp*.). *Immunogenetics* **57**, 108-115.

Schaschl H, Wandeler P, Suchentrunk F, Obexer-Ruff G, Goodman SJ (2006) Selection and recombination drive the evolution of MHC class II DRB diversity in ungulates. *Heredity* **97**, 427-437.

Scheffler K, Martin DP, Seoighe C (2006) Robust inference of positive selection from recombining coding sequences. *Bioinformatics* **22**, 2493-2499.

Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**, 502-504.

Schulze ED, Williams RJ, Farquhar GD*, et al.* (1998) Carbon and nitrogen isotope discrimination and nitrogen nutrition of trees along a rainfall gradient in northern Australia. *Functional Plant Biology* **25**, 413-425.

Sefc KM, Payne RB, Sorenson MD (2001) Characterization of microsatellite loci in village indigobirds *Vidua chalybeata* and cross-species amplification in estrildid and ploceid finches. *Molecular Ecology Notes* **1**, 252-254.

Sehgal RNM, Jones HI, Smith TB (2001) Host specificity and incidence of Trypanosoma in some African rainforest birds: a molecular approach. *Molecular Ecology* **10**, 2319-2327.

Sehgal RNM, Jones HI, Smith TB (2005) Molecular evidence for host specificity of parasitic nematode microfilariae in some African rainforest birds. *Molecular Ecology* **14**, 3977-3988.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629.

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* **69**, 82-90.

Shulmeister J (1999) Australasian evidence for mid-Holocene climate change implies precessional control of Walker Circulation in the Pacific. *Quaternary International* **57-8**, 81-91.

Shulmeister J, Lees BG (1995) Pollen evidence from tropical Australia for the onset of an ENSO-dominated climate at c. 4000 BP. *Holocene* **5**, 10-18.

Šimková A, Ottová E, Morand S (2006) MHC variability, life-traits and parasite diversity of European cyprinid fish. *Evolutionary Ecology* **20**, 465-477.

Slatkin M (1986) Interchromosomal biased gene conversion, mutatiuon and selection in a multigene family. *Genetics* **112**, 681-698.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* **236**, 787-792.

Sommer S (2005) The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology* **2**, 16.

Sorenson MD, Payne RB (2001) A single ancient origin of brood parasitism in African finches: implications for host-parasite evolution. *Evolution* **55**, 2550-2567.

Southern H (1945) Polymorphism in *Poephila gouldiae*. *Journal of Genetics* **47**, 51-57.

Spaulding AW, Mock KE, Schroeder MA, Warheit KI (2006) Recency, range expansion, and unsorted lineages: implications for interpreting neutral genetic variation in the sharp-tailed grouse (Tympanuchus phasianellus). *Molecular Ecology* **15**, 2317-2332.

Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **64**, 583-616.

Sunnucks P, Wilson ACC, Beheregaray LB, Zenger K (2000) SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology* **9**, 1699-1710.

Swofford DL (1998) *PAUP*: phylogenetic analysis using parsimony (*and other methods)* Sinauer Associates, Sunderland, Mass., USA.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595.

Takahata N, Nei M (1990) Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**, 967-978.

Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP (2003) A plea for DNA taxonomy. *Trends in Ecology & Evolution* **18**, 70-74.

Tidemann SC, Boyden J, Elvish R, Elvish J, Ogorman B (1992a) Comparison of the breeding sites and habitat of two hole-nesting estrildid finches, one endangered, in northern Australia. *Journal of Tropical Ecology* **8**, 373-388.

Tidemann SC, Lawson C, Elvish R, Boyden J, Elvish J (1999) Breeding biology of the Gouldian Finch *Erythrura gouldiae*, an endangered finch of northern Australia. *Emu* **99**, 191-199.

Tidemann SC, Mcorist S, Woinarski JCZ, Freeland WJ (1992b) Parasitism of wild Gouldian Finches (*Erythrura gouldiae*) by the air-sac mite *Sternostoma tracheacolum*. *Journal of Wildlife Diseases* **28**, 80-84.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

Vincek V, Ohuigin C, Satta Y*, et al.* (1997) How large was the founding population of Darwins Finches. *Proceedings of the Royal Society of London - Series B: Biological Sciences* **264**, 111-118.

Wandeler P, Hoeck PEA, Keller LF (2007) Back to the future: museum specimens in population genetics. *Trends in Ecology & Evolution* **22**, 634-642.

Wang ZS, Farmer K, Hill GE, Edwards SV (2006) A cDNA macroarray approach to parasite-induced gene expression changes in a songbird host: genetic response of house finches to experimental infection by Mycoplasma gallisepticum. *Molecular Ecology* **15**, 1263-1273.

Wares J, Barber P, Ross-Ibarra J, Sotka E, Toonen R (2006) Mitochondrial DNA and Population Size. *Science* **314**, 1388-1390.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**, 1358-1370.

Westerdahl H, Hansson B, Bensch S, Hasselquist D (2004a) Between-year variation of MHC allele frequencies in great reed warblers: selection or drift? *Journal of Evolutionary Biology* **17**, 485-492.

Westerdahl H, Waldenström J, Hansson B*, et al.* (2005) Associations between malaria and MHC genes in a migratory songbird. *Proceedings of the Royal Society B: Biological Sciences* **272**, 1511-1518.

Westerdahl H, Wittzell H, Von Schantz T, Bensch S (2004b) MHC class I typing in a songbird with numerous loci and high polymorphism using motif-specific PCR and DGGE. *Heredity* **92**, 534-542.

Wilson DJ, McVean G (2006) Estimating diversifying selection and functional constraint in the presence of recombination. *Genetics* **172**, 1411-1425.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177-1191.

Woinarski JCZ, Williams RJ, Price O, Rankmore B (2005) Landscapes without boundaries: wildlife and their environments in northern Australia. *Wildlife Research* **32**, 377-388.

Woodroffe CD (1993) Late quaternary evolution of coastal and lowland riverine plains of Southeast-Asia and northern Australia - an overview. *Sedimentary Geology* **83**, 163-175.

Woodroffe CD, Mulrennan ME, Chappell J (1993) Estuarine infill and coastal progradation, southern Van-Diemen Gulf, northern Australia. *Sedimentary Geology* **83**, 257-275.

Worthington-Wilmer J, Hall L, Toop J (1994) Extreme Population Structuring in the Threatened Ghost Bat, Macroderma gigas: Evidence from Mitochondrial DNA. *Proceedings of the Royal Society B: Biological Sciences* **257**, 193-198.

Wüster W, Ferguson JE, Quijada-Mascarenas JA*, et al.* (2005) Tracing an invasion: landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes : Viperidae : *Crotalus durissus*). *Molecular Ecology* **14**, 1095-1108.

Yang Z, Nielsen R, Goldman N, Pedersen A-MK (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**, 431-449.

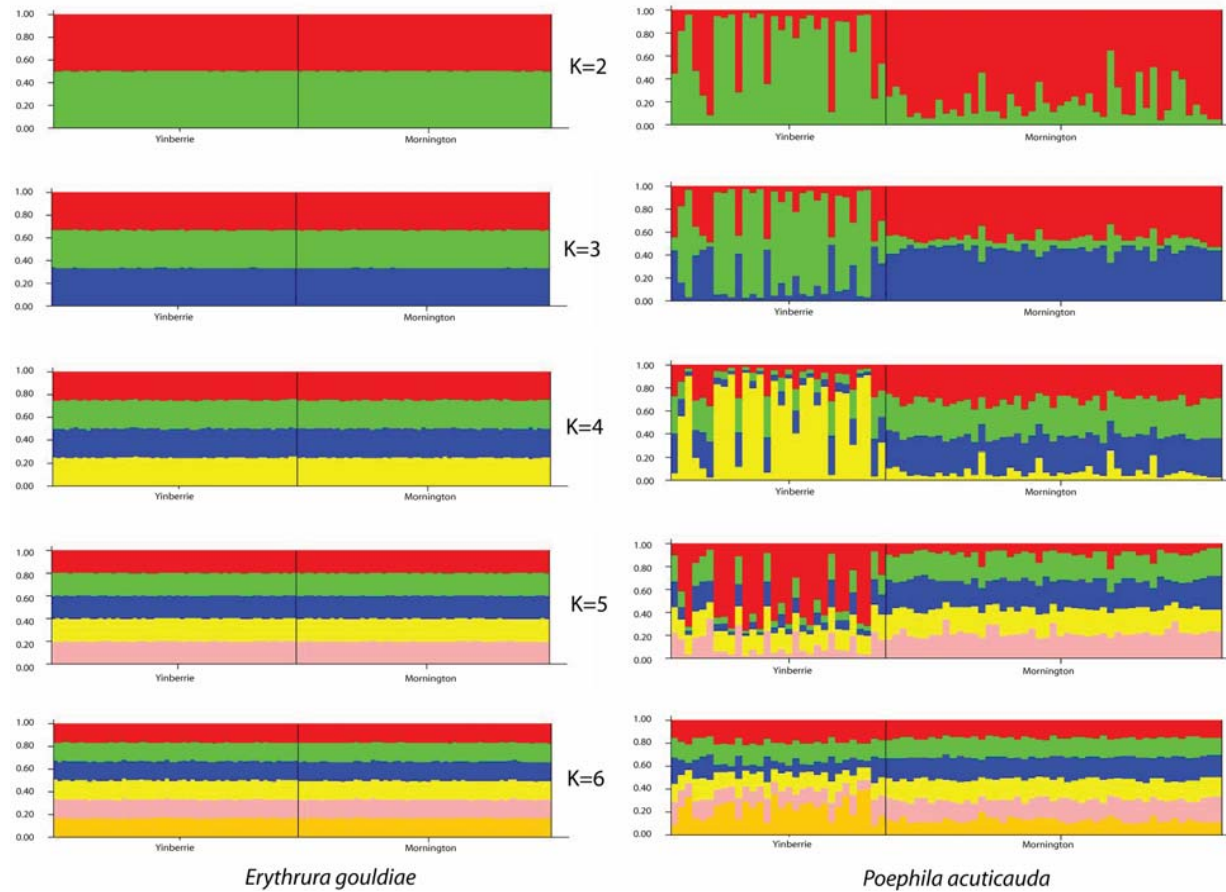
Yodogawa Y, Nishiumi I, Saito D, Okanoya K (2003) Characterization of eight polymorphic microsatellite loci from the Bengalese finch (*Lonchura striata* var. *Domestica*). *Molecular Ecology Notes* **3**, 183-185.

APPENDICES

Appendix I Posterior means and 95% credible intervals (CI) of panmictic heterozygosity (Hs), *f* and θ^{II} for *E. gouldiae* and *P. acuticauda* under three different models.

		E. gouldiae				*P. acuticauda*		
Model	Parameter	Yinberrie + Mornington	Yinberrie	Mornington	Mareeba	Yinberrie + Mornington	Yinberrie	Mornington
Full	Hs	0.7531 (0.7371 - 0.7687)	0.7600 (0.7386 - 0.7813)	0.7463 (0.7238 - 0.7681)		0.7987 (0.7778 - 0.8173)	0.7948 (0.7626 - 0.8221)	0.8026 (0.7777 - 0.8250)
	f	0.0535 (0.0271 - 0.0853)				0.0838 (0.0437 - 0.1297)		
	θ^{II}	0.0957 (0.0795, 0.1132)				0.1244 (0.0983 - 0.1547)		
f = 0	Hs	0.7541 (0.7394 - 0.7693)	0.7608 (0.7405 - 0.7817)	0.7475 (0.7264 - 0.7686)		0.7931 (0.7742 - 0.8106)	0.7905 (0.7607 - 0.8159)	0.7957 (0.7712 - 0.8176)
	θ^{II}	0.0800 (0.0663 - 0.0959)				0.0595 (0.0414, 0.0927)		
$\theta = 0$	Hs	0.8016 (0.8016 - 0.8016)			0.6627 (0.6254 - 0.6965)	0.8082 (0.7873 - 0.8224)		
	f	0.1382 (0.1047 - 0.1744)			0.0847 (0.0308 - 0.1578)	0.1757 (0.1306 - 0.2239)		

Appendix II Assignment of E. gouldiae and P. acuticauda individuals to populations under hypothetical scenarios of two to six populations (K). Vertical bars represent individual birds. The proportion of colours in the bars corresponds to the probability of an individual to be assigned to the respective hypothetical population. Individuals from different sampling sites are separated by a vertical black line.



Appendix III Amino-acid sequences of MHC class IIβ haplotypes. A) *Erythrura. gouldiae*; B) *Poephila acuticauda*. Dots indicate sequence identity with the consensus sequence. Amino-acid positions identified as being under positive selection are indicated with an asterisk. Predicted peptide binding amino-acids in human DRB1 genes (Brown *et al.* 1993)are indicated with a cross.

A)

```
                       10        20        30        40        50
               ----|----|----|----|----|----|----|----|----|----|----|--
CONSENSUS      TEKVRFVERYIYNREQFMMFDSDVGVYEGFTPLGEKNARRWNNNPEIMEYARTAVDW
Ergo02.01      .....YA..........V.................R..N.F.S......N.......
Ergo02.02      ......................................K.......W.........N
Ergo02.03      ....K..H.Q....QTYA........F.....Y...Q..HR...Q...DHV.GE..R
Ergo02.04      .........S............................K.......W.........N
Ergo03.01      .........H..........N...............V..HR.....R..........
Ergo03.02      ..R......H..........................V..HR.....R..........
Ergo03.03      .........H............................K.V................
Ergo03.04      .......A.H............................K.V................
Ergo03.05      .....S...H....................A.......K.V................
Ergo03.06      .........H...................L........K.V................
Ergo04.01      ....K.........QTYA..............Y...V..H....Q....SR......
Ergo04.02      .........H................F.....Y...Q..Y......L...K.GQ..N
Ergo04.03      ...................L............F...Q..HR.....R..........
Ergo04.04      ................................F...Q..HR.....R..........
Ergo04.05      .....L...F....L.YS.......E.V....R...V.QK...D..R...R......
Ergo04.06      .........H................F.....Y...Q..H......L...K.GQ..N
Ergo05.01      .........H......................F........................
Ergo05.02      .........H......................F........................
Ergo05.03      ....K.........QTYA..............F...V..Y......W...K.GQ..N
Ergo05.04      .........H......................F........................
Ergo08.01      ..........................F....SF.....K.V.....W...R......
Ergo08.02      ....K.........QTYA........F....SF...V..Y......W...K.GQ..N
Ergo08.03      ..........................F....SF.....K.V.....W...R......
Ergo08.04      ..............QTYA........F....SF...V..Y......W...K.GQ..N
Ergo08.05      .......G.H............................K.V................
Ergo08.06      ..............QTYA........F....SF...V..Y......W...K.GQ..N
Ergo09.01      ....K..H.Q......I......E..F.....I...Q..Y......R..........
Ergo09.02      ....K..H.Q......I.........F.....I...Q.GY......R..........
Ergo09.03      .........H............................K.V................
Ergo09.04      .........H............................K.V................
Ergo11.01      .....L...F....L.YS.......E.V....R...V.QK...D..R...R......
Ergo11.02      ......M.......Q..L........HV....Y...Q.QYR........HK.....R
Ergo11.03      .........H....QTYA........F.....F...V..Y..........K.GQ..N
Ergo11.04      .........L......YL.......EFV....H.....K.L..D..WL..K.GQ..T
Ergo11.05      ..............L.Y........E.V....F......QA.....W..HK.GQM.N
Ergo11.06      ..R....L......L.HL.......E.V....F...Q..Y...D..R.........R
Ergo11.07      ....K.........QTYA........F.........V..H.......I..K.GQ..N
Ergo12.01      ....K.........QTYA..............I...V..H.........SR......
Ergo12.02      .........................................................
Ergo12.03      .........H...............................................
Ergo12.04      ....................................V..H......W...R......
Ergo12.05      ....................................V..H.........SR......
Ergo14.01      ....K.........QTYA..............I...V..H......R..........
               ----|----|----|----|----|----|----|----|----|----|----|--
                       10        20        30        40        50
                        *       *        *     *     ***         *     *
                      + + +    ++        +        +   ++      + ++  +
```

Appendix III (Continued)

A) (Continued)

```
                        10        20        30        40        50
                ----|----|----|----|----|----|----|----|----|----|----|--
CONSENSUS       TEKVRFVERYIYNREQFMMFDSDVGVYEGFTPLGEKNARRWNNNPEIMEYARTAVDW
Ergo14.02       .................................................D.......
Ergo14.03       ..........T..............................................
Ergo15.01       ....K.........QTYA........F.....Y...Q..Y......L...K.GQ..N
Ergo15.02       ................................FR..V..HR.....WT..M..E..T
Ergo15.03       ....K.........QTYA........L.....Y...Q..Y......L...K.GQ..N
Ergo15.04       ....K.........QTYA.S............I...A..H.........SR......
Ergo15.05       .......H.H......................F...V..HR.....W...M..E..T
Ergo15.06       ................L...............F...V..HR.....W...M..E..T
Ergo15.07       .......H.H......I.....................K.V.D...R..........
Ergo15.08       .......H.H......I.....................K.V.....R..........
Ergo16.01       .........H......................F...V..HR.........K.GQ..N
Ergo16.02       .........H......................F...V..H..........K.GQ..N
Ergo16.03       .........H......................F...V.LH..........K.GQ..N
Ergo16.04       ..........................C.....F.............W..........
Ergo16.05       .........H................F.........V..H....Q.R...R......
Ergo16.06       ....K.........QTYA..............F...V..H..........K.GQ..N
Ergo18.01       ....K.........QTYA..............Y...V..H.........SR......
Ergo18.02       ....K.........QTYA..............Y...V..H.........SR....G.
Ergo18.03       ....K.........QTYA..............F...V..H......W..........
Ergo18.04       .........H......I.....................K.V.....R..........
Ergo18.05       .........H......................F.....K.L.....R...V......
Ergo18.06       ....K.........QTYA..............Y...V..H.........SR......
Ergo18.07       .........H......................F.....K.L.....R...V......
Ergo18.08       .........H......I.....................K.VS....R..........
Ergo18.09       .........R......I.....................K.V.....R..........
Ergo19.01       ....K..H.Q......I.........F.....I...Q..Y......R..........
Ergo19.02       ....K.........QTYA........F.....F...V..H......R.......L..
Ergo19.03       .........H............................K.V................
Ergo19.04       .........H......S.....................K.V................
Ergo20.01       ........................................L................
Ergo20.02       ....K.........QTYA........F....SF...V..Y......W...K.GQ..N
Ergo20.03       ....K.........QTYA..............Y...V..H......R..........
Ergo20.04       ....K.........RTYA..............Y...V..H......R..........
Ergo20.05       .......H.H......I...................S.K.V.....R..........
Ergo20.06       ....K.........QTYP..............Y.V.V..H......R..........
                ----|----|----|----|----|----|----|----|----|----|----|--
                        10        20        30        40        50
                         *       *        *     *     ***         *     *
                       + + +    ++        +        +   ++      + ++   +
```

Appendix III (Continued)

B)

```
                    10        20        30        40        50
            ----|----|----|----|----|----|----|----|----|----|----|--
CONSENSUS   TEKVRFVERYIYNREQYLMFDSDVGEYVGFTPYGEKNAKRWNSDPEWMENRRTAVDW
Ergo04.05   .....L...F....L..S..............R...V.QK......R..........
Poac03.01   ..R.............F.........H.........Q..YR..N..I..HK.....R
Poac03.02   .......H.R.........YE..........LF...Q.RYK....K...YT......
Poac03.03   .......D........V...............L.......F..N..GLQYK.A...R
Poac03.04   .......Q......Q...........H........EF.RQA..N.....QQ..T...
Poac03.05   ..R....D.H......V..L............L.......F..N..GLQYK.A...R
Poac03.06   ..W......L......F.........F....AF..R....F..N..I......V..T
Poac03.07   .......Q........F.........H.....L.......L........YI.....T
Poac03.08   .....Y...........V..............H.GR....F..N..GLQYK.A...R
Poac03.09   M......Q........F.........H.....L.......L........YI.....T
Poac04.01   .......Q.L.......M............I.F...F.G.V..N.....YI..V..T
Poac04.02   .....YLQ.L.......M..............F...F.G.V..N.....YI..V..T
Poac04.03   ..R....Q.L......DV.Y...................N.........YI.....T
Poac04.04   .......Q.L.......M......E.......F...F.G.V..N.....YI..V..T
Poac04.05   .....YLQ.........M........H........EF.RQA........QQ..T...
Poac04.06   .......Q.L.......M........H........EF.RQA........QQ..T...
Poac04.07   .....YLQ.........M........H........EF.RQA........QQ..T...
Poac04.08   .......N........FV...........LNR....W.RER........ET.....T
Poac05.01   .....Y.Q.R....Q..V.Y......H.........Q.RYR.....I..R.......
Poac05.02   ..R....D.C.......A........F.........Q.RYR.....I..R.......
Poac05.03   .....C.Q.R....Q..V.Y......H.........Q.RYR.....I..R.......
Poac05.04   .....Y.Q.H.........................R..........Y..YK.....R
Poac05.05   ..R....D.........A........F...................I..RK.AE..I
Poac05.06   ..R....L........F...............L..E....L..N..L..S..AE..R
Poac06.01   .....Y.Q.R....Q..V.Y......H.........Q.RYR.....I..R.......
Poac06.02   .......Q...H..Q..A........F............N...N..I..YI.....T
Poac06.03   .....YR..L.........................R....L.G..........S..R
Poac06.04   .....YR..L..........N..............R....L.G..........S...
Poac06.05   ..............Q..A........F................N...L.HK.....R
Poac06.06   .....YR..L.........................R....L.G..........S...
Poac06.07   .....YR..L.........................R....L.G..........S...
Poac06.08   .......H.R.........YE..........LF...Q.RYK....K...YT......
Poac06.09   .....Y..........DV..............A...W.RD.........ET....GT
Poac06.10   .G.....A........DV........H.....A...W..E.........ET......
Poac06.11   .....Y..........FV..............H..R....L........S...E..R
Poac06.12   .................A........F................N..F..HK.....R
Poac07.01   ..R....D.H.......V..................F.R.L........YT.....T
Poac07.02   ...........C..Q.F........LH.........Q.RYR..N..I..HK.....R
Poac07.03   .....Y.Q.H....................................Y..YK.....R
Poac07.04   .......D........F.IY..........N.....W.RDR....K...YQ..S...
Poac07.05   .....Y.Q.H.........................EF.R.L........YT.....T
Poac09.01   ..R..Y...H......................F.............Y..YK.....R
Poac09.02   ..R....Q........DV........H.....A...W..E.........YI.....T
Poac09.03   .......D.H......V...............L.......F..N..GLQYK.A...R
Poac09.04   ..R...A..........V........F................N...L.HK.....R
Poac09.05   .......D.H......V...............L.......F..N..GLQYK.A...R
Poac09.06   ..............Q..A........F.....A....T........L..........
            ----|----|----|----|----|----|----|----|----|----|----|--
                    10        20        30        40        50
                     *                      *  **  **     *  **  *
                   + + +    ++        +        +   ++      + ++  +
```

Appendix III (Continued)

B) (Continued)

```
                    10        20        30        40        50
            ----|----|----|----|----|----|----|----|----|----|----|--
CONSENSUS   TEKVRFVERYIYNREQYLMFDSDVGEYVGFTPYGEKNAKRWNSDPEWMENRRTAVDW
Poac09.07   ..R..............V.................EF.RQA........HK.GQ..N
Poac09.08   .......D.H......V...............L.......F..N..GLQYK.A...R
Poac09.09   ..R..............V.................EY.RQA........HK.GQ..N
Poac10.01   ..R..YR..L.................E.......R....L.G..............
Poac10.02   ..R..YR..L.........................R....L.G..............
Poac10.03   ..R..YR..L....................................Y..YK.....R
Poac10.04   .......D.H......V...............L.......I..N..GLQYK.A...R
Poac10.05   ..R.GYR..L.........................R....L.G..............
Poac12.01   .......Q.L.......M...C..........F...F.G.V.GN.....YI..V..T
Poac12.02   .......Q.........V........H.........C.RQA.....F..S......T
Poac12.03   ................F.........H.........Q..YR..N..I..HK.....R
Poac12.04   .....Y..........DV.Y............A...W.RD.........YA.....T
Poac12.05   ..............Q.F........VH.........Q..YR.NN..I..HK.....R
Poac12.06   .......Q.L.......M..............F...F.G.V..N.....YI..V..T
Poac13.01   .....Y..........................L.......L.....L..QQ..T...
Poac13.02   .....Y.Q.R....Q.....................Q.RYR..N..I..R.......
Poac13.03   .....Y..........F...............L..E....F..N.....QQ..T...
Poac13.04   .......Q........F.........H.....L..E....L.....L..Y..AE..R
Poac13.05   .....YA.........F...............L..E....F..N.....QQ..T...
Poac13.06   .......Q.........V.Y.....M.L...RL..E....F..N..L..RKQ....R
Poac14.01   .G.....A........DV........H.....A...W..E.........ET......
Poac14.02   .G.....A........DV........H.....A...W..E.........ET......
Poac14.03   .................V........H........EF.RQA..N.....QQ..T...
Poac14.04   ..R..Y.Q.R....Q.....................Q.RYR..N..I..R.......
Poac15.01   .......Q......Q.DV.Y......H.....F...F.R.L..N..F..S.......
Poac15.02   ..R.............F.........H.........Q.RYR..N..I..HK.....T
Poac15.03   ..............Q.F...............L..G....L..N....KH..GQ..N
Poac15.04   ..R..Y.N.........A........F........IQ.RY......F..RK.AE..R
Poac16.01   .......Q.R....Q.LV........F.....F..EW..E...N..F..S......R
Poac16.02   .......Q........DV........H.....A...W..E...N..F..S.......
Poac16.03   .......Q........DV........H.....A...W..E...N..F..S.......
Poac16.04   .......Q.R....Q.LV........F.....F..E...SL.N...Y...T.....T
Poac16.05   .......D.H......V...............L.......F.....GLQYK.A...R
Poac16.06   .......D.........A........F...................I..RK.AE..I
Poac16.07   .....Y...H......H...............F.............Y..YT.....T
Poac17.01   .....YR..L.........................R....L.G..............
Poac17.02   .....YR..L.........................R....L.G..............
Poac17.03   .....Y..........FV..............H..R....L........S...E..R
Poac17.04   .G...............V........H.....A...W..E.........YQ..S..T
Poac19.01   .......V.H.......V..............H...F.R.L.........TQ.....
Poac19.02   .......V.H.......V..............H...F.R.L......T..TQ.....
Poac19.03   .......V.H.......V............A.H...F.R.L.........TQ.....
Poac19.04   .......Q.........V........H.........C.RQA.....F..S......T
Poac19.05   .......Q.........V.Y.....M.L....L..E....F..N..L..RKQ....R
Poac19.06   .......V.H.......V...........LNR....W.RDR........YA.....T
Poac19.07   ....G..N........FV...........LNR....W.RDR........YA.....T
Poac19.08   .......N........FV...........LNR....W.RDR........YA.....T
            ----|----|----|----|----|----|----|----|----|----|----|--
                    10        20        30        40        50
                     *                      *  **  **     *  **  *
                   + + +    ++        +        +   ++      + ++  +
```

1