The place of functional genetics and population genomics within conservation genetics

Vincenzo Pasquale Repaci,
BSc (Hons)

Presented for the degree of

Doctor of Philosophy

Department of Biological Sciences
Faculty of Science and Engineering
Macquarie University
New South Wales, Australia

Submitted : December 2015

<The place of functional genetics and population genomics within conservation genetics>

Vincenzo Pasquale Repaci, BSc. Hon

Conservation Genetics Lab

Department of Biology Sciences, Macquarie University

Date Submitted: 31 December, 2015

- a title page giving the title of the thesis in full, the names and degrees of the candidate, the name of the organisation, institute or laboratory in which the research was carried out (if applicable), the name of the Department of the University associated with the work and the date when submitted or re-submitted for the degree;
- a table of contents;
- a summary of approximately 200 words for a Master degree or 300 words for a Doctorate;
- a statement signed by the candidate to the effect that the work has not been submitted for a higher degree to any other university or institution. The candidate shall also indicate in the thesis the sources of information used and the extent to which the work of others has been utilised, and specify if Ethics Committee approval has been obtained (and the protocol number);
- acknowledgments (if applicable);
- pages will be numbered consecutively.

Dedication & Thanks

You, me, or nobody is gonna hit as hard as life. But it ain't how hard you hit; it's about how hard you can get hit, and keep moving forward. How much you can take, and keep moving forward.

That's how winning is done.

Rocky Balboa

This work is primarily dedicated to two people: the first is a boy who, in a small town in Italy, left school, and, with it, his dream of becoming an engineer, to work on a farm to keep his brothers in school. The second is a girl who, in a classroom in inner Sydney, raised her hand when asked who wanted to matriculate from high School; she was told to put her hand down because her kind would never finish school. She did finish high school and has never turned away from a challenge or backed down from a bully.

My education was a gift from my parents, who prioritized and celebrated learning, who watched David Attenborough with me as a child and came on my school excursions to the zoo. Whatever success I have achieved, or will achieve, is the result of the opportunities afforded to me because they got up at midnight to go work, because they spent countless hours helping to make my handwriting legible and because they made sure that their children could do the things they missed out on. You earned this far more than I did. Thank you.

When I count my many blessings, I will always remember and be thankful for my supervisors, Associate Professor Jenny Donald and Associate Professor Adam Stow. In a difficult project you went above and beyond the call of duty continuously. I continue to be amazed not only by your intellect and insight but also by your patience, your kindness and your character. As educators, mentors and TBGBTLDTH (the big giant brain that lives down the hallway) I will forever be in your debt.

I am also grateful to my fellow students, both in the Conservation Genetics Lab and beyond, who were a joy to work with, to tease and to drink coffee with. To Shan, Prof. Hotstuff, Dr. Cucumbers,

SuperTed, C-bomb, H-Dogg, Marisa, Kat, Team Jacob, Johno, Liette , Miranda, Ben, Jackets and Teddy Bear Pete, who I count as friends as well as colleagues, I owe an enormous debt of gratitude for making the better part of the last decade fun. Further, sincere thank you's are owed to those who have been my professional colleagues at the university, especially Sharyon, Marie, Laura, Veronica, Kat, Anne-Marie and Teresa. You have been a joy to work with and learn from, and I will greatly miss working with you.

As a researcher, I make a pretty good teacher, and I am especially grateful to my students who have taught me so very much over the course of this degree. In Claire, Peter, Siobhan and Keira I was privileged to teach, at least for a little while, scientific minds that far and away surpass my own. I would also like to dedicate a portion of this work to Brendan and Sam. Samantha Dickson was the bravest brightest and toughest student I have ever taught. The world is poorer for losing her so early, and her insight and enthusiasm is sorely missed. I learned more from you than you from me.

Finally I must acknowledge, thank and dedicate this work to my soon-to-be wife Laura. My editor, sounding board, emergency data analyst, problem solver, coach and muse. Without you I would not have finished. Thank you for motivating me, managing me and not murdering me. The next chapter will be even better.

The beginning of wisdom is this: Get wisdom.

Though it cost all you have, get understanding.

Cherish her, and she will exalt you;

embrace her, and she will honour you

.

Proverbs 4:7-8

Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis. Where people have assisted with sections of this thesis they have been acknowledged in the acknowledgement section below.

Ethics approval was obtained for procedures in this thesis in accordance with Macquarie University Animal Ethics procedures.

Research Presented in Chapter 2 ARA 2010/038 (letter presented in appendix I)

Research Presented in Chapters 4&5 ARA 2013/015 (letter not presented as I was not part of the project but instead used the publically accessible data produced by the project)

Vincenzo Repaci 31st December 2015

Acknowledgements

Whilst the majority of work presented in this thesis is my own, I would like to acknowledge the contributions of the following people whose time, effort, experience and expertise were vital to the project.

CHAPTER	EXPERIMENTAL DESIGN	SAMPLING	MOLECULAR ANALYSIS	DATA ANALYSIS	WRITING & REVIEW
1	n/a	n/a	n/a	n/a	A/Prof Jenny Donald A/Prof Adam Stow Peter Mahoney
2	A/Prof Adam Stow	Nick Atchison	Kiera Beattie Miranda Christopher	A/Prof Adam Stow	A/Prof Jenny Donald A/Prof Adam Stow Dr. Stephen Hoggard
3	n/a	Claire Burford Sarah Houlahan Dr. Stephen Hoggard	n/a	n/a	A/Prof Jenny Donald A/Prof Adam Stow Peter Mahoney Dr. Katherine McCellan
4	A/Prof Jenny Donald A/Prof Adam Stow	Dr. Stephen Hoggard Dr. Shannon Smith Dr. Kate Umbers Benjamin Oforio Dr. Siobhan Dennison	Martia Holley Dr. Liette Waldron Dr.Phyllis Farmer AGRF	Ben Christopher Katy Breakwell	A/Prof Jenny Donald A/Prof Adam Stow Peter Mahoney
5	A/Prof Jenny Donald A/Prof Adam Stow	Benjamin Oforio	Benjamin Oforio DArT	Jennifer Rowlands Laura O'Laney	A/Prof Jenny Donald A/Prof Adam Stow Peter Mahoney
6	n/a	n/a	n/a	n/a	A/Prof Jenny Donald A/Prof Adam Stow Peter Mahoney

(AGRF denotes Australian Genome Research Facility, DArT denotes Diversity Array Technologies)

I also acknowledge Emeritus Professor Richard Frankham who has been an ongoing sounding board and source of excellent advice and helped to develop my understanding of the concepts contained within the thesis.

Thesis Abstract

This thesis explores questions about genetic markers within the field of conservation genetics. It first engages accepted conservation genetic theory and applies 'traditional' genetic techniques including microsatellite and mtDNA analysis to inform conservation management of the Australian native species mala (Lagorchestes hirsutus). These genetic markers suggest multiple founder effects within the captive population and that captive management may benefit from translocation. However, they are poorly suited to answering other important questons such as the degree of adaptation to captivity. The thesis then investigates the types of genetic markers commonly used in conservation and explores the factors that may affect one alternate marker, the major histocompatibility complex (MHC) with a meta-analysis. It finds the relationship between MHC and microsatellite markers is significant but not simply predictive and that more work may be required in order to understand MHC diversity in threatened species

As MHC may be an ideal genetic marker for a number of conservation scenarios, we then attempt to develop a novel method for genotyping MHC that can be used on a large number of vertebrate species. This method is based on tagged degenerate primers and genotyping by 454 sequencing. During this chapter both the molecular biology and bioinformatics components of early next-generation sequencing techniques are explored. This method fails.

Finally, the thesis explores the emerging field of population genomics and attempts to use a population genomic technique, RAD-TAG sequencing, to look for disruptive selection associated with land clearing in native scincid lizard Cunningham's skink (Egernia cunninghami). This is a question pertinent to wider conservation efforts because of possible associations between disruptive selection and outbreeding depression. Using RAD-TAG methods we were able to find some evidence of disruptive selection occurring at small temporal and spatial scales. This final analysis allows for a reflection on the changes in conservation genetics stemming from the sequencing revolution, which coincided with the period of this thesis, and for suggesting future directions based on the outcomes of these studies.

Contents

Table of Figures

1 General Introduction: The assessment of genetic diversity within Conservation Genetics

1.1 The importance of genetic factors in conservation

The effects of a growing and resource hungry human population have put unsustainable pressure on natural ecosystems. Ecosystems are underpinned by the ecological services provided by resident biotic communities (Balvanera *et al*. 2006). High levels of biodiversity in those communities promote ecosystem security and resilience (Loreau *et al*. 2001; Elmqvist *et al.* 2003). Human impacts are causing a loss of biodiversity on a scale rarely seen in earth's history; this is termed the sixth great extinction (Barnosky *et al.* 2011). This loss is driven by five key threats to global biodiversity, identified by the IUCN as: 1) habitat loss and degradation; 2) invasive alien species; 3) over-exploitation of natural resources; 4) human induced climate change and; 5) pollution and disease (IUCN 2015). In order to ensure the survival of humankind and other species, understanding and conserving biodiversity must become global priorities (IUCN 2015, WHO 2015). This biodiversity includes animal and plant species and the ecosystems they inhabit (IUCN 2015, WHO 2015). Furthermore, genetic diversity is now recognised as part of biodiversity (IUCN 2015). Therefore, conservation aims to maintain not only the diversity of environments and species on the planet, but also the genetic diversity within species. This is, if understatement is permitted, a substantial challenge.

Conserving genetic diversity safeguards long term evolutionary potential. Genetic diversity, the differences in the sequence of DNA at the same location in the genome between individuals (loci), is the raw material on which selection acts to drive evolution (Reed & Frankham 2003; Whitehead & Crawford 2006). The level of genetic diversity within a species may range from zero to extremely high (Hamrick *et al.* 1992; Ward *et al.* 1994; Ellstrand & Roose 1987). The action of selection, differential survival and reproduction due to environmental pressures, causes the frequency of alleles at loci within the population to change (Frankham *et al.* 2002). This change is responsible for the evolutionary changes observed in the fossil record (Wandeler *et al.* 2007), and extant populations (Parmesan 2006). However, if a population does not have genetic variants capable of

surviving the pressures created by the environment, extinction occurs (Bonnell & Selander 1974; Sexton *et al.* 2009). In the context of the threats currently facing biodiversity, this means that if we are unable to preserve genetic diversity many species will be unable to evolve and will face extinction. As a well as this longer term risk, a lack of genetic diversity can also reduce population viability in the immediate future.

Low genetic diversity in the short term is correlated with loss of fitness due to inbreeding depression. Most animal species of conservation concern, which comprise up to 38% of some taxa (IUCN 2015), are diploid, meaning they have two copies of each chromosome and thus each gene. Therefore deleterious recessive mutations do not cause a loss of fitness in heterozygotes. However, in species of conservation concern, where populations are small, genetic diversity is depleted (Frankham *et al.* 2002) and many remaining individuals may be related. This is important because the mating of related individuals produces offspring that are more likely to have alleles which are identical by descent (Fig. 1). Where these alleles are deleterious, the organism's fitness may be reduced. This effect, termed inbreeding depression, and occurs because inbreeding creates homozygotes at rare recessive deleterious mutations (Wright 1977). Inbreeding depression has been shown to have a large effect on fitness in laboratory bred species including squinting bush brown butterflies (*Bicyclus anynana) as well as Drosophila sp.* models (Ehiobu *et al.* 1989, Bijlsma *et al.* 1999; Saccheri *et al.* 1996), swine (*Sus scrofa domesticus*) and guinea pigs (*Cavia porcellus*) (Wright 1977). Furthermore, it has been suggested to be the cause of a loss of fitness in animal and plant species in the wild (Lande & Schemske 1985; Crnokrak & Roff 1999; Höglund *et al.* 2002). For example, Fredrickson *et al.* (2007) found a decrease in litter size associated with increased inbreeding in the Mexican wolf (*Canis lupus baileyi*) (Fig.1.1). The relationship shown in Figure 1 depicts a significant correlation between inbreeding coefficient, a measure of the proportion of alleles identical because of inbreeding, and reproductive fitness outside of the model species used in the lab. Therefore, inbreeding depression is accepted as a serious risk to small populations in short

timescales (O'Grady *et al.* 2006; Nieminen *et al.* 2001). However, inbreeding is not the only genetic factor contributing to extinction risk, as is summarized in table 1.1.

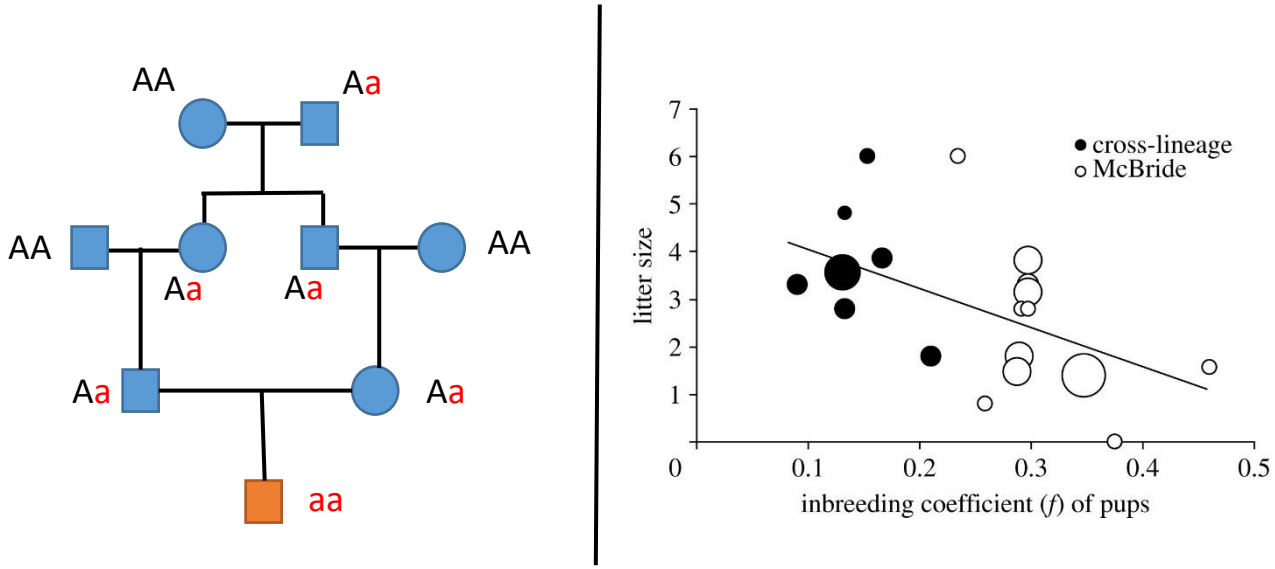


Figure 1.1: The figure on the left illustrates cause of inbreeding depression, the inheritance of a rare recessive allele which is exposed by inbreeding. The figure on the right demonstrates the importance of inbreeding depression in wild populations of conservation concern. It shows the significant ($p<0.0001$) reduction in litter size associated with inbreeding in the McBride and cross-bred lineages of Mexican wolves (right figure from Fredrickson et al. 2007 Fig. 4)

Table 1.1: The different types of genetic diversity and the way that they are affected by mechanisms that cause loss of genetic diversity.

	Genetic Bottleneck	Inbreeding	Genetic drift (in small populations)
Allelic Diversity	Large effect	Small effect	Large effect
Heterozygosity	Small effect	Large effect	Small effect

Genetic factors contribute to the extinction vortex, the self-propelling loss in population size that results in extinction (Gilpin & Soule 1986) (Fig.1.2). This occurs when population size has been reduced either by anthropocentric pressures or environmental pressures, resulting in smaller, more

fragmented and more isolated populations. Smaller populations are susceptible to inbreeding and a loss of genetic diversity associated with chance (henceforth genetic drift) which reduce reproductive fitness, adaptability and survival, further reducing population size. The extinction vortex theory is consistent with the pathway to extinction shown in wild vertebrate populations over short (12-21 year) timescales (Fagan & Holmes 2006). However, the vortex is not immutable. Genetic management and intervention to reduce pressures such as over-exploitation can arrest, and even reverse, population declines (Bijlsma *et al.* 1999; Saccheri *et al.* 1996). Nevertheless, the loss of genetic diversity associated with small populations significantly increases the risk of extinction (O'Grady *et al.* 2006).

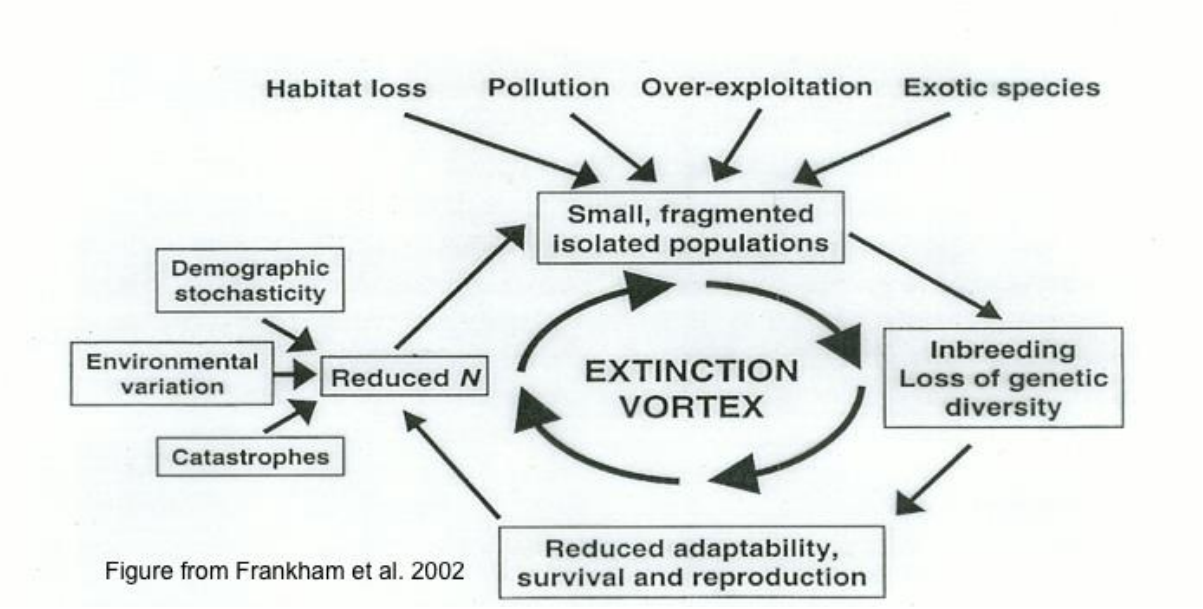


Figure 1.2 reduced remaining populations become smaller, more fragmented and more isolated. These factors drive inbreeding and increase genetic drift which reduce genetic diversity. The reduction in genetic diversity reduces adaptability and reproductive fitness. In turn these losses reduce population size in the next generation increasing the effects of the cycle. (Figure Frankham et al. 2002)

Small populations may lose genetic diversity rapidly. Sexual reproduction, which is utilised by most animal species of conservation concern, involves resampling of genetic material. This resampling causes genetic drift which, depending on the size of the population, may cause loss of alleles in a relatively small number of generations (Tremblay & Ackerman 2001). To illustrate, if a diploid

individual is heterozygous at a locus the chance that they will fail to pass an allele to the next generation is given by $(1/2)^n$ where n is the number of offspring produced. In small populations this creates a significant chance an allele will not be passed on at all, and this chance increases with rare allelels which are already at low frequency in the population. Alternatively, in very large population with large numbers of offspring the chance an allele is not passed on is negligible. The effects of these scenarios on allele frequencies, and thus genetic diversity are modelled in figure 1.3b and 1.3c. The loss of genetic diversity demonstrated in figure 1.3, and widely accepted as a consequence of small population size (Ellstrand & Elam 1993; Frankham 1996; Hauser *et al.* 1992) pushes populations further into the extinction vortex.

Small population size affects intra-population processes and the genetic integrity of the population. Selection, differential survival and reproduction depending on environmental suitability, acts to increase the frequency of beneficial alleles in the next generation and reduce the frequency of deleterious alleles (Frankham 2005; Whitlock & Agrawal 2009). However, the strength of selection is dependent on population size as described in the equation below which shows that in small populations selection may not be effective (Equation 1).

Equation 1

$$S < \frac{1}{N_e}$$

This equation states that when the coefficient of selection (S), (proportional fitness compared to the population maximum), is less that the inverse of the effective size(Ne) (the idealised size of the population) of the population then the allele is effectively selectively neutral. This can result in the situation modelled in Figure 1.3a & 1.3b where a mildly beneficial allele increases in frequency in a large population but is entirely removed in a small population (Whitlock 2000). Therefore, small populations will accumulate mildly deleterious alleles which will then reduce the overall fitness of individuals, which is called mutational meltdown or Muller's ratchet (Felsenstein 1974, Gabriel *et al.*

1993). Although there has been some dispute (see review in Felsenstein 1974, Frankham 2005), Muller's ratchet is consistent with observations in asexual populations (Gabriel *et al.* 1993) and predicted to be significant in small sexually reproducing populations (Gabriel *et al.* 1991). In order to manage a population's susceptibility to mutational meltdown population size must be considered.

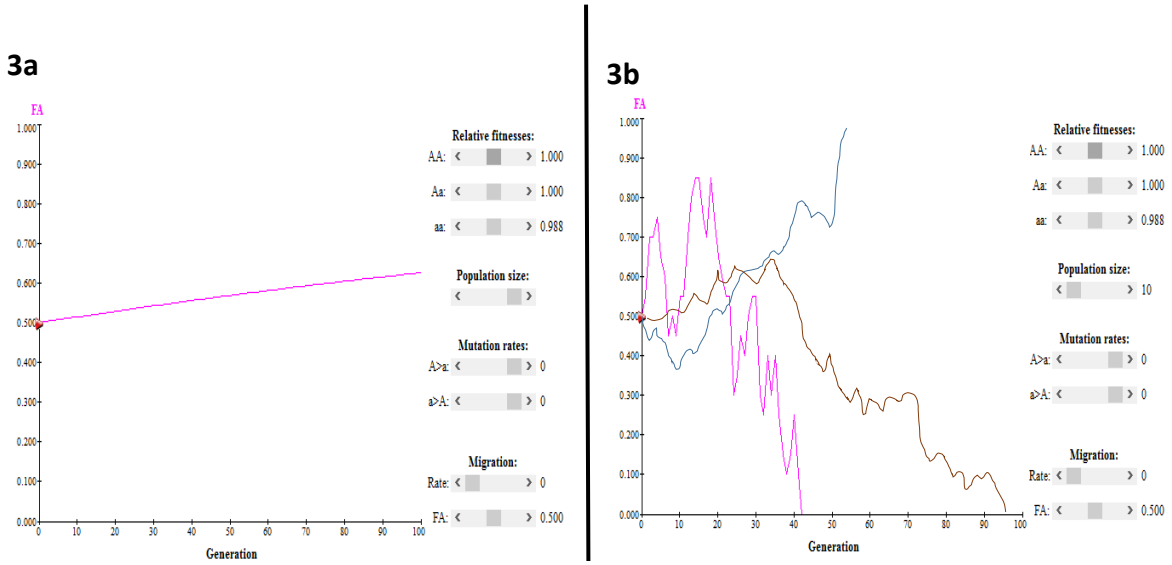


Figure 1.3 Simulations run using Selection 3.1 (Goodnight) Figure 1.3a shows the increasing frequency of a mildly beneficial allele over 100 generations in an infinite population. Figure 1.3b shows the results of simulation where the same allele becomes either extinct or fixed with the same parameters but a greatly reduced population size (population size 10).

It is the effective size (termed N_e), rather than census size of the population, which affects genetic diversity, and small effective sizes are a significant threat to species survival. The effective size of a population, as defined by Wright (1931), and further developed by Wright (1938) and Crow (1954), is the size of an idealized population that would experience the same genetic characteristicsas the population in question, though the equation used to determine effective size will depend on which genetic characterisitic is being examined (e.g. Waples 1989, Waples 2006). Nevertheless, the measure is significant as it is the effective size that interacts with mutation and selection to determine the amount of genetic diversity in a population (Lande & Barrowclough 1987). Therefore effective size predicts susceptibility to inbreeding depression (Frankham 1995) and decreased

evolutionary potential (Ellstrand & Elam 1993). For example, the decline in the population of New Zealand snapper (*Pagrus auratus*) due to over-exploitation is mirrored by a decrease in effective population size which was predicted to reduce population viability and population productivity (Hauser *et al.* 2002). Using N_e, Franklin (1980) suggested the 50/500 rule, that is, populations required an effective size of 50 to avoid inbreeding depression and an effective size of 500 to retain evolutionary potential in the longer term. These population sizes have been generally considered as a useful guide to minimum viable population size (Shaffer 1981) but have recently been revised to 100/1000 in the light of further research showing that populations with the original numbers have not retained genetic diversity in the wild (Frankham *et al.* 2014). However, research suggests that the effective sizes for vertebrate populations are, on average, 0.11 of census size (Frankham 1995). Therefore a population with census size of 250, an IUCN criteria to assign vulnerable, endangered or critically endangered status (IUCN 2001), has an average effective size of 27.5 well below the number required to avoid inbreeding depression. This has been applied to the endangered grassland daisy (*Rutidosis leptorrhynchoides*) where Young et al. (1999) suggest five populations of greater than 5000 to preserve long term population viability. These data suggest that without active management of genetic diversity many species are in danger of extinction, even if other conservation action is being taken (Holsinger & Gottlieb 1991; Dawson *et al.* 2011). Conservation genetics, as a field, enables this management with the aim of maintaining and even restoring genetic diversity and, thus, population viability.

1.2 Genetic management within conservation programs in order to reduce extinction risk

Captive breeding is a common conservation strategy that requires genetic data in order to be most effective. Captive breeding programs have been used to increase the numbers of endangered, or critically endangered species such as Mexican wolves (*Canis lupus baileyi*) and red wolves (*canis rufus*) (Hedrick & Fredrickson 2008). They are the last remaining hope of species extinct in the wild including the Socorro dove *(Zenaida graysoni*) (Carlson *et al.* 2013), and have successfully been used

to re-establish viable wild populations e.g. the Plains bison (*Bison bison*) (Freese *et al.* 2007, Seddon *et al.* 2007). These programs aim to preserve species that would not survive in the wild, with the aim of eventual reintroduction and the establishment of self-sustaining wild populations (Griffiths & Pavajeau 2008; Lerp *et al.* 2014). There have been notable captive breeding success stories including Golden Lion Tamarins (*Leontopithecus rosalia*) (Ballou & Lacy 1995) (Butler & Merton 1992). To facilitate these successes, captive breeding requires extensive genetic management. The first step in this management revolves around species identification. In small populations, suitable mates may not be available for all individuals and this may lead to hybridisation with closely related species (Avise & Nelson 1989, Garnet *et al.* 2011). This hybridisation may cause extintion (Rhymer & Simberloff 1996). Therefore, if a population containing hybrids is bred in captivity, it is important to manage the impact of the hybrid genotypes in order to preserve the original species (e.g. identification of hybrids in Przewalski's horse (Bowling *et al.* 2003). To do this hybrids must be identified, a task that is often morphologically difficult and may be impossible with subsequent generations (Smith & Fonseca 2004; Mallet 2005). Genetic markers have been successfully used to identify such hybrids in captive breeding programs e.g. (Wayne & Jenks 1991). This allows for the contribution of hybrid genes to be minimized in the captive program and species integrity to be preserved. However, most captive breeding programs do not suffer from too much genetic diversity, but rather need genetic management to prevent them suffering from too little.

Genetic management is used in captive breeding programs to avoid inbreeding and thus inbreeding depression. In early captive breeding programs deliberate inbreeding was suggested as a methodology of 'purging' deleterious alleles (Hedrick 1994,) and increasing captive fitness. Further analysis (e.g. Frankham *et al.* 2001) discovered that these apparent gains in fitness were related to improvements in animal husbandry rather than purging and that the effects of inbreeding were almost universally detrimental (Frankham 2005), though the scale of those effects varied between species. Modern captive breeding programs avoid mating closely related individuals by using studbooks systems (e.g. Cassinello 2005, also shown in the results of Chapter 2) and/or genetic

analysis of relatedness (e.g. Tzika *et al.* 2009 work on the Jamaican boa (*Epicrates subflavus*). This concern must be weighed against biasing the distribution of genetic diversity in the population, Lacy (1989) notes that preferential breeding of some individuals in a population will bias the genetic composition even several generations later. Loebel *et al.* (1992) modelled equalising founder representation using Drosophila melanogaster and found that it had a beneficial effect on inbreeding coefficient over several genetions. Therefore, population management involves compromise between goals that have competing interests at both generational and intergenerational. This is particularly important where programs aim to reintroduction populations, an aim of 52% of all amphibian captive breeding programs (Griffiths & Pavajeau 2008). However, genetic diversity must be managed, rather than simply maximized, for optimum outcomes.

Populations in captivity adapt to their altered habitats and empirically work shows these adaptations have large detrimental effects on fitness in wild conditions (Latter & Mulley 1995, Frankham 2008). Adaptation to captive conditions has been documented in a single generation (Christie *et al.* 2012) and the speed and amount of adaptation increases with the levels of genetic diversity. Therefore, careful genetic management is required to balance minimisation of inbreeding depression against minimisation of adaptation to captive conditions (Gilligan & Frankham 2003). Ultimately, captive breeding, even with informed genetic management, is a solution that can only be applied to a handful of species for reasons of cost and space. Fortunately, the understanding of the role genetic factors play in extinction also suggests another method of reducing extinction risk, one which also relies on genetic knowledge but is much less resource intensive and does not cause the problems associated with adaptation to captive conditions, genetic rescue or outcrossing.

Genetic rescue, which is migration of genetically distinct individuals into populations depauperate of genetic diversity, can relieve inbreeding depression and reduce extinction risk. The work of Sewall Wright in the 1930's established that small isolated populations were at higher risk of extinction due to the effects of inbreeding depression (Provine 1986). Widespread habitat fragmentation has

broken large populations for many vertebrates into several small isolated populations single populations into isolation of populations that were previously joined (e.g. Peres 2001).

These isolated population fragments are more likely to go extinct than a single large population (Wilcox & Murphy 1985) partially because each may be susceptible to genetic factors affecting small populations even when the species effective size is still large. Genetic diversity in each fragment is acted upon independently by drift. Furthermore, in small populations power of selection to remove deleterious alleles from the population is overcome by genetic drift (Lacy 1987, also see figure 1.3) . This means that fragmented populations lose genetic diversity and accumulate deleterious alleles faster than a single population with the same number of individuals. However, migration of genetically distinct individuals reintroduces genetic diversity and reverses inbreeding depression (Lacy 1987, Ingvarsson 2001). Of additional significance for conservation programs that aim for species conservation in the short and long term, migration of genetically distinct individuals restores evolutionary potential and thus improves species viability (Ingvarsson 2001). As figure 1.4, shows this is true even when both fragments suffer from the effects of inbreeding. In theory a single effective migrant per generation is enough to minimise the loss of both heterozygosity and allelic diversity (Spieth 1974, Frankham *et al.* 2002). However, Mills and Allendorf (1996) conclude that in non-ideal populations the optimal number varies between 1-10 migrants per generation. Still these are very low numbers of individuals, and translocation is very resource efficient compared to other conservation initiatives such as captive breeding programs (Rahbek1993, Snyder *et al.* 1996). This has been demonstrated to reverse the decline of severely inbred populations of Swedish Adders (*Vipera berus*) (Madsen et al 1999), and Greater Prairie Chickens (*Tympanuchus cupido pinnatus*) (Westemeier *et al.* 1998).Therefore, it is not surprising that translocation is often recommended as part of conservation programs (e.g. Lacy 1987, Burkey 1989, Mills & Allendorf 1996).

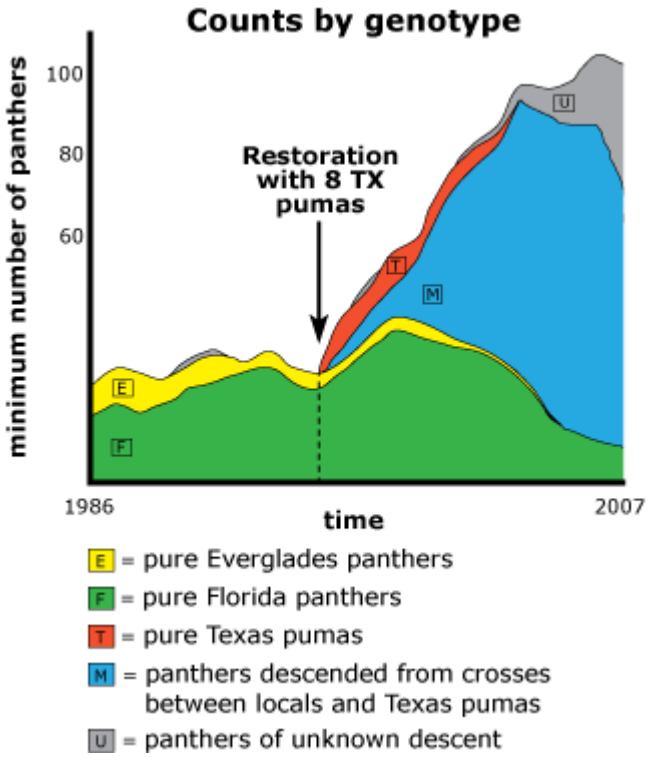


Figure 1.4: The effects of genetic rescue are shown in increase in Florida panther population size since an outcrossing event in 1995 (Figure adapted from Johnson *et al.* 2010 by Gishlick et al)

Outbreeding depression has been a significant impediment to the adoption of outcrossing practices. Outbreeding depression occurs where outcrossing between populations destroys beneficial local adaptations and reduces fitness (Templeton *et al.* 1986, Frankham *et al.* 2002). Outbreeding depression has been demonstrated with model species including the round worm *Caenorhabditis elegans* (Dolgin *et al.* 2007). It has also been found in greenhouse experiments where plants from different localities were cross bred (Montalvo & Ellstrand 2001). Additionally, it is held responsible for the total loss of the re-established Tatra Mountain ibex (*Cupru ibex ibex*) population where outcrossing occurred as part of a conservation program (Grieg 1979). Repeated analyses, including those by Frankham *et al.* (2011) and Tallmon *et al.* (2004), suggest, however, that the risk posed by outcrossing is manageable by comparison with the known effects of inbreeding depression and loss of genetic diversity. In spite of this, when making conservation decisions where the consequences of potential failure (such as with the Ibex) are large and very public, while success (long term species survival) is likely to be cryptic, wariness on the part of conservation managers is expected. To

address this problem, it is import to codify decision making practices (Frankham *et al.* 2011) and also to ensure conservation managers have the best information possible available to them.

The effectiveness of conservation genetics and a portion of the effectiveness of conservation rests on our ability to categorise the amount and distribution of genetic diversity. The level of genetic diversity within a species is often cryptic (Gibson & Dworkin 2004). Genetic diversity can be heavily influenced by ancient population bottlenecks such is in the Serengeti Cheetah (*Acinonyx jubatus*) where genetic diversity levels are extraordinarily low (He = possibly as low as 15% of ~43 individual adults in 1994 according to Kelly 2001), presumably as a result of a bottleneck ~10000 years ago (Menotti-Raymond & O'Brien 1993). Additionally, the effect of human civilisation, including those of ancient hunter-gatherer societies, has shaped the amount and distribution of genetic diversity present in species today (Vitousek *et al.* 1997). The distribution of genetic diversity within a species may also be cryptic, with source populations containing large proportions of the overall genetic diversity and sink populations (with immigration but no emigration), which are often less genetically valuable (Dias 1996). Landscape factors, including those reaching back to the last glacial maximum, have also been shown to have large effects on distribution of extant genetic diversity within species (Storfer *et al.* 2007). It is impossible to predict which of these factors will affect a species of conservation concern, and what level of influence each exerts. Therefore, in order to manage genetic diversity, and reduce extinction risk, it is first necessary to have tools capable of examining genetic diversity in individuals, populations and species.

1.3 Molecular tools used to understand genetic diversity

While the theoretical underpinnings of population and conservation genetics are relatively well established, the molecular methods used to inform conservation managers undergo near constant change (Schlötterer 2004). The crux of the problem, when trying to apply conservation genetics theory to populations, is deciding which genetic variation to examine, and how to do so. Whole

genome sequencing is a relatively recent innovation (reviewed in Shedure & Hanlee 2008) and although the field of population genomics is burgeoning (see Ellegren 2014), it is not yet widespread. Certainly in many species of conservation concern relatively little is known about the genome structure and therefore genetic markers are still used as a surrogate for whole genome diversity.

Using genetic markers to represent whole genome genetic diversity may result in inaccurate or misleading representations, depending on the limitations of the marker. Any analysis, except whole genome studies, will fail to recognise some types of genetic diversity. How important this loss of information is will depend on the biological questions being asked. Depending on the question of interest a relatively small number of genetic markers will generate highly accurate data, e.g. police forensics use 11 microsatellite markers and the chance of a mismatch is 1 in a trillion (Human Genetics Commission 2009). Furthermore, using two different genetic markers may result in different understandings of genetic diversity such as in the San Nicolas Island fox (*Urocyon littoralis dickeyi*) which has very little genetic diversity at microsatellite markers (He = 0) whereas studies using MHC markers reveal substantially more genetic diversity (He = 0.62) (Aguilar *et al.* 2004). This has also been shown in Atlantic herring (*Clupea harengus*) where microsatellite markers were used to discern population structure which mtDNA and allozyme markers did not show (Shaw *et al.* 1999). It is for this reason that choosing the correct genetic marker to answer conservation questions is important (Wan *et al.* 2004).

The choice of genetic marker is determined by practical reasons (such as invasiveness of the sampling required, financial or time constraints, and the technology available) as well as by the specific question being asked. In conservation genetics the welfare of organisms is a primary concern. For this reason, invasive genetic analyses, such as allozyme electrophoresis requiring large amounts of muscle or liver tissue (e.g. Richardson *et al.* 1986), are typically avoided. These invasive first-generation genetic marker tests were typically phased out (described in figure 1.5) in animal species, although not entirely (e.g. Habel *et al.* 2011, Alam *et al.* 2014), following the widespread use

of genetic markers based on the polymerase chain reaction (PCR). Using PCR researchers were able to amplify large amounts of DNA from small samples. This facilitated the creation of minimally invasive (e.g. Wasko *et al.* 2003, Caudron *et al.* 2007) and non-invasive (e.g. Valière *et al.* 2003) genetic sampling methods. Both microsatellite and MHC techniques (discussed in detail in this thesis) belong to this second generation of genetic markers. Subsequently, next generation sequencing has facilitated the creation of additional genetic markers by reducing the time and cost involved in sequencing large quantities of DNA (Hayden 2014). The progression in use of genetic markers is described by Figure 1.5. Wan *et al.* (2004) argue that inappropriate marker choice can lead to incorrect outcomes and that one the most popular genetic markers, microsatellites, are ineffective for conservation genetics. Thus, the choice of genetic marker and what that marker is measuring is crucially important to interpreting results generated by that marker. In this thesis, I engage with this issue by first briefly discussing some of the most commonly used genetic markers in conservation genetics, in chronological order, and then by applying various markers to conservation problems.

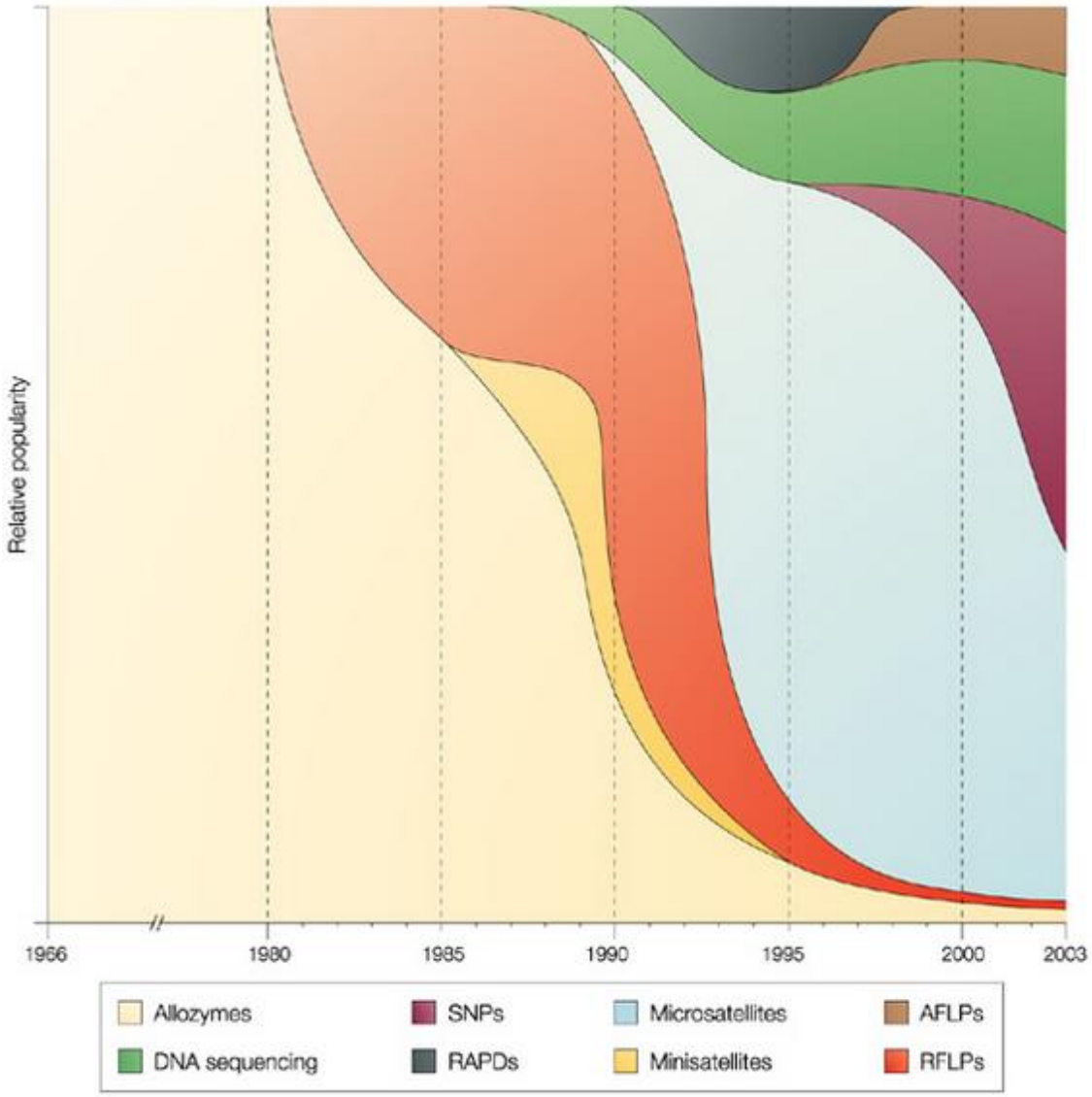


Figure 1.5: The use of different types of genetic markers in published research from 1966 to 2003 (figure from Schlötterer 2004)

1

2 **First generation genetic markers: allozymes**

3 The first generation of genetic markers revolutionised the level of information available about the
4 genetic composition of a species and changed our understanding of within-species processes. This
5 revolution was achieved by using biochemistry to visualise genetic variation as genetic markers that
6 could be used to test genetic theory. The field of population genetics precedes genetic markers,

being founded on the work of Wright, Fisher, Haldane and Malécot (Epperson 1999) and Dobzhansky (O'Brien pers comm). However, widespread use of genetic markers, as opposed to more obvious phenotypic traits, did not occur until the development of electrophoresis (Smithies 1955) which enabled the development of the first widespread population genetic marker system, allozyme electrophoresis.

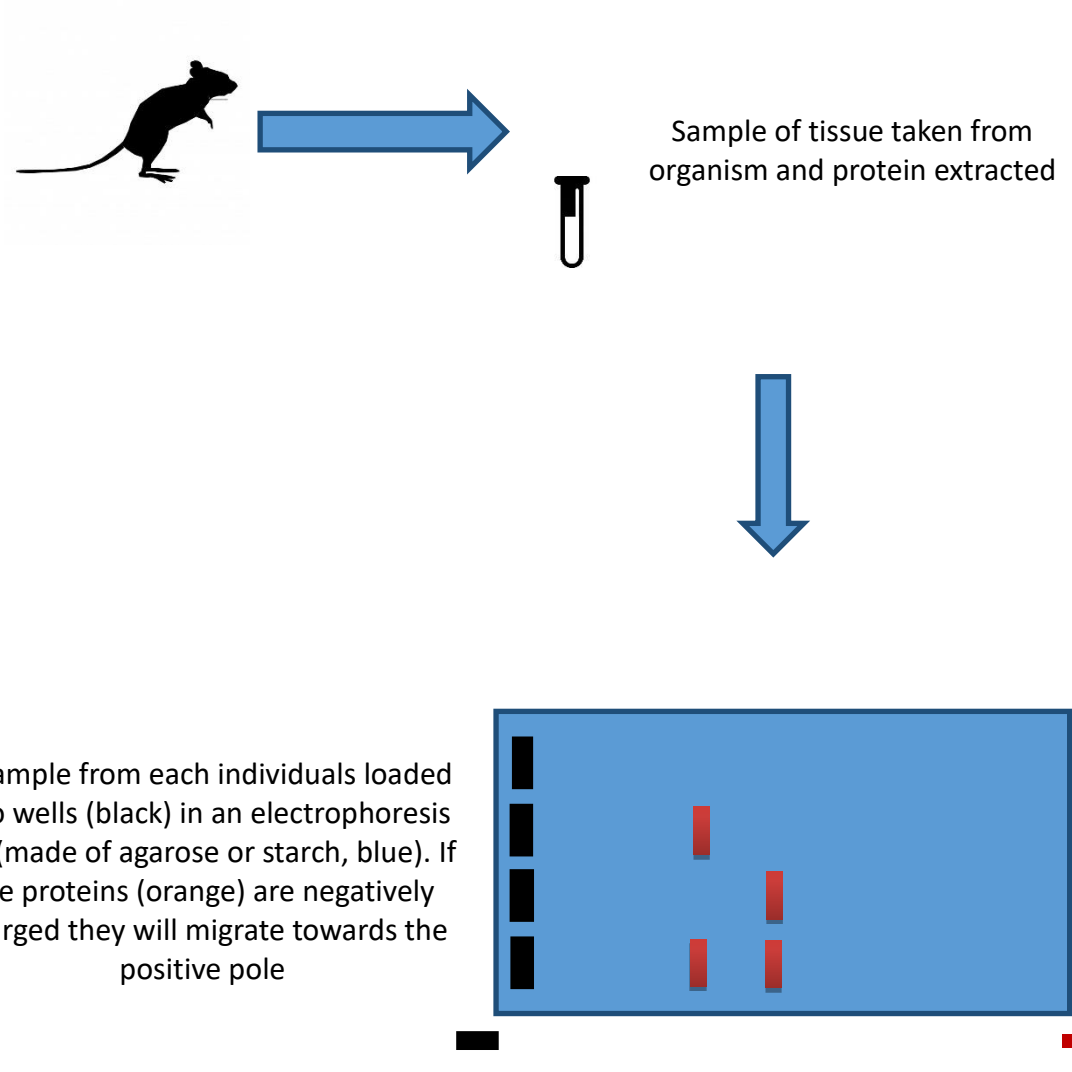


Figure 1.6: A brief explanation of the process of allozyme electrophoresis.

Allozyme electrophoresis allowed for differentiation of different versions of an enzyme and had large benefits for population and, thus, conservation genetics. As is described in Figure 1.6, gel electrophoresis is based on the side chains of amino acids changing either the shape or charge of the overall polypeptide. In order to exploit these differences, polypeptides are placed into a gel matrix and the pH is adjusted so that the COO- groups on the amino acids are neutralized and the polypeptide has a net positive charge, stemming from NH4+ (Richardson & Baverstock 1986). This net positive charge means that polypeptides migrate towards the negative pole. The rate at which a particular polypeptide will move through the gel matrix was described by Richardson & Baverstock (1986)

$$\mu = \frac{Qd}{4\pi r^2 \eta}$$

Equation 2

Where: μ represents mobility; Q represents charge on the protein molecule; d the distance of the electrical double layer from the molecule; r the radius of the molecule and n viscosity of solution. Given that gel electrophoresis uses a solution with equal viscosity and that gel rig apparatus are designed so that distance to the electrical double layer is constant across all lanes of the gel, the rate of protein movement will be determined by charge and radius of the polypeptide. The charge of the polypeptide is determined by the amino acid side chains and four of the amino acids; Arginine; Lysine; Aspartic acid and Glutamic acid are charged. Further, the shape of the polypeptide can be affected by the interactions of hydrophilic and hydrophobic amino acids. Therefore, changes in allele sequence could result in detectable changes in polypeptide structure. Visualisation techniques based on enzyme action meant that allozyme electrophoresis created co-dominant genetic markers (Richardson & Baverstock 1986).

The use of allozyme markers presented distinct advantages and disadvantages. As the first widely used co-dominant genetic marker, allozymes were able to be used to exclude paternity to a high

degree of certainty and facilitated a revolution in our understanding of bird mating systems (e.g. Gowaty &Karlin 1984). They were also instrumental in testing population genetics theory in model species such as *Drosophila pseudoobscura* (Lewontin & Hubby 1966) and were applied extensively in human research (Harris & Hopkinson 1976). Furthermore, as functional markers, they are able to show evidence for local adaptation by identifying differences in enzymes associated with changed environmental conditions (McKay *et al.* 2001, Dhuyvetter *et al.* 2004). However, as nearly neutral genetic markers, their ability to discover local adaption is rare (Reed & Frankham 2001). Furthermore, allozyme electrophoresis is limited in resolution as it cannot detect any genetic changes that do not affect mobility of an enzyme. For this reason allozymes are currently widely considered to lack the resolution (number of scoreable alleles) necessary for examining genetic diversity in populations (e.g. Tessier *et al.* 1995, Estoup *et al.* 1998). Additionally, allozymes, as protein markers, require large amounts of tissue be collected, often from vital organs such as the liver (Richardson *et al.* 1986). For this reason allozymes are unsuited to conservation programs where animals cannot be sacrificed though they are still used in plant studies (e.g. Mckay *et al.* 2001) where removing large amounts of tissue for study does not create the same risks. As figure 1.5 shows, the use of allozyme markers has decreased dramatically since the creation of microsatellite markers. Therefore, allozymes are particularily interesting as they mark the end of the widerspread use of functional genetic markers in population studies. The neutral markers that replaced allozymes (detailed below) all require the assumption that neutral genetic diversity mirrors functional genetic diversity to be useful within conservation genetics.

The molecular revolution and its effect on genetic markers

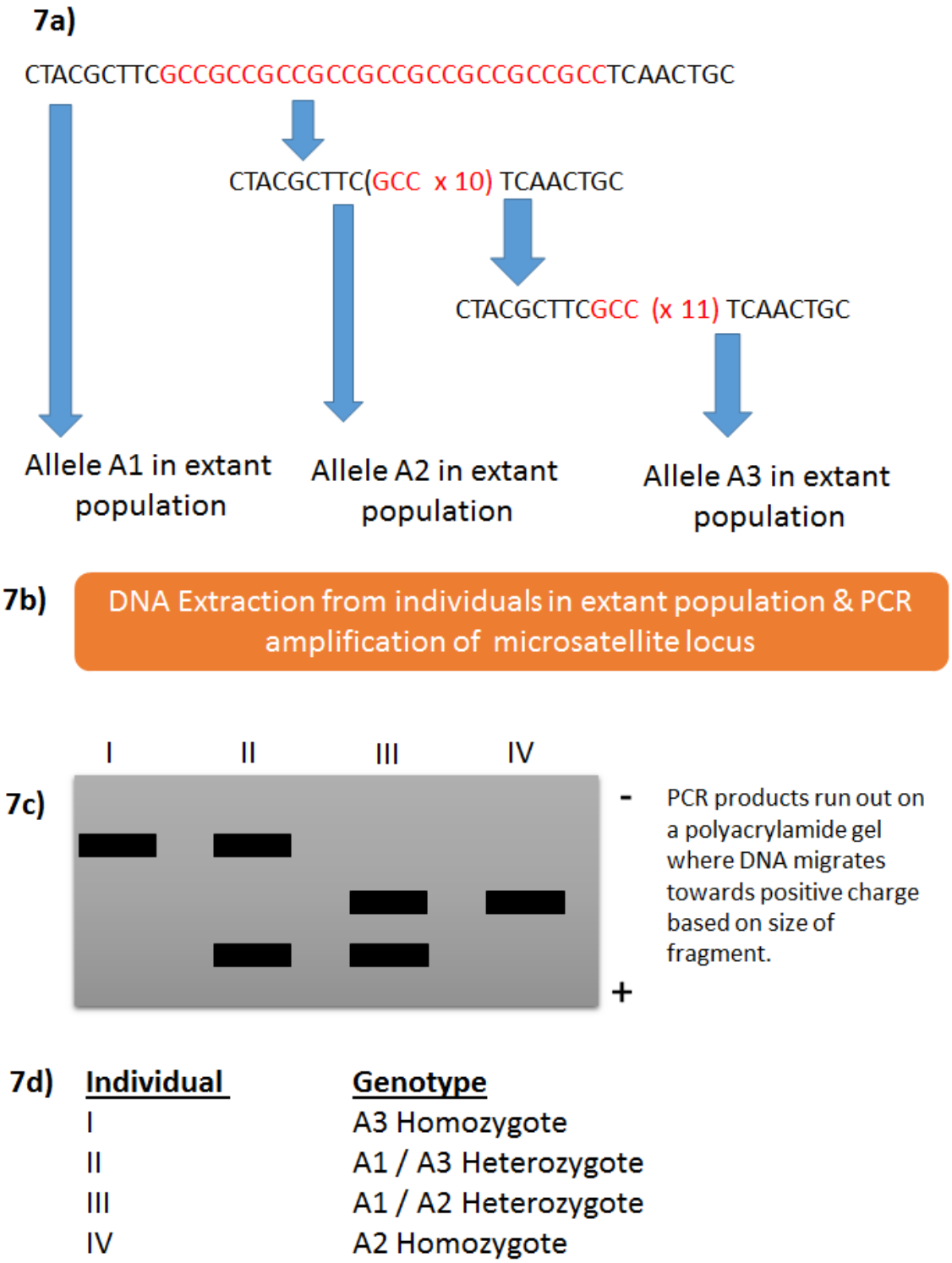
The molecular revolution enabled the creation of genetic markers based on DNA and their widespread adoption in conservation studies. Restriction Fragment Length Polymorphism (RFLP) is a type of genetic marker based on variation in the cutting sites associated with restriction enzymes and allowed for (relative to other techniques available at the time) low cost, fast and informative

analysis of large populations (Ragot & Hoisington 1993). RFLPs have been widely used to study genetic diversity (eg. Yuhki & O'Brien 1990) and parentage (Smouse & Chakraborty 1986). However, RFLPs can only detect variation which changes the ability of a restriction enzyme to cut DNA and therefore have limited resolution. Additionally early RFLP techniques also required large amounts of DNA (Jiang 2013) which can be difficult to obtain without harm to animals. Therefore, genetic markers based on the polymerase chain reaction (PCR) became widely used in conservation genetics.

The widespread use of PCR enabled the creation of a suite of genetic markers that were relatively inexpensive and minimally invasive;. PCR enabled the exponential amplification of genetic markers quickly and relatively cheaply (Morgan *et al.* 1998). This meant that genetic markers could be typed from small samples of tissue (e.g. Aljanabi & Martinez 1997), blood (e.g. McCabe 1991), saliva (e.g. Ng *et al.* 2004) or even faecal or hair samples (e.g. Taberlet *et al.* 1996). This has several benefits for conservation studies, such as enabling detailed genetic typing without causing significant harm to animals or when animals are difficult to find or capture (Sloane *et al.* 2000). This techonology also enabled the use of genetic sequence as a marker.

Genetic sequence of mitochondrial (mtDNA) as a type of genetic marker was enabled by PCR and remains a key component of population genetics. Mitochondrial DNA has several advantages as a marker including being inherited materally (Giles *et al.* 1980) and having a slow mutation rate (Palmer and Herbon 1988). For these reasons mtDNA is used to power molecular clocks that time the divergence between species (e.g. Hasegawa *et al.* 1985). For example Weir and Schulter (2008) calibrated an mtDNA molecular clock against fossil and biogeographic data to produce a 12 million year history of avian evolution across 12 taxonomic orders . However, mtDNA only shows the maternal genetic lineage and does not show fine scale genetic portioning, as it evolves slowly, therefore non-sequence genetic markers were commonly employed or employed in conjunction with mtDNA in population studies (e.g. Lorenzini *et al.* 2003, Feulner *et al.* 2004).

The widespread use of PCR created several new types of genetic markers, including minisatellites, Single stranded Conformational Polymorphisms (SSCP), terminal Restriction Fragment Length Polymorphism (t-RFLP) and Amplified Fragment Length Polymorphism (AFLP). However, the most commonly used type of second generation molecular maker between 1993 and 2003 was microsatellites (Schlötterer 2004).



7d)

Individual	Genotype
I	A3 Homozygote
II	A1 / A3 Heterozygote
III	A1 / A2 Heterozygote
IV	A2 Homozygote

Figure 1.7: A description of the process of typing microsatellite markers. Variation in microsatellite length is created as DNA polymerase is more likely to make errors when reading repeating sequence (Jarne & Lagoda 1996). These distinct alleles are amplified by PCR and like allozyme electrophoresis are visualised by running out on a gel where smaller fragments move faster towards a positive charge than larger fragments. The result is a set of co-dominant markers that are simple to genotype and have high levels of variability.

1

Microsatellites were the most commonly used genetic marker between 1993 -2003 (Figure 1.5) because they provided a relatively low cost, high-resolution method for examining genetic diversity. As shown in Figure 1.7, microsatellites are short repeating sections of DNA (e.g. CACACACACA) with mutation rates that are higher than surrounding sequence and related to length of the repeating block (Schug *et al.* 1998). This high level of genetic diversity gives the markers high resolution (Jarne & Lagoda 1996). Like allozymes, microsatellites are simple to visualize (Frankham *et al.* 2010), keeping costs relatively low. Additionally, due to sequence similarities, microsatellite markers can be used in closely related species (e.g. Moore *et al.* 1991; Primmer *et al.* 1996; Cordeiro *et al.* 2001). Thus, the advantages of microsatellite markers have led to their extensive use in species of conservation concern for examining the amount of genetic diversity (e.g. Hokanson *et al.* 1998; Fabuel *et al.* 2004), examining intra-population processes such as mating systems (e.g. Awadalla & Ritland 1997; Jones & Avise 2001), understanding dispersal and population structure (Balloux & Lugon-Moulin 2002; Margaritopoulos *et al.* 2009), and proposing management plans (e.g. Maudet *et al.* 2002) as well in stopping the trade in endangered species (e.g. Wasser *et al.* 2008). However, despite their many advantages, microsatellites do have some important limitations that mean they may not always represent the best choice of genetic markers.

The limitations of microsatellites are discussed in more detail in Chapter 3 and by other authors (Wan *et al.* 2004). Briefly, microsatellites, which are scored based on size, miss variation where mutation has produced alleles of the same size in different ways (Wan *et al.* 2004). Although this is of negligible importance in studies where microsatellites are used to identify individuals (e.g. Taberlet & Luikart 1999) it can be important when attempting to understand the relationship between traits and alleles. Furthermore, as a neutral marker, microsatellites may not reflect the effects of selection on functional areas of the genome (e.g. Cousyn *et al.* 2001). Alternatiely, when microsatellite loci are involved in gene regulation and therefore underselection or linked to loci under selection (Kashi & King 2006) the distribution of genetic diversity at microsatellite markers may not match the distribution of genetic diversity at other functional (e.g. Beacham *et al.* 2005) or

even neutral (e.g. Vàli *et al.* 2008) loci. These uncertainties complicate management decisions. Using microsatellite markers developed for another species may result in lower levels of genetic variation being detected (Primmer *et al.* 1996). These limitations may help to explain the trends visualised in Figure 1.5. Unlike the first generation of genetic markers, which was dominated almost completely by allozymes, the second generation included a number of alternative markers along with microsatellites.

1.4 MHC, the major histocompatibility complex

The major histocompatibility complex (henceforth MHC) is an area of the genome that has long been studied and has recently been used as genetic marker in conservation. MHC was first described by Gorer (1936) and linked to tissue rejection in mice by Little in the late 1940's (Auchincloss & Winn 2003). The first major breakthrough, for which Snell *et al.* were awarded the 1980 Nobel Prize, was identifying that tissue rejection in mice was based on MHC incompatibility (Snell 1951). This led to an understanding of how MHC, termed HLA (human leukocyte antigen) in humans, was similarly correlated with human survival after organ transplants (e.g. Terasaki *et al.* 1965). This discovery led to extensive research (e.g. Bernatchex & Landry 2003, Forsberg *et al.* 2007, Agbali *et al.* 2010, Babik 2010 among many others) aimed at understanding the structure and function of MHC in humans and mice (Fig. 1.8).

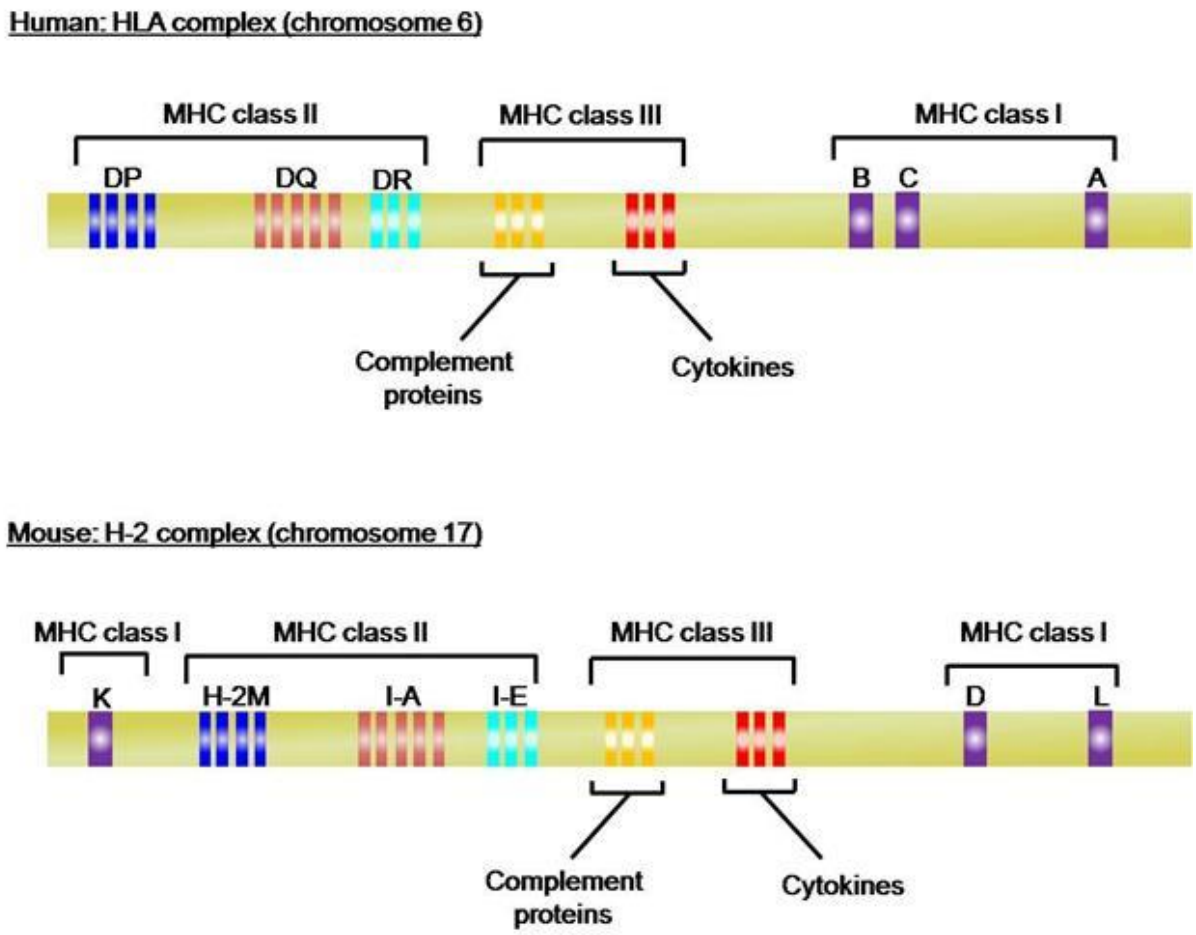


Figure 1.8: The structure of MHC on human chromosome 6 and mouse chromosome 17. Figure from National Programme on Technology Enhanced Learning, India.

The genotyping of MHC has evolved with molecular technology and has allowed for greater insight into the gene complex. Early work genotyping MHC used the mixed leukocyte reaction whereby leukocytes from two donors would react if they contained different MHC derived antigens. This technique was used on mice, rats, chicken and humans (Amos & Batch 1986) and was able to show that as related individuals had more similar MHC a less powerful reaction would occur, which could be measured by measuring cellular uptake of tritiated thymidine (Bain & Lowenstein 1964). The importance of MHC to organ transplantation led to the micro-cytotoxicity assay whereby lymphocytes are tested against serum with known anti-HLA antibodies in order to match MHC types for better survival rates (Dyer & Middleton 1993). Whilst this method is still used it is no longer

considered the 'gold standard' as molecular methods are more widely practised (Erlich *et al.* 2011). There are now a number of molecular methods for the typing of MHC in humans including SSCP (single strand conformational polymorphism), PCR-RFLP (restriction fragment length polymorphism) and sequencing based approaches (Erlich *et al.* 2011). However, these approaches, though successful, are only widely used in humans and model-vertebrates (Babik 2010). Nevertheless, the typing and genotyping of MHC has created a better understanding of its role in the immune system.

MHC is now understood to play a key role in the adaptive immune system. MHC allows vertebrates to identifying cells bearing pathogens and mark them for destruction (Janeway *et al.* 2001). Zinkernagel & Doherty (1974) demonstrated that this ability was based on genetic variation in MHC and were awarded the Nobel Prize in Medicine (1996) for their discovery. Specifically, different genetic variants at MHC are able to chaperone different intracellular proteins to the surface of cells where they are displayed and then can be recognised as antigens by the immune system (Janeway *et al.* 2001). MHC is divided into three major classes with distinct functions (Fig. 8). Class I MHC genes mediate cellular immunity (Kobayashi & van den Elsen 2012). These genes are expressed in all nucleated cells and mediate the expression of epitopes (Hewitt 2003) These epitopes are recognized by Killer T-cells (Townsend *et al.* 1986) which can lead to apoptosis, the activation of macrophages and the production of cytokines (Janeway *et al.* 2001). In contrast, MHC Class II is involved in the activation of T-cells (Germain 1994). MHC class II genes encode proteins that mediate the presentation of epitopes on antigen presenting cells (APC's). These epitopes are recognised by, and activate, T-cells (Janeway *et al.* 2001). The activated T-cells are then able to stimulate the activation of B & T cell responses to the antigen threat. Within the Class II genes most research in vertebrates has focused on the DRB genes that code for antigen binding sites (Sommer 2005). The role of MHC class III is distinct from the other types of MHC as it encodes proteins that are secreted and have immune functions including controlling the inflammation response as well as proteins with other unknown functions including mRNA processing and genes associated with a variety of diseases (Lehner *et al.* 2004). However, not all MHC class III genes have immune functions (Lehner *et al.*

2004). Whilst the initial understanding of the structure and function of MHC classes was based on research in mice and humans (and this research continues), additional studies outside of these model vertebrates suggest a more complicated and more variable structure of MHC in other vertebrates (compared in Figure 1.9).

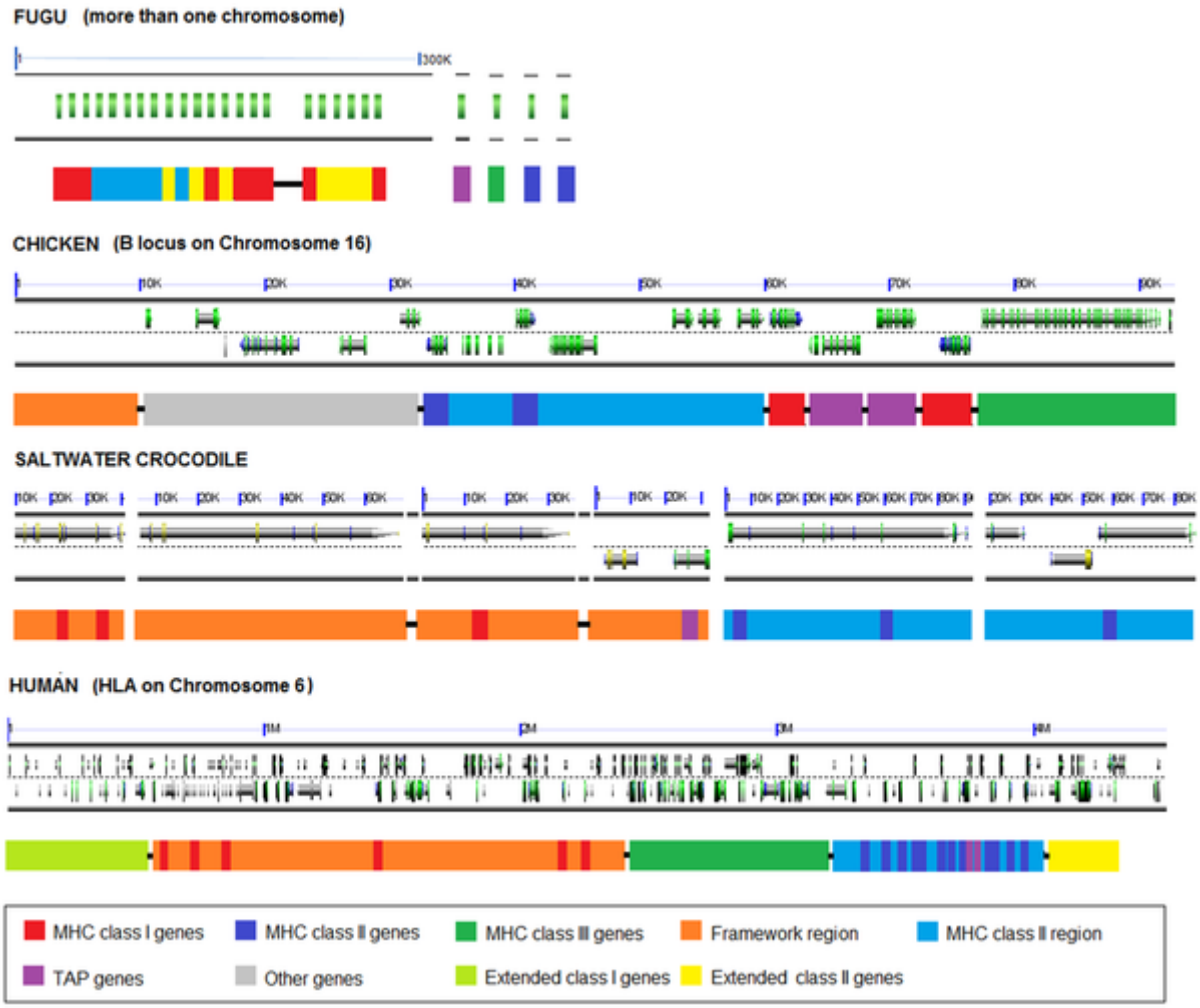


Figure 1.9: The structure of MHC compared between FUGU, chicken, saltwater crocodile and humans (figure from Jaratlerdsiri et al. 2014)

In order to extend our knowledge of MHC beyond model vertebrates and use it in conservation, the scale of differences in MHC between vertebrates must be understood. However, the overall structure of MHC is not highly conserved across all types of vertebrates (Fig. 1.9). While chicken (*Gallus gallus domesticus*) MHC is comprised of two sub-regions, physically located next to each on microchromosome 16 (Delany *et al.* 2009) in a similar way to humans, these two regions are inherited independently (Briles *et al.* 1993). Conversely, in bovine MHC a large part of MHC class II

has moved in the genome as a result of an inversion event, yet retains the same functional gene content (Childers *et al.* 2006). The *Xenopus* genus of frogs exhibit non-classical MHC class I genes on a different chromosome than other MHC genes (Flajnik *et al.* 1993) and in bony fishes MHC class I and II genes are not located on the same chromosomes (Sato *et al.* 2000). Kulski *et al.* (2002) identifies major structural differences between MHC in mammalian and non-mammalian vertebrates, a finding that is significant considering the proportion of MHC knowledge that revolves around studies of humans and mice. Of concern to conservation geneticists, the structure of MHC in most vertebrates remains unknown (Kulski *et al.* 2002). However, Grossberger & Parham (1992) identify highly conserved structures within MHC across species and these structures have, in part, enabled researchers to understand the unique position of MHC in the vertebrate genome.

As a functional genetic marker MHC possesses unique advantages relating to both the amount of, and role of, its genetic diversity. This is discussed in detail in later chapters. Briefly, MHC is the most variable of all known functional genes in vertebrates (Hedrick 1994, Sommer 2005) with >7,500 common alleles found in humans (de Bakker *et al.* 2006). As a gene complex this variability extends to the number of genes in the complex (Hosomich *et al.* 2006) and the transcription intensity of MHC genes (Hosomich *et al.* 2006) within and between species. This vast genetic variability is of interest to conservation geneticists in part because specific MHC variants have been correlated with susceptibility to specific diseases (Westerdahl *et al.* 2005, Pitcher & Neff 2006), important in some conservation programs. MHC variability is also correlated with fitness (e.g. Radwan *et al.* 2012) (discussed in more detail in chapter 3), important in all conservation programs. Furthermore, genetic diversity at MHC, including MHC dissimilarity (e.g. Forsberg *et al.* 2007, Agbali *et al.* 2010) and allele counting (e.g. Reusch *et al.* 2001) has been shown to affect mate choice in a number of vertebrates (discussed in more detail in chapter 3) which is important to some captive breeding programs. Finally, MHC has been suggested as a good genetic marker for examining local adaptation (Eizaguirre & Lenz 2010) which is important when outcrossing between population fragments (Frankham 1995). However, MHC is not a widely used genetic marker, particularly when contrasted with

microsatellites. The comparison of MHC and microsatellites forms the basis for Chapter 3.1. One of the reasons that MHC is not used widely, as discussed in chapter 4, is that it is difficult to genotype. These difficulties however may be lessened by the widespread adoption of next generation sequencing technologies

1.5 Genetic Markers based on Next Generation Sequencing

The human genome project was a collaboration between governments of the UK, USA, Japan, France, Germany and China and industry that in 2001 published a 90% complete sequence of the human genome. In this process, new sequencing tools were developed that radically changed the speed at which genetic sequences could be generated. Prior to the genome project Sanger *et al.* (1977) had developed the most widely used DNA sequencing method, which is described in Figure 1.10a and b. This method used the polymerase chain reaction to generate numerous sequences which terminated at random positions depending on the incorporation of fluorescent nucleotide bases (Smith *et al.* 1986). These PCR products were separated by size and the specific base at a position detected based on fluorescence. Although Sanger sequencing produces highly accurate sequences (Sanger *et al.* 1977) and is still widely used in sequencing applications (e.g. Tedersoo *et al.* 2010), it produces too few sequences, at too slow a speed, to be a cost effective solution for whole genome sequencing in conservation projects. Pyrosequencing, developed in the 1990's (Ronaghi *et al.* 1996), commercialized in the 2000's (Brenner *et al.* 2000) and described in Figure 1.10 address these short comings. These methods use bead based PCR to amplify many DNA sequences in parallel and a series of washes to detect bases and then shorten the PCR product by one base in order to enable the next base to be detected. Though more expensive for a single run than Sanger sequencing pyrosequencing approaches produce exponentially more sequence data (Marqulies *et al.* 2005). These advances have made generating sequence data vastly more practical. By comparison the human genome project cost $2.7 billion (NHGRI 2010) whereas the cost of sequencing a genome in 2001 was approximately $100 million (Hayden 2014). Figure 1.11 shows the impact of next

generation technology on reducing the cost of sequencing; the cost of one megabase of sequence data was reduced from near $10million in 2000 to approximately $5,000 in 2013.

These technological advances have allowed for practical and cost effective genotyping techniques which exploit the resolution of sequence data and speed of next generation platforms.

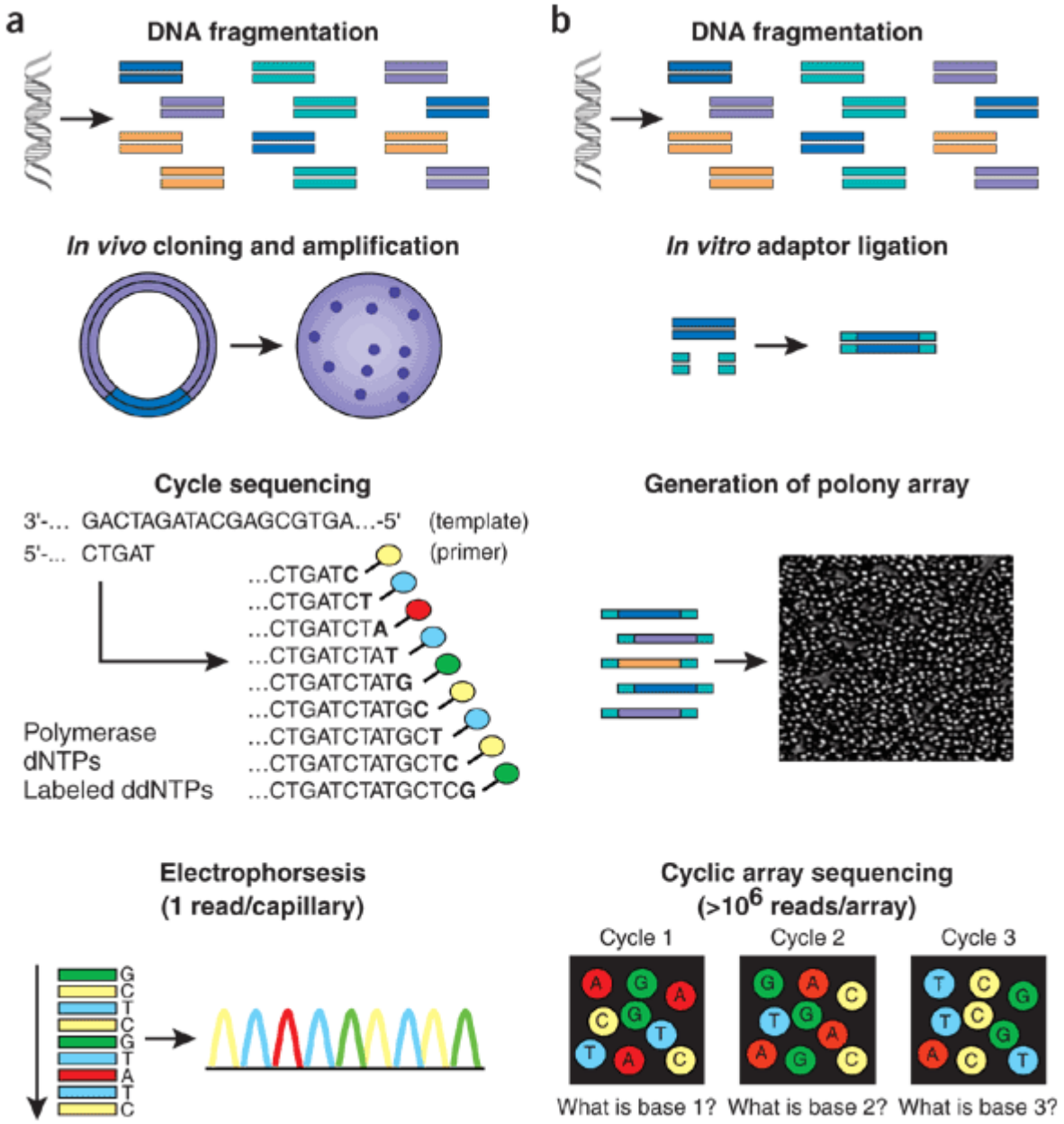


Figure 1.10a & b: comparison of Sanger sequencing which is figure 1.10a and next generation sequencing which is figure 1.10b (Shendure & Hanlee 2008)

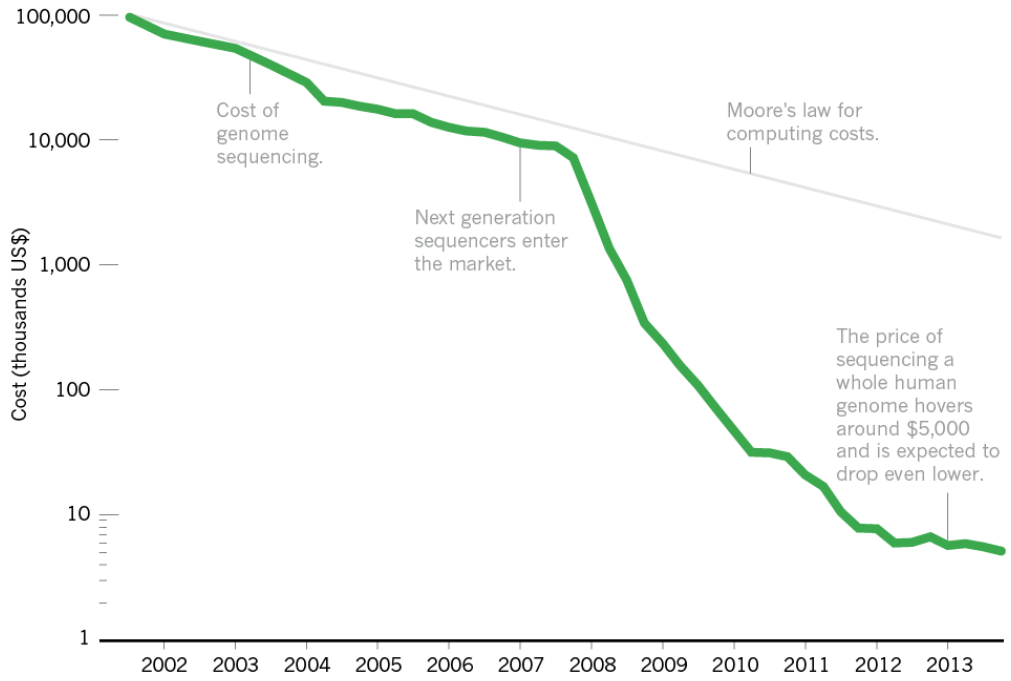


Figure 1.11: The cost of sequencing compared to Moore's law (predicting a halving cost every year). Figure from Hayden 2014

The amount of data generated by next generation sequencing has enabled the use of sequence variation as a genetic marker. Several specific genotyping by sequencing methodologies will be discussed in later chapters (Chapter 4 and Chapter 5) though numerous other methodologies exist (reviews include Davey *et al.* 2011, Nielsen *et al.* 2011). Briefly, these technologies sequence large amounts of either genomic DNA or PCR product and detect single nucleotide polymorphisms (henceforth SNPs) (Davey *et al.* 2011). These SNPs are the alleles genotyped, thus the creation of, and genotyping of, genetic markers for a species can now occur in a single step (De Pristo *et al.* 2011). Further, genotyping by sequencing approaches are the only method to accurately determine all of the genetic variation at a locus. These advantages make genotyping by sequencing enticing for conservation genetics. However, in addition to opportunities, genotyping by next generation sequencing creates unique challenges.

Challenges specific to each technique used are discussed in the appropriate chapters (Chapter 4 & Chapter 5) in detail. Briefly, all massive parallel sequencing technologies upon which genotyping by

sequencing is based are single step procedures (Davey *et al.* 2011). As well as exponentially increasing the amount of data produced this decreases the ability of a researcher to optimise the molecular approaches. This is important as optimising molecular procedures for the type of sample (e.g. Austin & Dillon 1997) and the type of sampling used (e.g. Sarks & Peters 2002) is commonly required to generate good data in conservation genetics. To mitigate this risk next generation sequencing methodologies have quality control steps before sequencing is carried out (Marqulies *et al.* 2005). However, the effectiveness of next generation methods for genotyping in species of conservation concern, where little genetic information is typically available, is difficult to assess as the publication bias in science (Young *et al.* 2008) means that failures may be cryptic. Further difficulties are presented by the amount of data next generation approaches generate.

The vast riches of data generated by next generation methodologies present bioinformatics challenges. Genetic markers used in conservation genetics have typically been scored by eye (e.g. Bentzen *et al.* 1996, Stow *et al.* 2001) where the researcher uses their experience to differentiate signal from noise. The amount of data next generation sequencing platforms produce makes this difficult. Genotyping is now typically an automated process carried out by computer programse.g. McKenna *et al.* 2010, Catchen et al 2001,, a practice which was less prevalent when work for this thesis began. Applying a researcher's insight requires translating ideas into rules applied by these programs. This is the basis of the data refining steps after sequences are generated in genotyping by next generation sequencing approaches (e.g. Babik *et al.* 2009). These technical steps may create challenges for researchers more used to working with visual data as the steps in generating and scoring alleles are now cryptic. Further, the sheer amount of data created by next generation approaches poses a processing problem and data sets, though highly informative, are long and arduous to analyse.

Given the plethora of available methodologies, technologies and techniques it is important to understand the uses and limitations of different genetic markers and the role of genetic diversity in

order to enable the best conservation outcomes. This thesis will explore these themes in six chapters.

Chapter 2: Using genetics to inform the management of a threatened species, the mala (Lagorchestes hirsutus)

In this chapter we use second generation genetic markers in order to understand the amount of, and distribution of, genetic diversity in the mala (*Lagorchestes hirsutus*), a species of wallaby extinct on mainland Australia and preserved in captive populations.

These data are being used to inform conservation management.

Chapter 3: Examining the factors that influence genetic diversity at MHC

In this chapter a meta-analysis is used to investigate a number of questions around the amount of genetic diversity at the major histocompatibility complex and its use as a genetic marker.

Chapter 4: A Novel method of genotyping MHC for conservation genetics

Chapter 3 suggests that MHC diversity may be accurately described by the amount of, and distribution of, microsatellite diversity. Herein we attempt to develop a novel method of genotyping MHC specifically for conservation genetics.

Chapter 5: Investigating disruptive selection due to land clearing in E.cunninghami

Chapter 4 suggests that genotyping by next generation sequencing approaches are promising, if challenging, tools for conservation genetics. To further explore this theme we attempt to re-examine

a skink population, previous investigated using microsatellite markers, using a RAD-TAG sequencing, a type of genotyping by next generation sequencing approach.

Chapter 6: General Discussion of the above

2 Using genetics to inform the management of a threatened species, the mala (*Lagorchestes hirsutus*)

The mala is a nationally important conservation success story. Before the last wild population went extinct in 1992, a small number of individuals were captured and used to begin a captive population. That population has grown to more than 500 individuals distributed across six sites. This study used microsatellite and mtDNA markers to investigate the genetic health of the captive population and make management recommendations. A pilot study was run to trial minimally invasive methods of obtaining genetic data (hair and blood spots) in order to provide the best balance between high quality data and low invasiveness. *The main study used 4 microsatellite and 1 mtDNA marker to investigate genetic diversity within the mala. A minimum of 6 unequally represented founders was discovered with 5 mtDNA sequences, and the moderate Fis and Fst values were consistent with the multiple founder effects experienced during the captive history. Recommendations were made for outcrossing programs and future research.*

2.1 Introduction

Global biodiversity faces significant threats that have resulted in extinction rates 1000 times greater than prehuman levels (Pimm *et al.* 1995). In Australia, invasive species, climate change, habitat destruction and overexploitation have been identified as key threats to biodiversity (EPBC 1999). These have contributed to Australia's mammal extinction rate which is currently the highest in the world (Hobbs *et al.* 1998, Johnson 2006) and is predicted to continue to remain high (Williams *et al.* 2003, Johnson & Isaac 2009) with 26% of Australian mammals restricted to less than 20% of their former ranges (Short and Smith 1994). Of particular concern are Australian mammals in the critical weight range which Burbidge & McKenzie (1989) defined as between 35g- 5500g, especially ground living species in low rainfall environments (Johnson & Isaac 2009). These species have suffered greater population declines and increased extinctions (Johnson & Issac 2009), in large part due to the introduction of cats (late 18th century) and foxes (1885). The persistence of introduced predator populations, as well as other continuing threats means that captive management programs are increasing required to preserve these threatened species (reviewed in Snyder *et al.* 1996).

The ultimate goal of species captive programs is a successful reintroduction of self-sustaining populations to the wild. Successful reintroductions, however, despite notable successes such as the black footed ferret (Dobson & Lyles 2000) are not easily achieved. In Australia, 29 of 54 documented species reintroductions to date have failed (Sheean *et al.* 2012), while worldwide the success rate is estimated at 11% (Beck *et al.* 1994). Several reasons for the low rate of success in reintroductions to the wild have been suggested. These include the continued presence of introduced predators and changed environments. The re-introduction area plays a role in reintroduction success, with reintroductions to the same area that the source population inhabited showing greater success in plants (e.g. Knapp and Dyer 1998) and animals. Furthermore, the loss of genetic diversity (e.g Ewing *et al.* 2008) and adaptation to captivity (Frankham 2008) in captive populations is a sometimes overlooked problem that may contribute to reintroduction failure.

Captive populations are often established with small numbers of individuals, resulting in genetic bottlenecks and inbreeding, even when census size subsequently increases (e.g. Hoelzel 1999, modelled in England *et al.* 2003 and reviewed in Frankham *et al.* 2009). The relationship between inbreeding and fitness is discussed in detail in chapter 1 and reviewed by Reed & Frankham (2003). Additionally, unequal founder representation may cause genetic differentiation between captive populations (e.g. Frankham *et al.* 2009). These effects can be mitigated by effective management of genetic diversity. As discussed in Chapter 1, successful genetic management reduces inbreeding and, therefore, inbreeding depression. Further, preserving genetic diversity within captive populations retains their evolutionary potential and long term viability (Frankham 1980, Frankham *et al.* 2010). However, in captive populations genetic diversity must be managed and not simply maximized. The management of genetic diversity must balance minimizing inbreeding against adaptation to captive conditions. Adaptation to captivity has been acknowledged since the era of Darwin but only recently have the underlying mechanisms have been understood. As evolution can be expressed simply as Genetic Diversity x Selection x Time, any measures taken to maximize genetic diversity, including minimizing inbreeding depression, also maximize adaptive potential. Furthermore, species with short generation lengths, such as the mala (with females maturity at 5-23 months and male reproductive maturity at 14-20 months (Langford 2009)), adapt more quickly to captivity. This was modelled by Woodworth *et al.* (2002) who demonstrated, using *Drosophila*, that adaptation to captivity increased with population size and that this adaptation reduced fitness in simulated wild conditions. Adaptation to captive conditions is believed to be one of the contributing factors to the high number of reintroduction failures (e.g. Araki *et al.* 2007, reviewed in Frankham *et al.* 2009). Therefore, genetic management is a key component of captive programs which aim to re-introduce animals into the wild.

The mala (*Lagorchestes hirsutus*) is one of numerous Australian mammal species that requires active management to avoid extinction. The mala (or rufous hare wallaby) is a small endemic wallaby that once inhabited much of the Hummmock Grasslands of the central Australian deserts (Maxwell *et al.*

1996). The species was first described by Europeans in 1884 (Gould 1884) but indigenous knowledge of the species is much older, with mala having both spiritual and dietary importance to many indigenous groups including the Anangu people of central Australia. The International Union for the Conservation of Nature (IUCN) Red List threat status of the species has escalated from *Rare* (1982), to *Endangered* (1994) and then to *Vulnerable* (1996). Addtionally, mala is listed as *Endangered* under the EPBC 1999 and *Extinct in the Wild* (Langford & Burbidge 2001). The trend towards extinction in the mala is a direct result of the European settlement of Australia.

The key historical threats to mala were the introduced predators and changes to fire regimes following European settlement. These two factors are causally linked to the extinction of the mala on mainland Australia. Foxes (*Vulpes vulpes*) were introduced into Australia in 1855 (Saunders & McLeod 2007) and wild populations established by 1870. Cats (*Felis catus*) were introduced in the late 18th century and feral populations bolstered by planned releases of thousands of animals into central Australia in the mid-19th century (Dickman 1996). Both foxes and cats have been shown to have catastrophic effects on native mammal populations (Dickman 1996, Moseby *et al.* 2009). The current range of both species completely encompasses the mala's historic range (Figure 2.1). It is therefore unsurprising that predation by foxes and cats has been identified as a key reason for the extinction of mala on the mainland (MALA reintroduction program fact sheet). Foxes, in particular, pose a threat to mala. A single fox has been known to kill dozens of mala in a few days once gaining access to a population (Pers. comm. Andrews 2009) and is believed to be responsible for destroying one of the two last wild populations (Lundie-Jenkins 1989). Furthermore, European settlement drove changes in the fire regime imposed by Aboriginal peoples, adversely affecting mala habitat (Mala reintroduction program fact sheet).

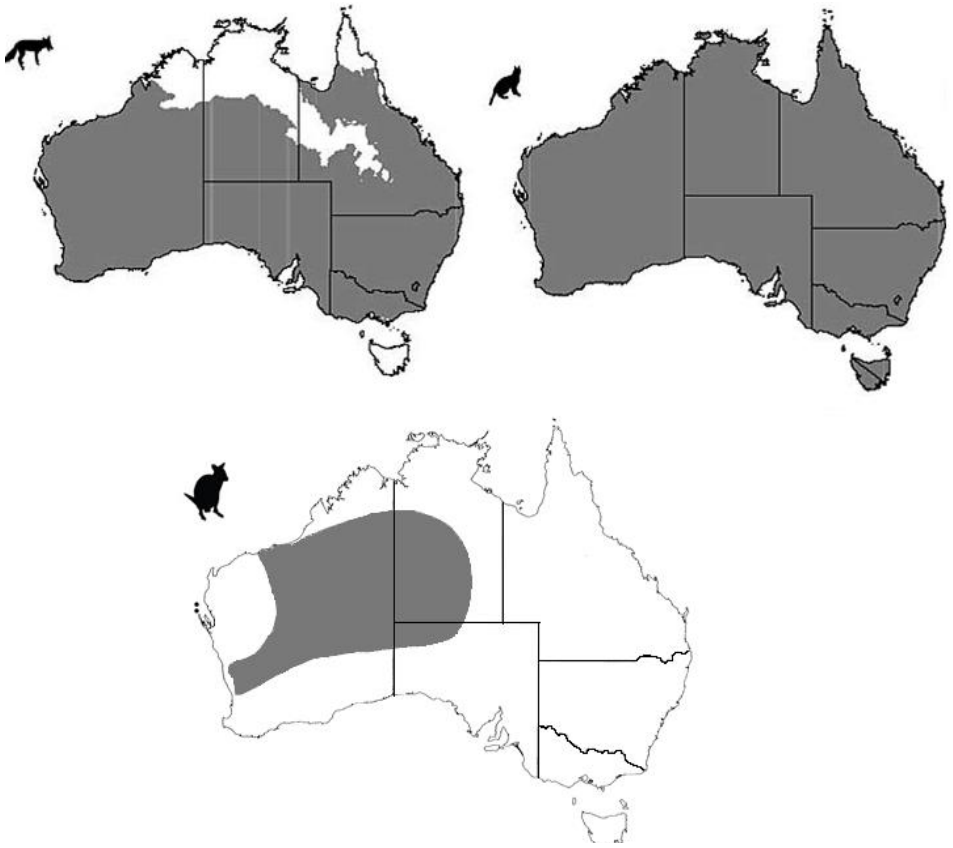


Figure 2.1: The current distribution of red fox and feral cats in Australia and the historic distribution of the mala (Adapted from figures by Langford 1999 and Reddiex & Forsyth.2004)

Predation and fire reduced mala numbers until the last mainland population was destroyed by a wildfire in 1992 (Pavey 2006). Prior to this extinction, however, a captive breeding program was launched using 7-22 wild caught mala, thenumber of founders varies depending on which source is citied (Hardman & Morrow 2006, Goodall 2009). The mala program is correctly hailed as 'an important success story' (Goodall 2009) because committed husbandry and forward-thinking policy (Langford 1999, Goodall 2009) have rapidly increased the census size of the population. Captive mala were initially housed at the Arid Zone Research Institute in Alice Springs, but as the program successfully increased the population size, animals were moved (Fig. 2.2). This has resulted in the existence of captive, semi-captive (supplementary feed and predator fencing) and wild island populations. The creation of these populations was based on logistic and practical requirements and

did not consider genetic factors. This may have resulted in a distribution of genetic diversity that is not optimal for species survival.

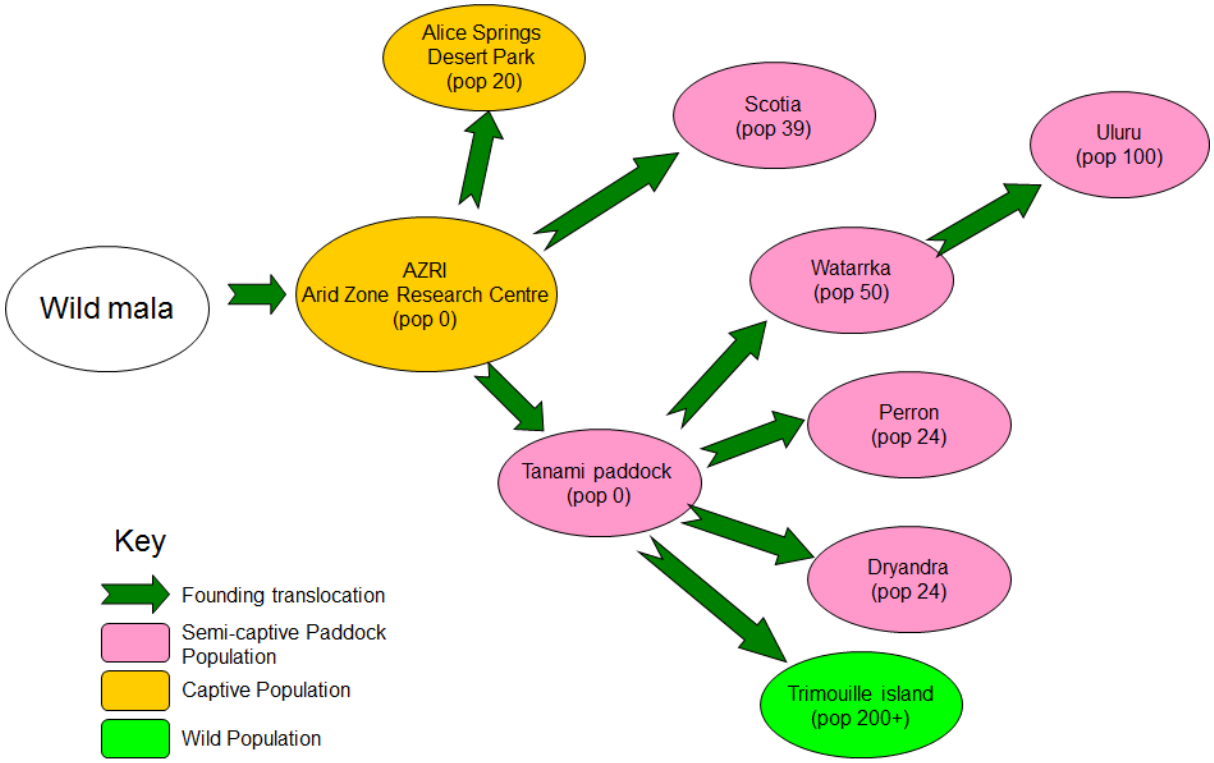


Figure 2.2 The order and source of founding from the wild population, current population size and level of captivity (completely captive or semi-captive populations kept in predator proof paddocks) of mala populations since the establishment of the captive breeding program (P.Comm Atchison).

Predation by foxes has thwarted several attempts to re-establish wild mala populations on the mainland (Mala reintroduction fact sheet, reviewed in Sheean *et al.* 2012). The success of the wild island population, however, suggests adaptation to captivity may not yet be a major problem for this species. However, as an effective fox control has yet to be conceived for mainland Australia, mala are likely to remain in managed populations for an extended period. The small size of the founder population in the mala program and their lengthy stay in small captive populations means that effective genetic management to avoid harmful adaptation will be an important consideration in conserving this species.

Therefore, this study aims to provide genetic data relevant to informing decision making in mala management. By determining the level of inbreeding in the population the likelihood of inbreeding affecting fitness can be assessed. Furthermore, by determining the amount of genetic diversity

partitioning between the sub-populations, this study will determine if outcrossing will reduce inbreeding. Additionally, understanding the amount and distribution of genetic diversity in a program where these factors were not considered at founding the mala captive breeding program can provide insights into other similar programs.

2.2 Methods

Pilot Study: DNA extraction techniques in a threatened species

Mala, like many endangered species can be difficult to both capture and sample and require intelligent sampling design. Mala in particular are prone to injury during capture and capture myopathy (Freeguard & Ritcher 2009). Therefore, the amount and invasiveness of handling needs to be minimised when sampling genetic material. For that reason tissue sampling (e.g. ear punch as per Spencer *et al.* 1991) in order to obtain DNA may not be optimal. However, invasiveness must be balanced with the quality of DNA obtained, as low quality DNA is unsuitable for many applications (Gagneux *et al.* 1997, Taberlet *et al.* 1999). For this reasons we undertook a pilot study with hair, tissue and blood samples already available (where hair and tissue samples were from the same individuals) to determine the best practice for sampling in this species.

DNA extraction from hair

Strands of hair with follicles attached are a non-invasive source of DNA. DNA extraction from hair has been used both in forensics (e.g. Walsh *et al.* 1991) and conservation programs (Sloane *et al.* 2000). In conservation it has two main advantages. Firstly by using double sided tape around burrows or feeding sites, it allows sampling of species that are difficult to capture by traditional methods (Sloane *et al.* 2000). Secondly, it is a non-invasive approach that does not cause significant stress to the animal, which is advantageous in endangered species where handling and/or capture can cause morbidity or mortality (Sloane *et al.* 2000). However, getting high quality DNA from hair can prove difficult (Gagneux *et al.* 1997, Taberlet *et al.* 1999) with high rates of allelic drop outs in

microsatellite markers which can bias genetic analysis (Gagneux *et al.* 1997) of particular concern to conservation geneticists, as they decrease observed heterozygosity and thus suggest elevated levels of inbreeding in the population.

DNA extraction was attempted using a Dithiothreitol (DTT) extraction methodology with a commercially available kit (DNAeasy Qiagen Kit 2006) following the manufactures instructions. Additionally, DNA extraction was attempted using a 5% Chelex suspension using modified by Baldwin *et al.* (2010) from Gagneux *et al.* (1997) (full protocol in appendix 2.1). Between five and ten hairs were used per extraction.

DNA extraction from blood stored on microcards

Bloodspots were taken from a small number of captive animals at Alice Springs Desert Park by park veterinarians as part of routine check-ups and from dead animals for the pilot study. DNA isolation from blood spots stored on microcards provides an alternative method of obtaining DNA that is minimally invasive. DNA extraction from blood stored on cards is a well-established process (e.g. McCabe *et al.* 1987) with several commercial kits available (e.g. Qiagen DNEasy kit). Obtaining blood spots is minimally invasive, the cards can be stored for long periods and yield high quality DNA (Mullen *et al.* 2009). DNA extraction from blood spots stored on cards was chosen as a compromise between invasiveness and proven performance. Extraction was carried out in accordance with the manufacturer's recommendations (Qiagen).

DNA Extraction from tissue

DNA extraction from tissue via salting out is a simple fast and cheap method of reliably producing high quality DNA. As two tissue samples were available (from deceased animals from the captive population) we attempted a DNA extraction using a modified salting out protocol (Sunnucks & Hales 1996) that has proven successful for a number of species (e.g. Stow *et al.* 2004a, 2004b, Hoggard *et al.* 2011, Repaci *et al.* 2006). The technique can be used with small tissue samples which do not cause significant distress to most animals, therefore it has applied to sampling in catch-sample-

release studies (e.g. Stow & Sunnucks 2004). However, tissue sampling is more invasive than other methods of obtaining DNA used in this study , so, even though small tissue samples have been shown to cause no greater medium term harm to mice (Cinelli *et al.* 2007), this was our least favoured option.

A sample of tissue was cut using a sterile razor blade and placed in 1.5ml microfuge tube with 300 µL TNES (50 mM Tris, pH 7.5, 400 mM NaCl, 20 mM EDTA, 0.5% SDS) with 100 pg/mL Proteinase K overnight. Proteins were precipitated by adding 5M NaCl and pelleted by microfuge at 14,000 rpm for 5 min. The supernatant was transferred to a new 1.5ml microfuge with 700 µL of ice cold 100% ethanol to precipitate DNA. DNA was pelleted by microfuging (as above) and washed in 70% ethanol and then air dried. DNA was resuspended in 30 µL ddH_20.

PCR amplification of markers and optimization:

In order to test quality of the DNA extraction the DNA was used to amplify both mitochondrial (Folmer *et al.* 1994) and microsatellite (Eldridge *et al.* 2004) markers, multiple rounds of testing were required and the final and most successful reactions are described below.

Direct amplification from blood cards

Where PCR was attempted directly from blood cards the suggested protocol was followed (Chum & André 2013). A sterile razor blade was used to cut a 3mm disc from the blood card which was added to 20 µl of ddH_20 and incubated at 50 $^{\circ}$C for 3minutes before being added to the reaction mixture.

PCR of mtDNA markers

The universal DNA primers, LCO1490 (5'-ggtcaacaaatcataaagatattgg-3) and HCO2198 (5'-taaacttcagggtgaccaaaaaatca-3'), were used to amplified a 710-bp region of the mitochondrial cytochrome oxidase subunit 1 gene (Folmer *et al.* 1994). PCRs were performed in a final volume of 10 mL containing 0.5U Taq DNA polymerase (Promega), 10 mM forward primer, 10 mM reverse primer, 8 mM dNTPs, 1 unit Taq Buffer (Promega) and 2.0 mM $MgCl_2$. PCR amplifications had an

initial denaturation at 94 C for 4 min followed by five 'touch down' cycles of 94 C denaturation for 15 s, annealing temperatures (55˚C, 53 ˚C, 51˚C, 49˚C, 47 ˚C) for 30 s and an extension step of 72 ˚C for 80 s. After the final touchdown cycle, another 30 cycles were carried out at 50 ˚C annealing temperature with a final extension of 5 min at 72 C. PCR products were visualised on 2% agarose gel 2 with PAGE GelRed (Biotium) with a 100 b.p. ladder as a positive control.

PCR amplification of microsatellite markers

This study attempted to amplify six microsatellite markers (see table 3.1). PCRs were performed using a positive control, from a tissue sample which had been sequenced successfully and a negative control of pure water. Reactions were carried out in a final volume of 10 mL containing 0.5U Taq DNA polymerase (Promega), 10 mM forward primer, 10 mM reverse primer, 8 mM dNTPs, 1 unit Taq Buffer (Promega) and 2.0 mM $MgCl_2$. PCR amplifications had an initial denaturation at 94˚C for 4 min followed by five 'touch down' cycles of 94 C denaturation for 15 s, annealing temperatures (60˚C, 58˚C, 56˚C, 54˚C, 52˚C, 50 ˚C) for 30 s and an extension step of 72 ˚C for 80 s. After the final touchdown cycle, another 30 cycles were carried out at 50 ˚C annealing temperature with a final extension of 5 min at 72 ˚C. PCR products were visualised on 2% agarose gel 2 with PAGE GelRed (Biotium) with a 100 b.p. ladder as a positive control.

2.3 Results

The extraction and amplification was successful using from both tissue and bloodspots stored on FTA microcards but all forms of DNA extraction from hair failed. A representative subset of these results is detailed in Figure 2.3 below. Conclusions from pilot study

The pilot study suggests that DNA extraction via hair did not produce DNA of high enough quality for the main study within the timeframes of the project, as all DNA extractions from hair failed to produce a band on an agarose gel or a PCR product when tested with known primers. As DNA extraction from blood spots and DNA extraction from tissue both produced high quality DNA, blood spots were used for the main study as this process of obtaining DNA is less invasive. This is despite the Bloodspot samples being of lower than expected volume.

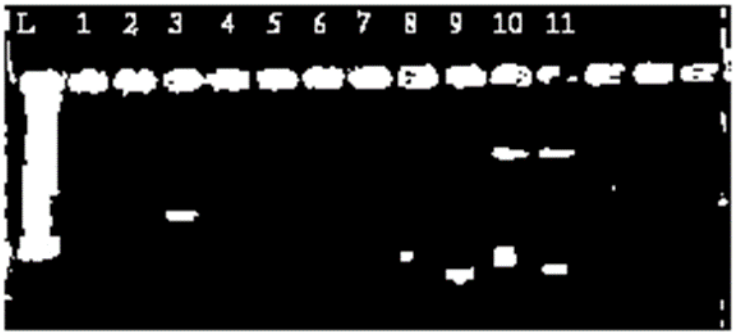


Figure 2.3 : The results of DNA amplification (Gel picture enhanced) of Me17 (microsatellite loci) and mtDNA CO1 region, only lanes 3, 10 and 11 show amplification in this gel. DNA extraction methods are abbreviated as: P.C.P, Phenol-Chloroform purification of a salting out DNA extraction; FTA, DNA extraction from bloodspots on FTA microcards; DTT DNA extraction using Dithiothreitol. Layout is shown below

Lane	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Loci	Me17	Me17	Me17	Me17	Me17	CO1	CO1	CO1	CO1	CO1	CO1
Source	Hair	Hair	Blood	Blood	Hair	Hair	Hair	Hair	Hair	Blood	Blood
Extract Method	Salting out	P.C.P.	FTA	FTA	DTT	Salting out	P.C.P	DTT	DTT	Direct PCR	FTA

Methods Main Study

Capture and Sampling

As mala are easily stressed, this study used Bromilow soft traps to minimize capture stress (Kinnear *et al.* 1988) and ran sampling as part of routine catch-ups wherever possible. The study aimed to sample 30 individuals (or as many as were available) from five of the six extant mala populations.

Timourille island, though the largest population, was not available as it is remote and rarely visited. Animals were caught in soft traps. A lance was used to obtain a blood spot, which was stored on Whatman FTA Microcards, and the animal was released within five minutes of capture. Actual sample numbers varied slightly based on capture rates and are recorded below.

Table 2.1 sample sizes obtained per population. A.S.D.P. represents Alice Springs Desert Park. Dryandra and Trimourille Island were not sampled.

Population	Uluru	Watarrka	A.S.D.P.	Peron	Scotia	Dryandra	Timourille Island
Sample Size	30	30	11	27	15	n/a	n/a

DNA extraction and marker amplification

Due to small blood spot volume, direct amplification was not used. As per the pilot study DNA was extracted from FTA microcards using the Qiagen mini DNAeasy kit (Qiagen). The details of this are recorded above. DNA extraction was confirmed by running the samples on 2% agarose gel and visualized with PAGE GelRed (Biotium) with a 100 b.p. ladder as a positive control. Full results of these analyses are presented in Appendix 2.

One mtDNA locus and a total of 8 microsatellite loci were amplified but only 4 produced scoreable genetic data (as the others did not produce clear scorable peaks) . The loci used and primer sequences are recorded below. Due to the size of blood spots the study had a limited amount of DNA from most samples and whilst microsatellites for the mala exist (Eldridge *et al.* 2004) they were optimised using radiotagged dNTP's on polyacrylamide gel and had to be re-optimised for flurochrome based genotyping, this involved varying magnesium ion and annealing temperature. The genetic markers scored are detailed in the table below.

Table 2.2: Details of genetic markers used in mala study, * denotes primer developed by Taylor & Cooper 1998, ** denotes primers developed by Folmer et al. 1994 and *** denotes primers developed by Pope et al. 1996. The use of an @ symbol denotes primers where genotyping was attempted in all 123 samples.

Locus	Annealing Temp	Size Range	Primer Sequence Forward	Primer Sequence Reverse
Me14*	60	160–210	ACTGGGGCAAAATACAGGG	CCAGTGGGAGTTGAGTCATATC
@Me15*	60-55	225–270	GGAGCCATCTTAGGAAGACT	CTTGTCTCACACAGCCTAGG
@Me16*	60-50	240–280	TTGGGCTGTCTCCTCATCTG	GGAATCCTCAGGCGCTATGA
Me17*	60-50	110–140	GGGGTATGAACTAGATGACC	CCAGACAAGTAAGGATGCTG
LCO1490 / HCO2198**	60-50	710	GGTCAACAAATCATAAAGATATTGG	TAAACTTCAGGGTGACCAAAAAATCA
@Y76***	58	163-172	AGAGTAGTAATTTCAGTCCTTTG	CTGAACCTTATTCTCCCACAT
@Y170***	55	145-177	GGACTCAAACCCAACACTAGC	TGCATGCCTTTGTCATACACG

mtDNA markers were amplified using the protocol described in the pilot study (above) and sequencing was carried out at MQ Sequencing Facility on a ABI 3130 Genetic Analyzer (Applied Biosystems). Microsatellite markers were amplified and visualised and scored as described in the pilot study.

Statistical Approaches to Determining founder number

The minimum founder number was determined from mtDNA data using MEGA 5.5 (Tamura *et al.* 2011), to construct a Neighbour joining tree (full method in appendix 2.2) f mtDNA sequences where each unique sequence must represent at least one female founder. Microsatellite alleles were counted as a single founder can contribute, at most, two alleles to a locus.

Examining inbreeding & Genetic partitioning

To determine the level of Inbreeding, inbreeding coefficients were calculated using Genepop (Rousset 2008) and F-Stat (Goudet 1995). These values range between 0 and 1 with higher values indicating more inbreeding.

To examine the partitioning of the genetic diversity in the population the programs GenePOP (Raymond & Rousset 1995) and TFPGA (Miller 1997) were used to determine Fst

values. Furthermore, UPGMA clustering (using TFPGA, full method in appendix 2.4) was created to visualize the distribution of genetic diversity using microsatellites and a phylogenetic tree was made using MEGA 5.5 (Tamura *et al.* 2013) for the mtDNA sequence.

Modelling the effect of migration on genetic diversity

To model the amount of migration that would have a significant effect on genetic differentiation of a population the program EASYPOP (Balloux 2001) was used with the following parameters: Diploid population, Two sexes, Random Mating, Island Model of Migration and population sizes from the most recent mala census.

RESULTS

Inbreeding in the mala population

There was significant inbreeding detected in the mala populations. Furthermore, several of the populations were found to have moderate to high Fis values. On the other hand, as table 2.3 shows (with full results in appendix 2.3), the Fis value for the Alice Springs Desert Park is much lower than other populations indicating that our results were consistent with inbreeding causing high Fis and not null alleles. Therefore, there is evidence of overall inbreeding in the mala population and multiple founder effects in the sub-populations as evidenced by increased Fis scores.

Table 2.3 The coefficient of inbreeding (Fis) for each population as calculated by F-Stat (Goudet 1995).

Population	Uluru	Watarrka	Alice Springs Desert Park	Peron	Scotia
Fis Value	0.1974	0.4173	0.0583	0.2311	0.2365

Genetic Partitioning in the mala population

Analysis revealed significant genetic partitioning between populations (95% confidence interval for whole population Fst 0.0469-0.1796 calculated using TFGPA). The genetic distance (Fst) between population pairs varied considerably as shown in table 2.4 below.

Table 2.4 Calculation of the fixation index Fst using F-stat (Goudet 1995) for population pairs, A.S.D.P. represents the Alice Springs Desert Park.

	Uluru	Watarrka	A.S.D.P	Peron
Watarrka	0.0387			
A.S.D.P	0.1162	0.0872		
Peron	0.0782	0.0329	0.0752	
Scotia	0.2383	0.1958	0.2555	0.2266

A second analysis using UPGMA clustering (using TFPGA) suggests that Scotia is the most genetically isolated group (full results in appendix 2.4).

Founder representation in the mala population Alignment and analysis using MEGA 5.5 revealed 5 unique mtDNA sequences which were split among all captive groups. However, as the figure below describes one mtDNA haplotype was found in 95% of samples sequenced, suggesting that a single female/ group of related females are over represented in the extant captive population.

The analysis of microsatellite data suggests the possibility of a smaller number of founders than mtDNA. As the number of alleles found per microsatellite locus ranged between 4 and 9, a 5 founder population is possible, if all except one were heterozygotes and all had unique allleles

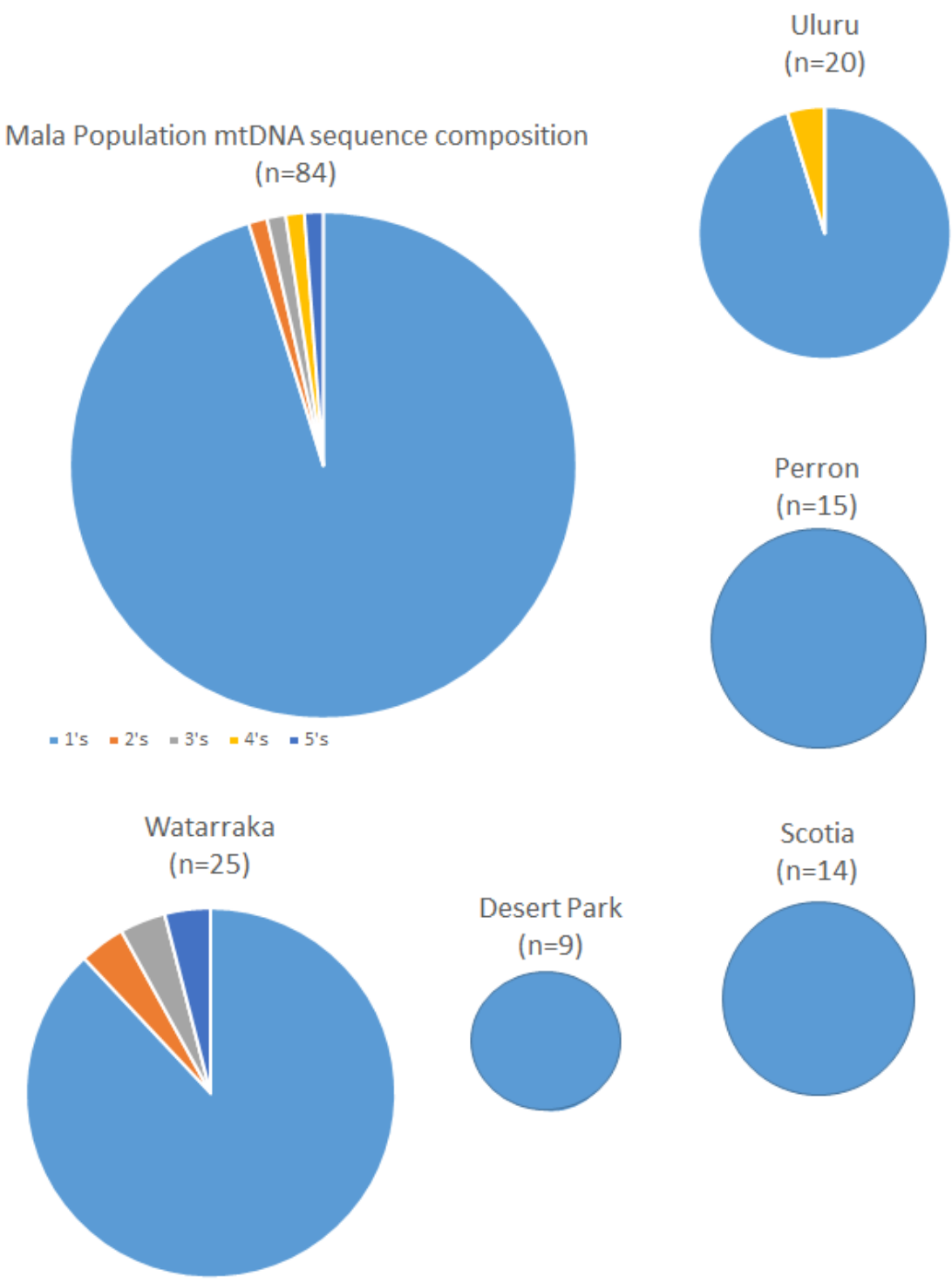


Figure 2.4: The mtDNA haplotypes present in the mala population. Only 5 sequences were present in the entire sample 4 of which only occurred once

Modelling the effects of migration

The results of the modelling carried out in Easypop are shown in figure 2.5 below. These were generated using a diploid model with population approximated as per figure 2.2, an island model of

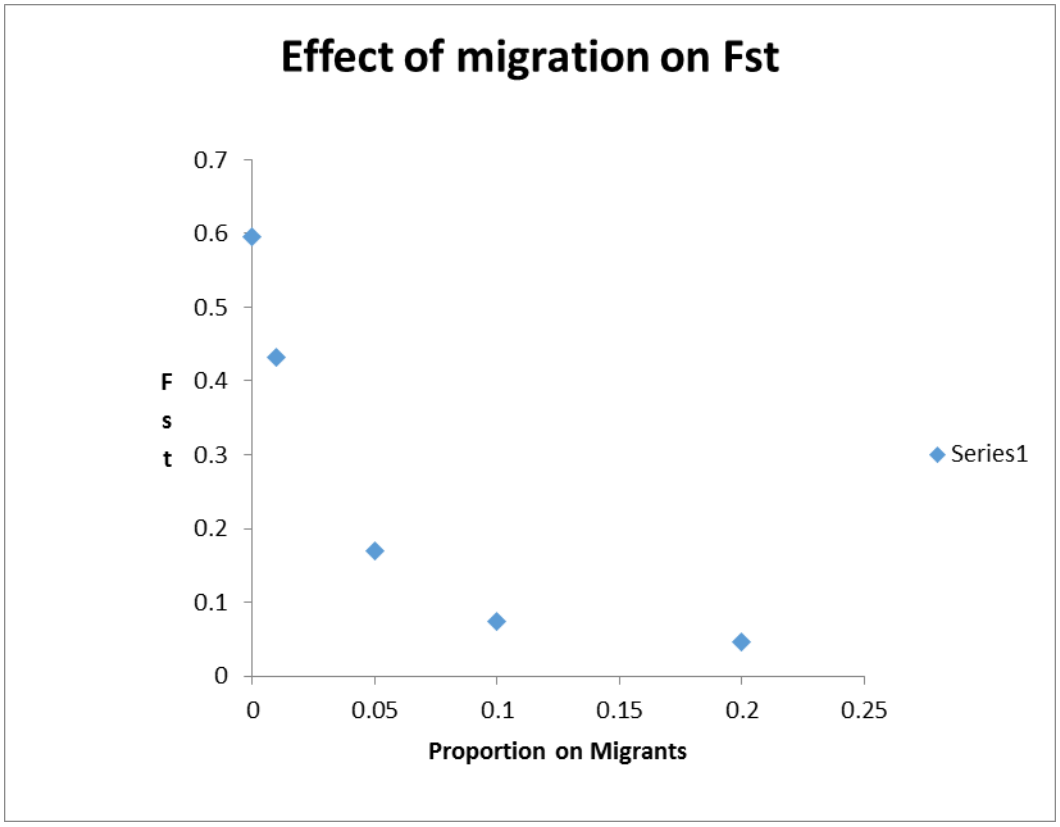


Figure 2.5*:* Results of the EasyPop simulation showing the effect of migration on Fst. The most significant reduction in Fst occurs when the proportion of migrants is increased from 0 to 0.05

migration was used with a unified mutation model and the number of loci genotyped successfully in this study. By adjusting the proportion of migrants this modelsuggested that over 100 generations a migration rate of between 10% and 5% of the overall population will reduce genetic differentiation between populations.

2.4 Discussion

Evaluation of the genetic health of the mala population

The captive mala population represents more than five founders. We estimate the least number of founders represented by taking the larger of the two minimum number estimates of founder representation (5 from microsatellite data and 6 from mtDNA haplotypes plus at least one male); this is still substantially less than the 22 founders for the population which would suggest a maximum of 44 alleles per microsatellite loci and 21 mtDNA haplotypes (as recorded by Hardman & Morrow 2006). Frankham *et al.* (2010) suggests that typically the genetic diversity of the founders of captive breeding programs is lower than would be expected as the population had been small before the beginning of the captive population. Furthermore, earlier work by Frankham (1995) suggests an order of magnitude difference between the census size of a population and the effective size of the population for genetic analysis (Ne). In light of this the number of genetically distinct founders represented in the mala population is higher than expected. This may be due to the rapid decline in mala numbers before the captive population was founded (from equations in Frankham *et al.* 2010). A rapid decline, as opposed to a gradual decline to small numbers, creates a genetic bottleneck but where numbers are quickly increased, as with the mala, results in a smaller loss of heterozygosity.

The founder number can not be definitely known but all analysis suggest it to be much lower than Frankham's recommendation of 50 genetically distinct individuals (Frankham 1980, Soule 1980) for short term population viability. However, this number is not atypical for other captive breeding programs. Of 17 captive breeding programs reviewed in Frankham (2009) 14 had fewer than 20 founders and 6 had fewer than 10. Therefore, this result supports previous suggestions (Frankham 2009) that in the future captive programs must act earlier and obtain larger numbers of founders to preserve genetic diversity, and thus evolutionary potential. Since there is no longer a wild population with which to supplement the captive population, it is important to maintain founding genetic diversity in the captive mala population.

The captive mala population suffers from unequal founder representation. The large proportion of mtDNA sequence from a single maternal lineage (92%) indicates that a single female, or group of related females is disproportionately represented among the surviving mala. A large bias in founder representation is not unusual in captive populations (Frankham 2009) as some individuals are more likely to breed in captive environments than others (Lacy 1989). This is likely to be the case here, but serial bottlenecks (Gautschi *et al.* 2002) as well the effects of drift in small populations (Kuehn *et al.* 2003) and skewed breeding success (Jamieson 2011) could also have created the skewed representation discovered. However, theoretically by equalising founder representation (Lacy 1989, Doyle *et al.* 2001) the effective size of the population grows. In practice this means that genetically less valuable individuals should be used as the first wave for reintroduction attempts (Earnhardt 1999, Frankham 2009), while the program should attempt to increase representation of the rarer genotypes. This over-representation contributes, in some part, to the significant levels of inbreeding in the population.

Inbreeding was detected in individual mala populations with the A.S.D.P. having the lowest level (Fis = 0.0583). Given the small number of founders and the number of generations, at least 15 given 5-23 months till sexual maturity, inbreeding in the captive population should not be surprising. However, the large variation in Fis values between the different populations (0.0583-0.4173) is informative. Whilst the other populations allow for uncontrolled mating, the A.S.D.P. houses a captive program with studbook management. Studbooks provide pedigree information and thus can used to avoid highly inbred matches (Fernandez & Caballero 2001) and are routinely used in captive breeding programs e.g. Golden Lion Tamarinds (Ballou & Lacy 1995). The lower Fis result is an endorsement of studbook management even when the genetic composition of the founding population is unknown. Nevertheless, inbreeding at similar levels to the overall mala population have been shown to reduce fitness in captivity (Laikre & Ryman 1991) and the wild (Crnokrak & Roff 1999). It is therefore reasonable to suppose the mala may be suffering from inbreeding depression which could be

reversed by migration between populations, or outcrossing. However, in order for migration to have this effect there needs to be genetic differentiation between the populations.

The genetic differentiation in the extant mala populations was caused by bias in foundation of captive populations and can be improved through management. The genetic differentiation in the extant mala population (95% C.I. 0.0469-0.1796) indicates some genetic partitioning, particularly between Scotia and the other populations. This is roughly what would be expected given the foundation of populations (summarised in figure 2) and therefore these findings can be extrapolated to the populations from which we were unable to acquire data from (Trimouille island and Dryandra). Without engaging in the SLOSS (single large or several small populations) debate begun by Diamond (1975), the current population structure is the only logistically feasible solution. Furthermore, it may minimise the risk of population collapse (Margan *et al.* 1998) and will reduce adaptation to captivity (Frankham *et al.* 2009). However, outcomes for the species can be improved by controlled migration between captive populations, as recommended for the species (Richards 2012), which increase population viability and future reintroduction success.

Recommendations for genetic management

Recommendations based on limited genetic data are difficult. On one hand, it, this study used a small number of individuals and a small number of loci, therefore any results can be misleading. Drastic action taken based on misleading results may result in species loss.On the other hand, conservation may require action before the complete genetic data can be generated. On the balance, we suggest that a more robust genetic analysis take place, ideally one such as RAD-TAG used in chapter 5 before any change in management take place. For the purposes of discussion we have included recommendation if, and only if, the outcomes of that more thorough investigation match the outcomes of this one. Without a more comprehensive understanding of the genetic structure any change in management is premature, particularily given the continued growth in the population currently.

We recommend immediate outcrossing to reduce inbreeding depression. The immediate outcrossing equalises the distribution of genetic diversity in order to minimize the effects of inbreeding (as discussed in chapter 1). According to EASYPOP modelling we estimate the movement of 25 animals between populations will counteract the genetic differentiation that has occurred due to current management strategies. By re-distributing the genetic diversity more evenly between sub-populations we reduce the risk that disease, predator entry or disaster could cause catastrophic damage (Richards 2012). However, the risk of moving disease with animals needs to be considered as does the risk to the population of movements of 5% of the animals, even if they are between secure captive and semi captive environments. Furthermore, it is also important to note that the numbers of mala to be moved would be different if we were able to survey the island population, and thus this study recommends that the genetic data from that population be included in the overall strategy as soon as possible.

In order to manage the mala population the best approach may be periods of isolation following outcrossing events.By isolating the populations from each other the speed at which adaptation to captivity occurs will be decreased (Frankham *et al.* 2009). Furthermore, as captive conditions are different in the different areas (e.g. Uluru, Scotia) different genes will be favoured. This helps to prevent alleles being lost. Additionally, genetic drift may also create unfavourable outcomes as the selective pressures on genes that are necessary for survival in the wild are relaxed in captivity (Bryant & Reed 1999) and selection is less effective in small populations (Willi *et al.* 2006). However, as genetic drift will create different outcomes by chance in the different populations, the isolation of populations will allow more diversity to be retained (Lacy 1987).

The periodic migration between captive groups allows for long term minimisation of loss of genetic diversity and inbreeding which is associated with groups containing few individuals. General consensus is that inbreeding should be allowed to build up in populations to an Fis level of 0.2 before outcrossing occurs (Frankham *et al.* 2009). In the mala, where sexual maturity is reached at

14 months and survival is up to 8 years in captivity (Richards 2012), this could take a number of years depending on the continued growth and maintenance of the captive populations. Therefore this system addresses the short and long term aims of the mala recovery program by balancing the need to minimize inbreeding and adaptation to captivity and it does so using relatively low-intensity management methods. To complement and enhance the isolation – periodic migration model we also suggest preferential breeding of individuals with rare alleles.

The unequal founder representation in the extant population can be partially redressed by taking advantage of the success of studbook managed breeding at the Alice Springs Desert Park (A.S.P.D). Mala at the A.S.D.P. facility have the least inbreeding (suggested by a low Fis value of 0.06). However, the A.S.D.P. sample contained none of the rare mtDNA genotypes and lacked several rare alleles present in other populations. We suggest moving 20 individuals with rare genotypes to the desert park facility and continuing the studbook management in order to increase the frequency of these rarer genetic variants. This approach has been suggested for Xalda sheep (Goyache *et al.* 2003). As the A.S.D.P. has an excellent record of reproductive success, it would be possibly to purposefully increase the numbers of these rare variants. Then the offspring of these individuals could be translocated to the other semi-captive environments. Thus by increasing the numbers of individuals carrying rare alleles the underrepresentation of some genetic lineages in the current population would be minimized. However, due to the small number of markers amplified it may be beneficial to first conduct a more detailed genome screen of the type carried out in chapter 5. In essence this approach would entail a supportive breeding program within the larger overall captive program.

Limitations of the study

The sampling regime used in this study creates significant limitations that need to be acknowledged and in the longer term redressed. The study did not sample from two populations, Dryandra with approximately 24 individuals and Trimouille Island with over 200 individuals. Unfortunately Dryandra

could not be surveyed and Trimouille Island is rarely surveyed as atomic testing was carried out on it in 1956 (Cooper *et al.* 1983) with the detonation of two bombs both on and off the island with kilo tonne range payloads (Moroney & Cooper 1982). Although the island, excepting ground zero for bomb detonation, is now considered not to have dangerous levels of radiation (Cooper *et al.* 1983) management teams rarely go to the island. This creates a large gap in our knowledge of the mala population.

The Trimouille Island population is the largest population of mala and also is entirely wild with no predator control or supplementary feeding. The population has presumably thrived due to the lack of foxes and cats on the island. An understanding of the genetic diversity in this population would allow for recommendations for the foundation of future island populations, e.g. the planned reintroduction to Dirk Hartog Island (Richards 2012). To this end we recommend redressing this lack of sampling. To complement additional sampling it is important to redress the limitations of the genetic markers used.

The number and type of genetic markers used created limitations for the study. Although 8 microsatellite markers were available for the mala (Eldridge *et al.* 2004), due to limitations in the amount of DNA obtained for each sample from the blood spots we were only able to optimise four microsatellite loci (and a mtDNA locus) for fluorochrome visualisation. However, small blood spots were considered safer for the animals which have the potential for significant capture morbidity and mortality (Freeguard & Ritcher 2009).Typically, conservation genetics studies use larger numbers of loci to understand populations e.g. Zenger & Cooper (2001) use 9 markers in a population study of the eastern grey kangaroo (*Macropus giganteus*). However, studies have been undertaken with small numbers of markers previously (Harris *et al.* 1991, Hughes & Queller 1993) and our approach was able to yield significant information and suggest improvements to management. Nevertheless more genetic markers would yield a more complete understanding of genetic diversity, particularly if those loci were functional.

The loci used in this study were neutral loci which may not be ideal for understanding captive populations and specifically adaptation to captive conditions. This is discussed both in Chapter 1 and Chapter 4 in detail. Although routinely used in conservation genetics programs, questions around the effectiveness of neutral markers in conservation situations, particularly in understanding adaptation (as discussed in detail in the introduction) prompted us to attempt RAD-TAG genotyping (see chapter 5) using mala DNA. Unfortunately the low amounts of DNA available meant that quality control processes failed on these samples. Subsequent chapters will examine alternative methodologies of incorporating this kind of data into future studies.

3 Examining the factors that influence genetic diversity at MHC

3.0 Introduction

Conservation genetics is a relatively recent but important field informing decision making with the larger framework of conservation. Influenced by the seminal work of Frankel (e.g. Frankel 1970, Frankel 1974) and Soule (e.g. Frankel & Soule 1981, Soule 1986, Soule 1991) and continued by Frankham *et al.* (e.g. Frankham 1995, Frankham 1980, Frankham *et al.* 2010, Frankham *et al.* 2014), the importance of genetic diversity to population viability has been demonstrated (Reed & Frankham 2003). Genetic diversity is a buffer against extinctions in the short term and the raw material for evolution in the long term (Laikre *et al.* 2009). Therefore, species depauperate of genetic diversity will have increased risk of extinction (Speilman *et al.* 2004, Frankham 2005). This has been demonstrated empirically with *Drosophila* (Bijlsma *et al.* 2000) and is suggested in many wild populations (Speilman *et al.* 2004). Furthermore, lack of genetic diversity is also associated with inbreeding depression, a loss of fitness caused by the mating of closely related individuals which exposes rare recessive deleterious mutations. Inbreeding depression has also been demonstrated both in lab (Bijlsma *et al.* 2001) and wild populations (e.g. Slate *et al.* 2000, reviewed in Crnokrak & Roff 1999). Both of these factors are important considerations as populations of concern are typically small. Therefore, genetic diversity is routinely, though not yet ubiquitously, considered as part of conservation management of species (Frankham *et al.* 2009, Rivers *et al.* 2014).

Until recently, amplifying an entire genome has been impossible and use of genomics for conservation purposes is still largely impractical, therefore conservation genetics has made use of genetic markers. As detailed in Chapter 1, the first group of genetic markers widely used in conservation genetics were allozymes (Crochet 2000). However, the resolution of these markers is low and therefore the next generation of genetic markers were based in the more variable non-expressed areas of the genome (henceforth termed neutral markers) with higher variability and thus resolution (e.g. Bowcock *et al.* 1994). Though there are a number of neutral markers (see chapter 1 for discussion) the most commonly used over the last decades are Microsatellites.

Changes in molecular technology have created a plethora of new genetic markers and conservation genetics must decide which, if any, should be used in conjunction with, or in place of, microsatellites. Jarne & Lagoda (1996) make an argument for why microsateliites are better genetic markers than those available previously. However, more recently a number of alternative methodologies have been developed including Amplified Fragment Length Polymorphisms (AFLP), Single Strand Conformation Polymorphism (SSCP) and increasingly Single Nucleotide Polymorphism (SNP) (Anne 2006). Deciding which markers to use is ultimately dependent on the specific conservation questions being answered (Sunnucks 2000, Anne 2006). However, a more general question persists: should conservation genetics aim to investigate functional genetic markers? Work has been done at a population level on multiple genes including Heat Shock Proteins (e.g. Krebs & Feder 1997) but the most commonly suggested functional gene complexes for study in conservation genetics is MHC (Edwards & Potts 1996).

The Major Histocompatibility Complex is an excellent candidate genetic marker as it has both high variability and a function important to conservation. The background of this genetic marker is discussed in Chapter 1. Briefly, MHC is a gene complex within the vertebrate genome that plays a role in disease resistance (Snell 1981). Proteins created by MHC genes mediate the recognition of antigens by the adaptive immune system. Therefore, greater genetic diversity at MHC enables greater pathogen resistance. This relationship between MHC diversity and fitness has been demonstrated in both human (e.g. Lie *et al.* 2008) and animal (e.g. Aguilar *et al.* 2004) studies. As a gene complex MHC is able to generate large amounts of genetic diversity by intragenic recombination (Richman *et al.* 2003, Hosomichi *et al.* 2008) as well as nucleotide mutations. This genetic diversity is maintained by multiple types of pathogen mediated selection (PMS) (Spurgin & Richardson 2010). Furthermore, MHC diversity is maintained by mate choice mechanisms in vertebrates that select for MHC dissimilar mates (Wedekind *et al.* 1995) (discussed in detail in Chapter 3.2.1) or optimal MHC allele numbers (e.g. Bonneaud *et al.* 2006, Eizaguirre *et al.* 2009). It is

therefore unsurprising that MHC is the most variable functional gene yet discovered (Sommer 2005). These characteristics give MHC the potential to be a powerful marker for conservation genetics.

In order to use MHC as a genetic marker it is important to understand the factors that affect its genetic diversity. In general, neutral genetic markers are used because they are widely dispersed throughout eukaryote genomes and are highly variable (Jarne & Lagoda 1996). However, several studies suggest that the distribution of genetic diversity at different markers is not always equal (e.g. Aguilar *et al.* 2004, Dionne *et al.* 2007, Hambuch & Lacey 2002). In fact, diversity within a single type of genetic marker may not be equal, for example microsatellite diversity may be different where markers were cross amplified across species (Barbara *et al.* 2007). Furthermore, for MHC alternative ecological and life-history traits may impact genetic diversity, and may vary across species as may differences in the genomic architecture of MHC (Pers comm Grueber). Therefore, an alternative approach to population genetics focuses on functional markers partly because they are under selection (Edwards & Potts 1996, Sommer 2005).

This chapter attempts to investigate some of the factors that influence genetic variation at MHC using a meta-analysis to integrate findings of genetic variation across multiple species (as per the definition of Johnson *et al.* 1981). This analysis included 70 species of which the majority (48) were mammals (full meta analysis data available in appendix 3.1 and 3.2). Section 3.1 investigates the effect of neutral genetic diversity, specifically microsatellite diversity, on genetic diversity at MHC. Section 3.2 examines the effect of within population processes, specifically monogamy and multiple mating on MHC diversity. Finally, Section 3.3 looks at population size itself, as well as group living, as a selective pressure that may affect genetic diversity at MHC.

3.1 Meta-analysis: Can neutral genetic diversity or IUCN status predict genetic diversity at MHC

3.1.1 Introduction

Genetic diversity at microsatellite markers is believed to mirror the abundance and distribution of genetic diversity genome wide in a population. Microsatellite markers are comprised of short tandem repeats (STR's) of DNA sequence e.g. ccgccgccg which have a high error rate when being copied by DNA polymerase (Jarne & Lagoda 1996, Bennett 2000). Further, they are typically located in non-transcribed regions of the genome, where selection has been considered less effective (Bennet 2000). Therefore, the genetic variation created by mutation is higher than other sequences. Microsatellites have been widely adopted in conservation genetics because they allow for a comparatively cheap and rapid method of creating a set of highly informative genetic markers for a specific species (Selkoe & Toonen 2006). These markers have been extensively used to understand mating systems (e.g. Jones & Avise 2001, Fessehaye *et al.* 2006), gene flow within species, (e.g. Racey *et al.* 2007), to identify cryptic species (e.g. Hoekzema & Sidlauskas 2014) and to investigate outcrossing (e.g. Marshall and Spalton 2000). Furthermore, the heterozygosity at microsatellite markers has been linked to fitness in toads (Rowe *et al.* 2004) and marmots (Da Silva *et al.* 2006). However, these markers may be poorly suited for some conservation questions.

Changes in frequencies of microsatellite alleles, and other neutral variation, are controlled by genetic drift rather than selection. Genetic drift is the process whereby fertilisation resamples all alleles in the population by chance every generation (Kimura 1968). This causes fluctuation in allele frequencies over time without selection and is responsible for the fixation or extinction of many alleles mainly in small populations (Kimura & Ohta 1969). As microsatellite diversity is usually selectively neutral, genetic drift is the major contributor to changes in the frequency of microsatellite alleles. However, while functional genes are affected by genetic drift, they are also

affected by selection. In order to produce adaptation in a species the effect of selection must prevail. Where a species spans different environmental conditions this adaptation may produce genetic differentiation at functional alleles.

Local adaptation is an important consideration in conservation genetics. Local adaptation has been widely demonstrated (e.g. Taylor 1991, SäNEN *et al.* 2003, Hoekstra *et al.* 2005) and is a key step in the process of speciation. It can be rapid (Cousyn *et al.* 2001, Partecke & Gwinner 2007) and affect development rates (e.g. Laugen *et al.* 2003), environmental tolerances (e.g. SäNEN *et al.* 2003) and reproductive cycles (e.g. Phillimore *et al.* 2010). This is concerning as crossing distinct and genetically differentiated populations could result in impaired fitness in the offspring of those crosses (Kawecki & Ebert 2004, Frankham *et al.* 2011). This loss of fitness, termed outbreeding depression, has been demonstrated empirically (Cooke & Philipp 2005). The level to which local adaptation can occur is dependent on several factors including gene flow.

Local adaptation, the change in populations to better exploit local environmental conditions is, in theory and practice, prevented and in fact reversed by gene flowgene flowgene flow between populations which is detectable using neutral markers. Speith (1974) predicted a single migrant per generation should be enough to prevent population divergence and thus local adaptation, a number that Mills & Allendorf (1996) revised to between 1 and 10. This has been supported by reviews which found gene flow constrained local adaptation (e.g. Garant *et al.* 2007). Furthermore, gene flow has been detected using neutral markers (e.g. Repaci *et al.* 2007 among numerous others) which should mean that the potential for, if not the evidence of, local adaptation is detectable by neutral markers. In brief, if neutral makers show gene flowgene flowgene flow and gene flowgene flowgene flow precludes local adaptation, then neutral markers will be able to demonstrate where local adaptation cannot occur, even if they cannot confirm it has occurred. For conservation this is problematic as the confirmation of local adaptation is more useful for outcrossing than ruling it out. and in addition disruptive selection may produce local adaptation even with gene flow.

Disruptive selection may also produce local adaptation, and thus potentially can allow for outbreeding depression, without a detectable change in genetic diversity at neutral markers. Disruptive selection favours both extremes of a trait but not the intermediate and can promote genetic polymorphism or genetic isolation (Mather 1955). Additionally, disruptive selection can create local adaptation even when gene flow occurs (e.g. McKay *et al.* 2001, Nielsen *et al.* 2009). However, this may not be reflected accurately in analyses of neutral diversity (Cousyn *et al.* 2001). However, disruptive selection may be uncommon enough in species of conservation concern that neutral markers are still an effective surrogate. Alternatively, there may be other scenarios where neutral and functional genetic variation does not correspond, specifically in small populations, which are crucial in conservation genetics.

The balance between selection and genetic drift is altered in small populations, so that selection may not be effective in small populations because selection may be overwhelmed by genetic drift in small populations according to .

S < 1/2Ne (Kimura 1983)

Equation 3:

In smaller populations therefore, the frequencies of functional alleles are increasingly determined by genetic drift like neutral loci (Kimura 1983). This ultimately results in the fixation of deleterious alleles which reduce fitness. The accumulation of these alleles across numerous loci results in mutational meltdown, where mutational load increases until individuals cannot survive in wild conditions, termed Mullers Ratchet (Felsenstein 1974). This would suggest that neutral markers are more effective surrogates in small populations compared to large populations. However, small populations also experience balancing selection, which may act on one or more loci, including heterozygote advantage, rare allele advantage and selection varying in direction, all of which act to preserve genetic diversity (Frankham *et al.* 2010). These types of selection act to slow the loss of

functional diversity which may explain discrepancies between levels of diversity at functional and neutral loci in populations of conservation concern.

Empirical work has shown differences between the amount and distribution of functional and neutral genetic diversity. The San Nicolas Island fox (*Urocyon littoralis dickeyi*) has retained genetic diversity at functional markers despite being genetically depauperate at microsatellite markers (Aguilar *et al.* 2004). On the other hand, a meta-analysis by Sutton *et al.* (2011a) found that bottlenecks reduce MHC diversity by approximately 15% more than they reduce neutral diversity in a species. Furthermore, some disparity between levels of genetic diversity at different neutral and functional loci has been discovered in the Northern Elephant Seal (*Mirounga angustirostris*) (Weber *et al.* 2004). These differences have also been found in sheep (Boyce *et al.* 1997) and wolves (Niskanen *et al.* 2014). Whilst a few examples in no way create a rule, they may suggest that the link between neutral and functional diversity is worth investigating further. This requires choosing a functional marker to compare with widely used microsatellites.

Chapter 3.0 and Chapter 1 outline the benefits of genetic markers based on MHC for conservation. Importantly, in studies where both MHC based and neutral genetic markers have been genotyped the resulting pictures can be inconsistent which may reveal gaps in our understanding of species history. Studies of the Galapagos hawk (*Buteo galapagoensis*) (Bollmer *et al.* 2011) revealed the same genetic bottleneck using both microsatellites and MHC markers. Further, studies of Tuatara (Miller *et al.* 2010) showed that genetic bottlenecks, rather than selection, was the main determinant of MHC diversity across a landscape. This is of particular interest in conservation where many, if not most, populations have recently experienced a genetic bottleneck. On the other hand, during a five year period within Charbonnel & Perberton's (2005) study of Soay sheep, MHC and neutral markers suggested different levels of genetic differentiation between populations, significantly this was related to parasite activity. Furthermore, Oliver *et al.* (2009) study of the Water vole (*Arvicola amphibious)* concludes that it is useful to contrast neutral and MHC variation as selective forces that affect the way MHC and functional genes vary across a landscape. Therefore, it

may be prudent for conservation genetics to add analysis of MHC to other genetic tools when examining a population.

The importance of incorporating information on MHC diversity into conservation programs has seen significant but unresolved debate (see Hughes 1991, Vrijenhoek & Leberg 1991).However, there have been relatively few cases of MHC data being used in conservation management (Sommer 2005). Nevertheless, Frankham *et al.* (2010) suggested that MHC management should be distinct from other genetic management in species of conservation concern. A recent review by Ujvari & Belov (2011) pointed to MHC genotyping as a measure of the immunological fitness of a population and a method of increasing breeding program success. This is possible as a number of molecular approaches enable analyses of MHC to facilitate this management (e.g. Babik *et al.* 2009). However, It may be difficult to justify directing resources to genotyping MHC, except in cases such as the Tasmanian devil where it is necessary to solve a very specific conservation problem, (Siddle *et al* 2010) if the amount and distribution of neutral diversity acts as an accurate surrogate. Simply put, if the inconsistency between MHC and neutral genetic markers is simply an exception, then MHC only needs to be genotyped if it contributes to solving a particular conservation problem. However, if neutral genetic markers can not be used to accurately estimate the levels of MHC diversity in a population, then it may be necessary to genotype MHC and perhaps other functional markers independently to conserve the functional variation which is important to species survival.

This meta-analysis aims to determine if genetic diversity at MHC can be predicted by genetic diversity at microsatellites. Although reviews of MHC studies exist, (e.g. Ujvari & Belov 2011) and meta-analyses have been performed to examine the relationship between neutral markers and fitness, (Reed & Frankham 2003) we know of no studies that have reviewed the relationship between these two markers.

A strong and straightforward relationship between microsatellite and MHC diversity across species would suggest that genotyping MHC is not necessary where information on neutral markers exists.

Further, it suggests that when diversity at neutral markers is low MHC diversity will be low and therefore local adaptation at MHC should be rare. Conversely, if the relationship is not present, or is weak, it adds further weight to the argument for incorporating MHC genotyping into common conservation practice.

In addition, the meta-analysis will compare genetic diversity and IUCN status. The IUCN, founded in 1948, was the world's first global conservation organisation (IUCN website). It has brought together the skills of 11,000 researchers to create the IUCN Red List which for 50 years has classified species based on their risk of extinction. While the list is widely accepted and used worldwide, it current lacks an explicit consideration of genetic diversity (Rivers *et al.* 2014). Although geneticists are now united in understanding the importance of genetic diversity to population viability (Frankham *et al.* 2010), the method or methods by which genetic diversity should be measured for consideration by the redlist are unclear. In order to include genetic diversity among the IUCN categories either a single measure must be decided upon, or alternative measures must be compatible.

Comparing IUCN status to genetic diversity allows for a test of the methodology. Low levels of neutral genetic variation have been found in a large number of species of conservation concern e.g. Chinese Alligator (*Alligator sinensis*) (Wang *et al.* 2006), Javan rhinoceros (*Rhinoceros sondaicus*) (Fernando *et al.* 2006) and Galapagos penguin (*Spheniscus mendiculus*) (Nims *et al.* 2008). Although there is no single genetic marker used by the IUCN the use of genetic markers has also led to recommendations for changed IUCN status (IUCN 2015). Therefore, it is predicted that as threat levels increase genetic diversity at MHC should decrease.However, this is a pilot meta-analysis designed to point to interesting research directions and can not rule out other possible explanations for this effect.

3.1.2 Methods

Data gathering

A meta-analysis was conducted using Google scholar and Web of Science to identify species where MHC genotyping had been carried out. From these the method of genotyping, class of MHC genotyped, sample size, number of MHC variants and MHC heterozygosity was recorded. For each of these species a literature search was carried out using the same databases looking for microsatellite genotyping. Where microsallite genotyping was found both observed and expected heterozygosity was recorded and then IUCN red list status was found (IUCN 2015) (full method available in appendix 3.3).

This is a limited comparison made only because the available data did not facilitate a more powerful analysis. The number of MHC loci vary between species which may bias the number of MHC variants whereas He is an average across a known number of microsatellite loci and thus not biased. Nevertheless, the results of this comparison are useful in determining whether species with higher neutral diversity also have higher MHC diversity.

Statistical approach

The literature was surprisingly sparse. Although many studies have been done on MHC and there has also been extensive work on genetic diversity in threatened species, there are few species for which both genetic diversity at neutral markers and genetic diversity at MHC are available at a population level. Nevertheless, the literature review revealed two main methods of genotyping MHC, sequencing based approaches and SSCP. SSCP has been estimated to detect 95% (Vidal-Puig & Moller 1994) to 97% (Hayashi 1991) of mutations. Comparatively, the sensitivity of cloning and sequencing approaches depends on the level of replication and technology used with rates of 85% - 95% accuracy at detecting mutations reported (Tsiatis *et al.* 2010). As these produce largely similar data and sample size was small, a Mann Whitney U-test was performed. (This and all other analyses were performed using SPSS 21 (IBM Corp. 2012)) to check for a difference in allele number generated by these two different genotyping approaches.. As no difference was found the datasets for cloned and sequenced MHC genotyping and MHC genotyping by SSCP were combined.

To facilitate further analysis Shapiro-Wilk tests for normality were carried out on the continuous variables about which data was collected: Number of MHC variants, Microsatellite Observed Heterozygosity and Microsatellite Expected Heterozygosity.

A Spearman rank-order correlation was used to investigate the relationship between neutral genetic diversity measured using microsatellite markers and functional genetic diversity at MHC. In this analysis expected heterozygosity (H_e) as determined using microsatellite markers was used as the measure of neutral genetic diversity. This measure was chosen as it was available more often than other measures in the data set. Additionally, heterozygosity can estimated accurately from a small number of samples (Nei 1978) and so it is often used in calculations to determine the amount and division of genetic diversity in a population (Nei 1977). Genetic diversity at MHC was measured using the number of alleles discovered as measures such as He were typically unavailable for MHC in the dataset. The lack of H_e for MHC is understandable as the number of MHC loci in non-model vertebrates is highly variable (see Chapter 1) and often unknown. In these tests, as per all tests in this chapter, a null hypothesis of no effect was used as is commonly practice.

As the data set was relatively small and contained data from combined molecular methods, we included an additional analysis as a proof of concept. Species were categorised by IUCN category into two groups, threatened (including critically endangered, endangered, threatened and vulnerable) and not- threatened (species of least concern). This categorisation was used as there were few threatened species for which genetic data were available. A Mann-Whitney U test was used to compare the number of MHC variants in each group.

3.1.3 Results

There was a significant positive correlation between genetic diversity at MHC and genetic diversity measured at microsatellite loci.

The data results of the Sharipo-Wilk test (shown below) indicated the data was not normal and thus that non-parametric tests should be used in further analysis. This was the case even if data was transformed using a square root transformation.

Table 3.1: The results of Shapiro-Wilk test for normality, df = degrees of freedom, Sig = p-value, ** denotes a significant result

	Shapiro-Wilk		
	Statistic	df	Sig.
Microsatellite H_e	.909	30	.014**
Microsatellite H_o	.916	30	.021**
MHC_Number of Alleles	.821	30	.000**

There was a significant relationship between both: Number of MHC variants and Microsatellite He and Number of MHC variants and Microsatellite H_o (observed heterozygosity). As described in the table below, the relationship between He and Number of MHC variants has both a higher Rho value (correlation coefficient) and a lower p-value.

Table 3.2: The results of Spearman rank order correlations, between number of MHC alleles and Microsatellite diversity ** denotes a significant correlation

Correlations

			Microsatellite He	Microsatellite Ho
Spearman's rho	MHC Number Of Alleles	Correlation Coefficient	.535**	.494**
		Sig. (2-tailed)	.001	.002
		N	38	37

Although small sample size (n=38 for the correlation between number of MHC variants and H_e, n=37 for the correlation between the number of MHC variants and H_o) precluded an analysis comparing the mean number of MHC variants in each IUCN category, a significant difference was found

between threatened species and species of least concern. Species of least concern had a higher mean number of MHC variants (see table 3) than species with an IUCN status of near threatened or worse. A Mann-Whitney U-test determined that this difference was significant (p=0.015).

Table 3.3: The summary statistics for the Number of MHC variants compared between IUCN status (where Least Concern = Least Concern and Threatened or Worse = Near Threatened, Threatened, Vulnerable, Endangered, Critically Endangered)

IUCN Status	95% Confidence Interval for Mean			
	Mean	Std. Error	Lower Bound	Upper Bound
Least Concern	20.32	4.25	11.54	29.10
Threatened or Worse	11.63	2.65	6.25	17.01

3.1.4 Discussion

There is a significant relationship between MHC and microsatellite diversity but not one that is simple and predictive. Both the relationship between IUCN and MHC diversity and the relationship between Microsatellite and MHC diversity are in the predicted direction but the scarcity of genotyping data restricts the ability to explore the relationship in more detail. This is concerning as MHC diversity is important in species of conservation concern.

Species with a threatened or worse IUCN status had significantly fewer MHC variants but the small number of samples offered limited resolution. As predicted MHC diversity was lower in species that the IUCN classified as Near Threatened or worse. Reed and Frankham (2003) showed a link between genetic diversity of a population and fitness in *Drosophila*, further Frankham (1996) correlated genetic variation at allozymes to population size. Therefore, it is unsurprising thatthreatened populationswhich are small and typically have lower fitness also have lower amounts of genetic diversity (e.g. Frankham *et al.* 2005). The well documented endangered populations with relatively high genetic diversity e.g. Greater One-Horned Rhinoceros (*Rhinoceros unicornis*) (Dinerstein & McCracken 1990), Copper Redhorse (*Moxostoma hubbsi*) (Lippe *et al.* 2006) would seem to be exceptions rather than the rule. However, there were not enough samples in each IUCN class to

investigate the effect of class in detail. This is illustrative of the difficulties associated with genotyping MHC diversity in non-model species discussed elsewhere (briefly in Chapter 1 and also in Chapter 4). This lack means that it is not possible to determine if IUCN classes reflect significant genetic differences at a finer scale. Whilst outside the scope of the main thrust of this research, it may be significant when determining whether to revise IUCN classification (e.g. Mace & Lande 1990). In order to do so better methods of genotyping MHC would be useful.

There is a significant but not straightforward positive correlation between genetic diversity at microsatellites and MHC. The significant relationship between genetic diversity at MHC and microsatellites is consistent with expectations. Heterozygosity at microsatellites has been found to correlate with fitness in individual vertebrates, e.g. Big Horn Sheep (Luikart *et al.* 2008) and Red Deer (Slate *et al.* 2000) and in a meta-analysis of vertebrate species (Chapman *et al.* 2009). MHC variability has also been correlated to fitness in the wild in species including the striped mouse (Froeschke & Sommer 2005) and three spined stickleback (Wegner *et al.* 2003). The positive correlation shows that in vertebrates as neutral genetic diversity increases functional diversity at MHC should be expected to increase. This finding in itself is valuable, for example when choosing to outcross to one of two populations it would be more reasonable to expect that the one with higher microsatellite diversity also had higher MHC diversity which is valuable in small populations. However, as the data were not normal, a regression could not be performed, and instead the rank of data was compared. This means that there is not a straightforward way to predict MHC diversity from microsatellite diversity. Further, the moderate Rho value suggests that there are other important factors affecting MHC diversity.

The inability to accurately predict MHC variability based on microsatellite data may be an artefact of the approach used. This analysis compared microsatellite heterozygosity to number of MHC variants which is not a straight forward comparison but was necessitated by the data available. Furthermore, it was often difficult to determine the number of MHC loci which were genotyped in a species and

variation in this may act to reduce the strength of the correlation observed. A standard method of MHC genotyping would redress some of these concerns.

In combination, these two results suggest interesting possibilities for neutral and functional genetic markers in conservation. The relationship between genetic diversity at genetic markers and population size was shown by Soule (1976) and later by Frankham (1996). However, the subsequent progression in development of genetic markers means that these correlations may need to be revisited. For example, although Reed & Frankham (2003) found in a meta-analysis that 19% of variation in fitness was explained by genetic diversity, a later study (Chapman *et al.* 2009) that included more microsatellite data found only a small effect though many of these studies examine variation across individuals rather than species. This may be because microsatellite diversity is not well correlated to the functional diversity, previously measured in allozyme studies, which does have a strong effect on fitness. Significantly, Chapman *et al.* (2009) notes that the majority of analyses that link heterozygosity measures to fitness do not focus on heterozygosity measures obtained using microsatellite data despite those being the most common marker set currently in use. This is significant because of the moderate Rho value correlating MHC and microsatellite markers. Although both markers are significantly related, MHC may be better correlated with fitness. The lower genetic diversity with near threatened or worse IUCN status is consistent with, but provides no evidence for, this possibility. However, the lack of data gives an insight into the challenge that this presents to conservation. Practical considerations including cost, ease of use and speed of genotyping have made microsatellite marker use common in population genetics (Chistiakov *et al.* 2006, Chapman *et al.* 2009). In contrast MHC is complex, unwieldy and difficult to genotype. Therefore, as well as finding the expected relationships between MHC diversity and neutral genetic variation, and MHC diversity and IUCN status, this meta-analysis suggests good reasons why more MHC data are needed. , (an attempting at producing such a method is detailed in chapter 4)

3.2 The Effect of Mating System of Number on MHC variants

3.2.1 Introduction

Falling in love is easy, staying in love is hard. Animal mating strategies are many and highly varied, ranging from scramble competition (Wells 1977) to lifelong monogamy (e.g. Stow & Sunnucks 2004). The powerful effects of these systems were first described by Darwin who proposed that traits unfavourable to fitness like bright colours and costly ornaments were maintained by the effects of sexual selection (Darwin 1871). However, there is still significant work to be done in understanding the role of mating strategies in shaping and being shaped by vertebrate evolution. Monogamy, a mating strategy based on long term commitment to a single partner, is especially enigmatic.

As with many life-history traits, numerous explanations have arisen for the evolutionary origins of monogamy. Monogamy is highly variable across taxa, occurring in only 3-5% of 4000 mammal species (Schuiling 2003) and is described as intrinsically unstable (Schuiling 2003). The theories for its evolution are many and varied, and include prevention of infanticide (reviewed by Palombit 1999, Opie *et al.* 2013), low mate availability (Emlen & Oring 1977), provision of offspring (Whiteman & Côté 2004) and territory defence (e.g. Piper *et al.* 1997, Whiteman & Côté 2004).The debate into the cause of monogamy continues unabated; for example Opie *et al.* (2013) concluded social monogamy in primates is driven by infanticide while , Lukas & Clutton-Brock (2013) conclude that infanticide cannot be the cause of monogamy in primates and point to the benefits of mate guarding. Trying to discover a single selective pressure directing the evolution of monogamy in vertebrates may not be possible.

Despite the continued debate it should be recognised that these competing theories are not necessarily mutually exclusive. Phylogenetic analysis suggests monogamy has arisen independently multiple times (Lukas & Clutton-Brock (2013), suggests 61 origins) and thus convergent evolution is likely. Even among vertebrates, assigning a single driver for monogamy is an oversimplification.

Generally, the pressures of offspring provisioning (Whiteman & Côté 2004), territory acquisition and defence (Morley & Balshine 2002,Whiteman & Côté 2004) and benefits associated with mate guarding (Lukas & Clutton-Brock 2013) have been suggested as the major drivers of monogamy in different species.

At this juncture it is important to clarify that social monogamy is not indicative of genetic monogamy. The differentiation between social monogamy and genetic monogamy was first made by Black (1996) and is that in genetic monogamy all offspring are produced by only two individuals. On the other hand, social monogamy is defined as a close sociospatial relationship between two individuals (Reichard 2003). Even where social monogamy is common, genetic monogamy is rare (Griffith *et al.* 2002). This difference is best demonstrated in bird species where 90% are socially monogamous, and were presumed to be genetically monogamous (Lack 1968). However, molecular tools discovered widespread extra-pair paternity and reduced this number to 14% (reviewed in Griffith *et al.* 2002). This finding is expected, as species that display social monogamy benefit from shared offspring provision, and presumably avoidance of infanticide (Opie *et al.* 2013) but these benefits do not require genetic monogamy.

True genetic monogamy is inherently risky and seldom favoured. Genetically monogamy has fitness benefits, including a decrease in disease transmission, increased longevity and greater incentive for paternal care (Xia 1992, Martin & Hosken 2003). However, genetic monogamy reduces genetic diversity in offspring compared to multiple mating e.g. Huo *et al.* (2010). In short, multiple partners allows an individual to hedge their reproductive bets, whereas a single partner increases the stakes of reproductive betting (Phillipi & Seger 1989). Phillipi & Seger (1989) show that under many conditions the reduction of variation in reproductive output is advantageous even if comes at the cost of mean reproductive success. In short, multiple mating is a safer evolutionary strategy. Furthermore, population genetics predicts that unequal reproductive outputs, such as those monogamy can create, lead to faster loss of genetic diversity in populations (Frankham *et al.* 2005).

This is because monogamous species make a single large reproductive bet on the genetic and reproductive quality of their mate whereas multiple mating species make several smaller bets on the quality of their multiple mates. In monogamous systems a mate's genetic material may be highly beneficial, increasing the fitness of all offspring or highly deleterious, reducing the fitness of all offspring Therefore, the scarcity of monogamy is unsurprising, as the evolutionary driver of sexual reproduction as a system is its ability to produce large amounts of genetic variation (Weismann 1889), a process genetic monogamy limits. Unsuprisingly, many monogamous systems contain mechanisms to optimise reproductive choices .

Monogamy, like all other mating systems has costs associated with it, and mate choice based on genetic cues has the ability to reduce these costs . By assessing the genetic material of a potential mate it is possible to choose 'good genes' that may increase the likelihood that offspring will be successful (reviewed by Moller & Alatalo 1999, Kokko *et al.* 2002). Although 'good genes' selection is by no means restricted to monogamous species, it may mean that their single reproductive event is better informed. Additionally, genetic cues may be used to find genetically dissimilar mates and thus increase genetic variability of offspring (discussed recently in Muehlenbein 2010). This 'compatible gene" theory (reviewed by Mays *et al.* 2008) increases the variability among offspring compared to that of a chance mating and has been demonstrated both in the lab (e.g. Parrott *et al.* 2007, Huchard *et al.* 2013) and field based studies (e.g. Hucahard *et al.* 2013). In genetically monogamous species it provides a mechanism of retaining genetic diversity and producing offspring with high levels of genetic diversity even when mating with only a single individual. It is for this reason that mate choice for dissimilar genes has been investigated extensively in monogamous vertebrates.

The genetic cue most often investigated in mate choice studies is the Major Histocompatibility Complex (MHC) (e.g. Wedekind *et al.* 1995, Reusch *et al.* 2001, Cutrera *et al.* 2012, Kamiya *et al.* 2014). Mate choice based on MHC has been extensively researched in vertebrates and is an accepted mate choice driver in vertebrates (review in Sommer 2005 and Zielger *et al.* 2005).

Although a number of mechanisms have been researched including secondary sexual traits such as bird spur-associated MHC variants (Von Schantz *et al.* 1996), MHC allele number optimisation (Reusch *et al.* 2001) and MHC similarity (Bichet *et al.* 2014), the most commonly suggested mechanism for MHC based mate choice is dissimilar MHC preference (Jordan & Bruford 1998, Zielger *et al.* 2005). This was described most famously by Wedekind et al (1995) who asked women to rate the attractiveness of the odours of men with both similar and dissimilar MHC. Women rated the t-shirts worn by MHC dissimilar men as more attractive, which is consistent with women choosing partners who to maximize genetic dissimilarity and disease resistance in offspring.

As MHC is involved in pathogen recognition, this preference for genetically dissimilar mates is beneficial in that it produces offspring that are able to resist a larger suite of diseases (Grimholt *et al.* 2003, Westerdhal 2007, Consuegra & de Leaniz 2008) as well as acting to prevent inbreeding (Potts & Wakeland 1993, Pusey & Wolf 1996) and creates high genetic variability among offspring. MHC based mate choice can be expected in monogamous vertebrates as it allows individuals to make better reproductive choices in systems where reproductive choice is crucial. Furthermore, a recent meta-analysis by Kamiya et al (2014) suggests there is a general trend for female preference for dissimilar MHC types in vertebrates. However, whether this expectation is a general rule in monogamous vertebrates is not known.

Although mate choice based on MHC has been found in numerous species it may not be a general strategy associated with monogamy. While a general preference for dissimilar mates is accepted (Jordan & Bruford 1998) and strong MHC based mate choice has been found in several species (e.g. Freeman-Gallant *et al.* 2003, Olsson *et al.* 2003,), in other species MHC cannot be shown to influence mate choice (e.g. Sommer 2005). Additionally, the well-known publication bias for positive results (Dickerson 1990) may mean that findings of no relationship are under reported. This is further complicated by the possibility of miss-assigning mate choice mechanisms to MHC when they may actually be determined by other parts of the genome (Sherborne *et al.* 2007). However, MHC

diversity is often incorporated into conservation programs based both on its importance to resisting disease and its importance as a mate choice (Sommer 2005).

Determining the relationship between MHC-based mate-choice-mechanism and mating system will help us understand how species reduce the costs associated with monogamy. As explained in figure 3.1, monogamy with less reproductive bet hedging should result in less genetic diversity than multiple mating because of a smaller pool of unqiue alleles to combine with an indiciduals DNA. Because of this, mate choice mechanisms, including those based on MHC, may become more important this is because with fewer individual reproductive bets it is more important to win individual bets. Further, this may, depending on the strength of selection, mean that multiply mating species have lower MHC diversity than monogamous species. To resolve this investigation across species is required as Sherborne *et al.* (2007) raises the possibility that positive results found for MHC based mate choice may be artefacts of the experimental design and not real. This will complement the work done on individual species (previously described) and provide further evidence for the theory. It is hoped that by understanding the ways that vertebrates manage the risks associated with monogamy we can gain insight into its evolution in vertebrates.

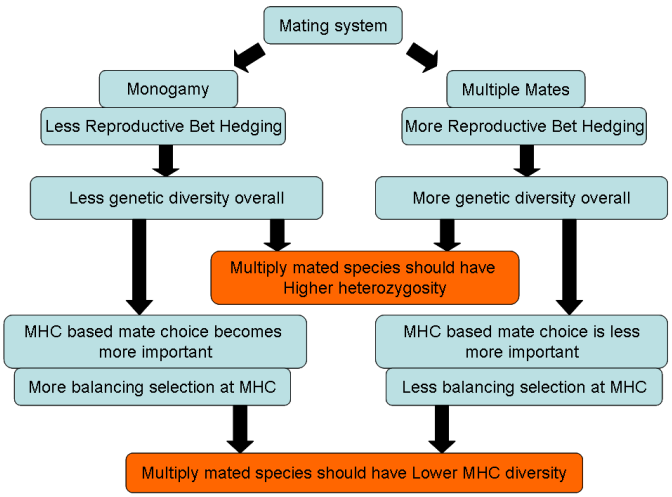


Figure 3.1 . Current predictions on the effect of mating system on genetic diversity overall and at MHC

3.2.2 Methods

Data Collection

The data collection strategy for this chapter is detailed in Section 3.1. For this analysis it was also important to classify species as monogamous or multiply mating. As there are many definitions of monogamy based both on social and genetic factors (Johnson & Ryder 1987) this was a non-trivial definition. For the purposes of this analysis monogamous species are defined as genetically monogamous species that typically produce offspring with a single mate for a season or more. Numbers in each category are shown in Table 3.4

As in previous chapters genetic diversity at MHC was measured by counting the number of MHC variants found for a species using either SSCP or cloning based approaches (explained in Section 3.1, full method avaialbe in appendix 3.4). In contrast neutral genetic diversity was measured by using expected heterozygosity (He) of microsatellites markers.

Analysis

Genetic diversity was compared using Mann-Whitney U-tests (IBM SPSS 21).

3.2.3 Results

There was no significant difference found for either Heterozygosity or MHC diversity between monogamous and multiply mating species. Table 1 shows that there is a very slight decrease in both MHC and microsatellite alleles in monogamous species but the results of both Mann-Whitney U-tests were non-significant: Mann-Whitney U-Test for differences in the number of MHC allleles $p = 0.373$, Mann-Whitney U-test differences in the microsatellite Heterozygosity $p=0.238$.

Table 3.4 Summary of results where N (MHC) is the number of individuals in each class for which there was MHC data, Avg. MHC is the mean number of MHC variants, N (µsat) is the subset of species with MHC data for which microsatellite data was also available and Avg. He is the average predicted heterozygosity at microsatellite loci.

	N (MHC)	Avg. MHC	St. Error	N (µsat)	Avg. He	St. Error
Monogamous	18	14.17	4.32	11	0.49	0.09
Multiply Mating	28	15.25	2.8	16	0.64	0.03

3.2.4 Discussion

This analysis found no significant difference in either microsatellite heterozygosity or genetic diversity at MHC in monogamous compared to multiply mating species. However, a weak trend was found for less genetic diversity both at Microsatellite and MHC loci in monogamous species. Whilst the analysis is constrained by the limitations of the study design and data availability, which may in itself produce artefacts, it does provide an interesting basis for future work.

The limitations of this study may have contributed to the lack of significant results. This study is hampered by small sample size (n=46) for reasons discussed in chapter 3.1. Furthermore, as MHC based mate choice is typically (Wedekind *et al.* 1995, Wedekind & Penn 2000), but not always (Reusch *et al.* 2001), a preference for genetically dissimilar mates, the relationship between mating system and MHC heterozygosity may have been more meaningful. Again, the available data did not facilitate this analysis. While these factors limit confidence in the results they provide excellent support for the need for development of a general MHC genotyping methodology (as developed in Chapter 4). Furthermore, a molecular experiment, as opposed to a meta-analysis, would not require the assumption that populations studied within a species for microsatellite diversity and MHC diversity were equivalent. This would allow for greater confidence in the results. However, preliminary studies are designed to examine the potential for more detailed studies and the lack of a significant result is interesting enough to warrant further investigation.

The approach used by Kamiya *et al* (2014) may have provided solutions to these problems. This study used a meta-analysis and meta-regression techniques to find female choice for dissimilarity across vertebrates only when multiple loci were used. To do this they converted measures in multiple studies to correlation coefficients. Whilst this carries the inheriant risk of violating the assumptions of correlations coefficients, particularily when it is carried out without referencing the original data, it enables more complex and powerful analysis that we could perform. However, without this methodological approach we did not find a significant different in MHC diversity.

The lack of a significant difference in genetic diversity at MHC provides further support for the work of Sherborne *et al.* (2007). Sherborne suggests that findings of MHC mate choice in mice are actually a result of mate choice based on a different gene product, MUP, and MHC dissimilarity mirrored MUP dissimilarily because of the way mice were bred. This is significant as the work on MHC based mate choice in mice was seminal (Yamazaki *et al.* 1976) and underpins later work on MHC based mate choice (Wedekind *et al.* 1995). If the apparent MHC based mate choice in mice is indeed an artefact of experimental design, then the lack of greater MHC diversity across monogamous vertebrates makes sense. If MHC is not driving mate choice in monogamous species, then rare MHC variants are not likely to confer a mate choice benefit and are more likely to be lost. Alternatively, MHC based mate choice may be operating but not in the way described by the seminal literature (Yamazaki *et al.* 1976, Wedekind *et al.* 1995) and more recently in grey mouse lemurs (*Microcebus murinus*) (Huchard *et al* 2013) and European badgers (*Meles meles*) (Sin *et al* 2015). Reusch *et al.* (2001) describes MHC based mate choice in three spined sticklebacks (Gasterosteus aculeatus) not as choice for dissimilarity but rather as counting for optimum allele number. If this mechanism is common in monogamous species it would not maximise genetic diversity at MHC but rather select for a stable amount of genetic diversity as too many new alleles would increase the number of alleles past the optimum. Indeed, the selection described could limit MHC diversity in such species in a way consistent with our (very weak) trend. Overall the results of this study are inconsistent with

the results of Jordan & Burford (1998) that mate choice for dissimilar mate choice is a general trend, there was not enough genetic data avialable to test Jordan & Burford's hypothesis.

Multiple mating is presumed to have an evolutionary advantage in that it produces more genetically variable offspring but this study found no difference in heterozygosity associated with mating system. This may be an issue of scale as the advantages of increased genetic diversity apply at an individual level and may not extend to a population. Alternatively, it may be a methodological issue as genetic diversity at microsatellites was measured using H_e and not a measure more sensitive to the loss of rare alleles such as allelic richness which may have shown a difference between monogamous and multiply mating species.. However, it is also possible that mate choice for high genetic diversity, be it signalled by MHC or otherwise, is powerful enough to offset any reduction in genetic diversity associated with monogamy. This is unlikely as it would mean that there is no disadvantage, in terms of genetic diversity, of monogamous systems and an advantage would presumably be required to offset the disadvantage created by not being able to hedge reproductive bets.

As monogamous species have not been shown to preferentially retain MHC it may be important to manage the MHC if it plays a role in mate choice for a species. Small populations with strong inbreeding avoidance mechanisms may lead to extinction via a lack of acceptable mates even when potential partners are available, termed the glass effect (Tainaka & Itoh 1996). The glass effect, proposed by Tainaka & Itoh (1996), describes a state where no mating occurs due to a lack of suitable mates, even where potential mates are present, which causes a population collapse. Further, Hughes (1991) shows that management of captive populations relying on effective size (Ne) alone cannot reliably conserve specific gene families such as MHC. Hughes went on to suggest management of captive populations to maximise diversity of few important functional genes such including MHC, a view that was and remains controversial (Gipplin & Wills 1991, Miller & Hedrick 1991, Ballou *et al*.2010). The results of this work may, if true, provide support for Hughes' view. If

MHC is important to mate choice but monogamous species do not have higher levels of it, then it can be assumed that endangered monogamous species have low levels of genetic diversity at MHC. Where mate choice is based on MHC, a glass effect becomes likely in the medium term even if general genetic diversity is managed. Therefore, these results may add further weight to the suggestion that genetic diversity at MHC be managed independently, at least for some species.

As a pilot study, these results suggest directions for future research. Although the tools required to genotype MHC for multiple non-model vertebrates did not exist when the literature reviewed herein was written, they soon will (e.g. Babik 2010, Chapter 4). This technical innovation should allow the question to be re-examined with more clarity and fewer limitations. By using techniques such as RAD-TAG genotyping (chapter 5) it should be possible to genotype both neutral and functional diversity in a single reaction for non-model species, again allowing for multiple species to be analysed with a single method. In addition, these methodologies could also be used to examine species such as longnose filefish (Kokita & Nakazono 1998) that display both monogamy and multiple mating. Alternatively, MHC variation may be so great that even these methods are unable to provide a general genotyping approach and each genotyping method must be species specific. As these questions are unable to currently be answered with a literature review, this work supports the need for the development of general MHC genotyping methods, as a greater sample size would enable the questions to be answered with confidence.

3.3 The effect of group living on genetic MHC diversity.

3.3.1 Introduction

Throughout evolutionary history, pathogens have applied selective pressures to their vertebrate hosts (Freeland 1976; Zuk & Stoehr 2002; Bernatchez & Landry 2003). The diseases caused by pathogens such as bacteria and viruses may be communicated via numerous and diverse vectors including respiration, fluid transfer, sexual transmission and in-utero transmission from mother to offspring (Lange & Ferguson 2009; Abrams & Miller 2011; Loke 2013). These infections may result in death (Daszak *et al.* 2000; Taylor *et al.* 2001; Lips *et al.* 2006), loss of condition (Ringø *et al.* 2007; Sorci & Faivre 2009), behavioural defects (Klein 2003) and/or competitive disadvantage in their hosts (Howard & Minchella 1990; Folstad & Karter 1992), all of which reduce fitness in vertebrates.

Pathogens have, however, co-evolved with their hosts. The development of immune responses, as well as phenotypic and behavioural plasticity in vertebrate hosts, has exerted reciprocal selective pressures on pathogen genomes (Agrawal 2001; Woolhouse *et al.* 2002). Host mortality rates have also driven pathogen evolution. Where excessive virulence has led to hosts dying too rapidly to allow transmission to a new host, many pathogens have decreased in virulence over time until an evolutionary stable state was reached (Bremermann & Thieme 1979; Baalen & Sabelis 1995). For example, the Myxomatosis virus in Australia evolved lower levels of virulence in response to the over-rapid death of its rabbit hosts (Fenner & Meyers 1978).

The foremost vertebrate evolutionary response to pathogens at the cellular level has been the development of the acquired immune system (Kimbrel & Beutler 2001). Genes responsible for the acquired immune system have been identified in the oldest known jawed vertebrates (Kasahara *et al.* 1995; Flajnik *et al.* 1999). These genes are, however, absent in the primitive jawless fishes whose evolution predates jawed vertebrates (Flajnik & Kashahara 2001), indicating an evolutionarily recent origin for acquired immunity. A significant problem for the acquired immune system, however, is the

discrimination of pathogens from self (Kimbrel & Beutler 2001). This problem is primarily managed by the major histocompatibility complex (MHC) (as described in chapter 1.4), a group of genes unique to vertebrates (Arala-Chavez & Sequeira 2000). Genes at the MHC code for proteins that, when expressed on the surface of cells, detect foreign bodies (Kashahara *et al.* 1995).

Increased genetic diversity at the MHC has been correlated with higher fitness in vertebrates (see Chapter 1). As MHC is involved in antigen presentation, greater heterozygosity at MHC loci should enable greater pathogen recognition (Ejsmond & Radwan 2009; Kumar *et al.* 2011). This is supported by numerous studies which link higher MHC diversity with increased fitness resulting from resistance to a greater number of pathogens (e.g. Reusch *et al.* 2001; Thoss *et al.* 2011). The relationship has been demonstrated both in captivity (Kubinak *et al.* 2012) and in the wild (Agudo *et al.* 2012; Knafler *et al.* 2012). Additionally, the relationship is robust, persisting even when the immune system is compromised (Carrington *et al.* 1999). The unusual evolution of this gene complex is extensively discussed in Chapter 1.

Many vertebrate taxa have adopted social strategies in order to obtain increased fitness from group-living. Potential fitness gains include mutual defence from predators, the ability to monopolise resources, increased foraging success, cooperative rearing of young, and increased swimming efficiency (Ebensperger *et al.* 2011; Shultz *et al.* 2011; Beauchamp 2013; Burgerhout *et al.* 2013; Stankowich *et al.* 2014). There are, however, significant potential costs of social living.

Species that live in social groups experience greater physical proximity to conspecifics, leading to an increase in the frequency and duration of contacts between individuals (Hughes *et al.* 2002; Altizer *et al.* 2003). Such physical contacts are key opportunities for the transmission of pathogens between hosts. Exposure to pathogens is, therefore expected to be higher in large social groups than in solitary species or taxa living in less dense populations (Alexander 1974; Freeland 1976; Arneberg *et al.* 1998; Møller *et al.* 2001). As exposure to the effects of pathogens is higher in taxa with social living strategies, it is reasonable to assume higher levels of defence against these threats to have

evolved in such species. A long term collaboration between the groups headed by Andrew Beattie and Adam Stow at Macquarie University continues to assess antimicrobial defences across a range of invertebrate taxa. As well as information on the defences of each species this collaboration also determined that anti-microbial strength, and presumably disease risk, increased with group size in invertebrates (Stow *et al.* 2007, Hoggard *et al.* 2011).

As MHC is a major defence mechanism in vertebrates, we would, therefore, expect a similar trend toward greater MHC diversity with larger group size in vertebrate taxa. Pathogen-mediated selection is believed to be the key driver of the high genetic diversity observed in major histocompatibility complex (MHC) genes (Spurgin & Richardson 2010). Therefore, it is reasonable to suggest that as disease risk increases so too will MHC diversity to counter that risk. Indeed, this has been shown in some cases e.g. clines of MHC diversity have been found within species that correspond to clines of disease risk (Dionne *et al.* 2007). However, we know of no work examining the role of group living on MHC diversity.

This study aims to determine if group living affects MHC diversity and if group size is correlated with MHC diversity.

3.3.2 Methods

Data collection

See Chapter 3.1 (full method available in appendix 3.5)

Data analysis.

In order to determine the effect of group living on MHC diversity, species were grouped as either Group living (group size 3 or more) or Solitary (solitary or pair living). A Mann-Whitney U-Test was performed using IBM SPSS 21 (IBM Corp 2012) to compare the number of MHC variants in each group.

In order to investigate the effect of group size on MHC diversity a Spearman Rank Order Correlation using SPSS 21 (IBM Corp 2012) was performed between maximum group size, as described in the literature, and number of MHC variants.

3.3.3 Results

Group living species were found to have a higher mean number of MHC variants than Solitary and Pair living species (Table 1); a Mann-Whitney U-Test found that this difference was significant (p=0.032). However, there was no relationship between the maximum size of the group and the number of MHC variants (Spearman's Rho = 0.029, p-value = 0.851)

Table 3.5: The number of MHC variants in Solitary and Group Living species

		N	Mean	Std. Deviation	Std. Error Mean
MHC_NumberOfAlleles	Solitary	16	7.3125	5.32565	1.33141
	Group Living	47	17.1277	20.37327	2.97175

3.3.4 Discussion

The analysis suggests that whilst group living significantly increased the number of MHC variants in a species, the maximum group size had no effect.

Group living dramatically increases the number of MHC variants found in a species. This is presumably a reflection of heightened disease risks associated with group living. Group living greatly increases the number of close interactions between individuals and thus the potential for disease transmission. This has been identified as a major cost associated with sociality (Coates *et al.* 1995, Hughes *et al.* 2002) and is the foundation of approaches attempting to understand disease based on social interactions (Wey *et al.* in 2008, Sih & Wey 2013). However, this relationship is not always as simple as is presented above, particularly in invertebrates which have been shown to have higher

pathogen survival rates when kept in groups (Hughes *et al.* 2002) and show horizontal transmission of pathogen resistance (Traniello *et al.* 2002). Nevertheless the relationship is largely accepted, particularly in vertebrates where previous studies have documented behavioural adaptations in social species to avoid sick animals both in the lab (Minchella 1985) and in natural environments (Behringer *et al.* 2006). These data add an additional layer to the understanding of adaptations to the increased pathogen threat posed by sociality. They suggest an increased ability to recognise pathogens in social species. This may be important to our understanding of the evolution of sociality.

It is assumed that increased MHC diversity is an adaptation to group living, however MHC diversity may constrain group size. The evolution of sociality is difficult to explain (e.g. Lin & Michener 1972, Gadagkar 1985, Schwarz & Hogendoom 1998, Fischman 2014) but one of the most widely accepted theories suggests a 'monogamy window' whereby species must evolve monogamy in order to then evolve sociality (Hughes *et al.* 2008, Boomsma 2009). Therefore the evolution of monogamy must predate the evolution of sociality. Similarly, it is possible that in order for species to live in a group it must first evolve enough diversity within an immune response to survive the increased disease threat. In vertebrates increased diversity at MHC would be a method of meeting this threat. Although investigating this question is beyond the scope of this study, it is an idea that may be worth further consideration, if difficult to test.

In both theory and experimental work a directly proportional relationship between group size and disease risk is described. Classical epidemiology focused on the interaction between group size and disease transmission (eg. Busenberg *et al.* 1983 from Busenberg & Driessche 1986). Stow *et al.* (2007) found a similar relationship demonstrating that as sociality, and thus group size, increased in bee species so did antimicrobial defence. Additionally, Frankham (2005) describes a reduction in disease benefit of the intentional fragmentation of captive populations. Therefore, we expected that as group size increased so would MHC diversity. However, our study found no evidence of this.

One explanation for the lack of relationship between group size and MHC diversity is an error in experimental design. Using maximum group size was a necessary compromise given the availability of data, but may have been unacceptable one. Established epidemiology (e.g. Busenberg & Driessche 1986) has established that it is not just the size of a group but also how that size fluctuates that determines the transmission of disease. Our study had no method for accounting for size fluctuation. Further, somewhat more recent studies, such as the Colorado Springs Study (Klovdahi *et al.* 1994), demonstrated that the types of connections within the group also affect disease transmission. Again, our data had no measurement for these. Unfortunately, these are human based studies and similar research is not available for the breadth of vertebrate species. Therefore, the lack of a definitive confidence in the methods is somewhat understandable, though not less frustrating

Alternatively, the lack of relationship may be real. In contrast to the overall work by Stow *et al.* (2007), a smaller study using a single lineage (Hoggard *et al.* 2013) was unable to find an effect of group size on antimicrobial defence and thus presumably pathogen threat. This similarity is interesting as both this study and that of Hoggard *et al.* 2013 use relatively few closely related taxa compared to those used by Stow *et al.* (2007). Given the strong evidence of an increasing pathogen threat with increasing group size in the literature these results, if real, may be the product of a second mechanism managing the increased disease threat associated with larger groups. Several prospective mechanisms have been found that may account for this in invertebrates including grooming (Evans & Spivak 2010) and removing sick individuals (Baracchi *et al.* 2012). In vertebrates these and other behavioural adaptations may be important (Penn & Potts 1998) as may other parts of the adaptive immune system. Indeed, a case could be made for numerous possibilities, including microbial symbiosis (as per Mazmanian 2008) and thus this puzzle, if borne out by more exhaustive research, suggests the picture of host-pathogen evolution in vertebrates is even more complex than we first suspected.

4 A Novel method of genotyping MHC for conservation genetics.

An understanding of the amount and distribution of MHC diversity is important for many species of conservation concern. However, current methods for genotyping MHC are poorly suited to informing conservation. This study attempted to develop a method for genotyping MHC that would be ideal for conservation genetics using next generation sequencing and tagged primers. Analysis using the next generation sequence program Mothur was inconclusive but further analysis using Microsoft Excel and biological and molecular controls determined that the method failed to produce reliable results. The insights gained during this attempt are discussed to benefit future work.

4.1 Introduction

The major histocompatibility complex (MHC), an area of the genome that has received considerable attention in immunobiology and evolutionary biology, has recently been incorporated into conservation decision making. As previously discussed, MHC was first described by Gorer (1936) and its role in organ transplant rejection determined by Snell & Higgins (1951). As our understanding of the role of MHC has grown (e.g. reviews by Klein & Figueroa 1986, Fernando *et al.* 2008), so has the importance of understanding MHC diversity when solving conservation problems. MHC diversity has been shown to have a key role in disease resistance (e.g. Sollid & Thorsby 1993, Siddle *et al.* 2010) with greater diversity resulting in greater disease resistance in many species (e.g. Radwan *et al.* 2012). Therefore, MHC has become important to managers of threatened and endangered species where disease epidemics pose a significant risk (e.g. Laurance *et al.* 1996, Laurenson *et al.* 1998). Furthermore, the gene complex has other roles which are important in conservation management of some species.

MHC has been shown to play an important role in mate choice in many species and, in these species, may be important in avoiding population collapse. The effect of MHC in mate choice has been demonstrated in numerous vertebrate taxa including mammals (Yamazaki *et al.* 1976, Wedekind *et al.* 1995), birds (e.g. Bonneaud *et al.* 2006), reptiles (Olsson *et al.* 2003) and fish (Reusch *et al.* 2003). Therefore, understanding MHC diversity may be important for captive breeding programs in which successful mating can be a significant challenge (Hughes 1991). Furthermore, preferences are often for MHC dissimilar mates (e.g. Wedekind *et al.* 1995). These data have led to suggestions that MHC imprinting, recognising familiar MHC scents and avoiding them in mates, may be a widespread method of inbreeding avoidance (e.g. Penn & Potts 1998). Should this be true, MHC diversity would be even more important for some species because 'the glass effect' based on MHC similarity may cause a population collapse. Therefore, if MHC is a mate choice signal in a species of conservation concern, understanding the level and distribution of genetic diversity is crucial. Even when mate

choice is not affected by MHC and disease epidemics are not an important consideration, knowledge of MHC diversity is useful to conservation managers as a functional genetic marker.

MHC is the most genetically diverse of all functional genes yet discovered in vertebrates (Potts *et al.* 1993) and is, therefore, an ideal genetic marker. As functional diversity is the fuel of evolution, it is of more immediate concern to conservation than neutral diversity (see chapter 3 for a more detailed analysis) (Crandall *et al.* 2000). Typically, neutral markers are used as a surrogate for functional diversity as their high levels of variation increases the resolution of analysis of populations (e.g. Bowcock *et al.* 1994, Rhodes *et al.* 1998). However, as MHC is highly variable it facilitates high resolution studies (e.g. Dionne *et al.* 2007) and MHC diversity can also provide information on local adaptation (Eizaguirre & Lenz 2010), a key concern for conservation (Ujvari & Belov 2011).

Distributions of MHC diversity would provide information about changing disease threats (Eizaguirre & Lenz 2010). This is significant as disease poses a potentially catastrophic risk to endangered populations (Snyder *et al.* 1996). Further, specific diseases form parts of the environment that species can adapt to. Therefore, using MHC to inform outcrossing will reduce the risk of destroying local adaptation. In conservation fear of destroying local adaptation has impeded the practice of outcrossing (Templeton 1986, Edmands 2007), despite such outcrossing being vital to species survival (Frankham *et al.* 2011). Therefore, MHC offers many benefits as a genetic marker. Accordingly, genotyping MHC is recommended for species of conservation concern (Crandall *et al.* 2000, Ujvari & Belov 2011) even as a replacement for genetic markers (Hughes 1991) However, that approach is not universally endorsed(Miller & Hedrick 1991, Sommer 2005). Nevertheless, in order to incorporate knowledge of MHC diversity into conservation it must first be genotyped.

How is MHC currently Genotyped?

There are a number of methodologies which have been used to genotype MHC. These have included single stranded conformational polymorphism (SSCP) (Kostia *et al.* 1998), systems based on microsatellites (e.g. Meagher & Potts 1999) and methods based on next generation sequencing (e.g.

Wegner 2009). None of the currently employed genotyping methods, however, provide the type and/or quantity of data that is needed to inform conservation in many vertebrate species. Most approaches to MHC genotyping, e.g. the approach used by Kostia *et al.* (1998), are single gene approaches similar to those used other functional genes of interest (e.g. Heils *et al.* 1996). However, because MHC is a multilocus complex it may not function in the same manner as a single locus . This is because genetic variation at MHC is created by both within- and between- locus recombination (Gu & Nei 1999, Reusch & Langefors 2005), and the number of loci is highly variable between (Bernatchez & Landry 2003) and within (Freeman *et al.* 2006, Siddle *et al.* 2010) species. Therefore, the genetic approaches that have focused on a single locus have missed much of the genetic variation in the complex. This is true even if the complete structure of MHC is already known for the species (Stewart *et al.* 2004). However, for most species of conservation concern, detailed genomic information is not yet available. In these cases because MHC is a gene complex the number of MHC loci is unknown and it is impossible to determine how much of the genetic variation at MHC is missed by single locus approaches. Therefore, a more general approach to genotyping MHC is necessary.

The gene complex nature of MHC can complicate approaches to understanding and scoring the molecular products of genotyping efforts. Highly similar alleles can exist at multiple MHC loci (Yuhki *et al.* 2003). Therefore, a single set of primers may amplify many loci at once. This can be problematic in visualisation techniques designed for single locus genotyping such as single strand conformational polymorphism (SSCP). Although SSCP has been used to genotype MHC (e.g. Binz *et al.* 2001) when multiple loci have been amplified false alleles may be introduced where single strands from different alleles reanneal with each other in a heteroduplex (Sunnucks *et al.* 2000). Additionally, the highly conserved nature of MHC (e.g. Yuhki *et al.* 2003, Schwensow *et al.* 2010) means that a single primer set may amplify multiple species. This in turn raises the risk of PCR contamination as more types of alien DNA can be amplified in the reaction. At the same time, MHC variants can have very different sequences, a level of divergence uncommon in other functional

genes because of the way that MHC evolves and is maintained in a population (Potts & Wakeland 1993, Spurgin & Richardson 2010). This increases the risk of null alleles as a single primer set may not amplify all alleles at a locus. Furthermore, cloning and sequencing approaches, which have also been used to successfully genotype MHC (e.g. Hordvik *et al.* 1993), become more laborious as more loci and alleles are present, particularly when alleles of the same length have different sequences, as can be the case in MHC. Furthermore, both of these problems are compounded in many species of conservation concern as sequence libraries do not exist and thus it is impossible to predict how many loci will be amplified until molecular reactions are carried out and sequencing is performed by a set of PCR primers. A newer method of genotyping MHC based on next generation sequencing may help to solve many of these problems.

Babik *et al.* (2009) have developed a method of genotyping MHC using 454 sequencing in bank voles (*Myodes glareolus*). The method, summarised in Fig. 4.1 below, relies on amplifying multiple loci with tagged forward primers and then parallel sequencing all individuals at once. The tagged forward primers allow for the sequences, which are the alleles, to be assigned to individuals. The method is advantageous in that it simultaneously genotypes multiple loci and is capable of processing multiple individuals in a single run. Furthermore, as this is a sequencing-by-genotyping approach, it is also able to capture all of the genetic variation present in alleles. Accordingly, the method is ideally suited to conservation genetics. However, this method relies on the existence of a large sequence library which, while available for the bank vole, is typically not available for species of conservation concern. Consequently, this study has attempted to modify the methodology developed by Babik et al. (2009) in order to enable it to better answer conservation questions.

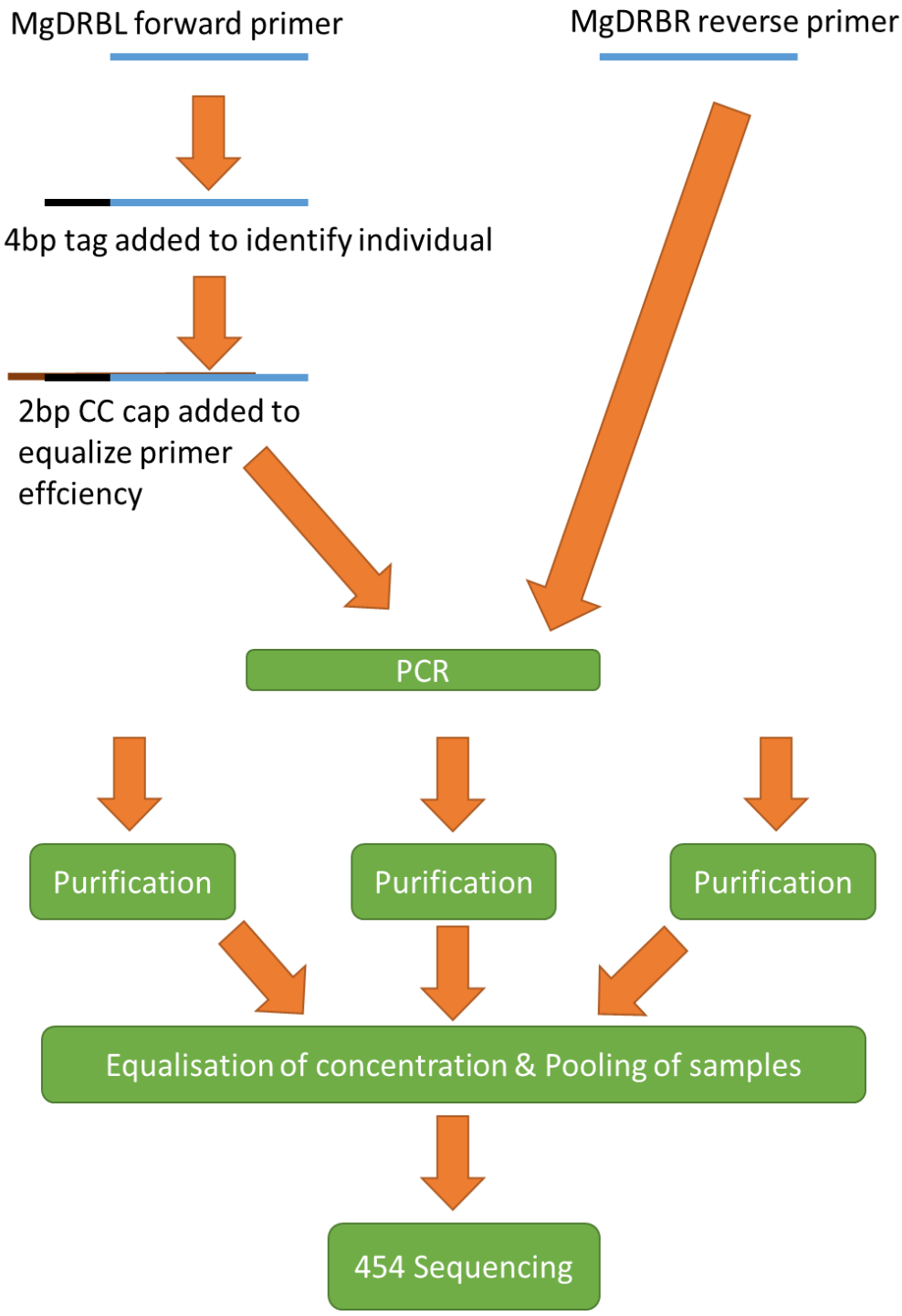


Figure 4.1: The genotyping method used by Babik *et al.* (2009). Degenerate primers were created and the forward primer tagged with a 4bp tag, a 2bp cap was then added to equalize primer efficiency before PCR was carried out. Each PCR product was purified using QIAGEN Mini Elute column before concentrations were equalized and products were pooled. 454 sequencing was then used to sequence PCR product.

1

What would a conservation genetics approach to MHC genotyping look like?

In order to design a method of genotyping MHC for conservation, attention must be paid to both theoretical and practical considerations. As conservation is usually concerned with overall levels of genetic diversity rather than diversity relating to a single disease threat (e.g. Reed & Frankham 2003), the method does not need to genotype a single locus. However, Sunnucks (2000) suggests multi-locus methods are not as flexible, informative or connectable as single locus approaches. Nevertheless, variable numbers of MHC loci (discussed in Chapter 1) and differences in the structure of MHC in different vertebrates (also discussed in Chapter 1) mean that genotyping multiple MHC loci separately is not practical. Therefore, the method for genotyping MHC for conservation must be a multi-locus method. Fortunately, the amount of genotyping data produced by next generation sequencing methods provides unprecedented resolution (Binladen *et al.* 2007) which may redress problems usually associated with multiple loci approaches. Significantly, species of conservation concern do not typically have large sequence libraries so the method cannot require a detailed genetic map. Furthermore, as conservation is time sensitive (e.g. Martin *et al.* 2012), the method should not require long optimisation periods and the genotyping process should be rapid. Finally, as conservation funding is limited and currently insufficient to address all human impacts in the Anthropocene (James *et al.* 2001, McCarthy *et al.* 2012), the method should be as cost effective as possible.

In this study we attempted to develop a method for genotyping MHC that meets the above requirements. Using degenerate primers allows for multiple loci to be amplified in multiple species (e.g. Villesen & Fredsted 2006). This method is tested both with traditional genotyping visualisation methods and with genotyping by 454 sequencing, a next-generation sequencing approach. When genotyping by next generation-sequencing, a modified system of forward and reverse tags on primers reduces costs and increases the numbers of individuals that can be processed in a single sequencing run. Methods and approaches for analysing the type of data produced by genotyping by next generation sequencing are also discussed.

4.2 Methods & Results

The methods and results have been combined for ease of interpretation. A flowchart describes the overall method and then each section is discussed in detail.

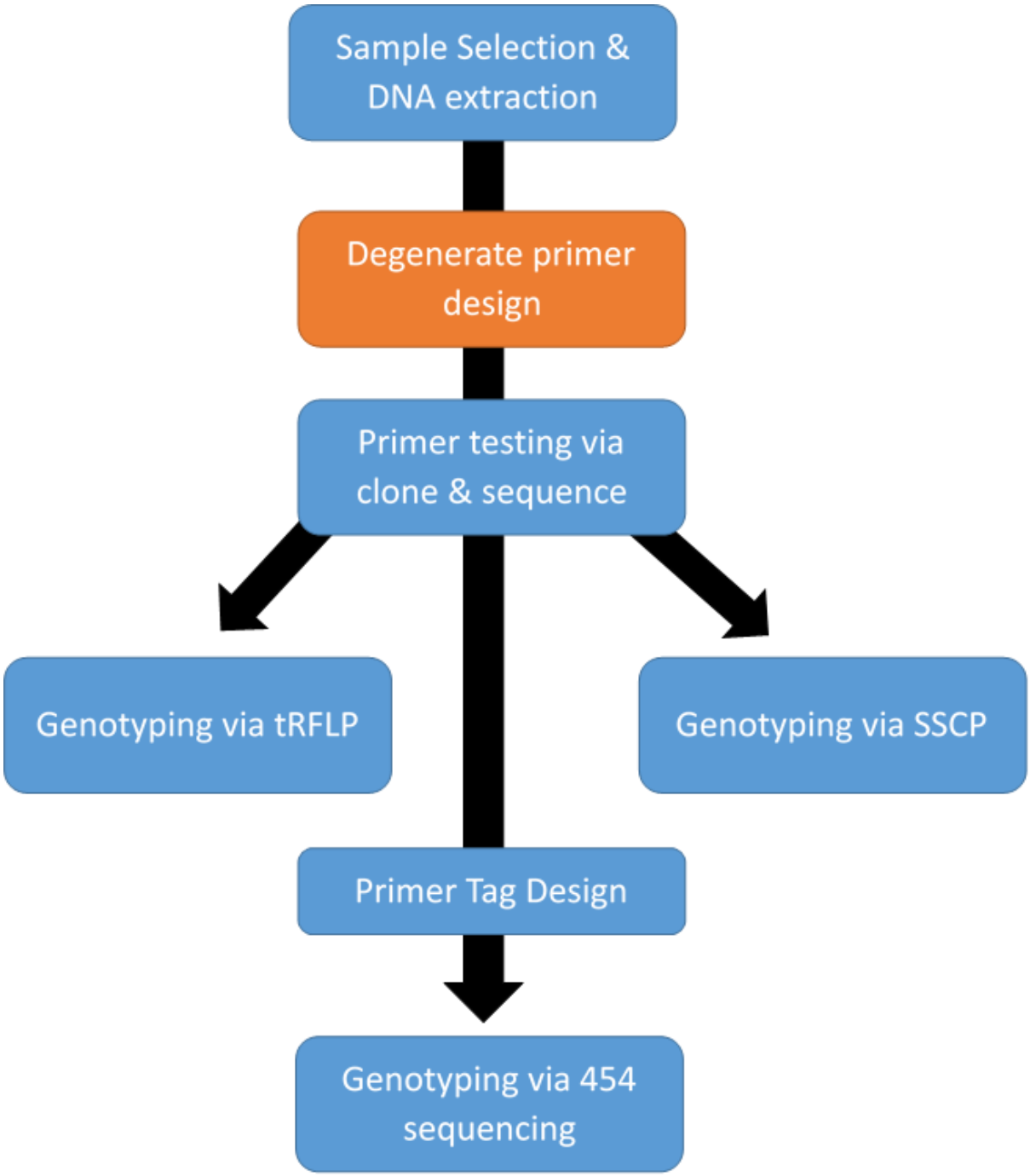


Figure 4.2*:* A flowchart describing a summary of the *in vitro* methods

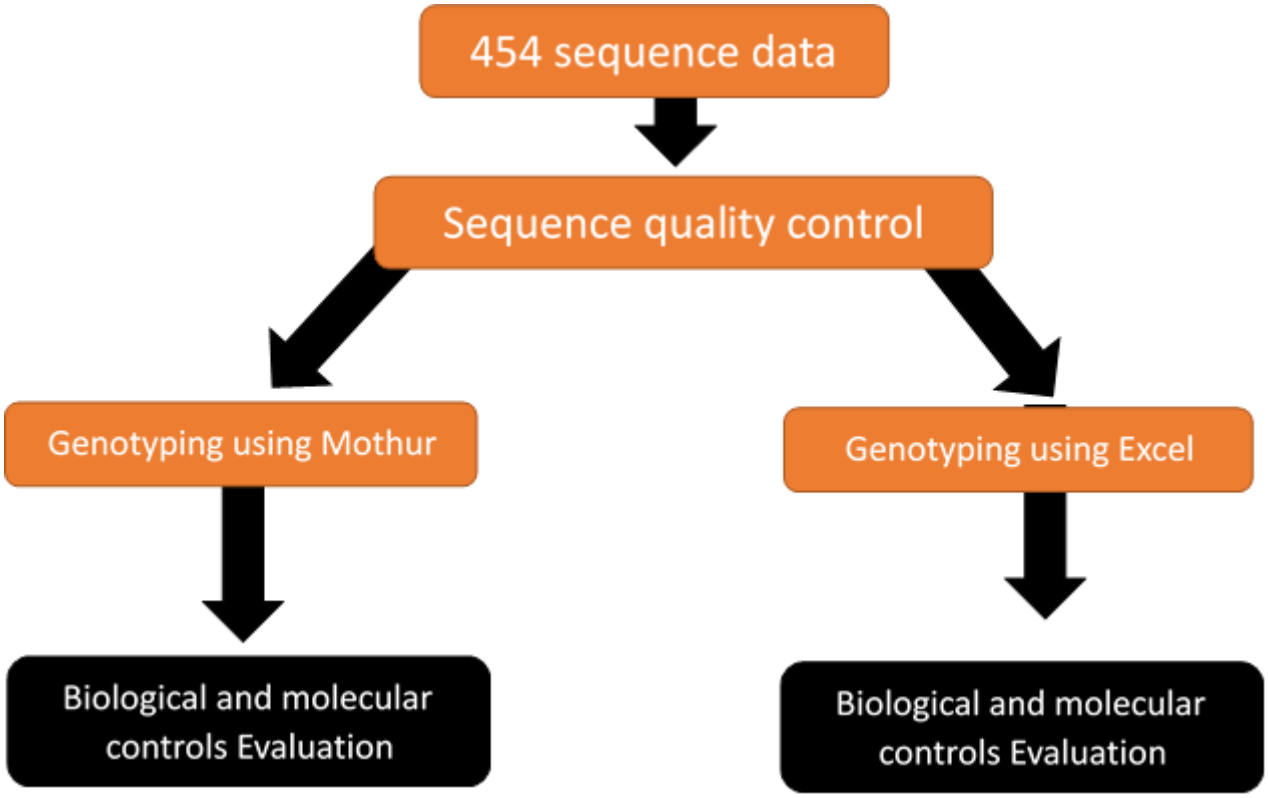


Figure 4.3 **:** A flowchart describing the in silico methods

1

2

Sample Selection & DNA Extraction

This study uses 208 samples, 191 *Egernia cunninghami* (Cunningham's skinks) (collection methods described in Stow *et al.* 2001) that were included to test the ability of this method to genotype populations, and 15 samples of other vertebrates made up of: 2 Western Spiny-tailed Skink *(Egernia stokesii)*, 3 Spotted dtella (*Gehyra punctata*) , 3 Tree skink (*Egernia striolata*), 3 Great Desert Skink, (*Liopholis kintorei),* 2 Red bellied black snake (*Pseudechis porphyriacus)* and 2 Australian Sea Lions (*Neophoca Cinerea*).

The use of *E.cunninghami*, a species for which some previous population genetics research exist, presented a number of advantages. Firstly, the populations used in this study had been used previously and genotyped using microsatellites (Stow *et al.* 2004a, Stow *et al.* 2004b). Therefore, the comparison between neutral and functional diversity (previously examined in chapter 2.1) could be re-examined within a single species. Additionally, previous studies (Stow *et al.* 2004a) found a loss in neutral genetic variation associated with land clearing. This analysis would therefore be able to compare the effect of land clearing on neutral genetic diversity to its effect on functional genetic diversity. Furthermore, as MHC is notoriously difficult to genotype (Strand & Höglund 2011, Sommer *et al.* 2013), the samples were chosen to include a biological control in the form of two family groups where parentage had previously been determined (Stow *et al.* 2004 b). This additional control allows for confirmation of Mendelian inheritance of alleles.

DNA extraction for all samples was performed using a modified salting out protocol (Sunnucks & Hales 1996). Briefly, a small sample of tissue is crushed and incubated with 10µ protenase K (10mg/ml) and 800µl of TNES overnight at 37 ° Celsius. Proteins are then precipitated with 170µl 5M NaCl and discarded. DNA is then precipitated with 300µl of ice cold ethanol, the supernatant discarded and the DNA washed with 400µl of 70% ethanol. The supernatant was again discarded and the pellet air dried. DNA was resuspended in 20 µl of sterile water.

In order to amplify MHC a set of degenerate primers were developed.

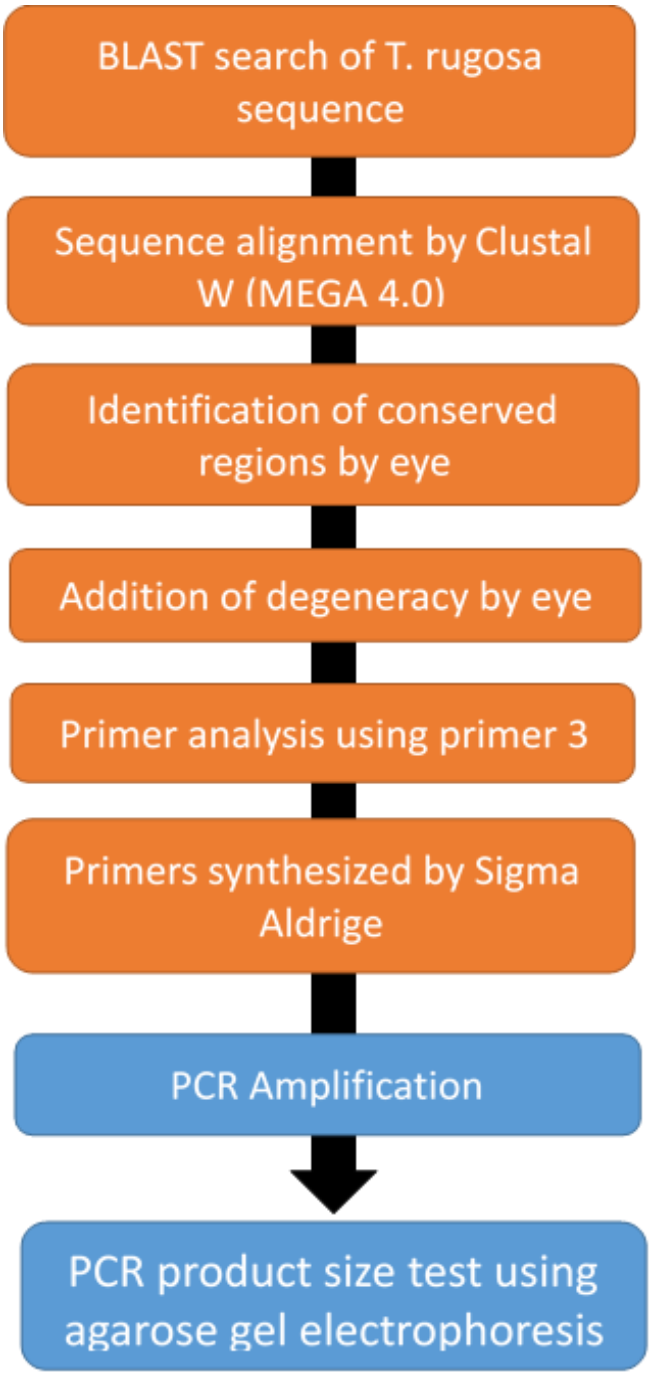


Figure 4.4: A flowchart describing the pathway to the creation of degenerate primers for MHC

Degenerate Primer Design

A partial sequence of Blue tongue lizard (*Tiliqua rugosa*) MHC (Gardener 2012 pers. comm.) was used to search the NCBI BLAST nucleotide database (Altschul *et al.* 1990) with the parameter 'somewhat similar matches'. The sequence matched most closely to a number of vertebrate MHC sequences, specifically MHC from *Anolis, Ameiva*, chicken and opossum. These sequences were exported to MEGA 4.0 (Tamura *et al.* 2007) and aligned by CLUSTAL W, using the default

parameters, in order to identify regions of the MHC that were conserved between these diverse vertebrate taxa.

The aligned sequences were examined by eye in order to determine potential primer sites. Highly conserved regions approximately 200-300 b.p. apart were noted, the short sequence length being necessary to facilitate subsequent genotyping by 454 sequencing (Margulies *et al.* 2005, Rothberg & Leamon 2008). Within these regions degeneracies were added to the potential primer sequences by eye so that all combinations of nucleotides in the priming sites of all vertebrates were included. Primer regions were chosen to limit degeneracy as increased degeneracy of PCR primers can increase the likelihood of secondary structure in PCR primers (Singh *et al.* 1998) and cause the primer to amplify unintended regions of the genome (Rose & Schultz 1998).

The primer sets designed by eye were then analysed using Primer 3 (Untergasser *et al.* 2012) to determine suitability. Primer sequence was examined to establish melting temperatures and determine the likelihood of secondary structure. Only primers with LOW/MODERATE likelihood of secondary structures were selected for further analysis. Primers were synthesised by Sigma-Aldrich and resuspended in ddH20. The details of the primer sets that were synthesized are summarized in the table below.

Table 4.1: Primers used in analyses. Primers MHC2Ex2F2 and MHCEx2R2 are from Miller et al. 2005, all primers with prefix ec were designed for this study

Primer Name	MHC class	size (b.p.)	5' Sequence	tm
MHC2Ex2F2	II	540	GCGCRGWGCCYCCMGARCATT	67.44
MHC2Ex2R2	II		GCTGGSGTGCTCCACCTGGCA	69.87
ecMHC-1f	I	190	TGTTGTGGGGTATGTGGATG	60.09
ecMHC-1r	I		CCCTCCAGTCTGGTTGTAGC	59.72
ecMHC-2f	I	190	ATGTGGACGACAAGCTCCTT	59.73
ecMHC-2r	I		CCCTCCAGTCTGGTTGTAGC	59.72
ecMHC-3f	I	170	AGCACTACGACAGCAACACG	60.12
ecMHC-3r	I		CCCTCCAGTCTGGTTGTAGC	59.72
ecMHC-4f	I	160	AGCGCTTCTTCCACTACGAC	59.64
ecMHC-4r	I		CCCTCCAGTCTGGTTGTAGC	60.21
ecMHC-5f	I	180	CRTNGTYGTGGGRTAYGTGG	42.45
ecMHC-5r	I		CCTCCAGTCTGGTTGTAGC	54.13

Reliable amplification of DNA using these primers proved to be challenging. A number of PCR conditions were trialled and the strongest PCR product on agarose gelwas achieved with the following concentrations: in a final volume of 10 mL containing 0.5U Taq DNA polymerase (Promega), 10 mM forward primer, 10 mM reverse primer, 8 mM dNTPs, 1 unit Taq Buffer (Promega) and 2.0 mM $MgCl_2$. PCR amplifications had an initial denaturation at 94 C for 4 min followed by five 'touch down' cycles of 94 C denaturation for 15 s, annealing temperatures (60 C, 58 C, 56 C, 54 C, 50 C) for 30 s and an extension step of 72 C for 80 s. After the final touchdown cycle, another 30 cycles were carried out at 50 C annealing temperature with a final extension of 5 min at 72 C. PCR products were visualised on 2% agarose gel 2 with PAGE GelRed (Biotum) with a 100 b.p.

ladder as a positive control. However, the PCR was prone to contamination, visible as a clear band on the negative control well of a 1.5% agarose gel electrophoresis of the same approximate size as target bands, which could not reliably be removed. In an effort to minimize the possible sources of contamination PCR was carried out using PROMEGA EasyTaq clear, an all-in-one master mix which minimized opportunities for contamination resulting from pipetting technique. Furthermore, the reaction was carried out in a negative pressure chamber to further minimize the likelihood of contamination by foreign DNA. Final PCR amplification was undertaken with using the same conditions and concentrations with the addition of the all-in-one master mix and the products were run out on 1.5% agarose gel.

The results of agarose gel electrophoresis suggested that a single primer set was the best candidate for further analysis. All primer sets except MHC2Ex2F2 / MHCEx2R2 (from Miller *et al.* 2005) produced PCR product but all primers except for ecMHC-5 also produced additional bands outside of the target size range. This suggested that the degeneracy in the primers was allowing the amplification of non-target, presumably non-MHC sequence. Furthermore, these additional longer PCR products could not reliably genotyped with 454 sequencing technology due to read length restrictions (Rothberg & Leamon 2008). Therefore, further analysis used only primer set ecMHC-5 which amplified only a single band of the predicted size.

1

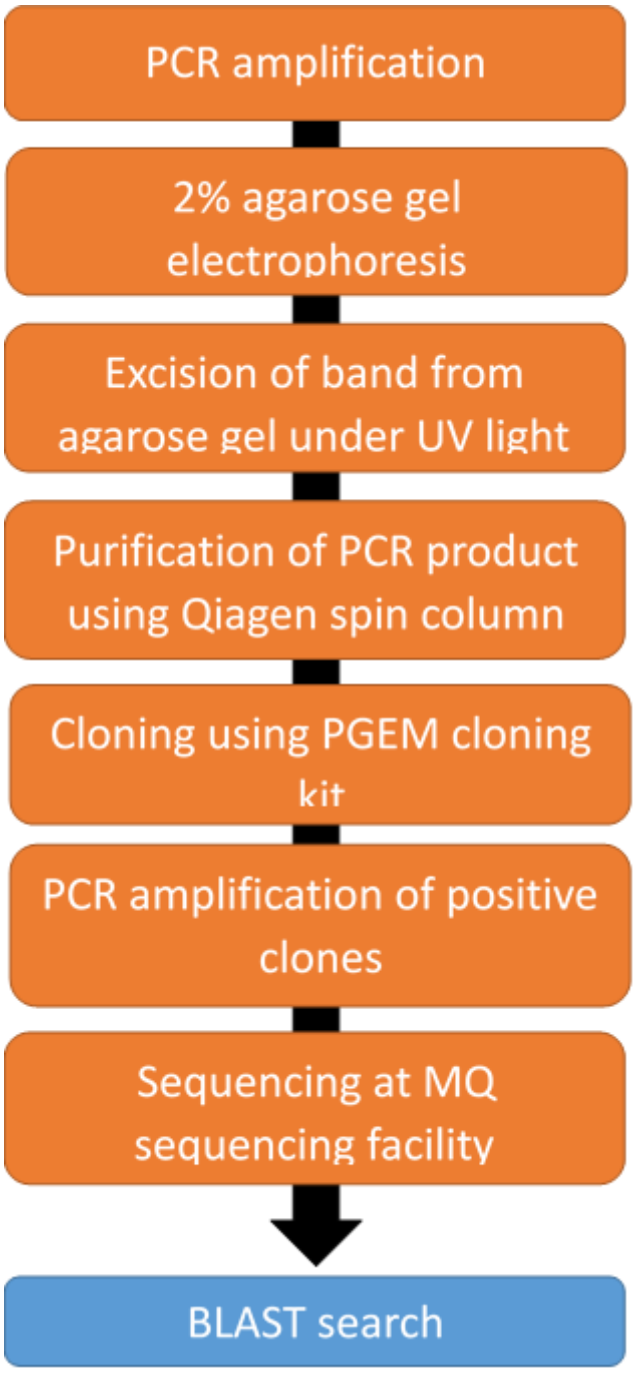


Figure 4.5: A flowchart describing the PCR amplification and testing of degenerate MHC primers by cloning and sequencing.

4.3 Methods: primer testing via cloning and sequencing

The PCR amplification followed the previously described method on 8 individual *E.cunninghami* animals. Cloning procedure was then carried out on the PCR products of a single individual using the PGEM kit (Promega) using the recommended procedure as per the manufacturer's instructions. A second PCR was carried out on the bacteria that had successfully incorporated the PCR product, which are visible as white colonies on the agarose plate. This PCR used the primers M13-f TGTAAAACGACGGCCAGT and M13-r CAGGAAACAGCTATGACC with the following concentrations: in a final volume of 10 mL containing 0.5U Taq DNA polymerase (Promega), 10 mM forward primer, 10 mM reverse primer, 8 mM dNTPs, 1 unit Taq Buffer (Promega) and 2.0 mM $MgCl_2$. PCR amplifications had an initial denaturation at 94 C for 4 min followed by 35 cycles of: 94 C denaturation for 15 s, annealing temperatures 50 C for 30 s and an extension step of 72 C for 80 s. PCR product was run out on 1.5% agarose gel purified using a MinElute Gel Extraction kit (QIAGEN) and sequenced on an ABI Prism 3130x1 genetic analyser (Applied Biosystems) by Macquarie University sequencing facility.

4.4 Results: primer testing via cloning and sequencing

The results of the cloning and sequencing project were a qualified success. The process itself gave yields much lower than similar studies using the same technology (Babik *et al.* 2009). However, the clones that were present were able to be amplified and resulted in high quality sequence. The sequences were analysed by BLAST search (Altschul *et al.* 1990) which resulted in 13 of 14 returning MHC matches (table 4.2). However, the relatively high E values, which denote the probable number of matches in the blast database (e.g. E-value of 1 suggests one match by chance a database of the current size), of these matches were concerning. A lower E-value (e.g. Palenik *et al.* 2003) would allow for more confidence in the matches. Nevertheless, as no MHC sequence for *Egernia* was

available, the high E-values may have indicated a knowledge gap rather than unsuccessful amplification. Additionally shorter sequences have higher E-values (blast.ncbi.nlm.nih.gov). Therefore, this result was interpreted as successfully amplification of MHC and genotyping was able to proceed.

Table 4.2: The results of BLASTsearch on amplified sections of MHC

Seq Name	Closest MHC match	Query Coverage	Max Ident	Max Score	E value
Clone 1	*Monodelphis domestica*	64%	76%	102	9e-19
Clone 2	As above				
Clone 3	*Gallus gallus*	93%	69%	77.0	4e-11
Clone 4	*Trichosurus vulpecula*	57%	84%	84.2	2e-13
Clone 5	*Amblyrhynchus cristatus*	40%	80%	80.6	3e-12
Clone 6	As above				
Clone 7	As above				
Clone 8	*Monodelphis domestica*	82%	72%	102	9e-19
Clone 9	*Macropus eugenii*	62%	74%	87.8	2e-14
Clone 10	*Gallus gallus*	92%	70%	83.4	9e-13
Clone 11	*Anas platyrhynchos*	82%	70%	75.2	1e-10
Clone 12	*Grus canadensis*	61%	71%	55.4	1e-04
Clone 13	*Ameiva ameiva*	72%	71%	71.6	1e-09
Clone 14	No MHC matches				

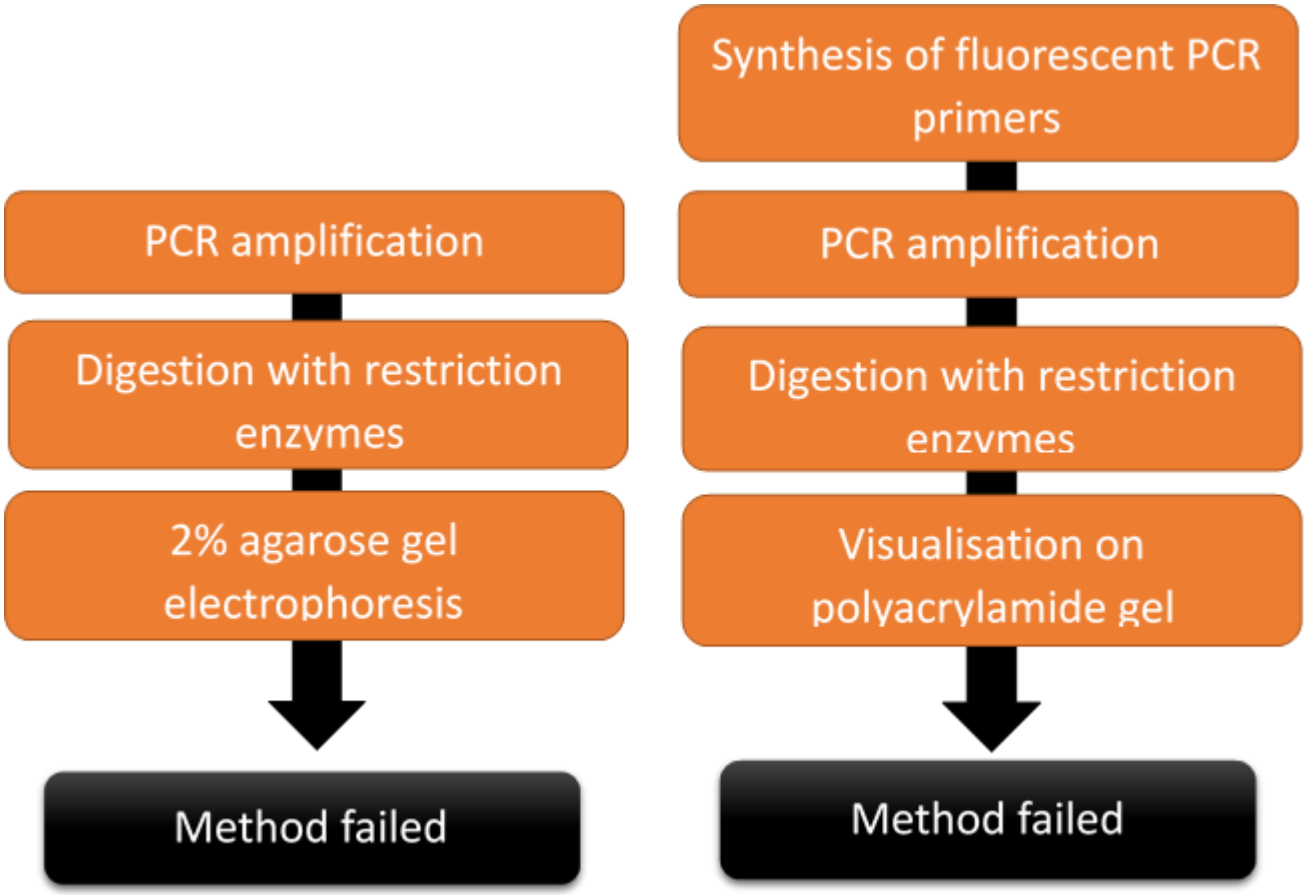


Figure 4.6: A flowchart describing the attempts to genotype MHC using RFLP & tRFLP techniques.

Genotyping via tRFLP

4.5 Methods: genotyping via tRFLP

The tRFLP protocol used is that described in Waldron *et al.* (2009) and in detail in appendix 4.1. The only modifications from this protocol is that a nested PCR was not required and the restriction enzymes used were BamH1, HindIII and EcoR1 (Promega). Further, the forward primer exMHC-5f had the fluorotag FAM welded to the 5' end to enable visibility on an ABI Prism 3130x1 genetic analyser (Applied Biosystems).

4.6 Results: genotyping via tRFLP

This method failed to genotype MHC. Though results were produced, the peaks on the electrophoretograms were not distinct. Furthermore, there was not clear and repeatable variability between samples. Presumably this is because the short PCR product did not contain restriction sites for the enzymes tested and thus the tRFLP and RFLP could not differentiate between alleles. However, the failure may have also been due to methological errors.

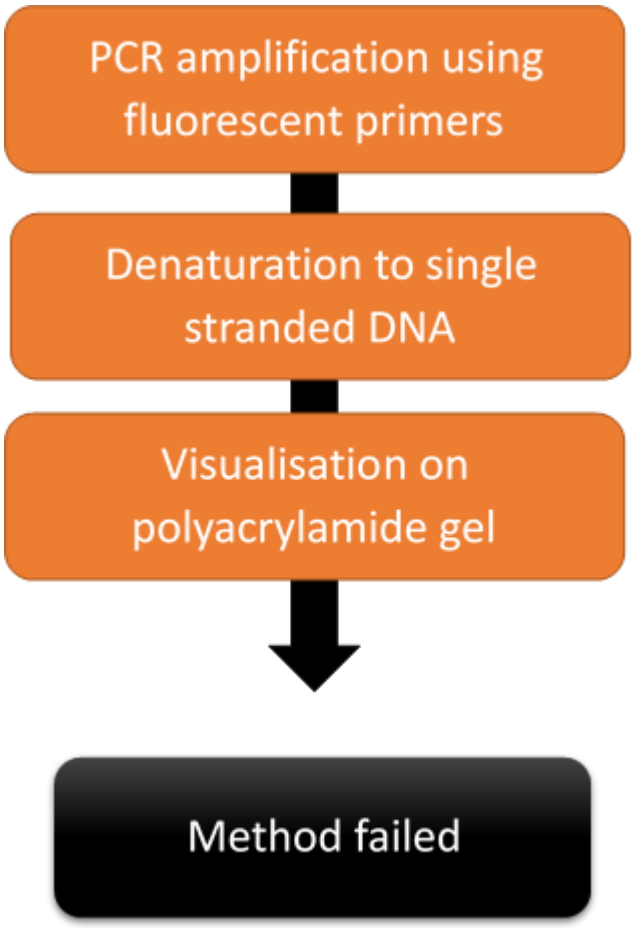


Figure 4.7: A flowchart describing the attempts to genotype MHC using SSCP.

4.7 Methods: genotyping via SSCP

SSCP of MHC was carried out as per Binz *et al.* (2001). The only alteration to their methodology was that labelled reverse primers were unnecessary for our reaction.

4.8 Results: genotyping via SSCP

This method failed to produce reliable genotypes. The electrophoretegrams produced were not consistent between reactions for the same individual, showing different peak patterns. Furthermore, 'alleles' were found present in offspring that were not present in parents. These results were presumably due to the presence of a large number of heteroduplexes which were increased by the number of loci amplified by the degenerate primers. However, it could also be due to errors in the methodology.

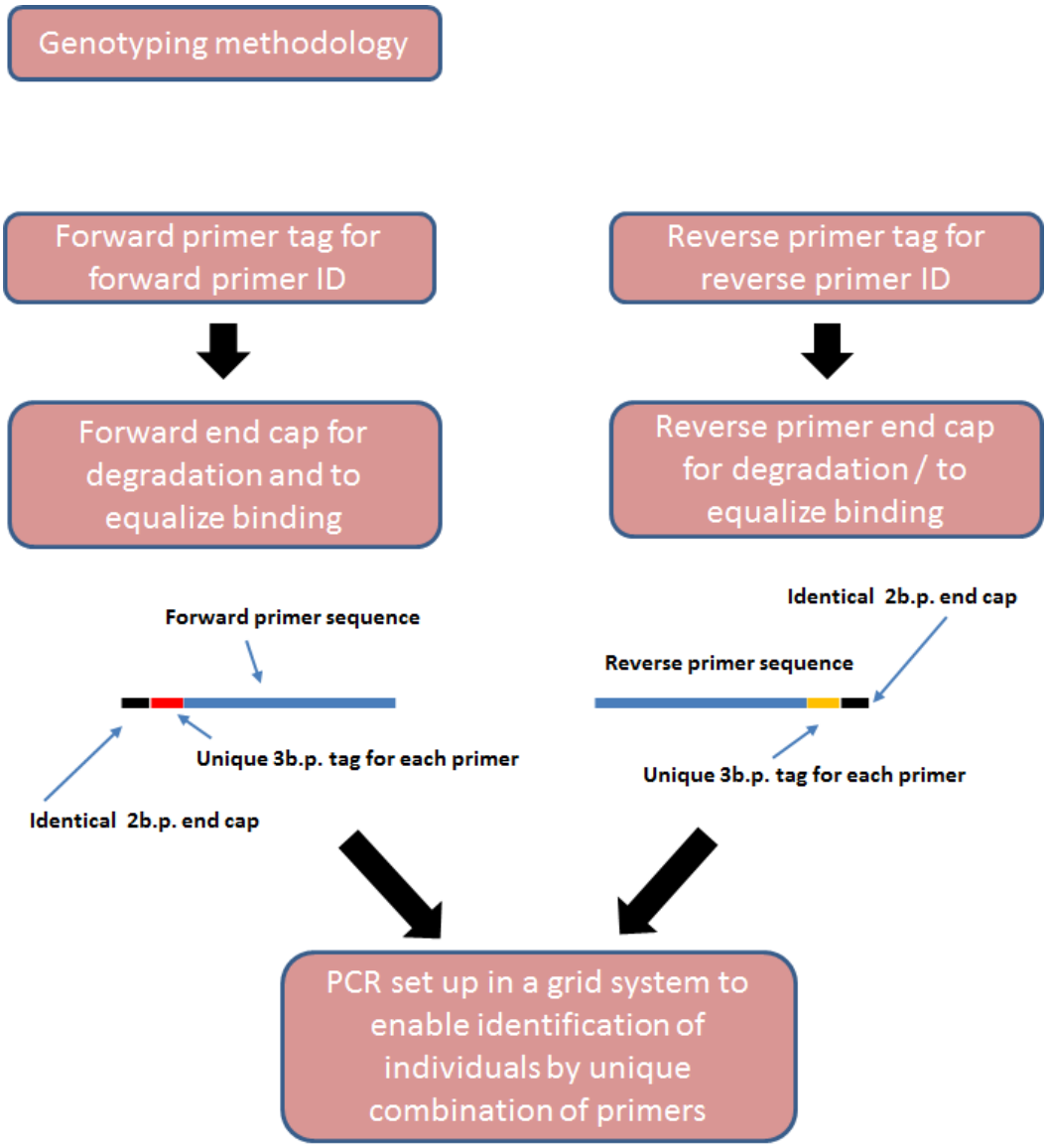


Figure 4.8: A flowchart describing the method of forward and reverse primer tagging and PCR half reactions. Unlike other methodologies it uses a tag on both forward and reverse primers.

4.9 Primer Tag Design

To accurately assign sequences to individuals genotyping by sequencing approaches use tagged primers. Our method is modified from the method developed by Babik and colleagues (2009) and is shown in figure 4.14 above. The method used in this study used 4 bp tags on the forward and reverse primers, and assigned a forward primer per row, and a reverse primer per column, on the PCR plate (shown below in figure 4.15). Therefore each individual, a single well on the PCR plate, had a unique combination of primer tag sequences. Furthermore, a 2 bp cap (CC on forward primers and GG on reverse) was used to buffer the effect of any degradation at the ends of the PCR product and to minimize preferential binding. This modification was necessary as it allowed for a much larger number of individuals to be included in the analysis with the same number of primers compared to the method of Babik *et al.* (2009). By using primer combinations to tag an individual as opposed to unique reverse primers an order of magnitude more individuals could be genotyped with the same number of primers e.g. 10 forward and 10 reverse primers creates 100 unqiue combinations, whereas using Babik's method 100 unique reverse primers are required. Therefore, this approach had substantial cost and time benefits.

As this experiment aimed to amplify 210 samples (208 samples and 2 negative controls) it used 15 uniquely tagged forward primers and 14 uniquely tagged reverse primers. This means that a total of 29 primers were used. By comparison, Babik et al's (2009) original method with only forward primers required 211 primers for the same number of samples. Additionally, as half master mixes were created (half the volume of a PCR without DNA) as described previously (under the Degenerate Primer design subheading) for each forward and each reverse primer, the errors in concentration associated with using very small volumes of solutions, as would be the case when adding an individual forward primer for each reaction, were minimized. Furthermore, this approach is beneficial in that it allowed the entire PCR process to be carried out in a single reaction.

	Forward Primer 1	Forward Primer 2	Forward Primer 3	Forward Primer 4	Forward Primer 5
Reverse Primer 1	Sample 1 amplified by F1/R1	Sample 2 amplified by F2/R1	Sample 3 amplified by F3/R1	Sample 4 amplified by F4/R1	Sample 5 amplified by F5/R1
Reverse Primer 2	Sample 6 amplified by F1/R2	Sample 7 amplified by F2/R2	Sample 8 amplified by F3/R2	Sample 9 amplified by F4/R2	Sample 10 amplified by F5/R2
Reverse Primer 3	Sample 11 amplified by F1/R3	Sample 12 amplified by F2/R3	Sample 13 amplified by F3/R3	Sample 14 amplified by F4/R3	Sample 15 amplified by F5/R3
Reverse Primer 4	Sample 16 amplified by F1/R4	Sample 17 amplified by F2/R4	Sample 18 amplified by F3/R4	Sample 19 amplified by F4/R4	Sample 20 amplified by F5/R4

Figure 4.9: The final PCR Plate showing a combination of forward primer half reactions added down columns and reverse primer half reactions added across rows. As shown this results in a large number of unique primer combinations from a relatively small number of unique primers.

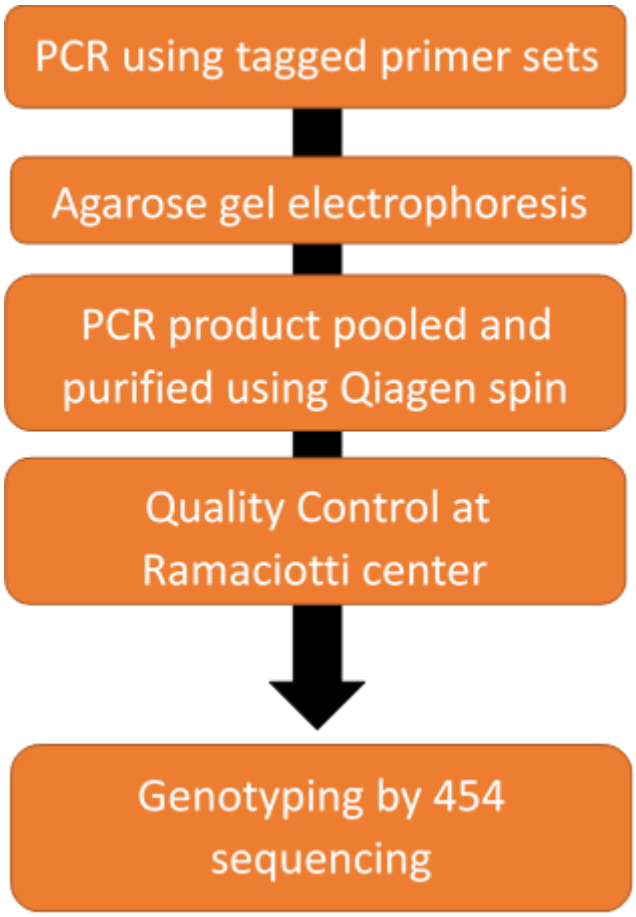


Figure 4.10: A flowchart describing the process of PCR amplification and sample preparation for 454 sequencing.

4.10 Methods: genotyping via 454 sequencing

The PCR is identical to that described earlier in this chapter except that it uses tagged primers (also described earlier in this chapter) and 454 sequencing was carried by the Ramaciotti Centre for Gene Function Analysis (University of New South Wales, Sydney, Australia) on a Roche 454 GS-FLX sequencer with titanium chemistry (Roche Applied Science, Indianapolis, Indiana, USA).

4.11 Results: genotyping via 454 sequencing

52986 sequences were generated with full results available in a publically accessible Google folder (https://drive.google.com/open?id=0B38dGYQ4egWAflNwaWRsWTVoUGczZ1JBenY1R09mQ1QzSUFROTFYelNBaFc5OWxUa1gxNmM)

Sequence quality control and analysis using Mothur

As this was a novel approach to genotyping at the time (less so now), genetics software had not been designed to deal with this method of genotyping. However, several programs had been created for the analysis of next generation sequencing data (e.g. Mothur, Schloss *et al.* 2009, The Genome Analysis Toolkit McKenna *et al.* 2010).The program Mothur (Schloss *et al.* 2009) was commonly used (Huse *et al.* 2010) and thus was chosen for this analysis. The analysis using Mothur is summarized in the figure below.

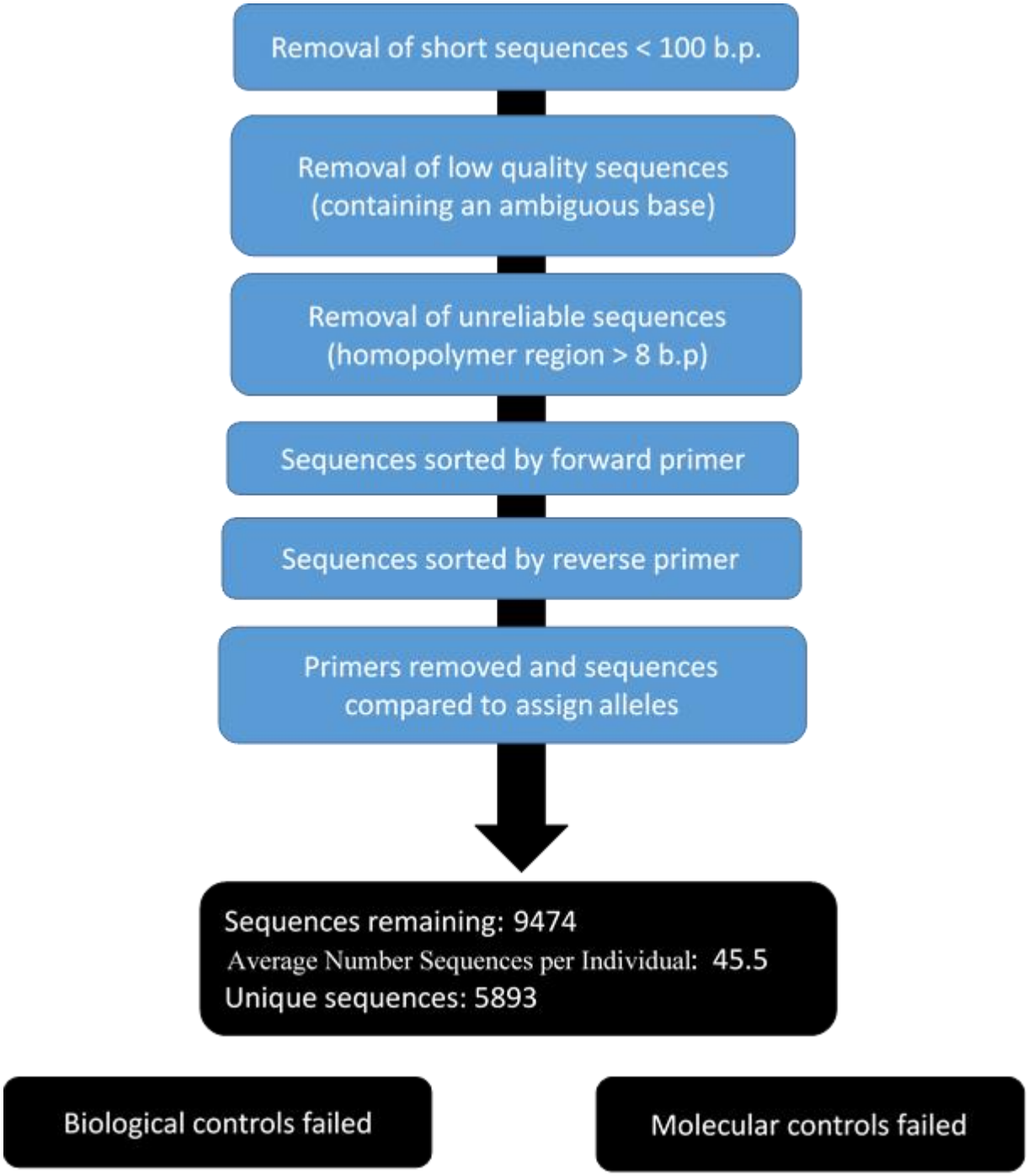


Figure 4.11: A flowchart describing the steps used to check data quality, assign sequences to individuals and genotype using Mothur.

The first step in the analysis was to remove low quality and unreliable sequences so as to exclude them from subsequent analysis. The Mothur software suite was used to merge these output files to produce a single fastA (.fna) file containing sequence information and a single quality (.qual) file containing the base call quality information for each sequence (merge.files script). The Mothur

software suite was used to identify and remove low quality sequences from the dataset. Sequences with an average base call accuracy of below 99%, a length of below 100bp, the presence of one or more ambiguous nucleotides or a homopolymer region greater than 8bp in length were removed using the script *trim.seqs(fasta=input.fna, qfile=input.qual, qaverage=20, minlength=100, maxambig=0, maxhomop=8)*. The high quality sequences that remained could then be sorted into individuals for genotyping.

The Mothur software suite was used to sort sequences to individuals based on their identification tag combination. An oligo (.oligos) file was created containing the primer sequence and the identification tag groups. The script *trim.seqs(fasta=Vince454Reads.trim.fasta, oligos=VinceForward.oligos)* was used to identify sequences containing forward identification tags and sort them in to groups. These sequences were then reversed and the script *trim.seqs(fasta=Vince454Reads.trim.trim.fasta, oligos=VinceReverse.oligos)* was run for each of the forward identification tag groups. This allowed for identification of all unique combinations of forward and reverse identification tags. As an additional quality control, each sequence was searched against the NCBI non-redundant database to eliminate non-MHC sequences from the dataset. A MHC sequence was identified as a sequence where at least one of the highest three matches in NCBI BLAST (Altschul *et al.* 1990) were MHC. Each sequence was then mapped back to the individual based on the combination of forward and reverse identification tags. However, the process though simple in conception was problematic in execution.

4.12 Results: analysis using Mothur

This analysis suggests that 9474 sequences of high quality were generated, with an average of 45.5 assigned per individual. However of 9474 sequences, 4383 were singletons, only appearing in a single copy. Babik *et al.* (2009) excluded such sequences from further analysis as they were deemed likely PCR artefacts. Therefore only 1510 sequences were left as potential alleles, an average of 7.25

per individual or 2.85% of the overall sequencing run, far fewer than anticipated. The reduction of sequences at each step is summarised in the table below. An alternative filtering process was described by Sommer (2013) which first removes shorter than expected sequences before removing sequences with incorrect primer sequences and selecting reads with the expected protein reading frame. However this could not be used due to the quality of the data and high number of singleton sequences. The results of processing, presented below, suggests the method failed to produce enough high quality data to genotype.

Table 4.3 The results of the Mothur analysis of sequences showing the number of sequences remaining after each step of the analysis

Step	Number of Sequences Remaining
All high quality Sequences Generated	52986
Only Sequences that contained forward primer	24723
Only Sequences that also contain reverse primer	19635
Removed low quality reads (107 removed)	19528
Removed ambiguous nucleotides (1918 removed)	17610
Homopolymer region >8 (8 sequences removed)	17602
Removed Short sequences (8075 removed)	9527
Sequences that didn't match MHC removed (53)	9474
Unique Sequences (potential alleles)	5893
Number of Unique sequences with > 1 copy	1510
Potential alleles as percentage of total	2.85%

The alternative explanation is that Mothur may not have provided reliable results. The script used in Mothur did identify specified forward or reverse tags but was not designed to do so in a genotyping context. It identified primer sequences even when they were in the middle of a sequences, and not at the correct position at the beginning or end and thus potentially created false alleles. Furthermore, the program altered the file, by adding a manual line break, after locating a primer tag. This meant that running the script through the same file multiple times changed the number of alleles. Therefore the program may have both created both false alleles and potentially destroyed real alleles. This could either be an artefact of our approach using Mothur in an atypical manner or an error associated with the particular download of Mothur or operating system on which it was

installed. As subsequent downloads on different systems yielded the same result, the decision was made to abandon further analysis in Mothur. The STACKS pipeline (Catchen *et al.* 2013) was also used in an attempt to analyse the data and also failed. To confirm these results an alternative analysis was carried out using Microsoft Excel (Microsoft 2013).

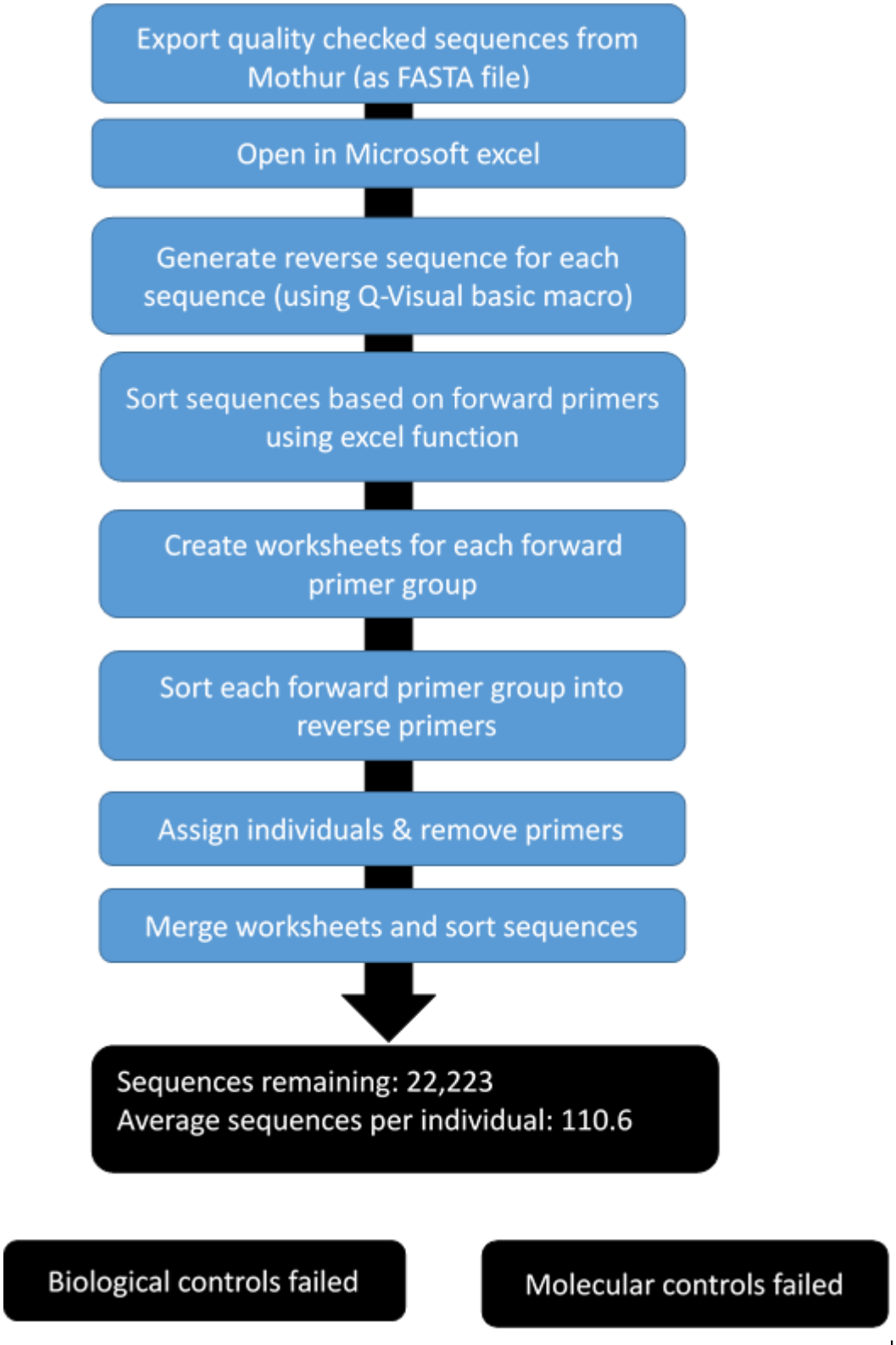


Figure 4.12: A flowchart describing the attempt at genotyping using Microsoft excel

Microsoft Excel was chosen for the next stage of the analysis as the visual nature of the program allowed each step to be checked before proceeding to the next. The individual steps, the commands used and the reasoning behind each step are detailed in the table below. The quality checked sequence file was exported from Mothur as a FASTA file and opened in Excel. A Q-visual basic macro was used to create a reverse sequence for each sequence; the code for this macro is included in appendix I. Excel functions were then used to sort the data into forward primers and then separate sequences belonging to each forward primer into a separate worksheet. Sequences in these worksheets were then sorted by reverse primer, resulting in a worksheet for each individual. Sequences that did not contain both a valid forward and reverse primer were eliminated. Additionally, the forward and reverse primer sequences and tags were designed in such a way that it was impossible for a sequence and the reverse of that sequence to both be recognised, therefore no sequence could be counted twice. Primers were then removed, as degeneracy in primer sequence could create multiple pseudoalleles from a single real allele and sequences named by individual. All sequences were combined into a single worksheet and sorted into alleles based on sequence. In so doing each instance of the same sequence was given the same allele number and thus individuals were genotyped. For ease of understanding alleles were counted in a pivot table. Each step in this program is detailed in the table below.

1 Table 4.4: A description of each step in the genotyping of MHC based on 454 sequence data using
2 Microsoft Excel

Step	Command	Reasoning
File opened in excel	Opened as txt file no special formatting required	This can be opened directly from the fasta file
A reverse copy of each sequence was created with a custom VBA macro	The macro was created with the assistance of a programmer and is detailed in the appendix	454 sequencing occurred randomly from a forward and reverse direction so to identify and analyse sequence from a forward direction both were needed. Differences between forward and reverse primers prevented sequence from being analysed twice
Sequences were searched for forward primers	Filter Command	Sequences without a forward primer were not complete
Sequences with forward primer were searched for a reverse primer	Filter Command	Sequences without both a forward and reverse primer were not complete
A count of sequences was done	Sort command	To determine what proportion of amplification to sequencing reactions were successful
Sequences without both forward and reverse primers were removed	Filter then Delete	These could not be assigned to individuals and thus could not be analysed
A copy of the sequences was created	Copy then Paste	One copy allowed for sequences to be assigned to individuals whilst the other allowed for sequences to be assigned to alleles
Sequences were sorted by forward primers tags	Filter Command then Sort Command	To assign sequences to one of 15 groups of individuals
Each group of forward tags was isolated and then sequences were sorted by reverse tag	Filter Command then Sort Command	Sorting by reverse tag within forward tag groups allowed sequences to be assigned to individuals
A count was performed of each combination of forward and reverse tags	CountIF Command	To count the number of sequences generated per individual
Primers and tags were removed from both copies of the data	Left and Right Command	Once sequences were assigned to individual tags were unnecessary and as primers had degenerate bases they could create false alleles
For the combined dataset duplicates were removed from the list of sequences	Remove Duplicate Command	This counted the number of unique sequences, and thus alleles
The list of unique sequences was used to count the number of times each sequence occurred in each individual	CountIF Command	This determined how many times each allele occurred in an individual sequencing results

A new worksheet was created combining the results of the allele count for each individual	Copy and Paste Command	This showed the allelic composition of the entire sample on a single table
The excel functions that created the data were removed leaving only the data	Paste Special Values Only Command	Excel cannot reliably analyse data that links to several spreadsheets. This command changed the data to simple numbers in excel
A replace function was carried out to remove any allele with a copy number of 1 with a 0	Replace Command	This removed singletons as per Babik et al
A replace function was carried out to replace any allele copy number greater than 0 with a 1	Replace Command	This created a binary matrix of 1 and 0's whereby 0 indicated absence of an allele and 1 indicated presence

4.13 Results: analysis using Microsoft Excel

Analysis with Microsoft Excel could not produce reliable genotypes for MHC. A total of 23,223 sequences were assigned to individuals with an average of 110.6 per individual (St. Dev 29.1). The analysis assigned 116 sequences to the negative control lanes.

Biological and Molecular Controls Evaluation

The experiment included two sets of controls in addition to the controls for sequence quality used by the Ramaciotti centre and those undertaken in Mothur: molecular controls and biological controls. The molecular controls were PCR reactions with no template DNA, which should have produced no results but generated 116 sequences. This was despite the molecular controls not showing amplification when run on an agarose gel after PCR but prior to the 454 process. Further, the sequences included the most common alleles found in the dataset.

The study also included a biological control in the form of two family groups of *E. cunninghami* where parentage had previously been determined using microsatellites. Numerous alleles in these family groups were not inherited in a Mendelian manner, where the majority of alleles present in the offspring in both families were not present in either parent. This could not be a case of mistaken paternity as paternity had been previously established using genetic methods (microsatellite

markers in Stow & Sunnucks 2004a). This is presumably due to the low copy number per sequences and large number of singleton sequences (30451 unique sequences in 52986 sequences total). Meaning that either the majority of sequences in an individual were missed by the method or that the majority of sequences assigned to an individual were created by molecular misadventure and not real genotypes. The molecular and biological controls suggested that further analysis would be meaningless as the data were unreliable.

4.14 Discussion

While this methodology did not create a method for reliably genotyping MHC, it did bring to light a number of issues that must be considered when building such methodology. These issues are discussed below, as well as the insights from this work for future MHC genotyping attempts.

Traditional approaches to genotyping do not work for MHC

In this study traditional methods of genotyping MHC were unable to produce data for a conservation genetics scenario. Traditional approaches developed to visualize single loci failed to produce reliable results for MHC amplified by degenerate primers. Visualisation using t-RFLP failed, although RFLP approaches have previously been used in some species to genotype MHC (e.g. Uni *et al.* 1993), likely because the short sequence of MHC amplified did not contain variability at restriction sites. Visualisation using SSCP also failed, presumably because a large number of alleles were amplified which, because of the increased probability of heteroduplexes and homoduplexes, created large numbers of false alleles (Sunnucks *et al.* 2000) which resulted in unreliable and unclear electophoretograms. The cloning and sequencing approach produced positive results, albeit after multiple attempts, however the costs and time associated with this approach make it untenable for large scale sequencing of multiple individuals and multiple loci.

It is noteworthy that these techniques have been successfully used on other species to genotype MHC though not in conjunction with degenerate primers and next generation sequencing. Although

this failure could be attributed to researcher error or molecular mishap they are more likely to be due to the ambitious approach. Traditional approaches to genotyping were not designed for multi-locus primers and we know of no methods that have successfully used traditional genotyping methods on a gene complex without removing loci from the analysis. For this reason genotyping by next generation sequencing was attempted, and although it ultimately failed, some encouraging data were produced.

2. Next generation sequencing methods have the capacity to genotype MHC

Despite not producing reliable genotypes, the results of the study show that next generation sequencing is able to produce large numbers of sequences for MHC genotyping. This study generated approximately 52,000 sequences (which should have equated to approximately 250 per individual). However, unlike Babik *et al.* (2009), sequences could not be converted into genotypes. Nevertheless, these data support next generation sequencing as a methodology with the potential to produce tagged sequences that can facilitate rapid genotyping. Realising this potential requires that several potential issues are addressed.

The molecular approach may have produced the poor results. Degenerate primers were used in an effort to create a method that could genotype multiple sequences. However, there were persistent, long term, intractable issues with obtaining uncontaminated PCR product. Over the course of six months we changed reagents and then changed brands of reagents, going to an all-in-one PCR master mix (Promega Gotaq), in order to remove this contamination. Furthermore, a negative control was sequenced in order to find and eliminate any sequences resulting from contamination, or PCR artefacts due to degenerate primers. Unfortunately, there were multiple sequences in the negative control and some were similar to the most common sequences. This suggests there may be a flaw in the design of the degenerate primers, although these same primers produced positive results in cloning and sequencing (previously undertaken). This suggests the primers are useful,

albeit problematic and prone to contamination. Further complications may have arisen due to the choice of next generation sequencing method.

In addition to molecular misfortune the sequencing method chosen for the genotyping may have contributed to the failure of the study. 454 sequencing, despite being successfully used in molecular ecology (e.g. Meyer *et al.* 2008, Jumpponen *et al.* 2010) is now planned for obsolescence (Nederbragt 2014). Partially this is because 454 sequencing has been superseded by newer platforms (van Dijk *et al.* 2014). However, there is a positionally dependent error rate which is now known to occur in 454 sequencing (Dohm *et al.* 2008, Gilles *et al.* 2011). This was unknown at the time these experiments were being undertaken, and these errors may have contributed to our poor results. Although the overall error rate in 454 data is low, approximately 0.5% (Rothberg & Leamon 2008), the error is contextually dependent with rates of up to 20% reported largely in short homopolymer regions (Prabakaran *et al.* 2011). Additionally, other factors, including position in the sequencing plate, can increase errors to above 50% (Gilles *et al.* 2011). Furthermore, 454 sequence quality decreases at the end of the sequence (Dohm *et al.* 2008, Fichot & Norman 2013) where the primer tag is located. This is not typically a problem in genome sequencing work where 454 sequencing has been used because multiple overlapping sequences are generated for each genome section and very few would contain the error. However, for genotyping-by-sequencing, where all sequences are the same length, this flaw has serious repercussions. Errors occur in the same places, creating numerous false alleles at particular bases. Alternative approaches e.g. MiSeq have not reported this problem (Loman *et al.* 2012). That there are other MHC data sets which have used 454 methodologies which did not yield useful data, further suggests that this method may have been an issue (Belov 2013 pers. comm.). Additionally, this study highlights some general limitations around next generation platforms.

Limitations of genotyping by next generation sequencing and workarounds

This study highlighted the limitations of next generation methods for genotyping by sequencing in conservation. Next generation sequencing methods use parallel sequencing, where all sequences are generated in a single run which makes them faster and cheaper per sequence than traditional Sanger sequencing (Schuster 2008, van Dijk *et al.* 2014). However, it also means that optimisation is more difficult, unlike Sanger based sequencing where the results of one sequencing reaction can be used to improve the next. In this study, even though samples passed quality control they could not produce useful data. As next generation sequencing runs are relatively expensive it is difficult to run multiple attempts in order to optimise procedure. These factors need to be considered when deciding on a method for genotyping. Furthermore, as optimisation problems can lead to erroneous data, stringent controls, such as the biological controls used in this study, must be in place.

The results of this study also support the use of multiple levels of control in experiments involving next generation sequencing. The sequence quality control, using Mothur, removed sequences with uncertain bases but could not remove errors introduced by either the molecular approach or the sequencing platform. Although a large number of sequences (17,610) were considered high quality, additional biological controls were able to show that the method did not produce useful data. The biological controls, in this case family groups identified by previous work (Stow *et al.* 2004b), showed that alleles were not inherited in a Mendelian manner and thus the data were unreliable. Therefore, this study echoes previous work (e.g. Babik *et al.* 2009, Prabakaran *et al.* 2011) in demonstrating the importance of molecular and biological controls, as well as the *in silica* controls typically used in studies involving next generation sequencing. The difficulties in understanding the data produced and identifying the errors in the process, also highlight the challenges associated with analysis of this kind of data.

Genotyping by next generation sequencing is a challenging process that requires repurposing of sequence analysis tools. This study used Mothur, a tool that is often used in sequence analysis, to

analyse genotyping by next generation sequencing data. Ultimately, it was unsuitable for the purpose, as the 'black box' nature of the program, where it is not immediately clear what the script is doing, made it difficult to understand where problems may be arising. Further, it was discovered that the searching function, used to identify tagged primers, changed the sequence file. In sequencing operations this is not an issue as sequences are usually still generated by a shotgun approach where overlapping sequences are generated then assembled (Shendure & Jir 2008). However, in our approach the same program had the potential to create false alleles and eliminate true alleles. Though more suitable programs have since been made available (e.g. STACKS, Catchen *et al.* 2011), these results highlight the issues with repurposing programs and using analysis tools without being able to identify how each action is carried out. Our solution was to re-process the data using Microsoft Excel, which was possible due to the modest number of sequences generated (for a next generation platform). Although this will not be applicable to all projects, our results suggest it is worth consideration, at least for a subset of the data. Excel allowed the inputs and outputs of each step to be visualised, and allowed for the same sorting to be undertaken without any knowledge of scripting.

4. Future directions in genotyping MHC for conservation

Understanding the diversity present in the major histocompatibility complex is still important for many conservation efforts so effort to genotype the gene complex remains important. Over the longer term, genotyping by whole genome sequencing is a real possibility (e.g. Gudbjartsson *et al.* 2015), even in conservation. This has a number of advantages including eliminating errors stemming from PCR and primer creation. Given the exponential increase in sequence generation and the decreases in cost (Hayden 2014), this is the direction all genotyping is predicted to move towards. Additionally, it will generate information on all functional genes to allow for the first time a picture of all genetic diversity in an individual and population. However, genotyping by genome sequencing will not be available for some time. In order to create a short – medium term solution, the timeframes at which conservation must act in order to be effective, next generation sequencing

remains the most promising tool. Future work could benefit from understanding the challenges of this project and planning to avoid or minimize the difficulties that we encountered.

5 Investigating disruptive selection due to land clearing in *E.cunninghami*

Disruptive selection is a powerful force in shaping the evolution of populations. It is of particular interest to conservation genetics because disruptive selection can, even in the presence of gene flow, lead to local adaptation. Further, concerns around local adaptation constrain outcrossing, which is necessary for the survival of many species. This study used a population genomics technique, DArTseq, to investigate the effects of disruptive selection due to recent land clearing in Egernia cunninghami. The discovery of disruptive selection in 5 loci using two different methods to discover loci under selection over small time and geographic scales caused by anthropogenic influence is discussed as well the implications of both the results and technology for conservation programs.

5.1 Introduction

The sequencing revolution has created a new field within genetics, population genomics. Over the last decade and the course of this PhD, advances in sequencing methods have exponentially increased the quantity of sequence data generated (Margulies *et al.* 2005) while decreasing the cost per sequence by orders of magnitude (Hayden 2014, also see chapter 1). This has meant that geneticists can not only produce complete genomic sequences for species (e.g. Li *et al.* 2010) and individuals (e.g. Wheeler *et al.* 2008) but also for populations (e.g. Liti *et al.* 2009). By sequencing entire genomes (e.g. Liti *et al.* 2009), or large sections of genomes (e.g. Hohenlohe *et al.* 2010), genetic diversity in populations is explored with unprecedented resolution.

Population genomics approaches have several advantages over traditional population genetics methods for conservation. Genomics approaches, unlike previous methodologies based on genetic markers, genotype thousands of loci, both functional and neutral, in a single step (e.g. Sansaloni *et al.* 2011). This increases the speed of processing and, thus, reduces the time required to make informed conservation decisions. Additionally, as the whole genome is sampled, genomics approaches remove the biases associated with choosing a genetic marker which can potentially affect conservation decision making (Wan *et al.* 2004). Furthermore, population genomics does not require molecular optimisation for individual species (e.g. Miller *et al.* 2006, Sansaloni *et al.* 2011), which is advantageous in non-model species where established genetic frameworks are not available.

The advent of population genomics has enabled researchers to ask new questions, and to re-examine older questions within conservation genetics (Ouborg *et al.* 2010, Narum *et al.* 2013). Consequently, population genomics has assisted in setting conservation priorities. For example, after the giant panda (*Ailuropoda melanoleuca*) genome was sequenced (Li *et al.* 2010), subsequent population genomics research by Zhao *et al.* (2013) suggested the existence of three ecologically significant units (ESU's) with local adaptations. Previously, researchers had divided the giant panda

into two sub-species (Wan *et al.* 2003). However, population genomics approaches have yet to be fully tested with less charismatic species and other conservation questions. This study aims to use population genomic techniques to engage with questions around disruptive selection.

What is Disruptive Selection?

First studied in the 1950's (Levene 1953, Mather 1955), in part because of its role in sympatric speciation (Maynard Smith 1962), disruptive selection favours both extremes of a trait simultaneously whilst selecting against the intermediate value. This results in a population where the intermediate version of a trait has lower frequency than either extreme, and as such is unique among types of selection (Figure 5.1).

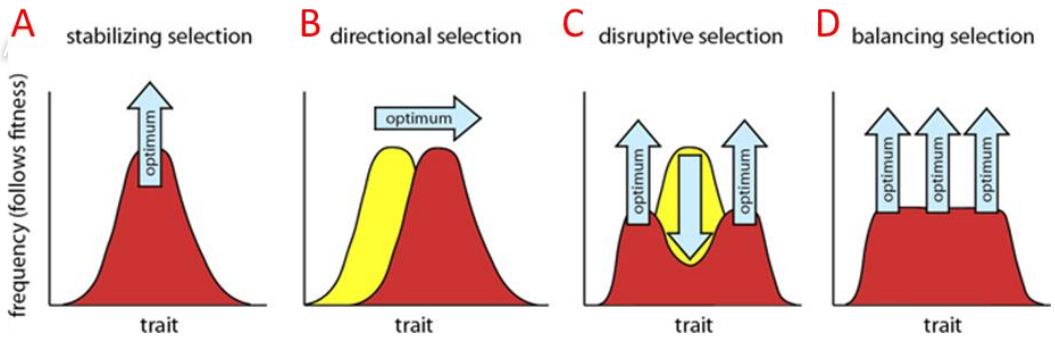


Figure 5.1 Types of selection Loewe (2008). 5.1 A describes stabilizing selection where selection acts against extreme variants of a trait increasing the frequency of the median variant. 5.1 B describes directional selection where selection acts for an extreme variant causing the distribution of variants in the population to shift. 5.1 C describes disruptive selection whereby selection acts to favour the both extreme variants of a trait at the expense of the intermediate version; in this case the distribution of variants changes but the mean variant of the trait is the same. Finally, 5.1 D describes balancing selection where selection acts to maintain a uniform distribution of variants at a trait. The remaining type of selection, frequency dependent selection is explained in Figure 5.2.

Disruptive selectionchanges the shape of the distribution without altering the mean value. Over time more individuals with extreme values at a trait are produced and fewer with intermediate values. Unlike directional selection, disruptive selection does not change the mean value of a trait over time,

but does change the shape of the distribution. However, in natural populations the various types of selection are more difficult to disentangle.

The difference between disruptive selection and directional selection can be based on the definition of a population, and the terms are often not clearly delineated (e.g. Lenormand 2002, Albertson *et al.* 2003). For example, if a small geographic area is defined as a population, selection favouring a skin pigment could be classified as directional selection. However, if the population definition was expanded to a larger area, and a different pigment was favoured in a subset of that environment, pigmentation would now be under disruptive selection. Additionally, negative frequency dependent selection, a type of disruptive selection, can produce the same effects as balancing selection in that it preserves genetic diversity in the population over time. Rueffler *et al.* (2006) presents an example of this, summarized by figure 5.2, whereby a herbivore population exploits seeds of varying sizes. As the population shifts to exploit the most common, intermediate size seeds, they become rarer. This will then create an advantage for extremes that focus on extremely large or extremely small seeds. However, this benefit is also transient, as the population shifts to take advantage of these different resources. Though empirical evidence of this is rare (Rueffler *et al.* 2006), intraspecific competition has been shown to drive disruptive selection in three spined sticklebacks (*Gasterosteus aculeatus*) (Bolnick 2004) demonstrating at least the first step in this process. The final result is that multiple variants are maintained in the population over time.

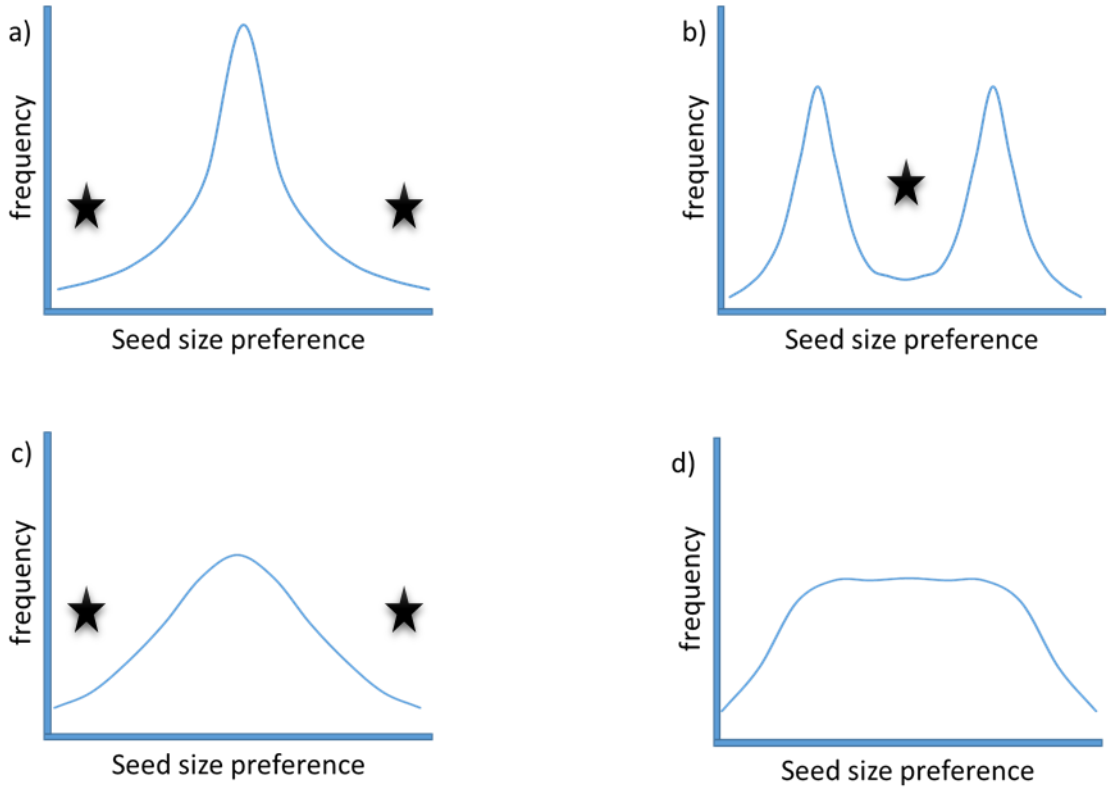


Figure 5.2: Negative frequency dependent selection. The figures a –d depict the changes in seed size in a population over time. The blue line indicates the distribution of seed size preference in a herbivore in the environment and the stars indicate the preferences with the highest fitness at that time. Here disruptive selection may result in a uniform distribution as in figure d.

1

Why is disruptive selection important to conservation?

Outbreeding, or genetic rescue, is the only practicable method to restore genetic diversity to fragmented populations and should be a key component of conservation programs (Frankham 2010, 2015). As discussed previously (Chapters 1 & 4), genetic rescue reintroduces genetic variability, reducing inbreeding depression and restoring population viability (Frankham *et al.* 2002, Weeks *et al.* 2011). The benefits of outcrossing have been demonstrated empirically in *Drosophila melanogaster* (Spielman & Frankham 1992), and in plant populations (Frankham *et al.* 2010). Further, the deliberate introgression between Texas puma (*Puma concolor stanleyana*) and Florida panther (*Puma concolor coryi*) increased survivability in both first and second generation offspring of the Florida panther (Hostetler *et al.* 2010). However, despite its advantages genetic rescue is rarely practised. A review by Frankham *et al.* (2011), which highlighted the importance of outcrossing, could only identify 19 occasions where outcrossing was used as part of conservation efforts.

The lack of outcrossing within conservation programs is often attributed to fear of outbreeding depression (Weeks *et al.* 2011). Outbreeding depression occurs when individuals from different, locally adapted, populations are crossed, and the mating destroys locally adapted gene combinations in the offspring by introducing alleles for traits unfavourable in the local environment (modelled in Edmands & Timmerman 2002). Spectacular failures such as the Tatra mountain ibex (*Capra ibex ibex*), which became extinct after managed outcrossing produced maladapted hybrids (Rhymer & Simberloff 1996), graphically illustrate the consequences of outbreeding depression. Further, the interplay between the loss of fitness associated with outbreeding depression and the increased fitness from the alleviation of inbreeding depression is difficult to predict (Tallmon *et al.* 2004). However, the risks of outbreeding depression may be negligible compared to the loss of fitness due to genetic factors stemming from a small population (Frankham 2010, Frankham 2015). Nevertheless, while outcrossing is often discussed (e.g. Frankham *et al.* 2001, Edmans 2006, Hedrick & Friedrickson 2010, Waller 2015), it is rarely implemented (Frankham 2015). Therefore, to facilitate more frequent and successful outcrossing our understanding of local adaptation must improve.

Disruptive selection is important to conservation because it is involved in local adaptation. Local adaptation is caused by disruptive selection acting to favour different alleles in different areas over the range of species. This is described in Figure 5.3 where soot from an industrial factory selects for a dark colour morph . Local adaptation has been observed at both the macrogeographic scales, for example embryonic development speed is different in common frog (*Rana temporaria*) associated with a 1600km latitudinal gradient (Laugen *et al.* 2003) and microgeographic scales, for example wood frog (*Rana sylvatica*) tadpoles have local adaptation associated with predator presence at a scale of 0.3-8 km (Relyea 2002). Additionally, local adaptation occurs between seasonal populations within the same habitat (Taylor 1991). Furthermore, local adaptation can occur rapidly (Colautti & Barrett 2013) and occur in response to anthropogenic influences (e.g. Råsånen *et al.* 2003).

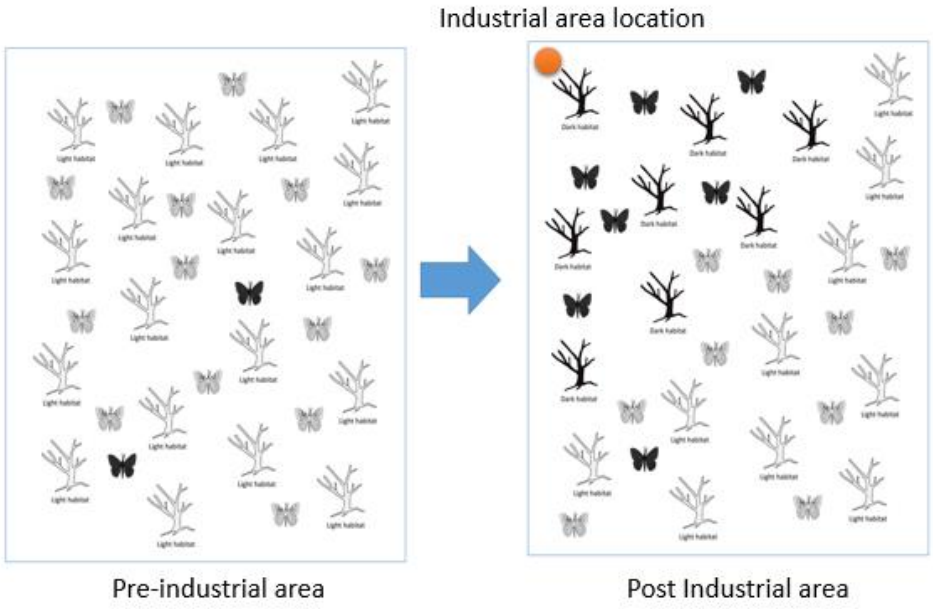


Figure 5.3 Disruptive selection causing local adaptation in the presence of gene flow. If soot from an industrial area turns the trees in one part of a forest black (as in the top left corner in the post industrial area) there is a selective pressure for black moths in that area whereas white moths are favoured in other areas where white trees predominate. If these pressures are strong enough the local adaptation will persist even when migration can occur between areas and so the moths near the black trees will remain black (heavily modified from Richardson *et al.* 2014).

Therefore, choosing a population to outcross to, without first understanding the distribution of genetic diversity and investigating disruptive selection, is a risky proposition. Unfortunately,

Hendrick & Fredrickson (2010) found that conservation managers may be too conservative in their decision making. Therefore, without an understanding of local adaptation, which relies on an understanding of disruptive selection, outbreeding is rarely implemented.

How is disruptive selection measured?

Local adaptation cannot be investigated using neutral markers as it may persist despite migration between populations (henceforth gene flow). In theory genetic differentiation between populations, which leads to local adaptation, is stopped and reversed by a single (Speith 1974), or at worst 1-10 (Mills & Allendorf 1996), effective migrants per generation. Further, this gene flow is usually measured with neutral markers (e.g. Waits *et al.* 2000, Spong *et al.* 2002, Henry *et al.* 2009). However, when disruptive selection removes alleles in a local population faster than they are added by migration local adaptation will persist. This will be undetectable by methods that focus on tracking the flow of neutral genetic markers through a population. Disruptive selection has been shown to produce local adaption even where gene flow occurs (e.g. Cousyn *et al.* 2001, Hoekstra *et al.* 2005, Gonzalo-Turpin & Hazard 2009). Hoekstra *et al.* (2005) studied the Rock pocket mouse (*Chaeotdipus intermedius),* finding colour variation was determined by local adaptation despite high levels of gene flow between neighbouring light and dark coloured mice. Additionally, this response has been shown to be rapid, even in wild populations (Cousyn *et al.* 2001). Therefore, functional genes (e.g. Dionne *et al.* 2007, Ekblom *et al.* 2007) and population genomics techniques (Stapley *et al.* 2010, Savolainen *et al.* 2013) have been used to examine local adaptation.

There have been studies that use functional genes e.g. Dionee *et al.* (2007), who investigated MHC diversity in Atlantic Salmon (*Salmo salar*) and found that diversity increased on a temperature gradiant. These techniquesrequire target genes to have been identified and primers for these genes to have been developed. This approach is not ideal for conservation where the genes responding to selection may not be known or there may not be sufficient genomic information available to design primers. As an alternative, Stapely *et al.* (2010) pointed to next generation sequencing technologies

as methods that would enable the study of adaptation, including local adaptation. This has been applied in conservation genetics, for example Kjeldsen *et al.* (2015) used population genomics to find 10 putative loci associated with local adaptation in Koalas (*Phascolarctos cinereus*) across 8 populations on the east coast of Australia. However, the methods used to identify loci under selection in population genomic data remain controversial.

There is no single best method to identify loci under selection. Rellstab *et al.* (2015) identify a number of statistical methods for the identification of loci under selection. The review identifies methods based on logistic regression, matrix correlation and mixed effect model. In methods based on logistic regression, e.g. SAM (Joost *et al.* 2007), the environmental effect is the predictor variable and the presence or absence of the allele the response variable. Further, Rellstab *et al.* (2015) also reviews outlier tests based on Fst such as BAYESCAN (Foll & Gaggiotti 2008). However, the review does not recommend a single best statistical approach. Furthermore, the results of the programs using the same statistical framework, e.g. logistic regression, can vary.

The effectiveness of methods for discerning loci under selection varies with the characteristics of the analytical program and the populations chosen. In testing methods to identify loci under selection against simulations, Narum & Hess (2011) found that the program BAYESCAN (Foll & Gaggiotti 2008) returned the lowest rate of false positives. Conversely, Lotterhos & Whitlock (2014) found that BAYESCAN had the highest false positive rate. This increased false positive rate is attributed to the population history, in our case the recent landclearing splittling one population into two areas with different environemnts, violating the assumptions of the BAYESCAN software (Lotterhos & Whitlock 2014). The programs FLK (Bonhomme *et al.* 2010) and Bayenv2 (Günther & Coop 2013) are recommended over BAYESCAN for populations with recent divergence (Lotterhos & Whitlock 2014). Of these two, Bayenv is only predicted to perform better under isolation by distance or more complex population history models (Günther & Coop 2013). Therefore, determining that a particular

1 locus is responsible for local adaptation is difficult and evidence of selection is *prima facie* only.
2 These factors affect the choice of study organism in which to investigate disruptive selection

3 **Cunningham Skink as a model species for investigating disruptive selection associated with land**
4 **clearing**

5 Here we investigate whether land clearing has caused disruptive selection in a population of
6 Cunningham's skink (*Egernia cunninghami*). Cunningham's skink is an Australian native scincid lizard
7 whose population persists as isolated fragments after land clearing (e.g. Stow *et al.* 2001). As
8 Cunningham's skink is relatively long lived, up to 20 years in captivity (A. Stow Pers. Comm.), and the
9 time since land clearing is known for many areas including the study site (e.g. approximately 100
10 years in Bathurst), the effect of fragmentation can be studied on multiple temporal and spatial
11 scales.

12 A genetic framework already exists for the wild populations used in this study (Stow *et al.* 2001,
13 Stow & Sunnucks 2004a, Stow & Sunnucks 2004b, Stow & Briscoe 2005). Stow et al (2001) used
14 microsatellites to determine high pairwise relatedness in lizards within rocky retreat sites and to
15 show lower dispersal in cleared than natural areas. In 2004 Stow and Sunnucks (Stow & Sunnucks
16 2004a, Stow & Sunnucks 2004b) used microsatellites to examine the mating system in the species,
17 showing inbreeding avoidance and high mate fidelity as well as site fidelity. Stow & Briscoe (2005)
18 used 10 microsatellite loci to determine that genetic diversity was lower in populations fragmented
19 by land clearing activities than in populations from unmodified areas in the same populations this
20 study examines. However, due to the limitations of the genetic markers available, they were unable
21 to investigate disruptive selection; this study aims to fill that gap in the knowledge.

22 There are two significant possible outcomes of this study:

23 1. If significant evidence is found for disruptive selection, this would suggest that
24 fragmentation, a common threat to biodiversity (Gibbons *et al.* 2009), may cause adaption in

species. This may be an important consideration for outcrossing within conservation programs.

2. If no evidence for disruptive selection is found it would suggest that, though previous studies on this population found a loss of neutral genetic diversity, functional diversity is unchanged. This would suggest that fragmented populations may accumulate genetic damage more slowly than neutral markers would suggest. Alternatively, this may be caused by a lack of power in the experimental design.

5.2 Methods

Sampling Strategy and DNA extraction

Tissue samples from Cunningham's skinks (*Egernia cunninghami*) were taken from paired populations at a single locality in Bathurst (33° 27' S, 149° 24' E), Central Tablelands of New South Wales, Australia, shown in figure 5.4, as part of a larger study conducted by Stow *et al*. At this site part of the adjacent areas of the landscape had been cleared for agricultural use and part had been left in its natural state. Therefore, what had been a single population of *E.cunninghami* is now two populations at a single location, a population living in the cleared area where 18 individuals were sampled and a population living in the naturally vegetated area where 9 individuals were sampled. This site was chosen as a single location with recent land-clearing largely precluding the possibility of differential adaptation before the fragementation occurred. Though this sample was small, a significant history of genetic work on this population already exists (Stow *et al.* 2001, Stow & Sunnucks 2004a, Stow & Sunnucks 2004b, Stow & Briscoe 2005, Stow & Briscoe 2005). Tissue samples were taken as previously described (Chapter 4, Stow *et al.* 2001) and DNA was extracted using a GenCatch TM spin-column (Epoch Life Sciences) as per the manufacturer's instructions.

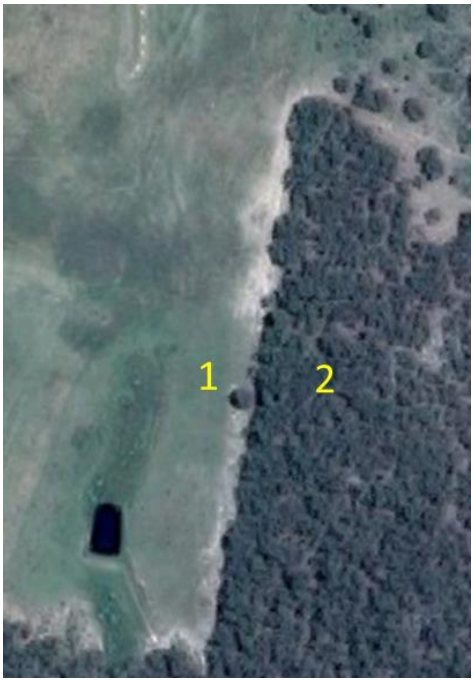


Figure 5.4: A map of the sampling area showing where populations were sampled. In the figure the cleared is labelled 1 and natural area is labelled 2.

Sequencing and SNP genotyping

Genotyping was carried out as part of a larger project (Oforio & Stow In prep. Sequencing, genotyping and SNP discovery and genotyping were performed at Diversity Arrays Technology Pty. Ltd. (Canberra, Australia), using standard DartSeqTM protocol. Firstly, the quality and quantity of genome DNA were checked visually on 0.8% agarose gel using GelRed (Biotium). Then 100 ng of DNA from each sample was using a combination of PstI and SphI restriction enzymes and ligated with unique barcoded adapters (P1 adapter) at the PstI /SphI digestion sites. Digested and ligated samples were cleaned using a spin-column (Qiagen) and then amplified using PCR with barcode and adapter specific primers (Protocol available from Diversity Array Technologies) and checked for size (<200 b.p.) on an 0.8% agarose gel using GelRed (Biotium). Equimolar amounts of all amplified samples were pooled and then denatured using NaOH and then hybridized to the flow cell.

The library was sequenced with Illumina HiSeq2500 (Illumina) for 77 cycles which resulted in fragments of 77 bp long. To assess the reproducibility of SNPs calls, ≥15% random replicates were carried through the protocol pipeline.

Preliminary Sequence analysis

Preliminary analysis was carried out as part of a larger project which included populations not used in this chapter (Oforio et al 2017) with previously established methods (Truszewski 2015). Once sequencing was completed, raw sequenced data were converted to .fastq files using the Illumina HiSeq2500 software. Sequences were stripped of barcodes, cleaned and filtered to include only those with a Phred score ≥ 25. SNPs were identified and called following standard protocols in DArTSoft14TM (Diversity Arrays Technology). All reads then underwent quality assessment, during which they were checked for foreign DNA sequence from bacteria and viruses using both GenBank and the in-house DArT database (Diversity Arrays Technology Pty. Ltd., Canberra), and any with PHRED (Ewing et al. 1998) quality scores 95% and read depth >5 were retained.Then all monomorphic sequence clusters were removed. To ensure the quality of individual samples all duplicate markers and those with minor allele frequencies < 5% were removed as were markers with average read depth < 10 or > 45, individuals genotyped at < 100% call rate (CR: proportion of genotyped SNPs), and individuals genotyped at < 94% reproducibility as per a previous project (Truszewski 2015). DNA sequences and statistics (i.e., call rate, polymorphic information content, heterozygosity, read depth and reproducibility for all loci and individuals) will be accessible from the Dryad Digital Repository (http://datadryad.org/) and the Diversity Array Technology Pty. Ltd., Canberra, Australia (Report-DEgs14-1547).

Preliminary data analysis

The DaRTseq process results in 4724 polymorphic SNP markers in the *Egernia cunninghami* genome. This study follows the recommendations of Rellstab *et al.* (2015) in using multiple methods that match both the available data and what is known about the populations. Microsoft Excel was used to determine the number of monomorphic alleles in each population that were not monomorphic in the other population. GenAlEx 6.5 (Peakall & Smouse 2012) was then used to ascertain the level of Heterozygosity present and Fst for each population. The frequency of the first allele for each locus

was then calculated using Microsoft Excel so that analysis for disruptive selection could proceed using FLK (Bonhomme *et al.* 2010). Further, the presence or absence of each allele was checked using Microsoft Excel so that analysis for disruptive selection could proceed using Samβada (Joost *et al.* 2007).

Investigation for disruptive selection using FLK in R

Testing for loci under disruptive selection was conducted using the program FLK (Bonhomme *et al.* 2010) according to the recommended methods with the default parameters. This program was chosen as it has been shown to be highly accurate when population divergence is recent (Lotterhos & Whitlock 2014). The program works via comparing the pattern of differences between populations against expectations under neutral genetic theory (Bonhomme *et al.* 2010). This required the use of R (Ihaka & Gentleman 1996), the R-package APE (Paradis *et al.* 2015), the R-package NumPY (Van der Walt *et al.* 2011) and the R-package simuPOP (Peng & Kimmel 2005). A sample of 27 individuals from a different *E.cunninghami* population was used as an outgroup, as stipulated by Bohomme *et al.* (2010).

Investigation for disruptive selection using SAMβADA

A second method of testing for loci under disruptive selection was carried out using the program Samβada (Joost *et al.* 2007). This program detects loci under selection based on a spatial analysis which uses GIS data (Joost *et al.* 2007). As is the default in the program, a univariate model was specified. However, Samβada was developed before population genomics studies became mainstream and the default options may not be optimal for this type of work. Therefore, the Bonferroni correction was disabled to prevent the small number of samples and the large number of comparisons resulting in alpha values that were unrealistically low. Instead a graph was produced of the Efron coefficient (Efron 1981) a pseudo r-squared estimator which describes the amount of variation in the result of the logistic regression (presence or absence of the allele) that is explained

by the predictor variable (the environment) of all alleles. The alleles are ordered by Wald Score, the results of a test for the significance of a predictable variable in a logistic regression. In order to have a biological meaning both the Efron coefficient and the Wald score must be high enough to be significant at a 5% level with the same parameters as the data (Stucki *et al.* 2016)..

Final analysis of loci under selection

Loci found to be under selection by both programs were tested for linkage disequilibrium using GenePop (Rousette 2008) and then the sequences for each loci searched on the BLAST database (Altschul *et al.* 1990).

5.3 Results

DaRT successfully sequenced and genotyped 27 individuals from two populations of *Egernia cunninghami,* full results of this are available in a publically accessible google folder https://drive.google.com/open?id=0B38dGYQ4egWAflNwaWRsWTVoUGczZ1JBenY1R09mQ1QzSUFROTFYelNBaFc5OWxUa1gxNmM). Furthermore, both analytical approaches, Samβada and FLK found loci under disruptive selection.

4274 loci were genotyped, of which 3039 were polymorphic for at least one of the two populations of interest (loci that were polymorphic only in the outgroup were still included in the FLK analysis). The results for each population are summarized in the table below and overall genetic differentiation was low with Fst = 0.023 (S.E = 0.0004).

Table 5.1 Summary statistics for genetic diversity of populations of Egernia cunninghami in cleared and natural environments in Bathurst Central NSW

	n	Avg. He	Number of fixed loci
Bathurst Natural	9	0.121	649
Bathurst Cleared	18	0.132	105

The analysis using FLK suggested candidate loci under disruptive selection at both the 95% and 99% confidence level. As shown in the figure below, 70 loci were identified outside of the 95% confidence envelope and 12 loci were identified outside of the 99% confidence envelope. Under the null scenario all loci should fall inside the confidence window envelope.

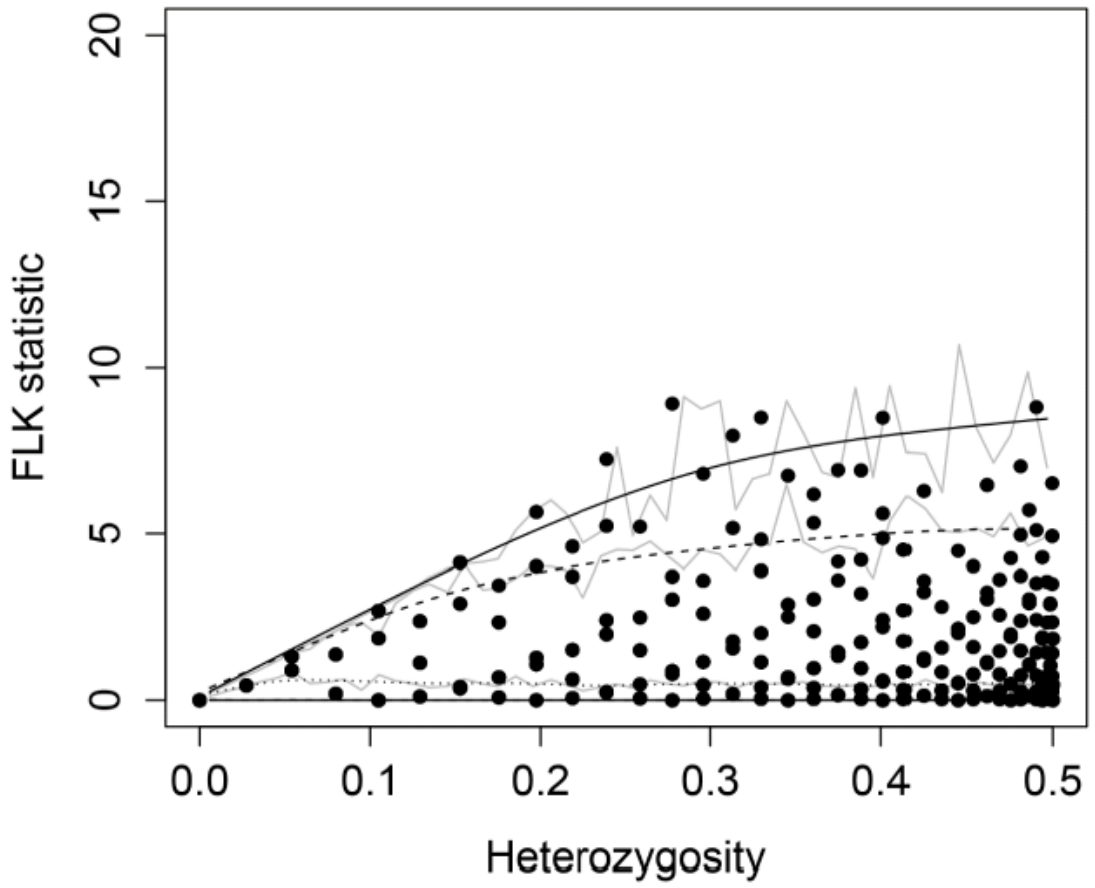


Figure 5.4 The results of an FLK analysis for loci under different selection in *E.cunninghami* cleared and natural populations in Bathurst. The solid lines indicate the envelope for 99% probability of no selective differences whereas the dotted lines indicate the envelope for 95% probability of no selective differences. The jagged lines represent a spline on the envelope data to take into account the simulation process and are considered more informative than the curvedlines (Bonhomme *et al.* 2010). Loci significant at the 95% level are shown in green and those significant at the 99% level are shown in orange.

The analysis using Samβada is also consistent with the action of selection on few loci. As shown in figure 5.6, the largest Efron coefficient was 0.48, six alleles are significant at the 1% level and 82 are

significant at the 5% level. A small number of non-significant alleles (shown on the right of the figure below) also have relatively large Efron coefficient but low Wald Scores.

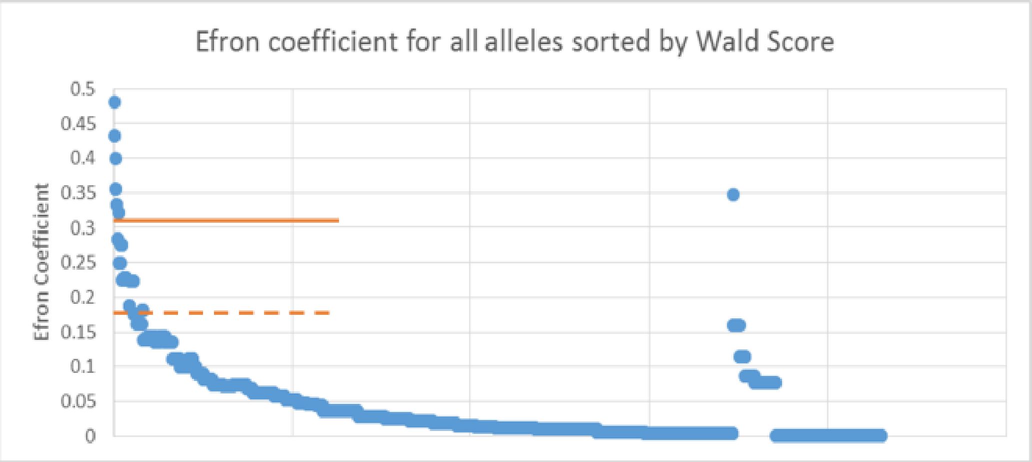


Figure 5.5 A graph, using Microsoft Excel, of the outputs from Samβada. The alleles have been sorted by Wald Score and then the Efron of each is graphed. Alleles above the full line on the left side of the graph are significant at the 1% level, alleles above the dotted line on the left side of the graph are significant at the 5% level.

As summarised in Table 5.2 below, a total of 5 loci were identified by both programs as being under increased likelihood of selection at the 5% level. Only these loci were considered for further analysis because each of the methods used can create false positives (Rellstab *et al.* 2015). Therefore eliminating all loci only identified by a single program is a powerful, if conservative, approach to remove these false positives. An analysis using Genepop revealed no linkage disequilbirium between pairs of loci (minimum p-value 0.28). These loci were then searched in the NCBI BLAST database (Altschul *et al.* 1990) and the results summarized in table 5.3.

Table 5.2: A description of the loci identified using both methods.

Method	Total Loci Identifed (α = 0.05)	Loci identified by both methods (α = 0.05)	Unique loci identified (α = 0.05)	Total loci identified (α = 0.01)	Loci identified by both methods (α = 0.01)	Unique loci identified (α = 0.01)
FLK	70	5	65	12	0	12
Samβada	82	5	77	6	0	6

Table 5.3: The results of a BLAST search of each of the loci identified as significant in both tests for disruptive selection.

Loci	Sequence	BLAST	E-value	Function
3275	TGCAGAAGGCCAAAGAGC AGGAACAAAATGTAAGAG GCCTTTATCCTGTTGGCTG TGGGGCCAGTGTGA	*Canis familiaris* chromosome X	0.013	No function assigned
3671	TGCAGATTCACTTACTCGC AGTTGGGTCCACCCCCGC GAGCATGAGATCGGAAGA GCGGTTCAGCAGGA	*Strongyloides stercoralis* genome assembly	0.004	No function assigned
1116	TGCAGTGAGGCTCCCTGC AAAACTTGGTGCTTTTGCA CCCCTTCTAGCTATACCAC TGGCCACAGCTAT	*Xenopus tropicalis* internexin neuronal intermediate filament protein	0.53	Expressed in brain
3608	TGCAGAAGAAAATTGTTA GCTGGATTGCTAACTTGTG GTAAATTAAAATGTCACAT GTTCCCATTTTTT	*Ovis canadensis canadensis* isolate 43U chromosome 2 sequence	0.53	No function assigned
3219	TGCAGTGTCCCTTCAATTT TTCAGTGTGGTACATTCTG ATATGGACTCAATGTCTTT AGCATGAGATCG	*Apteryx australis mantelli* genome assembly	1.9	No function assigned

5.4 Discussion

Significant evidence of disruptive selection was found using both FLK and Samβada but only a small number of loci were identified as being under disruptive selection by both methods. However, given that both the number of samples (19 and 8) and the number of populations (2) was small, any discovery of significant difference is important. A discussion of these results follows with the proviso that more complete sampling could provide more loci under selection and further strengthen the arguments presented.

There is some evidence of disruptive selection resulting from land clearing. Firstly, it must be noted that this evidence is suggestive only as the sample size was small and the effect of each locus is unknown. Further, only a few loci under disruptive selection were identified using both methods.The most of interesting of these is locus 1116 as it is expressed in the brain (Guérette *et al.* 2007). However, confirmation of loci that are involved in adaptation and the specific roles that each allele plays is outside the scope of this study. Rellstab *et al.* (2015) recommends mutational analyses using

knock out models to confirm the effects of the loci identified. However, this is complicated as the lack of genomic frameworks for skinks make the identification of locus function difficult. This is evidenced by the low E-scores in the BLAST results. Alternatively, this could be used as the basis of a second larger experiment with replicate populations. The logical population for such an experiment would be Tarana, also in the Central Tablelands of NSW, which has been used in previous investigations of the species (Stow *et al.* 2004a, Stow *et al.* 2004b). Nonetheless, in total this study identified 142 loci under disruptive selection at p=.05, of which 5 were identified by both programs. If even one of these is correctly identified, disruptive selection is occurring.

The results of this study are consistent with rapid adaptation in response to land clearing. Land clearing occurred at the Bathurst site approximately 100 years ago for grazing. As *E. cunninghami* reaches maturity after 5 years (Barwick 1965 from Lunney *et al.* 2009) there have only been approximately 20 generations since land clearing. Additionally, adaptation should be restrained by gene flow from neighbouring naturally vegetated areas. The findings suggesting rapid disruptive selection occurring over small distances. This is consistent with previous findings showing that adaptation can occur rapidly (Stockwell & Ashley 2004, Fraser *et al.* 2011), even in a single generation (Christie *et al.* 2012). Furthermore, the results are consistent with other findings of local adaptation over small spatial scales (Fraser *et al.* 2011). These findings may be important to conservation practice.

The results of this study are consistent with, but do not prove, increased risk of outbreeding depression. Previous work, done largely on model organisms (e.g. Woodworth *et al.* 2002), has shown that populations adapted to altered conditions suffer fitness losses, often dramatic losses, when reintroduced to wild conditions (Woodworth *et al.* 2002, Jule *et al.* 2008). Further, adaptation to captive environments is held partly responsible for the widespread failure of reintroduction efforts (Frankham 2008, Jule *et al.* 2008). In the same way, individuals adapted to cleared conditions may suffer reduced fitness in natural conditions. Furthermore, the offspring of individuals from two

populations where disruptive selection has produced local adaptation may suffer reduced fitness compared to their parents. Therefore, the results of this study may be taken as support for the concerns around outcrossing raised by several authors (e.g. Templeton *et al.* 1986, Edmans & Timmerman 2003). However, though disruptive selection creates the possibility for outbreeding depression we have no data to test for it and cannot confirm that it would occur. Additionally, the loss of fitness from outbreeding depression may be concealed by a short term fitness gain associated with relieving inbreeding depression in populations which have little genetic diversity. Nevertheless, the identification of disruptive selection between population pairs using this technology inform may conservation management of outcrossing.

A recent review by Frankham (2015) states that, "There are no scientific impediments to the widespread use of outcrossing … provided potential crosses have a low risk of outbreeding depression." This study demonstrates the usefulness of population genomics to determining, and reducing, that risk. Using the DaRTseq technology and a combination of analyses, this study was able to examine disruptive selection in a wild population for a species where little genomic information is available. For the first time, it was possible to identify areas where a wild population was missing genetic diversity. By examining potential outcrossing targets, it would be possible to determine which population could best reintroduce this diversity. Furthermore, the ability to identify potential disruptive selection enables conservation managers to avoid outcrossing to populations adapted to different local environments. This method would allow conservation managers to pick the better of several source populations, which is important given that outcrossing is recommended as a vital part of conservation managent (Frankham 2015). Further, services such as DArTseq, which perform preliminary bioinformatics, offer a reduced analytical load. This is particularly important in light of the difficulties associated with processing data including processing large numbers of sequences and understanding the causes of errors (further discussed in chapter 4). Therefore, the molecular and analytical approaches used herein can be considered successful, although they are not without opportunities for improvement.

Methods of detecting disruptive selection using population genomics can be refined further. Of 152 loci identified as under selection only 5 were identified by both methods. This means that the study may significantly underestimate the number of loci under selection .. In addition, the analytical process is still largely unwieldy, one approach required knowledge of three different programming languages and each required a unique input format. Ultimately, this is to be expected given the newness of this approach but the creation of simpler, and more universally accepted, analytical tools would only speed the adoption of these approaches in conservation practice.

6 General Discussion: Reflections on the incorporation of functional genetics into conservation genetics

"I have not failed, I have found 10,000 ways that won't work"

Edison

The central question that this thesis looked to answer was around the advisability and feasibility of incorporating functional genetic markers into conservation genetics. These questions have, in large part, been answered. The evolution of methodologies and technology in the field during the last decade has meant that functional markers can now be incorporated into most conservation genetics programs. As the period of this thesis and the breadth of technology and analytical approaches in this project mirrors the spread of next generation sequencing technologies, it has become possible to review the effects of the sequencing revolution on conservation genetics. Therefore, this discussion will examine the issues that arise out of the move towards use of next generation sequencing technology and reflect on the changes and challenges in the practice of conservation genetics.

6.1 The revolution was televised

The way that sequencing is carried out has fundamentally changed. Massive parallel sequencing now generates orders of magnitude more genetic data than in any time in history (Figure 1). Parallel sequencing has largely driven the sequencing of 51388 complete genomes (NCBI 2015). It is being used to discover the genetic diversity that underlies disease and pharmaceutical response in humans (The International Hapmap 3 Consortium 2010) and has facilitated personal genomic sequencing (Drmanac 2011, Peters *et al.* 2012). Furthermore, it has enabled the sequencing of entire transcriptomes in order to study gene expression and structure (Morozova & Marra 2008). The technology has improved the study of evolution of antibiotic resistance, non-coding RNA's, regulatory protein binding, chromatin packaging and metagenomics (Mardis 2008). Within population genetics, parallel sequencing has enabled high throughput methods of developing and

scoring genetic markers (e.g. Davey & Blaxter 2010) including the identification of functional markers

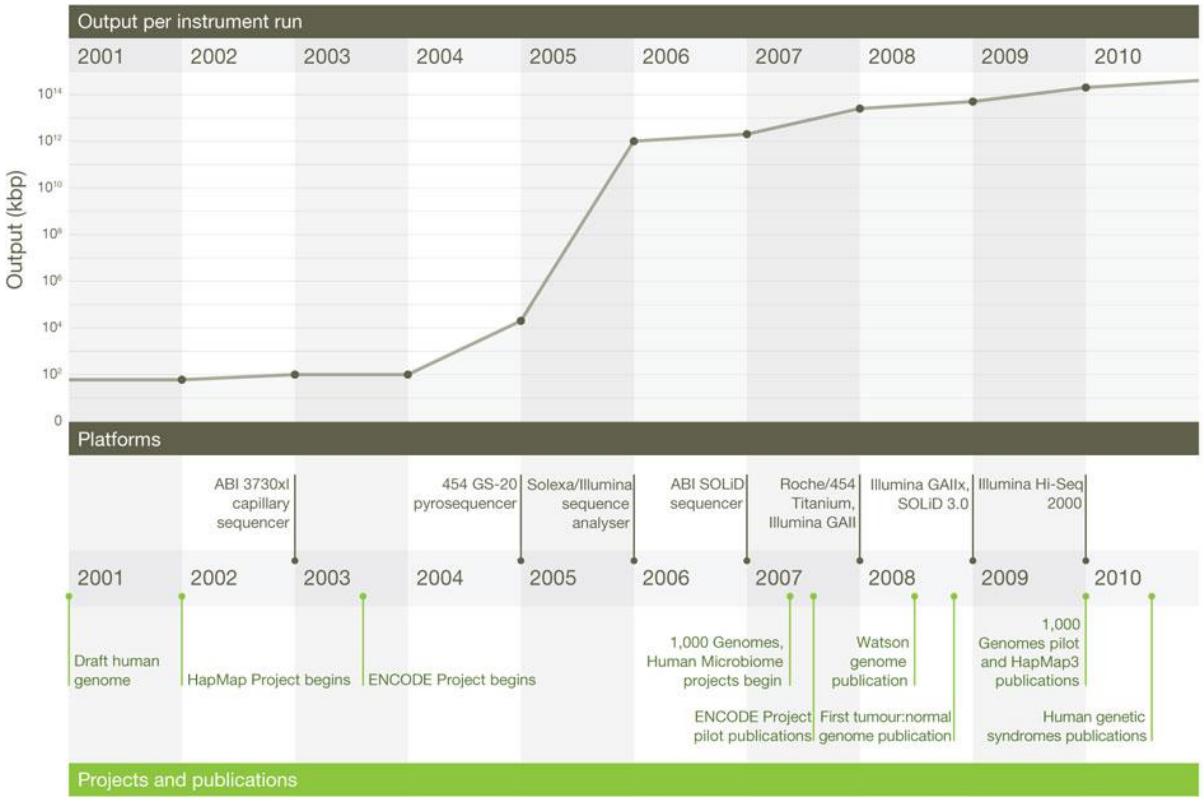


Figure 6.1: A description of the changes in output of DNA sequences and major innovations in sequencing technology from Mardis (2011). The beginning of parallel sequencing with 454 pyrosequencing in 2004 marked a period of rapid growth in the production of genetic data that continues: .

(e.g. Hohenlohe *et al.* 2010).

Within conservation genetics this technology has changed, fundamentally, the information available for conservation decision making. The process used by Babik *et al.* (2012) (also the basis of chapter 4) was able to genotype the MHC in an entire population in a single run. Previously no techniques allowed for genotyping of MHC in multiple individuals at one time. Hohenlohe *et al.* (2010) were able to determine SNP variation that was linked to adaptive variation in Threespine Stickleback populations. This allowed for the mapping of local adaptation in the species. Both of these methods required a single parallel sequencing reaction. These advances have directly impacted on conservation decision making. As an example, Hohhenlohe *et al.* (2011) used parallel sequencing

technology to produce markers capable of discerning hybrids in Westslope Cutthroat Trout (*Oncorhynchus clarkii lewisi*), a crucial step in attempting to prevent introgression from driving the species to extinction. Ouborg *et al.* (2010) stated conservation genetics is now becoming conservation genomics and our results in Chapter 5 support this. Thus, discussing the incorporation of functional genetics in conservation genetics is moot. Parallel sequencing technology which genotypes both functional and neutral genetic markers (e.g. Davey & Blaxter 2011) is now a standard practice in conservation. However, this thesis still provides an excellent opportunity to discuss the advantages and challenges for conservation genetics that have accompanied this sequencing revolution.

6.2 The streets are paved with gold: Using parallel sequencing to inform conservation genetics

Massive parallel sequencing provides amounts and types of data that were previously impossible within the constraints of conservation timescales and budgets. Parallel sequencing has removed the impediments that constrained the generation of genetic data for conservation, as shown in figure 2. Practically, this means that markers no longer need to be developed or optimised for individual species and large numbers of individuals can be genotyped in a single run (Davey *et al.* 2011). The genotyping of individual samples with multiple reactions as carried out in chapter 2 is no longer necessary.

Furthermore, the laborious process of marker optimisation required in both chapter 2 and chapter 4 is also no longer necessary. The Rad-Tag process, as used in chapter 5, and other population studies (e.g. Barchi *et al.* 2011, Scaglione *et al.* 2012) are examples of a parallel sequencing methodology that more than replaces the hundreds of individual reactions carried out for chapter 2. Scaglione *et al.* (2012) studied the globe artichoke (*Cynara cardunculus L. var. scolymus*) which had a poorly explored genome. Through a single RAD-Tag process they generated (approximately) 34,000 SNP

markers. Therefore, this methodology can reduce the time taken to furnish conservation decision makers with genetic data. Further, it can reduce the time spent generating data and the costs, including staff costs, associated with this. Additionally, the batching inherent in parallel sequencing techniques allows for larger sample sizes to be analysed, which provides more power for statistical analysis. This is illustrated by the differences in genotyping in chapters 4 and 5. Even though the MHC genotyping approach in chapter 4 was batched and high throughput compared to more traditional approaches, the approach used in chapter 5 is much faster and higher throughput. Additionally, the continued decrease in cost of these technologies (Watson 2014) means that more individuals can be genotyped and more genomic data obtained than with any other method.

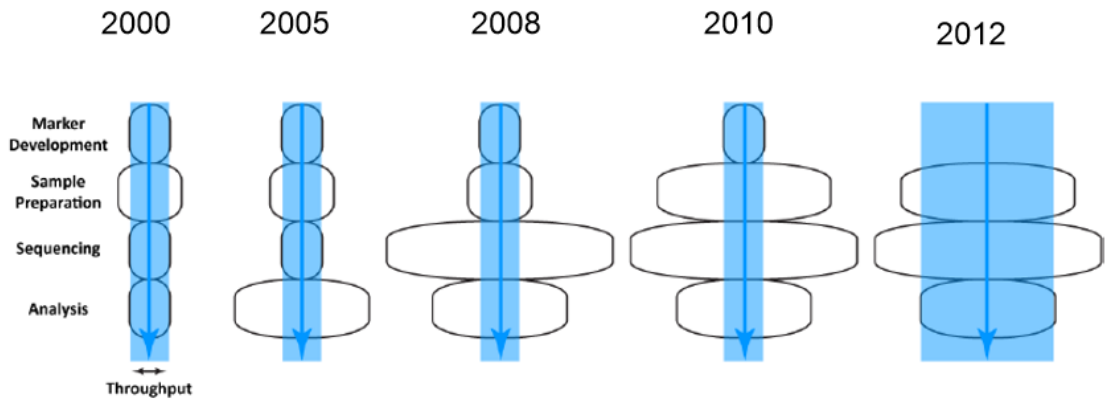


Figure 6.3 : A representation of the workflow associated with genotyping in population genetics.The direction of the process is shown by the arrow, the blue section indicates the volume of samples that can be processed in the narrowest point of the workflow and the bubbles represent the volume that each step in the workflow can process. The advances created by parallel sequencing have eliminated marker development and increased throughput in sample preparation and sequencing. Analysis is now the lowest throughput step of the process. Figure adapted from throughput figure Lemmon *et al.* (2012) by Stow (2015).

Massive parallel sequencing produces both more data and more types of data than any genetic markers previously used in conservation. Conservation geneticists have had to choose the correct genetic marker to answer a specific question (Wan *et al.* 2004), balancing resolution with practical genotyping concerns and appropriateness of data. The advantages and disadvantages of different marker types are touched on chapter 1 and discussed in greater detail by reviewers (e.g. Smith &

Wayne 1996, Want *et al.* 2004, Flinger & Klank 2009). In the same vein, this thesis compared (in Chapter 3.1) the information provided by functional and neutral markers in order to serve conservation decision making. Parallel sequencing makes this question largely irrelevant as both types of data can be generated in a single run (e.g. Stapley *et al.* 2010, Davey & Blaxter 2011, Zhao *et al.* 2013) and distinguished and compared as part of data analysis rather than experimental design. For example, Davey & Blaxter (2011) report selection signatures associated with oceanic and freshwater populations in 31 candidate genes previously linked to skeletal or osmotic traits in three spined sticklebacks. Further, resolution of genetic markers is no longer a constraint to conservation genetics as parallel sequencing can provide almost any level of resolution (e.g. Chiang *et al.* 2009, Tennessen *et al.* 2012). Additionally, the sequence data generated can be used to answer multiple questions and is being made available to the whole scientific community through databases such as the Genome Expression Omnibus (NCBI 2015). Therefore, the ability to generate good genetic data for a population of conservation concern, traditionally a difficult and important part of conservation genetics (e.g. Aliah *et al.* 1999, Stow *et al.* 2002), can now be assumed for new research on populations of conservation concern. However, as evidenced by this thesis, the generation and analysis of this data is not without challenges and potential pitfalls.

The major technical challenge for researchers around generating data with parallel sequencing technologies is around quality control. Previous genetic techniques used in conservation genetics such as microsatellites and RFLPs are relatively robust techniques. Using these, a conservation geneticist can produce results including electrophoretograms and gel results that, though imperfect, are easily scoreable. However, parallel sequencing requires high quality DNA (AGRF 2014). These increased requirements for DNA quality may have played a role in the failure of chapter 4 to produce results even though DNA passed quality control. Further, the nature of parallel techniques means that there is little or no opportunity for molecular optimisation. These optimisation steps have traditionally allowed for biochemistry to be tailored for a specific species at a relatively low cost. Therefore, when parallel sequencing techniques fail to produce data they do so at a much larger

scale. This was illustrated in chapter 4 and continues to be difficult for many studies focusing on the Major Histocompatibility Complex.

The Major Histocompatibility Complex remains a difficult area to work in spite of, and in some cases because of, the advances in parallel sequencing technology. Among functional genes the structure and evolution of MHC is unique (as discussed in chapter 1 and 4). Duplications and inter- and intra-locus and well as intra class gene conversions shape the genetic diversity of MHC (Beck *et al.* 1996). This results in a large number of similar or repeated sequences that can be either within or removed from other MHC genes. It is difficult to separate and resolve these types of sequences with parallel sequencing. Unlike earlier Sanger sequencing, parallel sequencing technology produces mainly short reads 50-150 b.p (Treangen & Salzberg 2013). Short reads may fail to correctly assign repeated sequences. Where the read length is shorter than the repeat length it is impossible to place the sequence. Further, when assembling genomes the presence of multiply highly similar sequences makes genome assembly difficult (Treangen & Salzberg 2013). The challenges associated with genotyping MHC are illustrated by the gene map of extended MHC in humans which required the efforts of a consortium and four years of research even after the publication of the human genome (Horton *et al.* 2004). Further, the diversity of gene structure in MHC across vertebrates (Kelley *et al.* 2005) makes it difficult to apply any lessons learned about MHC structure from one species to another. For this reason research on MHC has been relatively rare outside of model species (Kelley *et al.* 2005, Babik *et al.* 2009).

The key to unlocking the secrets of MHC in many vertebrates may rest in third generation sequencing. Current parallel sequencing technologies, though powerful, are not able to answer all genomic questions (Schat *et al.* 2010) and they are not ideal for understanding MHC. Single-molecular real-time sequencing (SMRT), termed third generation sequencing, requires no PCR step to amplify DNA (Liu *et al.* 2012). Nanopore sequencing is another single molecular approach that can differentiate between methylated and unmethylated bases (Clarke *et al.* 2009). Both are predicted

to further reduce the cost of sequencing (Clarke *et al.* 2009, Liu *et al.* 2012). Further, these and other third generation sequencing methods have longer DNA reads (Schat *et al.* 2010, Liu *et al.* 2012). Specifically, SMRT read length is approximately 1300 b.p. whilst Nanopore sequencing has the potential to produce 5000 b.p. sequences (Liu *et al.* 2012). These approaches have two benefits for MHC investigations: firstly by eliminating PCR one potential site of error is eliminated from the approach used in chapter 4, and secondly the long read length makes it easier to separate and position MHC sequences. This read length is crucial as it makes it easier to understand gene duplication events and differentiate between genes and pseudogenes. However, as with parallel sequencing, the real challenge of third generation methods will be how to interpret and understand the data they produce.

6.3 An embarrassment of riches: Challenges associated with parallel sequencing

Massive parallel sequence produces orders of magnitude more data than other methods used in conservation genetics and unsurprisingly this much data is hard to analyse. The sequencing revolution shifted one of the challenges in all genetics from data generation to data interpretation (Miller *et al.* 2010, Treangen & Salzberg 2012). This is evident within this thesis. Chapter 2 was carried out by traditional methods and the majority of time was invested into molecular optimisation and processing with data analysis being a relatively simple process. In chapter 4 the magnitude of the challenge of molecular optimisation was duplicated by the challenge of understanding, interpreting and attempting to troubleshoot the data generated. Chapter 5 however followed a pattern that was more typical of contemporary projects where molecular processes were a very small amount of the total time investment. This type of challenge is not limited to genetics, or even biological sciences; challenges associated with processing and analysing 'big data' are currently

constraining many industries (e.g. Lynch 2008, Silberstein *et al.* 2011, Laurila *et al.* 2012). Within conservation genetics this shift has also changed the type of data produced.

Analysing massive parallel sequencing data requires a particular skillset which is not ubiquitous among conservation geneticists. High throughput techniques and the rise of cost effective biomolecular services e.g. macrogen (www.macrogen.com), the AGRF (http://www.agrf.org.au/), Illumina (ww.illumina.com/services.html) indicate a new direction for the field of conservation genetics. The molecular skills required for the analysis carried out in chapter 2 may be becoming less valuable. A bioinformatics skillset is becoming more important. Many of the major analysis programs including Mothur (Schloss *et al.* 2009), R (R Core Team 2015) and MUSCLE (Edgar 2004) assume knowledge of multiple and diverse coding languages. However, bioinformatics skills are uncommon in conservation genetics and often self-taught (Zauhar 2001). Further, the programs themselves are clunky and unwieldy compared to those produced by specialist programmers (Christopher 2013, pers. comm.). Therefore, despite its fundamental importance (Kanehisa & Bork 2003), the ability to analyse the data produced by parallel sequencing technology remains a challenge within conservation genetics.

Conservation genetics may benefit from a different approach in order to effectively utilize data generated by parallel sequencing. The individualistic approach common in conservation genetics where a single scientist collects samples, generates data and analyses data may no longer be optimal. In order to fully harness the diverse opportunities created by parallel sequencing, a team based approach may be necessary. Human genetics research is typically dominated by large multi-disciplinary teams working co-operatively on a single problem e.g. (International Hapmap Consortium, International Inflammatory Bowel Disease Genetics Consortium, International Headache Genetics Consortium). This allows for big problems to be addressed by teams which incorporate diverse specialist skillsets including genetics, biomolecular and bioinformatics (e.g. Rivas *et al.* 2011, Jostins *et al.* 2012). Such an approach is distinct from that used in Chapter 4 and 5 where

specialist advice was sought to fill gaps in a researcher's knowledge. The experiences in this thesis suggest the latter approach may suffer because of the difficulties communicating the questions and study system across disciplines in short time periods. Multi-disciplinary teams guided by conservation genetics questions allow for long term collaboration. These teams have proven effective in other disciplines and would allow for conservation geneticists to reap the benefits of parallel sequencing data without having to develop an entirely new skillset in a markedly different field.

A major challenge currently facing conservation genetics lies in integrating new sequence data with older sequence and genotype data. As chapter 3 demonstrates, different types of genetic data are difficult to compare. Whilst significant relationships were found between genetic diversity at MHC and microsatellite loci, these relationships are not straightforward. This mirrors the wider body of research that often compares genetic markers on the same population, finding both the same (e.g. Hedrick *et al.* 2001, Aguilar & Garza 2006) and different (e.g. Aguilar *et al.* 2004, Ekblom *et al.* 2007) outcomes depending on the markers used. Overall, both genome wide genetic diversity (Reed & Frankham 2003) and genetic diversity at MHC (e.g. Bonneaud *et al.* 2004, Westerdahl *et al.* 2005, Pitcher & Neff 2006) are correlated with fitness but using one to predict the other accurately is difficult. This may be symptomatic of a larger problem around using one type of genetic diversity to predict anotherthat will continue. Genetic data collected using genetic markers, the primary method of collecting data from the 1980's to the mid-2000's, is difficult to compare to genetic data generated by parallel sequencing. Although within-species studies have been carried out (e.g. Babik & Radwan 2007 compared to Babik *et al.* 2009), we know of no meta-analysis to compare the data. This is primarily because of differences in both breadth (few loci in genetic markers vs many loci in parallel sequencing) and depth (many amplifications in genetic markers vs few amplifications in parallel) of the data. This means that we do not know how to effectively use the new data with the old to answer conservation genetics questions. Either, an effort must be made to estimate a general relationship for variation at sequence based and traditional genetic markers or much of the older

work on populations of interest must be repeated. On first glance repeating this work appears daunting and unnecessary. However, upon further reflection, should the trend towards decreased cost of sequencing reactions continue (Gullapalli *et al.* 2012), and with techniques for the extraction (e.g. Rohlan *et al.* 2004) and amplification (e.g. Freeman *et al.* 2003) of DNA after long term sample storage, it may become an attractive option so that a single standard type of data is produced.

Ultimately the challenges involved in using parallel sequencing data must be overcome. Conservation genetics is somewhat unusual within genetics in that inquiry is driven not by curiosity but rather impending doom. The depth and breadth of sequence data allows for new methods of pushing back that impending doom for many populations. For others genetic management may avert extinction altogether. Therefore, projects such as ours and others that attempt to integrate technological advancement into conservation practice should continue. In short, the prize is worth the effort.

References

Abrams, E. T., and Miller, E. M. (2011). The roles of the immune system in Women's reproduction: Evolutionary constraints and life history trade-offs. *American journal of physical anthropology*, 146(S53): p.134-154.

Act, E. P. B. C. (1999). Environment Protection and Biodiversity Conservation Act 1999 (Cth).

Agbali, M., Reichard, M., Bryjová, A., Bryja, J., and Smith, C. (2010). Mate Choice For Nonadditive Genetic Benefits Correlate With Mhc Dissimilarity In The Rose Bitterling (Rhodeus Ocellatus). *Evolution*, 6(64):p.1683-1696.

Agrawal, A.A. (2001). Phenotypic plasticity in the interactions and evolution of species. *Science*, 294(5541): p.321-326.

AGRF (Australian Genome Research Facility) (2014). Sequenom Genotyping Sample Submission Guidelines. *Australian Genome Research Facility, St Luca, Queensland*,

Agudo, R., Carrete, M., Alcaide, M., Rico, C., Hiraldo, F., and Donázar, J. A. (2012). Genetic diversity at neutral and adaptive loci determines individual fitness in a long-lived territorial bird. *Proceedings of the Royal Society B: Biological Sciences.* 279(1741): p.3241-3249..

Aguilar, A., Garza, Garza, J.C. (2006). A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss Walbaum*). *Molecular Ecology*, 4(15):p.923-937.

Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D., and Wayne, R.K. (2004). High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences*, 10(101):p.3490–3494.

Alam, M.A., Siddik, M.A.B., Nahar, A., Mahmud, S., and Parvez, I. (2014). Genetic Divergence in Natural Populations of Yellowtail Catfish. *Pangasius pangasius* (Hamilton-Buchanan, 1822) Using Allozyme Marker. *World Applied Sciences Journal*, 29(7):p.926-932.

Alberston, R.C., Streelman, J.T., Kocher, T.D. (2003). Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proceedings of the National Academy of Sciences*, 9(100):p.5252–5257.

Alexander, R. D. (1974). The evolution of social behavior. *Annual review of ecology and systematics*, p.325-383.

Aliah, R.S., Takagi, M., Dong, S., Teoh, C.T., and Taniguchi, N. (2008). Isolation and Inheritance of Microsatellite Markers in the Common Carp *Cyprinus carpio*. *Fisheries science*, 65(2):p.235-239.

Aljanabi, S.M., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acid Research*, 25 (22):p.4692-4693.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *Journal Molecular Biology* , 215:p.403-410 .

Amos, D.B., and Bach, F.H. (1968). Phenotypic expressions of the Major Histocompatibility Locus in man (HL-A). Leukocyte antigens and mixed leukocute culture reactivity. *The Journal of experimental medicine, 128(4)*, pp.623-637.

Anne, C. (2006). Choosing the right Molecular Genetic Markers for studying biodiversity: from molecular evolution to practical aspects. *Genetica,* 127(1-3):p.101-120.

Araki, H., Cooper, B., and Blouin, M.S. (2007). Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science*, 318(5847):p.100-103.

Arala-Chaves, M., and Sequeira, T. (2000). Is there any kind of adaptive immunity in invertebrates? *Aquaculture*, 191(1):p.247-258.

Arneberg, P., Skorping, A., Grenfell, B., and Read, A. F. (1998). Host densities as determinants of abundance in parasite communities. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1403): p.1283-1289.

Auchincloss, H., and Winn, H.J., (2003). Clarence Cook Little (1888–1971): The Genetic Basis of Transplant Immunology. *American Journal of Transplantation*, 4:p.155-159.

Austin, A.D., and Dillon, N. (1997). Extraction and PCR of DNA From Parasitoid Wasps That Have Been Chemically Dried. *Australian Journal of Entomology*, 36:p.241-244."

Awadalla, P., & Ritland, K. (1997). Microsatellite variation and evolution in the *Mimulus guttatus* species complex with contrasting mating systems. *Molecular Biology and Evolution*, 14(10): p.1023-1034.

Barbara, T., Palma-Silva, C., Paggi, G.M., Bered, F., Fay, M.F. and Lexer, C., (2007). Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Molecular Ecology*, 16(18): p.3759-3767.

Babik, W. (2010). Methods for MHC genotyping in non-model vertebrates. *Molecular Ecology Resources*, 10(2):p.237–251.

Babik, W. and Radwan, J. (2007). Sequence diversity of MHC class II DRB genes in the bank voleMyodes glareolus. *Acta theriologica*, 52(3): p.227-235.

Babik, W., Taberlet, P., Ejsmond, M.J. and Radwan, J., 2009. (2009). New generation sequencers as a tool for genotyping of highly polymorphic multilocus MHC system. *Molecular Ecology Resources*, 9(3):p.713-771.

Bain, B., and Lowenstein, L. (1964). Genetic Studies on the Mixed Leukocyte Reaction. *Science*, 145(3638):p.1315-1316.

Baldwin, H. J., Hoggard, S. J., Snoyman, S. T., Stow, A. J., & Brown, C. (2010). Non-invasive genetic sampling of faecal material and hair from the grey-headed flying-fox (*Pteropus poliocephalus*). *Australian Mammalogy*, 32(1): p.56-61.

Ballou, J. D., & Lacy, R. C. (1995). Identifying genetically important individuals for management of genetic variation in pedigreed populations. *Population management for survival and recovery*, p.76-111.

Ballou, J.D., Lees, C., Faust, L.J., Long, S., Lynch, C., Lackey, L.B. and Foose, T.J. (2010). Demographic and genetic management of captive populations. In Wild mammals in captivity: principles and techniques for zoo management, *University of Chicago Press*. p219.

Balloux, F. (2001). EASYPOP (version 1.7): a computer program for population genetics simulations. *Journal of Heredity*, 92(3): p.301-302.

Balloux, F., & Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular ecology*, 11(2): p.155-165.

Balvanera, P., Pfisterer, A.B., Buchmann, N., He, J.S., Nakashizuka, T., Raffaelli, D., & Schmid, B. (2006). Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology letters*, 9(10): p.1146-1156.

Barchi, L., Lanteri, S., Portis, E., Acquadro, A., Vale, G., Toppino, L.and Rotino, G.L. (2011). Identification of SNP and SSR markers in eggplant using RAD tag sequencing. *BMC Genomics*, 12(304):p.1471-2164.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O., Swartz, B., Quental, T.B., ...& Ferrer, E.A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, 471(7336): p.51-57.

Beacham, T.D., Candy, J.R., McIntosh, B., Macconnachie,C., Tabata, A., Kaukinen, K., Deng, L., Miller, K.M., and Withler, R.E. (2005). Estimation of Stock Composition and Individual Identification of Sockeye Salmon on a Pacific Rim Basis Using Microsatellite and Major Histocompatibility Complex. *Variation Transactions of the American Fisheries Society*, 134:p.1124–1146.

Beauchamp, G. (2013). *Social Predation: How group living benefits predators and prey*. Elsevier, Philadephia.

Beck, B. B., Rapaport, L. G., Price, M. S., & Wilson, A. C. (1994). *Reintroduction of captive-born animals*. In Creative conservation (pp. 265-286). Springer Netherlands.,

Beck, S., Abdulla, S., Alderton, R.P., Glynne, R.J., Gut, I.G., Hosking, L.K., Jackson, A., Kelly, A., Newell, W.R., Sanseau, P., Radley, E., Thorpe, K.L., and Trowsdale, J. (1996). Evolutionary Dynamics of Non-coding Sequences within the Class II Region of the Human MHC. *Journal of Molecular Biology*, 255(1):p.1-13.

Bennett, P. (2000). Demystified... microsatellites. *Journal of Clinical Pathology*, 53(4),p.177.

Bentzen, p., Taggart, C.T., Ruzzante, D.E., and Cook, D. (1996). Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences*, *53*(12): pp.2706-2721.

Bernatchez, L., and Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, 16(3):p.363–377.

Bichet, C., Penn, D.J., Moodley, Y., Dunoyer, L., Cellier-Holzem, E., Belvalette, M., Grégoire, A., Garnier, S. and Sorci, G. (2014). Females tend to prefer genetically similar mates in an island population of house sparrows. *BMC evolutionary biology,* 14(1), p.1.

Bijlsma, R., Bundgaard, J. and Boerema, A.C. (2000). Does inbreeding affect the extinction risk of small populations? predictions from Drosophila. *Journal of Evolutionary Biology*, 13(3):p.502-514.

Bijlsma, R., Bundgaard, J., and van Putten, W.F. (1999). Environmental dependence of inbreeding depression and purging in Drosophila melanogaster. *Journal of Evolutionary Biology*, 12(6):p.1125-1137.

Binladen, J., Gilbert, M.T.P., Bollback, J.P., Panitz, F., Bendixen, C., Nielsen, R. and Willerslev, E., (2007). The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PloS one*, 2(2):p. e197.

Binz, T., Reusch, T.B.H., Wedekind, C., and Milinski, M. (2001). SSCP analysis of Mhc class IIB genes in the threespine stickleback. *Journal of Fish Biology*, 58(3):p.887-890.

Black, J.M., (1996). *Partnerships in Birds: The Study of Monogamy.* The Study of Monogamy Oxford University Press Oxford,

Bollmer, J.L., Hull, J.M., Ernest, H.B., Sarasola, J.H. and Parker, P.G., (2011). Reduced MHC and neutral variation in the Galápagos hawk, an island endemic. *BMC evolutionary biology*, 11(1), p.1.

Bolnick, D.I. (2004). Can intraspecific competition drive disruptive selection? An experimental test in natural populations of sticklebacks. *Evolution*, 58(3): p.608-618.

Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S., and SanCristobal, M. (2010). Detecting Selection in Population Trees: The Lewontin and Krakauer Test Extended. *Genetics*, 186(1):p.241-262.

Bonneaud, C., Chastel, O., Federici, P., Westerdahl, H. and Sorci, G. (2006). Complex Mhc-based mate choice in a wild passerine. *Proceedings of the Royal Society of London B: Biological Sciences*, 273(1590):p.1111-1116.

Bonnell, M.L., & Selander, R.K. (1974). Elephant seals: genetic variation and near extinction. *Science,* 184(4139): p.908-909.

Borghans, J. A., Beltman, J. B., and De Boer, R. J. (2004). MHC polymorphism under host-pathogen coevolution. *Immunogenetics,* 55(11): p.732-739.

Bowcock, A.M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J.R. and Cavalli-Sforza, L.L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, 368(6470):p.455-457.

Bowling, A.T., Zimmermann, W., Ryder, O., Penado, C., Peto, S., Chemnick, L., Yasinetskaya, N., and Zharkikhd, T. (2003). Genetic variation in Przewalski's horses, with special focus on the last wild caught mare, 231 Orlitza III Cytogenet. *Genome Research*, 101:p.226-234.

Boyce, W.M., Hedrick, P.W., Muggli-Cockett, N.E., Kalinowski, S., Penedo, M.C.T. and Ramey, R.R., (1997). Genetic variation of major histocompatibility complex and microsatellite loci: a comparison in bighorn sheep. *Genetics,* 145(2), p..421-433.

Bremermann, H. J., and Thieme, H. R. (1989). A competitive exclusion principle for pathogen virulence. *Journal of mathematical biology,* 27(2):p.179-190.

Brenner, S., Williams, S.R., Vermaas, E.H., Storck, T., Moon, K., McCollum, C., Mao, J., Luo, S., Kirchner, J.J., Eletr, S., DuBridge, R.B., Burcham, T., and Albrecht, G., (2000). In vitro cloning of complex mixtures of DNA on microbeads: Physical separation of differentially expressed cDNAs. *Proceedings National Academy of Sciences*, 97(4):p.1665-1670.

Briles, W. E., Goto, R.M., Auffray, C., and Miller. M.M. (1993). A polymorphic system related to but genetically independent of the chicken major histocompatibility complex. *Immunogenetics*, 37:p.408–414.

Bryant, E. H., & Reed, D. H. (1999). Fitness decline under relaxed selection in captive populations. *Conservation Biology*, 13(3): p.665-669.

Burbidge, A. A., & McKenzie, N. L. (1989). Patterns in the modern decline of Western Australia's vertebrate fauna: causes and conservation implications. *Biological conservation*, 50(1): p.143-198.

Burgerhout, E., Tudorache, C., Brittijn, S. A., Palstra, A. P., Dirks, R. P., and van den Thillart, G. E. (2013). Schooling reduces energy consumption in swimming male European eels, *Anguilla anguilla* L. *Journal of Experimental Marine Biology and Ecology*, 448: p.66-71.

Burkey, T V., (1989). Extinction in Nature Reserves: The Effect of Fragmentation and the Importance of Migration between Reserve Fragments. *OIKOS,* 55:p.75-81.

Butler, D., and D.Merton. (1992). *The black robin: saving the world's most endangered bird.* Oxford Univ. Press, Auckland.,

Carlson, J.S., Mart´ınez-Go´mez, J.E., Valkiunas, G., Loiseau, C., Bell, D.A.and Sehgal, R.N.M. (2013). Diversity and Phylogenetic relationships of hemospordian parasites in birds of Socorro Islandm Mexico, and their role in the re-introduction of the Socorro Dove (*Zenaida graysoni*). Journal of Parasitology, 99(2):p.270-276.

Cassinello, J. (2005). Inbreeding depression on reproductive performance and survival in captive gazelles of great conservation value. *Biological Conservation,* 122:p.453-464.

Catchen, J.M., Amores A., Hohenlohe P., Cresko W., and Postlethwait, J.H. (2011). Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3: Genes, Genomes, Genetics*, 1(3):p.171-182.

Caudron, A.K., Negro, S,S., Muller, C.G., Boren, L.J., and Gemell, N.J. (2007). Hair sampling and genotyping from hair follicles: A minimally-invasive alternative for genetics studies in small, mobile pinnipeds and other mammals *Marine Mammal Science*, 23(1):p.184–192.

Chapman, J.R., Nakagawa, S., Coltman, D.W., Slate, J. and Sheldon, B.C., (2009). A quantitative review of heterozygosity–fitness correlations in animal populations. *Molecular Ecology*, 18(13), p..2746-2765.

Charbonnel, N. and Pemberton, J. (2005). A long-term genetic survey of an ungulate population reveals balancing selection acting on MHC through spatial and temporal fluctuations in selection. *Heredity,* 95(5), p..377-388.

Chiang, D.Y., Getz, G., Jaffe, D.B., O'Kelly, M.J.T., Zhao, X., Carter, S.L., Russ, C., Nusbaum, C., Meyerson, M., and Lander, E.S. (2009). High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nature Methods*, 6:p.99-103.

Childers, C.P., Newkirk, H.L., Honeycutt, D.A., Ramlachan, N., Muzney, D. M., Sodergren, E., Gibbs, R.A., Weinstock, G.M., Womack, J.E., Skow, L.C. (2006). Comparative analysis of the bovine MHC class IIb sequence identifies inversion breakpoints and three unexpected genes. *Animal Genetics*, 37(2): p.121–129.

Chistiakov, D.A., Hellemans, B. and Volckaert, F.A. (2006). Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. *Aquaculture,* 255(1),p. .1-29..

Christie, M.R., Marine, M.L., French, R.A., and Blouin, M.S. (2011). Genetic adaptation to captivity can occur in a single generation. *Proceedings of the National Academy of Sciences,* 109(1)p.238-242.

Cinelli, P., Rettich, A., Seifert, B., Bürki, K., & Arras, M. (2007). Comparative analysis and physiological impact of different tissue biopsy methodologies used for the genotyping of laboratory mice. *Laboratory Animals,* 41(2):p. 174-184.

Clarke, J., Wu, H.C., Jayasinghe, L., Patel, A., Reid, S., and Bayley, H. (2009). Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology,* 4:p.265-270.

Colautti, R.I., and Barrett, C.H. (2013). Rapid Adaptation to Climate Facilitates Range Expansion of an Invasive. *Plant Science,* 342:p.364-366.

Consuegra, S. and de Leaniz, C.G., (2008). MHC-mediated mate choice increases parasite resistance in salmon. *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1641), p..1397-1403.

Cooke, S.J. and Philipp, D.P. (2005). Influence of local adaptation and interstock hybridization on the cardiovascular performance of largemouth bass *Micropterus salmoides*. *Journal of experimental biology,* 208(11), p..2055-2062.

Cooper, M. B., Lokan, K.H., Williams, G.A., and Toussiant, L. (1983). *The Radiological status of the Monte Bello Islands, May 1983*. Australian Radiation Laboratory. Victoria,

1 Cordeiro, G.M., Casu, R., McIntyre, C.L., Manners, J.M., & Henry, R.J. (2001). Microsatellite markers
2 from sugarcane (Saccharum spp.) ESTs cross transferable to erianthus and sorghum. *Plant science,*
3 160(6): p.1115-1123.

4 Coughlan, J.P., Galvin, P.T., Fitzgerald, R.D., Naevdal, G. and Cross, T.F. (1998). Microsatellite DNA
5 variation in wild populations and farmed strains of turbot from Ireland and Norway: a preliminary
6 study. *Journal of Fish Biology*,52(5): p.916-922.

7 Cousyn, C., Meester, L.D., Colbourne, J.K., Brendonck, L., Verschuren, D., and Volckaert, F. (2001).
8 Rapid, local adaptation of zooplankton behavior to changes in predation pressure in the absence of
9 neutral genetic changes. *Proceedings of the National Academy of Sciences,* 98(11):p.6256–6260.

10 Crandall, K.A., Bininda-Emonds, O.R., Mace, G.M., and Wayne, R.K., (2000). Considering evolutionary
11 processes in conservation biology. *Trends in Ecology and Evolution,* 15(7):p.290-295.

12 Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity,* 83(3): p.260-270.

13 Crochet, P.A. (2000). Genetic structure of avian populations—allozymes revisited. *Molecular*
14 *Ecology*, 9(10):p.1463-1469.

15 Crow, J.F. (1954). *Breeding structure of populations II. Effective population number.* In
16 O.Kempthorne, T.A., Bancroft, J.W., Gowen, and J.L. Lush (eds.), Statistics and Mathematics in
17 Biology. Iowa State College Press, Ames.

18 Cutrera, A.P., Fanjul, M.S. and Zenuto, R.R. (2012). Females prefer good genes: MHC-associated
19 mate choice in wild and captive tuco-tucos. *Animal Behaviour,* 83(3): p.847-856.

20 Da Silva, A., Luikart, G., Yoccoz, N.G., Cohas, A. and Allaine, D.,(2006). Genetic diversity-fitness
21 correlation revealed by microsatellite analyses in European alpine marmots (*Marmota marmota*).
22 *Conservation Genetics,* 7(3):p.371-382.

23 Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. John Murray, London, UK,

Daszak, P., Cunningham, A. A., and Hyatt, A. D. (2000). Emerging infectious diseases of wildlife--threats to biodiversity and human health. *Science*, 287(5452):p.443-449.

Davey, J, W., and Blaxter, M.L. (2010). RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, 9 (5-6):p.416-423.

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., and Blaxter, M.L., (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics,* 12:p.499-510.

Dawson, T.P., Jackson, S.T., House, J.I., Prentice, I.C., & Mace, G.M. (2011). Beyond predictions: biodiversity conservation in a changing climate*. Science,* 332(6025): p.53-58.

De Bakker, P.I.W., McVean, G., Sabeti, P.C., Miretti, M.M., Green, T., Marchini, J., Ke, X., Monsuur, A.J., Whittaker, P., Delgado, M., Morrison, J., Richardson, A., Walsh, E.C., Gao, X., Galver, L., Hart, J., Hafler, D.A., Pericak-Vance, M., Todd, J.A., Daly, M.J., Trowsdale, J., Wijmenga, C., Vyse, T.J., Beck, S., Murray, S.S., Carrington, M., Gregory, S., Deloukas, P., and Rioux, J.D. (2006). A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nature Genetics,* 38:p.1166-1172.

Delany., M.E., Robinson, C.M., Goto, R.M., and Miller, M.M (2009). Architecture and Organization of Chicken Microchromosome 16: Order of the NOR, MHC-Y, and MHC-B Subregions. *Journal of Heredity,* 100(5):p. 507-514.

DePristo, M.A., Banks, E., Poplin, R.E., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., Angel, G.D., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., and Daly, M.J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics,* 43(5):p.491-498.

Dhuyvetter.H., Gaublomme, E., and Desender, K. (2004). Genetic differentiation and local adaptation in the salt-marsh beetle *Pogonus chalceus*: a comparison between allozyme and microsatellite loci. *Molecular Ecology,* 13(5):p.1065-74.

Diamond, J. M. (1975). The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. *Biological conservation*, 7(2): p.129-146.

Dias, P.C. (1996). Sources and sinks in population biology. *Trends in Ecology & Evolution,* 11(8): p.326-330.

Dickersin, K. (1990). The existence of publication bias and risk factors for its occurrence. *Jama,* 263(10), p.1385-1389.

Dickman, C. R. (1996). *Overview of the impacts of feral cats on Australian native fauna* (pp. 1-92). Canberra: Australian Nature Conservation Agency.,

Dinerstein, E., and McCracken, G.F. (1990). Endangered Greater One-Horned Rhinoceros Carry High Levels of Genetic Variation. *Conservation Biology*, 4(4):p.417-422.

Dionne, M., Miller, K.M., Dodson, J.J., Caron, F., and Bernatchez, L., (2007). Clinal variation in MHC diversity with temperature: evidence for the ole of host-pathogen interaction on local adaptation in Atlantic Salmon. *Evolution,* 61(9):p.2154–2164.

Dohm, J.C., Lottaz, C., Borodina, T., and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acid Research*, 36(16):p.e105.

Doyle, R. M., Perez-Enriquez, R., Takagi, M., & Taniguchi, N. (2001). Selective recovery of founder genetic diversity in aquacultural broodstocks and captive, endangered fish populations. *Genetica,* 111(1-3): p.291-304.

Dobson, A. and Lyles, A. (2009). Black-footed ferret recovery. *Science,* 288(5468): p.985-988

Drake, J. W., Charlesworth, B., Charlesworth, D., and Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics,* 148(4): p.1667-1686.

Drmanac, R. (2011). The advent of personal genome sequencing. *Genetics in Medicine*, 13:p.188–190.

Dyer, P. and Middleton, D. (1993). *Histocompatibility testing: a practical approach* (Vol. 118). Oxford University Press.,

Earnhardt, J. M. (1999). Reintroduction programmes: genetic trade-offs for populations. *Animal Conservation*, 2(4):p. 279-286.

Ebensperger, L. A., Chesh, A. S., Castro, R. A., Tolhuysen, L. O., Quirici, V., Burger, J. R., ... and Hayes, L. D. (2011). Burrow limitations and group living in the communally rearing rodent, *Octodon degus*. *Journal of Mammalogy*, 92(1): p.21-30.

Edgar, R.C., (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acid Research*, 32 (5):p.1792-1797.

Edmands, S. (2006). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology*, 16(3):p.463-475.

Edmands, S., and Timmerman, C.C. (2002). Modeling Factors Affecting the Severity of Outbreeding Depression. *Conservation Biology*, 17(3): pp.883-892.

Edwards, S.V. and Potts, W.K., 1996. (1996). *Polymorphism of genes in the major histocompatibility complex (MHC): implications for conservation genetics of vertebrates*. In Smith, T.B., and Wayne, R.K. (eds). Molecular Genetic Approaches in Conservation. Oxford University Press, New York.,

Efron, B. (1981). Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. *Biometrika,* 68(3): p.589-599.

Ehiobu, N. G., Goddard, M.E., Taylor, J.F. (1989). Effect of rate of inbreeding on inbreeding depression in Drosophila melanogaster. *Theoretical and Applied Genetics*, 77(1): p123-127.

Eizaguirre, C., and Lenz, T.L., (2010). Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. *Journal of Fish Biology,* 77(9):p.2023-2047.

Eizaguirre, C., Yeates, S.E., Lenz, T.L., Kalbe, M. and Milinski, M. (2009). MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Molecular Ecology,* 18(15):p.3316-3329.

Ejsmond, M. J., and Radwan, J. (2011). MHC diversity in bottlenecked populations: a simulation model. *Conservation Genetics,* 12(1):p. 129-137.

Ekblom, R., Saether, S.A., Jacobsson, P., Piske, P., Sahlman, T., Grahn, M., Kalas, J.A., and Hoglund, J. (2007). Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Molecular Ecology,* 16(7):p.1439–1451.

Eldridge, M. D. B., Kinnear, J. E., Zenger, K. R., McKenzie, L. M., & Spencer, P. B. S. (2004). Genetic diversity in remnant mainland and "pristine" island populations of three endemic Australian macropodids (Marsupialia): *Macropus eugenii, Lagorchestes hirsutus* and *Petrogale lateralis*. *Conservation Genetics,* 5(3): p.325-338.

Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, 29(1):p.51-63.

Ellstrand, N.C., & Elam, D.R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annual review of Ecology and Systematics,* p.217-242.

Ellstrand, N.C., & Roose, M.L. (1987). Patterns of genotypic diversity in clonal plant species. *American Journal of Botan,*, p. 123-131.

1 Elmqvist, T., C.Folke, M.Nyström, G.Peterson, J.Bengtsson, B.Walker, and J.Norberg. (2003).
2 Response diversity, ecosystem change, and resilience. *Frontiers in Ecology and the Environment,*
3 1:p.488-494.

4 Emlen, S.T. and Oring, L.W. (1977). Ecology, sexual selection, and the evolution of mating
5 systems. *Science,* 197(4300):p.215-223.

6 England, P. R., Osler, G. H., Woodworth, L. M., Montgomery, M. E., Briscoe, D. A., & Frankham, R.
7 (2003). Effects of intense versus diffuse population bottlenecks on microsatellite genetic diversity
8 and evolutionary potential. *Conservation Genetics*, 4(5):p. 595-604.

9 Epperson, B.K. (1999). Gustave Male´cot, 1911–1998: Population Genetics Founding Father.
10 *Genetics,* 152: p.477–484.

11 Erlich, R.L., Jia, X., Anderson, S., Banks, E., Gao, X., Carrington, M., Gupta, n., DePristo, M., Henn,
12 M.R., Lennon, N.J., and Bakker, P.IW. (2011). Next-generation sequencing for HLA typing of class I
13 loci. *BMC Genomics*, 12(42):p.42.

14 Estoup.A., Rousset, F., Michalakis.Y., Cornuet.M., Adriamanga.M., and Guyomard, R. (1997).
15 Comparative analysis of microsatellite and allozyme markers: a case study investigating
16 microgeographic differentiation in brown trout (*Salmo trutta*). *Molecular Ecology*, 7:p.339-353.

17 Ewing, S. R., Nager, R. G., Nicoll, M. A., Aumjaud, A., Jones, C. G., & Keller, L. F. (2008). Inbreeding
18 and loss of genetic variation in a reintroduced population of Mauritius kestrel. *Conservation Biology*,
19 22(2):p. 395-404.

20 Fabuel, E., Barragán, C., Silió, L., Rodriguez, M.C., & Toro, M.A. (2004). Analysis of genetic diversity
21 and conservation priorities in Iberian pigs based on microsatellite markers. *Heredity*, 93(1): p.104-
22 113.

23 Fagan, W.F. and Holmes, E.E. (2006). Quantifying the extinction vortex. *Ecology Letters*, 9:p.51-60.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics*, 78(2):p.737-56.

Fenner, F., and Myers, K. (1978). *Myxoma virus and myxomatosis in retrospect: the first quarter century of a new disease. Viruses and the Environment*. Academic Press, London, 539-570.,

Fernández, J. and Caballero, A. (2001). A comparison of management strategies for conservation with regard to population fitness. *Conservation Genetics*, 2(2):p.121-131.

Fernández, J., Toro, M. A., & Caballero, A. (2001). Practical implementation of optimal management strategies in conservation programmes: a mate selection method. *Animal biodiversity and Conservation,* 24(2): p.17-24.

Fernando, M.M., Stevens, C.R., Walsh, E.C., De Jager, P.L., Goyette, P., Plenge, R.M., Vyse, T.J. and Rioux, J.D. (2008). Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genetics,* 4(4): p.e1000024.

Fernando, R., Polet, G., Foead, N., Ng, L.S., Pastorini, J. and Melnick, D.J. (2006). Genetic diversity, phylogeny and conservation of the Javan rhinoceros (*Rhinoceros sondaicus*). *Conservation Genetics,* 7:p.439-448.

Feschotte, C., and Gilbert, C. (2012). Endogenous viruses: insights into viral evolution and impact on host biology. *Nature Reviews Genetics*, 13(4)p.283-296.

Fessehaye, Y., El-bialy, Z., Rezk, M.A., Crooijmans, R., Bovenhuis, H. and Komen, H., (2006). Mating systems and male reproductive success in Nile tilapia (*Oreochromis niloticus*) in breeding hapas: a microsatellite analysis. *Aquaculture,* 256(1),p. .148-158.

Fichot, E.B., Norman, R.S (2008). Microbial phylogenetic profiling with the Pacific Biosciences sequencing platform. *Microbiome*, 1(1):p.1.

Finger, A., and Klank, C., (2009). *Review Molecular Methods: Blessing or Curse?* In Relict Species, p.309-320. Springer Berlin Heidelberg.

Flajnik, M. F., and Kasahara, M. (2001). Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. *Immunity,* 15(3):p. 351-362.

Flajnik, M., Ohta, Y., Namikawa-Yomada, C., and Nonaka, M. (1999). Insight into the primordial MHC from studies in ectothermic vertebrates. *Immunological reviews*, 167(1): p.59-67.

Flajnik, M.F., Kasahara, M., Shum, B.P., Salter-Cid, L., Taylor, E., and Du Pasquier, L. (1993). A novel type of class I gene organization in vertebrates: a large family of non-MHC-linked class I genes is expressed at the RNA level in the amphibian Xenopus. *EMBO Journal*, 12(11): p.4385–4396.

Foll, M., and Gaggiotti, O. (2008). A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics*, 180(2):p.977-993.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology,* 3:p.294–299..

Folstad, I., and Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *American Naturalist*, p.603-622.

Forsberg, L.A., Dannewitz, J., Petersson, E., and Grhan, M. (2007). Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout – females fishing for optimal MHC dissimilarity. *Journal of Evolutionary Biology*, 20:p.1859-1869.

Forterre, P., and Prangishvili, D. (2009). The origin of viruses. *Research in microbiology*, 160(7):p.466-472.

Frankel, O.H., (1970). Variation - the essence of life. *Proceedings of the Linnean Society*, 95(2):p.158-169.

Frankel, O.H., (1974). Genetic conservation: our evolutionary responsibility. *Genetics,* 78(1):p.53-65.

Frankel, O.H., and Soule, M. (1981). *Conservation and evolution.* Cambridge University Press, Cambridge, UK,

Frankham R., Bradshaw C. J. A. & Brook B.W. (2014). Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, 170: p.56–63.

Frankham, R. (1995). Annual Review of Genetics. *Conservation Genetics*, 29:p.305-327.

Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology,* 10(6):p.1500-1508.

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

Frankham, R. (2007). Genetic adaptation to captivity in species conservation programs. *Molecular Ecology,* 17(1):p.325–333.

Frankham, R. (1995). Conservation genetics. *Annual review of Genetics*, 29(1): p.305-327.

Frankham, R. (2005). Genetics and extinction. *Biological conservation*, 126(2): p.131-140.

Frankham, R. (2008). Genetic adaptation to captivity in species conservation programs. *Molecular Ecology*, 17(1): p.325-333.

Frankham, R. (1980). *founder effect and response to artificial selection in Drosophila*. In Selection experiments in laboratory and domestic animals: proceedings of a symposium held at Harrogate, UK, on 21st-22nd July 1979, edited by Alan Robertson. Slough, England, Commonwealth Agricultural Bureau, c1980..,

Frankham, R. (1995). Inbreeding and extinction: a threshold effect. *Conservation biology*, 9(4): p.792-799.

Frankham, R., 2009. *Genetic considerations in reintroduction programmes for top-order, terrestrial predators.* In Hayward, M.W., and Somers, M. (eds) Reintroduction of Top-Order Predators, Blackwell Publishing, Oxford, UK, 7:p.371.

Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, 24(11): p.2610-2618.

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological conservation*, 143(9): p.1919-1927.

Frankham, R. (2008). Genetic adaptation to captivity in species conservation programs. *Molecular Ecology,* 17(1):p. 325-333.

Frankham, R., Ballou, J.D., Briscoe, D.A. (2010). *Introduction to Conservation Genetics*, second ed. Cambridge University Press, Cambridge, UK.,

Frankham, R., Ballou, J.D., Eldridge, M.D., Lacy, R.C., Ralls, K., Dudash, M.R. and Fenster, C.B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology*, 25(3): p.465-475.

Frankham, R., Briscoe, D.A., & Ballou, J.D. (2002*). Introduction to conservation genetics*. Cambridge University Press.,

Frankham, R., Gilligan, D.M., Morris, D. and Briscoe, D.A. (2001). Inbreeding and extinction: effects of purging. Conservation *Genetics*, 2(3): p.279-284.

Fraser, D.J., Weir, L.K., Bernatchez, L., Hansen, M.M., and Taylor, E.B., (2011). Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity,* 106:p.404-420.

Freegard, C. and Richter, V. (2009). Animal Handling / Restraint using soft containment, Standard Operating Procedure 10.1. *Department of Environment and Conservation, Conservation Service Biodiversity.* Perth, Western Australia. ,

Freeland, W. J. (1976). Pathogens and the evolution of primate sociality. *Biotropica,* p.12-24.

Freeman, J.L., Feuk, L., Redon, R., McCarroll, S.A., Altschuler, D.M., Aburatani, H., Jones, K.W., Tyle-Smith, C., Hurles, M.E., Carter, N.P., Schere, S.W., and Lee, C. (2006). Copy number variation: New insights in genome diversity. *Genome Research*, 16:p.949-961.

Freeman-Gallant, C.R., Meguerdichian, M., Wheelwright, N.T. and Sollecito, S.V. (2003). Social pairing and female mating fidelity predicted by restriction fragment length polymorphism similarity at the major histocompatibility complex in a songbird. *Molecular Ecology,* 12(11),p. .3077-3083.

Freese, C.H., Aune, K.E., Boyd, D.P., Derr, J.N., Forrest, S.C., Gates, C.C., Gogan, P.J., Grassel, S.M., Halbert, N.D., Kunkel, K. and Redford, K.H. (2007). Second chance for the plains bison. *Biological Conservation,* 136(2):p.175-184.

Froeschke, G. and Sommer, S., (2005). MHC class II DRB variability and parasite load in the striped mouse (Rhabdomys pumilio) in the southern Kalahari. *Molecular biology and evolution*, 22(5), p.1254-1259.

Gabriel, W., Lynch, M., & Burger, R. (1991). *Population extinction by mutational load and demographic stochasticity*. In Species Conservation a Biological Approach, Birkhauser Verlag Basel, p..

Gabriel, W., Lynch, M., & Burger, R. (1993). Muller's ratchet and mutational meltdowns. *Evolution,* p.1744-1757.

Gagneux, P., Boesch, C., & Woodruff, D. S. (1997). Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular ecology*, 6(9): p.861-868.

Garnett, S.T., Olsen, P.D., Butchart, S.H.M., Hoffman, A.A. (2011). Did hybridization save the Norfolk Islandboobook owlNinox novaeseelandiae undulata? *Cambridge Journals*, 45(4):p.500–504.

Gautschi, B., Widmer, A., Joshi, J. and Koella, J.C. (2002). Increased frequency of scale anomalies and loss of genetic variation in serially bottlenecked populations of the dice snake, *Natrix tessellata*. *Conservation Genetics*, 3(3): p.235-245.

Germain, R.N. (1994). MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. *Cell,* 76(2):p.287-99.

Gibbons, P., Briggs, S.V., Ayers, D., Seddon, J., Doyle, S., Cosier, P., McElhinny, C., Pelly.V. and Roberts, K. (2009). An operational method to assess impacts of land clearing on terrestrial biodiversity. *Ecological Indicators*, 9:p. 26–40.

Gibson, G., & Dworkin, I. (2004). Uncovering cryptic genetic variation. *Nature Reviews Genetics*, 5(9): p.681-690.

Gilles, A., Meglécz, E., Pech, N., Ferreira, S., Malausa, T., and Martin, J.F. (2011). Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics*, 12(1):p.245.

Gilligan, M., and Frankham, R. (2003). Dynamics of genetic adaptation to captivity. *Conservation Genetics,* 4(2):p.189-197.

Gilpin, M. and Wills, C. (1991). MHC and captive breeding: a rebuttal. *Conservation Biology*, 5(4), p..554-555.

Gonzalo-Turpin, H., and Hazard, L. (2009). Local adaptation occurs along altitudinal gradient despite the existence of gene flow in the alpine plant species *Festuca eskia*. *Journal of Ecology*, 97(4):p.742-751.

Goodall, J. (2009). *Hope for animals and their world: how endangered species are being rescued from the brink*. Grand Central Publishing, New York City, New York.

Goodnight, K. (unknown). Selection 3.1 Avialable from https://strassmannandquellerlab.wordpress.com/home/teaching/,

Gorer, P.A. (1936). The Detection of Antigenic Differences in Mouse Erythrocytes by the Employment of Immune Sera. *British Journal of Experimental Pathology,* 17(1):p.42–50.

Goudet, J. (1995). FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, 86(6): p.485-486.

Gould, A. P. (1884). A case of spreading obliterative arteritis . *Trans Clin Soc London*, 17:p.95.

Gowaty, P.A., Karlin, A.A. (1984). Multiple maternity and paternity in single broods of apparently monogamous easter bluebirds (*Sialia Sialis*). *Behavioral Ecology and Sociobiology*, 15:p.91-95.

Goyache, F., Gutiérrez, J. P., Fernández, I., Gómez, E., Álvarez, I., Díez, J., & Royo, L. J. (2003). Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. *Journal of Animal Breeding and Genetics*, 120(2):p. 95-105.

Grieg, J.C. (1979). Principles of genetic conservation in relation to wildlife management in Southern Africa S. *Afr. Tydskr. Natuurnav*, 9(3/4):p.57-78.

Griffith, S.C., Owens, I.P. and Thuman, K.A. (2002). Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Molecular Ecology*, 11(11):p. .2195-2212.

Griffiths, R.A., and Pavajeau, L. (2008). Captive Breeding, Reintroduction, and the Conservation of Amphibians. *Conservation Biology*, 22(4):p.852–861.

Grimholt, U., Larsen, S., Nordmo, R., Midtlyng, P., Kjoeglum, S., Storset, A., Saebø, S. and Stet, R.J., (2003). MHC polymorphism and disease resistance in Atlantic salmon (Salmo salar); facing pathogens with single expressed major histocompatibility class I and class II loci. *Immunogenetics,* 55(4), p.210-219.

Grossberger, D. and Parham, P. (1992). Reptilian class I major histocompatibility complex genes reveal conserved elements in class I structure. *Immunogenetics*, 36(3):p.166-174.

Gu, X. and Nei, M. (1999). Locus specificity of polymorphic alleles and evolution by a birth-and-death process in mammalian MHC genes. *Molecular biology and evolution*, 16(2), p.147-156.

Gudbjartsson, D.F, Helgason, H., Gudjonsson, S.A, Zink, F., Oddson, A., Gylfason, A., Besenbacher, S., Magnusson, G., Halldorsson, B.V., Hjartarson, E., Sigurdsson, G.T., Stacey, S.N., Frigge, M.L., Holm, H., Saemundsdottir, J., Helgadottir, H.T., Johannsdottir, H., Sigfusson, G., Thorgeirsson, G., Sverrisson, G.T, Gretarsdottir, S., Walters, G.B, Rafnar, T., Thjodleifsson, B., Bjornsson, E.S., Olafsson, S., Thorarinsdottir, H., Steingrimsdottir, T., Gudmundsdottir, T.S, Theodors, A., Jonasson, J.G., Sigurdsson, A., Bjornsdottir, G., Jonsson, J.J., Thorarensen, O., Ludvigsson, P., Gudbjartsson, H., Eyjolfsson, G.I., Sigurdardottir, O., Olafsson, I., Arnar, D.O., Magnusson, O.T, Kong, A., Masson, G., Thorsteinsdottir, U., Helgason, A., Sulem, P. and Stefansson, K. (2015). Large-scale whole-genome sequencing of the Icelandic population. *Nature Genetics*, 47(5):p.435–444.

Gullapalli, R.R., Desai, K.V., Santana-Santos, L., Kant, J.A., Becich, M.J., (2012). Next generation sequencing in clinical medicine: Challenges and lessons for pathology and biomedical informatics. *Journal of Pathology Informatics*, 3:p.40.

Günther, T., and Coop, G. (2013). Robust Identification of Local Adaptation from Allele Frequencies. *Genetics*, 202(1):p.205-220.

Habel, J.C., Rodder, D., Schmitt, T., and Neves, G. (2011). Global warming will affect the genetic diversity and uniqueness of Lycaena helle populations. *Global Change Biology*, 17:p.194-205.

Hambuch, T.M., Lacey, E.A. (2007). Enhanced Selection for MHC Diversity in Social Tuco-Tucos. *Evolution,* 56(4):p.841-845.

Hamilton, W. D. (1980). Sex versus non-sex versus parasite. *Oik*os, p.282-290.

Hamrick, J.L., Godt, M.J.W., & Sherman-Broyles, S.L. (1992). *Factors influencing levels of genetic diversity in woody plant species. In Population genetics of forest trees* (pp. 95-124). Springer Netherlands.,

Hardman, B. and Moro, D. (2006). Optimising reintroduction success by delayed dispersal: is the release protocol important for hare-wallabies. *Biological Conservation*, 128: p.403-411.

Harris, H. and Hopkins, D. A. (1976). *Handbook of Enzyme Electrophoresis in Human Genetics*. North-Holland Publishing, Amsterdam,

Harris, P. C., Thomas, S.. Ratcliffe, P. J., Breuning, M. H., Coto, E., and Lopez-Larrea, C. (1991). Rapid genetic analysis of families with polycystic kidney disease 1 by means of a microsatellite marker. *Lancet,* 338: p.1484-1487.

Hauser, L., Adcock, G.J., Smith, P.J., Ramírez, J.H.B., & Carvalho, G.R. (2002). Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (Pagrus auratus). *Proceedings of the National Academy of Sciences*, 99(18): p.11742-11747.

Hayashi, K. and Yandell, D.W., (1993). How sensitive is PCR-SSCP? *Human mutation,* 2(5): p.338-346.

Hayden E. C. (2014). The $1000 genome. *Nature*, 507: p.294–295.

Hedrick P.W., and Fredrickson R.J. (2008). Captive breeding and the reintroduction of Mexican and red wolves. *Molecular Ecology*, 17:p.344–350.

Hedrick, P. W., and R. Fredrickson. (2010). Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. *Conservation Genetics*, 11:p.615–626.

Hedrick, P.W (1994). Purging inbreeding depression and the probability of extinction: full-sib mating. *Heredity,* 73:p.363-372.

Hedrick, P.W., & Fredrickson, R.J. (2008). Captive breeding and the reintroduction of Mexican and red wolves. *Molecular Ecology*, 17(1): p.344-350.

Heils, A., Teufel, A., Petri, S., Stöber, G., Riederer, P., Bengel, D., and Lesch, K.P. (1996). Allelic variation of human serotonin transporter gene expression. *Journal of Neurochemistry*, 66(6):p.2621-2624.

Henry, P., Caccone, A., and Miquelle, D., (2009). In situ population structure and ex situ representation of the Amur tiger. *Molecular Ecology*, 18:p.3173-3184.

Hewitt, E.R., (2003). The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology*, 110(2):p.163-169.

Ho, S. Y., Lanfear, R., Bromham, L., Phillips, M. J., Soubrier, J., Rodrigo, A. G., and Cooper, A. (2011). Time-dependent rates of molecular evolution. *Molecular ecology*, 20(15): p.3087-3101.

Hobbs, R. J., & Mooney, H. A. (1998). Broadening the extinction debate: population deletions and additions in California and Western Australia. *Conservation Biology*, 12(2):p. 271-283.

Hoekstra, H.E., Krenz, J.G., and Nachman, M.W. (2005). Local adaptation in the rock pocket mouse (Chaetodipus intermedius): natural selection and phylogenetic history of populations. *Heredity*, 94:p.217-228.

Hoekzema, K. and Sidlauskas, B.L., (2014). Molecular phylogenetics and microsatellite analysis reveal cryptic species of speckled dace (*Cyprinidae: Rhinichthys osculus*) in Oregon's Great Basin. *Molecular Phylogenetics and Evolution*, 77,p. .238-250.

HOELZEL, A. (1999). Impact of population bottlenecks on genetic variation and the importance of life-history; a case study of the northern elephant seal. *Biological Journal of the Linnean Society*, 68(1-2): p.23-39.

Hoggard, S.J., Wilson, P.D., Beattie, A.J. and Stow, A.J., (2013). The effectiveness of antimicrobial defenses declines with increasing group size and genetic similarity. *Annals of the Entomological Society of America,* 106(1):p.53-58.

Hoggard, S. J., Wilson, P. D., Beattie, A. J., & Stow, A. J. (2011). Social complexity and nesting habits are factors in the evolution of antimicrobial defences in wasps. *PloS one*, 6(7): p.e21763.

Höglund, J., Piertney, S.B., Alatalo, R.V., Lindell, J., Lundberg, A., & Rintamäki, P.T. (2002). Inbreeding depression and male fitness in black grouse. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1492): p.711-715.

Hohenlohe, P.A., Amish, S.J., Catchen, J.M., Allendorf, F.W. and Luikart, G., (2011). Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular ecology resources*, 11(s1):p. 117-122.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A., and Cresko, W. (2010). Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *PLOS Genetics*, 6(2):p.e1000862..

Hokanson, S.C., Szewc-McFadden, A.K., Lamboy, W.F., & McFerson, J.R. (1998). Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a Malus× domestica Borkh. core subset collection. *Theoretical and Applied Genetics*, 97(5-6): p.671-683.

Holsinger, K.E., & Gottlieb, L.D. (1991). Conservation of rare and endangered plants: principles and prospects. Genetics and conservation of rare plants. *Oxford University Press, New York*, p.195-208.

Hordvik, I., Grimholt, U., Fosse, V.M., Lie, O., and Endresen, C. (1993). Cloning and sequence analysis of cDNAs encoding the MHC class II β chain in Atlantic salmon (Salmo salar). *Immunogenetics,* 37(6):p.437-441.

Horton, R., Wilming, L., Rand, V., Lovering, R.C., Bruford, E.A., Khodiyar, V.K., Lush, M.J., Povey, S., Talbot, C.C., Wright M.W., Wain, H.M., Trowsdale, J., Ziegler, A., and Beck, S. (2004). Gene map of the extended human MHC. *Nature Reviews Genetics*, 5:p.889-899.

Hosomichi, K., Miller, M.M., Goto, R.M., Wang, Y., Suzuki, S., Kulski, J.K., Nishibori, M., Inoko, H., Hanzawa, K. and Shiina, T., 2008. (2008). Contribution of mutation, recombination, and gene conversion to chicken MHC-B haplotype diversity. *The Journal of Immunology*, 181(5):p.3393-3399.

Hosomichi, K., Shiina, T., Suzuki, S., Tanaka, M., Shimizu, S., Iwamoto, S., Hara, H., Yoshida, Y., Kulski, J.K., Inoko, H. and Hanzawa, K., (2006). The major histocompatibility complex (Mhc) class IIB region has greater genomic structural flexibility and diversity in the quail than the chicken. *BMC genomics,* 7(1):p.1.

Hostetler, J.A., Onorato, D.P., Nichols, J.D., Johnson, W.E., Roelke, M.E., O'Brien, S.J., Jansen, D. and Oli, M.K., (2010). Genetic introgression and the survival of Florida panther kittens. *Biological Conservation,* 143(11): p.2789-2796.

Howard, R. D., and Minchella, D. J. (1990). Parasitism and mate competition. *Oikos,* , p.120-122.

Huchard, E., Baniel, A., Schliehe-Diecks, S. and Kappeler, P.M. (2013). MHC-disassortative mate choice and inbreeding avoidance in a solitary primate. *Molecular ecology*, 22(15),p. .4071-4086.

Hughes, A.L., 1991. MHC polymorphism and the design of captive breeding programs. *Conservation Biology,* 5(2), p..249-251.

Hughes, C. R., & Queller, D. C. (1993). Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology*, 2(3):p. 131-137.

Hughes, W. O., Eilenberg, J., and Boomsma, J. J. (2002). Trade-offs in group living: transmission and disease resistance in leaf-cutting ants. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1502): p.1811-1819.

Human Genetics Commission. (2009). *Nothing to hide, nothing to fear Department of Health*. Whitehall, London

Huo, Y.J., Wan, X.R., Wolff, J.O., Wang, G., Thomas, S., Iglay, R.B., Leopold, B.D. and Liu, W. (2010). Multiple paternities increase genetic diversity of offspring in Brandt's voles. *Behavioural Processes*, 84(3), p..745-749..

Huse, S.M., Welch, D.M., Morrison, H.G., and Sogin, M.L. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology*, 12(7):p.1889-1898.

IBM Corp (2012). IBM SPSS Statistics for Windows, Version 210 Armonk, NY: IBM Corp

Ihakaa, R., and Gentlemen, R. (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, 5(3):p.299-314.

Ingvarsson, P.K. (2001). Ingvarsson PK. Restoration of genetic variation lost - the genetic rescue hypothesis. *Trends in Ecology & Evolution*: 16(2):p.62-63.

IUCN (2001). 2001 Categories & Criteria (version 3.1) IUCN website (http://www.iucnredlist.org/ accessed Dec 2015) Huntingdon Road, Cambridge

IUCN (2015). The IUCN Red List of Threatened Species. Version 2015-4.

James, A., Gaston, K.J. and Balmford, A. (2001). Can we afford to conserve biodiversity? *BioScience,* 51(1):p..43-52..

Jamieson, I.G., (2011). Founder effects, inbreeding, and loss of genetic diversity in four avian reintroduction programs. *Conservation Biology*, 25(1): p.115-123.

Janeway, C.A., Travers, P., Walport, M., and Shlomchik, M.J. (2005). *Immunobiology: The Immune System in Health and Disease*. 5th edition. Garland Science New York, NY.

Jaratlerdsiri, W., Deakin, J., Godinez, R.M., Shan, X., Peterson, D.G., Marthey, S., Lyons, E., McCarthy, F.M., Isberg, S.R., Higgins, D.P., Chong, A.Y., St John, J., Glenn, T.C., Ray, D.A., and Gongora, J. (2014). Comparative Genome Analyses Reveal Distinct Structure in the Saltwater Crocodile MHC. *PLoS ONE* , 9(12): p.e114631.

Jarne, P., and Lagoda, P.J.L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution,* 11(10):p.424-429.

Jiang, G. (2013). Plant Marker-Assisted Breeding and Conventional Breeding*: Challenges and Perspectives Advances in Crop Science and Technology*, 1(3):

Johnson, C. (2006). *Australia's mammal extinctions: a 50,000-year history*. Cambridge University Press.,

Johnson, C. N., & Isaac, J. L. (2009). Body mass and extinction risk in Australian marsupials: the 'critical weight range' revisited. *Austral Ecology*, 34(1): p.35-40.

Johnson, D.W., Maruyama, G., Johnson, R., Nelson, D. and Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological bulletin*, 89(1): p.47.

Johnston, V.H. and Ryder, J.P. (1987). Divorce in larids: a review. *Colonial Waterbirds*, p.16-26.

Jones, A.G., & Avise, J.C. (2001). Mating systems and sexual selection in male-pregnant pipefishes and seahorses: insights from microsatellite-based studies of maternity. *Journal of Heredity*, 92(2):p. 150-158.

Joost, S., Bonin, A., Bruford, M.W., Despres, L., Conord, C., Erhardt, G., and Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, 16(18):p.3955–3969.

Jordan, W.C. and Bruford, M.W., (1998). New perspectives on mate choice and the MHC. *Heredity*, 81(2), p..127-133.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., Anderson, C.A. and Essers, J., (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*, 491(7422): p.119-124.

Jule, K.R., Leaver, L.A., and Lea, S.E.G. (2008). The effects of captive experience on reintroduction survival in carnivores: A review and analysis. *Biological Conservation*, 141:p.355-363.

Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A. and Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Molecular ecology*, 23(21): p.5151-5163.

Kanehisa, M., and Bork, P. (2003). Bioinformatics in the post-sequence era. *Nature Genetics*, 33:p.305-310.

Kasahara, M., Flajnik, M. F., Ishibashi, T., and Natori, T. (1995). Evolution of the major histocompatibility complex: a current overview. *Transplant immunology*, 3(1):p. 1-20.

Kawecki, T.J. and Ebert, D., (2004). Conceptual issues in local adaptation. *Ecology letters*, 7(12), p..1225-1241.

Kashi, Y., and King, D.G., (2006). Simple sequence repeats as advantageous mutators in evolution. *TRENDS in Genetics*, 22(5): p.253-259.

Kelley, J., Walter, L., and Trowsdale J. (2005). Comparative genomics of major histocompatibility complexes. *Immunogenetics,* 56(10): p.683-95.

Kelly, M.J. (2001). Lineage loss in Serengeti cheetahs: consequences of high reproductive variance and heritability of fitness on effective population size. *Conservation Biology*, 15(1):p.137-147.

Kimbrell, D. A., and Beutler, B. (2001). The evolution and genetics of innate immunity. *Nature Reviews Genetics,* 2(4): p.256-267.

Kimura, M. and Ohta, T. (1969). The average number of generations until fixation of a mutant gene in a finite population. *Genetics,* 61(3), p.0.763.

Kimura, M., (1968). Evolutionary rate at the molecular level. *Nature*, 217(5129), p..624-626.

Kjeldsen, S.R., Zenger, K.R., Leigh, K., Ellis, W., Tobey, J., Phalen, D., Melzer, A., FitzGibbon, S., and Raadsma, H.W. (2015). Genome-wide SNP loci reveal novel insights into koala (Phascolarctos cinereus) population variability across its range. *Conservation Genetics*, p.1-17.

Klein, J. and Figueroa, F. (1986). The evolution of class I MHC genes. *Immunology today*, 7(2):p. 41-44.

Klein, S. L. (2003). Parasite manipulation of the proximate mechanisms that mediate social behavior in vertebrates. *Physiology and Behavior*, 79(3): p.441-449.

Knafler, G. J., Clark, J. A., Boersma, P. D., and Bouzat, J. L. (2012). MHC diversity and mate choice in the magellanic penguin, Spheniscus magellanicus. Journal of Heredity,, p. ess054.

Knapp, E. E., & Dyer, A. R. (1998). *When do genetic considerations require special approaches to ecological restoration?.* In Conservation Biology (pp. 345-363). Springer US.,

Kobayashi, K.S., & van den Elsen, P.J. (2012). NLRC5: a key regulator of MHC class I-dependent immune responses. *Nature Reviews Immunology*, 12(12): p.813-820.

Kokita, T. and Nakazono, A., (1998). Plasticity in the mating system of the longnose filefish, Oxymonacanthus longirostris, in relation to mate availability. *Journal of Ethology*, 16(2), p..81-89.

Kokko, H., Brooks, R., McNamara, J.M. and Houston, A.I. (2002). The sexual selection continuum. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1498), p.1331-1340.

Kostia, S., Kantanen, J., Kolkkala, M., and Varvio. S.L. (1998). Applicability of SSCP analysis for MHC genotyping: fingerprinting of Ovar-DRB1 exon 2 alleles from Finnish and Russian breeds. *Animal Genetics,* 29(6):p.453–455.

"Krebs, R.A., and Feder, M.E. (1997). Natural Variation in the Expression of the Heat-Shock Protein HSP70 in a Population of Drosophila melanogaster and Its Correlation with Tolerance of Ecologically Relevant Thermal Stress. *Evolution*, 51(1):p.173-179.

Kubinak, J. L., Ruff, J. S., Hyzer, C. W., Slev, P. R., and Potts, W. K. (2012). Experimental viral evolution to specific host MHC genotypes reveals fitness and virulence trade-offs in alternative MHC types. *Proceedings of the National Academy of Sciences*, 109(9): p.3422-3427.

Kuehn, R., Schroeder, W., Pirchner, F. and Rottmann, O. (2003). Genetic diversity, gene flow and drift in Bavarian red deer populations (*Cervus elaphus*). *Conservation Genetics*, 4(2): p.157-166.

Kulski, J.K., Shiina, T., Anzai, T., Kohara, S. and Inoko, H., (2002). Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunological reviews*, 190(1):p.95-122.

Kumar, H., Kawai, T., and Akira, S. (2011). Pathogen recognition by the innate immune system. *International reviews of immunology*, 30(1): p.16-34.

Lack, D. (1968). Ecological adaptations for breeding in birds. Methuen, London, UK

Lacy, R. C. (1987). Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology*, 1(2): p.143-158.

Lacy, R. C. (1989). Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. *Zoo Biology*, 8(2): p.111-123.

Lacy, R.C. (1987). Loss of Genetic Diversity from Managed Populations: Interacting Effects of Drift, Mutation, Immigration, Selection, and Population Subdivision. *Conservation Biology*, 1(2):p.143-158.

Laikre, L., & Ryman, N. (1991). Inbreeding depression in a captive wolf (*Canis lupus*) population. *Conservation Biology*, 5(1): p.33-40.

Laikre, L., Allendorf, F.W., Aroner, L.C., Baker, C.S., Gregovich, D.P., Hansen, M.M., Jackson, J.A., Kendall, K.C., Mckelvey, K., Neel, M.C. and Olivieri, I. (2009). Neglect of genetic diversity in implementation of the convention on biological diversity. *Conservation Biology*, 24(1):p.86-88.

Lande, R., & Schemske, D.W. (1985). The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution*, p.24-40.

Lange, A., and Ferguson, N. M. (2009). Antigenic diversity, transmission mechanisms, and the evolution of pathogens. *PLoS computational biology*, 5(10): p.e1000536.

Langford, D., (1999). *The mala (Lagorchestes hirsutus) recovery plan*. Parks and Wildlife Commission of the Northern Territory, Arid Zone Research Institute.,

Langford, D., & Burbidge, A. A. (2001). Translocation of mala (Lagorchestes hirsutus) from the Tanami desert, Northern Territory to Trimouille island, Western Australia. *Australian Mammalogy*, 23(1):p. 37-46.

Latter, B.D.H., and Mulley, J.C. (1994). Genetic Adaptation to Captivity and Inbreeding Depression in Small Laboratory Populations of Drosophila melanogaster. *Genetics,* 139:p.255-266.

Laugen, A.T., Laurila, A., Räsänen, K. and Merilä, J. (2003). Latitudinal countergradient variation in the common frog (Rana temporaria) development rates–evidence for local adaptation. *Journal of evolutionary biology*, 16(5), p..996-1005.

Laurance, W.F., McDonald, K.R. and Speare, R. (1996). Epidemic disease and the catastrophic decline of Australian rain forest frogs. *Conservation Biology*, 10(2): p.406-413.

Laurenson, K., Sillero-Zubiri, C., Thompson, H., Shiferaw, F., Thirgood, S. and Malcolm, J. (1998). Disease as a threat to endangered species: Ethiopian wolves, domestic dogs and canine pathogens. *Animal Conservation*, 1(4):p. 273-280.

Laurila, J.K., Gatica-Perez, D., Aad, I., Blom, J., Bornet, O., Trinh-Minh-Tri Do, Dousse, O., Eberle, J., and Miettinen, M. (2012). *The Mobile Data Challenge: Big Data for Mobile Computing Research.* In Pervasive Computing No. EPFL-CONF-192489.,

Lemmon, A.R., Emme, S.A. and Lemmon, E.M., (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, p.sys049.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *TRENDS in Ecology & Evolution*, 17(4):p.183-189.

Lerp, H., Plath, M., Wronski, T., Bärmann, E.V., Malczyk, A., Resch, R.R., ...& Pfenninger, M. (2014). Utility of island populations in re-introduction programmes–relationships between Arabian gazelles (Gazella arabica) from the Farasan Archipelago and endangered mainland populations. *Molecular Ecology*, 23(8):p. 1910-1922.

Levene, H. (1953). Genetic Equilibrium When More Than One Ecological Niche is Available. *The American Naturalist*, 87(836):p.331-333.

Lewontin, R.C., and Hubby J.L. (1966). A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura. *Genetics*, 54(2):p.595-609.

Li, R., Fan, W., Tian, G., Zhu, H., He, L., Cai, J., Huang, Q., Cai, Q., Li, B., Bai, Y., Zhang, Z., Zhang, Y., Wang, W., Li, J., Wei, F., Li, H., Jian, W., Li, J., Zhang, Z., Nielsen, R., Li, D., Gu, W., Yang, Z., Xuan, Z., Ryder, O.A., Leung, C.F., Zhou, Y., Cao, J., Sun, X., Fu, Y., Fang, X., Guo, X., Wang, B., Hou, R., Shen, F., Mu, B., Ni, P., Lin, R., Qian, W., Wang, G., Yu, C., Nie, W., Wang, J., Wu, Z., Lian, H., Min, J., Wu, Q., Cheng, S., Ruan, J., Wang, M., Shi, Z., Wen, M., Liu, B., Ren, X., Zheng, H., Dong, D., Cook, K., Shan, G., Zhang, H., Kosiol, C., Xie, X., Lu, Z., Zheng, H., Li, Y., Steiner, C.C., Lam, T.T., Lin, S., Zhang, Q., Li, G., Tian, J., Gong, T., Liu, H., Zhang, D., Fang, L., Ye, C., Zhang, J., Shi, Y., Li, Z., Liu, Q., Chen, Y., Zhao, J., Qu, N., Zhao, S., Tian, F., Wang, X., Wang, H., Xu, L., Xia, L., Vinar, T., Wang, Y., Lam, T.W., Yiu, S.M., Liu, S., Zhang, H., Li, D., Huang, Y., Wang, X., Yang, G., Jiang, Z, Wang, J., Qin, N., Li, L., Li, J., Bolund, L., Kristiansen, K., Wong, G.K.S., Olson, M., Zhang, X., Li, S., Yang, H., Wang, J., and Wang, J. (2010). The sequence and de novo assembly of the giant panda genome. *Nature,* 463:p.311-317.

Lie, H.C., Simmons, L.W. and Rhodes, G. (2009). Does genetic diversity predict health in humans?. *PloS one*, 4(7):p.e6391.

Lippe, C., Dumon, P., and Bernatchez, L. (2006). High genetic diversity and no inbreeding in the endangered copper redhorse, Moxostoma hubbsi (Catostomidae, Pisces): the positive sides of a long generation time. *Molecular Ecology*, p.1769–1780.

Lips, K. R., Brem, F., Brenes, R., Reeve, J. D., Alford, R. A., Voyles, J., ... and Collins, J. P. (2006). Emerging infectious disease and the loss of biodiversity in a Neotropical amphibian community. *Proceedings of the national academy of sciences of the United States of America*, 103(9): p.3165-3170.

Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., James, S.A., Davey, R.P., Roberts, I.N., Burt, A., Koufopanou, V., Tsai, I.J., Bergman, C.M., Bensasson, D., O'Kelly, M.J.T., Oudenaarden, A.V., Barton, D.B.H., Bailes, E., Nguyen, A.N., Jones, M., Quail, M.A., Goodhead, I., Sims, S., Smith, F., Blomberg, A., Durbin, R., and Louis, E.J., (2009). Population genomics of domestic and wild yeasts. *Nature*, 458:p.337-341.

Liu, Li, Y,Hu, N, He, Y Pong, R, Lin, D, Lu, L & Law M. (2012). Comparison of Next-Generation Sequencing Systems. *Journal of Biomedicine and Biotechnology*, 2012:p.1-11.

Loke, Y. W. (2013*). Life's Vital Link*: *The astonishing role of the placenta*. Oxford University Press.,

Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J., and Pallen, M.J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology,* p.434-439.

Loreau, M.*et al.* (2001). Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*, 294(p.804-.

Lotterhos, K.E. and Whitlock, M.C., (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests*. Molecular ecology*, 23(9): p.2178-2192.

Luikart, G., Pilgrim, K., Visty, J., Ezenwa, V.O. and Schwartz, M.K. (2008). Candidate gene microsatellite variation is associated with parasitism in wild bighorn sheep. *Biology Letters*, 4(2),p. .228-231.

Lukas, D. and Clutton-Brock, T.H. (2013). The evolution of social monogamy in mammals. *Science*, 341(6145):p. .526-530.

Lundie-Jenkins, G. W. (1989). *The Ecology and Management of the Rufous Hare-wallaby (Lagorchestes Hirsutus) in the Tanami Desert.* University of New England.,

Lunney, D., Hutchings, P., and Hochuli, D. (2009). *The Natural History of Sydney*. Royal Zoological Society of NSW, Mosman, NSW,

Lynch, C. (2008). Big data: How do your data grow? *Nature,* 455:p.28-29.

Lynch, M., and Conery, J. S. (2003). The origins of genome complexity. *Science*, 302(5649): p.1401-1404.

Mace, G.M., and Lande, R. (1991). Assessing extinction threats – toward a reevaluation of IUCN Threatened Species categories *Conservation Biology*, 5(2): p.148-157.

Madsen, T., Shine, R., Olsson, M. and Wittzell, H. (1999). Conservation biology: restoration of an inbred adder population. *Nature*, 402(6757): p.34-35.

Mala Reintroduction Fact Sheet. Australian Governement & Uluru Kata Juta National Park. (Date Unknown). Fact Sheet - Uluru – Kata Tjuta National Park Mala Reintroduction Project (accessed at http://olr.npi.gov.au/parks/publications/uluru/pubs/factsheet-mala.pdf)

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, 20(5): p.229-237.

Mardis, E.R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, 24(3):p.133-141.

Mardis, E.R. (2011). A decade's perspective on DNA sequencing technology *Nature,* 470:p.198–203.

Margaritopoulos, J.T., Kasprowicz, L., Malloch, G.L., & Fenton, B. (2009). Tracking the global dispersal of a cosmopolitan insect pest, the peach potato aphid. *BMC ecology*, 9(1): p.13.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z. and Dewell, S.B., (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* 437(7057), p.pp.376-380..

Margulies, M., Egholm., M., Altman., W.E, Attiya, S., Bader, J.S, Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P, Myers, E.W, Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T, Roth, G.T., Sarkis, G.J., Simons, J.F, Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A, Volkmer, G.A., Wang, S.H, Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057):p.376-80.

Marshall T C & Spalton J A (2000). Simultaneous inbreeding and outbreeding depression in reintroduced Arabian oryx. *Animal Conservation*, 3, p.241-248.

Martin, O.Y. and Hosken, D.J., (2003). Costs and benefits of evolving under experimentally enforced polyandry or monogamy. *Evolution*, 57(12), p..2765-2772.

Martin, T.G., Nally, S., Burbidge, A.A., Arnall, S., Garnett, S.T., Hayward, M.W., Lumsden, L.F., Menkhorst, P., McDonald-Madden, E., and Possingham, H.P. (2012). Acting fast helps avoid extinction. *Conservation Letters*, 5(4):p.274–280.

Mather, K., (1955). Polymorphism as an Outcome of Disruptive Selection. *Evolution*, 9(1):p.52-61.

Maudet, C., Miller, C., Bassano, B., Breitenmoser-Würsten, C., Gauthier, D., Obexer-Ruff, G., ...& Luikart, G. (2002). Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [Capra ibex (ibex)]. *Molecular Ecology*, 11(3): p.421-436.

Maxwell, S., Burbidge, A.A., Morris, K., (Eds), (1996). *The 1996 Action Plan for Australian Marsupials and Monotremes*. Wildlife Australia: Canberra,

Maynard Smith, J. (1962). Disruptive selection, polymorphism and sympatric speciation. *Nature*, 195:p.60-62.

Mays Jr, H.L., Albrecht, T., Liu, M. and Hill, G.E. (2008). Female choice for genetic complementarity in birds: a review. *Genetica*, 134(1):p. .147-158.

McCabe, E. R., Huang, S. Z., Seltzer, W. K., & Law, M. L. (1987). DNA microextraction from dried blood spots on filter paper blotters: potential applications to newborn screening. *Human genetics*, 75(3): p.213-216.

McCabe, E.R.B. (1991). Utility of PCR for DNA Analysis from Dried Blood Spots on Filter Paper Blotters. *PCR Methods and Applications,* p.99-106.

McCarthy, D.P., Donald, P.F., Scharlemann, J.P.W., Buchanan, G.M., Balmford, A., Green, J.M.H., Bennun, L.A., Burgess, N.D., Fishpool, L.D.C., Garnett, S.T., Leonard, D.L., Maloney, R.F., Morling, P, Schaefer, H.M., Symes, A., Wiedenfeld, D.A., and Butchart, S.H.M. (2012). Financial Costs of Meeting Global Biodiversity Conservation Targets: Current Spending and Unmet Needs. *Science*, 338(6109):p.946-949.

McDonald, B. A., and Linde, C. (2002). Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, 40(1): p.349-379.

McKay, J.K., Bishop, J.G., Lin, J.Z., Richards, J.H., Sala, A., and Olds, T.M. (2001). Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rokress. Proceedings of *The Royal Society of London B: Biological Sciences* 268(1477):p.1715-1721.

McKay, J.K., Christian, C.E., Harrison, S., and Rice, K.J. (2005). ‘‘How Local Is Local?’’—A Review of Practical and Conceptual Issues in the Genetics of Restoration. *Restoration Ecology*, 13(3):p.432–440 .

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A., (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20:p.1297-1303.

Meagher, S., and Potts, W.K. (1997). A microsatellite-based MHC genotyping system for house mice (Mus domesticus). *Hereditas*, 127:p.75-82.

Menotti-Raymond, M., & O'Brien, S.J. (1993). Dating the genetic bottleneck of the African cheetah. *Proceedings of the National Academy of Sciences*, 90(8): p.3172-3176.

Miller, H.C., Belov, K., and Daugherty, C.H. (2005). Characterization of MHC class II genes from an ancient reptile lineage, Sphenodon (tuatara). *Immunogenetics*, 57(11):p.883-891.

Miller, J.R., Koren, S., Sutton, G., (2010). Assembly algorithms for next-generation sequencing data. *Genomics*, 95(6):p.315-327.

Miller, M. P. (1997). *Tools for population genetic analyses (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data.* Computer software distributed by author, 4, 157.,

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., and Johnson, E.A., (2006). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, 17:p.240-248.

Miller, P.S. and Hedrick, P.W., (1991). MHC polymorphism and the design of captive breeding programs: simple solutions are not the answer. *Conservation Biology*, 5(4):p. .556-558.

Mills, L.S., Allendorf, F.W. (1996). The One-Migrant-per-Generation Rule in Conservation and Management. *Conservation Biology*, 10(6):p.1509–1518.

Møller, A. P., Merino, S., Brown, C. R., and Robertson, R. J. (2001). Immune defense and host sociality: a comparative study of swallows and martins. *The American Naturalist*, 158(2): p.136-145.

Moller, A.P. and Alatalo, R. V. 1 (1999). Good genes effects in sexual selection. *Proceedings of the Royal Society of London Series B Biological Sciences*, 266:p.85-91..

Montalvo, A.M., and Ellstrand, N.C. (2001). Nonlocal transplantation and outbreeding depression in the subshrub Lotus scoparius (*Fabaceae). American Journal of Botany*, 88(2):p.258-69.

Moore, S.S., Sargeant, L.L., King, T.J., Mattick, J.S., Georges, M., & Hetzel, D.J.S. (1991). The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species*. Genomics*, 10(3): p.654-660.

Morgan, U.M., Pallant, L., Dwyer, B.W., Forbes, D.A., Rich, G., and Thompson, R.C.A. (1998). Comparison of PCR and Microscopy for Detection of *Cryptosporidium parvum* in Human Fecal Specimens: *Clinical Trial Journal of Clinical Microbiology*, 38(4):p.995-998.

Morley, J.I. and Balshine, S., (2002). Faithful fish: territory and mate defence favour monogamy in an African cichlid fish. *Behavioral Ecology and Sociobiology*, 52(4):p.326-331.

Moroney, J. R., & Cooper, M. B. (1982). *Environmental radiation at the Monte Bello Islands from nuclear weapons tests conducted in 1952 and 1956*. Australian Radiation Lab..,

Morozova, O., and Marra, M.A., (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics*, 92(5):p.255-264.

Moseby, K. E., Hill, B. M., & Read, J. L. (2009). Arid Recovery–A comparison of reptile and small mammal populations inside and outside a large rabbit, cat and fox-proof exclosure in arid South Australia. *Austral Ecology,* 34(2): p.156-169.

Moxon, E. R., Rainey, P. B., Nowak, M. A., and Lenski, R. E. (1994). Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current Biology*, 4(1): p.24-33.

Muehlenbein, M.P. (2010). *Human evolutionary biology*. Cambridge University Press, UK,

Mullen.M.P., Howard, D.J., Powell, R., and Hanrahan, J.P., (2009). A note on the use of FTA™ technology for storage of blood samples for DNA analysis and removal of PCR inhibitors. *Irish Journal of Agricultural and Food Research*, 48:p.109-113.

Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R., and Hohenlohe, P.A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, 22(11):p. 2841–2847.

Nederbragt, A.J. (2014). On the middle ground between open source and commercial software - the case of the Newbler program. *Genome Biology*, 15(4):p.1-2.

Nei, M. (1977). F-statistics and analysis of gene diversity in subdivided populations. *Annals of Human Genetics*, 41(2):p.225-233.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89(3), p.583-590.

Ng, D.P.K., Koh, D., Choo, S.G.L., Ng.V., and Fu, Q. (2004). Effect of storage conditions on the extraction of PCR-quality genomic DNA from saliva. *Clinica Chimica Acta*, 343(1-2):p.191-194.

NHGRI (National Human Genome Research Institute) (2010). The Human Genome Project Completion: Frequently Asked Questions Available from (http://guides.is.uwa.edu.au/c.php?g=324809&p=2178312) (viewed Dec 2015),

Nielsen, E.E., Hemmer-Hansen, J., Poulsen, N.A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G.L., Ogden, R. and Carvalho, G.R. (2009). Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (Gadus morhua). *BMC Evolutionary Biology*, 9(1), p.0.276.

Nielsen, R., Paul, J.S., Albrechtsen, A., and Song, Y.S., (2011). Genotype and SNP calling from next generation sequencing data. *Nature Reviews Genetics*, 12:p.443-451.

Nieminen, M., Singer, M.C., Fortelius, W., Schöps, K., & Hanski, I. (2001). Experimental confirmation that inbreeding depression increases extinction risk in butterfly populations. *The American Naturalist*, 157(2): p.237-244.

Nims, B.D., Vargas, H., Merkel, J., and Parker, P.G. (2008). Low genetic diversity and lack of population structure in the endangered Gala´pagos penguin (Spheniscus mendiculus). *Conservation Genetics,* 9:p.1413-1420.

Niskanen, A.K., Kennedy, L.J., Ruokonen, M., Kojola, I., Lohi, H., Isomursu, M., Jansson, E., Pyhäjärvi, T. and Aspi, J., (2014). Balancing selection and heterozygote advantage in major histocompatibility complex loci of the bottlenecked Finnish wolf population. *Molecular Ecology*, 23(4), p..875-889.

O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W., & Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation*, 133(1): p.42-51.

Ofori, B. Y., Beaumont, L. J., & Stow, A. J. (2017). Cunningham's skinks show low genetic connectivity and signatures of divergent selection across its distribution. *Ecology and Evolution*, 7(1): 48–57.

Oliver, M.K., Lambin, X., Cornulier, T. and Piertney, S.B., (2009). Spatio-temporal variation in the strength and mode of selection acting on major histocompatibility complex diversity in water vole (Arvicola terrestris) metapopulations. *Molecular Ecology,* 18(1), p..80-92.

Olsson, M., Madsen, T., Nordby, J., Wapstra, E., Ujvari, B. and Wittsell, H., (2003). Major histocompatibility complex and mate choice in sand lizards. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(Suppl 2),p. .S254-S256.

Opie, C., Atkinson, Q.D., Dunbar, R.I. and Shultz, S., (2013). Male infanticide leads to social monogamy in primates. *Proceedings of the National Academy of Sciences*, 110(33): p.13328-13332.

Ouborgemail, N.J., Pertoldi, C., Loeschcke, V., Bijlsma, R., and Hedrick, P.W. (2010). Conservation genetics in transition to conservation genomics. *Trends in Genetics*, 26(4):p.177–187.

Palenik, B., Brahamsha, B., Larimer, F.W., Land, M., Hauser, L., Chain, P., Lamerdin, J., Regala, Q., Allen,E.E., McCarren, J., Paulsen, I., Dufresne, A., Partensky,F., Webb,E.A., and Waterbury, J. (2003). The genome of a motile marine Synechococcus. *Nature*, 424:p.1037-1042.

Palombit, R.A. (1999). Infanticide and the evolution of pair bonds in nonhuman primates. *Evolutionary Anthropology Issues News and Reviews*, 7(4),p. .117-129.

Paradis E., Claude J. & Strimmer K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20:p.289-290.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, p. 637-669.

Parrott, M.L., Ward, S.J. and Temple-Smith, P.D., (2007). Olfactory cues, genetic relatedness and female mate choice in the agile antechinus (Antechinus agilis). *Behavioral Ecology and Sociobiology*, , 61(7), p..1075-1079.

Partecke, J. and Gwinner, E., (2007). Increased sedentariness in European Blackbirds following urbanization: a consequence of local adaptation? *Ecology,* 88(4), p.882-890.

Pavey, C. (2006). *Threatened Species of the Northern Territory-Mala, Rufous Hare-wallaby* *Lagorchestes hirsutus* (*central mainland form*). Department of Land and Resource Management, NT,

Peakall, R. and Smouse P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. *Bioinformatics*, 28:p.2537-2539.

Peng, B., and Kimmel, M. (2005). simuPOP: a forward-time population genetics simulation environment. *Bioinformatics*, 21 (18): p.3686-3687.

Penn, D. J., Damjanovich, K., and Potts, W. K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences*, 99(17): p.11260-11264.

Penn, D., and Potts, W. K. (1998). Chemical signals and parasite-mediated sexual selection. *Trends in Ecology and Evolution,* 13(10):p.391-396.

Peres, C.A. (2001). Synergistic Effects of Subsistence Hunting and Habitat Fragmentation on Amazonian Forest Vertebrates. *Conservation Biology*, 15(6):p.1490-1505.

Peters, B.A., Kermani, B.G., Sparks, A.B., Alferov, O., Hong, P., Alexeev, A., Jiang, Y., Dahl, F., Tang, Y.T., Haas, J., Robasky, K., Zaranek, A.W., Lee, J.H., Ball, M.P., Peterson, J.E., Perazich, H., Yeung, G., Liu, J., Chen, L., Kennemer, M.I., Pothuraju, K., Konvicka, K., Tsoupko-Sitnikov, M., Pant, K.P., Ebert, J.C., Nilsen, G.B., Baccash, J., Halpern, A.L., Church, G.M., and Drmanac, R. (2012). Accurate whole-genome sequencing and haplotyping from 10 to 20 human cells. *Nature,* 487:p.190-195.

Philippi, T. and Seger, J., (1989). Hedging one's evolutionary bets, revisited. *Trends in Ecology & Evolution,* 4(2), p.41-44.

Phillimore, A.B., Hadfield, J.D., Jones, O.R. and Smithers, R.J. (2010). Differences in spawning date between populations of common frog reveal local adaptation. *Proceedings of the National Academy of Sciences,* 107(18): p..8292-8297.

Pimm, S. L., Russell, G. J., Gittleman, J. L., & Brooks, T. M. (1995). The future of biodiversity. *Science*, 269(5222):p.347.

Piper, W.H., Evers, D.C., Meyer, M.W., Tischler, K.B., Kaplan, J.D. and Fleischer, R.C., (1997). Genetic monogamy in the common loon (Gavia immer). *Behavioral Ecology and Sociobiology*, 41(1), p.25-31.

Pitcher, T.E., and Neff, B.D. (2006). MHC class IIB alleles contribute to both additive and nonadditive genetic effects on survival in Chinook salmon. *Molecular Ecology*, 15:p.2357–2365.

Pope, L. C., Sharp, A., & Moritz, C. (1996). Population structure of the yellow-footed rock-wallaby *Petrogale xanthopus* (Gray, 1854) inferred from mtDNA sequences and microsatellite loci. *Molecular Ecology*, 5(5): p.629-640.

Potts, W.K., and Wakeland, E.K. (1993). Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference. *Trends in Genetics*, 9(12):p.408–412.

Potts, W.K., Manning, C.J., Wakeland, E.K. and Hughes, A.L. (1993). The role of infectious disease, inbreeding and mating preferences in maintaining MHC genetic diversity: an experimental test [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 346(1317):p.369-378.

Prabakaran P., Streaker E., Chen W., and Dimitrov D. S. (2011). 454 antibody sequencing - error characterization and correction. *BMC Research Notes* , 4:p.404 .

Primmer, C.R., Møller, A.P., and Ellegren, H., (1996). A wide-range survey of cross-species microsatellite amplification in birds. *Molecular Ecology*, 5(3):p.365-378.

Provine W.B. (1986). *Sewall Wright and Evolutionary Biology.* The University of Chicago Press, Chicago,

Pusey, A. and Wolf, M. (1996). Inbreeding avoidance in animals. *Trends in Ecology & Evolution*, 11(5), p..201-206.

Queney, G., Ferrand, N., Weiss, S., Mougel, F. and Monnerot, M. (2001). Stationary distributions of microsatellite loci between divergent population groups of the European rabbit (*Oryctolagus cuniculus*). *Molecular Biology and Evolution*, 18(12): p.2169-2178.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna Austria,

Racey, P.A., Barratt, E.M., Burland, T.M., Deaville, R., Gotelli, D., Jones, G. and Piertney, S.B. (2007). Microsatellite DNA polymorphism confirms reproductive isolation and reveals differences in population genetic structure of cryptic pipistrelle bat species. *Biological Journal of the Linnean Society*, 90(3), p.539-550.

Radwan, J., Zagalska-Neubauer, M., Cichon, M., Sendecka, J., Kulma, K., Gustafsson, L., and Babik, W. (2012). MHC diversity, malaria and lifetime reproductive success in collared flycatchers. *Molecular Ecology,* 21(10):p.2469–2479.

Ragot, M., and Hoisington, D.A. (1993). Molecular markers for plant breeding: comparisons of RFLP and RAPD genotyping costs. *Theoretical and Applied Genetics*, 86(8):p.975-984.

Rahbek, C. (1993). Captive breeding—a useful tool in the preservation of biodiversity? *Biodiversity & Conservation,* 2(4):p.426-437.

RaSäNEN, K.R., Laurila, A. and Merilä, J. (2003). Geographic variation in acid stress tolerance of the moor frog, Rana arvalis. I. Local adaptation. *Evolution,* 57(2): p.352-362.

Raymond, M., & Rousset, F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86(3): p.248-249.

Reed, D. H., Lowe, E. H., Briscoe, D. A., & Frankham, R. (2003). Fitness and adaptation in a novel environment: effect of inbreeding, prior environment, and lineage. *Evolution*, 57(8): p.1822-1828.

Reed, D.H., & Frankham, R. (2001). How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution,* 55(6):p. 1095–1103.

Reed, D.H., & Frankham, R. (2003). Correlation between Fitness and Genetic Diversity. *Conservation Biology,* 17(1):p.230–237.

Rellstab, C., Gugerli, F., Eckert, A.J., Hancock, A.M., and Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24(17):p.4348–4370.

Relyea, R.A. (2002). Local population differences in phenotypic plasticity: predator-induced changes in wood frog tadpoles. *Ecological Monographs*,72(1): p.77-93.

Repaci, V., Stow, A.J. and Briscoe, D.A., (2006). Fine-scale genetic structure, co-founding and multiple mating in the Australian allodapine bee (*Exoneura robusta*). *Journal of Zoology*, 270(4): p..687-691.

Reusch, T. B., HaÈberli, M. A., Aeschlimann, P. B., and Milinski, M. (2001). Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature,* 414(6861): p.300-302.

Reusch, T.B.H., and Langefors, A. (2005). Inter- and Intralocus Recombination Drive MHC Class IIB Gene Diversification in a Teleost, the Three-Spined Stickleback *Gasterosteus aculeatus. Journal of Molecular Evolution*, 61(4):p.531-541.

Reusch, T.B.H.,Häberli, M.A., Aeschlimann, P.B., and Milinski, M., (2001). Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature*, 414:p.300-302.

Rhodes, M., Straw, R., Fernando, S., Evans, A., Lacey, T., Dearlove, A., Greystrong, J., Walker, J., Watson, P., Weston, P. and Kelly, M., (1998). A high-resolution microsatellite map of the mouse genome. *Genome Research,* 8(5):p.531-542..

Rhymer, J.M., and Simberloff, D.S., (1996). Extinction by hybridization and introgression. *Annual Review of Ecological Systems*, 27:p.83-109.

Richards J.D. (2012). *Rufous Hare-Wallaby (Lagorchestes hirsutus) National Recovery Plan.* Department of Environment and Conservation Western Australia,

Richardson, B.J., Baverstock, P.R., and Adams, M. (1986). *Allozyme electrophoresis : a handbook for animal systematics and population studies.* Academic Press, San Diego, California,

Richardson, J.L., Urban, M.C., Bolnick, D.I., and Skelly, D.K., (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology and Evolution*, 29(3):p.165-176.

Richman, A.D., Herrera, L.G., Nash, D. and Schierup, M.H. (2003). Relative roles of mutation and recombination in generating allelic polymorphism at an MHC class II locus in Peromyscus maniculatus. *Genetical Research,* 82(02):p.89-99.

Ringø, E., Salinas, I., Olsen, R. E., Nyhaug, A., Myklebust, R., and Mayhew, T. M. (2007). Histological changes in intestine of Atlantic salmon (*Salmo salar L.*) following in vitro exposure to pathogenic and probiotic bacterial strains. *Cell and tissue research*, 328(1): p.109-116.

Rivers, M.C., Brumitt, N.A., Lughadhab, E.N., and Meaghera, T.R., (2014). Do species conservation assessments capture genetic diversity? Global Ecology and *Conservation*, 2:p.81-87.

Rivas, M.A., Beaudoin, M., Gardet, A., Stevens, C., Sharma, Y., Zhang, C.K., Boucher, G., Ripke, S., Ellinghaus, D., Burtt, N. and Fennell, T. (2011). Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nature genetics*, 43(11): p.1066-1073.

Robinson, J., Waller, M. J., Parham, P., de Groot, N., Bontrop, R., Kennedy, L. J., ... and Marsh, S. G. (2003). IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. Nucleic acids *research*, 31(1): p.311-314.

Ronaghi M., Karamohamed S., Pettersson B., Uhlen M. and Nyren P. (1996). Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry*, 242: p.84–89.

Rose, T.M., and Schultz. E.R. (1998). Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences *Nucleic Acid Research*, 26(7):p.1628-1635.

Rothberg, J.M., and Leamon, J.H. (2008). The development and impact of 454 . *Nature Biotechnology*, 26(10): p.1117-1124.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular ecology resources*, 8(1):p. 103-106.

Rowe, G., Beebee, T.J.C. and Burke, T., (1999). Microsatellite heterozygosity, fitness and demography in natterjack toads Bufo calamita. *Animal Conservation*, 2(2): p..85-92.

Rueffler, C., Van Dooren, T.J.M., Leimar, O., and Abrams, P.A. (2006). Disruptive Selection then What? *Trends in Ecology and Evolution*, 21(5):p.238-245.

Saccheri, I.J., Brakefield, P.M., & Nichols, R.A. (1996). Severe inbreeding depression and rapid fitness rebound in the butterfly Bicyclus anynana (Satyridae). *Evolution,* p.2000-2013.

SäNEN, K.R., Laurila, A. and Merilä, J. (2003). Geographic variation in acid stress tolerance of the moor frog, Rana arvalis. I. Local adaptation. *Evolution,* 57(2), p.352-362..

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Biochemistry*, 74(12):p.5463-5467.

Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., and Kilian, A., (2011). Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proceedings*, 5(7):p.54.

Sato, A., Figueroa, F.,and Murray, B.W., (2000). Nonlinkage group of major histocompatibility complex class I and class II loci in bony fishes. *Immunogenetics*, 51:p.108–116.

Saunders, G., & McLeod, L. (2007). *Improving fox management strategies in Australia*. Bureau of Rural Sciences, Canberra.

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics,* 14:p.807–820.

Scaglione, D., Acquadro, A., Portis, E., Tirone, M., Knapp, S.J., and Lanteri, S., (2012). RAD tag sequencing as a source of SNP markers in *Cynara cardunculus L B.MC Genomics*, 13(3):p.3.

Schloss, P.D., Westcott, S.L., Ryabin, T, Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A, Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., and Weber, C.F. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmetnal Microbiology,* 75(23):p.7537-7541.

Schlötter, C. (2004). The evolution of molecular markers-just a matter of fashion? *Nature Reviews Genetics,* 5:p. 63-69 (2004).

Schug, M.D., Hutter, C.M., Wetterstrand, K.A., Gaudette, M.S., Mackay, T.F.C., and Aquadro, C.F., (1998). The Mutation Rates of Di-, Tri- and Tetranucleotide Repeats in *Drosophila melanogaster*. *Molecular Biology Evolution*, 15(12):p.1751–1760.

Schuiling, G.A., (2003). The benefit and the doubt: why monogamy? *Journal of Psychosomatic Obstetrics & Gynecology*, 24(1):p.55-61.

Schuster, S.C. (2008). Next-generation sequencing transforms today's biology. *Nature Methods*, 5(1):p.16-18.

Schwensow, N., Dausmann, K., Eberle, M., Fietz, J., and Sommer, S. (2010). Functional associations of similar MHC variants and shared parasite species in two sympatric lemurs Infection. *Genetics and Evolution*, 10:p.662–668.

Seddon, P.J., Armstrong, D.P., and Maloney, R.F. (2007). Developing the Science of Reintroduction Biology. *Reintroduction Biology*, 21(2):p.303-312.

Selkoe, K.A. and Toonen, R.J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters*, 9(5):p.615-629..

Semple, J., and Brown, S. (2004). Analysis of a high-throughput yeast two-hybrid system and its use to predict the function of intracellular proteins encoded within the human MHC class III region. *Genomics*, 83:p.153-167.

Sexton, J.P., McIntyre, P.J., Angert, A.L., & Rice, K.J. (2009). Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, 40: p.415-436.

Shaffer, M.L. (1981). Minimum Population Sizes for Species Conservation. *BioScience*, 31(2):p.131-134.

Shaw, P. W., C. Turan, J. M. Wright, M. O'Connell and G. R. Carvalho. (1999). Microsatellite DNA analysis of population structure in Atlantic herring (Clupea harengus), with direct comparison to allozyme and mtDNA RFLP analyses. *Heredity*, 83:p.490–499.

Sheean, V.A., Manning, A.D. and Lindenmayer, D.B., (2012). An assessment of scientific approaches towards species relocations in Australia. *Austral Ecology*, 37(2), p.204-215.

Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. *Nature Biotechnology*, 26:p.1135 - 1145.

Sherborne, A.L., Thom, M.D., Paterson, S., Jury, F., Ollier, W.E., Stockley, P., Beynon, R.J. and Hurst, J.L., (2007). The genetic basis of inbreeding avoidance in house mice. *Current Biology*, 17(23),p. .2061-2066.

Short, J., & Smith, A. (1994). Mammal decline and recovery in Australia. *Journal of Mammalogy*, 75(2): p.288-297.

Shultz, S., Opie, C., and Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature*, 479(7372): p.219-222.

Siddle, H.V., Marzec, J., Cheng, Y., Jones, M., and Belov, K., (2010). MHC gene copy number variation in Tasmanian devils: implications for the spread of a contagious cancer. *Proceedings of the Royal Society of London Series B: Biological Science*s, 277:p.2001-2006.

Silberstein, A., Machanavajjhala, A., and Ramakrishnan, R. (2011). *Feed Following: The Big Data Challenge*. In Social Applications In Databases and Social Networks (pp. 1-6). ACM.,

Singh, V.K., Mangalam, A.K., Dwivedi, S., and Naik, S. (1998). Primer Premier: Program for Design of Degenerate Primers from a Protein Sequence. *BioTechniques*, 24:p.318-319.

Slate, J., Kruuk, L.E.B., Marshall, T.C., Pemberton, J.M. and Clutton-Brock, T.H. (2000). Inbreeding depression influences lifetime breeding success in a wild population of red deer (*Cervus elaphus).* *Proceedings of the Royal Society of London B: Biological Sciences,* 267(1453):p.1657-1662.

Sloane, M. A., Sunnucks, P., Alpers, D., Beheregaray, L. B., & Taylor, A. C. (2000). Highly reliable genetic identification of individual northern hairy-nosed wombats from single remotely collected hairs: a feasible censusing method. *Molecular Ecology*, 9(9): p.1233-1240.

Sloane, M.A., Sunnucks, P., Alpers, D., Beheregaray, L.B., and Taylor, A.C. (2000). Highly reliable genetic identification of individual northern hairy-nosed wombats from single remotely collected hairs: a feasible censusing method. *Molecular Ecology*, 9(9):p.1233–1240.

Smith, J.L., & Fonseca, D.M. (2004). Rapid assays for identification of members of the Culex (Culex) pipiens complex, their hybrids, and other sibling species (Diptera: Culicidae). *The American Journal of Tropical Medicine and Hygiene*, 70(4): p.339-345.

Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B., and Hood, L.E., (1986). Fluorescence detection in automated DNA sequence analysis*.* *Nature*, 321(6071):p.674-679.

Smith, T.B., and Wayne, R.K. (1996). *Molecular Genetic Approaches in Conservation.* Oxford University Press, Oxford, England,

Smithies, O. (1955). Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults. *The Biochemical Journal*, 61:p.629-641.

Smouse, P.E., and Chakraborty, R. (1986). The use of restriction fragment polymorphisms in paternity analysis. *American Journal of Human Genetics*, 38:p.918-939.

Snell, G.D. (1981). Studies in histocompatibility. *Science,* 213(4504):p.172-178.

Snell, G.D., and Higgins, G.F., (1951). Alleles at the histocompatibility-2 locus in the mouse as determined by tumor transplantation. *Genetics,* 36(3):p.306-310.

Snyder, N. F., Derrickson, S. R., Beissinger, S. R., Wiley, J. W., Smith, T. B., Toone, W. D., & Miller, B. (1996). Limitations of captive breeding in endangered species recovery. *Conservation Biology*, 10(2):p. 338-348.

Snyder, N.F.R., Derrickson, S.R., Beissinger, S.R., Wiley, J.W., Smith, T.B., Toone, W.D., and Miller, B., (1996). Limitations of Captive Breeding in Endangered Species Recovery. *Conservation Biology*, 10(2):p.338-348.

Sollid, L.M., and Thorsby, E., (1993). HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. *Gastroenterology,* 105(3):p.910-922.

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, 2(16):p.1-18.

Sommer, S., Courtiol, A., and Mazzoni, C.J. (2013). MHC genotyping of non-model organisms using next-generation sequencing: a new methodology to deal with artefacts and allelic dropout. *BMC Genomics,* 14:p.542.

Sorci, G., and Faivre, B. (2009). Inflammation and oxidative stress in vertebrate host–parasite systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1513): p.71-83.

Soule, M. E., & Wilcox, B. A. (1980). *Conservation biology. An evolutionary-ecological perspective*. Sinauer Associates Inc, Sunderland, MA.

Soulé, M.E., (1976). *Allozyme variation, its determinants in space and time*. Pages 60-77 in F. J. Ayala, editor. Molecular evolution. Sinauer Associates, Sunderland, MA.

Soulé, M.E., (1986). *Conservation Biology: The Science of Scarcity and Diversity* Sinauer Associates, Sunderland, MA

Soulé, M.E., (1991). Conservation: tActics for a Crisis Science. *Science (Washington),* 253(5021): p.744-750.

Spencer, J.A., Sinclair, A.H., Watson, J.M. and Graves, J.A.M., (1991). Genes on the short arm of the human X chromosome are not shared with the marsupial X. Genomics*,* 11(2):p.339-345.

Spielman, D. and Frankham, R., (1992). Modeling problems in conservation genetics using captive Drosophila populations: improvement of reproductive fitness due to immigration of one individual into small partially inbred populations. Zoo *Biology*, 11(5):p. 343-351.

Spielman, D., Brook, B.W. and Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America*, 101(42):p.15261-15264.

Spieth, P. T., (1974). Gene flow and genetic differentiation. *Genetics,* 78(3):p.962-965.

Spong, G., Stone, J., Creel, S., and Björklund, M. (2002). Genetic structure of lions *(Panthera leo L.)* in the Selous Game Reserve: implications for the evolution of sociality. *Journal of Evolutionary Biology*, 15(6):p.945-953.

Spurgin LG, Richardson DS (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 277:p.979–988.

Spurgin, L. G., van Oosterhout, C., Illera, J. C., Bridgett, S., Gharbi, K., Emerson, B. C., and Richardson, D. S. (2011). Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations*. Molecular ecology*, 20(24): p.5213-5225.

Spurgin, L.G., and Richardson, D.S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society of London Series B: Biological Sciences,* 277:p.979–988.

Stankowich, T., Haverkamp, P. J., and Caro, T. (2014). Ecological drivers of antipredator defenses in carnivores*. Evolution*, 68(5):p. 1415-1425.

Stapley, J., Reger, J., Feulner, P.G.D., Smadja, C., Galindo, J., Ekblom, R., Bennison, C., Ball, A. D., Beckerman, A. P., and Slate, J. (2010). Adaptation Genomics: the next generation. *Trends in Ecology and Evolution*, 1295:p.1-8.

Starks, P. T., and Peters, J. M., (2000). Semi-nondestructive genetic sampling from live eusocial wasps, *Polistes dominulus* and *Polistes fuscatus Insectes. Sociaux*, 49:p.20-22.

Stewart, C. A., Horton, R., Allcock, R.J.N., Ashurst, J.L., Atrazhev, A.M., Coggill, P., Dunham, I., Forbes, S., Halls, K., Howson, J.M.M., Humphray, S. J., Hunt, S., Mungall, A.J., Osoegawa, K., Palmer, S., Roberts, A.N., Rogers, J., Sims, S., Wang, Y., Wilming, L.G., Elliot, J.F., de Jong, P.J., Sawcer, S., Todd, J.A., Trowsdale, J., and Beck, S. (2004). Complete MHC Haplotype Sequencing for Common Disease Gene Mapping. *Genome Research,* 14:p.1176-1187.

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S., Spear, S.F., ...& Waits, L.P. (2007). Putting the 'landscape' in landscape genetics. *Heredity*, 98(3): p.128-142.

Stow AJ, Briscoe DA, Gillings M, Holley M, Smith S, Leys R, Silberbauer T, Turnbull C, Beattie AJ (2007). Antimicrobial defenses increase with sociality in bees. *Biology Letters,* 3: p.422-424.

Stow, A. J. (2002). Microsatellite loci from the Cunningham's Skink (*Egernia cunninghami).* *Molecular Ecology Notes*, 2(3): p.256-257.

Stow, A. J., & Sunnucks, P. (2004). Inbreeding avoidance in Cunningham's skinks (Egernia cunninghami) in natural and fragmented habitat. *Molecular Ecology*, 13(2): p.443-447.

Stow, A.J. and Briscoe, D.A. (2005). Impact of habitat fragmentation on allelic diversity at microsatellite loci in Cunningham's skink (*Egernia cunninghami)*; a preliminary study. *Conservation Genetics*, 6(3), p.pp.455-459..

Stow, A.J. and Sunnucks, P. (2004). High mate and site fidelity in Cunningham's skinks *(Egernia cunninghami*) in natural and fragmented habitat. *Molecular Ecology*, 13(2):p.419-430.

Stow, A.J. and Sunnucks, P. (2004b). Inbreeding avoidance in Cunningham's skinks (*Egernia cunninghami)* in natural and fragmented habitat. *Molecular Ecology*, 13(2):p.443-447.

Stow, A.J., Sunnucks, Briscoe, D.A., and Gardner, M.G. (2001). The impact of habitat fragmentation on dispersal of Cunningham's skink (*Egernia cunninghami*): evidence from allelic and genotypic analyses of microsatellites. *Molecular Ecology*, 10(4):p.867–878.

Strand, T.M., and Höglund, J. (2011). Genotyping of black grouse MHC class II B using reference Strand-Mediated Conformational Analysis (RSCA). *BMC Research Notes*, 4:p.183.

Stucki, S., Orozco-terWengel, P., Forester, B.R., Duruz, S., Colli, L., Masembe, C., Negrini, R., Landguth, E., Jones, M.R., Bruford, M.W. and Taberlet, P. (2016). High performance computation of landscape genomic models including local indicators of spatial association. *Molecular Ecology Resources. In Press*

Sunnucks, P. (2000). Efficient genetic markers for population biology. *Trends in Ecology & Evolution*, 15(5):p.199-203.

Sunnucks, P. and Hales, D.F., (1996). Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus Sitobion (*Hemiptera: Aphididae). Molecular Biology and Evolution*, 13(3), p.510-524.

Sunnucks, P., Wilson, A.C.C., Beheregaray, L.B., Zenger, K., Fench, J., and Taylor, A.C. (2000). SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology*, 9:p.1699–1710.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P., and Bouvet, J. (1996). Reliable Genotyping of Samples with Very Low DNA Quantities Using PCR. *Nucleic Acids Research,* 24(16):p.3189-3194.

Taberlet, P., Luikart, G. (1999). Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society*, 68(1-2): p.41–55.

Taberlet, P., Waits, L. P., & Luikart, G. (1999). Noninvasive genetic sampling: look before you leap. *Trends in Ecology & Evolution*, 14(8):p. 323-327.

Tainaka, K.I. and Itoh, Y., (1996). Glass effect in inbreeding-avoidance systems: minimum viable population for outbreeders*. Journal of the physical Society of Japan*, 65(10), p.3379-3385.

Tallmon, D.A., Luikart, G., and Waples. R.S. (2004). The alluring simplicity and complex reality of genetic rescue. *TRENDS in Ecology and Evolution*, 19(9):p.490-496.

Tamura K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using Maximum Likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution,* 28:p.2731-2739.

Tamura, K., Dudley, J., Nei, M. and Kumar, S., (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular biology and evolution*, 24(8):p.1596-1599..

Taylor, A. C., & Cooper, D. W. (1998). Microsatellite markers for the Phalangerid marsupial, the common brushtail possum (*Trichosurus vulpecula). Molecular Ecology*, 7(12): p.1780-1782.

Taylor, E.B. (1991). A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. *Aquaculture,* 98:p.185-207.

Taylor, L. H., Latham, S. M., and Mark, E. J. (2001). Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* 356(1411): p.983-989.

Tedersoo, L., Nilsson, R.H., Abarenkov, K., Jairus, T., Sadam, A., Saar, I., Bahram, M., Bechem, E., Chuyong, G., and Kõljalg, U. (2010). 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist*, 188(1):p.291-301.

Tellier, A., Moreno-Gámez, S., and Stephan, W. (2014). Speed of adaptation and genomic footprints of host–parasite coevolution under arms race and trench warfare dynamics. *Evolution*, 68(8):p.2211-2224.

Templeton, A.R., Hemmer, H., Mace, G., Seal, U.S., Shields, W.M., and Woodruff, D.S. (1986). Local Adaptation, Coadaptation, and Population Boundaries. *Zoo Biology*, 5:p.115-125.

Tennessen, J.A., Bigham, A.W., O'Connor, T.D., Fu, W., Kenny, E.E., Gravel, S., McGee, S., Do, R., Liu, X., Jun, G., Kang, H.M., Jordan, D., Leal, S.M., Gabriel, S., Rieder, M.J., Abecasis, G., Altshuler, D., Nickerson, D.A., Boerwinkle, E., Sunyaev, S., Bustamante, C.D., Bamshad, M.J., Akey, J.M., Go, B., and Go, S. (2012). Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. *Science* , 337(6090):p.64-69.

Terasaki, P.I., Machioro, T.L., and Starzl, T.E. (1965). Sero-Typing of Human Lymphocoyte Anitgens: Preliminary Trials on Long-Term Kidney Homograft Survivors In Amos, D.B., and Russell, P.S. (eds). *Histocompatibility Testing*. National Academies of Sciences, Washington,

Tessier, N., Bernatchez, L. Presa, P and Angers, B. (1995). Gene diversity analysis of mitochondrial DNA, microsatellites and allozymes in landlocked Atlantic salmon. *Journal of Fish Biology*, 47(A):p.156-163.

Thoss, M., Ilmonen, P., Musolf, K., and Penn, D. J. (2011). Major histocompatibility complex heterozygosity enhances reproductive success. *Molecular ecology,* 20(7): p.1546-1557.

Townsend, A.R.M., Rothbard, J., Gotch, F.M., Bahadur, G., Wraith, D., & McMichael, A.J. (1986). The epitopes of influenza nucleoprotein recognized by cytotoxic T lymphocytes can be defined with short synthetic peptides. *Cell*, 44(6): p.959-968.

Treangen, T.J., and Salzberg, S.L. (2013). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nature Reviews Genetics*, 13(1):p.36-46.

Tremblay, R.L., & Ackerman, J.D. (2001). Gene flow and effective population size in Lepanthes (*Orchidacea*e): a case for genetic drift. *Biological Journal of the Linnean Society*, 72(1): p.47-62.

Tsiatis, A.C., Norris-Kirby, A., Rich, R.G., Hafez, M.J., Gocke, C.D., Eshleman, J.R. and Murphy, K.M. (2010). Comparison of Sanger sequencing, pyrosequencing, and melting curve analysis for the detection of KRAS mutations: diagnostic and clinical implications. *The Journal of Molecular Diagnostics*, 12(4): p.425-432.

Tzika, A.C., Remy, C., Gibson, R., Milinkovitch, M.C. (2009). Molecular genetic analysis of a captive-breeding program: the vulnerable endemic Jamaican yellow boa. *Conservation Genetics*, 10:p.69–77.

Ujvari, B., and Belov, K., (2011). Major Histocompatibility Complex (MHC) markers in conservation biology. *International journal of molecular science*, 12(8):p.5168-5186.

Uni, Z., Gutman, M., Leitner, G., Landesman, E., Heller, D., and Cahaner, A. (2000). Major Histocompatibility Complex Class IV Restriction Fragment Length Polymorphism Markers in Replicated Meat-Type Chicken Lines Divergently Selected for High or Low Early Immune Response. *Poultry Science*, 72(10):p.1823-1831.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G., (2012). Primer3—new capabilities and interfaces. *Nucleic acids research*, 40(15):p.e115-e115..

Väli, U., Einarsson, A., Waits, L., and Ellegren, H. (2008). To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology*, 17:p.3808-3817.

Valière, N., Fumagalli, L., Gielly, L., Miquel, C., Lequette, B., Poulle, M.L., Weber, J.M., Arlettaz, R., and Taberlet, P. (2003). Long-distance wolf recolonization of France and Switzerland inferred from non-invasive genetic sampling over a period of 10 years. *Animal Conservation*, 6:p.83-92.

van Baalen, M., and Sabelis, M. W. (1995). The dynamics of multiple infection and the evolution of virulence. *The American Naturalist*, p.881-910.

Van Der Walt, S., Colbert, S.C. and Varoquaux, G., (2011). The NumPy array: a structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2): p.22-30.

van Dijk, E.L., Auger, H., aszczyszyn, Y., and Thermes, C., (2014). Ten years of next-generation sequencing technology. *Trends in Genetics,* 30(9):p.418-426.

Van Valen, L. (1972). A new evolutionary law. *Evolutionary Theory*, 1:p.1-30.

Vidal-Puig, A. and Moller, D.E. (1994). Comparative sensitivity of alternative single-strand conformation polymorphism (SSCP) methods. *Biotechniques,* 17(3), p..490-492.

Villesen, P., and Fredsted, T., (2006). Fast and non-invasive PCR sexing of primates: apes, Old World monkeys, New World monkeys and *Strepsirrhines. BMC Ecology*, 6:p.8.

Vitousek, P.M., Mooney, H.A., Lubchenco, J., & Melillo, J.M. (1997). Human domination of Earth's ecosystems. *Science,* 277(5325): p.494-499.

Von Schantz, T., Wittzell, H., Goransson, G., Grahn, M. and Persson, K., (1996). MHC genotype and male ornamentation: genetic evidence for the Hamilton-Zuk model. *Proceedings of the Royal Society of London B: Biological Sciences*, 263(1368), p.265-271.

Vrijenhoek, R.C. and Leberg, P.L. (1991). Let's not throw the baby out with the bathwater: a comment on management for MHC diversity in captive populations. *Conservation Biology*, 5(2), p..252-254.

Waits, L., Taberlet, P., Swenson, J.E., Sandegren, F., and Franzén, R. (2000). Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear *(Ursus arctos). Molecular Ecology,* 9(4):p.421-431.

Waldron, L.S., Ferrari, B.C., Gillings, M.R. and Power, M.L. (2009). Terminal restriction fragment length polymorphism for identification of Cryptosporidium species in human feces. *Applied and environmental microbiology*, p.108-112.

Walsh PS, Metzger DA, Higuchi R (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*, 10(4): p.506-513.

Wan, Q.H., Fang, S.G., Wu, H., Fujihara, T. (2003). Genetic differentiation and subspecies development of the giant panda as revealed by DNA fingerprinting. *Electrophoresis*, 24(9):p.1353–1359.

Wan, Q.H., Fang, S.G., Wu, H., Fujihara, T. (2004). Which genetic marker for which conservation genetics issue? *Electrophoresis,* 25(14):p.2165–2176.

Wandeler, P., Hoeck, P.E., & Keller, L.F. (2007). Back to the future: museum specimens in population genetics. *Trends in Ecology & Evolution*, 22(12): p.634-642.

Wang, M., Jiang, Y. Y., Kim, K. M., Qu, G., Ji, H. F., Mittenthal, J. E., ... and Caetano-Anollés, G. (2011). A universal molecular clock of protein folds and its power in tracing the early history of aerobic metabolism and planet oxygenation. *Molecular biology and evolution,* 28(1): p.567-582.

Wang, Y., Zhu, W., Huang, L., Zhou, K., and Wang, R. (2006). Genetic diversity of Chinese alligator (Alligatorsinensis) revealed by AFLP analysis: an implication on the management of captive conservation. *Biodiversity and Conservation*, 15: p.2945-2955.

Wang, Y.Q., Zhu, W.Q., Huang, L., Zhou, K.Y. and Wang, R.P. (2006). Genetic diversity of Chinese alligator (Alligator sinensis) revealed by AFLP analysis: An implication on the management of captive conservation. *Biodiversity & Conservation*, 15(9), p.2945-2955.

Waples, R.S., (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics*, 121(2), p.379-391.,

Waples, R.S., (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, 7(2), p.167-184.,

Ward, R.D., Woodwark, M., & Skibinski, D.O.F. (1994). A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *Journal of fish biology*, 44(2): p.213-232.

Wasko, A., Martins, C., Oliveira, C., and Foresti, F. (2003). Non-destructive genetic sampling in fish. An improved method for DNA extraction from fish fins and scales. *Hereditas*, 138(3):p.161-165.

Watson, M. (2014). Illuminating the future of DNA sequencing. *Genome Biology*, 15(108):p.108.

Wayne, R.K., & Jenks, S.M. (1991). Mitochondrial DNA analysis implying extensive hybridization. *Nature*, 351:p.565-568.

Weber, D.S., Stewart, B.S., Schienman, J. and Lehman, N., (2004). Major histocompatibility complex variation at three class II loci in the northern elephant seal. *Molecular Ecology*, 13(3), p.711-718.

Wedekind, C. and Penn, D. (2000). MHC genes, body odours, and odour preferences. *Nephrology Dialysis Transplantation,* 15(9): p1269-1271.

Wedekind, C., Seebeck, T., Bettens, F. and Paepke, A.J. (1995). MHC-dependent mate preferences in humans. *Proceedings of the Royal Society of London B: Biological Sciences,* 260(1359):p.245-249.

Weeks, A.R., Sgro, C.M., Young, A.G., Frankham, R., Mitchell, N.J., Miller, K.A., Byrne, M., Coates, D.J., Eldridge, M.D., Sunnucks, P. and Breed, M.F., (2011). Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications*, 4(6): p.709-725.

Wegner, K. (2009). Massive parallel MHC genotyping: titanium that shines. *Molecular Ecology*, 18(9): p.1818-1820.

Wegner, K.M., Kalbe, M., Kurtz, J., Reusch, T.B. and Milinski, M. (2003). Parasite selection for immunogenetic optimality. *Science,* 301(5638): p.1343-1343.

Wells, K.D., (1977). The social behaviour of anuran amphibians. *Animal Behaviour*, 25: p.666-693.

Westerdahl, H. (2007). Passerine MHC: genetic variation and disease resistance in the wild. *Journal of Ornithology,* 148(2): p.469-477.

Westerdahl, H., Waldenstrom, J., Hansson, B., Hasselquist, D., Von Schantz, T., and Bensch, S. (2005). Associations between malaria and MHC genes in a migratory songbird. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1571): p.1511-1518

Westemeier, R.L., Brawn, J.D., Simpson, S.A., Esker, T.L., Jansen, R.W., Walk, J.W., Kershner, E.L., Bouzat, J.L. and Paige, K.N. (1998). Tracking the long-term decline and recovery of an isolated population. *Science*, 282(5394): p.1695-1698.

Whitehead, A., & Crawford, D.L. (2006). Variation within and among species in gene expression: raw material for evolution. *Molecular ecology*, 15(5):p.1197-1211.

Whiteman, E.A. and Côté, I.M. (2004). Monogamy in marine fishes. *Biological Reviews*, 79(2), p.351-375.

Whitlock, M.C. (2000). Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. *Evolution,* 54(6):p.1855-1861.

Whitlock, M.C., & Agrawal, A.F. (2009). Purging the genome with sexual selection: reducing mutation load through selection on males. *Evolution*, 63(3):p.569-582.

Wilcox, B.A., and Murphy, D.D. (1985). Conservation Strategy: The Effects of Fragmentation on Extinction. *The American Naturalist*, 125(6):p.879-887.

Willi, Y., Van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, p.433-458.

Williams, S. E., Bolitho, E. E., & Fox, S. (2003). Climate change in Australian tropical rainforests: an impending environmental catastrophe. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1527): p.1887-1892.

Woodworth, L. M., Montgomery, M. E., Briscoe, D. A., & Frankham, R. (2002). Rapid genetic deterioration in captive populations: causes and conservation implications. *Conservation Genetics*, 3(3):p. 277-288.

Woolhouse, M. E., Webster, J. P., Domingo, E., Charlesworth, B., and Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature genetics,* 32(4): p.569-577.

Xia, X. (1992). Uncertainty of paternity can select against paternal care. *The American Naturalist*, 139(5): p.1126-1129.

Yamazaki, K., Boyse, E.A., Mike, V., Thaler, H.T., Mathieson, B.J., Abbott, J., Boyse, J., Zayas, Z.A. and Thomas, L., (1976). Control of mating preferences in mice by genes in the major histocompatibility complex. *The Journal of experimental medicine*, 144(5): p.1324-1335.

Young, N.S., Loannidis, J.P.A., and Al-Ubaydli, O. (2008). Why Current Publication Practices May Distort Science. *PLoS Med, 5(*10):

Yuhki, N., and O'Brien, S.J. (1990). DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. *Proceedings of the National Academy of Sciences of the United States of America,* 87:p.836-840.

Yuhki, N., Beck, T., Stephens,R.M., Nishigaki, Y., Newmann, K.,and O'Brien, S.J. (2003). Comparative Genome Organization of Human, Murine, and Feline MHC Class II Region. *Genome Research,* 13:p.1169-1179.

Zauhar, R.J. (2001). University bioinformatics programs on the rise. *Nature Biotechnology*, 19:p.285-286.

1 Zenger, K.R. and Cooper, D.W., (2001). A set of highly polymorphic microsatellite markers developed
2 for the eastern grey kangaroo (Macropus giganteus). *Molecular Ecology Notes*, 1(1-2):p.98-100.

3 Zhao, S., Zheng, P., Dong, S., Zhan, X., Wu, Q., Guo, X., Hu, Y., He, W., Zhang, S., Fan, W., Zhu, L., Li,
4 D., Zhang, X., Chen, Q., Zhang, H., Zhang, Z., Jin, X., Zhang, J., Yang, H., Wang, J., Wang, J., and Wei,
5 Fuwen (2013). Whole-genome sequencing of giant pandas provides insights into demographic
6 history and local adaptation. *Nature Genetics*, 45: p.67-71.

7 Ziegler, A., Kentenich, H. and Uchanska-Ziegler, B. (2005). Female choice and the MHC. *Trends in*
8 *immunology*, 26(9), p.496-502.

9 Zinkernagel, R.M., ad Doherty, P.C. (1974). Restriction of in vitro T cell-mediated cytotoxicity in
10 lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature*, 248:p.701-702.

11 Zuk, M., and Stoehr, A. M. (2002). Immune defense and host life history. *The American Naturalist,*
12 160(S4): p.S9-S22.

13

Appendix I Ethics Letter, Final Approval

Research Office
Research Hub, Building C5C East
MACQUARIE UNIVERSITY NSW 2109

Animal Ethics
Phone +61 (0)2 9850 7758
Email animal.ethics@ro.mq.edu.au

Macquarie University
Animal Ethics Committee
Wildlife Protocol Application Form
(Version 1.3 December 2009)

Reference No:
(To be filled in by AEC)

OFFICE USE ONLY

Date received:

AEC meeting date:

Answer all questions below; Double Click the Yes, No, or N/A box to indicate your answer.

Submit your SIGNED form along with 14 extra copies (double-sided) to: Animal Ethics Secretariat, Research Office, Level 3 C5C, Macquarie University NSW 2109.

Please also submit a signed electronic copy (PDF preferred) with all appendices included to: Animal.Ethics@ro.mq.edu.au

For further information on how to complete this form, view the Animal Ethics Committee Application Guidelines at http://www.research.mq.edu.au/researchers/ethics/animal_ethics/documents/aecguidelines.doc or contact the Animal Welfare Officer leanne.gillespie@ro.mq.edu.au.

SECTION 1: ADMINISTRATION

(1-1) Title of Proposed Project

Mala recovery plan: genetic analysis to inform conservation management

Teaching: ☐
Research: ☒

(1-2) Personal Details of Principal Investigator* (PI)

NB as of 2010, for Honours, Postgraduate and Higher Degree Research students, the **Supervisor must be named as the Principle Investigator*

Name and Title (e.g. Dr, Prof, Ms)	:	**Dr. Adam Stow**
MQ Staff or Student Number	:	**99004372**
MQ Department	:	**Biology**
Address for correspondence (if not MQ)	:	
Relationship with institution (e.g. PhD, Masters, etc.)	:	**Senior Lecturer**
Email	:	**adam.stow@mq.edu.au**
Tel (work)	:	**8153**
Tel (home)	:	**98751824**
Mobile	:	
Fax	:	**98509395**

(1-3) Proposed Licensing Period of Project

Commencement date (dd/mm/yy)	:	**31st / June / 2010**
Completion date (dd/mm/yy)	:	**31st / June / 2011**
Duration of study (in months)	:	**12**

(1-4) Financial Support of Proposed Project

a) Is the project being supported by an internal or external grant? Yes ☒ No ☐

b) **If YES**, please state

Funding organisation/application date	:	**Alice Springs Desert Wildlife Park**
Status (pending or approved)	:	**Approved**
Title as it appears on grant application	:	**Mala recovery plan: genetic analysis to inform conservation management**

(1-5) Ethics Approval of Grant Funded Project

a) Will this project still go ahead if funding is not successful? Yes☐ No☐ N/A☒

b) Is Animal Ethics approval a condition of your grant funding? Yes☐ No☒ N/A☐

c) Will this Animal Ethics approval cover all or part of the work conducted under an internal or external grant? Yes☒ No☐ N/A☐

d) Are the procedures set out in your grant application the same as those set out in this Animal Ethics application? Yes☒ No☐ N/A☐

(1-6) Status of AEC Application

Please indicate whether this application is a

a) New project? ☒

b) Re-submission? ☐

c) Project that has previously or simultaneously been submitted to this or other AEC(s)? ☐

d) Significantly revised current protocol? ☐

e) **If you ticked 1-6b or 1-6c**, provide name of the AEC(s), approval reference number(s) and reasons for re- or simultaneous submission. N/A☐

Name of AEC(s)	:
Reference number(s)	:
Reasons	:

f) **If you ticked 1-6d**, quote the approval reference number, the species and number of animals used to date. N/A☐

Reference number	:
Species	:
Animals used to date	:

g) If an application has been made to another AEC for a *similar* project, indicate whether the application was passed or refused by that committee and provide a copy of this application or briefly describe the contents of the project.

Copy attached	☐	Description follows	☐	N/A☒
Passed	☐	Refused	☐	

(1-7) Location and Licensing

a) Identify where the field study will take place, including all the locations at which research using animals will be conducted outside of the field (full street addresses where applicable).

Samples will be collected at the following field sites

Area or Park of study : Alice Springs Desert Park
Address(es) : Larapinta Drive Alice Springs NT 0871
Phone: : 08 8951 8788

Area or Park of study : Watarrka National Park
Address(es) : 450km southwest of Alice Springs
Phone (ranger Station) : 08 8956 7460

Area or Park of study : Uluru Kata Tjuta National Park
Address(es) : 440 km south-west of Alice Springs
Phone (Central Management): 08 8956 1100

Area or Park of study : Dryandra Conservation Reserve
Address(es) : 180 kilometres south-east of Perth (22 kilometres north-west of Narrogin.)
Phone : 08 9881 9200

Area or Park of study : Scotia Sanctuary
Address(es) : 150 km south of Broken Hill, 65,000 ha located within the Murray Mallee subregion of the Murray Darling Depression Bioregion.

Area or Park of study : François Peron National Park
Address(es) : Adjacent to Denham WA 400km North West of Gelaldton

Genetic analysis will be performed at
The conservation genetics lab, Macquarie University
Faculty of Science, Department of biological Sciences Building E8A room 240

b) Does the project involve native, imported or protected species? Yes☒ No☐ N/A☐

c) Does the project require a relevant licence from the National Parks and Wildlife Service or other authorities? Yes☒ No☐ N/A☐

d) **If Yes to 1-7c**, indicate whether licence applications remain to be submitted (and if so when), are pending approval, names of issuing authorities, and if approved, permit number(s).

Approval has been obtained to sample at the following localities (Administered by the desert wildlife park)

Alice Springs Desert Wildlife Park

Approval is currently being sort from:

Site	Governing body
Watarrka National Park	**NRETAS (Natural Resources Environment The Arts and Sport NT Government)**
Uluru Kata Tjuta National Park	**Uluru-Kata Tjuta National Park Management Board**
Dryandra Conservation Reserve	**Department of Environment and Conservation Western Australia**
François Peron National Park	**Department of Environment and Conservation Western Australia**
Scotia Sanctuary	**Australian Wildlife Conservancy**

Sampling will not commence until all approvals have been obtained and supplied to the ethics committee.

A copy of all licences/permits must either accompany this application or be supplied to the Animal Welfare Officer as soon as issued by the regulatory authority.

Date of submission :
Authorities :
Permit number(s) :

(1-8) Biohazard and Safety Licensing

a) Does the project involve hazards to humans or other animals (e.g., recombinant DNA technology, infectious,

toxic, radioactive or carcinogenic agents; see 3.3.50, 3.3.51 and 3.3.53 Code of Practice for details)? Yes☐ No☒ N/A☐

b) If so, is Biosafety Committee and/or Gene Technology Regulator and/or O.H.S. approval required for the work? Yes☐ No☐ N/A☒

c) If so, have approval and/or licence(s) been obtained? Yes☐ No☐ N/A☒

d) Have appropriate measures for containment, disposal and decontamination been established as part of this protocol, in accordance with (but not limited to) Section 3.3.52 of the Code of Practice? Yes☐ No☐ N/A☒

e) Have relevant personnel been informed of the above risks and measures to be taken? Yes☐ No☐ N/A☒

f) Confirm that if Biosafety Committee and/or Gene Technology Regulator approval or risk assessment by an O.H.S. representative is required, work will *not* commence until such approval has been granted or received. Yes☐ N/A☒

SECTION 2: JUSTIFICATION FOR USE OF WILDLIFE

(2-1) Aims, Goals, and Significance of the Project

a) Describe the aims of the project *in lay terms* (≈100 words).

The Mala is a nationally important captive breeding success story. The Northern Territory government (through its Biodiversity Conservation Division and the Alice Springs Desert Park) plays a large role in achieving the goals of the 1999 Mala recovery plan and continues to be committed to maintaining and reintroducing mala populations in order to upgrade the species IUCN status from extinct to vulnerable. In order to facilitate this recovery an understanding of genetic diversity in the species is essential. The Conservation Genetics Lab (CGL) at Macquarie University proposes a detailed analysis of the genetic health of Mala populations in order to provide management and translocation strategies in accordance with the recovery plan.

b) Explain your research goals and outline strategy for achieving them (≈1000 words).

This project addresses the overall and specific aims of 1999 Mala Recovery Plan and the draft 2005 Mala Recovery Plan by providing genetic information to the mala recovery team. Specifically the research goals are:

1. **To assess the genetic contribution of founders to the current population in order to inform the mala project conservation team**
2. **To assess the genetic diversity of the paddock populations which have been established and suggest management strategies to maximise long term viability**
3. **To develop effective management and translocation prescriptions for the species.**

Methodology / Strategy

1. **Sampling design: 30 individuals from each of the 6 semi-captive populations of mala will be captured (method described below) and a blood spot will be made on an FTA microcard for each individual then they will be released.**

2. **Blood samples: Blood drops will be taken using a microlancet and stored on FTA microcards. DNA will be extracted from blood samples using the established protocol (commercial kit). This sampling regime is necessary to provide sufficient statistical power to understand population genetic diversity and genetic differentiation between mala populations.**

2. **Genetic assays: Microsatellite and mitochondrial (mtDNA) markers have already been developed for the Mala and similar species. These markers will be amplified from extracted DNA using the polymerase chain reaction (PCR) and genotypes resolved.**

3. **These data will allow even subtle patterns in genetic structure to be detected and categorised. Genotyping at both microsatellite and mtDNA loci will be undertaken at the Macquarie University facility for DNA and the Australian Genome Research Facility (AGRF) analysis using an ABI 3130 sequencer. Well established analytical approaches involving F statistics and estimates of effective population size will be undertaken.**

4. **The results will be used in conjunction with the latest research on adaptation to captivity and consequences of genetic bottlenecks in order to provide a detailed strategy for the management of genetic diversity within the species.**

c) Indicate the significance of the research and the expected benefits in accordance with the Code of Practice 2.2.16 (v).

This project is designed to provide information to the mala recovery team in order to aide the conservation of the species. The IUCN status of Mala is currently extinct and thus survival of the five captive populations is crucial for the survival of the species.

The project will provide data on the management of genetic diversity which is crucial as low genetic diversity increases extinction risk. Active management of genetic diversity is common and crucial to the success of captive breeding programs such as the mala conservation effort.

d) If the project repeats previously reported experiments, give reasons why they are to be repeated.

n/a

e) If the project is based wholly or in part upon your previous work, please provide information about the outcome of these studies. List up to ten of the most relevant

publications in peer-reviewed journals (including papers under review), book chapters, conference presentations, public lectures or any other output that demonstrates benefits accruing from your research. For projects that will take some time to come to fruition, or that have an applied focus (e.g., captive-breeding programs), briefly summarise progress to date.

This is a new project for the conservation genetics lab.

(2-2) Reasons for Use of Wildlife

a) Are there specific target species in this study? Yes☒ No☐ N/A☐

b) If so, what are the applicable characteristics of the specie(s) of animal(s)? N/A ☐

Scientific name	:	**Lagorchestes hirsutus**
Common name	:	**Mala**
Strain	:	**n/a**
Age, sex and weight (if known)	:	**n/a**

c) Why is it necessary to observe and/or capture these animals in the proposed studies and, in particular, why is this specie(s) of animal(s) (strain, age, sex and weight if applicable) the most appropriate specie(s)/model(s) for these studies?

This study aims to provide information on the genetic diversity of these specific populations of mala and therefore the individual animals from these populations must be sampled.

NOTE: In accordance with the Code of Practice, it is incumbent on the individual researcher to determine whether alternatives to using animals in their studies exist (e.g., Code of Practice Appendix 5 “Alternatives to the use of animals”, see also www.animalethics.org.au/).

d) Do alternatives to animal use exist for these studies? Yes☐ No☒

e) What steps have been taken to ascertain the alternatives in accordance with the Code of Practice? (e.g. literature review, etc.)

A literature review suggested that DNA extraction from hair samples was a possible alternative method of non-invasive sampling therefore we conducted a pilot study attempting to extract DNA from pre-existing mala hair samples (provided by the Desert Park) but this was unsuccessful.

f) If alternatives to use and/or capturing these animals exist for these studies, why is it not possible to use them? N/A☐

The alternate methods did not produce DNA of sufficient quality for the genetic analysis that this study proposes.

g) If relevant, what precautions will be taken if lactating animals or animals with pouch young are captured? N/A☐

The field team (Kim Branch and Chris Pavey) have experience zoo keeping, nursing animals and taking blood with the method that will be used for this study. They also have extensive experience with the capture of mala for population health surveys. This study will use the already established methods of mala capture and handling to minimize stress to animals and release animals quickly to avoid stress to pouch young.

(2-3) Number of Animals

a) If the design of the study requires a certain number or estimated number of animals to be observed and/or captured, how many will be required? If the request is for multiple years, provide a year by year breakdown of animals that will be required. Inclusion of a table to show how animal numbers have been calculated is highly recommended.

Number of animals **: 180**
Year by year use **: 1**

b) Explain, on the basis of experimental design and statistical considerations, why this number of animals will be required.

Six semi-captive populations of mala exist which must all be sampled to understand the way that genetic diversity is partitioned among these populations. A sample size of 30 individuals per population is necessary in order to provide statistical power for the approaches that will be used to investigate the distribution of this genetic diversity e.g. Fst, Ne calculations and.

Smaller sample numbers were considered and rejected as the mala captive populations though relatively large were founded by few individuals and thus are likely to have somewhat limited genetic diversity. If genetic diversity is missed due to insufficient sample numbers incorrect management strategies may be suggested.

c) Is opportunistic sampling of other, similar species likely to occur in this protocol?
Yes☐ No☒

If YES, please provide details (eg. possible species, procedures etc). If such opportunistic involvement of other species occurs, a log must be kept, and details reported to the AEC in both progress and final reports.

SECTION 3: ETHICAL CONSIDERATIONS

NOTE: In accordance with the Code of Practice, it is incumbent on the individual researcher to follow the guidelines for wildlife surveys issued by NSW Dept of Primary Industries in an effort to minimise any unnecessary impact on wildlife (e.g., see Code of Practice Section 5 Wildlife Studies; also www.animalethics.org.au/policies-and-guidelines/wildlife-research)

(3-1) Sequence of Events and Impact on Wildlife

a) Outline (sequentially) *what happens to the animal(s)* from the time they are observed, captured, and/or sampled to the time the project is completed. A flow chart or sequence of events is recommended.

1. Mala populations are semi-captive paddock populations thus the field work area will be of limited size (enclosures are 100 to 170 ha in area)

2. Mala will be captured either using the hand net method or via soft -capture Thomas traps as these methods have replaced the Bromilow trap that were previously considered best practice for reducing stress to mala (Langford & Burbidge 2001).

    a. 20 Thomas traps will be baited with fresh lucerne or chopped apples and carrots (depending on seasonal availability). They have previously been used for this species successfully and are endorsed as appropriate traps under the guidelines for trapping small macropods published by the Western Australia Department of Environment and Conservation.

3. Traps will be checked every two to three hours (any non-targeted species will be released immediately)

4. Animal will be placed into a dark bag for processing to block the animals vision and minimise stress

5. A Blood sample (blood drop) will be taken by piercing the animals ear with a microlancet and placing a single drop of blood (<100 µl) onto FTA microcard.

6. Animal will be released.

7. Blood cards will be transported to Macquarie University for genetic analysis

References

Langford D.G. and Burbidge A.A. 2001. Translocation of mala (*Lagorchestes hirsutus*) from the Tanami Desert, Northern Territory to Trimouille Island, Western Australia. *Australian Mammalogy* **23**: 37-46.

b) Describe all factors and procedures that may have an impact on an animal's well-being and how any adverse pain/distress impact will be minimised. This includes methods for passive observation, capture of target and non-target species (where relevant, include the type of traps, how many, over what time period they will be set,

how often, and what times they will be checked and/or cleared), handling, restraint, identification tags and tracking devices (if used, weight, how attached, and retrieved), release, and impact on the general population if known. Also include (where relevant) details of the type and frequency of body tissue and/or fluid sampling, treatment with substances, concentration of solutions (mg/ml) and dose rates (mg/kg), routes of administration (include range of needle gauge to be used), detailed anaesthetic and analgesic regimes, etc. Refer to the CHECKLIST in the **Application Guidelines** to ensure all details have been considered.

Procedure	Risk Minimisation
Animal Capture and handling exposes mala to risk of distress, injury and capture myopathy	**Animals will only be captured using techniques that have been successfully used on the species previously and have had minimal impacts on both individual animals and the population.** **Animal capture and processing will only be undertaken by individuals experienced in both the technique and in handling mala and animals that appear to be stressed will be immediately released.** **To reduce the chance of injury Thomas soft traps will be used.** **To reduce stress to the animal from being in the traps the traps will be checked every two hours and any non-mala animals immediately released.** **To reduce stress animals will be handled in dark bags to block the animal's vision.**
Blood sampling exposes mala to risk of distress and through stress capture myopathy	**We have reduced the volume of blood sampled to a single drop to minimize the stress to the animal. This method is widely used in both Forensics and agriculture.**

c) If transportation is necessary for animals, how will animals be transported, over what period of time and what precautions will be taken against cold/heat stress?
Note: check with the Animal Welfare Officer and/or Fauna Park Manager for Standard Operating Procedures and Guidelines relating to transportation

N/A ☒

d) What precautions, if required, will be taken in the event of inclement weather? N/A ☐

Due to the isolated nature of several of the sampling sites field work will not proceed in the event inclement or dangerous weather patterns (or high fire danger)

(3-2) Animal Monitoring

a) Who will monitor the animals during weekdays and who will be on duty on weekends and holidays? Include the species being monitored. N/A ☒

Names **:**
Weekdays **:**
Weekends/holidays **:**

b) How will animals be monitored during capture, handling and post release? N/A ☐

Animals will be closely monitored during capture and handling for signs of stress (including drooping head and neck, labored breathing, tremors, lethargy and lack of coordination or paralysis) by Kim Branch and Chris Pavey who have extensive experience working with the species.

Animals will not be monitored post release.

d) Please attach a copy of monitoring checklists you will be using.
Checklist attached ☐ N/A ☒

(3-3) Animal Housing, Management and Source

a) Will wildlife animals be housed for any reason? Yes☐ No☒ N/A ☐

b) If so, justify why animals need to be housed and not released immediately, and what type of housing, duration, density (separate or group housed), and feeding arrangements (what and how often) will be provided, and by whom. N/A ☒

(3-4) Emergencies

a) Identify possible emergencies which may arise, (a) animal injured in a trap; (b) other e.g. bushfire, hailstorm.

Animal Emergency

1. **Animal injured during capture or handling**

Staff Emergency

1. **Staff injury e.g. twisted ankle,**
2. **Injury Caused by animal e.g. snake bite**
3. **Environmental risks. E.g. bushfire**

b) List the procedures you have in place to deal with these emergencies, including emergency contacts in the field, e.g. contact number of local veterinary surgeon.

Northern Territory Sites (Alice Springs Wildlife Park, Kata Tjuta National Park, Watarrka National Park)

Emergency	Procedure / Contingency Plan	Emergency Contacts
Animal Injured	The field staff have experience pertinent to the emergency care of animals. Any injured animals will be taken to Alice springs Veterinary Clinic (which is not associated with Macquarie University or the Desert Park)	Alice Springs Veterinary Clinic 75 Bath St, Alice Springs, NT 0870 p: (08) 8952 9899
Staff Member injured	Field staff are experienced in first aid. Field staff will administer first aid on the scene and the injured staff member will be transported to Alice Springs Hospital.	Alice Springs Hospital PO Box 2234, Alice Springs, NT 0871 Gap Road, Alice Springs, NT 0870 ph: (08) 8951 7777 Watarrka Ranger Station (For Watarrka Population Only) Ph: (08) 8956 7460
Injury Caused by animal	Field staff are experienced in first aid including responding to animal bites. Field staff will administer first aid on the scene and the injured staff	Alice Springs Hospital Gap Road, Alice Springs,

	member will be transported to Alice springs Hospital.	NT 0870 ph: (08) 8951 7777
Environmental Risks	Field work will not be carried out on days with high or extreme fire risk	Alice Springs Fire & Rescue Services (for information on fire risk) (08) 8951 6688

Western Australia Site (Dryandra Conservation Reserve, François Peron National Park)

Emergency	Procedure / Contingency Plan	Emergency Contacts
Animal Injured	The field staff have experience pertinent to the emergency care of animals. Any injured animals will be taken to Perth Vet Emergency (Dryandra) Or Chapman Animal Hospital (François Peron)	Perth Vet Emergency 305 Selby St North, Osborne Park (08) 9204 0400 Chapman Animal Hospital 74 West Coastal Highway GERALDTON WA, 6530 (08) 99642828
Staff Member injured	Field staff are experienced in first aid. Field staff will administer first aid on the scene and the injured staff member will be transported to Narrogin Regional hospital (Dryandra) or Geraldton Regional Hospital (François Peron)	Narrogin Regional Hospital Williams Road, Narrogin (08) 9881 0333 Geraldton Regional Hospital Shenton Street, Geraldton (08) 9956 2222
Injury Caused by animal	Field staff are experienced in first aid including responding to animal	Narrogin Regional Hospital

	bites. Field staff will administer first aid on the scene and the injured staff member will be transported to Narrogin Regional Hospital.	Williams Road, Narrogin (08) 9881 0333 Geraldton Regional Hospital Shenton Street, Geraldton (08) 9956 2222
Environmental Risks	Field work will not be carried out on days with high or extreme fire risk	**Bunbury Fire Station** (for information on conditions) 36 Forrest Avenue Bunburry (08) 9721 4644 Geraldton Fire Station (for information on conditions) Durlacher Street Geraldton WA (08) 9921 2222

New South Wales Site (Scotia Sanctuary)

Emergency	Procedure / Contingency Plan	Emergency Contacts
Animal Injured	The field staff have experience pertinent to the emergency care of animals. Any injured animals will be taken to RSPCA Veterinary hospital Broken Hill	RSPCA Veterinary Hospital Broken Hill South Road Broken Hill (08) 8087 7753 After hours 0427 272 549
Staff Member injured	Field staff are experienced in first aid. Field staff will administer first aid on the scene and the injured staff member will be transported to Mildura Base Hospital	Mildura Base Hospital Ontario Avenue Mildura Vic 3500 (03) 5022 3333
Injury Caused by animal	Field staff are experienced in first aid including responding to animal bites. Field staff will administer first aid on the scene and the injured staff member will be	Mildura Base Hospital Ontario Avenue Mildura Vic 3500 (03) 5022 3333

	transported to Mildura Base Hospital.	
Environmental Risks	Field work will not be carried out on days with high or extreme fire risk	Murraylands Regional Office (responsible for the Danggali Conservation park adjacent to Scotia) (08) 8595 2111. (For information regarding conditions)

(3-4) Duration and Fate of the Animals

a) What will happen to animals at the completion of the project?

Animals will be released.

b) Does this project involve the use of any animals or fauna survey sites that have been the subject of previous research? Yes☐ No☒ N/A☐

c) **If YES**, Include project name(s), ARA(s) animal identification number(s), and what has previously been done to these animals ***(FLAG)***. N/A ☐

Project name(s) **:**
Identification No(s). **:**
Animal History **:**

d) Will voucher specimens be taken? And, if so, justify the taking number of voucher specimens, and where the voucher specimens be lodged? N/A ☐

No

e) If animals are to be euthanised, how will this be done? The possibility that euthanasia *may* be required in the field under emergency circumstances must be considered. Please detail how such a situation would be managed and the method by which euthanasia would be carried out. N/A ☐

Euthanasia would only be administered under emergency circumstances in the field in which case it would be administered in accordance to the Guidelines established by the NSW department of primary industries for the destruction of wallabies that being a single shot to the brain with a .22 caliber rifle. If the first shot does not produce a loss of consciousness then a second shot would be administered.
In this case any pouch young would be transported back to the desert wildlife park to be cared for.

It is however important to point out that the investigators consider this scenario extremely unlikely given their extensive experience capturing mala for population monitoring.

f) Where will euthanasia be carried out, who will do it, and what is their experience in the technique to be used? N/A ☐

Location(s)	**:**	At the field site
Name(s)	**:**	Kim Branch
Qualification(s)	**:**	Experience as a zookeeper which has including humane destruction of animals.

g) Could animal tissue be shared with other investigators?

No, as no tissue will be collected.

(3-5) Personnel Qualifications and Technical Experience

a) List names, staff/student numbers, academic and/or licensing qualifications and technical experience, and contact numbers of all personnel participating in the animal components of the project.

	Name	MQ Student/ Staff No.	Relevant qualifications and technical experience in procedures to be undertaken**..	Telephone numbers work and home
Principal Investigator / Supervisor	Dr. Adam Stow	99004372	Adam has studied the population genetics of numerous species including mammals, reptiles, birds as well as several invertebrate species. He will supervise the population genetics work in the laboratory.	Work (02) 9850 8153 Email adam.stow@mq.edu.au
Associate investigator	Vincenzo Repaci	40102262	Vincenzo is currently undertaking a PhD on conservation genetics and has worked on mala genetics previously. He will be carrying out genetics work in the laboratory.	Work (02) 9850 8331 Email vincenzo.repaci@mq.edu.au Home (02) 9747 3082
	Nick Atchison	Staff of Desert Wildlife Park	Curator Zoology, Master of Wildlife Management at Macquarie University – in progress, Nick Atchison has been working as a zookeeper, manager and curator in the zoological parks industry for 25 years within Australia and abroad. He has been Curator Zoology at Alice Springs Desert Park since July 2009. Nick will be responsible for capturing animals and obtaining blood samples.	Phone (08) 8951 8756 Email: nick.atchison@nt.gov.au Fax: (08) 8951 8720
		Staff of		

	Kim Branch	Desert Wildlife Park	Specialist Zookeeper Animal Health, Alice Springs Desert Park Kim Branch is a qualified veterinary nurse with animal care and nursing experience for 16 years. Kim has experience collecting blood samples for research and veterinary purposes. Kim has been responsible for managing specimen collection and data base for greater than 10 years. Kim will be responsible for capturing animals and obtaining blood samples.	Phone (08) 8951 8761 Email kim.branch@nt.gov.au Fax: (08) 8951 8720
	Dr. Chris Pavey	Regional Manager, Biodiversity South, Biodiversity Conservation Division, NRETAS	Chris has extensive experience working with mala in the field as has been involved in numerous mala capture events and translocations. He chairs the national Mala recovery team.	Work (08) 8951 8249 Email chris.pavey@nt.gov.au

***Include details of experience with the <u>actual species</u> used in the protocol* (e.g. details of previous degree, experience in animal handling and procedures, training in standard operating procedures). *If none, detail how*

b) Have any of the people participating in the project had any animal research authority or animal supplier's licence cancelled? Yes☐ No☒

d) **If YES**, provide name of the person, the date on which the authority/licence was cancelled, who cancelled the authority/licence and reason for the cancellation N/A ☒

(As section 4 is a declaration requiring signatures I'll run a hard copy of this to the ethics office)

Appendix 0 A record of the excel macro used to reverse compliment DNA

The source code for the Excel Macro which reverse complements DNA, the Code written by Leo Heuser on http://www.pcreview.co.uk forum.

```
Excel Macro for reverse complementing sequence Function RevDna(Cell As Range) As Variant
'Leo Heuser, 3 Aug. 2004
Dim CellValue As String
Dim Counter As Long
Dim DNA As String
Dim Dummy As Long

RevDna = CVErr(xlErrValue)

CellValue = UCase(Cell.Value)
DNA = "ACGT"

For Counter = 1 To Len(CellValue)
Dummy = InStr(DNA, Mid(CellValue, Counter, 1))
If Dummy = 0 Then Exit Function
Mid(CellValue, Counter, 1) = _
Mid(DNA, Len(DNA) + 1 - Dummy, 1)
Next Counter

RevDna = StrReverse(CellValue)
End Function
```

Appendix 1 Paper Published during PhD candidature

Pages 265-273 of this thesis have been removed as they contain published material under copyright. Removed contents published as:

Duckett Paul. E., Repaci Vincenzo (2015) Marine plastic pollution: using community science to address a global problem. *Marine and Freshwater Research* vol. 66, pp. 665-673.

https://doi.org/10.1071/MF14087

Appendix 2 Additional information and details for Chapter 2

Appendix 2.1 DNA extraction using a 5% Chelex suspension

DNA was extracted from between 5-10 plucked hairs as per Baldwin *et al.* (2010) modification of Gagneux *et al.* 1997., their methodologies have been reproduced below.

Gagneux *et al.* 1997

"DNA extraction from shed (telogen) hair Following current standard methodology, two extraction methods were used. The first is an extraction in 200 pL 5% Chelex-100 (BioRad) suspension: the hair is first washed with 70'X ethanol and de-ionized distilled water and then 3-5 mm of the proximal end with the hair follicle is cut off and allowed to drop into the Chelex solution. All handling is carried out in a closed UV isolation cabinet wearing a face mask and using sterile forceps and scissors that are washed in bleach, and rinsed in alcohol and (DI) water between each sample treatment. The solution is then heated at 56 "C for 2 h, vortexed at high speed for 10 s, put in boiling water for 10 min, vortexed again for 10 s and finally centrifuged for 3 min at 16 000 g. The extracts are stored in the dark at 4 "C. The second method of extraction involves a digestion of the cut off, proximal 3-5 mm portion of the hair with protease K overnight at 56 "C in a hair lysis buffer (10 nM Tris pH 8.0, 50 mg/mL Protease K, 35 mM DTT, 0.9% Laureth lo), before Chelex is added (C. Orrego, personal communication). Subsequent procedures follow the outline above. Extractions were all made from one hair at a time. Sample preparation and PCR set-up were performed in two different closed UV hoods. Each batch of DNA extractions was controlled for contamination by amplifying at least two Chelex blanks"

Baldwin *et al.* 2010 modification

"In a 5% Chelex suspension following Gagneux et al. (1997), with the following modification: hair was added to the solution without washing in ethanol and water (n = 4). Between four and 10 hairs were used per extraction".

Appendix 2.2 Constructing a Neighbour joining tree using MEGA 5.5

The methodological justification produced by Tamura *et al.* 2011 is reproduced below.

"This command is used to construct (or Test) a neighbor-joining (NJ) tree (Saitou & Nei 1987). The NJ method is a simplified version of the minimum evolution (ME) method, which uses distance measures to correct for multiple hits at the same sites, and chooses a topology showing the smallest value of the sum of all branches as an estimate of the correct tree. However, the construction of an ME tree is time-consuming because, in principle, the S values for all topologies have to be evaluated and the number of possible topologies (unrooted trees) rapidly increases with the number of taxa.

With the NJ method, the S value is not computed for all or many topologies. The examination of different topologies is imbedded in the algorithm, so that only one final tree is produced. This method does not require the assumption of a constant rate of evolution so it produces an unrooted tree. However, for ease of inspection, MEGA displays NJ trees in a manner similar to rooted trees. The algorithm of the NJ method is somewhat complicated and is explained in detail in Nei and Kumar (2000).

For constructing the NJ tree, MEGA may request that you specify the distance estimation method, subset of sites to include, and whether to conduct a test of the inferred tree through an Analysis Preferences dialog box."

Given this justification the following parameters were used (Justifications from Tamura *et al.* 2011)

Parameter	Input value
Tree Type	Neighbour Joining tree (Which is more accurate than Minimum Evolution tree with relatively short DNA sequences)
Distance estimation method	We have no specific information on this for mala so system defaults for MEGA 5.5 were used
Subset of sites to include	We have no specific information on this for mala so system defaults for MEGA 5.5 were used
Is a test of the inferred tree included	Given the small sample size MEGA cannot calculate a Bootstep test on this data

Appendix 2.3 The results of molecular analysis of mala populations

Table 2.31: The genotypes of mala used in the studies at 4 microsatellite loci, the genotype at each loci is given by a four digit number where two digits represent an allele. Therefore, 3434 is a homozygote of allele 34 whereas 2334 is a heterozygote with allele 23 and allele 34. In this dataset unsuccessful amplifications are shown using a 0.

Uluru Samples	Loci 1	Loci 2	Loci 3	Loci 4
UU1,	0	0	0	0
UU2,	5261	6262	7373	7179
UU3,	5555	7982	7575	7070
UU4,	5056	7982	7070	7179
UU5,	5356	6262	7175	7171
UU6,	5261	0	7373	7179
UU7,	5061	7979	7375	7171
UU8,	0	6279	7373	7171
UU9,	5256	6279	7373	7171
U10,	5052	6267	7575	7179
U11,	5256	6262	7575	7171
U12,	5361	6279	7575	7171
U13,	5256	6262	7575	7171
U14,	5661	6268	7373	7171
U15,	5052	6279	0	7171
U16,	5256	6279	7373	7179
U17,	5656	6262	7373	7171
U18,	5252	6278	0	6471
U19,	5656	6278	7575	7179
U20,	5056	6278	7575	6471
U21,	5356	6282	7575	7179
U23,	5252	6282	7575	7179
U24,	5052	6278	7373	7171
U25,	5256	6262	7575	7171
U26,	5261	7878	7373	6471
U27,	5056	6278	7575	6471
U28,	5656	6262	7575	7171
U29,	5353	6262	7373	7171
U30,	5662	6267	7375	7171
Watarka Samples				
AA1,	5656	6278	7373	7979
AA2,	0	7878	6975	7171
AA3,	5661	6278	7398	7171
AA4,	5656	6262	7373	7179
AA5,	5662	6267	7373	7171
AA6,	5362	6278	7575	7171

AA7,	5656	7878	7575	7171
AA8,	5356	6278	7373	7171
AA9,	5356	7878	7373	7171
A10,	6161	6284	7575	7171
A11,	5050	7878	7575	7171
A12,	5661	6262	7373	7179
A13,	5356	6262	7373	7171
A14,	5353	6278	7373	7171
A15,	5662	6278	7373	7171
A16,	0	7878	7373	7171
A17,	6161	6262	7575	7171
A18,	5656	6278	7575	7171
A19,	5261	6778	7373	7171
A20,	5252	6278	7575	7171
A21,	5356	7878	7575	7171
A22,	5353	7878	7575	7171
A23,	5062	6278	7575	7171
A24,	5353	7878	7575	7171
A25,	5361	6282	7575	7171
A26,	6161	6267	7575	7179
A27,	5261	7884	7575	7171
A29,	0	6284	7575	7179
A30,	5656	0	0	7979
Alice Springs Desert Park Samples				
A31,	5053	6284	7575	7171
A32,	5053	7884	7575	7171
A33,	5662	6267	7175	7179
A34,	5053	6278	7373	7171
A35,	5252	6284	7575	7171
A36,	5362	6284	7575	7171
A37,	5053	6284	7575	7171
A38,	5053	7884	7575	7171
A39,	4951	7884	7575	7171
A40,	5353	8484	7575	7171
A41,	5353	6784	7373	7179
Peron Samples				
PP1,	5662	7878	7375	7979
PP2,	5656	6778	7575	7171
PP3,	5053	7878	7575	7179
PP4,	5362	6278	7070	7171
PP5,	5356	6267	7575	7171
PP6,	5656	7884	7575	7171

PP7,	5062	7878	7175	7979
PP8,	5361	6262	6969	7171
PP9,	5261	6767	6975	7171
P10,	5656	6278	6975	7171
P11,	5353	6778	6969	7071
P12,	5362	6267	7070	7071
P13,	5656	6778	7575	7171
P14,	5662	6278	7575	7179
P15,	5056	7884	7575	7171
P16,	5656	6267	7575	7171
P17,	5656	6267	7575	7171
P18,	5256	7884	7575	7179
P19,	5656	6262	7575	7171
P20,	5356	7884	7373	7071
P27,	5062	7884	7575	7179
Scotia Samples				
SS1,	5356	6278	6275	7171
SS2,	6262	6262	7375	7171
SS3,	5353	6262	7171	7171
SS4,	5356	7878	7171	7179
SS5,	5353	6278	7171	7171
SS7,	5353	6262	7373	7171
SS8,	5361	6278	7171	7179
SS9,	5353	6278	7171	7171
S10,	5662	7878	7171	7179
S11,	5353	7878	7373	7171
S12,	5362	7878	7171	7179
S13,	5353	6278	7171	7171
S14,	5353	6278	7171	7179
S15,	5353	6278	7171	7171

Table 2.32 The results of mtDNA sequencing on 107 mala. The first column is the name of the individual samples. The second column, Quality, denotes sequence quality, where 1 is high quality sequence, 0.5 is generally high quality sequence with areas of low quality and 0 is low quality sequence. Only high quality sequence (1) was used in subsequent analysis. The third column is the sequence and the fourth column gives the population and a sample number within the population for each individual.

Names	Quality	Sequence	Population
A40	1	CACGCAAAATTTCAGCGTTAGGTAATTATTGGTAGGGCCTGCTCTGAAGGATAGG ATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACT AATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAA ATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTA TATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTT ATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATG CTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATT TAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGT TTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGAT TATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGA GTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTA TGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGG GGGTTGGAAGGGAAAAATGGTTA	Desert Park 10
A41	1	CCCTAATTTGCTAGGTCGACCGCAGTTTATGGGCCTGCTCTGAAGGATAGGATGC CTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATG GAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGT CTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATG TACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGT ATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAG GGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAAT GTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAG TAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATG TAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGT GATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCA AATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTGTGGTTTT AAGAGGAAAAAAAAAAAAATGGAA	Desert Park 11
A32	1	GCGCCAGTATTTGCGATGTGATGGTACAGTTTATAGGCCTGCTCTGAAGGATAGG ATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACT AATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAA ATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTA TATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTT ATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATG CTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATT TAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGT TTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGAT TATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGA GTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTA TGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGG GGTTTTGAAGGGAAAAAAGGGT	Desert Park 2
A33	1	CCCAAAAATTGCTAGGTCGACTGACATTTATAGGCCTGCTCTGAAAGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGTTT TTTAATGGAAAAAAAGGAAAAGT	Desert Park 3
A34	1	GGGCATCATTGGGATGTGATGGTACATTTATAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG	Desert Park 4

		TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTGGGGTT TTGAAGGGAAAAA	
A36	1	CCCAAAATTTCTAGGGTCGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTG TTGATGGTAAAATCCCCACCCCCAGGGGCTCTCCTGTCGTTTTTCCACCTTAACCC CCCTTCTTCCTTCCCTCCC	Desert Park 6
A37	1	CAAAATACTATATACCAAATTGCTGGTCGACCGCAATTTATGGGCCTGCTCTGAA GGATAGGATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGG TATTACTAATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCAT TGATGAAATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGT ATGTTTATATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAG CTGGTTTATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGC TTATATGCTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAAC ATTGATTTAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTT AATGTGTTTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTA CTAGGATTATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGAT GGGTTGAGTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATC GAGGCTATGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCA GCTTTGGTGTTTGATGGT	Desert Park 7
A38	1	CCCCAAAATTTTTCTAGGCACGACCGCAATTTATGGGCCTGCTCTGAAGGATAGG ATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACT AATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAA ATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTA TATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTT ATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATG CTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATT TAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGT TTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGAT TATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGA GTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTA TGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGG GTGTTGATGGTAAAAAAAAATCACTCCACCTCCTTTTTTCCTTCCTAAAACCCTCC TCCTCTCTCTTCCTCCTCTCCTCTTCGTCCTCTTT	Desert Park 8
A39	1	CACCAAAAATTGCTAGGTCGACCGCATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGG TTTGATGGTAAA	Desert Park 9
P01	1	GACGTAGGTCTCATCGTTGAGTGTTGATGGAAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC	Perron 1

		AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTGGGTGT TTTGGATGGAAAAAGGGGGAAAA	
P10	1	CCCAAAATTCCTAAGCTCGACCGCTATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TGATGGTA	Perron 10
P11	1	ACCCAAAATTGCTAGGTCGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTG TTTGATGGTAAAACCCCTTTTTTTTTG	Perron 11
P14	1	TACCCCCAAATTGCTAGGTCGACCGCATTTATGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCTGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGG GTTTGATGGTAAAACTTTTTTTAT	Perron 14
P15	1	CCCAAAATTGCTATGTCGACAACATTTATGGGCCTGCTCTGAAGGATAGGATGCC TAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGG AAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTC TATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGT ACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTA TTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGG GGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATG TGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGT AATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGT AATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTG ATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAA ATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGGTTGT GAAGGGAAAAAACCAAAAAACACAAAAAA	Perron 15
P16	1	ACCAAAATTGCGTTGCTACTAGTACATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGGTT TTTAAAGGAAAAAAAAAAAAGTTAA	Perron 16
P17	1	GGCCGAGATGCTGCTCGGTCGAGTGTAGATGGTAGGCCTGCTCTGAAGGATAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT	Perron 17

		AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGT GTTTTGATGGAAAAATGGGAAA	
P18	1	GCCCCAAATTGCTAGGCACGACCGCTATTTATGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTG GTTTGAAGGGTAAAAATTA	Perron 18
P19	1	CCCCAAATTGCAGGGTCGACCGCATTTATGGGCCTGCTCTGAAGGATAGGATGCC TAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGG AAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTC TATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGT ACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTA TTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGG GGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATG TGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGT AATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGT AATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTG ATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAA ATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTTTT GATGGTAAAA	Perron 19
P20	1	CCCCAAATTTGCTAGGATCGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGT GTTTGATGGTAAAATTTTTTTTTTTAAATT	Perron 20
P03	1	CCCCCAAAATTGCCTAGCACGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGG TGTTGATGGTAAAATGGGTTGTGCCAAGCTTCCCTGTTTACTCCTACTGACTCGCG AAGATCCGTCCGCAGATTA	Perron 3
P04	1	CTATCCCCAAATTTCTAGGTACGACCGCAATTTATGGGCCTGCTCTGAAGGATAG GATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTAC TAATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGA AATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTT ATATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTT TATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATAT GCTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGAT TTAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTG TTTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGA TTATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTG AGTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCT ATGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTG GGTGTTGATGGTAAAATTCCT	Perron 4
P08	1	CGCTAATTTGCGTAGGTGTACTGACATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT	Perron 8

		GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGGTT TTGAATGGAAA	
P09	1	GAGCAAATCTCATCGGTCGAGTGTTGATGGTAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TTGAATGGAAAAAAAGGAAAAGAATATAAA	Perron 9
S01	1	ACAACTCCCCAAATTGCTAGGCTCGACCGTCATTTATGGGCCTGCTCTGAAGGAT AGGATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATT ACTAATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGAT GAAATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGT TTATATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGG TTTATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTAT ATGCTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTG ATTTAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATG TGTTTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAG GATTATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGT TGAGTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGG CTATGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTT GGGTGTTGATGGTAAAA	Scotia 1
S11	1	GGCCTAATCTCATCGTCGACTGTAGATGGTAGGCCTGCTCTGAAGGATAGGATGC CTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATG GAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGT CTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATG TACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGT ATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAG GGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAAT GTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAG TAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATG TAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGT GATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGGCTATGC AAATGGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTG TTTTTAAAGGGAAAAAGGGAAAAAG	Scotia 10
S02	1	ACCATAATTTGCTAGGTCGACTGACATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTGGGGGG GGGGGGGGAAAAAAAAACCAA	Scotia 2
S03	1	GGGGCATCTTTGCTTGTTGATGGTACATATTATGGACCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG	Scotia 3

		CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGTTTGGGGT GTTTAGGGGAAA	
S07	1	GCGCCTAATCTGCTACGTCGACTGTAGATGGTAGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTG TTTTGGATGTAAAAACGGG	Scotia 6
S10	1	GGCCTAATCTCATACGTCGACTGTAGATGGTAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TTGATGGTAAAAAGGGGTAAAA	Scotia 9
U10	1	ACCGTCCCCCCAATTTCTACGTCGACCGCATTTATGGGCCTGCTCTGAAGGATAGG ATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGAATTGAGGTATTACT AATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAA ATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTA TATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTT ATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATG CTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATT TAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGT TTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGAT TATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGA GTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTA TGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGG GTGTTTGATGGTAAAATTTTT	Uluru 10
U12	1	CCCCCAAATTTGCTAGGTACGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGT GGTTGGATGGTAAAATTTTTA	Uluru 12
U13	1	CCAAAATTATGCTAGTTAATACCACAATTTATGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGT GTTTGATGGTAAAA	Uluru 13
U15	1	GGGGCACTTTTCGGTGTGATGGTACAATGGTAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA	Uluru 15

		TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TGATGGT	
U16	1	GGGCACATTTGCGTGGTGATGGTACATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGG TTTGAATGGAAAGG	Uluru 16
U17	1	AGGGGGATCTTTGGGTGTGATGGTAGCTTCTAGGCCTGCTCTGCGGGAGCGGATG CCTCGCTGGGCGGGATGGTGGTTTCTCCTGCGTGGGGGATTAGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGT GTTGATGGTAAACCCCCCCCCCCCCCCCCGGCGCCGGGGGCGCCGCCCTTTTTTTT GGTTTTTTTTGT	Uluru 17
U19	1	CCAAAAATTTCTAGGGTCGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGGT TTGATGGTAAAAGGAAAGGGGGGGGTGCG	Uluru 19
U02	1	AACCCCTAATTGCTAGGTCGACCGCATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGG TTTGAATGGAAAAAGGTC	Uluru 2
U20	1	CCACAAAATTGCCAGGTACGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGT GTTGATGGTAAAAAAAAAAGGGAGGCCCGGTTCGTCTCTTTTTCCCTTATCCTCCT CATCGCTTTCTCCCTCCCTCGTCTTACCTGGCCCCCTCTT	Uluru 20

U21	1	ACCCAAATTTGCTATGTCGATAGCTATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTG TTTGATGGTAAATTGGGGGTTGTTGTT	Uluru 21
U23	1	GCGCTAATTTGCTAGTTCGGGTACAGTTTATAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTGGGGTT TGTGGGGGGAAAAAAGGGAAAAAATAATTTAATTTC	Uluru 23
U25	1	GGGGCATCATTCCGGTGTGATGGTACATATATGAGGCCTGCTCTGAATGATAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGGAAACTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGG GGTTTGATGGTAAACC	Uluru 25
U29	1	CCCCAAAAATTGCTTGAGTACTACCGCACTTTATGGGCCTGCTCTGAAGGATAGG ATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACT AATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAA ATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTA TATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTT ATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATG CTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATT TAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGT TTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGAT TATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGA GTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTA TGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGG TGGTTTGATGGTAAAAAAATAATTTTTTTTA	Uluru 29
U03	1	CCCCAAATTGCTAGGTCTACCACATTTATGGGCCTGCTCTGAAGGATAGGATGCC TAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGG AAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTC TATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGT ACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTA TTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGG GGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATG TGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGT AATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGT AATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTG ATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAA ATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTGTTG ATGGTACTTTCCTTCTTCCTTTTTCTTCTTCTCTCCTCTTTTTTTCTTCCTTTTTCTTC CCCCTTCCTTCCTTCCTTTCTCCCCTCTTCCCTTTTTCTCTCTTTTTTTCTTTCTTCCC TCCTCTCCTCTCCCCTCCCCTCTTCCTTCCCCTTCCTTCTCTTTTTCCCTTTTCTCTTT TNCTTCTTTTCCCCCCCTCTCCCTCTTCCTTCTCTCCCCCCCCCCCCCCTCCCCCCTC CCCCCTCTCTTCCCCCCCTCCTTCCCCTTCTCCTCTCTCCTCTCTTTTCTCTTTTTTCT TCCTCTCTTTTTCTCTTCCTCCTTCCTCCCTTTCCCTTCTCTCCCTTTCTCTCTCTCTC TCCCTCTCCTCTCCCCCCTCTTTTCTCTTCTTCTTTCCTTCCTCCTCTCTCTCTCTC CTTTCTTTCTCCCCCCCCCCCCTCCTCCCCTTCCTCTTCCCTCCCTCTCCCTTCCCCC CTCTCCTCTTTTCNCCTCCTCTCTTNCCTCTTCTTCCTCTTCCCCTCTTTCCCTCTCTT TTTCTTTCTTCTTCTTTCTTCTCTTCTCCCTCTCTTCTTCTCTTTTCCCCTCCTCCTCT TCTTCCCCCCTCTCCTCTCTCTTCCCTTCTTTCTCCCTTCTCTCCCCTCTCCCCT	Uluru 3

U05	1	GACCCGAGATTGCTAGGTACGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGG GTTTGAAATGGAAAAA	Uluru 5
U06	1	CAACCCCCAAATTCATGGTACGACCGCAATTTATGGGCCTGCTCTGAAGGATAGG ATGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACT AATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAA ATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTA TATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTT ATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATG CTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATT TAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGT TTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGAT TATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGA GTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTA TGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGG TGTTTGATGGTAAAAATT	Uluru 6
U07	1	CCCAAGATTGCGTGGTTCGACTGCATTTATAGGCCTGCTCTGAAGGATAGGATGC CTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATG GAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGT CTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATG TACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGT ATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAG GGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAAT GTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAG TAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATG TAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGT GATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCA AATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTGTGGTTG GGGGGGGGAAAAAAAAGGAAAAG	Uluru 7
U08	1	CCCTAAATTGCTACGGTCGACCGCAGTTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGTTGGGGGTG AAAAGGGGAAAAAAAGGG	Uluru 8
U09	1	CCTAAATTGCTAGGTCGACCGCAGTTTATGGGCCTGCTCTGAAGGATAGGATGCC TAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGG AAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTC TATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGT ACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTA TTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGG GGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATG TGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGT AATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGT AATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTG ATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAA ATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGTTTTG TTGATGAAAA	Uluru 9
A10	1	GACAAATTTGCGATGGTCGTATGACATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT	Watarraka 10

		GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGGTT TGAAAGGGAAAAAAAGGGAAAGGAAAG	
A11	1	ACAGAAATTGCTAGGTCGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGATGC CTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATG GAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGT CTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATG TACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGT ATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAG GGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAAT GTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAG TAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATG TAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGT GATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCA AATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGGTTT TAAAAGGAAAAAAATTTT	Watarraka 11
A12	1	CCGAAATTGTCTAGTCGACCGCATTTATGGGCCTGCTCTGAAGGATAGGATGCCT AGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGGA AGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTCT ATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGTA CTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTAT TCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGGG GTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATGT GGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGTA ATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGTA ATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTGA TAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAAA TGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTGTTTG ATGGTAA	Watarraka 12
A14	1	GCCAAAATTTGGTTGTTCTAGTACATTTATGGGCCTGCTCTGAAGGATAGGATGCC TAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGG AAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTC TATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGT ACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTA TTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGG GGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATG TGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGT AATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGT AATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTG ATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAA ATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTGTTT GATGGTAAAAG	Watarraka 14
A15	1	CCCAAATTTTCATAGTCCTACCGCAATTTATGTGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCTTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTG TTGATGGTAAAACCCCCCCCCCCCCCCCCCCCCCCTCTCCTTTTCTCCTTTATCCTCT CCACACCTTAATCACTAACA	Watarraka 15
A16	1	GCCAAAATCTGCTAGTTCATAGTACTATTTATGAGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGG TTTGGAAGGGAAAAAAAGAAAAGTTTGAGTGTA	Watarraka 16
A17	1	GCCCATCATTGCGTGTTGATGGTACATTTATAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT	Watarraka 17

		GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTGGGGGT TGGAAGGGAAAAACGGAAAAC	
A18	1	GGGCAACATTGCGTTGGTGATGGTACACTTTATGGGCCTGCTCTGATAGAAAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGT GTTTTGATGGAAAAAAGGAAACCTAAAA	Watarraka 18
A19	1	CCCCCAAATTTGCTAGGGACGACCGCATTTATGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGG TTTTTAATGGAAAAATTT	Watarraka 19
A02	1	AGCAAACTCTGCTGCGTAGAGAGACATTTATAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGG TTTGAATGGAAAAATGGTAAATT	Watarraka 2
A23	1	ACCCCAATTTGCTTATGACGACAGCATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TGATGGTAAAATTTTTTTTTTTTT	Watarraka 23
A24	1	CCCAAAATTTGCTAGGTCGACCGCAGTTTATAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGGTT TGTAAGGGAAAAAAAGAAAAAGGGAAAGGT	Watarraka 24

A25	1	CCCTAAATTGCTTGTTCGACGACATTTTATGGGCCTGCTCTGAAGGATAGGATGCC TAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGG AAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTC TATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGT ACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTA TTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGG GGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATG TGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGT AATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGT AATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTG ATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAA ATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGTGGTTTT GAAGGGAAAAAAAGAAAAAAAAGTGTTAATA	Watarraka 25
A26	1	CCCACAAATTGCGTGTTCATGATACATTTATGGGCCTGCACTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGG TTTGAATGTAAAAAAAGGAAAGGTTATTGTTTG	Watarraka 26
A27	1	ACAACCCAAAATGCTAGGTCGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGG TGTTTGATGGTAAAAA	Watarraka 27
A29	1	CCCAAAAATTGCTGGTACGAACCGCATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGG TTTGATGGTAAAATTTTTTTTTTTTTGTTTT	Watarraka 29
A03	1	CCCTAATTGCGTAGTCTACTACATTTATGGGCCTGCTCTGAAGGATAGGATGCCTA GATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGGAA GGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTCTA TGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGTAC TTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTATT CTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGGG GTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATGT GGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGTA ATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGTA ATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTGA TAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAAA TGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGGTTTG AAAGGAAAAAAAGTGAA	Watarraka 3
A30	1	CACCAAAATTTGCTGGTACGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA	Watarraka 30

		TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTG TTTTGATGGTAAAATATTTTATTTTTTT	
A04	1	CCCAAATTTTTAGGCACCACCAGCATCTTATTGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGT GTTGATGGTAAATTCCCTCTCCCTCCTGTCTTCCGCCCCGCCTCCACGGGGAGGGG CCAGCGCCGCTTTCCTTCGTTCAGCAA	Watarraka 4
A06	1	CCCAAAATTTGCTAGGTCGACAGCAATTTATGGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTG TTTGATGGTAAAAGGTTGTGTTTGTTTTTTT	Watarraka 6
A07	1	CCCCAAAATTTGCTAGGTACGACCGCAATTTATGGGCCTGCTCTGAAGGATAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGG GTTTTAATGGAAAAATTTTAATTGT	Watarraka 7
A09	1	GCGGCATTTTTGCGATGTGATGGTATAGATGGTAGGCCTGCTCTGAAGGATAGGA TGCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGG GTTTGGAAGGGAATAATGGGAA	Watarraka 9
S11	1	GCGATAATTTCGACGTTCGACTACAGTTTATAGACCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGG TTTTGAAGGGAAAAAAAGGAAAACCAAAAGGGCAAACA	Scotia 11
P02	0.5	ACCCCACCCCCCTCGCCTATACCCGCCCCCCTTTTATGGCCCTGCTGTGAGGATAG GATGCCTACATGGGCGGGATGGCGGCCTCTCGTGAGTAGGCACAGTAAGGTATTA CTAATGGAAGGGATATGCTAGAACGGTACGCTTTAAGATTCATAACTCATTGATG AAATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTT	Perron 2

		TATATGTACTTTGTGAGTTTTTATGTAATTAGGAGATATGCTAGTGTTGAAGCTGG TTTATGTCTTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTAT ATGCTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTG ATTTAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATG TGTTTTAGTAATATGTGAAGTTGTTAGTCTTTTGTTATCATTGAATTTTATGTATTA CTAGGGATTATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGA TGGGTTGAGTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAAT CGAGGCTATGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATC AGCTTTGGGTGTTGATGGTAAAAGGCCGCGCACCTTTTCCGACCTCCGCCGGTCGT CTCTCTTTCTCCTTGGT	
S12	0.5	GGCCATATCTGCTACGTCGACTGTAGATGGTAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TTTAATGGAAANNNNNNNNANGNNTATGNTTTNNNNTTNTTTTTGATNTGTNNTT ANTGTATGTTTTNTTTTTTTTNTTTTTTTTNNTTTNTTTTTTTTNTTTTTTTTTTTTTTN TTTTTTTTTTTTTTTTTNTTTTNNNGTTNTTTTTTTTTTTTTTTTNTNTTNGGTTTTNT TTTNTTTTTTTNNTTGGTNTTGTTTTTGGGGGGTGGTTTGTTGTTTTGGTGTGTGGT TTGGTTTTTTGTTTGTTGNTTTTTGGGTTTNTNTTGGTTTGGTNTTTTGGGGGNNNN NNNNNGNNNGNNGNNNNNGGNGGNNNNGGGNNGNGGNNNNNTNNTGNNNN GGNGNNNNGGNNNNNNGGTGTNGTGNGGNGNGGNGGGNNGGNNNNNNGNNGGT GGGGGGGNNGGNNGGGGNNNNNGGGGNGGNNNNGGNGGGNNGNGGGNNNGGGG GGGGNNNNNGGGNGGGGNNNNGGTGGNNGNNGNGNGGGNNNGNNNNGNGTGN NNNTGNNGNGTNNGGGGGGNNGGGGTNNGGGTNNNNNGNGNTNGGGTTGTGNG TTTGTGTGNNNNGNNTTG	Scotia 12
S13	0.5	CCCCATCTGCTCATACGTCGAGTGTAGATGGTAGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTG NTTTGATGGAAANNNGNNNAGGGNATGNTNTGCNTTNTNCTTGATGTGTGTTGGG GGGTTTTTTTTTTTTTTTTTTTGGGGTTTTGGGGTTTGTGTTTTGTTTTTGGGTGNGT TGGTTTTGTTGGTGTGGGGGTGTNGTGTGTGNTTTTGNTGTGGGGTTGGGTTTTGT TGTGTGTGTGGTTTTGGTGGTTGGGGTTGTGTNTTTGGTGGTGGGTGGGTGGGGTG GGGGTTGTTGGTTGGGGGGTGTGNNGGGGNTGTGTNGTGGGTGGGGGGGGGGTT GNTNGNGGGGTGGGGGGGGGGGTTGGTGTGGGGTGTGGGTGGGGGGGGGGGGG GTGGGGGGGGTGGTGGGGGGGGGTTNGGNGTGGNGTGTGGGGTNGGGGGGNGGTG GNNGGGGGGGGGGGGGTGGGGGNGGGTGNGTGTGGGGGTTGGGGGTTGGGGTGGT NGGTGGGGTGGGNGGTGGTGTGTTGGGNGGTTGGGTGGGGGGTTGTTGGNGGGN TGGNGTGGGGGGGTGGGGGGTGGGGGGGTGTGGTGGGTGGGGGGGGGGGGGGG G	Scotia 13
S14	0.5	CCGCATAGTCTGCTACGTCGAGTGCAGATGGTAGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTG TGTTTAATGTAAANANNANAAANNNNAATNNTTNNNATTNTTATTAATTTTTTAT AAAAAAATTTATTAATANTTTAATTTTTANNNTTTATAANAAATATTATTTTTTTA TTAATTAATTTNTTTTNTTAATTTTNAATTTAANATTTTTTNTTTTTATTTTTAATTT NNTATTNTTTTAAAANNANANNAATAATTTTNNANATNTAATTANNTAAATTTTT NAATTNNNTAATNNTTTATTANTTTANNAANNNNNNNNNANNNNANNNNNANNN NANNNNNNNNNNNNANNNNNNNNANNNNNNNNNNNNNANNNNNNNNNNNNT NNNNNNNNNNANNNNNNNNNNNNNNNNNAANNNNNNNNNANNNNNNNNNNN	Scotia 14

		NNNNNNNNNNNNNNNTNNNNNNNNNNNANNNANAAANNANNNNNNANNNNNN NNNNNNNNNNNNNNNNANNNNNNNTNNNNNNTNNNANNTNNNNNNNNNNN NNNNNNNNNNNNNNTNNNNNNNNNNNNNNNNNNNTNNNNNNNNATNNNNNNN NNNNANNNNNNNNNNNNNNNAANNNNTTNNNNNNNNNNANNNNNNNAANNNN NAANNANNANNANTNNNANNNNNANNNNNNANANNANNANN	
S15	0.5	CCCAGAAGCTCATCTGTCGACTGTTGATGGTAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TTGGATGTAAAANNNGGNAAANNNNNNNNNNNNNNNNNTNNNNNNTNANNGTN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNGNNNNNNGGGNNGGNNGGNGNGGG GGGGGGNNNGGGGNNGGNNGNNNNNGNGNNGNNNNGGNGNNGNNNGGNNGGN GGNNNGGNNNGNGGGNGNGGNNNGNNGGNNNNNGGNNNGGGNGGGNNNNNNG NGGGNNNGGGNGNGGNNNGGGNNGGNNNNNNGGNNGGGNGGNGNGNGNGGNN GGNNNGGGNNGGGGGGGGGGGGGNNGGNNGGNNNNGNGGNGGGNGNGGGGG NGGNGNNGGNGGGGNGNGGGNNNGNNGNGGGGNNNNNGNNGNGNGNNGNNNG GNNNNNGGGGGGGNNNNNGNNNNNNNNNNGGNNGNNNNGNGNNNNNGNGG NNNNNGGNNNNNNNNGNNGNGGGGGNNNGGGNNNGGGGNGGGGNNGGGGGG NGGNNGGNGGGGGNNNGGNNNGNNGGNNGGNGNNNNNNNGNGGNGGGNNNNG GNNNGNNNGGGGNGNNGGGNGNNNGGNGGGGGGGGGGNGNNNGGNNGGGNNG NNNG	Scotia 15
S04	0.5	CGCCTAATCTGCTACGTCGACTGCAGATGGTAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TTTGATGTAAANANNGGAAAGNGNTNNNNNNTGNNTTNTTNTTGATGTGTTGTTA GTAAGGTGTGATTGTTTAAATTTTTGGGGTTTATAGAAAATATATTTTTATTTTTA ATAGNTTGAAATTGTTAATTTTNGAGTAAAGATATATGGTGTNGAATTTAGGTAT GGAATAGGGTTAAAAAAGGGTGATGAATTAGGAGGTAAGTAAATAGTAAATTGT TTAAGGGGGTAATTGGTGGTAGATTTGAGGAAGANANNAGGNTGANNNNGTAGG GNNNNGNNGNNGNGGNGNNGGGNGNNNNTNNNNNANNTNNNNNNGGNNNNNG NGGNNNNNGATTGGGNNGNNNNNNGGAGNGGNNGGAGGAAATGNGNAGAAGN NGGNNNGANANNNNNGNANGNGNNNGNNAGTGGGNGAAGGNNGNNGTNGANG GGGGNNGGGGGGNGGGNGTGAGGATGGGGNNGGNNGGTNNAGTGNNGGGGGGG NAGGTNNGNGGNGNGGGANNGNTGNGNNGNNGATGGNNNGGATGNNTNANAGN NGGNGGAAGGGGAGAGNNGGGGAGGGGTATGNGNGNGAGGAGGNNTAGGGGA GGGGNGGAGNGGNNGAAATGGNGGTGGAAGGGGGAGGANAGGAAGGAANAGG GGAGGAANGTGGGN	Scotia 4
S05	0.5	GCATAATCTGCTCTGTCGACTGCAGATGGTAGGCCTGCTCTGAAGGATAGGATGC CTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATG GAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGT CTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATG TACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGT ATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAG GGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAAT GTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAG TAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATG TAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGT GATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCA AATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTGGTGNTTT TAATGTAAAANNGGAAAAANNNNNNNNNNGNNNTNNTNNNNNANNNNNNNAA NNNAANNNNNNANNNNNANNANNNANNNANNNNAAANANNAAATTTTANAANN AAAAANNNANNANNNNNANNNNNAANAAAAAANNNANNNNNNNNNANNNNNN ANNNNAAANNNAANAAAAANNNNANAAAAAANNANNNAAANNAAANNNAANN NNNNAANNNNNAANNNNNNNAAAANANANAAANNNANANNNNNNNANNNNN NNNNNANNNNNNNANAANNAAAANANNNNNNNNAANNANNNNNNNNNNNNN	Scotia 5

		NNNANNNANNNNNNNNANNNNNNANNNNANNNNNNNNNNANNNNNNNNNNN NNNNNNNNNNNNNNNANNNANNNNNNNNNNNANNNANNNNNNNNNNANNNN NNNNNANNNAANANNNNAAANNNNNNNNANNNAAANNNNNNNNANNNNNNNN NNNNNANNNNNNAANNNNNNNNNNNNNNANNNNNNNNANNNNNNNNNNAANN NANNNNNNNNANNNANNNNANAANNANNNNNNNNNNNNNNAAANNNAAANNN NNNANNNNNNNNNNANANNNNNNNANNNANANNANNANNNAAAANNNNNAAN NNAAANAANAANNNNAAANNANAANNANNAAAANAANNAAAAAAAANNAANN AANNNANNNNNNNNNANANNNNANNNAANNNAAAANNNNA	
S08	0.5	GACCTAATCTCATCGTTGAGTGTAGATGGTAGGCCTGCTCTGAAGGATAGGATGC CTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATG GAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGT CTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATG TACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGT ATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAG GGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAAT GTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAG TAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATG TAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGT GATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCA AATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTGG TTTGGTGTATAAANAGGAAAAANNNNATNNTNTNCNTTNTTTCTTTGATTTTTTCC TAANNAANNTNTTNATNNNTTNNNNTTTTTNNNTTNNNNNNNTNTNNNNTNTTNT TTNNNNTTNNNNTTNNTCCCNNTNNTTNNNNNNNNANNNNNNNNNNTTNNNNN NNNNNNNNNNNTANNNNNTNNAANANNNNTAANANAAATTNNNNNNNNNAANN TAANNNAATTNNNTNNANNNNNNTAANNNTTNNATTNTTTANNNANNNANTNNN NNNNNTNNNNAANNNNANNNNNNNNNTNNNNNNNNNNNNNNNNNNTAANNN NNAANANNNNNNNANNNNNANNNNNNNAANNNNNNNNTNNNNNNNANNNNNN NTNNNNNNNATTNNNNNNNNNNNANANNNNNNNANNNNNNNNNNNANNNANA NNNNNNANNAAATNNNNNNNANNANTANNNNNNNNAAANATNAANNNNTANNN NNNANANNNTNTNNTNNNAANNNANNNANNANNANAAANTAANANNTNNANNN ANNANNANANNNAANNNNTNNNNANNANNNTAANANANNNNNANNNNAATAA AAATANNAAAAAAAANNNANNNNNANANNNAANNNANNNATNNNTANNTTNAN NNNANNTAAANNNANNTNNNNTANNNNANNNNTNNAAANAANNNNNA	Scotia 7
S09	0.5	GGCATAGTCTCATACGTCGACTGTAGATGGTAGGCCTGCTCTGAAGGATAGGATG CCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAAT GGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATG TCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATAT GTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATG TATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTA GGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAA TGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTA GTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTAT GTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTG TGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGC AAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTGTT TTGGATGGTAANNNNNGAAAGGNNTATGTTTNCNNTTATTTTTTGATGTTTTNTTA GGAAAGTATTGATTTTTATAGTTTTGGGGTTATAGAGATATNTGATTGTTTTAGAA ATTAATTTGGTGGTGTTAATGATGAAAAAANAATTTTTGGGTTTTGATGTTAGGGT TGGGTATGNATTTAAGGGGGGTTGTTAATATGGGGGATGGTAGTAGGAATAATTT TTGAATTGGGTAGTGGGTTGTTAATTTAGAAAAGGAAGANANNGNANGGNNGNG GGGNAAGTGANATGTNGNGGGTNNNGNNNGNNGNNNAANGAANNNANGAGANN NNGNAAAGAAGANAGAGGGGGGAGAGGGGGGGGANGGGGAGGGGGGGGGGG GAAAGANGAAGGGGGANGGGGGANNNNTNGNGGGGGAGGGNGAAGNGNTNGA ANTGAGTANAGTNNNGTGNNGGGGGGGGGNGNNGAGGGNNGGNNNAAGGNNNN NGTGGAAGGATAGGGGNNNNTGNNGGNGAGGNGGTGAGANAGAATGNNNNAAG AAAANGGAANGATNGAANGATANNAGGAAANNGGGGGGATGGNTANNTTANNN AGAAGNGGAAAAANTANNGAGAANGAAGGTAAAGAGGGGTNNNANNANGNATG GGGGGTGGGGTNGATGGGAGGTGGNNGGANANGNGTNGTGAGAAGAGGGGGGG G	Scotia 8
U11	0.5	TACCTCCAAATCTCCTCAGTTCGACCGCAATTTTAGGGCCTGCTCTGAGGGATAGG ATGCCTAGATGGGCGGGATGGTGGTCCCTTATGAGTAGGGGATTGAGGTATTACT AATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAA ATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTA TATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTT ATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATG CTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATT TAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGT TTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGAT TATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGA GTGTGATACGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTA TGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGG GTGTTGATGGTAGAATCTAAAAAATATATTAACTTAATTTAAATATATAATATAAT TTCCTTTCAATTATATATAGACTTGCTTTACTATTTTTGTTTTATATGATTTTTAGA CATTATCTTCTACTATTAGACACAAACTCGTTAGAAATCTCCATCTCCTTAAATCA TAGTAGAATGTATTCTGAGTAAGAAACATATAATATATACAATATACTTATTCTAT	Uluru 11

		TAATTTCCATTCGACATTAATTTATTTAAACTTATCACATATTATTATGAATAACCT CGATATAAACACCAAACCCCGACCTGTGCACTTTACCATCTCCCTGAAAGTCTTA AAATGCTCCGTCGAGTTCATCTATCTATTGCTCTCATACTGTATATCCCACCTCTTT CACTATACCCAAATCAACCTACCCACATTACACTTCACGAACCAATGTACTTTTTT TAACTTCTACAACCAGGCTAATAAAATTAAACACCTCTTATCTAATATCTTCTTTA GCTGTCATTAATGCACTCTTCATCCACTAGTACCTACTGCTCATCTCAGGTACAGC CTCTCTTTTCCCATGATCAACATCACATCTCCATCACTCCT	
U18	0.5	CACCCCCGACTCGCTGGTACTCCCCCCCTTTATCGGATCTGCTCTTGGTGCAGGAT GCCTAGACGGGTGTGATGGTGCACCTCTCGTGAGTCTGCTAAAGAGGTATTACTA ATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAA TGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTAT ATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTA TGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGC TAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTT AATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTT TAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATT ATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAG TGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTAT GCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGT GTTTGATGGTACCCCCCCNCCCCCCCCCCCCCCCCCNCNCNCCCCNCCNCCCCCC CCCCCNCCCCCCCCCCCCNCCCCCCCCCCCCCCCCCNCCCNNCCCCCCCCCCNCCN NCCCCCCCCCCCCCCCCCCCCCCCCCCCCNCCCCCCCCCCCCCCNNCCCCCNCCCC CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCNCCCCCNCCCCCC CCCCCNNCCCNCCCCCCCCCCCCCCCCCCCCCCCCCCCCNNCCCCCCCCCCCCCC CNCNCCCCCCCCCCCCCCCNCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCNNCCCCC NCNCNCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCNCCCCNNCCCC CCCCCNNCCCCCCCCCCCCCCCCCCCCCCCC	Uluru 18
A20	0.5	TTCCTACCCAACTTTCTTTCTCCTACCTCATTTTATCGGCCCTGCTCTGAAGGATAG GATGCCTAGATGGGCGGGATGGTGCTCCCTTTTGAGTAGGGGATTGAGGTATTAC TAATGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGA AATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTT ATATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTT TATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATAT GCTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGAT TTAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTG TTTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGA TTATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTG AGTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCT ATGCAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTG GGTGTTGATGGTAGTTTTTTTTGGGGGGGTTTGGTGTCCTCTCGCCGCACTCGCTC CTTCAATTTCCTTTACCCACTCCTCACACTA	Watarraka 20
A28	0.5	CCCCCCAACCTCCCCTCCTCCCACCGCCCTTTATAGCGCCCGCTCTGAAGGATAGG ATGCCTAGATGGGCGGGATGGTGCGCCCCTCTTGAGTAGGGGATTGAGGTATTAC TAATGGAAGGGATATGCTAGAACGGTCCGCTTTAAGATTCATAAATCATTGATGA AATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTT ATATGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTT TATGTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATAT GCTAGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGAT TTAATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTG TTTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGA TTATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTG AGTGTGATAGGGTAGCTTAATACTGACGTAGGTAAAGAATTTTGATAAAAATCGA GGGCTATGCCAAATGGATTTGTTGTGCTGTCCGGAAGTAGTTTAAGTAGAATATC AGCCTTTGGGGTGGTTGATGGAAAAAGGGGGGGGGGGGAGAAAAGAAACCCCCA AGCATACACTATGGTCCACATCCGTACCAAATACGAGC	Watarraka 28
A05	0.5	CCAAAATTTCATGGGTCGACCGCATTTATGGGCCTGCTCTGAAGGATAGGATGCC TAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGG AAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTC TATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGT ACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTATGTA TTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCTAGG GGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTAATG TGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTTAGT AATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTATGT AATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGTGTG ATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATGCAA ATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGGTGTTG ATGGTATTTTCTTTCTCCTTCTTCCCCTCTCCCNCCCTCTCTCTCCTTCCCTCTCTTT TTTTCTCCTTCTCTTTTTCCCCTTTTTCTTTCTTCTCTTTTCTTTTTCTCTTTTCTTTTT TCTTTTTCTCTTTCTCCTTCTCTCTCCCTTTCTCTTTTCTTCCTCCCTTTTTCTTTCCT CTCTCCTCCCTCTCCTCCTTCCTCCTCCTCTCTCTCCCCTCCCTCCCTCTCCCCCTCT TCCTCTTCCCTCTCTCTTTTCCTCCCTTCCCTCCCTCCCCCCTCCCCTCCTCTCTTCT CTTTCTTTCCCCTTTTCCTTTTCTTCCTCCTTCTTCTCCTCTCCCCTTCTCTCTTCCCT CCCCCCTTTTCCTCTCCTCTTTCCCCTTTTCCTCNCTTCTTCTCTTCCCTCCTCTTCTT	Watarraka 5

		TTTCTCTTCCCTTCCTTTTTTTTTTTCTCTTTTCTCTTCCTCTTTCTTTTTCCTCCTTCTT CCTTCCCTTCCTTTTCTCTCTCCCCCTCTTCCTTCTCCTTTTCTTCTCTTCTCCTTCCT TCCTCCCTCTTCTCTTCTCCCCTCTCTTTTTCCTCTCTTCTCTCCCTTCTTCCTTTCTC TCTTCTCTCTCTCCCCCTTCTTTCCTCTCTTCTCTTCTCTCCCTCTCTCTTCTTCGTTC TCTTCTCTTTTTCCCCTTCTTTCTT	
A31	0	CCAAATTTTGCCTTTGGATGGCACGCTATATTTTGCGACCGAGATTTGTAGGACTG CTCTCCAGGATAGGATGCCAAGATGGTCTCTTTGGTGGTTTCTCTTCGGTAGGGGA TTGGGGTAGTACTAACGGAAGGGATATGCTATGAACGGCAAGCTTTAATAATCAT AAGTCATTGATGAAGTGTCATTTTTGTTGTAGGATTAA	Desert Park 1
A35	0	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCTCATCTCCCCCGCGTGGGCCGGA AGCCCCCCCGGCCTCGGTCCCGCCCCCCTCTTCACACTCCCCAAAGGAGTACTACT AATGGAAAGGATCTGCTACGAACGGCTAGTCTTACACATTCCTCAATCATTGATG AACTGTCTATGTCATTGGAGGATTGTATGTATTA	Desert Park 5
P12	0	CATACACCTAATTCTCACCTCCACATCCTTATAGTACGCCCTCCACGGGAAGGATC CCCACAAGGGGCGGCATGGGGGTCTCCCAGGGTGGGCCCTTAAGTGTCACTACAA AGGGGTATATACTAAAGGAAAGCACTTGATAGACTAGCACATTGTAAAATGGCTA CATCATTGAGGAAATGTCTATTATAATAGGAGGAT	Perron 12
P13	0	CCCAAAATTTTTGTGGTCAAGCCAACCTCTTTGTGGCACGCTCCCAAAGAAGGAT GCCCCCAAGGGCGGGATGGTGGTTCCTTCTGAGTGGGGGATTGAGGTGTTTCCTA TGGAAGGGATATGCTACAAAGGTAAGCTCTTTGAATCATAAATCATTGATGAAAT GTCTATGTATTTGGAGGATATAATGTATTATGTAA	Perron 13
P27	0	GCCCCAAATTCCAGGTTCGACCAACAATATATAGTCCTGCTCCCCAGGAAGGATG CCTCCATGGGCGGCATGGGGGTTTCTCCCCAGTAGGGGAGTGAGGTATTACTAAT GGAAGGGAGAGGCTAGAACGGTAAGCTTTAATAATCATAAATCATTGATGAGAT GACTATGTAATTGGAGGATAAATGTATTAGGTAAG	Perron 21
P05	0	TGCTCACCTTCCCCGCCCATTTTTATGGTGACCCCCCCCTCAGAGAAGAGAGAGG CCCTCCTAGGGGGCGGGAGGGGGTTTCCCTCCCGGGTAGGGGGGTTGTGGTGTCC CCTCCTAGCGGGGGAGATGCCGCTAGGGGTGAGCTCTATGATACACAAATCTCCG AGATAAAGTCTCTGTGTTTGGGGGAGATATAGTGTTGTGAGAAGTGATTGTGGT	Perron 5
P06	0	CAAATAACTCTCTACTTTCCGTAGCAAGATTTTAAGGACCCGCTCCTCAAGGAAA GGAGGCCACCATGGGCGGGATGGGGGTTTCTCCTGCGTAGGGGAGTGAGGTGTTA CTAAAGGAAGGGAGGGGCTAGAACGGAAAGCTTTAATATTCATAAATCATTGAT GAAATGACTGTGTAATTGGAGGATAGAATGTATTGTGTAA	Perron 6
P07	0	CGCCTAAATTGCTCGTAGACTGCATTTATGGGCCTGCTCTGAAGGATAGGATGCC TAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAATGG AAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAATGTC TATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATATGT ACTTTGTGAGTTTTTATGTAATTGGAAATATGCTAGTGTTGAACCGGGTTTATATA TTTTTATTTTTATGTTTTTTTAAGTGTATGTGATTTAGCTTCCTTATTACTAGGGGGG GAGGGTGATAATTAAATAATATAAAAATAATTTTATTAAAAATAATTTTATTGTG GGAGGTTAATTCATTTTTTTGTTTTTATTTTAATAAAAATTTAAAGTGGATAGGGA ATATAAAAATGGAAATCTTAGGCTACAATGACATTGAGTTTTACTCAGTGAATGG AATATTTTAAAATATAATGTACAAGGTAAAACTCTCAATTATAACAGCAGAGTTT GATAGTGAGTGAAAAACCTCATTCCGGCCCATTTAGGAATCCTATCCTTCAAAGC AGGGCCATGAATTGGGGTCGTACCGCTGAAAACGCAGACTGGCTACTGGTTGTAC TTTCGGGACGCCGGGAGGCAAAAAAGGGGAAAAGAAGAGCAGCGGGGGGGGGG GGGGGGGGGGGGGGGGCGGGGGGGGGGGGAAGAG	Perron 7
U01	0	AACGGCCCGGCCCGGTGAACGCGCGCCCCTTTAGCGTATCATGCCTTGTGTGAGC GGTAAGCCACCCCCCGCGCGCGGCTCCTTTTACACCGCCTCCTCCCACCACTAGA AATATTGCTAAGGGAGGGGATATGCTAGAAGGGTAACCTTTAAGATTCATAAATC ATTGAAGAAATGTCTATGTAAGTGGGGGATTTAATGTGTTATAT	Uluru 1
U14	0	CCAAAATTAGCATGGCACAGACACATTAAAAGGCACGCCCCTAAAGAAGAGGAG GCCACGAAGGGGGGGGAGGGTGGTCCCTCTTGAGTGGGGGGGTGAGGTGTTACT ATGTGGGAGGGGAGATGCTATAACGGAGAGCTCTTATATACATAAATCATTGTTG AGATGTCTCTGTAATTGGAGGAGAATGTGTATTATGTAAA	Uluru 14
U22	0	AAACACATCAACGTCACACAAAACTTTATAGAATACCGCCCCCCACAGAAAAGA AGCCCCCCAAGGGGGGCCTTTGGGTTCCTCCCCGGTGGGGGATTGGGGTGTCCTA CCGGAGGGGGTGTCTTAAACGGGGAGCTTTAAATATAAAATAAAATTTTTGAAAA GTCTTTGTATTTGTAGGAGTAAATTGTATTGTGTAGAGTTGAGA	Uluru 22
U24	0	GACCTGCATTGCTATGGTCATGGTACATTTATGGGCCTGCTCTGAAGGATAGGAT GCCTAGATGGGCGGGATGGTGGTTTCTCGTGAGTAGGGGATTGAGGTATTACTAA TGGAAGGGATATGCTAGAACGGTAAGCTTTAAGATTCATAAATCATTGATGAAAT GTCTATGTAATTGGAGGATTTAATGTATTATGTAAAGTTTGAGTTGTATGTTTATA TGTACTTTGTGAGTTTTTATGTAATTGGAGATATGCTAGTGTTGAAGCTGGTTTAT GTATTCTTATAGTTATGTTTTTAGTAATGTTTAGGTTAATGTACTTGCTTATATGCT AGGGGTAAAGGTTATTAATGTACTATATACATATTAATGTTATAAACATTGATTTA ATGTGGGTAGTTCATTGATTTTTATGTTTTTAGTATTATTAAGAGTTTAATGTGTTT AGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAGGATTA TGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGTTGAGT GTGATAGGTAGCTTAATACTGACGTAGGTAAGAATTTGATAAAATCGAGGCTATG CAAATGGATTGTTGTGCTGTCAGGAAGTAGTTTAAGTAGAATATCAGCTTTGGTG GTTGGATGGTAAANNNNNNNANNNTATNGNTTTGNNTTNTTTTTGATGTTTNNTT NNGANTTTTTTNNTNTTTNNNTTTTTNNNTTTTTNNNNNNTNNTTTTTNTTTNNTNN TNNNTTNNTTNTNTNNTTNNNNNNNTNNNNNNNNTTNTNNNTNTTTNNNTTTNNT	Uluru 24

		TTTNTTTTNTNNNNNNNNTTTNNTTTGGNNNNTNTNNNNTNNNNNTNTNTNNNTT NNTTTNGGTTTTGTNNTTTNGGNNTNNNTTNNNNNTNNTNTTNNNNNNNNNTTNN NNNNNNNNTTNNNNNNNNTNNNNNNTNTNTNTNNNNTTTNNTTTTTTNNNNNNN NNNNNNNNNNNNNNTNNNGGTGGNNTGNNNNNNNGGGGGGGGGGGGNNTGGNG GGGNNTNNGNNGNNGGGGGNNTGGNNGTGNNNGGNNNGGGNNGTGGGGNGGNG GGGGTTGTGGTGTNGGGGNGGNNNGGTNNNGGGGNGNTTGGNNTGGTNGTGTGG TGGGGNGGGGTGTNNTNNGGGGNGTGTTGNNGTGTGTTTTGNNNGTTGGTTGTGG GGGNGNGTGGNGTNGNNGNNGNGNGNTGGGNTNNNGTNGTGGGGGTNGGGGTG NGTGTGGGTGGGGGNGGGGGNGGGGTNGTNTTGGTGGGGGGTGGTGTTGGTGNN GNGGGGGTGTNGGGGGGTNGNTGGGGNNNTTNTGNGGTTGGGGG	
U26	0	CCATCCAAAACCAACTCCTATATGAGACCCCCCCCCCCAAGAGAAAGAAAGGCCCC AAGGGCGGGCGGGGGGGGTTCCCTTAAGGGGCCGGTGACCGGGCCTCCCCACGA GGGGGGAGGGGCCCTTAAGAGGGGCTCTTTTATACTAACTCACTATGGAGAAACC GCCGTGTTTGGGAGGGGAATTTAGTTTGTAAAGAGTTTGTGTAATT	Uluru 26
U27	0	CCTCTCTTCCCCGCTGGTAATATAAGACGCCCGCCCTCGGAGGATAGGCGCCCAG GGGGGCGCGGGTGGGGGTCCCCTAGGGGGGGGGGATTGGGGTTTCTACAGGGGA AGGGATTCGCCTAAAGGGAGCGCTCTTAAGATTCACAAATCCTTGAAGAAGTGTC TCTGTTATTGGGGGATTTGATGTTTTATGGAAAGGTTGGGTTG	Uluru 27
U28	0	CCTCAGCGGTCGCAGCCTCCTTTAAGGAACGCCCCCCTACGGAAAGGAAGCCCAC CCGGGGGGGGAGGGGGGGGCCCCCCCCTTTGTGTGGTGTGGTTTCCTTGGTAGGG GGATGGCTACCGCTTAAGCTTCTGAATACATAGCACAGAGGTGAGGGTCTGTACT ATAGGGGGAGAAGGAGCTTATATGGAGTGGGTGTAGT	Uluru 28
U30	0	CAAAAAAGCGAAATTGACGGGCGAGGTATATAAGGAGGCTCTCAAGGAAGGAGA GGCTACCTAGAGCGGGAGGGAGGGTTTTCCTTTCGGGAGGGGGGTAGGGAGTGA CAACTGAGGGGGGGGTAGCTGTAAGGGGGGGCTTTATAAAAATTAAAAAATTGT TGAGAAGTCTCTGTTTTTGTGGGAGAATTTGTATTATGAAG	Uluru 30
U04	0	CACCCCTCCCCGCCTCTTTTCGGTCCCCGCCCCCAAGGGGGAGAGGCGCCCCCAA GGGGGCGGCATGAGGGGTTTCCCCCGGGGAGGGGGGTGGAGGTGTTTCTACTGG GGGGAGAGGCTCGAGTAGGGGTGCGCTCCTATAAACATCAAACCCCTGTGGAGG CGTCTCTCGTAGTAGGGGGAAGAAATTGTTTTATAAGATGTGTGTGTAGTG	Uluru 4
A01	0	CAAAATTTTTCGGTGTTGAGACGAGAGATGATAGGCAGGCTATGAGGGAGAGGA TGCCTAGAAGGGCGGCATGGTGGTTTCTCGTGAGTGGGGGATTGAGGTAGTGCTA ATGGGAGGGATAAGCTAGAGGCGGTAAGGCTTTAAGAGTCATGCATCATTGATG AAATGTCTATGTAATTGGAGGATTTAATGTATTATGTAAAATTTGAGTTGTATGTT TATATGTACTTTGTGAGTTTTTATGTTATTGGAGATATGCTAGTGTTGAAGCTGGT TTATGTATTCTTATAGTTATGTTTTTACTAATGTTTAGGTTAATGTACTTGCTTATA TGCTAGGGGTAAAGGTTATTAATGTACTATAGACATATTAATGTTATAAACATTG ATTTAATGTGGGTAGTTCATTGATTTTTGTGTTTTTAGTATTATTAATAGTTTAATG TGTTTAGTAATATGTGAGTTGTTAGTCTTTGTTATCATTGATTTTATGTATTACTAC GATTATGTAATATTATACATATGATGTAAAACGTCAAATTTTAAGCAGGATGGGT TGAGTGTGATAGGTAGCTTAATACTGACGTAGGTAAGAATATGATAAAATCGAGG CTATGCAAATGGATTGTTGTGCTGTCATGAAGTAGTTTAAGTAGGATATCAGCTTT GGGTGTTGATGGTAGGACAGGAAGGTGTAAAATAGTGTAGTAGCATTGTGGGGG AATATCAGGCAAAGAGCACAGCAGGACGCAGACGTTGTGTT	Watarraka 1
A13	0	CGGCCTTCTACCCTCTAGCTTAAAGTACGCACGCCCTCAAGGATAAGGAGGACCC GATGGGCGGGAGGGGGGTCTCTCTTGCGTAGGGGATTGAGGTATTACTAAAGGA AGGGAAATGCTATAAGGGAAAGCTTTAAGATTCCTAAAACATTGAGGAGATGTCT ATGTAATTGGAGGATTTAATGTTTTATGTTAAGTTTGAGTTGT	Watarraka 13
A21	0	CCCGCGGTCGCCTCCCAATCTCTGTAGAAGCACGCCCTCCCCCTAAACAAACGGC AACGTTCGTCGGAGGGACCCCCTTTAGGGGCGTGAGTATGATTGATCTTAAGGAG AGGCGGAGCCATGGAATAGCCGCCTTATCTTCATAGCTATATGCAGAGGCTCCCG CTTTATCAGTGTATTATTTTTGGTGTAGTTAGTGGTGTA	Watarraka 21
A22	0	CAATTAGTTAAGTTCTACTAAAACATTTATATTAGGCCCCCACAGAAAAAAGGCT CCAATGGTGGGTGGGGTGTTTTCCCTTGTGTGGGGGGGTAGGTGTGTTATTAAGT GAGGGGGATGCTATATCAGGGAGCTCTTCTATATTCATATCACATCGTAGAAGAG TATGTGTTATTGGAGTAGAATGTGTATTATTAAGAGTATGT	Watarraka 22
A08	0	CCCCGGAATATGAATGGTCAACAGCAAGGTATGGGCAGGCTCTCAAAGAAAGGG AGGCTCGATGGGCGGGATGGTGGTTTCCCTTGAATAGGGGATTAAGGTAGTACTA AGTGGAGGGGATATGCTAGAACGGAGAGCTTTAAGATTCATAAAACATTGATGA AAAGTCTATGATATTGTGGGAGATAAATGTATTATGAA	Watarraka 8

Appendix 2.4

The results of UPGMA clustering

Analysis of D:\TFPGA\MALATG~1.DAT

Data set contains genotypes of individuals sampled from populations.

Organism Type: Diploid

Marker Type: Codominant

0.0819 2 4

0.1311 1 2 4

0.1821 1 2 3 4

0.4087 1 2 3 4

UPGMA Cluster using Nei's (1972) original distance

Node	Distance	Includes Populations
1	0.0819	2 4
2	0.1311	1 2 4
3	0.1821	1 2 3 4
4	0.4087	1 2 3 4 5

Results from bootstrapping

1000 permutations conducted

Node	Proportion of similar replicates
1	0.5040
2	0.4920
3	0.6890
4	1.0000

0 bootstrap replicates produced trees containing ties.

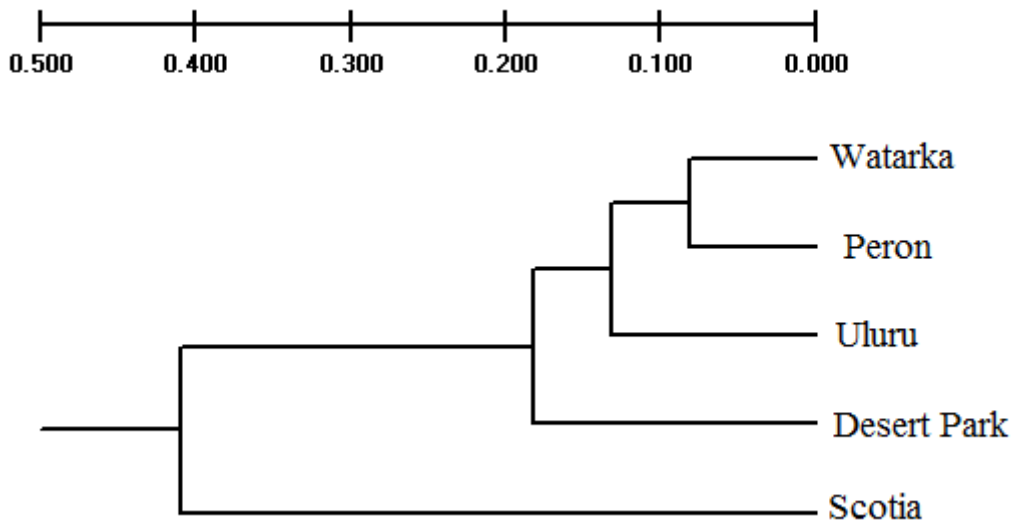


Figure A2.1: The results of UPGMA analysis clustering

Appendix 3 Additional Information and Statistics detail for Chapter 3

Appendix 3.1 Summary Statistics

Table A3.1: Summary counts of species types used in the meta-analysis

	Fish	Amphibian	Reptile	Bird	Mammal
Number genotyped	4	0	4	14	48
Genotyped using sequence based approaches	2	0	0	7	18
Genotyped using SSCP	2	0	2	5	22
Genotyped using other methods	0	0	2	2	8
Confirmed solitary	2	0	4	0	6
Social and living in group size 100 or greater	0	0	0	7	3

Appendix 3.2 Meta analysis dataset

Table A3.2: The data used in the metaanalysis. To fit the table into the page column headings have been given by numbers where 1- Species, 2- Taxa, 3- IUCN status, 4-Solitary or group living, 5- Maximum population size, 6- Mating System, 7- Microsatellite expected heterozygosity averaged across loci and populations, 8- Microsatellite observed heterozygosity averaged across loci and populations, 9- Microsatellite alleles per loci averaged across populations, 10 – Number of individuals used in for microsatellite genotyping he study, 11- MHC genotying method, 12- Class of MHC genotyped, 13- Number of individuals used for MHC genotyping, 14- Number of MHC alleles, 15- Number of MHC loci, 16- MHC heterozygosity.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Chatham Island black robin Petroica traversi	bird	endangered	group		monogamous					SSCP	II	10	4	2	
South Island Robin	bird	critically endangered	group		monogamous	0.48		4.2		SSCP	II	16	41	7	
Crested ibis Nipponia nippon	bird	endangered			monogamous	0.506	0.414	2.56		SSCP	II	36	5		
Galapagos penguin Spheniscus mendiculus	bird	endangered	group	100+	monogamous	0.038	0.0304	2		Clone and sequence	II	30	3		

Humboldt penguin Spheniscus humboldti	bird	unclassified	group	100+	monogamous	0.758	0.704	7.857		Clone and sequence	II	20	8		
Adelie penguin Pygoscelis adeliae	bird	near threatened	group	100+	monogamous					Clone and sequence	II	4	4		
Chinstrap penguin Pygoscelis antarctica	bird	unclassified	group	100+	monogamous					Clone and sequence	II	2	3		
Gentoo penguin Pygoscelis papua	bird	near threatened	group	100+	monogamous	0.111	0.118	1.33		Clone and sequence	II	6	10		
Little penguin Eudyptula minor	bird	least concern	group	100+	monogamous	0.613	0.545	7.14		Clone and sequence	II	4	4		
Gouldian finch Erythrura gouldiae	bird	near threatened	group	100+	monogamous					SSCP	II	14	78		
Long-tailed finch Poephila acauticada	bird	least concern	group					16.4		SSCP	II	14	94		
Great reed warbler Acrocephalus arundinaceus	bird	least concern	group		polygynous		0.578	10.1		DGGE	I	354	67		

Seychelles warbler Acrocephalus sechellensis	bird	vulnerable	group	5	promiscuous	0.48	0.51			DGGE	I	486	10		
Lesser kestrel Falco naumanni	bird	least concern	group	50	monogamous		0.66	11.4		Clone and sequence	II	21	26		
Chinook salmon (Oncorhynchus tshawytscha)	fish	threatened (regional Canada)	group		Promiscuous	0.8054	0.7738	39.08		SSCP	I	342	37		0.8852
Brown trout (salmo trutta)	fish	least concern	solitary	1	promiscuous		0.649	11.7		Clone and sequence	I	485	13		0.762
Gila Trout (Oncorhynchus gilae gilae)	fish	endangered	solitary	1	promiscuous	0.6265	0.4835	4.35		Clone and sequence	II	143	5		0.59
Sonoran topminnow Poeciliopsis occidentalis	fish	endangered	group			0.2367	0.2443	8		SSCP	II	338	17		0.3018
African elephant Loxodonta africana	mammal	vulnerable	group	10	polygynanous	0.81	0.81	8.9		Clone and sequence	II	30	6		0.58
Asian elephant Elephas maximus	mammal	endangered	group	10	polygyanous	0.57	0.52	4.11		Clone and sequence	II	3	4		

African green monkey Chlorocebus sabaeus	mammal	least concern	group	16	polygynaous (but only one offspring per year)		0.66	4.88		Clone and sequence					
African wild dog Lycaon pictus	mammal	endangered	group	20	monogamous	0.68	0.69			Clone and sequence	II	368	17		
American bison Bison bison	mammal	near threatened	group	231	promiscuous	0.616	0.611			SSCP	II	20	9		
Arabian oryx Oryx leucoryx	mammal	vulnerable	group	30	polygamous	0.565	0.601	3		SSCP	II	57	3		
Australian bush rat Rattus fuscipes	mammal	least concern			promiscuous	0.78	0.72	8.69		TGGE	II	?	36		
Aye-aye Daubentonia madagascariensis	mammal	near threatened	solitary	1	polygynous					Clone and sequence	II	7	16		
Baiji the Chinese river dolphin Lipotes vexillifer	mammal	critically endangered	group	4	probably promiscuous					Clone and sequence	II	18	43		
Bengal tiger Panthera tigris tigris	mammal	endangered	group	4	polygynous	0.74	0.449	6.4		Clone and sequence	I	14	14		

Brown bear Ursus arctos	mammal	least concern	group	5	promiscuous	0.71	0.66	6.8		Clone and sequence	II	38	19		
Domestic cat felis catus	mammal	domestic			promiscuous	0.8		14.3		RSCA	II	71	23		
California sea lion Zalophus californianus	mammal	least concern			polygynous	0.5892	0.602	6.375		DGGE	II	19	1		
California sea otter Enhydra lutris nereis	mammal	endangered	group	100	polygynous	0.426	0.433	4.14		SSCP	II	70	5		
Cheetah Acinonyx jubatus	mammal	vulnerable	group	5	promiscuous		0.471			RSCA	II	25	5		
West african Chimpanzee Pan troglodytes versus	mammal	endangered	group		promiscuous					RSCA	I	30	4		
Common hamster Cricetus cricetus	mammal	least concern	solitary	1		0.707	0.716	8.36		Clone and sequence	II	15	13		
Coyote Canis latrans	mammal	least concern	group	6	monogamous	0.792	0.67	9.27		SSCP	II	38	18		

Desert bighorn sheep Ovis canadensis	mammal	least concern	group	10	polygynous	0.861	0.566	6.1		SSCP	II	206	21		0.742
Domestic dog Canis lupus familiaris	mammal	least concern					0.618			clone and sequence	II	50	22		
Ethiopian wolf Canis simensis	mammal	endangered		13	monogamous (probably)	0.278	0.242	2.4		Clone and sequence	II	99	11 (over 3 loci)	3	
Eurasian beaver Castor fiber	mammal	least concern			monogamous	0.545	0.519	3.4		RFLP	II	25	1		
European Moose	mammal	least concern								SSCP	II	19	4		
North American Moose	mammal	least concern			polygynous		0.324	5.6		SSCP	II	30	7		
European bison Bison bonasus	mammal	vulnerable	group	13	polygynous		0.37	2.46	22	SSCP	II	110	4		
European mink Mustela lutreola	mammal	critically endangered	solitary	1	polygynous					Clone and sequence	II	20	9		

European wolf Canis lupus lupus	mammal	regionally threatened	group							SSCP	II	167	36	3	
Giant panda Ailuropoda melanoleuca	mammal	endangered	solitary	1	polygynous	0.62	0.57	5.5	20	SSCP	II	60	7		
Gray mouse-lemur Microcebus murinus	mammal	least concern			see malagasy mouse lemur					SSCP	II	145	12		0.78
Hawaiian monk seal Monachus schauinslandi	mammal	critically endangered	group		promiscuous (probably)	0.48	0.49	3.5		Clone and sequence	I	6	6		
Iberian red deer Cervus elaphus hispanicus	mammal	unclassified	group	40	polygynous	0.728	0.731	8.67		SSCP	II	94	18		
Lion-tailed macaque Macaca silenus	mammal	endangered	group		polygynous					DGGE	II	10	9		
Malagasy mouse lemur Microcebus murinus	mammal	least concern	group	20	promiscuous					SSCP	II	228	14		
Mountain goat Oreamnos americanus	mammal	least concern	group	50	polygynous	0.5	0.51	3.44	215	Clone and sequence	II	25	2		

Bastard big-footed mouse Macrotarsomys bastardi	mammal	least concern	pair	2	Monogamous					SSCP	II	22	9		
Western Tuft-tailed rat Eliurus myoxinus	mammal	least concern	group		Promiscuous					SSCP	II	75	20		
Malagasy giant rat Hypogeomys antimena	mammal	endangered	pair	2	Monogamous					SSCP	II	139	2		0.49
Mexican wolf Canis lupus baileyi	mammal	unclassified (but extinct in wild?)	group	8	monogamous	0.548	0.507	5.3		SSCP	II	36	5		0.49
North American gray wolf Canis lupus	mammal	least concern	group	7	monogamous?	0.647	0.538	5.03		Clone and sequence	II	175	44		
Northern elephant seal Mirounga angustirostris	mammal	least concern	group	100+	polygynous			1.46		SSCP	II	110	5		
Przewalski's horse Equus ferus	mammal	endangered	group		polygynous	0.474				SSCP	II	14	6		
Red Wolf Canis rufus	mammal	critically endangered	group	8	monogamous?	0.548		5.3		SSCP	II	48	4		

Rhesus macaque Macaca mulatta	mammal	least concern	group	80	polygynous	0.71		9.2		USAT	II	5891	21		0.81
Striped mouse Rhabdomys pumilio	mammal	least concern	group		polygynous					SSCP	II	58	20		
Black-footed rock-wallaby Petrogale lateralis lateralis	mammal	near threatened	group		polygamous	0.4378		2.88		SSCP	II	79	15		
Tammar wallaby Macropus eugenii	mammal	least concern	group		polygynous	0.82	0.71	10		Clone and sequence					
Tasmanian devil Sarcophilus harrisii	mammal	endangered	solitary	1	promiscuous	0.44	0.42	3.63		Clone and sequence					
Western barred bandicoot Perameles bougainville	mammal	endangered	Solitary	1	promiscuous					Clone and sequence	II	142	2		
Hungarian meadow viper Vipera ursinii rakosiensis	reptile	endangered	solitary	1	polygynous					RFLP	I		1		
Sand lizard Lacerta agilis	reptile		solitary	1	promiscuous					RFLP			Band sharing used here		

North Brother Island Tuatara Sphenodon guntheri	reptile	vulnerable	solitary	1	monogamous	0.06		1.2		OSCP (modified SSCP)	I	27	3		0.527
Tuatara Sphenodon punctatus	reptile	least concern	solitary	1	monogamous	0.78		7.7		OSCP (modified SSCP)	I	26	15		

Appendix 3.3 Methods for Chapter 3.1 Can neutral genetic diversity or IUCN status predict genetic diversity at MHC.

A literature of Google Scholar (www.scholar.google.com) and Web of Science (www.webofknowledge.com) was carried using the keywords MHC population genetics, Major histocompatibility complex populations, MHC genetic diversity and MHC variation. From the results of these searches a literature review was conducted to determine the following where possible; MHC genotyping method, Class of MHC genotyped, Number of individuals used for MHC genotyping, Number of MHC alleles discovered, Number of MHC loci found and MHC heterozygosity.

For species where MHC genotyping had taken place an additional literature review was conducted to determine the following; vertebrate taxa, IUCN status, Solitary or group living, Maximum population size recorded in the literature, Mating System, Microsatellite expected heterozygosity averaged across loci and populations, Microsatellite observed heterozygosity averaged across loci and populations, Microsatellite alleles per loci averaged across populations, Number of individuals used in for microsatellite genotyping study. This resulted in the dataset presented as Appendix 3.2.

For Chapter 3.1 IUCN status was determined by checking the IUCN redlist (www.iucnredlist.org). Then only species which had data on both microsatellites and MHC were used for further analysis. This resulted in 37 species with information on both MHC and expected heterozygosity (He) and 38 species with information on both MHC and observed heterozygosity (Ho) that could be used in statistical analysis.

Appendix 3.4 Methods for Chapter 3.2 The Effect of mating system on the number of MHC alleles

A literature of Google Scholar (www.scholar.google.com) and Web of Science (www.webofknowledge.com) was carried using the keywords MHC population genetics, Major histocompatibility complex populations, MHC genetic diversity and MHC variation. From the results of these searches a literature review was conducted to determine the following where possible; MHC genotyping method, Class of MHC genotyped, Number of individuals used for MHC genotyping, Number of MHC alleles discovered, Number of MHC loci found and MHC heterozygosity.

For species where MHC genotyping had taken place an additional literature review was conducted to determine the following; vertebrate taxa, IUCN status, Solitary or group living, Maximum population size recorded in the literature, Mating System, Microsatellite expected heterozygosity averaged across loci and populations, Microsatellite observed heterozygosity averaged across loci and populations, Microsatellite alleles per loci averaged across populations, Number of individuals used in for microsatellite genotyping study. This resulted in the dataset presented as Appendix 3.2.

For Chapter 3.2 monogamy was defined as species that typically produce offspring with a single mate for a season or more (Johnson & Ryder 1987). This was used as true genetic monogamy is difficult to ascertain and simple social monogamy may include large proportions of extra pair mating. Species that were not monogamous were classified as multiply mating for the purposes of this investigation. Whilst it is true that polygamous and polyandrous species have different characteristics to each other and are different again from species with mating systems like scramble competition they were simplified for the purposes of this analysis. This is because any multiple mating would act as a hedging strategy for reproduction bets may decrease the pressures on mate choice mechanisms. For this analysis a total of 18 species with MHC data were classified as monogamous by the literature and 28 species were classified as multiply mating.

Appendix 3.5 Methods for Chapter 3.3 The Effect of group living on genetic diversity at MHC

A literature of Google Scholar (www.scholar.google.com) and Web of Science (www.webofknowledge.com) was carried using the keywords MHC population genetics, Major

histocompatibility complex populations, MHC genetic diversity and MHC variation. From the results of these searches a literature review was conducted to determine the following where possible; MHC genotyping method, Class of MHC genotyped, Number of individuals used for MHC genotyping, Number of MHC alleles discovered, Number of MHC loci found and MHC heterozygosity.

For species where MHC genotyping had taken place an additional literature review was conducted to determine the following; vertebrate taxa, IUCN status, Solitary or group living, Maximum population size recorded in the literature, Mating System, Microsatellite expected heterozygosity averaged across loci and populations, Microsatellite observed heterozygosity averaged across loci and populations, Microsatellite alleles per loci averaged across populations, Number of individuals used in for microsatellite genotyping study. This resulted in the dataset presented as Appendix 3.2.

For Chapter 3.3 species were classified as either solitary, where the literature recorded them as either living alone or living in pairs, or group living where the literature classified them living in groups of three or more. Furthermore, for species classified as group living the literature was searched for maximum group sizes occurring in natural environments (not counting zoo or wildlife park populations). For ease of analysis species where maximum group size exceeded 100 individuals were all classified as 100+ individuals as analysed as a single group. This methodology was limiting in that for many species group size fluctuates greatly over the course of a year or a lifetime and thus any risks associated with group living fluctuates with it. We chose to use maximum group size as it was an indicator of greatest disease risk and thus more likely to be associated with differences in MHC variability. A total of 16 species were classified as solitary and 47 were classified as group living for this analysis.

Appendix 4 Additional detail on methods for Chapter 4

Appendix 4.1

tRFLP protocol

The protocol used for tRFLP PCR reaction in Waldron *et al.* (2009) is reproduced below.

"The primary and secondary reaction mixtures contained 6 mM $MgCl_2$, 200 μM deoxynucleoside triphosphates, 200 nM of each primer, 1 U of Red Hot Taq*(ABgene, Surrey, United Kingdom), and 2 μl of DNA template. After an initial denaturation at 94°C for 3 min, a total of 35 cycles, each consisting of 94°C for 45 s, 56°C for 45 s, and 72°C for 1 min, were performed, followed by a final extension step of 72°C for 7 min. The secondary PCR mixture (total volume, 50 μl) contained 1 μl of the primary PCR product. The cycling conditions for the secondary reactions were identical to those used for the primary PCR. All PCRs were performed with a negative control containing only PCR water and a positive control containing* C. parvum *DNA. Reactions were performed with Eppendorf Mastercycler Personal instruments (Eppendorf, North Ryde, Australia), and products were resolved by electrophoresis on 1.5% (wt/vol) agarose gels using Tris-borate-EDTA. Secondary product fragments that were the correct size (832 to 835 bp depending on the species) were purified using a QIAquick PCR purification kit (Qiagen, Melbourne, Australia) by following manufacturer's instructions for the spin protocol."*

Our protocol was modified from this in that nested PCR was unnecessary as a large amount of product was produced from the first PCR which was confirmed by visualisation on 1.5% agarose gels as per the protocol above. In addition nested PCR was not suitable the method relies on a second set of primer sequences being available within the target amplification zone (Roux 1995), this was not possible for MHC as the method was intended to amplify a large number of diverse sequences. The primers and annealing, denaturation and extension temperatures used in this PCR were previously described in Chapter 4.

The protocol used for tRFLP analysis in Waldron *et al.* (2009) is reproduced below.

"…The constituents used for primary and secondary reactions were identical to those described above except that Red Hot Taq *was replaced by* Accutaq *(1 U) and dimethyl sulfoxide was added at a final concentration of 2%. The cycling conditions were initial denaturation at 94°C for 3 min, followed by 35 cycles, each consisting of 94°C for 45 s, 56°C for 45 s, and 68°C for 1 min, and then a final extension at 68°C for 7 min. For secondary PCRs that generated the expected products, the products were purified and subjected to restriction digestion with VspI. The amount of DNA in the restriction digest used for T-RFLP analysis was 10-fold less than the amount used for RFLP analysis due to the sensitivity of capillary electrophoresis.*

Samples were analyzed at the Macquarie University DNA sequencing facility. Prior to capillary electrophoresis, 10 μl of HiDi/Standards master mixture comprising 9.9 μl of formamide and 0.1 μl of the internal DNA standard LIZ500 (Applied Biosystems) was added to 1 μl of template, and this was followed by denaturation at 95°C for 5 min. The fluorescence of the T-RFs was detected using an ABI Prism 3130x1 genetic analyzer (Applied Biosystems) in Genescan mode (8.5 kV; 40-s injection; 60°C for 100 min) with a G5 filter. T-RF sizes were determined using GeneScan software, version 4.0 (Applied Biosystems)."

Instead of Red Hot Taq or Accutaq our study used Gotaq all in one master mix (Promega) as the MHC work undertaken in this chapter encountered significant challenges with PCR contamination. Additionally, instead of Vspl restriction enzyme this study used BamH1, HindIII and EcoR1 both singularly and in combination (chosen based on availability). Both the conditions for running the samples at Macquarie University DNA sequencing facility and the software used to analyse the data was the same in both studies.

The strong products produced on agarose gel in the target range suggested that this method successfully amplified MHC sequence. However, no further analysis could be performed as the electrophoretograms did not produce distinct or repeatable peaks.

Appendix 4.2

The method used by Binz et al (2001) to genotype MHC using SSCP is reproduced below

"PCR reactions (10 µl) were carried out in either a Geneamp 2400 thermal cycler (PE Biosystems) or a Biometra Thermocycler and employed: 1 µl of spine extract, 1X buffer (Amplitaq buffer II, PE Biosystems), 0·5UofAmplitaq Gold polymerase (PE Biosystems), 2·5m MgCl2, 200 µl dNTPs, 0·5 µl fluorescent labelled forward primer (5 ' label: FAM), 0·5 µl labelled reverse primer(5' label: HEX). Some PCR reactions were also run using Promega Taq polymerase and buffer system 'B' (PE Biosystems). PCR cycling parameters were the following: 10 min at 95° C (denaturation of the template and activation of the Amplitaq Gold polymerase),40 cycles (33 when using the Promega Taq) of 95° C (denaturation) for 30 s, 15 s at 56° C (annealing) and 75 s at 72° C (primer extension) followed by a final step of 72° C for7 min. 1 µl of PCR product was mixed with 9 µl loading mix (for 48 reactions: 396 ml lformamide, 12·5 ml of commercial DNA standard ROX 350 (PE Biosystems), 12·5 µl ROX-labelled PCR products of particular alleles used as additional internal standards and 22 µl of 0·3 M NaOH. The mixture was denatured at 95° C for 5 min, snap cooled on ice and analysed by capillary electrophoresis on a ABI 310 Genetic Analyzer (PE Biosystems) using the following conditions: the polymer concentration was 5% (Genescan polymer, PE Biosystems) solubilized in 5% glycerol and 0·5 X TBE. The running buffer consisted of 0·9 µl TBE and 10% glycerol. Negative controls were always included. Run conditions were 12 kV for 25 min at 30° C using the short capillary from PE Biosystems (green band, 34 cm)."

In our method as we had unique forward primers (as described previously in Chapter 4) using individually labelled reverse primers was not necessary. Further, all analysis was undertaken at Macquarie University DNA sequencing facility.

References

Roux, K.H. (1995). Optimization and troubleshooting in PCR. *Genome Research*, *4*(5): pp.S185-S194.