Spatial and temporal genetic structuring in yellow-eyed penguins

Sanne Boessenkool

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Otago, Dunedin,
New Zealand

April 2009

Abstract

Improving our understanding of the forces driving population decline and the processes that affect the dynamics of threatened populations is central to the success of conservation management. The application of genetic tools, including our ability to examine ancient DNA, has now revolutionised our ability to investigate these processes. The recent human settlement of the Pacific, particularly in New Zealand, provides a unique, accessible system for revealing anthropogenic impacts on native biota. In this thesis I use genetic analyses from modern, historic and subfossil DNA to investigate temporal and spatial genetic structuring of the endangered yellow-eyed penguin (*Megadyptes antipodes*), and use these analyses to answer questions related to the conservation of this species.

The yellow-eyed penguin is endemic to the New Zealand region and currently breeds on the subantarctic Auckland and Campbell Islands and the southeast coast of the South Island. The current total population size is estimated around 6000-7000 individuals, of which more than 60% inhabit the subantarctic. Despite intensive conservation measures by governmental and local community agencies, population sizes have remained highly unstable with strong fluctuations in numbers on the South Island. The species was believed to be more widespread and abundant before human colonisation of New Zealand, thus current management assumed the mainland population to be a declining remnant of a larger prehistoric population.

Genetic and morphological analyses of subfossil, historic and modern penguin samples revealed an unexpected pattern of penguin extinction and expansion. Only in the last few hundred years did *M. antipodes* expand its range from the subantarctic to the New Zealand mainland. This range expansion was apparently facilitated by the extinction of *M. antipodes*' previously unrecognised sister species, *M. waitaha*, following Polynesian

settlement in New Zealand. The demise of *M. waitaha* is the only known human-mediated extinction of a penguin species.

Despite *M. antipodes*' recent range expansion, genetic analyses of microsatellite markers reveal two genetically and geographically distinct assemblages: South Island versus subantarctic populations. We detected only two first generation migrants that had dispersed from the subantarctic to the South Island, suggesting a migration rate of less than 2%. Moreover, the South Island population has low genetic variability compared to the subantarctic population. Temporal genetic analyses of historic and modern penguin specimens further revealed that the harmonic mean effective population size of the *M. antipodes* South Island population is low (<200). These findings suggest that the South Island population was founded by only a small number of individuals, and that subsequent levels of gene flow have remained low.

Finally, we present a novel approach to detect errors in historic museum specimen data in cases where *a priori* suspicion is absent. Museum specimens provide an invaluable resource for biological research, but the scientific value of specimens is compromised by the presence of errors in collection data. Using individual-based genetic analysis of contemporary and historic microsatellite data we detected eight yellow-eyed penguin specimens with what appear to be fraudulently labelled collection locations. This finding suggests errors in locality data may be more common than previously suspected, and serves as a warning to all who use archive specimens to invest time in the verification of specimen data.

Overall, yellow-eyed penguins have a remarkable dynamic history of recent expansion, which has resulted in two demographically independent populations. These results reveal that anthropogenic impacts may be far more complex than previously appreciated.

Acknowledgements

Many wonderful people gave me support, advice and encouragement during my PhD. Without their contribution this thesis would not have been possible. I would like to start with thanking my supervisors, Jonathan Waters and Philip Seddon. You two gave me great space and freedom to design and fulfill my research, but at the same time you were always there when I needed help and guidance. Thank you both for your encouragement and trust. Jon, you motivated me to focus on publishing my papers and taught me how to sell a story. Your way with words is an inspiration, as is your astonishingly fast feedback on drafts – I was shocked many times when a corrected draft was back in my inbox within just hours after sending it. It sure made progress fast and efficient. Phil, you have the truly amazing ability to feel when I needed some extra encouragement and support. Your words made me smile, sometimes brought tears to my eyes, they picked me up when I was feeling low and helped me recover my confidence when it had temporarily departed me. Thank you for believing in me as a person.

My time in the genetics lab would not have been the same without Tania King. You have taught me innumerable laboratory techniques and showed me how to manage a lab with all its users. You've made me appreciate knowing the basics and made me understand that learning things the slow, hard way can actually be beneficial in the long term. Thank you for your friendship and for making life in the lab so enjoyable.

I am indebted to numerous people at the Australian Centre for Ancient DNA who taught me the necessary skills and the persistence needed to work with old samples, and who made my time in Adelaide so much fun. I am particularly grateful to Jeremy Austin, who made time in his busy schedule to teach me, it was really great to work with you. Nic and Maria, you two are hilarious, thank you for looking after me. Baby swans, chocolate penguin,

botanic gardens... Sarah, you're lovely, stay in touch and come and visit any time. Finally, Rene and Emily, thank you for letting me stay in your lovely little home.

Paul Scofield and Trevor Worthy provided invaluable help with the identification of penguin bones and the species description. I truly enjoyed collaborating with you guys and hope we will continue to do so in the future.

Many people helped me in the field, especially Ursula Ellenberg, Thomas Mattern and Ryan Clark who taught me how to blood sample and monitor penguins. It was great working with you all. Department of Conservation staff provided a lot of support during field work, particularly Melanie Young, Pete McClelland, Bruce McKinlay, Sheryl Pullar and Kevin Pierce. I'd also like to thank Rosalie Goldsworthy. A big thanks to Sandy King and Joanna Wright from the Yellow-eyed Penguin Trust for organizing the field trip to Codfish and Stewart Island. Samples on offshore islands were collected by Jon Potter, Matt Charteris and Annabel Gosset (Campbell Island), Louise Chilvers and the sea lion team (Enderby Island), Kath Walker and Graeme Elliott (Adams Island), Thomas Mattern, Tiffany Browne and Andrew Hill (Stewart and Codfish Islands). Thanks also to Lloyd Davis, Alvin Setiawan and Melanie Massaro for letting me use their samples collected on Otago Peninsula in 2001. Finally I'd like to thank everybody else who assisted me walking up and down the beautiful cliffs of Boulder Beach.

I am very grateful to the Auckland Museum, American Natural History Museum, Australian Museum, Canterbury Museum, Museum of Comparative Zoology, Natural History Museum Geneva, Natural History Museum Tring, Natural History Museum Paris, Museum of New Zealand Te Papa Tongarewa, Natural History Museum Vienna, Swedish Museum of Natural History, Otago Museum, South Australian Museum, Smithsonian Institution, Museum für Naturkunde Berlin for supplying tissue samples of historic specimens. I was further permitted to sample penguin bones held at Canterbury Museum, Museum of New Zealand Te Papa Tongarewa, Otago Museum and the

Archaeology Department at the University of Otago. Without your support this work would not have been possible.

The Department of Zoology is a truly amazing place thanks to all the lovely people who work there. I think back to the days that I was part of the 'zoo' family with great fondness. I am especially indebted to Ronda Peacock and Wendy Shanks, who make the reception a joyous place and were never tired to help out with any problems (and thanks for the dozens of complements on my dressing up efforts!); Vivienne McNaughton, who was forever helpful in the store; Ken Miller, whose understanding of graphics, and particularly photoshop, is invaluable; Murray McKenzie, who made my special penguin bone holder for successful drilling; Karen Judge, for help sorting a variety of lab issues. I also thank Dave Wharton for allowing us to invade his lab for the museum specimen DNA analysis, and Hamish Spencer for letting me use one of his computers. I would further like to thank Martyn Kennedy for teaching me a number of phylogenetic programs, and John Darby for sharing his penguin knowledge with me.

I am grateful for scholarship support, awards and grants from the University of Otago, the New Zealand Federation for Graduate Women and the New Zealand Educated. Conference funding from the Ecological Research Group, the New Zealand Society for Biochemistry and Molecular Biology, the Division of Sciences and the Department of Zoology allowed me to share my research results with the national and international scientific community.

A big thanks to everybody at the Centre for Ecological and Evolutionary Synthesis at the University of Oslo for welcoming me into their department and for providing me desk space and a computer during the last six months of my PhD. Your support was truly amazing.

Ian Jamieson, over the years you have become a mentor to me. You made me appreciate the importance of learning skills rather than simply getting stuck into my own research. You taught me lots about communication with

stakeholders, showed me how to review papers, and most importantly, encouraged that combining a research career with other dreams in life is perfectly doable. I miss talking over a beer on Friday afternoons. Our farewell dance party at your and Frances' beautiful house was unforgettable. You two are always welcome in our home.

My dearest friends in Dunedin, you made the PhD such a joyous ride. Catherine Grueber, you're a never-ending inspiration, by far the smartest girl I've ever met, and I greatly admire your down-to-earth nature and your impressive perseverance to tackle all the challenges you are faced with. I missed having you around the past year, but am thankful we can talk over the phone almost every week and can't wait to see you again. My a-little-bit-crazy darling, Crid Fraser, the best office mate one can wish for. The aquarium with your little baby anemones, your art and photos on the wall, the decorated office door, and most importantly your happy presence, made sharing an office with you a truly unforgettable experience. Thank you for the hours of coffee drinking at EG, the relaxing wine (in half liter glasses☺) in front of your fire place and for feeding us Thai 'Crid' curries. You're a great friend. Frenchies, the international lunch club was awesome and I really look forward to catching up soon in Europe. There are many others and I can't mention you all by name, but I assure you that you are all in my thoughts.

I'd like to dearly thank my dad, mum, brother, Marijke, Willem, all the other members of my family and my family-in law, for supporting our move to the other side of the world and for appreciating the joy we found there. Dad, mum, thanks for showing me the world is out there full of places to explore and thanks for your encouragement and friendship.

Finally, my deepest thanks go to Bastiaan Star, my husband, my partner, my very best friend. I couldn't have done this PhD without you. Especially in the past six months, after we moved back to Europe, you're help with proof reading, figures and discussing ideas has been truly amazing. Thank you for

your endless encouragements, support, patience, and your love. I am a lucky girl to have you by my side.

Table of contents

List of Figures

List of Tables

Chapter 1:

General introduction

Human settlement throughout the globe has initiated a dramatic species extinction crisis and is leading to an increasing number of small and declining wild populations. Conservation managers are faced with the daunting task of protecting, maintaining and eventually restoring the world's biodiversity where possible. Understanding the forces driving population decline, and the processes and dynamics that affect small populations, are central to the success of these conservation efforts. In this thesis I apply temporal and spatial genetic analyses to investigate the current and historical genetic structuring of yellow-eyed penguins, *Megadyptes antipodes*, and apply these results to improve our understanding of the dynamics that affect this endangered species.

Genetics in conservation biology

The rapid development of genetic technologies and associated analytical tools during the last few decades has hugely expanded the role of genetics in conservation biology, and led to the emergence of the field of conservation genetics. The ability to sequence ancient DNA has enabled a critical expansion of this field by adding a temporal dimension that allows direct comparisons of past and present patterns of genetic diversity. The role of genetics in managing endangered populations can essentially be divided into two components (Avise 1994; Frankham *et al.* 2002; Allendorf & Luikart 2007). First, there have been numerous studies assessing direct genetic consequences of population decline, which primarily include the loss of genetic diversity and inbreeding depression. Second, genetics can be used as a tool to elucidate systematics, units for conservation, and a variety of demographic processes and dynamics of populations. The following section discusses the key scientific issues emerging within these two distinct subdisciplines of conservation genetics.

Direct genetic consequences of population decline

As a population declines, it becomes increasingly affected by stochastic dynamic processes that are of a demographic (e.g. differential survival and reproduction of individuals), an environmental (e.g. variation in weather, food supply, predators, competitors) and/or a genetic nature (Shaffer 1981; Shaffer 1987). The relative importance of stochastic genetic processes is determined by the population's effective size (N_e), which is consequently one of the most important parameters in conservation genetics (Waples 2002). In a genetic context, N_e is defined as the size of an ideal population experiencing the same rate of genetic drift as the actual population under consideration, where ideal populations are those with no selection, constant population size, random mating, equal sex ratios, discrete generations, and random variation in reproductive success (Wright 1931; Frankham 1995; Palstra & Ruzzante 2008). Although it is difficult to predict specific minimum N_e thresholds required to avoid the genetic problems that typically arise in small populations, theoretical studies have suggested that approximately $N_e = 50$ is needed to minimize inbreeding depression and $N_e = 500$ is required to maintain sufficient evolutionary potential (although thresholds as high as 5000 have been proposed; Franklin 1980; Franklin & Frankham 1998; Lynch & Lande 1998). Despite this controversy, these numbers provide rough but valuable guidelines in conservation management.

Two of the most critical conservation genetic processes – genetic drift and inbreeding – often coincide in small populations, but they may also occur independently of each other, affecting the genetic diversity of a population in different ways (Jamieson *et al.* 2008). Genetic drift is the random fluctuation in allele frequencies due to sampling effects, which can lead, for instance, to the erosion of beneficial alleles and the accumulation of slightly deleterious alleles. The biological consequences (e.g. reduced fitness) associated with such a 'mutational meltdown' may occur gradually and therefore may not be immediately apparent (Keller & Waller 2002). In the long term, however, this loss of genetic diversity can reduce a population's adaptive potential to a

changing environment (Lacy 1987). In contrast to genetic drift, inbreeding is a swift process that can have immediate effects on individual and population-level fitness (Keller & Waller 2002). These fitness effects, termed inbreeding depression, typically result from an increase in homozygosity for recessive deleterious alleles. Overall, the relative importance of genetic drift and inbreeding becomes increasingly important as a population declines in size.

The importance of loss of genetic diversity and inbreeding depression in determining population viability and extinction probability was hotly debated in the 1980s and 1990s (e.g. Soulé 1987; Lande 1988; Caughley 1994; Hedrick *et al.* 1996) and continues to generate discussion today (e.g. Jamieson 2007). Genetic factors rarely act alone, however, implying that a more holistic approach is needed to integrate the demographic, environmental and genetic factors that affect populations. Despite the ongoing debate, population managers are aware of genetic problems affecting wildlife populations and genetic assessment is now an important element in the conservation management of small and declining populations.

Genetics as a tool in conservation

The second, major role of genetics in conservation biology lies in the relatively broad use of genetic data as a tool to elucidate conservation units, demographic processes and dynamics of populations. Such information is used to guide a wide variety of management decisions and adds to our overall understanding of population dynamics through space and time. I will not attempt to provide an exhaustive overview of the applications of genetic data to conservation biology, but will briefly touch on a few examples that are relevant to this thesis.

First, analyses of genetic data can help resolve systematic issues and establish associated priorities for conservation (Haig 1998). For example, the taxonomic boundaries of albatrosses (family Diomedeidae) have been much

debated and the designation of species or subspecies status to certain groups has been questioned (Abbott & Double 2003). Mitochondrial control region analyses by Abbott and Double (2003) confirmed the full species status of Salvin's and Chatham albatrosses as well as the taxonomic separation of these two species from the shy and white-capped albatrosses, whereas the status of shy and white-capped albatrosses remained problematic (Abbott & Double 2003). A second example comes from the Australasian teals, whose taxonomic status has also been subject of much debate (Kennedy & Spencer 2000). Phylogenetic analyses based on three mitochondrial DNA genes shed light on this issue and confirmed that the Brown, Auckland Island and Campbell island teals should be recognised as separate species. This recognition supported the development of individual conservation programs for these rare New Zealand teals (Kennedy & Spencer 2000).

Second, genetic data are used in the identification of populations, information that plays a key role for determining appropriate scales for management (Cegelski *et al.* 2003; Waples & Gaggiotti 2006). This process is not always straightforward, however, as populations can be defined within either 1) an evolutionary paradigm, in which case they can be considered evolutionary significant units (ESUs), or 2) an ecological paradigm, in which case they may form separate management units (MUs) (Moritz 1994; Waples & Gaggiotti 2006; Palsbøll *et al.* 2007). There are numerous examples for the use of genetic data to delineate units for conservation. Classic work comes from the early 1990s, when phylogeographic patterns revealed a continuum between deep evolutionary differentiation and more recent population subdivision among populations of freshwater, coastal and marine species (Avise 1992; Avise 1994). The discussion of deep versus shallow divergence, or in other words the recognition of ESUs versus MUs, has also played an important role in the conservation of Scandinavian brown bears. Initial work on mitochondrial DNA suggested that these bears should be managed as two ESUs (Taberlet *et al.* 1995), but later analysis of nuclear microsatellite data revealed that recognition as a single ESU with four (Waits *et al.* 2000) or three (Manel *et al.* 2004) MUs was more appropriate. Recently, genetic

analyses from radiated tortoises (*Geochelone radiata*), a species subject to strong anthropogenic pressures as a result of habitat destruction and poaching, revealed three distinct MUs and identified two rivers as major barriers to dispersal (Paquette *et al.* 2007). Importantly, the identification of management units hinges on obtaining estimates of dispersal rates, but the estimates of dispersal rates themselves also provide significant insights into the vulnerability of populations.

Third, genetic data are applied to quantify levels of migration or gene flow among populations. Immigration and associated gene flow can have a strong influence on the survival probability of small populations by 1) increasing numerical abundance (the 'rescue effect'; Brown & Kodric-Brown 1977), 2) enhancing population growth rate by reducing the effects of inbreeding, and maintaining or increasing genetic variability (the 'genetic rescue effect'; Madsen *et al.* 1999; Richards 2000; Vilà *et al.* 2003; Hedrick 2004), or 3) recolonising habitat patches where populations have gone extinct (Hanski 1998). The estimation of dispersal rates has become particularly important for species that are faced with increasingly fragmented habitats, as exemplified by work on wolverine (*Gulo gulo*; Cegelski *et al.* 2003) and gorilla (*Gorilla gorilla diehli*; Bergl & Vigilant 2007).

Finally, genetic data have proven to be extremely valuable as a tool in wildlife forensics such as the identification of illegal trade or poaching (Manel *et al.* 2002; Manel *et al.* 2005). This novel research area was initiated by the early work from Baker and Palumbi, who used mitochondrial DNA markers to identify meat from protected whales and other mammal species that was being sold illegally at commercial markets in Japan (Baker & Palumbi 1994). More recently, protocols have been developed to determine species of origin from detached shark fins (Shivji *et al.* 2002) and the geographic origin of poached ivory (Wasser *et al.* 2004).

Overall, the examples discussed above reflect the enormously wide utility of genetic analysis as a tool to investigate patterns and processes relevant to

conservation biology. The above examples all focus on the use of genetic data from contemporary samples. Over the last two decades, however, great progress has been made in our ability to obtain DNA from historic or more ancient source material. This use of ancient DNA (aDNA) techniques to characterise historic genetic diversity and elucidate ancient population dynamics has potential to revolutionise approaches to conservation of extant populations.

In addition to the advent of new DNA techniques, the development of advanced statistical methods has considerably broadened the horizons of conservation genetic research. Traditionally, many of the analytical methods depended on idealised population models (e.g. Wright 1931; Nei 1978; Slatkin 1985) based on biologically unrealistic assumptions (e.g. infinitely large populations). Reliance on such assumptions has proven particularly problematic in conservation biology, a field that typically focuses on small and declining populations (Pearse & Crandall 2004). Although these traditional methods remain valuable as basic descriptors of genetic diversity (Frankham *et al.* 2002; Allendorf & Luikart 2007), exponential increases in computer power over the past decade, together with the development of likelihood and Bayesian statistical approaches, has led to a revolution in relatively advanced genetic analytical methods (e.g. Beaumont & Rannala 2004; Excoffier & Heckel 2006). The advent of these new techniques has enabled researchers to retrieve new insights from existing datasets, and provides enormous scope for novel research directions. For example, the development of admixture analyses and assignment tests has opened up the possibility to identify the origins of individuals, to objectively identify genetic assemblages, to estimate dispersal rates, and to study admixture or even hybridization among populations or species (reviewed in Manel *et al.* 2005). Furthermore, the use of Bayesian statistics has the great advantage for conservation biologists, providing direct probability assessment rather than null hypothesis testing. Such actual probabilities are more tangible and easier to interpret for conservation managers when research outcomes are used to guide management decisions.

Most current applications of genetics in conservation rely on the use of 'neutral' genetic markers (but see Gemmell *et al.* 2004) such as nuclear microsatellite loci or sequences of the mitochondrial genome (e.g. from the rapidly evolving control region). Microsatellites consist of repetitive simple sequences that are interspersed throughout the eukaryotic genome (Tautz 1989; Schlötterer 2000). These fast-evolving loci typically show high levels of genetic variation due to their high mutation rate, a feature that enhances statistical power and allows for identification of separate entities up to the individual level (Urquhart *et al.* 1995). In contrast, the mitochondrial control region (or d-loop) is a more powerful tool to study the ancestry of species or populations, primarily due to its maternal inheritance and lack of recombination. This mtDNA region is typically divided into three sections, with a highly conserved central region surrounded by two hypervariable sequences (HVI and HVII) (Howell *et al.* 1996; Lambert *et al.* 2002). Importantly, mitochondrial DNA is present in many copies throughout the cell, which is advantageous when working with low quality degraded samples (see below).

The use of ancient DNA

The field of ancient DNA research was initiated by the successful retrieval of DNA from the quagga, an extinct member of the horse family (Higuchi *et al.* 1984). One year later, one of the main pioneers of ancient DNA study, Svante Pääbo, reported the successful extraction of DNA from an Egyptian mummy (Pääbo 1985). These first two studies, however, used labour intensive cloning techniques to investigate DNA from their ancient samples, which severely restricted the utility of ancient genetic samples for several years. Only after the invention of the Polymerase Chain Reaction (PCR) in the late 1980s did the field of ancient DNA research expand to include a wide variety of historic material. In the next decades an ever increasing range of extinct taxa were studied using aDNA techniques, including the marsupial wolf, the moa and

the mammoth (see review by Hofreiter *et al.* 2001). Several studies even reported spectacular achievements such as the successful extraction from insect and plant DNA out of amber from the Oligocene, and even dinosaur DNA from the Cretaceous (see references in Austin *et al.* 1997). However, subsequent analyses using more rigorous and reproducible methodologies revealed the inherent difficulties of recovering truly ancient DNA, and showed many of the early spectacular claims to be non-authentic (Austin *et al.* 1997; Willerslev & Cooper 2005).

The difficulties that arise when working with ancient DNA result from 1) the degraded and modified nature of ancient DNA and 2) the very low quantity of endogenous DNA and subsequent high risk of contamination with exogenous DNA (see reviews by Wayne *et al.* 1999; Hofreiter *et al.* 2001; Pääbo *et al.* 2004; Willerslev & Cooper 2005). Natural DNA repair systems no longer function after an organism dies and DNA is slowly degraded by enzymes and micro-organisms, resulting in an accumulation of DNA damage over time. The speed with which this degradation occurs depends largely on the environment in which the DNA is preserved. For example, in cold climates this process is significantly slower than in warmer climates. DNA damage hinders successful amplification of and/or leads to nucleotide misincorporations during amplification (Pääbo *et al.* 2004). As a result, aDNA protocols now require replicate amplifications from individual samples before a consensus sequences can be scored. In addition, the inherently low quantities of endogenous DNA that characterise ancient samples equate to a high risk of contamination. Common sources of 'foreign' DNA include human cells (e.g. from archaeologists who excavated the material, museum curatorial staff, biologists, or even human DNA that enters the laboratory via plastic ware), microbial organisms, DNA from closely related species (e.g. domestic animals), or PCR product back-contamination. To minimize the risk of contamination strict laboratory guidelines and criteria have been developed (Cooper & Poinar 2000; Willerslev & Cooper 2005), and researchers need to take a cognitive and critical approach when assessing the authenticity of their data (Gilbert *et al.* 2005).

Initially, most ancient DNA research focused on extinct species (Hofreiter *et al.* 2001), but in the last decade the focus has widened and there have been an increasing number of studies on extant species. Importantly, technical advances improved the feasibility of analysing larger numbers of samples, and have thus allowed consideration of ancient DNA in population genetic studies. Such research typically involves a comparison of genetic diversity over time, facilitating inferences about temporal changes in population size and population connectivity. For example, Paxinos and others (2002) analysed DNA from the nene (Hawaiian goose, *Branta sandvicensis*) sampled over four different time periods (modern, 1900s, 400 years before present, and up to 2500 years before present). Comparison of the genetic diversity revealed that the nene lost most of its genetic diversity in prehistoric times during a period of early human population growth and expansion of settlements in Hawaii, rather than during the known recent population bottleneck. Other recent examples of aDNA research on extant taxa include the work on brown bears (Leonard *et al.* 2000; Barnes *et al.* 2002) and wolves (Leonard *et al.* 2005; Leonard *et al.* 2007; Leonard & Wayne 2008) in North America, which have led to the appreciation that species that survived the great megafaunal extinctions suffered genetic and ecological declines across these periods (Hofreiter 2007).

Historic museum specimens collected by early naturalists in the late 1800s and early 1900s have also proven to be a valuable resource for population genetic research, particularly for evaluating the more recent anthropogenic impacts on natural populations (e.g. Miller & Waits 2003; Johnson *et al.* 2004; Larsson *et al.* 2008; Taylor *et al.* 2008). The application of these ancient and historic DNA analyses to endangered species potentially has major implications for conservation management. Comparisons of contemporary and (pre)historic genetic diversity can, for example, reveal causes of population decline, help prioritise and identify units for conservation, guide captive breeding programs and reintroductions, reveal historic population sizes, and reveal historic population connectivity (see

Leonard 2008 for an extensive review on this topic). Furthermore, aDNA studies can reveal previously unrecognised impacts of human settlement in pristine wildlife areas. This field holds particular promise in the Pacific – the last frontier of human colonisation – where many islands were settled only in the last few thousand years.

New Zealand – the last major landmass to be colonised by humans

Around 65-80 million years ago, the New Zealand landmass split off from Gondwanaland and slowly became the isolated island archipelago it is today. As a result of this isolation, a most remarkable flora and fauna evolved with a high degree of endemism. Notably, New Zealand was characterised by a complete lack of mammals, with the exception of two species of bats (Worthy & Holdaway 2002; but see Worthy *et al.* 2006), which allowed birds to become the dominant fauna of these islands. A relatively large proportion of the avifauna was flightless or had reduced ability to fly, large bodied land birds were common and seabirds were numerous (Worthy & Holdaway 2002). The long isolation and the lack of mammalian predators had resulted in a naïve fauna that was particularly vulnerable to human colonisation.

Polynesian settlers arrived in New Zealand from a tropical east Polynesian homeland *c.*1280 AD (Whyte *et al.* 2005; Wilmshurst *et al.* 2008). The livelihood of these early settlers depended primarily on hunting, supplemented with horticulture of Polynesian food crops such as the *kumara* (Walter *et al.* 2006). Southern New Zealand was, however, too cold for the cultivation of these tropical Polynesians crops and people relied on bracken fern root and cabbage tree to supplement resources from hunting (Anderson & Smith 1996; Anderson 2002). The early settlements were so-called 'transient villages' that were regularly relocated as local resources became depleted (Walter *et al.* 2006). On the South Island of New Zealand these settlements were concentrated along the east coast where a wide variety of

marine resources was available, along with easy accessibility to large birds such as the moa (Anderson 1989).

The reliance on hunting for subsistence led to rapid depletion of large game. This transition has been documented clearly by the studies on stratified middens, which reveal a marked shift from big game (e.g. moa, seals) to small game (e.g. small birds, fish, shellfish) within just decades of human settlement (Anderson *et al.* 1996; Nagaoka 2001). In addition to the hunting pressure, large scale habitat destruction was brought about by forest clearance due to anthropogenic fires (McGlone 1983). Finally, the introduction of the Pacific rat (the *kiore*) and the Polynesian dog (the *kuri*) has played a role in the demise of a fauna that was naïve to mammalian predators (Worthy 1999). Overall, Polynesian settlement in New Zealand lead to the extinction of approximately 33 taxa of resident birds, and severe declines in another ~30 species (Worthy 1999).

The first Europeans to arrive in New Zealand were Abel Janszoon Tasman and his crew in 1642, followed by several voyages lead by Captain James Cook in the late 1700s. Substantial settlement by European migrants did not commence until after the signing of the Treaty of Waitangi in 1840, when the country became a British colony (King 2003). For the fauna of New Zealand, the settlement by Europeans had two major consequences: 1) further habitat destruction as a result of forest clearance for large scale farming and the harvest of timber; 2) immense predation pressure with the introduction of predatory mammals such as cats, rats, stoats, ferrets, weasels and possums (see references in Moors 1983; O'Donnell 1996; Craig *et al.* 2000). Current estimates show that at least 41% of the endemic bird species have become extinct in New Zealand since human settlement commenced, and of those species remaining, 35% are now classified as threatened (Worthy & Holdaway 2002). These numbers are likely to increase, as the impact of European settlement is ongoing and is likely to eventually exceed that of Polynesian settlement (Worthy 1999).

The subantarctic islands of New Zealand (Antipodes, Auckland, Bounty, Campbell and Snares Islands) are not permanently inhabited by humans. Nevertheless, sealers and whalers visited most of these islands in the past and introduced various exotic species (e.g. pigs, rabbits, mice, cats, rats). The Auckland Islands were settled by Polynesians for a few years in the 13th century, and a second settlement period took place in the 1840s and 1850s. Both of these settlements probably ceased as a result of the harsh weather conditions (Anderson 2005). Farming was introduced on Auckland and Campbell Islands in the late 19th century but abandoned again in the 20th century. In 1998, New Zealand's subantarctic islands were given Natural World Heritage recognition for their high levels of biodiversity and endemism, and a conservation plan was adopted focusing on the full restoration of the terrestrial ecosystems (UNEP 2008). All recently introduced animals have now been removed from Campbell Island, while rabbits and mice have also been eradicated from Enderby Island (part of the Auckland Islands). Eventual removal of all introduced species from all the subantarctic islands is planned (UNEP 2008). Overall, therefore, these subantarctic islands can be regarded as relatively unmodified and pristine.

New Zealand is well-known throughout the world for its expertise relating to the management of introduced species and the translocation of endemic biota onto offshore islands or into predator-free reserves on the mainland. In addition to strong government-driven conservation policy, the New Zealand public has invested heavily in community-led restoration projects. Certainly, the presence of large iconic species has helped to mobilise public support and many of these species are now flagships for conservation programs. One such conservation icon is the yellow-eyed penguin, *Megadyptes antipodes*.

The yellow-eyed penguin

Megadyptes antipodes, also known as *hoiho*, was until recently considered the sole species within its genus (but see Boessenkool *et al.* 2009a, Chapter

2). This penguin is most closely related to the *Eudyptes* penguin species, from which it diverged some 15 million years ago (figure 1.1; Bertelli & Giannini 2005; Baker *et al.* 2006; Ksepka *et al.* 2006; Clarke *et al.* 2007). The species is endemic to the New Zealand region and currently breeds on the subantarctic Auckland and Campbell Islands and the southeast coast of the South Island (including surrounding islands such as Stewart Island; Marchant & Higgins 1990; McKinlay 2001). The current total population size is estimated around 6000-7000 individuals, of which more than 60% inhabit the subantarctic (McKinlay 2001). The species has been classified as endangered by the IUCN (EN B2b(iii)c(iv)) based on its confined breeding range, the decline in suitable habitat and the extreme fluctuations in numbers (Birdlife International 2008).

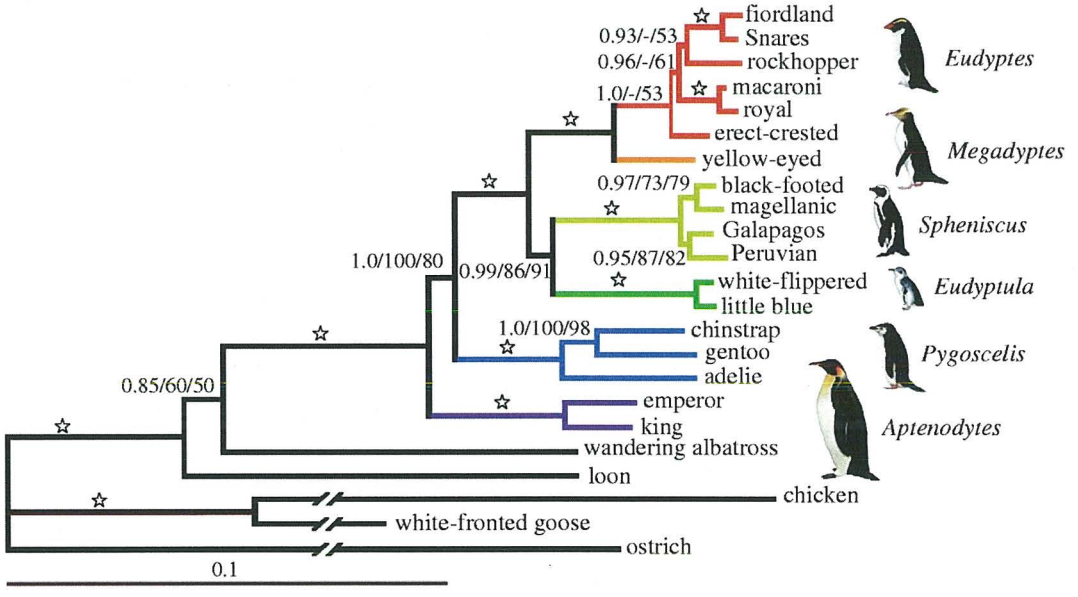


Figure 1.1 Phylogeny of extant penguins from Baker *et al.* 2006. The tree represents a Bayesian phylogeny based on sequences from one nuclear gene and four mitochondrial DNA regions. Numbers above branches are Bayesian posterior probabilities/ML bootstrap support/MP bootstrap support which are presented as open star when 1.0/100/100.

Yellow-eyed penguins are socially monogamous and adults stay at their breeding area year round. They have been described as the least colonial of all penguins: patchily distributed breeding sites consist of loose aggregations of nests that are usually visually isolated from each other (Darby & Seddon 1990). The breeding cycle starts around August-September, when pairs select a nest site and one or two eggs are laid. Chicks hatch synchronously (in nests with two eggs) after an incubation period of 40-50 days (Darby & Seddon 1990). The roles of incubation, chick feeding and chick guarding are shared equally between both parents. Following fledging in late January-March juvenile birds disperse, but almost 90% of the surviving juveniles eventually return to breed near their natal area (Richdale 1957). Females reach reproductive maturity when they are 2-4 years old, whereas males usually commence breeding a little later at 3-4 years (Richdale 1957). First year survival is low (~40%), but once adult age has been reached survival rates are high (~85-90%) and individuals may live for up to 20 or even 25 years (Richdale 1957; Department of Conservation unpublished data).

Conservation efforts for yellow-eyed penguins on and around the South Island of New Zealand have largely focused on predator trapping – chiefly targeting mustelids and rats – and revegetation of coastal habitat. Despite these measures, however, population sizes have remained unstable and fluctuated strongly over recent decades (McKinlay 2001; Moore 2001). This demographic instability has been attributed to changes in food supply (van Heezik & Davis 1990), climatic variations (Peacock *et al.* 2000) and disease epidemics (e.g. Gill & Darby 1993; Department of Conservation unpublished data). Although ongoing research is increasing our knowledge in these areas, such threats are inherently difficult (if not impossible) to control. The long term survival of *M. antipodes*, therefore, will likely depend on the species' intrinsic resilience.

Thesis outline

In this thesis I apply genetic analyses from modern, historic and subfossil DNA to investigate temporal and spatial genetic structuring of yellow-eyed penguins, and apply results to questions related to the conservation of this species. Chapter two tests for temporal changes in *M. antipodes* genetic diversity associated with human settlement of New Zealand by assessing mitochondrial DNA variation of prehistoric, historic and modern penguin samples. Chapter three describes the isolation and characterization of twelve microsatellite DNA markers from enriched genomic libraries. Chapter four assesses the hypothesis that *M. antipodes* comprises a single demographic population across its New Zealand – subantarctic range. Chapter five uses microsatellite analyses of contemporary and historic South Island samples to test for temporal changes in genetic diversity over the last century, and to calculate genetic estimates of the effective population size of South Island yellow-eyed penguins. Finally, chapter six describes how individual based genetic analyses can reveal previously unsuspected inaccuracies in the geographic origin of museum material, and uses this method to demonstrate historic falsification of archive *M. antipodes* specimens.

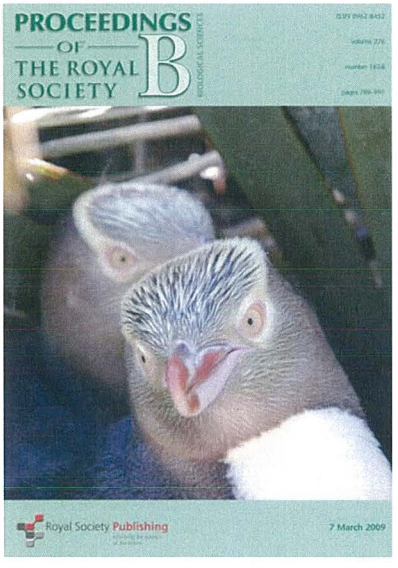
I have used the first person plural for all data chapters because these are co-authored manuscripts (of which I am the first author) that are either published (Chapter 2, 3 and 4), intended for publication (Chapter 5) or submitted (Chapter 6). I designed, performed, analysed and wrote all first drafts of the research presented in this thesis, but my co-authors performed various crucial roles of support and help. Jonathan Waters and Phil Seddon advised me on the study design and the analyses, helped interpret results, read and commented on the manuscripts and the general introduction and discussion of this thesis. Jeremy Austin taught me the techniques for genetic analysis of the bone samples, helped with the analyses of these samples and read and commented on the manuscript of Chapter 2. Alan Cooper provided laboratory space for the ancient DNA analysis and read and commented on the manuscript of Chapter 2. Trevor Worthy and Paul Scofield advised me on

the identification of penguin bones, helped interpret the penguin extinction and colonisation, provided the qualitative morphological description of the new penguin species and read and commented on the manuscript of Chapter 2. For Chapter 6, Paul Scofield provided important feedback and information about H. H. Travers and read and commented on the manuscript. Bastiaan Star provided feedback, helped interpret results, helped design and make some of the figures, and read and commented on the manuscripts. Tania King helped with the development of the microsatellite libraries and read and commented on the manuscript of Chapter 3.

Since all data chapters have been written in the style of scientific research papers there may be some overlap in the introduction sections of the various chapters. Permits to conduct this research included Department of Conservation collection permits for Otago (OT-19097-RES) and Southland (SO-17933-FAU); University of Otago Animal Ethics approval 69/06; Environmental Risk Management Approval GMO05/UO028.

Chapter 2:

Relict or colonizer? Extinction and range expansion of penguins in southern New Zealand



This chapter has been published as:

Sanne Boessenkool, Jeremy J Austin, Trevor H Worthy, Paul Scofield, Alan Cooper, Philip J Seddon & Jonathan M Waters (2009) Relict or colonizer? Extinction and range expansion of penguins in southern New Zealand. Proceedings of the Royal Society of London B: Biological Sciences 276: 815-821.

Abstract

Recent human expansion into the Pacific initiated a dramatic avian extinction crisis, and surviving taxa are typically interpreted as declining remnants of previously abundant populations. As a case in point, New Zealand's endangered yellow-eyed penguin (*Megadyptes antipodes*) is widely considered to have been more abundant and widespread in the past. In contrast, our genetic and morphological analyses of prehistoric, historic and modern penguin samples reveal that this species expanded its range to the New Zealand mainland only in the last few hundred years. This range expansion was apparently facilitated by the extinction of *M. antipodes*' previously unrecognised sister species following Polynesian settlement in New Zealand. Based on combined genetic and morphological data we describe this new penguin species, the first known to have suffered human-mediated extinction. The range expansion of *M. antipodes* so soon after the extinction of its sister species supports a historic paradigmatic shift in New Zealand Polynesian culture. Additionally, such a dynamic biological response to human predation reveals a surprising and less recognised potential for species to have benefited from the extinction of their ecologically similar sister taxa and highlights the complexity of large-scale extinction events.

Introduction

Colonization of the Pacific — the 'final frontier' of human expansion — has left a trail of vertebrate extinctions readily discernible from archaeological and paleontological data (Steadman & Martin 2003), providing an accessible system for revealing anthropogenic impacts on indigenous biota (Hurles *et al.* 2003). Subsistence hunting by early Polynesians is typically implicated in early extinctions (Worthy 1999; Holdaway & Jacomb 2000), and any surviving taxa are usually interpreted as declining remnants of previously abundant populations. With the advent of ancient DNA techniques we now have a means to test the timing and severity of species and population declines by directly characterizing temporal changes in genetic diversity (Paxinos *et al.* 2002; Shapiro *et al.* 2004; Leonard *et al.* 2007; Valdiosera *et al.* 2008).

In New Zealand, Polynesian expansion southwards (*c.*1280 AD), followed by European colonisation (1769 AD onwards), destroyed much of an indigenous biota that was naïve to terrestrial mammalian predators (Higham *et al.* 1999; Wilmshurst *et al.* 2008). At least 41% of the endemic bird species have become extinct, and of those remaining, 35% are now classified as threatened (Worthy & Holdaway 2002). The endangered yellow-eyed penguin (*Megadyptes antipodes*), also known as *hoiho*, is one of New Zealand's most publicised threatened species and is the focus of extensive conservation effort, including strong community involvement. The species is considered *taonga* (sacred) by the local Māori, is of high economic importance for local tourism industries and has been ecologically well studied over the last decades. The total population of ~7000 individuals breeds on the subantarctic Auckland and Campbell Islands and the southeast coast of the South Island of New Zealand (Marchant & Higgins 1990; McKinlay 2001, figure 2.1). Previous analysis of the fossil records and anecdotal evidence suggest that this penguin was more abundant and widespread in the past and

consequently current management assumes yellow-eyed penguins on the mainland are a declining remnant of the prehistoric population (Worthy 1997; Moore 2001). The presence of penguin bones in archaeological middens from early Polynesian settlers in New Zealand, ancestors of modern Māori, indicates that penguins have been subject to human hunting pressure, but to date this finding has not been considered significant. To test for temporal changes in *M. antipodes* genetic diversity associated with human settlement of New Zealand we assessed mitochondrial DNA variation of prehistoric, historic and modern samples of yellow-eyed penguin. Based on the results of our genetic analysis we further performed detailed morphological comparisons between prehistoric and modern *Megadyptes* bones, which lead us to describe a new penguin species that became extinct only a few hundred years ago and revealed the unsuspected recent range expansion of *M. antipodes*.

Material and methods

DNA extraction and sequencing

Yellow-eyed penguin blood samples were collected in 2005-2007 by wing venipuncture of the brachial vein from six different locations throughout the species' breeding range (N = 15-20 for each location; M5-M8, M10, M11 in figure 2.1). DNA was extracted and purified using 40 µg proteinase K in 5% Chelex (BioRad; Walsh *et al.* 1991). A 813 bp fragment of the first hypervariable region of the mitochondrial control region was amplified using primers: L-Man-CR4 (5'-CTGTGCACTGCTTTATGTACGC-3') and H-Man-CR7 (5'-GTGCATCAGTGTTAAGATGATTCC-3'). PCRs (15 µl) containing 0.5 µM of each primer, 0.8 mM dNTPs, 1.5 mM $MgCl_2$ and 0.75 U *Taq* polymerase (Mango *Taq*, Bioline) were amplified for 2 min at 94°C, 35 cycles of 20 s at 94°C, 20 s at 50°C and 1 min at 72°C, followed by 10 min at 72°C. Purified PCR products were sequenced with H-Man-CR7.

Historic toepad samples were obtained from 55 museum specimens collected between 1840 to 1944 across the breeding range of *Megadyptes antipodes* and currently held in 15 museum collections worldwide (Appendix 9.1). Tissue samples were rehydrated by a 24 h wash in 1 ml 10 mM Tris-HCL (pH 8.0), and DNA was subsequently extracted using the Chargeswitch Forensic DNA Purification Kit (Invitrogen) or the DNeasy Tissue Kit (Qiagen) following manufacturers' instructions. No differences were observed in extraction or amplification success between either of these kits. Two overlapping fragments were amplified using primer pairs L-Man-CR4 (5'-CTGTGCACTGCTTTATGTACGC-3') and H-Man-CR12 (5'-ACAAACGATACCAACCTATGGG-3') (299 bp); and L-Man-CR11 (5'-GAGTAATGGTATGAGGATTAGCTCC-3') and H-Man-CR14 (5'-CGGGTTGCTGATTTCACGTG-3') (287 bp), yielding a total of 402 bp. For some samples a single 444 bp fragment was amplified using primers L-Man-CR4 and H-Man-CR14. Primers H-Man-CR12, L-Man-CR11 and H-Man-CR14 were designed in conserved regions that did not show any polymorphisms in the sequences obtained from modern samples. PCRs (25 µl) containing 0.4-0.8 µM of each primer, 0.8 mM dNTPs, 2.0 mM $MgCl_2$ and 0.5-1.0 U *Taq* polymerase (Mango *Taq*, Bioline) were performed as above with cycles increased to 50. Purified PCR products were sequenced with the same primers used for amplification.

A total of 69 prehistoric *Megadyptes* bones from the South Island and the Auckland Islands, New Zealand, were obtained from museum collections (Appendix 9.2). Morphological descriptions from Worthy (1997) were used for identification of *Megadyptes* bones. All but two bones (NMNZ S.42156.1 and NMNZ S.42156.2) were indirectly dated to 600-1700 AD based on associated archaeological remains (references for radiocarbon dates per sampling site can be found in Appendix 9.2). With the exception of one specimen found at a site dated to 1700 AD (Mapoutahi Pa, see Appendix 9.2), no specimens have been found from the period 1500-1800 AD. Independence of individual bones was achieved by either sampling the same bone type

within a location or by sampling bones from different strata within the archaeological site. Bones were sampled using a hand drill and powdered in a Mikro-Dismembrator S (Sartorius). A total of 50-80 mg of bone powder was decalcified in 2 ml 0.5 M filtered EDTA for 24 hrs. DNA was extracted using the DNeasy Tissue Kit (Qiagen) following manufacturer's instructions with the following modifications: 1) double volumes were used for proteinase K, AL and ATL buffers, and 2) 2-4 µl carrier RNA were added to each sample following proteinase K digestion. Samples were amplified for two overlapping fragments as described above. PCRs (25 µl) containing 2 µl of non-diluted or 1:10 diluted DNA, 0.8 µM of each primer, 1.0 mM dNTPs, 2.0 mM $MgSO_4$, 1 mg/ml BSA/RSA, and 0.5-1.0 U *Taq* polymerase (Platinum *Taq* DNA Polymerase High Fidelity, Invitrogen) were performed with 1 min at 94°C, 50 cycles of 15 s at 94°C, 15 s at 55°C and 30 s at 68°C, followed by 10 min at 68°C. Where necessary, 1 µl of the PCR was used as a template for a second PCR to improve amplification success. PCR products were purified and sequenced with the same primers used for amplification. All sequences are deposited in GenBank (accession numbers FJ391944 - FJ391968).

Authenticity of (pre)historic DNA

Precautions for the analysis of historic and prehistoric DNA were adhered to. Historic sample DNA extractions and PCR set-up were performed inside a UV hood in a laboratory where no contemporary yellow-eyed penguin DNA or any vertebrate PCR products have ever been present. Genetic analyses of prehistoric bone samples were all performed at the Australian Centre for Ancient DNA where extractions and PCR set-up were carried out in a physically isolated, designated ancient DNA laboratory. Contamination was monitored by negative extraction and PCR controls. All historic and prehistoric samples were amplified and sequenced at least twice for both fragments. When conflict was observed among sequences, a third amplification was performed and a majority rule consensus applied (Brotherton *et al.* 2007).

Authenticity of prehistoric sequences was further confirmed by 1) extraction replications, 2) the use of different primers to amplify fragments within the target region and 3) cloning. First, for nine successfully amplified samples re-extractions and amplifications were performed following the same protocols as described above. Secondary extracts were amplified and sequenced twice and sequences were compared to those obtained for primary extracts. All consensus sequences of eight secondary extracts agreed with consensus sequences of primary extracts. One secondary extract could not be amplified successfully. Second, for two samples three overlapping fragments within the 402bp target sequence were amplified using primers L-Man-CR4 (5'-CTGTGCACTGCTTTATGTACGC-3') and H-Man-CR10 (5'-TCGTTTAGTCAATGTAATAGGAGC-3') (201 bp); L-Man-CR11 (5'-GAGTAATGGTATGAGGATTAGCTCC-3') and H-Man-CR12 (5'-ACAAACGATACCAACCTATGGG-3') (142 bp); L-Man-CR13 (5'-GACTAAACCCATAGGTTGGTATCG-3') and H-Man-CR14 (5'-CGGGTTGCTGATTTCACGTG-3') (174 bp). Purified PCR products were sequenced using L-Man-CR4, L-Man-CR11 and L-Man-CR13. Sequences of these fragments all agreed with sequences obtained with the primer pairs described above. Third, for three samples the amplification products obtained using primer pairs L-Man-CR4 with H-Man-CR12 and L-Man-CR11 with H-Man-CR14 (see above) were cloned using the TOPO TA Cloning Kit® with One Shot® Top 10 Competent Cells (Invitrogen) following the manufacturers instructions. Cells were plated on LB-amp agar plates and grown overnight at 37°C. A total of 48 colonies were picked per clone (six clones total) and placed into 96-well plates containing 20 µl 10mM Tris (pH = 8). Cells were lysed by heating to 95°C for 10 min and stored at -20°C. Amplifications were performed on eight colonies per clone. PCRs (25 µl) contained 0.2 µM of M13 forward and reverse primers, 1.0 mM dNTPs and 0.5 U *Taq* polymerase (HotMaster, Eppendorf). Thermocycler conditions were 2 min at 94°C, 50 cycles of 20 s at 94°C, 10 s at 55°C and 45 s at 65°C, followed by 10 min at 65°C. PCR products were purified and sequenced with

the M13 reverse primer. Clone sequences were compared to sequences obtained from the original PCR product. Sequences of clones all agreed with original sequences, or reflected ambiguous bases in the original sequence as expected (e.g. ~50% of the clones showed and A and ~50% a G at the site where the original sequence showed an A/G ambiguity).

Finally, an important observation supporting the authenticity of our sequences is the phylogenetic consistency of our results. Specifically, prehistoric sequences of *Megadyptes antipodes* and *M. waitaha* were consistent with geographic and morphological observations (see Results below). It is highly unlikely that such patterns would result merely from DNA damage.

Genetic analyses

Sequences were aligned using Sequencher (Schneider 1998) and analyses were restricted to the 402 bp region sequenced for all specimens. Applying the AIC criterion of Modeltest (Posada & Crandall 1998), we obtained HKY+I as most appropriate models of evolution for our dataset. Maximum-likelihood (ML) analyses were performed in PAUP* (Swofford 2003). Model parameters were estimated by a heuristic search, with 100 repetitions of stepwise addition. Using the estimated parameters, node support was calculated with 10 000 bootstrap replicates. Bayesian trees were estimated by MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003) in two independent runs, using 20 000 000 generations, sampling every 1000th generation, and discarding 25% as burnin. Convergence diagnostics of Bayesian analyses were explored using Tracer (Rambaut & Drummond 2007) and AWTY (Nylander *et al.* 2008). The topology of the ML and Bayesian trees were very similar and therefore only the Bayesian tree is shown. The shallow divergence within *Megadyptes* in relation to other penguin species made accurate model selection through Modeltest and rooting of the trees difficult, and rooted phylogenetic analyses were therefore only performed using a Neighbour-

Joining algorithm with a Kimura-2 distance parameter. Genealogical relationships among samples were reconstructed using a parsimony-based haplotype network with a 94% parsimony criterion in TCS (Clement *et al.* 2000). Observed genetic divergence (p-distance) was calculated among clades in PAUP* (Swofford 2003). Haplotype and nucleotide diversity indices were determined using DnaSP (Rozas *et al.* 2003).

Morphometric measurements and analyses

Qualitative osteological comparisons were made for coracoid, femur, humerus, tarsometatarsus and tibiotarsus and described using terminology from Baumel and Witmer (1993). Morphometric measurements (to the nearest 0.1 mm) of four different bone types (coracoid, femur, humerus and tarsometatarsus) were obtained from genetically analysed specimens (complete bones only), 26 contemporary skeletons (collected 1970-1990) and an additional 47 single bones from prehistoric sites (specimens listed in Appendix 9.3), using Vernier callipers. It was unknown whether prehistoric specimens represented single or multiple skeletons, and each bone type was therefore analysed separately. Difference in average bone length among modern and prehistoric samples was determined using ANOVA followed by posthoc analyses (Scheffe) in SPSS 16.0. Normality and homoscedasticity assumptions were met and Bonferroni corrections were applied where necessary.

Genetic relationships

We successfully amplified and sequenced DNA from 100 modern, 43 historic and 42 prehistoric samples, yielding a total of 23 distinct haplotypes (Appendix 9.4). Bayesian, maximum likelihood and distance analyses all reveal a previously unrecognised and well-supported genetic split among *Megadyptes* samples. Specifically, all South Island specimens from before 1500 (sampling sites P1-P9), with the exception of three specimens, form a

well-supported distinct genetic group (figure 2.1, figure 2.2). Within this group, a further genetic split is observed between the northern (P1 and P3) and southern (P4-P9) South Island samples. None of these prehistoric South Island haplotypes is found in the historic or modern samples. On the other hand, all prehistoric subantarctic sequences (P10) cluster with the historic and modern yellow-eyed penguins that now inhabit southern New Zealand and the subantarctic (M4-M11). Currently *Megadyptes* penguins are absent from the northern parts of the South Island (i.e. north of M4). The haplotype network clearly visualises the substantial divergence between haplogroups, the relatively close relatedness of haplotypes within each group and the presence of two highly common haplotypes (figure 2.3). Genetic divergence between the two identified *Megadyptes* groups was d = 2.24 - 4.23% and diversity indices were found to be low for both the prehistoric South Island penguins (h = 0.834, π = 0.009) and the group comprising prehistoric subantarctic and modern penguins (h = 0.547 and π = 0.004). Based on the observed unique genetic composition and the consistent morphological distinctness (presented below) of the prehistoric South Island penguins we describe these penguins as a new species.

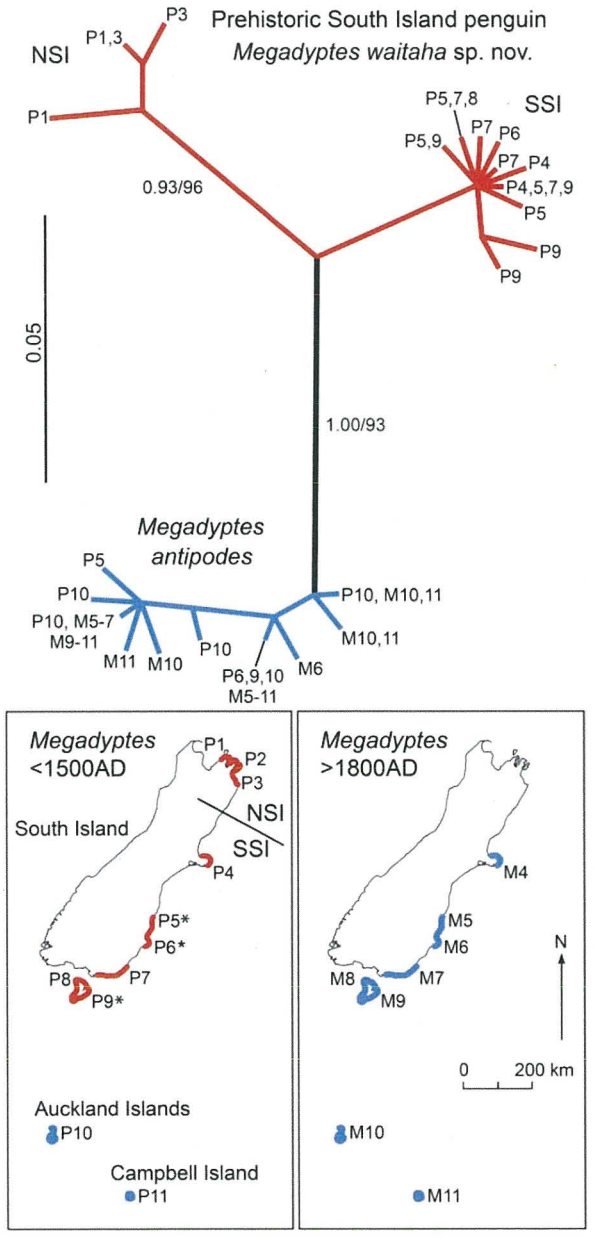


Figure 2.1 Spatiotemporal genetic relationships and distribution of *Megadyptes* penguins. Prehistoric South Island sequences (*M. waitaha*) are shown in red, *M. antipodes* sequences are shown in blue. Numbers on the main branches in the unrooted Bayesian phylogram represent posterior probabilities and ML bootstrap support. Maps show the South Island and subantarctic Campbell and Auckland Islands of New Zealand. Sampling sites are indicated with labels for prehistoric (P1 - P10) and modern (M4 – M11) samples, and prehistoric sites are further split in northern and southern South Island (NSI and SSI respectively). Number of samples possessing each haplotype varied between 1 and 94 (see figure 2.2). It is assumed modern *Megadyptes* inhabited Campbell Island prior to 1700AD, as they do at present, but there is currently no palaeontological evidence to support this.

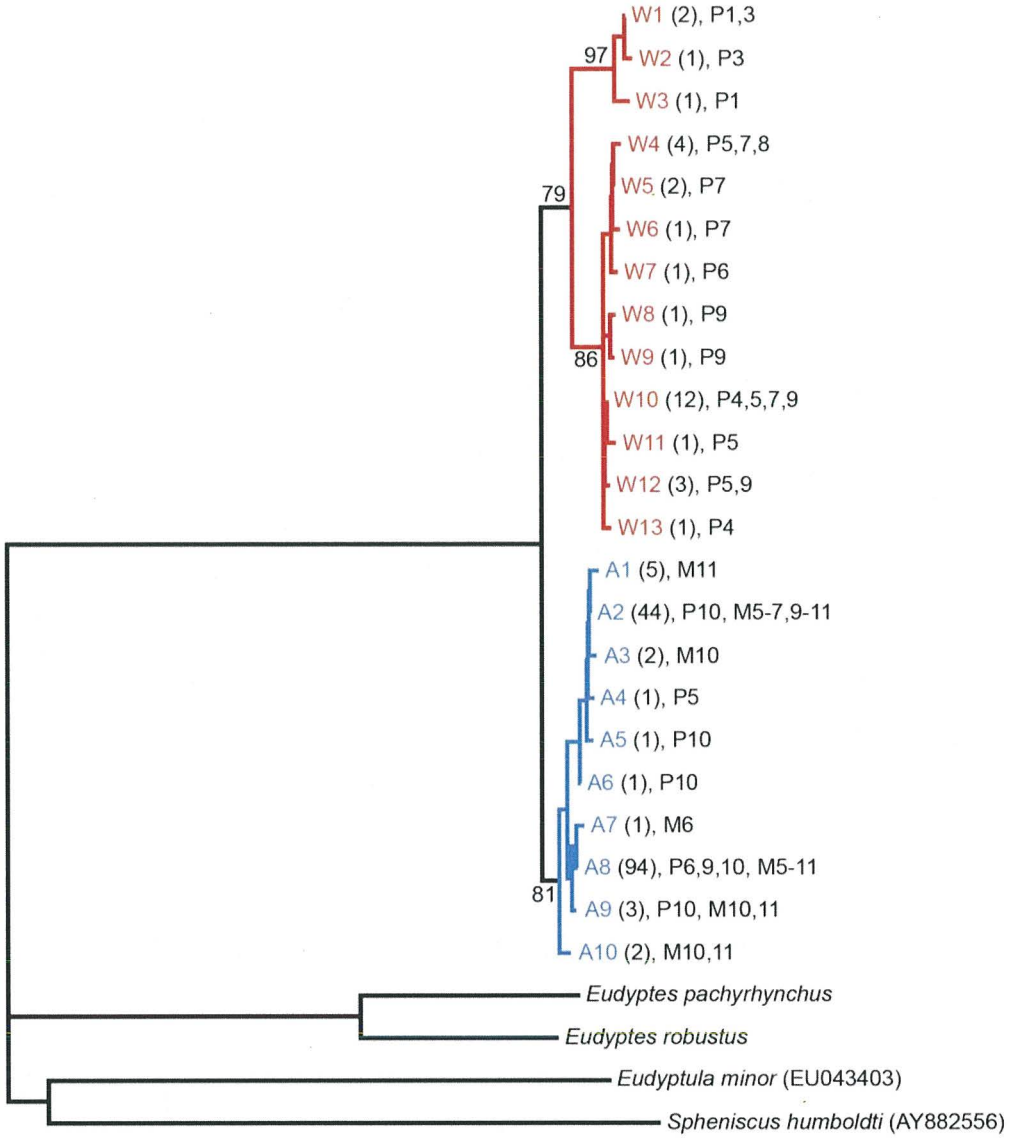


Figure 2.2 Neighbour-Joining phylogeny of the prehistoric South Island penguin *Megadyptes waitaha* (W1-W13, red) and *M. antipodes* (A1-A10, blue). Haplotype labels are followed by sample size (in brackets) and sampling region as indicated in figure 1. The bootstrap support for nodes are based on 1000 pseudo-replicates. Sequences for *Eudyptula minor* and *Spheniscus humboldti* were downloaded from GenBank. The *E. pachyrhynchus* blood sample was kindly provided by A. Paterson and DNA extraction and amplification were performed as described for *Megadyptes* modern blood samples (see Material and Methods). DNA from *E. robustus* was kindly provided by C. Millar (sample obtained from Auckland Museum specimen LB12869), and DNA was extracted and amplified as described for *Megadyptes* museum skin samples (see Material and Methods). GenBank accession numbers for sequences are FJ391944 - FJ391968. Note that eight of the historic *M. antipodes* specimens purportedly of subantarctic origin were in fact collected on the South Island (see Chapter 6).

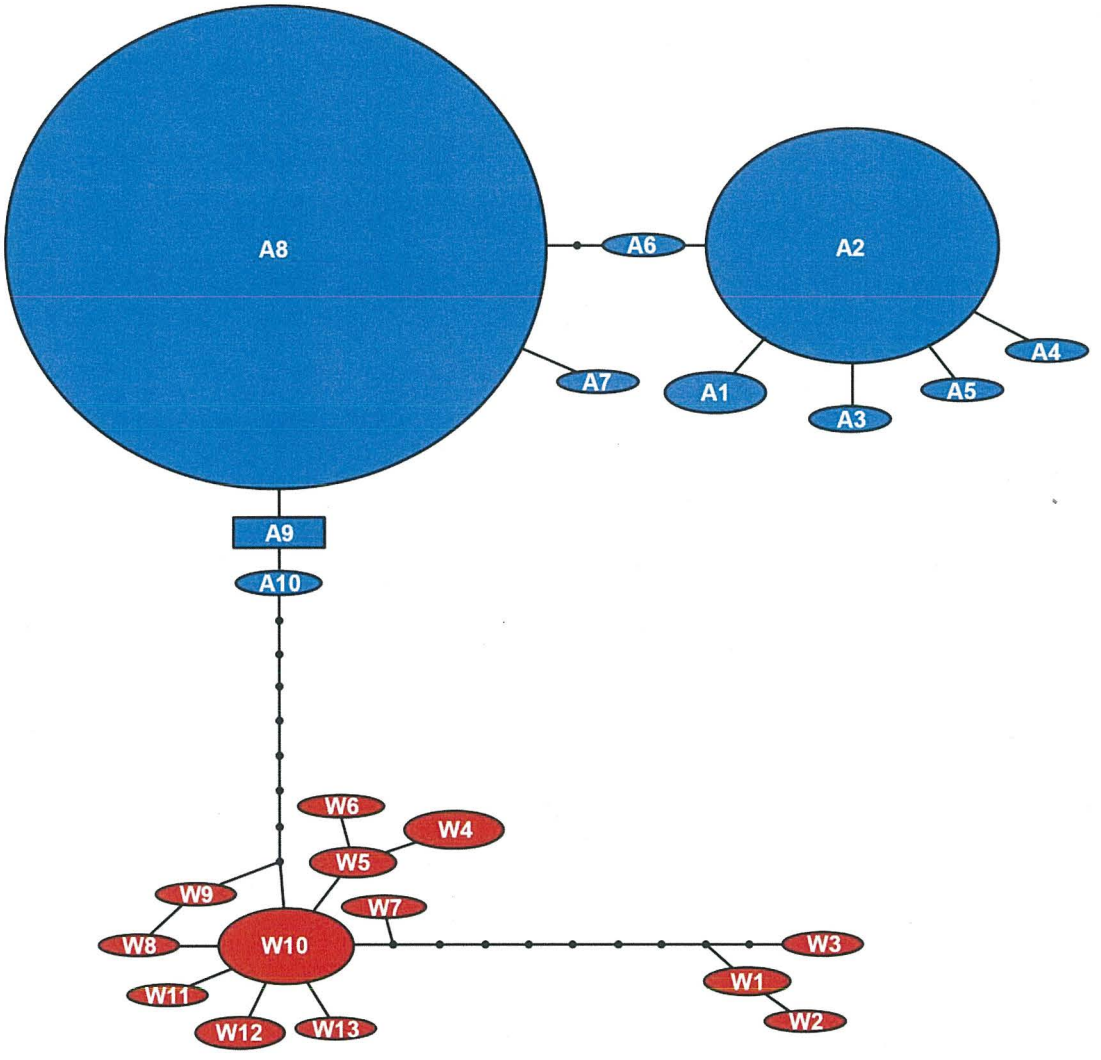


Figure 2.3 Haplotype network of *Megadyptes antipodes* (A1-A10, blue) and the prehistoric South Island penguin *M. waitaha* (W1-W13, red) sequences estimated using a 94% parsimony criterion. Haplotype frequency is indicated by symbol area. The putative ancestral haplotype is indicated by a rectangle. Black dots represent hypothetical intermediate haplotypes not detected in the current study.

Systematic palaeontology

SPHENISCIFORMES Sharpe, 1891

SPHENISCIDAE Bonaparte, 1831

MEGADYPTES Milne-Edwards, 1880

Megadyptes antipodes (Hombron & Jacquinot, 1841)

Megadyptes waitaha sp. nov.

Etymology

From Waitaha (Māori): the first Polynesian tribe that occupied much of the South Island, New Zealand, before they were displaced by Ngāti Māmoe, who in turn were later dominated by Ngāi Tahu.

Holotype

Canterbury Museum, CM AV13269 (figure 2.4 and 2.5). Left femur, complete. Measurements of Holotype: 77.1 mm length, 8.5 mm shaft width, 18.6 mm proximal width, 16.0 mm distal width.

Locality and horizon

CM AV13269 was collected from the dunes along Lake Grassmere, Marfells Beach, Marlborough on the South Island (41°43'21'S, 174°11'42'E, Site P3 in figure 2.1), by J. Britton and R. Britton in 1954. Material from these dunes has been widely studied and has been dated to the late Holocene, between 600-1500 years before present (Worthy 1998; Duncan *et al.* 2002).

Paratypes

CM AV11995, right femur, complete. CM AV16258Z, right femur, complete. CM AV34941, left femur, complete.

Referred material

All specimens from the northern South Island to Codfish and Stewart Island, just south of the South Island, that are listed in Appendix 9.2 and Appendix 9.3.

Diagnosis

Megadyptes waitaha bones are more slender and smaller than those of *M. antipodes* and differ for a range of characters described below. *M. waitaha* further forms a distinct genetic group based on hypervariable region I (HVI) of the mitochondrial control region. Genetic divergence from *M. antipodes* in HV1 mtDNA is 2.24 – 4.23% with the following fixed character states (character for *M. waitaha*/character for *M. antipodes*, position corresponding to *Eudyptes chrysocome* mitochondrial genome sequence, GenBank accession number AP009189): T/C (15829), A/G (15855), G/T (15910), T/G (16006) and A/G (16072).

Description and comparisons

M. waitaha bones (figure 2.4) are distinguished from *M. antipodes* as follows: (i) femur: lacks a prominent vascular foramen in the *fossa poplitea*; *linea intermuscular caudalis* more pronounced; *crista trochanteris* shorter and narrower; *impressiones obturatoriae* squarer and more pronounced; *condylus medialis* less robust, (ii) tibiotarsus: *crista cnemialis* more pointed, (iii) tarsometatarsus: *crista medialis hypotarsi* more flattened; *cotyla*

lateralis laterally less prominent; *crista lateralis hypotarsi* less pronounced; medial *foramina vacularia proximalia* more heavily occluded plantaroproximally; medial margin more concave giving whole bone more slender appearance, (iv) coracoid: *facies sternalis* proportionally narrower; dorsal *facies articularis sternalis* less robust; medial process above medial angle less robust; *cotyla scapularis* rounder and smaller; *processus procoracoideus* smaller and less ventrally curved; *processus glenoidalis* more robust, narrower coracohumeral surface (neck) between *processus glenoidalis* and *processus acrocoracoideus*; *foramina procorocoideus* absolutely and relatively larger, (v) humerus: *impressio coracobrachialis* proportionally deeper, especially proximally; ventrally located secondary fossa within *fossa pneumotricipitalis* deeper and orientated more anterior-ventrally, *sulcus transversus* dorsal pit relatively deeper, ventral bit shallower; *sulcus tendinis musculus humerotricipitalis* (sesamoid groove) deeper, and the proximal trochlear process caudally bounding the humerotricipital sulcus is more pointed and bent ventrally near tip.

Bones from *M. waitaha* are significantly smaller than bones from *M. antipodes* (figure 2.5, figure 2.6, Table 2.1). There is, however, no size differentiation between *M. waitaha* bones from the northern and the southern South Island of New Zealand (figure 2.6). The similar size of northern and southern populations of *M. waitaha* occurred over a geographic range greatly exceeding the distance from the South Island to the subantarctic islands and thus the geographical distance between the *M. waitaha* and *M. antipodes* populations in prehistoric times.

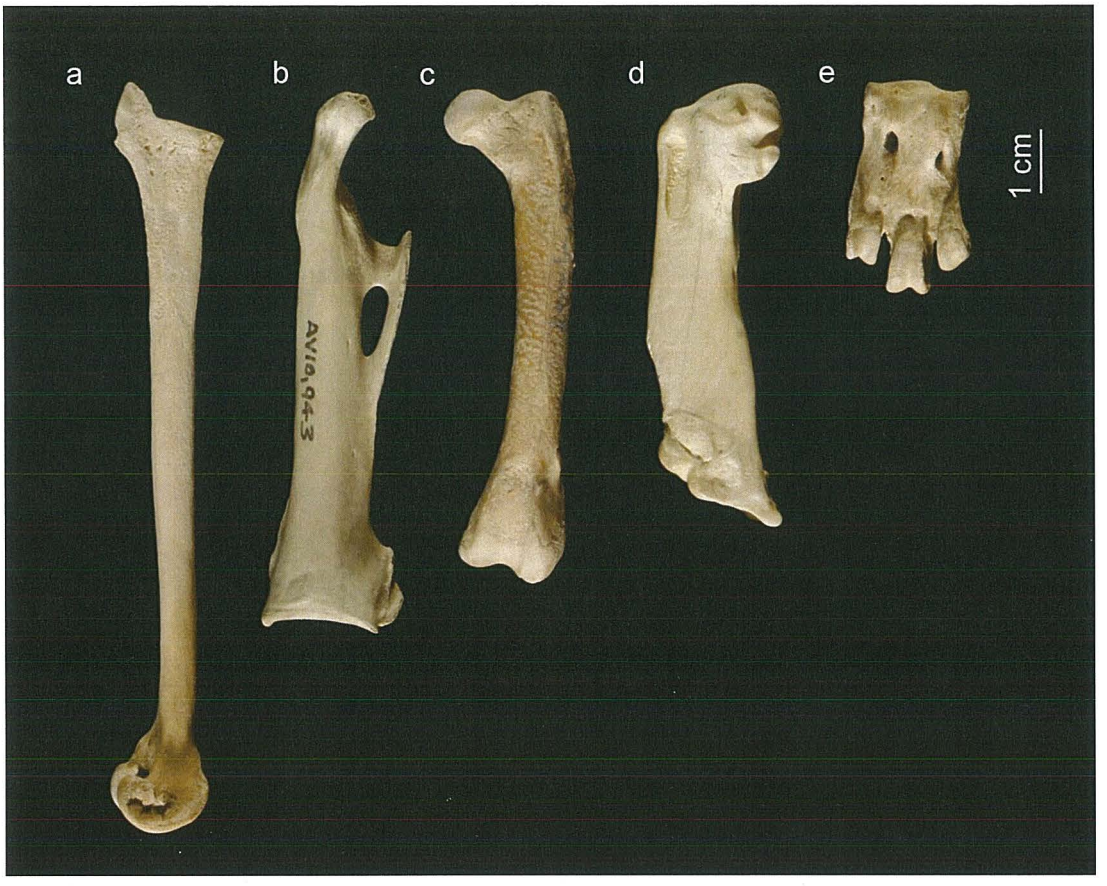


Figure 2.4 Bones of *Megadyptes waitaha*. Right tibiotarsus CM AV14316 in (a) medial view, (b) right coracoid CM AV10943 in ventral view, (c) left femur CM AV13269 in dorsal view, (d) right humerus CM AV9654B in cranial view and (e) left tarsometatarsus CM AV10944 in plantar view.

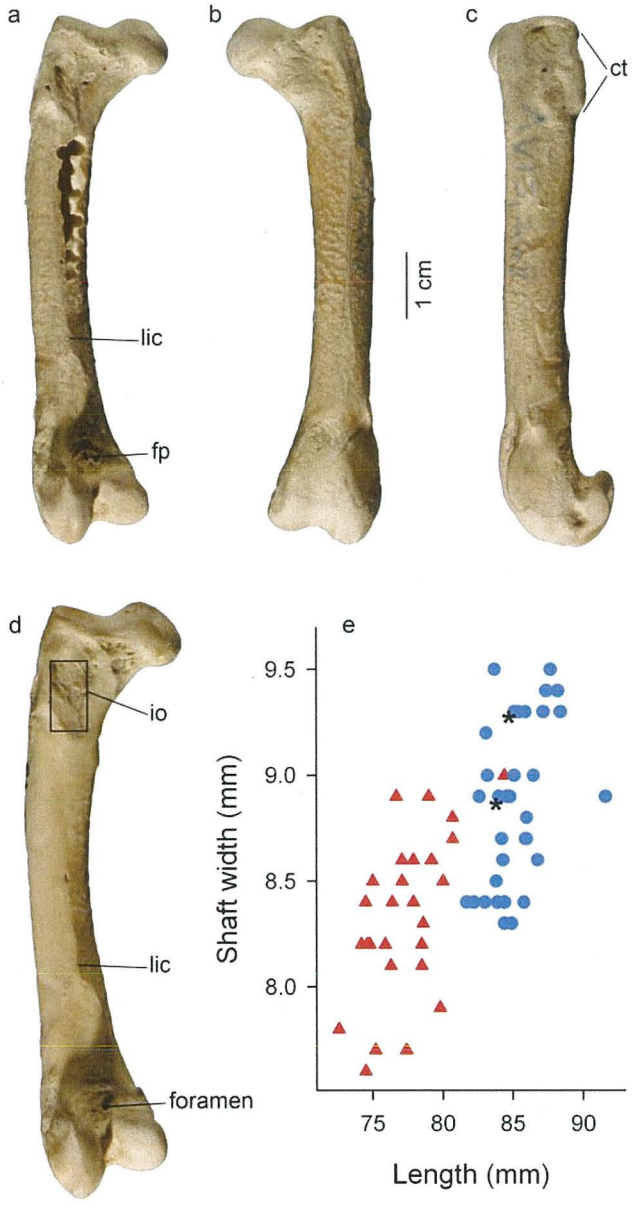


Figure 2.5 Holotype left femur of *Megadyptes waitaha* (a-c), *M. antipodes* left femur (d) and a plot showing the size differences of *M. waitaha* and *M. antipodes* femora. (a) Ventral view; (b) dorsal view; (c) lateral view. (d) Ventral view of *M. antipodes* left femur (CM AV32415). (e) Plot showing the size difference of *M. waitaha* (red triangles) and *M. antipodes* (blue circles) femora. Asterisks indicate two of the three prehistoric South Island samples (i.e. the two femora from P6 and P9 in figure 2.1) that cluster genetically with *M. antipodes*. The data revealed support the consistent genetic and morphological differences between *M. antipodes* and *M. waitaha*. Anatomical abbreviations: ct, crista trochanteris; fp, fossa poplitea; io, impressiones obturatoriae; lic, linea intermuscular candalis. The ventral view of *M. waitaha* femur (a) shows several drill holes resulting from the sampling of the bone.

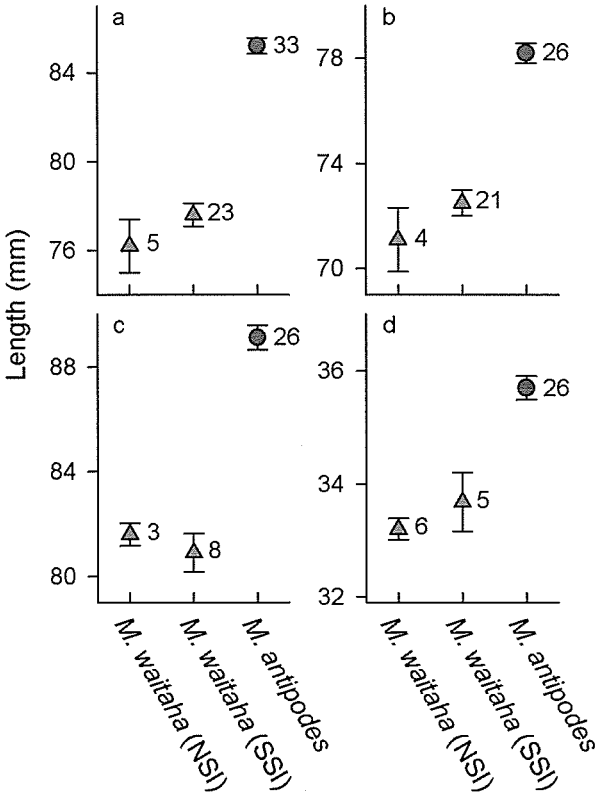


Figure 2.6 Average length of *Megadyptes waitaha* (red triangles) and *M. antipodes* (blue circles) (a) femur, (b) humerus, (c) coracoid and (d) tarsometatarsus. *M. waitaha* bones are divided into southern and northern South Island (see figure 2.1). Error bars are standard error intervals; numbers next to symbols represent sample sizes (n). Four separate single factor ANOVA showed significant differences among the groups (femur: $F_{2, 58} = 91.2$, humerus: $F_{2, 48} = 52.2$, coracoid: $F_{2, 34} = 48.6$, tarsometatarsus: $F_{2, 34} = 18.9$, all p-values ≤ 0.0001). Post hoc analysis (Scheffe) revealed significant differences between *M. antipodes* and both southern and northern *M. waitaha* (all $p \leq 0.0015$), but not between southern and northern *M. waitaha*.

Table 2.1 Morphometric measurements (in mm) of *Megadyptes antipodes* and *M. waitaha bones*. All measured bones are listed in Appendix 9.3. Not all measurements could be taken for each bone due to damage of proximal and/or distal ends of some bones.

		M. antipodes			*M. waitaha*		
Bone type	Measurement	*N*	Mean	s.e.	*N*	Mean	s.e.
Femur	Length	33	85.2	0.4	28	77.3	0.5
	Shaft width	33	8.9	0.1	28	8.3	0.1
	Distal width	32	19.7	0.2	22	17.8	0.2
	Proximal width	31	18.4	0.2	18	16.4	0.2
Humerus	Length	26	78.3	0.4	25	72.2	0.5
	Width	26	13.5	0.1	25	12.6	0.1
Coracoid	Length	26	89.1	0.5	11	81.1	0.5
Tarsometatarsus	Length	26	35.7	0.2	11	33.4	0.3
	Distal width	26	18.8	0.1	10	17.3	0.3
	Proximal width	26	23.1	0.2	11	21.6	0.4

Discussion

Human-mediated extinction of a new penguin species

Genetic and morphological analyses reveal a previously unrecognised penguin species in the *Megadyptes* genus. None of the haplotypes of this species are found in any of the historic or modern samples analysed, indicating that *M. waitaha* no longer survives. The presence of its bones in archaeological context implies that its extinction was likely caused by overexploitation (Jones *et al.* 2008). This finding is consistent with the fact that large-bodied species were particularly vulnerable to extinction by hunting in prehistoric New Zealand (Holdaway & Jacomb 2000; Duncan & Blackburn 2004). Indeed, the marked transition from big game (including large penguins) to small game and fish observed in stratified middens reflects the population decline of the larger species within just decades of human settlement (Nagaoka 2001). This previously described "blitzkrieg" was obviously not only directed against the well-known moa (Diamond 2000), but also other species such as the overlooked penguin we describe here. It is thus likely that *M. waitaha* went extinct within a few hundred years of human settlement in New Zealand. The recognition of two species in *Megadyptes* reveals an original taxon distribution similar to that of *Eudyptes*, which displays noticeable speciation within the genus (Jouventin *et al.* 2006) including different species inhabiting the South and subantarctic islands of New Zealand.

The phylogeographic split between northern and southern South Island samples of the extinct *M. waitaha* is concordant with biogeographic disjunctions observed around an upwelling zone at latitude 42°S in a number of coastal invertebrate taxa in New Zealand (e.g. Apte & Gardner 2002; Ayers & Waters 2005). This upwelling and associated longitudinal change in currents and water temperature may have also presented a barrier to

geneflow for *M. waitaha*. Currently, *M. antipodes* does not breed above 43°S, although occasional vagrants are found as far north as New Zealand's North Island (Marchant & Higgins 1990).

Recent range expansion of the yellow-eyed penguin

Our findings demonstrate that yellow-eyed penguins are not a declining remnant of a previous abundant population, but instead went through a recent range expansion following the extirpation of *M. waitaha*. Therefore, it seems almost certain that the entire extant yellow-eyed penguin population on the South Island is derived from a subantarctic stock. Only three of the prehistoric penguin specimens on the South Island were identified genetically and morphologically as *M. antipodes*. These specimens probably represent non-breeding vagrants from the subantarctic, as now commonly occurs with *Eudyptes* species. The observation of these inferred vagrants apparently attests to the ability of *M. antipodes* to disperse to the South Island and thus provides a clear mechanism for the suggested range expansion.

The rapid replacement of *M. waitaha* by *M. antipodes* suggests that competition between the two species previously prevented *M. antipodes* from expanding northwards, especially considering *M. antipodes* vagrants were present on the South Island even before the extinction of *M. waitaha*. The successful expansion of *M. antipodes* into the South Island, prior to the increase of European settlers and their commensals in the late 1800s and soon after the anthropogenic extinction of *M. waitaha*, may imply that a paradigmatic shift in Māori culture took place. Indeed, it has been suggested that cultural change including new forms of resource monitoring and conservation in Māori culture may have developed from the early 16th century, possibly forming the basis of modern Māori environmental management (Anderson 2002). Alternatively, the archaeological record shows a marked lack of coastal South Island village sites from the early 16th century, in the period following the extinction of big game, suggesting a local

temporary reduction of the human population (Anderson & Smith 1996). Environmental changes such as the severe decline in populations of sea lions (*Phocarctos hookeri*), known predators of penguins, might also have facilitated *M. antipodes* colonising the South Island (Childerhouse & Gales 1998; Lalas *et al.* 2007). We suggest that a similar extinction-colonisation process such as observed in *Megadyptes* might also explain the previously reported arrival of an Australian *Eudyptula minor* lineage in southern New Zealand (Banks *et al.* 2002; Overeem *et al.* 2008).

Ancient DNA analyses are proving to be an extremely valuable tool in wildlife conservation, providing an ability to directly characterise temporal changes in population sizes and connectivity (reviewed in Leonard 2008). The yellow-eyed penguin provides an unusual case in which prehistoric data support a recent range expansion, instead of the previously assumed decline in numbers. Although the conservation status of South Island *M. antipodes* might be questioned on the basis of these results, the species remains in a vulnerable state with a low total population size, a very confined breeding range, and ongoing threats from the marine and terrestrial environment (Birdlife International 2008). Although the observed range expansion provides evidence of this species' ability to colonise new habitats, the impact of European settlement such as the introduction of predatory mammals in New Zealand and surrounding islands might preclude any additional range expansion of *M. antipodes*. As such, the ongoing security of the species would seem to depend largely on the continued health of subantarctic populations. The New Zealand Department of Conservation's existing policy focuses on the security of a species as a whole, rather than the detailed history of a particular population. Overall, therefore, the yellow-eyed penguin's high conservation status should remain unaffected by our findings.

Complexity of large-scale extinction events

Our study reveals a new level of biogeographic and ecological complexity potentially associated with large-scale extinction events that afflicted, for example, the Pacific prehistoric avifauna and North American Pleistocene megafauna. Whereas conventional wisdom suggests that surviving species — like their extinct counterparts — suffered major genetic and ecological declines (Hofreiter 2007), we propose that in some instances native species benefited from the extinction of their ecologically similar sister taxa. For example, we suggest that this extinction-expansion interaction might have had a particularly strong influence on seabird distributions: as numerous colonies went extinct (Steadman 1995), newly vacated habitats would have facilitated rapid range expansion in this highly mobile group of species, e.g. as in *Pterodroma nigripennis* (Worthy & Holdaway 2002). Such dynamic anthropogenic processes may turn out to be far more common and important than previously understood.

Acknowledgements

We are very grateful to the Auckland Museum, American Natural History Museum, Australian Museum, Canterbury Museum, Museum of Comparative Zoology, Natural History Museum Geneva, Natural History Museum Tring, Natural History Museum Paris, Museum of New Zealand Te Papa Tongarewa, Natural History Museum Vienna, Swedish Museum of Natural History, Otago Museum, South Australian Museum, Smithsonian Institution, Museum für Naturkunde Berlin, I. Smith (Department of Archaeology, University of Otago), C. Lalas, C. Millar and A. Paterson for supplying material and the New Zealand Department of Conservation for sample collection. Thanks to Ken Miller and Bastiaan Star for help with the figures. This research was supported by the Department of Zoology, University of Otago, including PBRF Research Enhancement Grants to PJS and JMW. SB received a New Zealand Federation for Graduate Women Brenda Shore Award and a New

Zealand Educated Postgraduate Study Abroad Award. Samples were collected under Department of Conservation permits SO-17933-FAU and OT-19097-RES and University of Otago Animal Ethics Approval 69/06.

Chapter 3:

Isolation and characterization of microsatellite loci from the yellow-eyed penguin (*Megadyptes antipodes*)

This chapter has been published as:

Sanne Boessenkool, Tania M King, Philip J Seddon & Jonathan M Waters (2008) Isolation and characterization of microsatellite loci from the yellow-eyed penguin (*Megadyptes antipodes*). Molecular Ecology Resources 8: 1043-1045.

Abstract

Twelve microsatellite loci were isolated and characterized in the endangered yellow-eyed penguin (*Megadyptes antipodes*) using enriched genomic libraries. Polymorphic loci revealed 2-8 alleles per locus and observed heterozygosity ranged from 0.21 to 0.77. These loci will be suitable for assessing current and historical patterns of genetic variability in yellow-eyed penguins.

Introduction

The yellow-eyed penguin (*Megadyptes antipodes*) is an endangered species endemic to New Zealand and it's subantarctic waters (Birdlife International 2008). The total population of this penguin is estimated to comprise fewer than 7000 individuals distributed around the southeast coast of New Zealand's South Island, Stewart Island and the subantarctic Auckland and Campbell Islands (McKinlay 2001). Habitat loss, predation by introduced mammals, variation in food supply and disease have all been identified as factors contributing to the species' precarious state (Darby & Seddon 1990; McKinlay 2001). In addition to its value for ecotourism, *M. antipodes* has local cultural significance in New Zealand, and consequently is the subject of extensive conservation efforts with strong community involvement. Here we describe twelve microsatellite loci developed to examine migration patterns, identify management subunits and assess historical demography of yellow-eyed penguins.

Material and Methods

Two enriched genomic libraries of *M. antipodes* were constructed using a modified version of the protocol described by Perrin and Roy (2000). Genomic DNA was isolated using a DNeasy Kit (Qiagen) and digested with the restriction enzyme *Mbo*I (Promega). Digested DNA was ligated into a pUC19 cloning vector, DNA inserts were amplified using M13 primers and subsequently annealed to biotinylated $(GT)_{12}$ and $(GA)_{12}$ probes. Microsatellite-containing products were selectively isolated using streptavidin magnetic particles (Roche Applied Science) after which the microsatellite-containing fragments were digested with MboI, ligated into a pUC19 cloning vector and used to transform One Shot® Top 10 competent cells (Invitrogen). A total of 2976 individual clones were transferred to

Biodyne B nylon membranes (Pierce) and probed with [γ^{32}P]ATP. Subsequently, 159 positive clones were amplified and sequenced with M13 primers using an ABI 3730 Genetic Analyser (Applied Biosystems) with the BigDye Cycle Sequencing Kit (Applied Biosystems).

We detected a low frequency of positive clones containing long repeats (>8) in the first two libraries screened, therefore a subsequent library was constructed using the protocol described by Glenn and Schable (2005). DNA was digested with RsaI and ligated to the SuperSNX linkers. Biotinylated $(GA)_{12}$, $(GT)_{12}$, $(AAC)_6$, $(AAG)_8$, $(ACT)_{12}$ and $(ATC)_8$ probes were hybridized to the linker-ligated DNA and microsatellite-containing fragments were retained using streptavidin magnetic particles (Roche Applied Science). We used the proofreading *Pwo* Superyield DNA polymerase (Roche Applied Science) for amplification of the microsatellite-enriched fragments. PCR products were A-tailed as described in the pGEM®-T and pGEM®-T Easy Vector Systems manual (Promega) and subsequently ligated into plasmids and transformed using the TOPO TA Cloning® Kit containing pCR®2.1-TOPO® with One Shot® TOP 10 competent cells (Invitrogen). A total of 1920 individual clones were transferred to Biodyne B nylon membranes (Pierce) and probed with [γ^{32}P]ATP-labelled $(GA)_{12}$ and $(GT)_{12}$, or $(AAC)_6$, $(AAG)_8$, $(ACT)_{12}$ and $(ATC)_8$ repeats to identify repeat-containing clones. Ninety-seven positive clones were amplified and sequenced as described above.

Thirty-eight primer pairs were designed from the first two libraries and an additional twenty pairs were designed from the third library. Amplification of primers was tested by screening DNA from twelve yellow-eyed penguin individuals. Blood samples (~0.05 ml) of these twelve individuals had been collected from the brachial vein and stored in 1 ml Queens lysis buffer (Seutin *et al.* 1991). Twenty µl of blood in buffer was used to extract and purify DNA using 40 µg proteinase K in 5% Chelex (Biorad; Walsh *et al.* 1991). PCR reactions (5 µl) contained 1 µl DNA (5-10 ng), 0.5 µM of each primer, 0.25 U *Taq* DNA polymerase (Mango *Taq*, Bioline), 1× *Taq* buffer, 0.8 µM dNTP and

1.5 mM $MgCl_2$. Betaine and DMSO (1.1 M and 2% respectively) were added to the PCR mixture where necessary (Table 3.1). PCR amplification was performed in a Mastercycler *ep*Gradient S (Eppendorf) with the following profile: 2 min at 94°C, 35 cycles of 15 s at 96°C, 15 s at 46-54°C and 30 s at 72°C, followed by a 4 min final extension at 72°C. DNA fragments were resolved on 7-9% vertical, non-denaturing polyacrylamide gels and visualised with 0.05× SYBR Green I (Invitrogen).

Results and Discussion

Twelve loci were found to be polymorphic and subsequently genotyped in 43 individuals from Campbell Island. Genetic diversity based on number of alleles, observed and expected heterozygosities, and deviations from Hardy-Weinberg were calculated in GENEPOP v. 4.0 (Rousset 2008). Loci possessed 2-8 alleles, with observed and expected heterozygosities ranging from 0.21 to 0.77 and 0.19 to 0.76, respectively (Table 3.1). Locus Man47 showed significant departures from Hardy-Weinberg expectations (p = 0.03) due to a deficiency of heterozygotes. This deficit was apparently not caused by null alleles as all individuals amplified at this locus. Linkage disequilibrium among loci was also calculated in GENEPOP v. 4.0 (Rousset 2008). Significance levels were adjusted for multiple pairwise comparisons using the sequential Bonferroni correction (Rice 1989), after which no significant linkage disequilibrium was detected between any pair of loci. These markers are currently being used to investigate spatial and temporal patterns of genetic variation in yellow-eyed penguins.

Acknowledgements

We would like to thank Jon Potter (Auckland Zoo), Matt Charteris and Annabel Gosset (Department of Conservation) for collecting samples on Campbell Island. This research was supported by the University of Otago and

by the Japan Penguin Fund. Cloning work was completed under the Environmental Risk Management Authority approval number GMO05/UO028. Samples were collected under Department of Conservation permits SO-17933-FAU and OT-19097-RES and University of Otago Animal Ethics Approval 69/06.

Table 3.1 Characteristics of twelve microsatellite loci developed for *Megadyptes antipodes*. Locus name, repeat motif of cloned allele, primer sequences, allele size range (bp), number of alleles (*A*), observed and expected heterozygosity (H_O/H_E) and Genbank Accession number are reported. All loci were tested on 43 individuals.

Locus	Repeat motif	Primer sequences (5'-3')	Allele size (bp)	*A*	T_a (°C)	H_O/H_E	Accession no.
Mano3[1,*]	$(TG)_{16}...(TG)_4$	**F:** GCCTGAGAGACCCGTGTG	116-120	2	50	0.21/0.19	EU267109
		R: CTCCCCAGTTGCCTCCTG					
Mano8[1]	$G_7A(CT)_3(AT)_2(GT)_9$	**F:** CCTGTCTTCTATTAAACCCTC	116-126	3	46	0.33/0.33	EU267110
		R: CCACATTTGCACCAGTTG					
Man13[1]	$(GT)_{10}$	**F:** AACACATTTGACAGCCTG	122-130	3	48	0.40/0.34	EU267111
		R: GTTATTCCAACACCAAGC					
Man21[1,*]	$G_5(GT)_9$	**F:** TACTGGTAGCATGGGGTG	128-138	3	50	0.51/0.53	EU267112
		R: CACTGAAAGATGACAACGG					
Man22[1]	$(AC)_{13}$	**F:** TTTCCACTTGAGAGTGTATG	126-138	3	50	0.47/0.46	EU267113
		R: CAAACAGAAAGGATTTGTG					
Man27[1,*]	$(GA)_3CA(GT)_3GA(GT)_4G(CA)_2(TG)_8CA(TG)_2CG(CA)_2(TG)_2A(GT)_3...(CA)_2...(GCA)_3$	**F:** GATCCTGAGAAGAGAGACAG	136-150	2	48	0.42/0.38	EU267114
		R: GGCTGTTCATTTTGTCAC					
Man39[2]	$(GT)_{19}$	**F:** GATCTTTCCAGAGACCTC	137-147	5	52	0.56/0.49	EU267115
		R: ACCCTGTGAGTATGAACC					

Table 3.1 continued

Locus	Repeat motif	Primer sequences (5'-3')	Allele size (bp)	*A*	T_a (°C)	H_O/H_E	Accession no.
Man47[3]	$(TGA)_{20}$	**F:** ATACCTCCAGAATGGCTG	123-135	4	48	0.67/0.70	EU267116
		R: CACTAAGGGTGACCAAGG					
Man50[3]	$(GTT)_{12}(ATT)_2$	**F:** CCTCCACTTAGTTTTGCC	103-112	2	48	0.44/0.37	EU267117
		R: TGGAAGCATAACCATAGC					
Man51[3]	$(TTC)_{22}$	**F:** CAGAGATATTGACTCTGACATC	134-167	8	54	0.77/0.76	EU267118
		R: CCTATCACACAGAAACACTG					
Man54[3]	$(AAC)_{10}$	**F:** GTTTCCTATTTCAGTCTGG	136-139	2	48	0.30/0.35	EU267119
		R: TTGTGCTTTCAGTGTTGG					
Man55[3]	$(TTG)_{14}$	**F:** TTGAACTAGCAAGCAGTGTAG	152-158	2	48	0.37/0.46	EU267120
		R: AAGGGCATTTCCATTCTG					

[1,2,3]Refers to first, second and third library respectively

*Betaine and DMSO added to PCR mix

Chapter 4:

Multilocus assignment analyses reveal multiple units and rare migration events in the recently expanded yellow-eyed penguin (*Megadyptes antipodes*)

This chapter has been published as:

Sanne Boessenkool, Bastiaan Star, Jonathan M Waters & Philip J Seddon (2009) Multilocus assignment analyses reveal multiple units and rare migration events in the recently expanded yellow-eyed penguin (*Megadyptes antipodes*). Molecular Ecology 18: 2390-2400.

Abstract

The identification of demographically independent populations and the recognition of management units has been greatly facilitated by the continuing advances in genetic tools. Managements units now play a key role in short term conservation management programs of declining species, but their importance in expanding populations receives comparatively little attention. The endangered yellow-eyed penguin (*Megadyptes antipodes*) expanded its range from the subantarctic to New Zealand's South Island a few hundred years ago and this new population now represents almost half of the species' total census size. This dramatic expansion attests to *M. antipodes*' high dispersal abilities and suggests the species is likely to constitute a single demographic population. Here we test this hypothesis of panmixia by investigating genetic differentiation and levels of gene flow among penguin breeding areas using 12 autosomal microsatellite loci along with mitochondrial control region sequence analyses for 350 individuals. Contrary to our hypothesis, however, the analyses reveal two genetically and geographically distinct assemblages: South Island versus subantarctic populations. Using assignment tests we recognize just two first-generation migrants between these populations (corresponding to a migration rate of <2%), indicating that ongoing levels of long-distance migration are low. Furthermore, the South Island population has low genetic variability compared to the subantarctic population. These results suggest that the South Island population was founded by only a small number of individuals, and that subsequent levels of gene flow have remained low. The demographic independence of the two populations warrants their designation as distinct management units and conservation efforts should be adjusted accordingly to protect both populations.

Introduction

Genetic tools have revolutionised our ability to identify populations and assess gene flow in species where collection of robust ecological and demographic data is challenging due, for example, to the inaccessibility of breeding sites, or the secretive behaviour or rarity of the species. The accurate identification of populations plays a key role in conservation, where populations represent focal points for determining appropriate scales for management (Cegelski *et al.* 2003; Waples & Gaggiotti 2006). Populations may be recognized within either an ecological or an evolutionary paradigm (Waples & Gaggiotti 2006). The latter paradigm defines populations as groups of individuals that are connected through demographic cohesion (Waples & Gaggiotti 2006) and provides a suitable framework for the identification of management units (MUs) for conservation. The recognition of MUs, defined as demographically independent populations in which population dynamics are primarily dependent on local birth and death rates rather than immigration, is particularly important for short-term (e.g. <20 years) wildlife management programs (Palsbøll *et al.* 2007). While the importance of defining MUs is widely accepted for species that have suffered recent declines, less attention has been paid to the potential importance of MUs for expanding taxa.

The iconic yellow-eyed penguin (*Megadyptes antipodes*) – an endangered New Zealand endemic – was until recently considered a declining remnant of a once widespread and abundant population. Genetic and morphological research, however, suggests that *M. antipodes* underwent a dramatic range expansion during the last few hundred years (Boessenkool *et al.* 2009a, Chapter 2). Indeed, up to ~1500 AD, yellow-eyed penguins were apparently restricted to the subantarctic Auckland and Campbell Islands (figure 4.1), and only after the anthropogenic demise of the endemic mainland *M. waitaha* is *M. antipodes* inferred to have expanded north to mainland New Zealand

(Boessenkool *et al.* 2009a, Chapter 2). The latter species is now an important and well-promoted element of the country's wildlife tourism industry. This penguin currently breeds along the southeast coast of the South Island (i.e. mainland New Zealand), around Stewart Island, and on the subantarctic Auckland and Campbell Islands (figure 4.1). Coastal breeding sites are patchily distributed, however, and consist of loose aggregations of nests rather than true colonies (Seddon & Davis 1989). The species has been classified as endangered based on its confined breeding range, destruction of local habitat and extreme fluctuations in numbers during recent decades (Birdlife International 2008). This classification, in combination with the species' high profile, has led to intense conservation efforts by governmental and local community agencies. Such conservation measures typically involve predator trapping, revegetation of coastal habitat, regular monitoring of nests throughout the breeding areas on the South Island, and – more recently – regulations restricting access to beaches have been implemented in an attempt to reduce the impact of tourism (McKinlay 2001; Ellenberg *et al.* 2007).

Although M. antipodes' inferred recent expansion clearly indicates strong potential for long distance dispersal, it is unclear whether this was a "one-off" colonisation event in response to anthropogenic extinction, or whether such oceanic dispersal is an ongoing ecological phenomenon in this species. Considering the penguin's putative recent arrival in South Island (Boessenkool et al. 2009a, Chapter 2), low fecundity, and slow rate of reproduction (Darby & Seddon 1990), it is noteworthy that South Island M. antipodes nowadays constitutes a sizeable proportion (40%; ~800 nests) of the total yellow-eyed penguin population (McKinlay 2001). This substantial increase of M. antipodes on the South Island might result from ongoing immigration from the subantarctic, but three lines of evidence suggest otherwise. First, of the ~10 000 yellow-eyed penguins that have been banded on the South Island in the past three decades, along with ~550 on the Campbell and Auckland Islands, only one individual is known to have crossed the ~500 km stretch of ocean separating these regions. This migrating bird

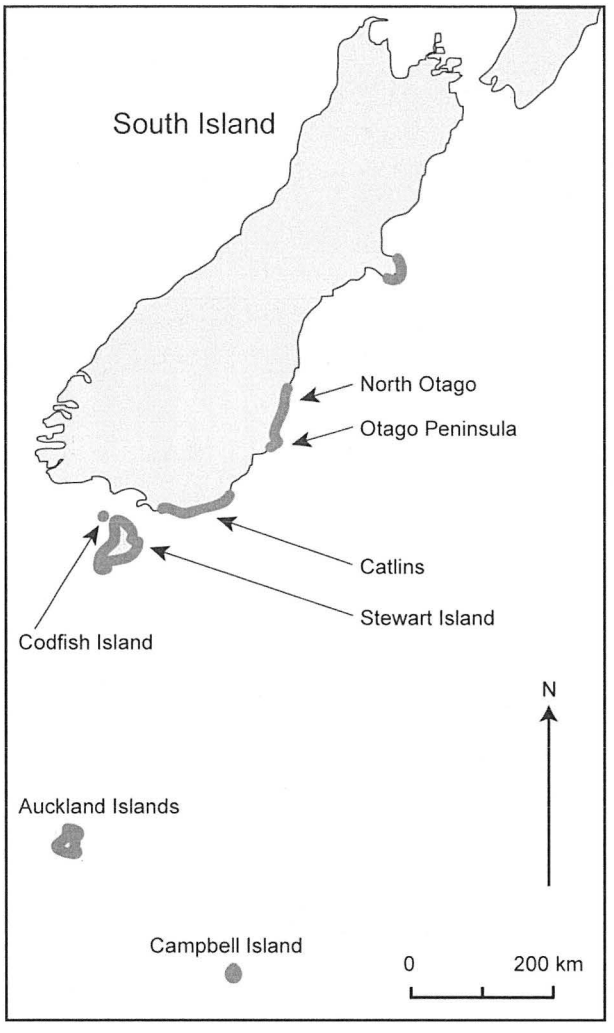


Figure 4.1 Map of the South and subantarctic islands of New Zealand. Arrows point to the geographical locations where samples were collected. The dark grey line represents the current breeding range of *M. antipodes*. Samples were not collected from the most northern part of the range (Banks Peninsula on the east coast of the South Island). In this area only ~30 birds are found and most are known migrants from Otago Peninsula.

was recovered dead on the South Island (Department of Conservation, unpublished data). Second, allozyme data show significant genetic differentiation between Auckland Islands and the South Island, and between Campbell Island and the South Island penguins (F_{ST} = 0.33 and 0.14, respectively; Triggs & Darby 1989). Finally, yellow-eyed penguins, like most penguin species, show strong reproductive philopatry with up to 98% of breeders returning to their previous breeding location (Richdale 1957; Ratz *et al.* 2004). Therefore, any dispersal between breeding areas most likely results from the movement of young adults before they reach maturity. Although

juveniles have been recorded travelling 600 km along the South Island (Marchant & Higgins 1990), almost 90% of the surviving juveniles eventually breed near the area where they hatched (Richdale 1957). Therefore, despite the inferred recent range expansion of *M. antipodes*, existing ecological and allozyme data imply that dispersal between the subantarctic and the South Island may be rare. In light of these apparent contradictions, an improved understanding of *M. antipodes* population structure and dispersal, including the identification of management units, is urgently needed to aid management of this endangered species.

The identification of demographically independent populations, and the subsequent designation of management units, is a two-step process. The first step requires identification of population boundaries. Such boundaries are not self-evident in species with considerable dispersal ability, and the separation of mobile individuals into groups based on geographical proximity is problematic. The physical barriers that restrict dispersal are particularly poorly understood for marine biota (Cassens *et al.* 2005). In such cases, genetic clustering analyses provide a promising means of inferring the number of populations, and these methods have already proven to be a valuable and reliable tool in conservation biology (Pearse & Crandall 2004; Manel *et al.* 2005; Waples & Gaggiotti 2006). Once populations have been identified, the second step requires quantitative estimates of dispersal rates among the different populations (Palsbøll *et al.* 2007). While the exact rate of migration required to prevent demographic independence is unknown (Waples & Gaggiotti 2006), it has been suggested that such independence occurs when dispersal rates drop below 10% (Hastings 1993). Determining dispersal rates from traditional genetic models is difficult, but recent advances in assignment methods now enable us to obtain critical information regarding genetic structure and gene flow (Berry *et al.* 2004; Pearse & Crandall 2004; Manel *et al.* 2005).

In this study we use microsatellite markers and mitochondrial DNA to test the hypothesis that *M. antipodes* comprises a single demographic population

across its New Zealand - subantarctic range. We discuss our results in light of the recent expansion and current conservation management of this species, and evaluate the designation of multiple management units.

Material and methods

Study area and sampling

Yellow-eyed penguin blood samples (N = 350) were collected at all major breeding areas in New Zealand and its subantarctic islands, covering the entire, known breeding range of the species (figure 4.1). In the subantarctic, samples were obtained from Campbell Island (52°32'S, 169°05'E, N = 49) and the Auckland Islands (50°29'S, 166°17'S, N = 52). On the South Island of New Zealand sampling was conducted in North Otago (45°23'S, 170°52'E, N = 35), Otago Peninsula (45°53'S, 170°37'E, N = 86), and the Catlins (46°34'S, 169°35'E, N = 38). Finally, samples were collected on Stewart Island (46°57'S, 168°80'E, N = 40) and Codfish Island (46°46'S, 167°38'E, N = 50), which lie just south of the South Island.

Samples were collected from 2005-2008, with the exception of 24 birds that were sampled on Otago Peninsula in 2001. These 24 birds were, however, all still breeding in this location in 2006 and can therefore be treated as part of the 2005-2008 sample. To avoid sampling dispersing birds, and to minimise sampling closely related individuals, samples were collected from either adults or chicks in each breeding area. Sampling of juveniles was avoided because juvenile birds are known to travel in their first year following fledging (Marchant & Higgins 1990). Juveniles can be recognised by the lack of a distinctive yellow eye and yellow crown across the head, features that develop during the first moult around 18 months after hatching. Resampling of individuals was avoided either by permanently tagging birds with flipper bands or transponders, or by visiting breeding sites during a single season only.

Penguins were captured on the nest (South Island and subantarctic islands) or on the beach (subantarctic islands only). Blood samples (~0.05 ml) were collected from the brachial vein using sterilised syringes and needles, and subsequently stored in 1 ml Queens lysis buffer (Seutin *et al.* 1991).

DNA extraction, microsatellite genotyping and mtDNA sequencing

DNA was extracted and purified using 40 µg proteinase K in 5% Chelex (Biorad; Walsh *et al.* 1991). All samples were genotyped at 12 microsatellite loci previously developed for yellow-eyed penguins (Man03, Man08, Man13, Man21, Man22, Man27, Man39, Man47, Man50, Man51, Man54, Man55; Boessenkool *et al.* 2008, Chapter 3). Microsatellite primer sequences and polymerase chain reaction (PCR) conditions are described in Boessenkool *et al.* (2008, Chapter 3).

Mitochondrial DNA sequence analyses targeted the first hypervariable region of the control region which was amplified using primers L-Man-CR4 and H-Man-CR7 as described in Boessenkool *et al.* (2009a, Chapter 2). The 813 bp fragment was sequenced for a subset of samples (N = 100), including 20 birds from each of the subantarctic breeding areas, and 15 from North Otago, Otago Peninsula, Catlins, and Codfish Island respectively. PCR products were sequenced one-way, using primer H-Man-CR7, resulting in a total of 731 bp used for further analyses. All haplotypes have been deposited in GenBank under accession numbers FJ822137 - FJ822143.

Hardy-Weinberg and linkage equilibrium of microsatellites

Deviations from Hardy-Weinberg proportions and linkage equilibrium were tested in each breeding area separately using GENEPOP 4.0 (Rousset 2008). Markov chain parameters employed 10 000 dememorizations, 1000 batches

and 10 000 iterations. Significance levels were adjusted for multiple comparisons using Bonferroni corrections (Rice 1989).

Population genetic structure

To identify genetically cohesive populations of yellow-eyed penguins we used the Bayesian clustering analysis implemented in STRUCTURE 2.2 (Pritchard *et al.* 2000). This model-based method uses a Markov chain Monte Carlo (MCMC) simulation to assign individuals to genetic clusters (K) on the basis of their genotypes, regardless of geographic sampling information. The analysis detects clusters under the assumptions of Hardy-Weinberg and linkage equilibrium within each cluster. Estimated membership coefficients per individual per cluster (Q) are calculated, allowing probabilistic assignment of individuals to clusters.

We performed five replicate runs for values $K = 1$ to $K = 7$ (seven being the number of sampled breeding areas), using the admixture model and assuming correlated allele frequencies (Pritchard *et al.* 2000; Falush *et al.* 2003). Exploratory runs showed that a burnin of 500 000 followed by 1 000 000 iterations was sufficient to achieve convergence in our dataset. We used a uniform prior for alpha, with a maximum of 10.0 and set ALPHAPROPSD to 0.05. Lambda was set at 1.0 and the prior for F_{ST} was left at default values with a mean of 0.01 and a standard deviation of 0.05. We estimated the optimal number of clusters for our data by comparing the log-likelihood of the data given the number of clusters (ln $P(X|K)$) (Pritchard *et al.* 2000), by examining the standardized second order rate of change of ln $P(X|K)$ (ΔK) (Evanno *et al.* 2005), and by evaluating individual membership coefficients for different values of K (Pritchard *et al.* 2000). The assumption of correlated allele frequencies can sometimes lead to overestimates of K (Pritchard *et al.* 2007), and we therefore repeated our analyses using the independent allele frequency model with lambda set to 1.0. The outcomes of

the analyses were, however, not affected by model choice and we report only results from runs in which correlated allele frequencies were assumed.

To evaluate levels of genetic variation within breeding areas and the populations identified from the clustering analyses in STRUCTURE, we calculated the total number of alleles, number of unique alleles, and expected and observed heterozygosity in GENETIX 4.05.2 (Belkhir *et al.* 1996-2004). When calculating levels of genetic variation within the identified populations, we placed samples into one of the identified populations according to their geographic sampling location rather than the population they were defined in genetically. In other words, individuals that had been placed genetically in a population different from their geographical sampling area were returned to their original location. This "correction" prevents biasing population allele frequencies due to the exclusion of possible migrants (Cegelski *et al.* 2003). For the mtDNA data we calculated haplotype and nucleotide diversity indices within each of the different breeding areas and the identified populations using DnaSP (Rozas *et al.* 2003).

To quantify levels of genetic differentiation we calculated F_{ST} values among both geographical breeding areas and between the populations identified by the clustering analysis. Calculations of Weir and Cockerham's (1984) F_{ST} were performed in GENETIX 4.05.2 (Belkhir *et al.* 1996-2004) for the microsatellite data, and F_{ST} values for mtDNA data were calculated with Arlequin 3.11 (Excoffier *et al.* 2005). Significance levels were tested using 10 000 permutations and adjustments for multiple comparisons were applied using Bonferroni corrections (Rice 1989). Genealogical relationships among samples were reconstructed using a parsimony-based haplotype network with a 95% parsimony criterion using TCS (Clement *et al.* 2000).

Detection of migrants

We employed two different methods to evaluate dispersal between the yellow-eyed penguin populations identified using STRUCTURE. We did not attempt to identify migrants between breeding areas, because assignment tests (see below) can be unreliable in identifying migrants when population units are only weakly differentiated (e.g. $F_{ST} < 0.05$; Berry *et al.* 2004).

First generation migrants were identified using the method specifically designed to detect F_0 immigrants, implemented in GENECLASS2 (Paetkau *et al.* 2004; Piry *et al.* 2004). We used the test statistic L_h/L_{max}, where L_h is the likelihood of drawing a genotype in the population from which it was sampled and L_{max} is the greatest likelihood of drawing this genotype in any of the sampled populations. This test statistic is appropriate when all source populations have been sampled (Paetkau *et al.* 2004), as is the case in the current study. We employed Rannala and Mountain's (1997) Bayesian criterion for likelihood estimation and Paetkau's *et al.* (2004) resampling method to generate critical values for rejecting the null hypothesis that an individual was born in the population where it was sampled. We simulated 10 000 individuals and set alpha at 0.001. This alpha value corresponds to an expected type I error rate of 0.35 migrants in our dataset comprising 350 individuals (see Paetkau *et al.* 2004). Since migration rates between the penguin populations appear to be low, we consider this an appropriate type I error rate for the current study.

The assignment test implemented in STRUCTURE 2.2 (Pritchard *et al.* 2000) was used to detect putative migrants along with any individuals with recent immigrant ancestry. The latter category was particularly important for our analysis, which includes 125 individuals that were sampled as chicks from the South Island. Although these individuals are obviously not migrants themselves, their genotypes may contain valuable information regarding the geographic origin of their parents. The assignment test implemented in STRUCTURE is a fully Bayesian method that uses geographical sampling

location as prior population information, and assumes with a user specified prior probability (v) that an individual is an immigrant (Pritchard *et al.* 2000). This specified probability that an individual is an immigrant (i.e. did not originate in its sampling location) can affect the outcome of the test (Pritchard *et al.* 2000). To account for uncertainty in v, we ran replicate analyses using distinct prior settings: v = 0.05 and v = 0.1. These settings correspond to individuals having a 5% or 10% probability, respectively, of being an immigrant or having migrant ancestry. Posterior probabilities (PP) of immigrant ancestry were calculated to two generations back, and models were run under the assumption that allele frequencies were correlated among populations with lambda set to 1.0. The MCMC simulation was performed with a burnin of 500 000 followed by 1 000 000 iterations.

Results

All 350 samples amplified at all 12 microsatellite loci with the exception of six samples, which have missing genotypes for one (three samples), three (two samples) or five loci (one sample), respectively. All 350 samples were included in all analyses. There was no evidence for linkage disequilibrium between any pair of loci in any of the breeding areas. Only one locus (Mano8) showed significant departure from Hardy-Weinberg proportions, and this was observed in just a single breeding area (Catlins).

Estimating the number of populations

Low variance in ln P($X|K$) across replicate runs, and visual inspection of the time series plots of the likelihood and the estimated parameters, confirmed convergence of the clustering analyses in STRUCTURE. Evaluation of ln P($X|K$), ΔK and Q for different values of K indicated that K = 2 captures the major genetic structure in our sample (figure 4.2). The likelihood of the data was lowest for K = 1, increased steeply for K = 2, and plateaued for higher values of K. When ln P($X|K$) only increases marginally above a certain value

of K (here $K = 2$), the lowest value of K is usually the best model choice for the data (Pritchard *et al.* 2007). The measure of ΔK also shows a clear mode for $K = 2$, but since it is not possible to estimate ΔK for $K = 1$ this measure confirms only that $K = 2$ is a better model for our data than higher values of K. At $K = 2$, the proportion of samples placed in each cluster was asymmetric (figure 4.3) and the average proportion of membership (average Q) of the seven breeding areas to one of the clusters ranged from 0.869 to 0.970. For models assuming $K = 3$, the average Q to the clusters dropped and many individuals were strongly admixed between two of the three clusters (figure 4.3). Taken together these results clearly support the $K = 2$ model for our data.

The two genetic clusters identified using STRUCTURE corresponded remarkably well with the geographic locations. The first cluster was made up of the subantarctic breeding areas Auckland and Campbell Islands. The average Q of these breeding areas to this cluster was 0.908 and 0.970, respectively. All the breeding areas on and near the South Island were placed in the second cluster. Average Q to this cluster was 0.976, 0.950, 0.925, 0.869 and 0.958 for North Otago, Otago Peninsula, Catlins, Stewart and Codfish Island, respectively. In all subsequent analyses, and throughout the discussion, we refer to the two clusters as *subantarctic* and the *South Island* populations. Since we are ultimately interested in geographic populations, and to prevent biasing allele frequencies by exclusion of migrants, we placed samples into one of the two populations (subantarctic or South Island) according to their geographic sampling location (see Materials and Methods).

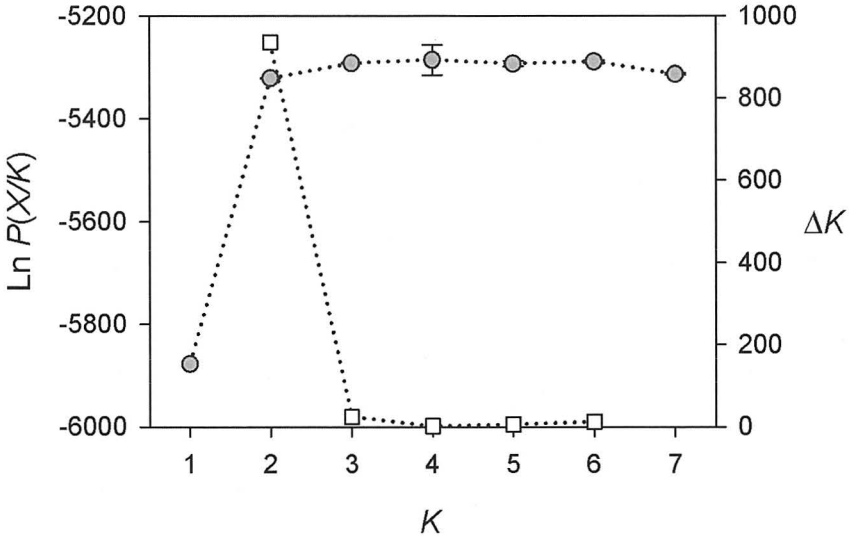


Figure 4.2 Inference of the number of genetic clusters (K) estimated using STRUCTURE. Both ln P($X|K$) (the likelihood of the data given K; grey circles) and ΔK (the standardized second order rate of change of ln P($X|K$); white squares) are plotted as a function of K. Error bars (where discernible) of ln P($X|K$) indicate standard deviations.

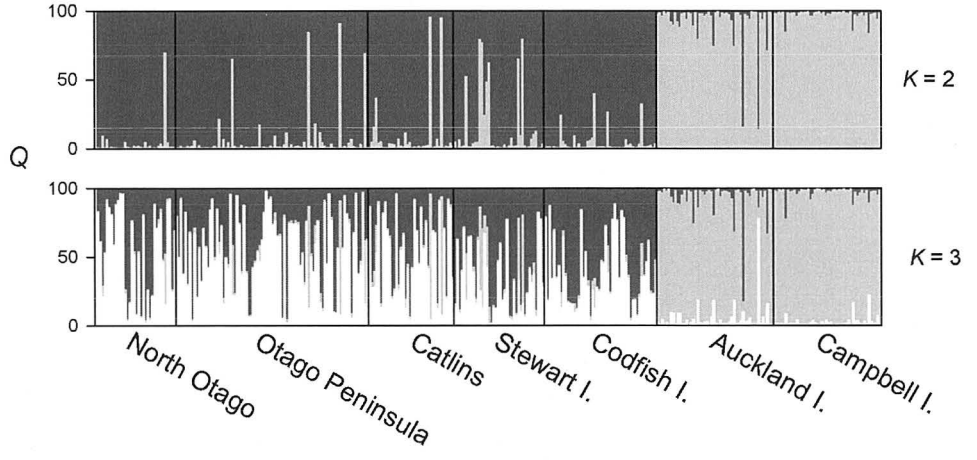


Figure 4.3 Proportional membership (Q) of yellow-eyed penguins to genetic clusters (K) for K = 2 (top graph) and K = 3 (bottom graph) as estimated using STRUCTURE. Each vertical bar represents a single individual and individuals are ordered by geographic sampling location. Colours correspond to genetic clusters.

Genetic variation within breeding areas and populations

The two populations South Island and subantarctic differ substantially in levels of genetic variation (Table 4.1). All microsatellite loci were polymorphic in the subantarctic, while only 9 out of 12 loci showed variation in the South Island. Consequently, the subantarctic population has on average more alleles per locus and higher mean heterozygosities. We detected 18 alleles that were unique to the subantarctic, whereas only two alleles were unique to South Island, and both of these were rare (frequency = 0.006 and 0.002, respectively). Within the South Island there was very little variation in levels of genetic diversity among the different breeding areas, whereas within the subantarctic, the Auckland Islands breeding area had more alleles per locus and more unique alleles than Campbell Island (Table 4.1). Allele frequencies per breeding area can be found in Appendix 9.5.

We observed low levels of genetic variation in the mitochondrial control region. Only seven haplotypes were detected, with a total of eight variable sites. Although overall genetic diversity was low, the subantarctic population possesses more genetic variation (h = 0.69 and π = 0.0028) than the South Island population (h = 0.31 and π = 0.0012; Table 4.1). All haplotypes were closely related and all rare haplotypes differed from two common haplotypes (C and E) by a single nucleotide substitution (figure 4.4). These two common haplotypes were the only haplotypes shared between the subantarctic and the South Island populations. A total of four haplotypes were unique to single breeding locations (A, D, F and G; Table 4.1, figure 4.4).

Table 4.1 Genetic variation at 12 microsatellite loci and 731 bp of the HVI mitochondrial control region in the yellow-eyed penguin populations South Island and subantarctic and each of the individual breeding areas. N = samples,size, L_{poly} = number of polymorphic loci, A = mean number of alleles per locus, A_{un} = number of unique alleles, H_E = expected heterozygosity, H_O = observed heterozygosity, H = number of haplotypes, h = haplotype diversity, π = nucleotide diversity. Mitochondrial data was not available for the Stewart Island samples.

	Microsatellite data						mtDNA			
Location	N	L_{poly}	A	A_{un}	H_E	H_O	N	H	h	π
South Island	**249**	**9**	**2.9**	**2**	**0.32**	**0.31**	**60**	**3**	**0.31**	**0.0012**
N. Otago	35	9	2.2	0	0.32	0.35	15	2	0.42	0.0017
Otago P.	86	9	2.6	0	0.31	0.30	15	3	0.45	0.0016
Catlins	38	9	2.6	1	0.31	0.30	15	2	0.34	0.0014
Stewart I.	40	9	2.3	0	0.33	0.30	-	-	-	-
Codfish I.	50	9	2.3	0	0.31	0.31	15	1	0.00	0.0000
Subantarctic	**101**	**12**	**4.3**	**18**	**0.46**	**0.44**	**40**	**6**	**0.69**	**0.0028**
Auckland I.	52	12	4.2	9	0.44	0.42	20	4	0.65	0.0025
Campbell I.	49	12	3.3	1	0.45	0.45	20	5	0.66	0.0028

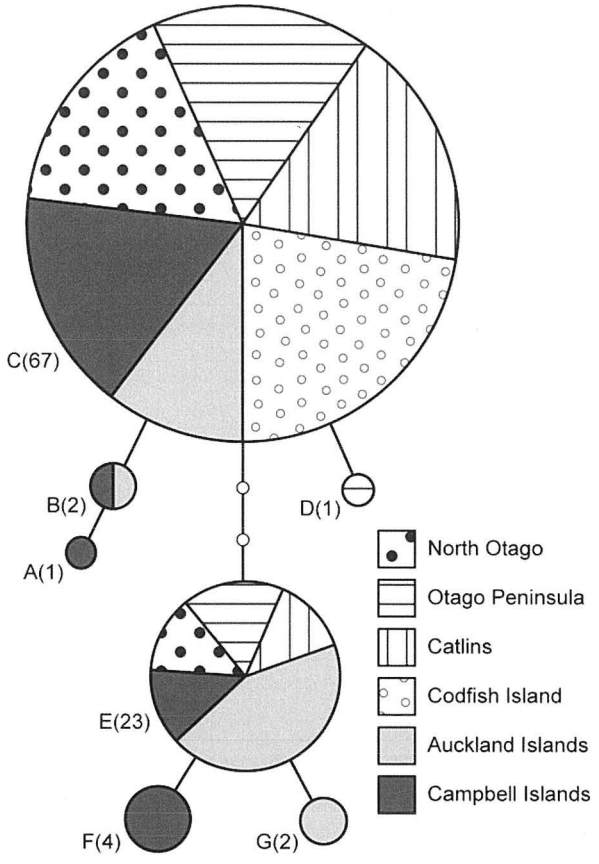


Figure 4.4 Haplotype network of yellow-eyed penguin control region haplotypes. Circle size corresponds to the number of samples possessing each haplotype. White dots represent hypothetical intermediate haplotypes not detected in the current study. Letters A-G represent haplotype names followed by the total number of samples possessing each haplotype in brackets. All haplotype sequences have been deposited in GenBank (accession numbers FJ822137 - FJ822143).

Genetic differentiation among breeding areas

Significant genetic differentiation between the subantarctic and the South Island populations was indicated by an F_{ST} value of 0.108 ($p < 0.01$) for microsatellite data. Comparison of microsatellite pairwise F_{ST} values among the seven breeding areas indicates the subdivision of the breeding areas into two populations: all F_{ST} values between South Island and subantarctic breeding locations (minimum 0.085) were substantially larger than values between breeding locations within a population (maximum 0.044; Table 4.2). Many of the microsatellite F_{ST} values among breeding locations within populations were, however, significantly different from zero ($p < 0.05$, Table

4.2). In contrast to the microsatellite data, all but one of the pairwise F_{ST} values for mtDNA data were non-significant (Table 4.2). When breeding areas were pooled into their respective populations, however, significant divergence between South Island and subantarctic is evident for the mtDNA control region (F_{ST} = 0.144, p < 0.001).

Table 4.2 Pairwise F_{ST} values among yellow-eyed penguin breeding areas for 12 microsatellite loci (above diagonal) and 731 bp of the HVI mitochondrial control region (below diagonal). Mitochondrial data was not available for the Stewart Island samples. Bold values indicate values that are significantly different from zero after Bonferroni correction.

Location	North Otago	Otago Peninsula	Catlins	Stewart Island	Codfish Island	Auckland Island	Campbell Island
N. Otago		0.005	**0.021**	**0.038**	0.010	**0.110**	**0.138**
Otago P.	-0.060		0.006	**0.031**	0.007	**0.120**	**0.146**
Catlins	-0.058	-0.059		**0.024**	**0.016**	**0.088**	**0.111**
Stewart I.	-	-	-		**0.026**	**0.085**	**0.118**
Codfish I.	0.214	0.161	0.143	-		**0.098**	**0.132**
Auckland I.	0.120	0.143	0.191	-	**0.463**		**0.044**
Campbell I.	0.024	0.016	0.044	-	0.228	0.102	

Detection of migrants

The analyses for detection of first generation migrants in GENECLASS identified only one individual as a migrant (p < 0.001). This penguin was sampled as an adult in the Catlins breeding area. For all other individuals we were not able to reject the null hypothesis that they were born in the area where they were sampled.

To visualize the results of the assignment test in STRUCTURE we ranked all individuals according to the posterior probability that they have no immigrant ancestry (figure 4.5). Two individuals, both from the Catlins

breeding area, were identified as potential migrants from the subantarctic. One of these individuals is the same individual identified as a first generation migrant by GENECLASS. In addition, two penguins from Otago Peninsula had moderate posterior probabilities of having migrant ancestry (PP > 40% for having immigrant parents and grandparents), and a further five penguins from South Island breeding sites showed signatures of being 2nd generation migrants (PP > 30% for having immigrant grandparents). In total only 13 individuals had a $PP_{\text{non immigrant}} < 0.80$, and of these only two were subantarctic birds. Increasing the migration prior (v) to 0.1 resulted in similar patterns, and we therefore only present the results for $v = 0.05$.

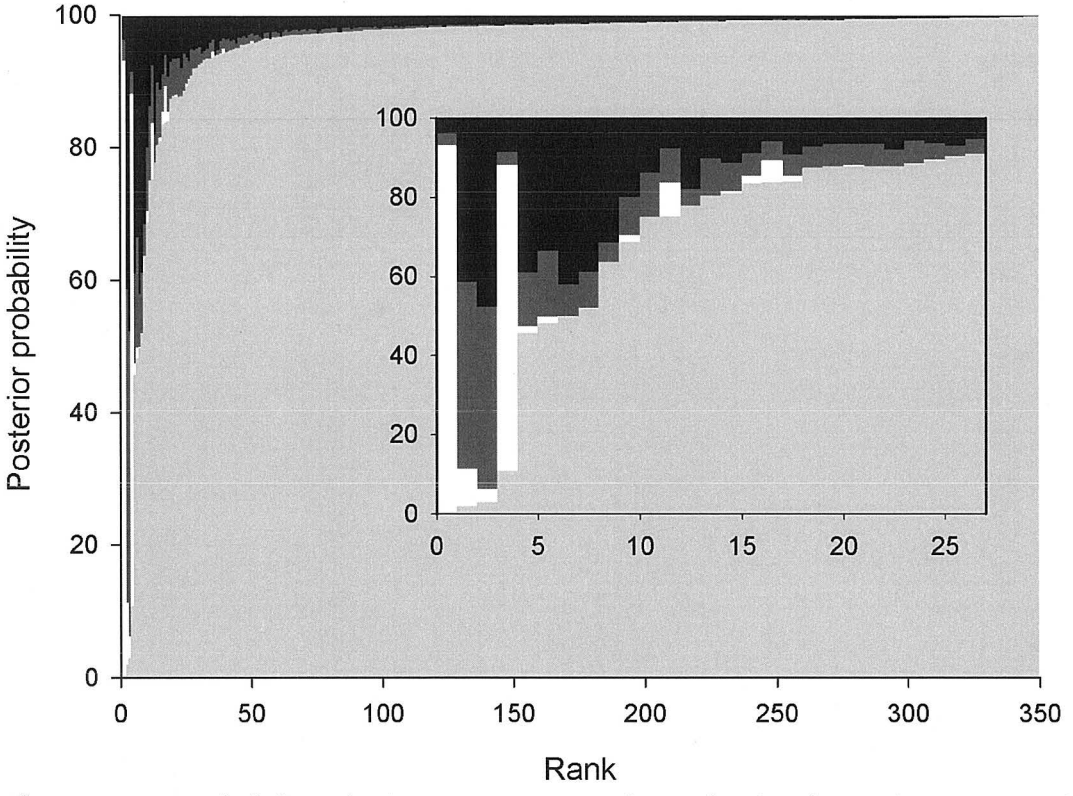


Figure 4.5 Probability of migrant ancestry estimated using the assignment test in STRUCTURE. Individuals are ranked according to the posterior probability that they have no immigrant ancestry. Colours correspond to the four different categories: no immigrant ancestry = light grey; immigrant = white; immigrant parent = dark grey; immigrant grandparent = black. The inset histogram shows an enlargement of the graph for the first 25 individuals.

Discussion

Contrary to our hypothesis that yellow-eyed penguins comprise one single population — based on known dispersal capabilities and inferred recent colonization history — our genetic analyses clearly indicate that *M. antipodes* comprises two broadly distinct genetic assemblages (South Island versus subantarctic). Furthermore, ongoing levels of migration among these two assemblages are sufficiently low to ensure demographic independence, and therefore warrant their designation as two distinct management units (MUs).

Population genetic structure in yellow-eyed penguins

Although the inference of the correct number of populations is potentially problematic (Pritchard *et al.* 2000; Evanno *et al.* 2005; Pritchard *et al.* 2007), our genetic clustering analyses revealed surprisingly strong evidence for the presence of two distinct genetic clusters ($K = 2$) in our *M. antipodes* data set. A biological interpretation of $K = 2$ is straightforward due to the evident concordance of the genetic clusters with geographic locations. The first cluster is made up of the breeding areas on the subantarctic Auckland and Campbell Islands (comprising the subantarctic population), whereas the second cluster contains all breeding locations sampled on the South Island of New Zealand and includes the near-shore islands Codfish and Stewart Island (comprising the South Island population).

Recognition of two yellow-eyed penguin populations is supported by the substantial microsatellite F_{ST} values detected among samples from the two geographic areas (Table 4.2). Although we note that many additional comparisons between breeding areas also yielded significant (albeit relatively low) values, the rejection of panmixia per se is not sufficient for the identification of populations or MUs (Palsbøll *et al.* 2007). Indeed, F_{ST} values are primarily useful as basic descriptors of genetic differentiation (Neigel

2002; Pearse & Crandall 2004) rather than providing a basis for the delineation of genetic assemblages or populations.

Interestingly, the strong population genetic differentiation observed for the microsatellite data is not reflected in mtDNA analyses when individual breeding areas were considered. This difference probably reflects a lack of resolution for mtDNA due to the detection of only seven haplotypes. Although five haplotypes were unique to either subantarctic (4) or South Island (1), two common haplotypes dominate the mtDNA variation in *M. antipodes* in both populations.

Long distance dispersal

The inferred recent colonisation of the South Island by yellow-eyed penguins from the subantarctic (Boessenkool *et al.* 2009a, Chapter 2) provides clear evidence of the species' ability to cross vast stretches of open ocean. Movement capabilities are, however, not always indicative of dispersal patterns (Milot *et al.* 2008). Only two individuals on the South Island were identified as migrants from the subantarctic using the assignment test in STRUCTURE (figure 4.5); one of these individuals was also recognised as an F_0 migrant by the analysis run in GENECLASS. The slight discordance between the two methods is not surprising, because we set alpha at a conservative value of 0.001 for the GENECLASS analysis. This alpha level reduced our Type I error rate (i.e. assigning migrant status while the individual is a non-migrant) to an acceptable value (see Material and Methods), but obviously increased our Type II error rate (i.e. failing to reject non-migrant status while the individual is a migrant). This arbitrary trade-off between Type I and Type II error rates is absent from the Bayesian framework that is implemented in STRUCTURE. Additionally, the STRUCTURE analysis allows for the investigation of migrant ancestry by identifying 1st and 2nd generation migrants in the assignment test, a feature that appears to be neglected by other studies that use the assignment test in

STRUCTURE (e.g. Cegelski *et al.* 2003; Bergl & Vigilant 2007). Performing this test without allowing for migrant ancestry doubled the number of migrant yellow-eyed penguins from the subantarctic to the South Island (data not shown). Specifically, the two individuals with high posterior probabilities for migrant ancestry (> 40%) would have been identified as F_0 migrants. Ignoring the possibility of migrant ancestry therefore leads to biased results overestimating the number of migrants, and researchers should be encouraged to incorporate migrant ancestry when estimating dispersal from assignment tests.

The strong genetic differentiation and low levels of gene flow observed in yellow-eyed penguins contrast with those found using microsatellite data in other penguin species. Adelie (*Pygoscelis adeliae*), Humboldt (*Spheniscus humboldti*) and blue penguins (*Eudyptula minor*) all show weak genetic heterogeneity and relatively high gene flow over large geographical areas (Roeder *et al.* 2001; Overeem *et al.* 2008; Schlosser *et al.* 2009). The surveyed populations of these species are, however, not separated by hundreds of kilometres of open ocean, and instead include populations along long stretches of coast lines.

Our data reveal that the large stretch of ocean separating the South Island from the subantarctic islands is a natural barrier that limits gene flow for yellow-eyed penguins. The water masses between the South Island and the subantarctic islands are separated by the Subtropical convergence, one of three major oceanic fronts in the South-West Pacific (Heath 1985). This convergence separates subtropical and subantarctic waters and is a well-known biogeographic and ecological boundary for marine taxa (Mackintosh 1960; Jouventin *et al.* 2006; de Dinechin *et al.* 2009). The current study shows that the Subtropical convergence apparently also acts as a barrier for dispersal in yellow-eyed penguins, impeding subantarctic-South Island gene flow.

The higher number of individuals with migrant ancestry on the South Island (11 individuals) compared to the subantarctic (two individuals) suggest that directional migration from south to north is more prevalent than vice versa. This result, however, might be confounded by the difference in level of genetic variation between the two populations. The subantarctic population has 18 unique alleles compared to only two in the South Island, and migrants carrying these unique alleles should be easier to identify genetically compared to migrants possessing shared alleles. Whether directional migration from south to north is more common than from north to south therefore remains uncertain.

Range expansion

Range expansion has been suggested as one of the factors reducing genetic structure in seabirds populations (Friesen *et al.* 2007). The genetic effects of range expansion are, however, not straightforward and depend strongly on the mode of expansion. Specifically, if range expansion is achieved by a relatively small number of individuals, founder effects can lead to a major reduction in genetic diversity in the newly founded population and strong genetic differentiation between source and founding population (Nei *et al.* 1975; Chakraborty & Nei 1977; Tarr *et al.* 1998; Pruett & Winker 2005; Fabbri *et al.* 2007). Random genetic drift will further enhance these processes if the newly founded population has a slow growth rate. The low genetic variation in the South Island yellow-eyed penguin population compared to the subantarctic population, together with the strong genetic differentiation between these two populations, suggests that the South Island yellow-eyed penguin population has been subject to such founder effects, and that this population was founded by only a small number of individuals. The low level of gene flow between the subantarctic and the South Island has subsequently been insufficient to homogenize the genetic structure generated by the founder event.

M. antipodes' inferred recent range expansion across the major ecological and biogeographic boundary of the Subtropical convergence is remarkable. By contrast, in rockhopper penguins, *Eudyptes chrysocome sensu lato*, range expansion to Amsterdam and Saint Paul Islands in the Indian Ocean appears to have occurred from southern Atlantic populations that lie in the same watermass (i.e. north of the Subtropical convergence), rather than from nearby Crozet and Kerguelen, which lie across the Subtropical convergence (de Dinechin *et al.* 2009). Ocean convergences could therefore be considered driving forces for vicariant speciation in southern ocean taxa, including penguins (Mackintosh 1960; de Dinechin *et al.* 2009). The recognition of two *Megadyptes* species –with one species now extinct– (Boessenkool *et al.* 2009a, Chapter 2) on either side of the subtropical convergence supports this speciation hypothesis. Whether the current levels of gene flow between the two populations of *M. antipodes* are sufficiently high to prevent speciation in the long term remains a question in the realm of speculation.

Identification of management units and conservation implications

The recognition of two management units in yellow-eyed penguins is warranted based on the two distinct genetic clusters and the identification of low numbers of migrants. Although we did not test the actual migration rate per se, we identified only two migrants out of 124 adult penguins sampled on the South Island, which approximates a conservative immigration rate in to the South Island population of 1.6%. If we apply the guideline that demographic independence occurs when migration rates drop below 10% (Hastings 1993; Palsbøll *et al.* 2007), it is evident that the migration rate between the subantarctic and South Island populations is sufficiently low to achieve demographic independence of each of these populations. Consequently, yellow-eyed penguins can no longer be regarded as one large, panmictic population. This finding has important implications for the conservation of this endangered species. Monitoring efforts have been intense on the South Island and full nest counts are made yearly in most

areas (McKinlay 2001). As a result, a reliable population estimate exists and changes in population size are detected rapidly. The demographic independence of the South Island population means, however, that the dynamics that are observed in the South Island may not reflect the dynamics in the subantarctic, and thus separate monitoring of this area is required. At the very least, a count of breeding yellow-eyed penguins on the subantarctic will have to be conducted, because data from this area are either incomplete or out of date (particularly on the Auckland Islands; Moore 1992; Moore *et al.* 2001).

Yellow-eyed penguins on the South Island have suffered several severe population crashes during the last decades that were attributed to food shortages or disease (Marchant & Higgins 1990; Gill & Darby 1993). Predation by re-colonising sea lions (*Phocarctos hookeri*) on the South Island and disturbance due to tourism are further, more recently identified threats that are likely to worsen in the coming years (Ellenberg *et al.* 2007; Lalas *et al.* 2007). Demographic independence and the low immigration from the subantarctic, mean that the South Island population has lower resilience to these threats than previously assumed. Although the species has expanded its range from the subantarctic to the South Island before, significant changes occurred to the habitat on the South Island since European settlement over the last 200 years (e.g. introduction of a diversity of predatory mammals) and it can not be assumed that such a range expansion would happen again. On a more positive note, the low gene flow from the South Island to the subantarctic makes it unlikely that diseases prevalent in the South Island population are transferred to the south. In fact, having two demographically independent populations prevents having 'all (penguin) eggs in one basket' and may protect the yellow-eyed penguin from extinction due to local, stochastic catastrophic events.

Acknowledgements

We are grateful to Jon Potter, Matt Charteris, Annabel Gosset, Louise Chilvers, Kath Walker and Graeme Elliott for sample collection on the subantarctic islands. We are also indebted to Sheryl Pullar, Sandy King, Rosalie Goldsworthy, Thomas Mattern, Ursula Ellenberg, Tiffany Brown, Andrew Hill and numerous field assistants for help with collecting samples around the South Island. Samples collected on Otago Peninsula during 2001 were kindly provided by Lloyd Davis. We would like to thank Tania King for her guidance and advice in the laboratory. Thanks also to Thomas Mattern and John Darby for providing information and ideas on yellow-eyed penguin dispersal. This research was supported by the Department of Zoology, University of Otago, including PBRF Research Enhancement grants to J.M.W. and P.J.S. Sample collection was supported by the Yellow-Eyed Penguin Trust and the Department of Conservation. Samples were collected under Department of Conservation permits SO-17933-FAU and OT-19097-RES and University of Otago Animal Ethics Approval 69/06.

Chapter 5:

Temporal genetic samples indicate small effective population size of the endangered yellow-eyed penguin

---

This chapter is in preparation for submission to Conservation Genetics as:

Sanne Boessenkool, Bastiaan Star, Philip J Seddon & Jonathan M Waters. Temporal genetic samples indicate small effective population size of the endangered yellow-eyed penguin.

Abstract

There is an increasing awareness that the long-term viability of endemic island populations in New Zealand is negatively affected by genetic factors associated with population bottlenecks and/or persistence at small population size. Here we use contemporary samples and historic museum specimens (collected 1888 – 1938) to estimate the effective population size (N_e) for the endangered yellow-eyed penguin (*Megadyptes antipodes*) on the South Island of New Zealand, and evaluate the genetic concern for this iconic species. The South Island population of *M. antipodes* — constituting almost half of the species census size — is thought to be descended from a small number of founders that reached New Zealand just a few hundred years ago. Despite intensive conservation measures, this population has shown dramatic fluctuations in size over recent decades. We compare estimates of the harmonic mean N_e for this population, obtained using one moment and three likelihood based-temporal methods, including one method that simultaneously estimates migration rate. Evaluation of the N_e estimates reveals a harmonic mean N_e in the low hundreds. Additionally, the inferred low immigration rates (m = 0.003) agree well with known migration rates between the South Island and subantarctic populations of *M. antipodes*. The low N_e of South Island *M. antipodes* is likely affected by strong fluctuations in population size, and high variance in reproductive success. These results show that genetic concerns for this population are valid and that the long-term viability of this species may be compromised by reduced adaptive potential.

Introduction

Untangling the relative roles of genetic and demographic factors that affect the persistence of endangered populations is a fundamental goal of conservation biologists and wildlife managers. In New Zealand, exotic mammalian predators have played a dramatic role in the decline and extinction of endemic fauna (Clout 2001; Duncan & Blackburn 2004), but intense conservation efforts have resulted in the eradication or control of these predators in localised mainland and offshore areas. Recent New Zealand conservation studies have also started to highlight the potential role of genetic factors in shaping the long-term viability of persisting endemic populations (Jamieson 2007; Jamieson *et al.* 2008). In particular, it is recognised that the loss of genetic diversity and increased levels of inbreeding — due to population bottlenecks and/or persistence at small population sizes — might have reduced mean population fitness and adaptive potential (Allendorf 1986; Lande & Shannon 1996; Frankham *et al.* 2002; Keller & Waller 2002).

Effective population size (N_e), defined as the size of an ideal population experiencing the same rate of genetic drift as the actual population under consideration (Wright 1931; Frankham 1995; Palstra & Ruzzante 2008), is a key parameter in studies of genetic diversity. Historically, estimation of N_e has been notoriously difficult, but this situation has been much improved by recent statistical developments facilitating the estimation of N_e from temporal genetic samples (Wang 2001; Berthier *et al.* 2002; Beaumont 2003; Wang & Whitlock 2003). These so-called temporal methods estimate the harmonic mean of a population's variance effective size based on the change in allele frequencies over the time interval separating the temporally spaced samples. The use of museum specimens is particularly promising in the estimation of N_e for species with long generation times (Wandeler *et al.*

2007). Here we use contemporary and historical samples to estimate N_e for the endangered yellow-eyed penguin (*Megadyptes antipodes*) on the South Island of New Zealand, and evaluate the genetic concern for this iconic species.

M. antipodes is thought to have expanded its range from the subantarctic islands to the South Island of New Zealand around 500 years ago, after the arrival of Polynesians but before settlement by Europeans and their commensals (figure 5.1; Boessenkool *et al.* 2009a, Chapter 2). Based on current low migration rates (<2%) between the South Island (including surrounding islands such as Stewart Island) and the subantarctic yellow-eyed penguins, and the relatively low levels of genetic variation of the current South Island population, it is thought that the South Island population descended from a small number of founders (Boessenkool *et al.* 2009b, Chapter 4). Nevertheless, around 40% (~800 nests, ~2200 individuals) of *M. antipodes* globally are now found on and around the South Island of New Zealand (McKinlay 2001).

By the 1980s, non-native predators — chiefly mustelids and cats, introduced by Europeans in the late 19th century — had caused major egg and chick predation (Darby & Seddon 1990) and prompted the implementation of intensive predator trapping around *M. antipodes* breeding areas. Despite these recent conservation measures, however, *M. antipodes* population sizes have remained highly unstable (McKinlay 2001; Moore 2001). This demographic instability has been attributed to changes in food supply (van Heezik & Davis 1990), climatic variations (Peacock *et al.* 2000) and disease epidemics (e.g. Gill & Darby 1993; Department of Conservation unpublished data). Regardless of their underlying causes, such fluctuations in population size are a primary factor leading to substantial reductions in N_e (Frankham 1995).

Based on the suggested recent founding of South Island *M. antipodes*, with subsequent fluctuations in population size, conservation biologists hold

genuine concerns for this population. In particular, the ongoing emergence of novel diseases (for example a diphtheria-like disease linked to infection by a strain of *Corynebacterium* led to 60% of chicks dying on the South Island in 2004; Department of Conservation, unpublished data) suggests that the adaptive potential of this population may be limited, a concern that may become increasingly important with predicted climate change. In this study we use microsatellite analyses of contemporary and historic South Island samples to test for temporal changes in genetic diversity over the last century, and to provide genetic estimates of N_e.

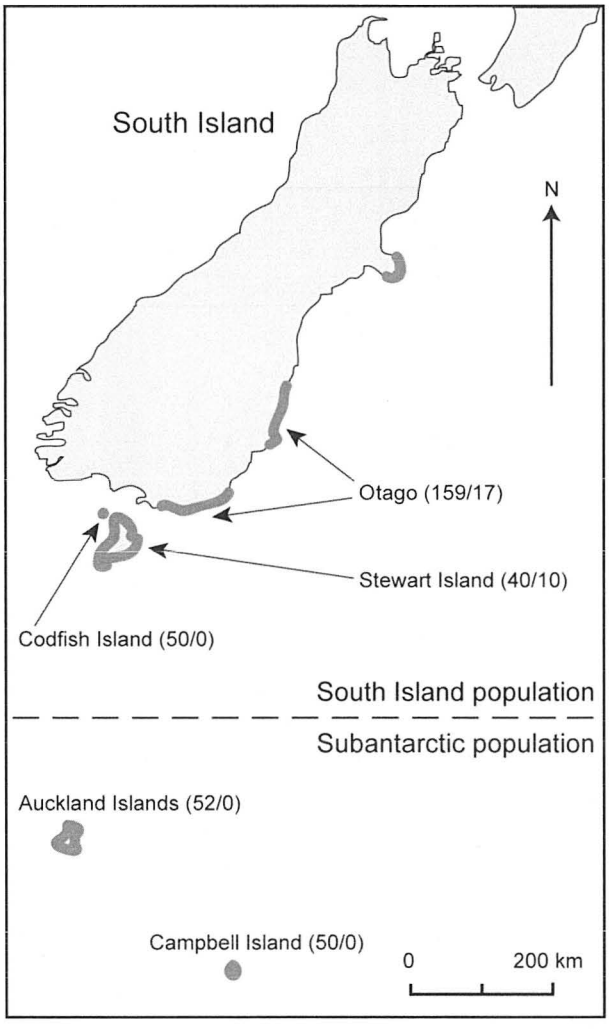


Figure 5.1 Map of the South and subantarctic islands of New Zealand. Arrows point to the geographical locations where samples were collected. Sample sizes for contemporary/historic samples are given in brackets. The dashed line refers to the genetic split between South Island and subantarctic populations.

Material and Methods

Study area and sampling

Yellow-eyed penguin blood samples (N = 249) were collected between 2005 and 2008 at five breeding areas on and around the South Island of New Zealand, including Stewart and Codfish Islands (figure 5.1). Together, these areas form the South Island yellow-eyed penguin population (Boessenkool *et al.* 2009b, Chapter 4). A total of 101 additional samples was collected from the subantarctic Auckland and Campbell Islands (genotypes of these samples are used for N_e estimates that allow for migration, see below). Details of blood sampling methods are described in Boessenkool *et al.* (2009b, Chapter 4).

To facilitate sampling of historic yellow-eyed penguin specimens we contacted a total of 128 museums around the world. Toe pad samples were obtained from 35 specimens collected between 1888 and 1938 at several locations on the South Island and on Stewart Island (figure 5.1, Appendix 9.6). These 35 samples included almost all yellow-eyed penguins specimens from the South Island with an explicit collection date (<1950) that are currently held in museum collections.

DNA extraction and genotyping

DNA from contemporary samples was extracted and purified using 40 µg proteinase K in 5% Chelex (Biorad; Walsh *et al.* 1991). All samples were genotyped at 12 microsatellite loci previously developed for yellow-eyed penguins (Man03, Man08, Man13, Man21, Man22, Man27, Man39, Man47, Man50, Man51, Man54, Man55; Boessenkool *et al.* 2008, Chapter 3). Microsatellite primer sequences and polymerase chain reaction (PCR)

conditions for contemporary samples are described in Boessenkool *et al.* (2008, Chapter 3).

For DNA extraction of historic toe pad samples a ~1 × 2 mm piece was rehydrated by a 24 h wash in 1 ml 10 mM Tris-HCL (pH 8.0). Following rehydration, toepad samples were finely cut with a sterile scalpel blade and DNA was extracted using the Chargeswitch Forensic DNA Purification Kit (Invitrogen) or the DNeasy Tissue Kit (Qiagen) following manufacturers' instructions. No differences were observed in extraction or amplification success between these two kits. Historic samples were amplified at the same 12 loci described above, with the exception of Man22 and Man27 which did not amplify consistently for the historic samples. These two loci were therefore omitted from all further analyses. PCR reactions (10 µl) contained 2 µl DNA, 0.5 µM of each primer, 0.5 U *Taq* DNA polymerase (Mango *Taq*, Bioline), 1× *Taq* buffer, 0.8 µM dNTP and 1.5 µM $MgCl_2$, with the addition of betaine and DMSO (1.1 M and 2% respectively) if necessary (see Boessenkool *et al.* 2008, Chapter 3). The amplification profile was 2 min at 94°C, 35-50 cycles of 15 s at 96°C, 15 s at 45-50°C and 30 s at 72°C, followed by a 4 min final extension at 72°C.

To prevent contamination of historic DNA with exogenous DNA or PCR products, all DNA extractions and PCR set-up of historic samples were performed inside a UV hood in a laboratory where no contemporary yellow-eyed penguin DNA or vertebrate PCR products have ever been present. Standard precautions for the analysis of historic DNA were closely adhered to, including the use of filter tips, UV radiation and cleaning of materials with bleach and/or 70% ethanol before and after each laboratory session, and maintenance of a one-way flow from the historic DNA laboratory to the modern/post-PCR laboratory. Historic samples were extracted in small batches of nine samples and potential contamination was monitored by negative extraction and PCR controls. To minimise the risk of erroneous genotypes due to allelic dropout and the amplification of false alleles

(Taberlet *et al.* 1996; Sefc *et al.* 2003), 2-7 successful amplifications were obtained for each historic sample before a genotype was scored, and genotypes were only scored when every allele was observed at least twice.

Disequilibrium and genetic diversity

Deviations from Hardy-Weinberg proportions and linkage equilibrium were tested using GENEPOP 4.0 (Rousset 2008) for contemporary South Island, contemporary subantarctic, and historic South Island samples separately. Markov chain parameters employed 10 000 dememorizations, 1000 batches and 10 000 iterations. Significance levels were adjusted for multiple comparisons using Bonferroni corrections (Rice 1989). Genetic diversity was quantified for the ten loci that amplified consistently in contemporary and historic samples, using the total number of alleles and expected and observed heterozygosity calculated in GENETIX 4.05.2 (Belkhir *et al.* 1996-2004). Calculations of allelic richness were performed using FSTAT 2.9.3 (Goudet 2002) to adjust for samples size differences. Statistical significance of differences in genetic diversity between historic and contemporary South Island samples was tested with a Wilcoxon signed rank test in SPSS (α = 0.05). The difference in genetic diversity between subantarctic and contemporary South Island is discussed extensively in Boessenkool *et al.* (2009b, Chapter 4), and subantarctic diversity is included here for comparative purposes only.

Effective population size

The quantification of N_e using temporal methods requires an estimation of the number of generations (T) separating the temporally spaced sampling points. We calculated average generation time using the formula $\Sigma(l_x\ b_x\ x)/\Sigma(l_x\ b_x)$, where x is age, l_x is the proportion of individuals surviving to age x and b_x is the reproductive output at age x (Begon *et al.* 2006; Table 5.1). Yearly adult survival of *M. antipodes* was set to 0.856 (Richdale 1957) and

reproductive output set to 1.16 fledglings per pair (Darby & Seddon 1990). Maximum age was set to 20 years (Department of Conservation unpublished data) and variation in age at first breeding as estimated by Richdale (1957) was incorporated in the analysis. Using these estimates, average generation time of *M. antipodes* was calculated at 7.7 years (see Table 5.1). The time span between the collection year of contemporary samples (2006) and the weighted average collection year for historic samples (1901) was 105 years, resulting in $T = 14$. To account for uncertainty in T we also present estimates of N_e using $T = 12$ and $T = 16$.

We used one moment and three likelihood-based approaches of the temporal method to obtain estimates of N_e. These methods typically assume discrete generations, no selection, no mutation, and a closed panmictic population. Although our dataset violates the first of these four assumptions, any bias due to overlapping generations can be minimised if samples are taken more than 10 generations apart (Waples & Yokota 2007), which is the case in our study. The effects of migration are more complex (Wang & Whitlock 2003; Fraser *et al.* 2007a; Palstra & Ruzzante 2008), and we therefore included an estimator of N_e that relaxes the assumption of a closed population (this estimator is referred to as N_{eOPEN}, in contrast to the other estimators, which are referred to as $N_{eCLOSED}$).

First, we calculated the moment estimator from Waples (1989) using the program NeEstimator (Peel *et al.* 2004). Second, we applied the coalescent-based likelihood method from Beaumont (2003) as implemented in the program TMVP (which is based on the program TM3 from Berthier *et al.* 2002). We assumed no change in N_e during the sampling interval and calculated N_e as the mode of the posterior distribution. The MCMC simulation was performed with 50 000 updates of which ten percent were discarded as burnin. The size of importance sampling was 100, the thinning interval was 10 and the size of the proposal distribution of parameter updates was 0.5. Third, we estimated N_e with the pseudo-likelihood based approach from Wang (2001) using the program MLNE. Finally, we applied the pseudo-

likelihood method from Wang and Whitlock (2003) that relaxes the assumption of no migration by jointly estimating N_e (N_{eOPEN}) and the migration rate m. This method requires allelic data from the source population (the subantarctic population) and at least two samples from the focal population (the contemporary and historic South Island population). The method assumes migration is constant, that all sources are sampled and that the source population is sufficiently large that allele frequencies are temporally stable (Wang & Whitlock 2003). For all likelihood-based methods, maximum N_e (N_{eMAX}) was set to 1000. Higher values of N_{eMAX} only affected the upper bound of the 95% confidence interval (CI) when this fell above 1000 in MLNE, but never influenced point estimates or the lower bound of the CI. For TMVP analyses, increasing N_{eMAX} only lead to marginal increases of the upper bound of the CI (data not shown).

Table 5.1 Calculation of the generation time for *M. antipodes* using the formula $\Sigma(x\ l_x\ b_x)/\Sigma(l_x\ b_x)$, where l_x is the proportion of individuals surviving to age x and m_x is the reproductive output at age x measured as the number of chicks that fledged per individual. The following parameters were used for the calculation of the life table: first year survival = 0.4 (Richdale 1957); adult survival = 0.856 (Richdale 1957); maximum age = 20 (Department of Conservation unpublished data); number of chicks fledged per nest = 1.16 (Darby & Seddon 1990). The proportion of males (m), females (f) and the total proportion of birds (average of m and f) breeding at a specific age is given in the table (data from Richdale 1957).

Age (x)	l_x	Prop. m breeding	Prop. f breeding	Prop. total breeding	b_x	$x\ l_x\ b_x$	$l_x\ b_x$
0	0.4000	0	0	0	0	0.0000	0.0000
1	0.3424	0	0	0	0	0.0000	0.0000
2	0.2931	0.08	0.48	0.28	0.1624	0.0952	0.0476
3	0.2509	0.55	0.96	0.755	0.4379	0.3296	0.1099
4	0.2148	0.88	1	0.94	0.5452	0.4684	0.1171
5	0.1838	1	1	1	0.58	0.5331	0.1066
6	0.1574	1	1	1	0.58	0.5476	0.0913
7	0.1347	1	1	1	0.58	0.5469	0.0781
8	0.1153	1	1	1	0.58	0.5350	0.0669
9	0.0987	1	1	1	0.58	0.5152	0.0572
10	0.0845	1	1	1	0.58	0.4900	0.0490
11	0.0723	1	1	1	0.58	0.4614	0.0419
12	0.0619	1	1	1	0.58	0.4309	0.0359
13	0.0530	1	1	1	0.58	0.3996	0.0307
14	0.0454	1	1	1	0.58	0.3683	0.0263
15	0.0388	1	1	1	0.58	0.3378	0.0225
16	0.0332	1	1	1	0.58	0.3085	0.0193
17	0.0285	1	1	1	0.58	0.2805	0.0165
18	0.0244	1	1	1	0.58	0.2543	0.0141
19	0.0208	1	1	1	0.58	0.2297	0.0121
20	0.0178	1	1	1	0.58	0.2070	0.0104

Results

All 350 contemporary samples amplified at all 12 microsatellite loci with the exception of six samples from the South Island, which have missing genotypes for one (three samples), three (two samples) or five loci (one sample), respectively. Of the 35 historic samples, DNA was successfully extracted from 27 samples and a total of 249 genotypes were scored at ten loci (historic samples did not amplify at loci Man22 and Man27). Eight historic samples had missing genotypes at one (four samples), three (one sample), four (two samples) and six loci (one sample) respectively. Allelic dropout was encountered in 16 out of 224 PCR amplifications of confirmed heterozygous historic samples. These 16 cases of allelic dropout were restricted to four of the 27 historic samples, with most instances occurring multiple times in replicate amplifications of the same locus (e.g. for one sample, dropout was observed in five out of seven replicate amplifications of locus Man47). The amplification of a false allele was detected in just one out of a total of 634 successful PCRs.

Disequilibrium and genetic diversity

There was no evidence for linkage disequilibrium between any pairs of loci, and no loci showed significant departure from Hardy-Weinberg proportions. Eight out of ten loci were polymorphic in the contemporary South Island *M. antipodes* samples, and these same eight loci showed variation in the historic samples. In contrast, all ten loci were polymorphic in the subantarctic population. Genetic diversity estimators were slightly lower historically compared to estimates from contemporary samples of the South Island population, but these differences were not significant (all p-values > 0.05, Table 5.2). Allele frequencies of contemporary and historic South Island, and contemporary subantarctic yellow-eyed penguins can be found in Appendix 9.7.

Effective population size

Point estimates of the harmonic mean of N_e for the South Island population of *M. antipodes* varied between 128 and 656 (T = 14) for the different methods applied (Table 5.3). Wang's pseudo-likelihood method gave the highest estimate with a large CI of which the upper bound was limited by our setting of N_{eMAX} =1000. The moment estimator (Waples 1989) gave a slightly lower point estimate and similar to Wang's estimator the CI were large (note that an upper bound cannot be set for the moment estimator). N_e estimates from Beaumont's (2003) likelihood-based method and the joint estimator of N_{eOPEN} and m from Wang and Whitlock (2003) were similar with highly congruent CIs. The N_e estimates are relatively robust to the number of generations (T) between sampling periods, showing only slight increases in N_e with increasing T (Table 5.3). Estimates of m were low (m = 0.003, CI 0.002 – 0.007) and consistent for different values of T (Table 5.3).

Table 5.2 Genetic diversity at ten microsatellite loci in contemporary and historic *M. antipodes*. L_{poly} = number of polymorphic loci, A/locus = mean number of alleles per locus, $A_{richness}$ = allelic richness, H_E = expected heterozygosity, H_O = observed heterozygosity.

Location	N	L_{poly}	A/locus	$A_{richness}$	H_E	H_O
South Island						
Contemporary	249	8	3.0	2.5	0.38	0.37
Historic	27	8	2.2	2.2	0.36	0.33
Subantarctic	101	10	4.5	3.7	0.47	0.45

Table 5.3 Effective population size estimates (N_e) and their confidence intervals (CI) for South Island *M. antipodes*, estimated using four different temporal methods. T = number of generations passed, m = migration rate.

	Estimated N_e (95% confidence interval)				
				Wang & Whitlock (2003)	
T	Waples (1989)	Beaumont (2003)	Wang (2001)	N_e	m
12	237	97	576	184	0.003
	(77 – 1141)	(55 – 405)	(200 – >1000)	(85 – 390)	(0.002 – 0.007)
14	277	124	656	196	0.003
	(90 – 1331)	(67 – 504)	(228 – >1000)	(92 – 431)	(0.002 – 0.007)
16	317	144	737	200	0.003
	(103 - 1521)	(73 – 559)	(255 – >1000)	(101 – 448)	(0.002 – 0.006)

Discussion

Effective population size estimates of South Island yellow-eyed penguins

Using microsatellite DNA analyses of historic (1888 – 1938) and contemporary samples we estimate the harmonic mean N_e of South Island *M. antipodes* between 128 and 656 with lower bounds of the CI varying between 67 and 228 and upper bounds between 431 and >1000. The evaluation of CIs in addition to point estimates of N_e is essential, because CIs generated by different analytical methods are often more consistent than point estimates (Fraser *et al.* 2007a). Additionally, the lower bound of the CI gives important insight into the status of a population with respect to critical conservation thresholds (Hansen *et al.* 2002). In the current study, the four methods applied to estimate N_e varied in their point estimates and their CIs (Table 5.3), a finding which raises questions about the relative accuracy of the different techniques.

Moment estimators such as the estimator from Waples (1989) are known to overestimate N_e and have low precision (resulting in large CIs), particularly when populations experience rapid genetic drift and allele frequencies are skewed (Wang 2001; Berthier *et al.* 2002; Palstra & Ruzzante 2008). Furthermore, the bias of this estimator seems to increase with increasing generations between samples (Tallmon *et al.* 2004). In contrast, simulations have shown Beaumont's estimator (Beaumont 2003) to be very accurate when as many as 10 generation have passed between samples (Tallmon *et al.* 2004). Our estimate of N_e from Beaumont's method, and in particular the associated CI (N_e = 124, CI 67 – 504), was very similar to the joint estimator (N_{eOPEN} = 196, CI 92 – 431) of Wang and Whitlock (2003), while Wang's N_e estimate (Wang 2001) was three times larger with an upper bound of the CI above 1000. The N_{eOPEN} estimator from Wang and Whitlock (2003) is considered to be superior to the closed population estimators ($N_{eCLOSED}$) and

expected to give more realistic values of N_e, because it relieves the assumption of no migration. Furthermore, the more confined CI of N_{eOPEN}, and its consistency with Beaumont's estimator (which is considered reliable with $T > 10$), suggests that our N_e estimates from these two methods are more accurate than those obtained from the moment or Wang's (2001) method.

The effect of migration on N_e is complex and should be addressed cautiously (Wang & Whitlock 2003; Fraser *et al.* 2007a; Palstra & Ruzzante 2008). Ignoring immigration can lead to either upward or downward biases of N_e depending on 1) the extent of gene flow, 2) the sampling interval and 3) the genetic differentiation between focal versus source population(s) (Wang & Whitlock 2003; Fraser *et al.* 2007a). Nevertheless, many studies have found $N_{eOPEN} < N_{eCLOSED}$, particularly in cases where spatial genetic structuring is weak or moderate, and associated migration rates (sometimes unrealistically) high (Fraser *et al.* 2007a). Interestingly, our study also found $N_{eOPEN} < N_{eCLOSED}$ (with the exception of Beaumont's estimator), despite the fact that *M. antipodes* has relatively low migration rates (CI 0.002 - 0.007; see also Boessenkool *et al.* 2009b, Chapter 4). Note that the m estimate from Wang & Whitlock applies to the population's effective size while Boessenkool *et al.*'s estimate (1.6%) applies to the population's census size. The observation that $N_{eOPEN} < N_{eCLOSED}$ — despite these low migration rates and strong population structuring (Boessenkool *et al.* 2009b) — further emphasizes the complex interaction between N_e and m, and reiterates that our understanding of the influence of m on N_e is currently incomplete (Fraser *et al.* 2007b).

The harmonic mean N_e is weighted towards the smallest values of N_e during the sampling interval in fluctuating populations (Leberg 2005). Close monitoring of yellow-eyed penguins on the South Island has revealed strong fluctuations in the total number of breeders during the last two decades, with the lowest population estimate recorded in 1990/1991 season when as few as 140 pairs bred on the South Island (Gill & Darby 1993), versus approximately 500 breeding pairs on the South Island in more recent years (Department of Conservation, unpublished data). Assuming N_e is lower in years with low

numbers of breeders, it is likely that our estimate of N_e is an underestimate of the contemporary N_e.

It is difficult to calculate the N_e/N_c (effective population size/census population size) ratio in *M. antipodes* because we cannot calculate the harmonic mean N_c over the time interval used to calculate the harmonic mean N_e. Dividing our best estimate of N_e (124 or 196) by the current census size (~2200 for the total South Island population, including surrounding islands) gives a ratio of 0.06 or 0.09, but this will probably be a slight underestimate of the actual ratio as the harmonic mean N_c over the time interval is likely to be less than the current census size. Natural populations of vertebrate taxa typically have N_e/N_c ratios of 0.10 – 0.11 on average and fluctuating population size is the most important factors reducing this ratio (Frankham 1995). The second most important factor leading to low N_e/N_c ratios is variance in reproductive success (Frankham 1995). Such variance has been shown to exist in yellow-eyed penguins, and parental 'quality' is likely an important component determining this variation (Efford & Edge 1998; Bull 2005). Our estimates of the N_e/N_c ratio of South Island *M. antipodes* appear close to the average ratio seen in many wildlife populations (i.e. 0.10). Unfortunately there are currently no comparable estimates of N_e/N_c for any other penguin species. Estimates of N_e for Galápagos and Magellanic penguins were calculated by Akst *et al.* (2002) but the method used to obtain these estimates is not considered reliable (e.g. this method results in N_e estimates of 3000 (data not shown) in *M. antipodes* which is higher than the current census size).

Conservation implications

The minimum N_e required to retain sufficient evolutionary potential is thought to approximate 500, although thresholds as high as 5000 have been proposed (Franklin 1980; Franklin & Frankham 1998; Lynch & Lande 1998). Coping with certain environmental challenges, such as the introduction of

disease and toxins, may only require an adaptive response at a few specific loci, and the population size needed to maintain sufficient genetic variation at such loci is more likely to lie in the thousands than in the hundreds (Willi *et al.* 2006). Our most reliable N_e estimates for South Island *M. antipodes* are well below such critical thresholds required to maintain adaptive potential. This finding is particularly notable in the context of the regular disease epidemics experienced by this population. Furthermore, South Island *M. antipodes* already have low genetic diversity compared to the subantarctic population at neutral loci (Boessenkool *et al.* 2009b, Chapter 4). Given predicted increases in rates of environmental variations due to climate change (NIWA 2008), the maintenance of adaptive genetic diversity in *M. antipodes* may become increasingly important. Ongoing monitoring of the population, in addition to continued predator trapping, is therefore essential.

Acknowledgements

We are very grateful to the Auckland Museum, American Natural History Museum, Australian Museum, Canterbury Museum, Museum of Comparative Zoology, Natural History Museum Geneva, Natural History Museum Tring, Natural History Museum Paris, Museum of New Zealand Te Papa Tongarewa, Natural History Museum Vienna, Swedish Museum of Natural History, Otago Museum, South Australian Museum, Smithsonian Institution, Museum für Naturkunde Berlin and C. Millar for supplying tissue samples of historic specimens. We thank the New Zealand Department of Conservation for help with collecting contemporary samples. We are indebted to Tania King for guidance and advice in the laboratory and we thank Stein Are Sæther for helpful discussions about analyses. This research was supported by the Department of Zoology, University of Otago, including PBRF Research Enhancement Grants to PJS and JMW. Samples were collected under Department of Conservation permits SO-17933-FAU and OT-19097-RES and University of Otago Animal Ethics Approval 69/06.

Chapter 6:

Genetic analyses suggest fraudulent origins of historic museum penguin specimens

This chapter has been submitted to Proceedings of the Royal Society of London B: Biological Sciences as:

Sanne Boessenkool, Bastiaan Star, R Paul Scofield, Philip J Seddon & Jonathan M Waters. Genetic analyses suggest fraudulent origins of historic museum specimens.

Abstract

Historic museum specimens are increasingly used to answer a wide variety of questions in biological research. However, the scientific value of these specimens depends on their authenticity. Here we use individual based genetic analyses in order to demonstrate historic falsification of archive specimen data from the late 19th Century. Specifically, using ten microsatellite markers, we analysed 350 contemporary and 43 historic yellow-eyed penguin (*Megadyptes antipodes*) specimens from New Zealand's South Island and subantarctic regions. Factorial correspondence analysis and an assignment test strongly suggest that eight of the historic specimens purportedly of subantarctic origin were in fact collected from the South Island. Interestingly, these eight specimens were all collected by the same collector, and all are currently held in the same museum collection. Further inspection of the specimen labels and evaluation of the collector's voyages did not reveal whether any accidental mistakes have been made or whether deliberate falsification was at play, but it seems clear that falsification of the specimens' geographic details would have increased their monetary value. This study represents a novel extension to the well-known applications of assignment tests in molecular ecology. In addition, our results serve as a warning to all who use archive specimens to invest time in the verification of specimen data.

Introduction

Museum collections, archived in natural history museums worldwide, provide invaluable resources of materials and knowledge that are of utmost importance to science and society (Suarez & Tsutsui 2004). These collections currently hold an estimated total of 3 billion specimens, of which 7-10 million are bird skins (Brooke 2000; Pennisi 2000). The crucial role of museum collections in defining species and their ranges started with the diligent efforts of 19th century collectors, and even today this wealth of information still plays a vital role in the documentation of species decline and conservation status assessment (Shaffer *et al.* 1998; Collar & Rudyanto 2003). During the last two decades, the use of museum specimens as sources of DNA samples has been facilitated by advances in molecular techniques, initiating a vast increase in the use of such archived specimens in population and evolutionary genetic studies (reviewed in Wandeler *et al.* 2007). Comparative studies now frequently compare levels of genetic diversity over time, thereby inferring changes in population size and population connectivity (e.g. Miller & Waits 2003; Johnson *et al.* 2004; Larsson *et al.* 2008; Taylor *et al.* 2008).

The potential problems arising from working with low quality DNA from historic specimens are well known and can be addressed using clear laboratory guidelines (Sefc *et al.* 2003; Pääbo *et al.* 2004; Sefc *et al.* 2007), but any additional pitfalls arising from errors in specimen data are potentially much more complex (reviewed in Rasmussen & Prŷs-Jones 2003). Specimen data are recorded on attached labels (and collector's notes when available) and the scientific value of specimens ultimately depends on the accuracy of these data. The minimum information typically associated with a specimen includes the identity, location, collection date and the name of the collector, but one or more of these entries may be missing. Additionally, these labels are also the most prone to error in specimen

collection (Winker 2000). Most errors are found in identity and location, which can lead to false representation of a species' distribution (Graham *et al.* 2004). Causes of inaccuracy in museum specimen data vary from simple mistakes or carelessness during collection or post-collection to serious cases of fraud (Rasmussen & Prŷs-Jones 2003). Neither inadvertent mistakes nor cases of deliberate fraud have received much attention in the literature, with a few exceptions such as the case of deceit by British Colonel Richard Meinertzhagen (Knox 1993; Rasmussen & Prŷs-Jones 2003; Dalton 2005).

Detecting errors in specimen data can be extremely challenging. Museum staff typically use collectors' field notes, information related to the voyages and travels of collectors and thorough examination of preparatory techniques (including X-rays) to identify errors (Knox 1993; Rasmussen & Collar 1999; Rasmussen & Prŷs-Jones 2003). An approach using geo-referencing of temporally collected samples was introduced to detect specimens with high probability of error without *a priori* suspicion (Peterson *et al.* 2004). Nevertheless, the above methods are limited by their focus on 1) specimens from suspicious collectors, 2) specimens that form outliers with respect to the species' natural range, or 3) specimens that form outliers with respect to collection date (e.g. collected years after a species was reported extinct or collected on dates that clash with collector's itineraries). When specimens do not fall into any of the above categories, error or fraud detection becomes near impossible. In the current study we present an unusual case in which individual based genetic analyses reveal previously unsuspected inaccuracies in the geographic origin of museum material. Specifically, our data suggest a case of possible fraud involving eight yellow-eyed penguin (*Megadyptes antipodes*) specimens purportedly from New Zealand's subantarctic islands. This detection not only presents a novel approach to detect errors in archive specimen data, but additionally implies that errors in museum collections are likely to be more abundant than previously anticipated.

M. antipodes was first described from an Auckland Island's specimen in 1841 (Hombron & Jacquinot 1841). The species is endemic to the New Zealand

region, where it inhabits the subantarctic Auckland and Campbell Islands, along with the southeast coast of South Island (Marchant & Higgins 1990; McKinlay 2001; figure 6.1). Genetic and morphological analyses of sub-fossil and historic specimens have shown that *M. antipodes* likely expanded its range from the subantarctic islands to the South Island of New Zealand after ~1500 AD, following the anthropogenic extinction of its sister species *M. waitaha* (Boessenkool *et al.* 2009a, Chapter 2). Despite the recent expansion event, the presence of significant microsatellite DNA structuring — and inferred low migration rates — among contemporary breeding sites support the genetic recognition of two separate populations, one on South Island and the other in the subantarctic (figure 6.1; Boessenkool *et al.* 2009b, Chapter 4). In the present study we analysed contemporary and historic museum specimens from both the South and subantarctic Islands using ten microsatellite markers. Our data strongly suggest that eight of the historic specimens believed to be of subantarctic origin were in fact collected from South Island.

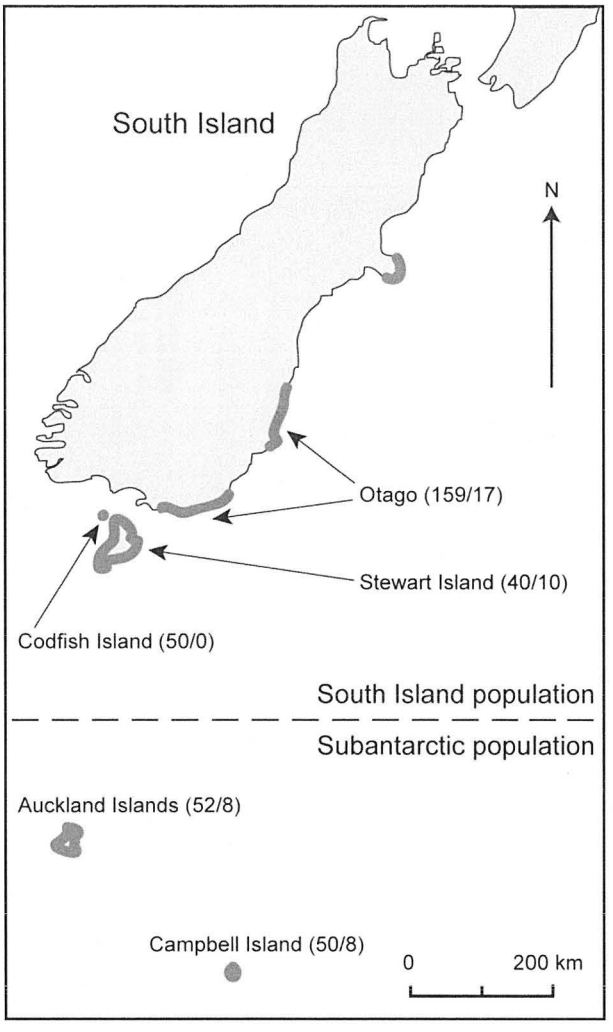


Figure 6.1 Map of the South and subantarctic islands of New Zealand. Arrows point to the geographical locations where samples were (purportedly) collected. Sample sizes for contemporary/historic samples are given in brackets. The dashed line refers to the split between the South Island and the subantarctic population.

Material and Methods

Sample collection, DNA extraction and genotyping

Yellow-eyed penguin blood samples were collected in 2005-2007 on the South Island (N = 249) and subantarctic Auckland and Campbell Islands (N = 101) of New Zealand (figure 6.1) as described in Boessenkool *et al.* (2009b, Chapter 4). DNA was extracted and purified using 40 µg proteinase K in 5% Chelex (Biorad; Walsh *et al.* 1991).

Historic toe pad samples were obtained from 55 specimens collected between 1840 and 1944 on the South Island (N = 35) and subantarctic Auckland and Campbell Islands (N = 20). These specimens are held in 15 museum collections around the world (Appendix 9.8). Following rehydration, toepad samples were finely chopped and DNA was extracted using the Chargeswitch Forensic DNA Purification Kit (Invitrogen) or the DNeasy Tissue Kit (Qiagen) following manufacturers' instructions.

All samples were genotyped at 10 microsatellite loci previously developed for yellow-eyed penguins (Mano3, Mano8, Man13, Man21, Man39, Man47, Man50, Man51, Man54, Man55; Boessenkool *et al.* 2008, Chapter 3). Microsatellite primer sequences and polymerase chain reaction (PCR) conditions for modern samples are described in Boessenkool *et al.* (2008, Chapter 3). PCR reactions for historic samples were performed in 10 µl volumes containing 2 µl DNA, 0.5 µM of each primer, 0.5 U *Taq* DNA polymerase (Mango *Taq*, Bioline), 1× *Taq* buffer, 0.8 µM dNTP and 1.5 µM $MgCl_2$, with the addition of betaine and DMSO (1.1 M and 2% respectively) if necessary (see Boessenkool *et al.* 2008, Chapter 3). The amplification profile was 2 min at 94°C, 35-50 cycles of 15 s at 96°C, 15 s at 45-50°C and 30 s at 72°C, followed by a 4 min final extension at 72°C.

Strict guidelines were followed in order to prevent contamination of historic DNA and to minimise the risk of erroneous genotypes due to allelic dropout and the amplification of false alleles (Taberlet *et al.* 1996; Sefc *et al.* 2003). DNA extractions and PCR set-up of historic samples were performed inside a UV hood in a separate laboratory where no contemporary yellow-eyed penguin DNA or vertebrate PCR products have ever been present. Historic samples were extracted in small batches of nine samples and potential contamination was monitored by negative extraction and PCR controls. A subset of historic samples (N = 4) from the subantarctic were re-extracted and genotyped to validate results. Secondary extract genotypes from three of these re-extracted samples agreed with those from primary extracts. The fourth re-extracted sample had failed to amplify successfully for the primary extraction. For all samples, 2-7 successful amplifications were obtained before a genotype was scored, and genotypes were only confirmed once every allele was observed at least twice.

Genetic analyses

Deviations from Hardy-Weinberg proportions and linkage equilibrium were assessed separately for modern South Island and modern subantarctic samples using GENEPOP 4.0 (Rousset 2008). Markov chain parameters employed 10 000 dememorizations, 1000 batches and 10 000 iterations. Significance levels were adjusted for multiple comparisons using Bonferroni corrections (Rice 1989). Allele frequencies and unique alleles were evaluated using GENETIX 4.05.2 (Belkhir *et al.* 1996-2004).

To evaluate the provenance of the historic museum specimens we employed two different methods. First, a two-dimensional factorial correspondence analysis (FCA) was performed using GENETIX 4.05.2 (Belkhir *et al.* 1996-2004). An FCA visualises genetic (dis)similarity of individual genotypes without grouping individuals a priori. Second, the assignment test implemented in STRUCTURE 2.2 (Pritchard *et al.* 2000) was used to infer

the probability that historic museum specimens originated from the South Island or the subantarctic populations respectively. This assignment test is a fully Bayesian method that uses geographical sampling location of individuals with confirmed geographic origin as prior information (Pritchard *et al.* 2000). The method assumes that all source populations have been sampled. In our analysis we specified the origin of the contemporary samples to be known, and the proportional membership coefficient (Q) of the historic samples to either of the two populations to be estimated by the program. An important assumption when applying this analysis to historic data is that the allele frequencies of the modern samples are representative of the allele frequencies of the historic populations (see discussion for further comments on this assumption). In the model, allele frequencies were assumed to be correlated among populations and parameters for priors of lambda and F_{ST} were left at default values. For the historic samples we applied the admixture model with a uniform prior for alpha, bounded by a maximum of 10.0, and we set ALPHAPROPSD to 0.025. The migration prior (v) for the assignment test was set to 0.01, but to account for uncertainty in v we ran replicate analyses using $v = 0.05$ and $v = 0.1$. The outcome of the analyses was unaffected by the migration prior and we only present the results from runs with $v = 0.01$. The MCMC simulation was performed with a burnin of 100 000 followed by 500 000 iterations.

Results

DNA was successfully extracted from 43 of the 55 historic samples. Twenty-seven of the successful extractions were from 'South Island' specimens, whereas 16 were from purportedly 'subantarctic' specimens (Campbell and Auckland Islands; figure 6.1). Thirteen historic samples had missing genotypes at one (six samples), two (one sample), three (two samples), four (two samples), five (one sample), or six loci (one sample). The amplification of a false allele was encountered on one out of a total of 1066 successful PCRs. Allelic dropout was detected in 36 out of 420 PCR amplifications of

confirmed heterozygous historic samples. In 26 out of the 36 cases dropout occurred in multiple replicate amplifications of samples at specific loci.

All 350 modern samples amplified at all microsatellite loci with the exception of six samples from the South Island, which had missing genotypes for one (three samples), three (two samples), or five loci (one sample), respectively. The data set revealed no evidence for linkage disequilibrium between any pairs of loci nor significant departure from Hardy-Weinberg proportions in the modern *M. antipodes* South Island and subantarctic populations.

Evaluation of the origin of historic samples

The two-dimensional factorial correspondence analysis (FCA) illustrates the clear genetic distinction between modern South Island versus subantarctic populations (figure 6.2). To evaluate the origin of the historic samples, we superimposed their genotypes over the modern samples. All historic samples which were reportedly collected on the South Island clustered genetically among the modern South Island samples, consistent with their geographic origins. In contrast, only eight of the sixteen historic 'subantarctic' samples were grouped with the modern subantarctic samples, whereas the remaining eight samples clustered among the modern South Island samples (figure 6.2).

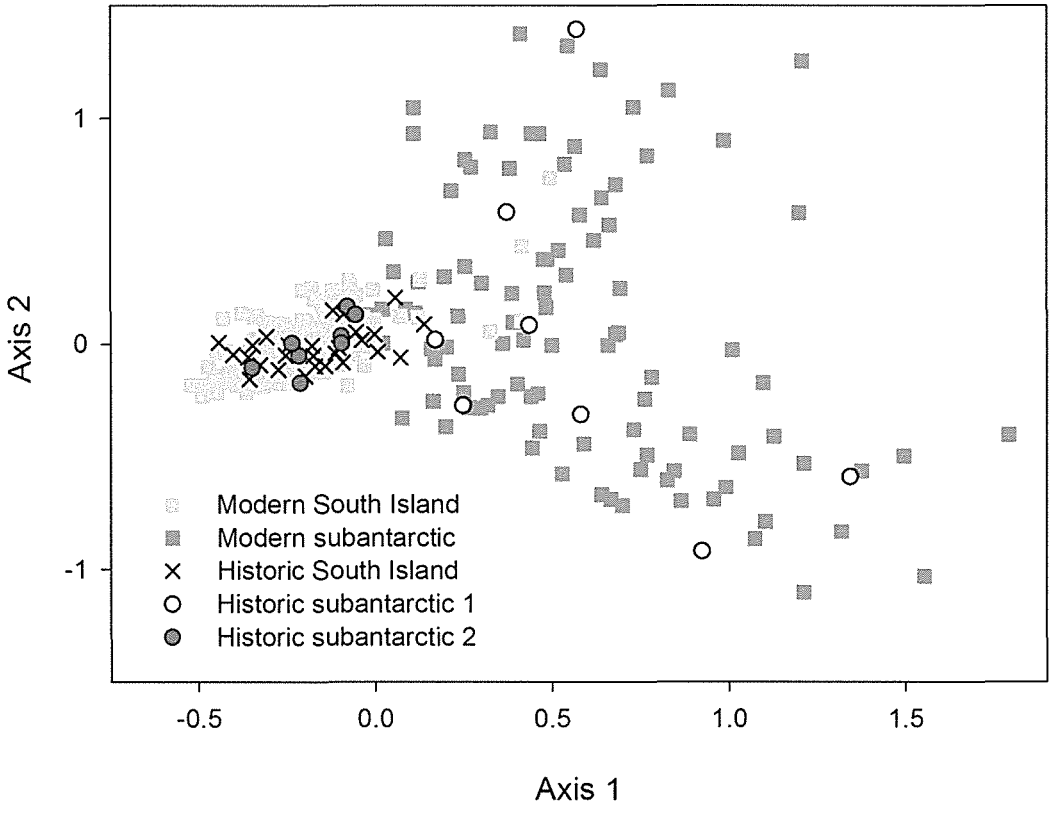


Figure 6.2 Plot of the two-dimensional factorial correspondence analysis based on genotypic variation at ten microsatellite loci of modern and historic *M. antipodes* samples. The axes explain 14.7% of the total variation.

Using the assignment test implemented in STRUCTURE we estimated the proportional membership coefficient (Q) for each of the historic samples to both the South Island and the subantarctic populations (figure 6.3). Of the 27 historic samples with South Island origin, 20 have high Q (i.e. $Q > 0.80$) to the South Island population. The other seven samples do not assign strongly to either of the populations; their Q-values lie between 0.20 - 0.80 indicating that they have mixed ancestry (Lecis *et al.* 2006; Bergl & Vigilant 2007). Note that of these individuals with mixed ancestry, two have missing data at four loci. Of the historic samples with purported subantarctic origins, eight have high Q to the subantarctic population ($Q > 0.80$), whereas seven show strong membership to the South Island population ($Q > 0.9$ for four samples and $0.80 < Q > 0.90$ for three samples respectively). One sample has weak

evidence for mixed ancestry (Q = 0.744 to the South Island population), but this individual lacked genotypic data at five of ten loci. The eight subantarctic samples that have strong membership to the South Island population are the same eight samples that were placed among the modern South Island samples in the FCA. Four of these eight specimens were reportedly collected on the Auckland Islands (three in 1893, one in 1894) and the other four were reportedly collected on Campbell Island (two in 1893, two in 1894). Interestingly, these eight specimens were all collected by the same collector, namely Henry Hamersley Travers (1844-1928), and are currently held in the same museum collection (American Natural History Museum).

The modern subantarctic population has 18 unique alleles, whereas only one unique allele was detected in the modern South Island population (see also Boessenkool *et al.* 2009b, Chapter 4). It is particularly noteworthy, therefore, that the eight historic samples supposedly collected from the subantarctic, but genetically categorised as South Island specimens (figures 6.2 and 6.3), do not possess any of the alleles unique to the subantarctic. In contrast, five of the eight historic samples with confirmed subantarctic origin possess a total of nine unique subantarctic alleles. The probability of recovering zero unique alleles (by chance) in our sample of eight suspect individuals would appear to be very low. We investigated this probability further by calculating the probability distribution of sampling unique alleles (figure 6.4), assuming these eight suspect individuals were randomly sampled from the subantarctic population. We simulated random resampling of the genotypes (allowing for the missing data present in the original sample) of the suspect eight individuals using probabilities based on the allele-frequency vectors of the modern subantarctic population and recorded the number of unique alleles in our sets of resampled genotypes. Our simulation confirmed that the probability of having zero unique alleles is extremely low (figure 6.4) and over 99.99% of our random draws resulted in genotype sets with one or more unique alleles.

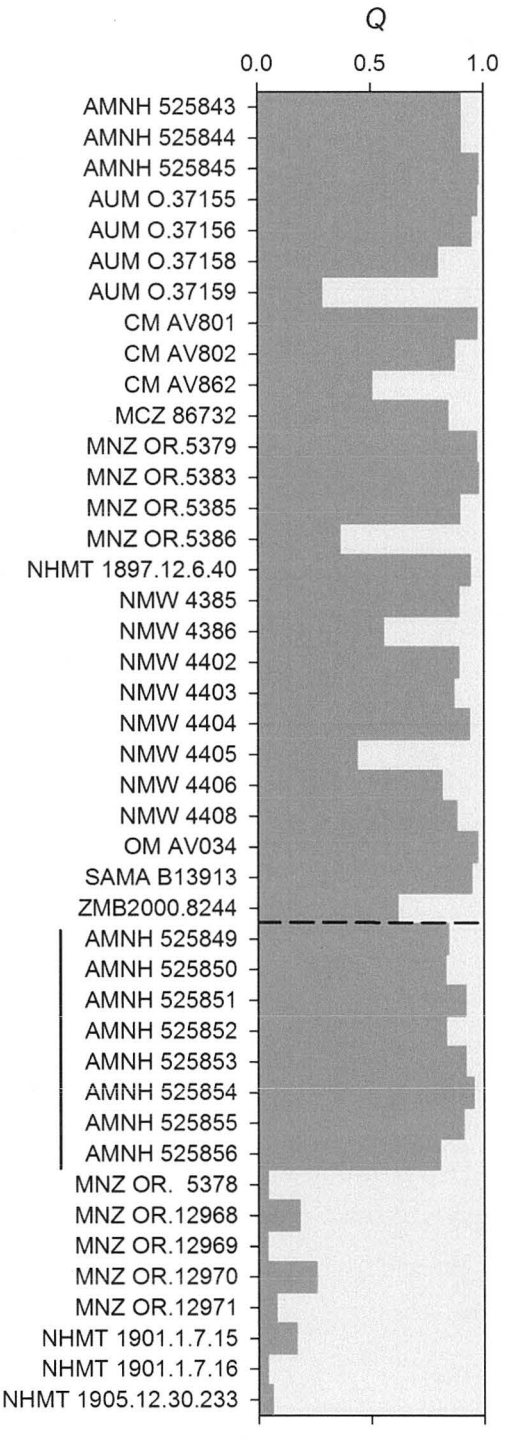


Figure 6.3 Proportional membership (*Q*) of historic *M. antipodes* specimens to the South Island (dark grey) and subantarctic (light grey) populations as estimated using the assignment test in STRUCTURE. Each horizontal bar represents a single specimen identified by its museum accession number and specimens are ordered by reported geographic sampling location and museum collection. Specimens above the dashed line where purportedly collected on the South Island, those under the dashed line on the subantarctic Auckland and Campbell Islands. Museum abbreviations can be found in Appendix 9.8.

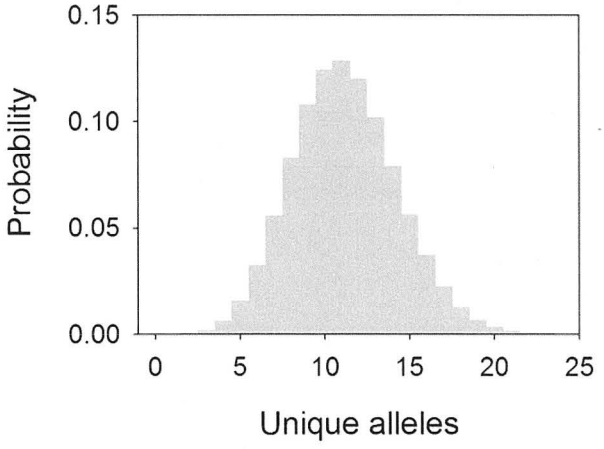


Figure 6.4 Probability distribution of sampling unique alleles in the genotype sets of eight individuals from the modern subantarctic population. The genotypes of the eight individuals were resampled (100 million iterations) based on the alleles-frequency distribution of the current subantarctic population, and the number of alleles that are unique to this population was recorded each iteration.

Discussion

Fraudulent origins of historic Megadyptes *specimens?*

Individual-based genetic analyses and evaluation of unique alleles reveal that eight historic *M. antipodes* museum specimens have incorrect specimen data with respect to geographic collection location. We suggest that these eight specimens were not collected on the subantarctic Auckland and Campbell Islands, as is stated on their specimen labels, but were in fact collected on the South Island of New Zealand.

The historic expansion of *M. antipodes* (Boessenkool *et al.* 2009a; Chapter 2) and the current pattern of migration (Boessenkool *et al.* 2009b; Chapter 4) both indicate an asymmetric pattern of rare dispersal in this species, involving migration northwards from the subantarctic to the South Island. It is not surprising, therefore, that a few South Island individuals show mixed

genetic ancestry (figure 6.3; see also Boessenkool *et al.* 2009b; Chapter 4). On the other hand, the single batch of eight purportedly 'subantarctic' birds that genetically have South Island origins, clearly conflicts with observed dispersal patterns for *M. antipodes*. Perhaps most suspiciously, all eight of these suspect birds are attributed to the same collector who purportedly obtained them within two years of one another.

The eight suspect specimens (see above) were obtained by the American Natural History Museum (AMNH) from Lionel Walter Rothschild the 2nd Baron Rothschild in 1932 (Rotschild 1983). The specimens have two different labels attached to them: one collector label which appears to be the original label from the collector Henry Hamersley Travers, and one Rothschild collection label (M. LeCroy, pers. comm.). Both labels state the collection location (i.e. Auckland or Campbell Islands) and date (i.e. 1893 or 1894). The original H.H. Travers labels are signed, and the Rothschild labels state "H.H. Travers coll." One specimen solely has a H.H. Travers label with "Rothschild coll." written on the back (M. LeCroy, pers. comm.). Based on this labelling, it appears that the specimens were obtained directly by Lord Rothschild from H.H. Travers without intervening dealers. Overall, the inspection of the specimen labels does not reveal any clues on whether any mistakes have been made on the labels or whether labels have been changed post-collection.

Henry H. Travers was a well-known professional collector and taxidermist in 19th century New Zealand (Cyclopedia Company Limited 1897). To date there have been no suspicions regarding Travers' reliability as a collector. Collector's field notes and itineraries can reveal clues that may help evaluate the veracity of specimen information (Rasmussen & Collar 1999; Rasmussen & Prŷs-Jones 2003; Graham *et al.* 2004), but unfortunately Travers' field notes are not readily available. Nevertheless, an assessment of the specific voyages that were made to the subantarctic islands in the 1890s revealed that it is unlikely that Travers' visited the Campbell and/or Auckland Islands in 1893, the date stated on some of the labels. The only registered voyage to the subantarctic in 1893 is one undertaken by the New Zealand government

steamer *Hinemoa* to the Antipodes Islands to rescue eleven survivors that were shipwrecked there (Headland 1989). It is certain, however, that Travers visited the Snares, Auckland and probably Campbell Islands in 1894, where he reportedly collected a large number of zoological specimens (Anonymous 1895; Headland 1989). Nevertheless, the above data suggest that a minimum of five of the eight 'suspect' collection dates are incorrect.

Monetary gain is considered the most evident motivation for the deliberate falsification of specimen information and cases within ornithology provide some well known examples of major specimen fraud (Rasmussen & Prŷs-Jones 2003). Many collectors, including H.H. Travers, made a living through the dealing of specimens. Since a specimen's value would likely increase according to its rarity or scarcity (Rasmussen & Prŷs-Jones 2003), falsification of a collection location could potentially increase the collector's income and perhaps also boost his/her reputation. While *M. antipodes* is not rarer on the subantarctic islands than on the South Island, specimens from geographically remote locations such as the subantarctic islands would nonetheless have been more valuable than readily available specimens from the South Island. The path from initial suspicion of specimen fraud to conclusive proof is, however, a long one (Rasmussen & Prŷs-Jones 2003). While the current results do not provide conclusive proof of deliberate fraudulence, they strongly suggest that H.H. Travers misrepresented the collection location for his *M. antipodes* specimens. Further investigation of Travers' specimens may shed more light on the possibility of repeated dishonesty.

Genetic analysis as a means to detect errors in specimen locality data

The detection of errors in archive museum specimen information can be an arduous and time-consuming process that often relies on the availability of historical data such as collector's field notes. The development of molecular techniques has now facilitated the use of DNA in this process. Genetic data

have successfully been used to detect sexing errors in museum bird skins (Lee & Griffiths 2003; Bantock *et al.* 2008) and to verify the identity of species (e.g. Hennache *et al.* 2003) or even eggs (Lee & Prŷs-Jones 2008). In this study we present a novel approach for using individual based population genetic analysis to detect errors in the locality data of archive museum specimens.

The use of individual based genetic analyses, including assignment tests, to identify an organism's geographic origin is well known (Waser & Strobeck 1998; Manel *et al.* 2005). The extension of these approaches to the verification of locality data of archive specimens is a promising new direction in museum science. This method, however, is not without its limitations. First, the reliability of the method relies on the availability of specimens with confirmed geographic origin that can be used to calculate population allele frequencies. Often such historic samples are not available, so one relies on the assumption that allele frequencies of modern samples are representative of the allele frequencies of the historic populations. In our study we fortunately had good knowledge of the study system, including the colonisation history and dispersal patterns of *M. antipodes*, and we can therefore be confident that this temporal assumption was not violated. Nevertheless, when populations are known to have suffered severe bottlenecks, for example, verifying the validity of this assumption will be challenging in the absence of historic specimen data. Second, successful assignment of specimens is contingent on the existence of sufficient genetic structuring among populations. If limited levels of dispersal exist, the method can still be applied to detect errors in a sample of multiple specimens which were collected by the same collector or for example on the same voyage, but the detection of single misinformed specimens will likely be unreliable. If researchers aim to verify the origin of a single specimen, the guidelines already developed for the use of assignment tests in the detection of wildlife poaching should be applicable (see Manel *et al.* 2002). Importantly, high thresholds (e.g. a probability of 0.999 that an individual belongs to a specific population) would have to be satisfied before a specimen can be confidently

assigned to a specific population (Manel *et al.* 2002). Furthermore, care has to be taken when choosing an assignment method, as some techniques require sampling of all potential source populations (Pritchard *et al.* 2000; Manel *et al.* 2002; Piry *et al.* 2004). The current study underlines the value of combining distinct genetic approaches to improve the reliability of error detection.

The approach we propose in the current study should be regarded complementary to the methods that are currently being applied for error detection in specimen data (see Knox 1993; Rasmussen & Collar 1999; Rasmussen & Prŷs-Jones 2003; Peterson *et al.* 2004). The strength of the genetic method, however, is that potential errors can be detected without *a priori* suspicion. With the continuously increasing number of population and evolutionary genetic studies that use archive specimens (Wandeler *et al.* 2007) there is considerable scope to apply the methods we outline here.

Finally, the results of the present study serve as a warning to all those who use museum specimens in population genetic studies. Especially when working on endangered species or populations, the reliability of specimen locality data is essential. Mistakes in locality data can among others affect the inferences we make when estimating historical population connectivity, defining conservation units, estimating effective population sizes or designing reintroductions. Researchers are encouraged to invest in the verification of specimen data to ensure that archive specimens remain a valuable resource for many years to come.

Acknowledgements

We are very grateful to the Auckland Museum, American Natural History Museum, Australian Museum, Museum of Comparative Zoology, Natural History Museum Geneva, Natural History Museum Tring, Natural History Museum Paris, Museum of New Zealand Te Papa Tongarewa, Natural History

Museum Vienna, Swedish Museum of Natural History, Otago Museum, South Australian Museum, Smithsonian Institution, Museum für Naturkunde Berlin and C. Millar for supplying tissue samples of historic specimens. We are greatly indebted Mary LeCroy and Paul Sweet from the American Natural history Museum. We thank the New Zealand Department of Conservation for help with collecting contemporary samples, and Tania King for guidance in the laboratory. This research was supported by the Department of Zoology, University of Otago, including PBRF Research Enhancement Grants to PJS and JMW. Samples were collected under Department of Conservation permits SO-17933-FAU and OT-19097-RES and University of Otago Animal Ethics Approval 69/06.

Chapter 7:

General Discussion

Improving our understanding of the causes that lead to population decline and the processes that affect the dynamics of small populations are vital for the protection and restoration of the world's biodiversity. The application of new genetic tools, supported by the continuing advances in statistical methodology, have now revolutionised our ability to investigate these biological processes. The ability to examine ancient DNA facilitates the direct reconstruction of population histories, which clearly enhances our understanding of the effects of anthropogenic environmental change on population dynamics. In the Pacific, and particularly in New Zealand, human settlement occurred recently and this area has one of the most complete archaeological and palaeontological records in the world. This unique record provides an accessible system for revealing anthropogenic impacts on native biota (Hurles *et al.* 2003). In this thesis, I have investigated the temporal and spatial genetic structuring of the yellow-eyed penguin (*Megadyptes antipodes*), an endangered New Zealand endemic, using genetic tools. Specifically, I have focused on questions directly relating to the conservation of this iconic species.

Yellow-eyed penguins were believed to be more widespread and abundant before human colonisation of New Zealand, thus current management has assumed the mainland population to be a declining remnant of a larger prehistoric population. In contrast, our genetic and morphological analyses of subfossil, historic and modern penguin samples revealed an unexpected pattern of penguin extinction and expansion (Boessenkool *et al.* 2009a, Chapter 2). Specifically, my work has shown that — at the time of human settlement — yellow-eyed penguins were present on the subantarctic Auckland and Campbell Islands, and only occasional vagrants reached the shores of South Island, New Zealand. Mainland shores were, instead, inhabited by the previously unrecognised sister species *M. waitaha*. Only after *M. waitaha* was hunted to extinction around 1500 AD, did *M. antipodes* expand its range northwards to the New Zealand mainland where it is found today. It is important to emphasize the key role of extinction in facilitating biogeographic expansion events, even in highly dispersive species such as

penguins. While abiotic processes such as glaciation are known to drive extinction-recolonization processes in many high-latitude regions (Hewitt 1996; Fraser *et al.* 2009), the current study is one of the first to demonstrate a role for anthropogenic impacts in initiating recolonization.

The recognition of *M. waitaha* as a full species was underpinned by congruent genetic and morphological analyses of modern, historic, and prehistoric *Megadyptes* specimens. The assignment of species status to diverged populations is a highly debated issue in the scientific literature (see for example Helbig *et al.* 2002; Sites & Marshall 2003; Meiri & Mace 2007). The difficulty lies, of course, in the fact that evolution, and hence the divergence of taxa, is a continuous process. Consequently, there is no definitive cut-off which allows unquestionable delineation of species boundaries at shallow phylogenetic levels. The mere presence of morphological and/or genetic differentiation between two allopatric populations may be insufficient to assign species status, and instead one has to take into account the degree of divergence of multiple characters. The designation of full species status to *M. waitaha* was based on a number of criteria: 1) clear qualitative character differences in five different bones, 2) significant morphological size variation, 3) fixed genetic differences at five nucleotides in the mitochondrial control region and 4) reciprocal monophyly in the Neighbour-Joining analyses (phylogenetic species concept). We can furthermore speculate that the two species were reproductively isolated (biological species concept) since their integrity was maintained despite *M. antipodes* migrants arriving on the South Island. Although the designation of species status in closely related taxa will always remain a source for discussion, the above mentioned criteria provide sufficient evidence to assign full species status to *M. waitaha*.

The recognition of a second species within the *Megadyptes* genus reveals a biogeographic pattern that is concordant with distributions of other penguin taxa from the New Zealand region. In particular, *Eudyptes* — the sister taxon of *Megadyptes* — shows evidence of localised speciation associated with

distinct island groups. For example, Fiordland crested penguins (*Eudyptes pachyrhynchus*) inhabit the southern South Island and several smaller surrounding islands (e.g. Stewart Island) of New Zealand, while the Snares penguin (*E. robustus*) is found on the subantarctic Snares Islands only, and the Royal penguin (*E. schlegeli*) breeds only on Macquarie Island (Marchant & Higgins 1990). Although Rockhopper penguins (*E. chrysocome sensu lato*) are found throughout the subantarctic, their populations show strong genetic structuring among either side of the subtropical convergence (de Dinechin *et al.* 2009). The historic distributions of *M. waitaha* and *M. antipodes* indicate a similar pattern, with the South Island versus subantarctic (Campbell and Auckland Islands) populations delineated by the subtropical convergence. Together, these multispecies data strongly support the role of this oceanographic boundary as an important force driving vicariant speciation (Mackintosh 1960; de Dinechin *et al.* 2009).

Considering the strong oceanographic boundary separating mainland New Zealand and the subantarctic, the range expansion of *M. antipodes* northwards across the subtropical convergence provides an unusual scenario. Although it involved a relatively short geographic distance (600 km), this northward expansion event — across the boundary — contrasts dramatically with, for instance, the inferred range expansion of rockhopper penguins, *E. chrysocome sensu lato*, to Amsterdam and Saint Paul Islands in the Indian Ocean. Rockhopper penguins are thought to have colonised Amsterdam and Saint Paul from populations in the southern Atlantic, separated by 6000 km but within the same watermass, rather than from the relatively proximate (1000 km away) source populations in the subantarctic Indian Ocean that lie immediately across the subtropical convergence (de Dinechin *et al.* 2009). The apparent difficulty of traversing the subtropical convergence is exemplified by the low levels of contemporary migration detected among the subantarctic and South Island populations of *M. antipodes* (Boessenkool *et al.* 2009b, Chapter 4). Specifically, genetic assignment analyses of modern samples based on 12 microsatellite loci developed for this species (Boessenkool *et al.* 2008, Chapter 3) detected only two first generation

migrants which had dispersed from the subantarctic to the South Island, suggestion a migration rate of less than 2% (Boessenkool *et al.* 2009b, Chapter 4). Notably, the subantarctic population possesses high diversity relative to mainland samples, including 18 unique alleles not detected on the South Island (Boessenkool *et al.* 2009b, Chapter 4). Taken together, these results suggest that the expansion from *M. antipodes* derived from a small number of individuals that crossed the subtropical convergence and settled in the niche released by the extinction of *M. waitaha* (Boessenkool *et al.* 2009a, Chapter 2).

The successful expansion of *M. antipodes* on the South Island so soon after the extinction of *M. waitaha* is remarkable and may have resulted from a unique combination of favourable environmental conditions. First, I hypothesise that a paradigmatic shift in Māori culture (following the extinction of large game) prevented the hunting of the newly arrived yellow-eyed penguins. This suggested cultural shift is thought to be associated with a temporal decline in the human population of southern New Zealand (Anderson & Smith 1996) and/or the development of conservation awareness and resource monitoring (Anderson 2002). Second, environmental changes such as the extinction of competitors (*M. waitaha*), and a severe decline of predators (e.g. sea lions; Childerhouse & Gales 1998; Lalas *et al.* 2007) may have facilitated the yellow-eyed penguin's settlement on South Island beaches. Finally, and perhaps most importantly, the expansion of *M. antipodes* took place before New Zealand was colonised by Europeans and their commensals. The clearing of some of the important coastal forest habitat, such as that on the Otago Peninsula (currently a stronghold for yellow-eyed penguins on the South Island) occurred only after the arrival of Europeans (see maps in McGlone 1983). Moreover, it appears that the current predation pressure imposed by introduced mammals — chiefly mustelids and cats — was not an issue in the first centuries following *M. antipodes*' expansion. The loss of penguin chicks and eggs due to predation was not noted in the 1940s during intensive monitoring by Lance Richdale (Richdale 1957) and it seems predation has become a significant threat only

in the second half of the 20th century (Darby & Seddon 1990). Overall, this collection of favourable circumstances demonstrates the stochastic nature of this inferred penguin expansion event and suggests that any additional expansions may be unlikely. Similarly, in the event that the South Island population goes extinct in the coming decades, there is no guarantee of another natural replacement event.

The recent history of the South Island yellow-eyed penguin population revealed by the current study may explain, in part, the apparent instability and sensitivity of this population. Since intensive monitoring of *M. antipodes* commenced in the 1980s the number of yellow-eyed penguin nests on the South Island has fluctuated annually from less than 200 to more than 600 (McKinlay 2001). These demographic fluctuations are thought to be caused by changes in food supply (van Heezik & Davis 1990), climatic variations (Peacock *et al.* 2000) and disease epidemics (e.g. Gill & Darby 1993; Department of Conservation unpublished data). Yellow-eyed penguins also show notable sensitivity to disturbance by unregulated tourism (Ellenberg *et al.* 2007). It now appears that this demographic instability and sensitivity of South Island *M. antipodes* may be attributed at least partly to 1) the population's recent founding and hence the lack of a long history of adaptation (Boessenkool *et al.* 2009a, Chapter 2), and 2) the reduced adaptive potential due to low effective population size (Chapter 5) and the low genetic variability as a result of the presumed founding bottleneck (Boessenkool *et al.* 2009a, Chapter 2; Boessenkool *et al.* 2009b, Chapter 4). Furthermore, the low immigration rate from the subantarctic attests to the population's demographic isolation (Chapter 4). Indeed, migration rates may be sufficiently low for the populations to be considered isolated in an evolutionary context. Importantly, consistently low effective population size estimates (<200; Chapter 5) imply that the South Island population will likely continue to lose genetic diversity due to random events, further eroding adaptive potential. Given predicted increases in environmental variability due to climate change (NIWA 2008), this suggests that the South Island population of yellow-eyed penguins will remain vulnerable and unstable.

The analyses of *Megadyptes* genetics presented in this thesis were limited to 'neutral' genetic markers, and did not test for divergence involving loci under selection. It could be argued that genetic diversity at neutral loci does not fully represent genome wide diversity, because balancing selection is known to help retain genetic diversity in adaptive genes of some organisms (Aguilar *et al.* 2004). At the same time, however, environmental stress together with suboptimal habitat and reduced fitness will limit heritability and thus reduce a population's response to selection response (Willi *et al.* 2006). It is therefore difficult at present to gauge the relative levels of adaptive genetic diversity of South Island versus subantarctic *M. antipodes*. Happily, 'next generation' sequencing technologies promise to revolutionise biologists' ability to study diversity at adaptive genes, or even to assess genome wide variation. One specific gene complex that may be particularly fruitful for further study is the major histocompatibility complex (MHC), which plays a crucial role in immune response (e.g. Edwards & Hedrick 1998; Penn *et al.* 2002). Given the disease epidemics observed in the South Island population, comparisons of MHC diversity between South Island and subantarctic yellow-eyed penguins would be of interest. Additionally, MHC analysis of historic museum samples may also shed light on the possible maintenance of variation at these loci during the last century. Finally, research on ecological fitness related traits such as hatching success may reveal divergence in levels of adaptive variation among the two *M. antipodes* populations and give insights into their relative fitness.

The goal of current conservation management of yellow-eyed penguins is to increase the species' population (McKinlay 2001). Until now, the species has been considered to be a single panmictic population, although conservation management efforts have admittedly focused primarily on the South Island population. This mainland focus is not surprising considering the assumption that the penguins on the South Island were a declining remnant from a previously widespread and abundant population. Moreover, the immediately tangible threats to *M. antipodes* persistence are most visible on the South

Island rather than in the relatively pristine subantarctic islands. In light of the present study, however, three important changes have to be made in the current conservation management of this penguin. First, the South Island and subantarctic parts of the distribution have to be considered as two, demographically independent populations and consequently managed as two separate management units (Boessenkool *et al.* 2009b, Chapter 4). Second, protection of the subantarctic population needs to be given high priority since the security of the species seems to depend largely on the health of this original source population. Third, the South Island yellow-eyed penguin population may not increase in numbers or become more widespread within the near future, despite current conservation measures. Therefore, conservation goals should be realistic and aim to maintain the population at its current size. Essentially, maintaining the population at its current size can be considered a management success rather than a failure. In practice, this last point will not reshape any policies with respect to predator trapping or revegetation, which continue to be an absolute necessity for the survival of yellow-eyed penguins on the South Island.

This study on (pre)historic penguins was facilitated by the availability of historic and subfossil specimens, and by the successful extraction of genetic material from these samples. Recent advances in molecular techniques have ensured that museum specimens will continue to provide an invaluable scientific resource for population and evolutionary genetic studies. Unfortunately, however, the scientific value of such specimens is sometimes compromised by the presence of errors in collection data, and the detection of such errors can be an arduous and time-consuming process. This thesis presents a novel approach to detecting errors in specimen data in cases where *a priori* suspicion is absent (Chapter 6). Specifically, individual-based genetic analyses of contemporary and historic microsatellite data were used to detect specimens with erroneous locality information. The finding of eight yellow-eyed penguin specimens with what appears to be fraudulently labelled collection locations was not only surprising but also disturbing. Importantly, errors may be more common than previously suspected. Researchers should

become aware of this hazard and are strongly encouraged to invest in the verification of collection data accompanying the specimens from which their samples originate.

The successful extraction of DNA from subfossil penguin bones holds great promise for further genetic studies on prehistoric material in New Zealand. Almost all the subfossil penguin bones used in the current study were originally excavated from archaeological sites in New Zealand and subsequently stored in a variety of collections. Preservation of DNA in such bones is largely dependent on environmental conditions, with low temperatures and dry conditions being the two most important factors slowing DNA decay (Pääbo *et al.* 2004). The conditions in coastal dune systems where most prehistoric villages were located, and hence the source of our penguin bones, are assumed to be far from ideal for the preservation of DNA. Furthermore, standard museum treatment and storage conditions of excavated material can have detrimental effects on DNA quality, whereas freshly excavated material has been shown to significantly improve the recovery of DNA (Pruvost *et al.* 2007). Despite these potential drawbacks, this study successfully extracted and amplified authentic DNA from penguin bones. Replicate extractions and amplifications were sufficient to overcome the problems of DNA damage and the phylogeographic consistency of our results clearly attests to the authenticity of the data. The success of this study highlights the potential of temporal genetic analyses to test for changes in population size and distribution following human settlement in New Zealand. Importantly, our study has revealed that anthropogenic impacts may be far more dynamic than previously recognised, with complex interactions among population processes including extinctions, declines, expansions and colonisations (see also for example Leonard *et al.* 2007). It will be most interesting to investigate whether extinction-expansion dynamics similar to those detected for *Megadyptes* penguins have also occurred in other New Zealand coastal vertebrates.

References

Abbott, C. L. & Double, M. C. (2003) Phylogeography of shy and white-capped albatrosses inferred from mitochondrial DNA sequences: implications for population history and taxonomy. *Molecular Ecology* **12**, 2747-2758.

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

Akst, E. P., Boersma, P. D. & Fleischer, R. C. (2002) A comparison of genetic diversity between the Galapagos Penguin and the Magellanic Penguin. *Conservation Genetics* **3**, 375-383.

Allendorf, F. W. (1986) Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology* **5**, 181-190.

Allendorf, F. W. & Luikart, G. (2007) *Conservation and the genetics of populations*. Blackwell Publishing, Malden.

Anderson, A. (1989) *Prodigious birds: moa and moa-hunting in prehistoric New Zealand*. Cambridge University Press, Cambridge.

Anderson, A. (1991) The chronology of colonization in New Zealand. *Antiquity* **65**, 767-795.

Anderson, A., Allingham, B. & Smith, I. W. G. (1996) *Shag River Mouth: the archaeology of an early southern Maori village*. ANH Publications, Canberra.

Anderson, A. & Smith, I. (1996) The transient village in Southern New Zealand. *World Archaeology* **27**, 359-371.

Anderson, A. (2002) A fragile plenty: Pre-European Maori and the New Zealand Environment. In: *Environmental Histories of New Zealand* (eds. Pawson E, Brooking T), pp. 19-34. Oxford University Press, Auckland.

Anderson, A. (2005) Subpolar settlement in South Polynesia. *Antiquity* **79**, 791-800.

Anonymous (1895) *Appendixes to the Journals of the House of Representatives for 1894, H25:1-4*.

Apte, S. & Gardner, J. P. A. (2002) Population genetic subdivision in the New Zealand greenshell mussel (*Perna canaliculus*) inferred from single-strand conformation polymorphism analysis of mitochondrial DNA. *Molecular Ecology* **11**, 1617-1628.

Austin, J. J., Smith, A. B. & Thomas, R. H. (1997) Palaeontology in a molecular world: the search for authentic ancient DNA. *Trends in Ecology & Evolution* **12**, 303-306.

Avise, J. C. (1992) Molecular population structure and the biogeographic history of a regional fauna - a case history with lessons for conservation biology. *Oikos* **63**, 62-76.

Avise, J. C. (1994) Molecular Markers, Natural History and Evolution. Chapman and Hall, New York.

Ayers, K. L. & Waters, J. M. (2005) Marine biogeographic disjunction in central New Zealand. *Marine Biology* **147**, 1045-1052.

Baker, A. J., Pereira, S. L., Haddrath, O. P. & Edge, K. A. (2006) Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. *Proceedings of the Royal Society B-Biological Sciences* **273**, 11-17.

Baker, C. S. & Palumbi, S. R. (1994) Which whales are hunted? A molecular genetic approach to monitoring whaling. *Science* **265**, 1538-1539.

Banks, J. C., Mitchell, A. D., Waas, J. R. & Paterson, A. M. (2002) An unexpected pattern of molecular divergence within the blue penguin (*Eudyptula minor*) complex. *Notornis* **49**, 29-38.

Bantock, T. M., Prŷs-Jones, R. P. & Lee, P. L. M. (2008) New and improved molecular sexing methods for museum bird specimens. *Molecular Ecology Resources* **8**, 519-528.

Barnes, I., Matheus, P., Shapiro, B., Jensen, D. & Cooper, A. (2002) Dynamics of Pleistocene population extinctions in Beringian brown bears. *Science* **295**, 2267-2270.

Baumel, J. J. & Witmer, L. M. (1993) Osteologia. In: *Handbook of avian anatomy: nomina anatomica avium* (ed. Baumel JJ), pp. 45-132. Nuttall Ornithological Club, Cambridge.

Beaumont, M. A. (2003) Estimation of population growth or decline in genetically monitored populations. *Genetics* **164**, 1139-1160.

Beaumont, M. A. & Rannala, B. (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics* **5**, 251-261.

Begon, M., Townsend, C. R. & Harper, L. L. (2006) *Ecology: from individuals to ecosystems*. Blackwell Publishing, Malden.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier. Available from.

Bergl, R. A. & Vigilant, L. (2007) Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). *Molecular Ecology* **16**, 501-516.

Berry, O., Tocher, M. D. & Sarre, S. D. (2004) Can assignment tests measure dispersal? *Molecular Ecology* **13**, 551-561.

Bertelli, S. & Giannini, N. P. (2005) A phylogeny of extant penguins (Aves: Sphenisciformes) combining morphology and mitochondrial sequences. *Cladistics* **21**, 209-239.

Berthier, P., Beaumont, M. A., Cornuet, J. M. & Luikart, G. (2002) Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: A genealogical approach. *Genetics* **160**, 741-751.

Birdlife International (2008) Species factsheet: *Megadyptes antipodes*. Downloaded from http://www.birdlife.org on 25 August 2008.

Boessenkool, S., King, T. M., Seddon, P. J. & Waters, J. M. (2008) Isolation and characterization of microsatellite loci from the yellow-eyed penguin (*Megadyptes antipodes*). *Molecular Ecology Resources* **8**, 1043-1045.

Boessenkool, S., Austin, J. J., Worthy, T. H., Scofield, P., Cooper, A., Seddon, P. J. & Waters, J. M. (2009a) Relict or colonizer? Extinction and range expansion of penguins in southern New Zealand. *Proceedings of the Royal Society of London Series B-Biological Sciences* **276**, 815-821.

Boessenkool, S., Star, B., Waters, J. M. & Seddon, P. J. (2009b) Multilocus assignment analyses reveal multiple units and rare migration events in the recently expanded yellow-eyed penguin (*Megadyptes antipodes*). *Molecular Ecology* **18**, 2390-2400.

Bonaparte, C. L. (1831) Saggio di una distribuzione metodica degli animali vertebrati. *Giornale Arcadico di Scienze Lettere ed Arti* **52**, 155-189.

Brooke, M. D. (2000) Why museums matter. *Trends in Ecology & Evolution* **15**, 136-137.

Brotherton, P., Endicott, P., Sanchez, J. J., Beaumont, M., Barnett, R., Austin, J. & Cooper, A. (2007) Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic Acids Research* **35**, 5717-5728.

Brown, J. H. & Kodric-Brown, A. (1977) Turnover rates in insular biogeography - effect of immigration on extinction. *Ecology* **58**, 445-449.

Bull, E. (2005) *Components influencing lifetime reproductive success in yellow-eyed penguins*. Bachelor of Science (Honours), University of Otago.

Cassens, I., Van Waerebeek, K., Best, P. B., Tzika, A., Van Helden, A. L., Crespo, E. A. & Milinkovitch, M. C. (2005) Evidence for male dispersal along the coasts but no migration in pelagic waters in dusky dolphins (*Lagenorhynchus obscurus*). *Molecular Ecology* **14**, 107-121.

Caughley, G. (1994) Directions in Conservation Biology. *Journal of Animal Ecology* **63**, 215-244.

Cegelski, C. C., Waits, L. P. & Anderson, N. J. (2003) Assessing population structure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. *Molecular Ecology* **12**, 2907-2918.

Chakraborty, R. & Nei, M. (1977) Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution* **31**, 347-356.

Childerhouse, S. & Gales, N. (1998) Historical and modern distribution and abundance of the New Zealand sea lion *Phocarctos hookeri*. *New Zealand Journal of Zoology* **25**, 1-16.

Clarke, J. A., Ksepka, D. T., Stucchie, M., Urbina, M., Giannini, N., Bertelli, S., Narvez, Y. & Boyd, C. A. (2007) Paleogene equatorial penguins challenge the proposed relationship between biogeography, diversity, and Cenozoic climate change. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 11545-11550.

Clement, M., Posada, D. & Crandall, K. A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.

Clout, M. (2001) Where protection is not enough: active conservation in New Zealand. *Trends in Ecology and Evolution* **16**, 415-416.

Collar, N. J. & Rudyanto (2003) The archive and the ark: bird specimen data in conservation status assessment In: *Why Museums Matter: Avian Archives in the Age of Extinction. Bulletin of the British Ornithologist's Club 123A* (eds. Collar NJ, Fisher CT, Feare CJ), pp. 95-113.

Cooper, A. & Poinar, H. N. (2000) Ancient DNA: Do it right or not at all. *Science* **289**, 1139-1139.

Craig, J., Anderson, S., Clout, M., Creese, B., Mitchell, N., Ogden, J., Roberts, M. & Ussher, G. (2000) Conservation issues in New Zealand. *Annual Review of Ecology and Systematics* **31**, 61-78.

Cyclopedia Company Limited (1897) *The cyclopedia of New Zealand (Wellington Provincial District)*. The Cyclopedia Company Limited, Wellington.

Dalton, R. (2005) Ornithologists stunned by bird collector's deceit. *Nature* **437**, 302-303.

Darby, J. T. & Seddon, P. J. (1990) Breeding biology of the yellow-eyed penguin (*Megadyptes antipodes*). In: *Penguin Biology* (eds. Davis LS, Darby JT), pp. 45-62. Academic Press Inc, San Diego.

de Dinechin, M., Ottvall, R., Quillfeldt, P. & Jouventin, P. (2009) Speciation chronology of rockhopper penguins inferred from molecular,

geological and palaeoceanographic data. *Journal of Biogeography* **36**, 693-702.

Diamond, J. (2000) Blitzkrieg against the moas. *Science* **287**, 2170-2171.

Duncan, R. P., Blackburn, T. M. & Worthy, T. H. (2002) Prehistoric bird extinctions and human hunting. *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**, 517-521.

Duncan, R. P. & Blackburn, T. M. (2004) Extinction and endemism in the New Zealand avifauna. *Global Ecology and Biogeography* **13**, 509-517.

Edwards, S. V. & Hedrick, P. W. (1998) Evolution and ecology of MHC molecules: from genomics to sexual selection. *Trends in Ecology & Evolution* **13**, 305-311.

Efford, M. G. & Edge, K. A. (1998) Can artificial brood reduction assist the conservation of yellow-eyed penguins (*Megadyptes antipodes*)? *Animal Conservation* **1**, 263-271.

Ellenberg, U., Setiawan, A. N., Cree, A., Houston, D. M. & Seddon, P. J. (2007) Elevated hormonal stress response and reduced reproductive output in yellow-eyed penguins exposed to unregulated tourism. *General and Comparative Endocrinology* **152**, 54-63.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier, L., Laval, G. & Schneider, S. (2005) ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.

Excoffier, L. & Heckel, G. (2006) Computer programs for population genetics data analysis: a survival guide. *Nature Reviews Genetics* 7, 745-758.

Fabbri, E., Miquel, C., Lucchini, V., Santini, A., Caniglia, R., Duchamp, C., Weber, J. M., Lequette, B., Marucco, F., Boitani, L., Fumagalli, L., Taberlet, P. & Randi, E. (2007) From the Apennines to the Alps: colonization genetics of the naturally expanding Italian wolf (*Canis lupus*) population. *Molecular Ecology* **16**, 1661-1671.

Falush, D., Stephens, M. & Pritchard, J. K. (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Frankham, R. (1995) Effective population size/adult population size ratios in wildlife - a review. *Genetical Research* **66**, 95-107.

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

Franklin, I. R. (1980) Evolutionary change in small populations. In: *Conservation Biology: An Evolutionary - Ecological Perspective* (eds. Soule ME, Wilcox BA). Sinauer Associates, Sunderland.

Franklin, I. R. & Frankham, R. (1998) How large must populations be to retain evolutionary potential? *Animal Conservation* **1**, 69-70.

Fraser, D. J., Hansen, M. M., Ostergaard, S., Tessier, N., Legault, M. & Bernatchez, L. (2007a) Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Molecular Ecology* **16**, 3866-3889.

Fraser, D. J., Jones, M. W., McParland, T. L. & Hutchings, J. A. (2007b) Loss of historical immigration and the unsuccessful rehabilitation of extirpated salmon populations. *Conservation Genetics* **8**, 527-546.

Fraser, C. I., Nikula, R., Spencer, H. G., & Waters, J. M. (2009) Kelp genes reveal effects of subantarctic sea ice during the Last Glacial Maximum. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 3249-3253.

Friesen, V. L., Burg, T. M. & McCoy, K. D. (2007) Mechanisms of population differentiation in seabirds. *Molecular Ecology* **16**, 1765-1785.

Gemmell, N. J., Metcalf, V. J. & Allendorf, F. W. (2004) Mother's curse: the effect of mtDNA on individual fitness and population viability. *Trends in Ecology & Evolution* **19**, 238-244.

Gilbert, M. T. P., Bandelt, H. J., Hofreiter, M. & Barnes, I. (2005) Assessing ancient DNA studies. *Trends in Ecology & Evolution* **20**, 541-544.

Gill, J. M. & Darby, J. T. (1993) Deaths in yellow-eyed penguins (*Megadyptes antipodes*) on the Otago Peninsula during the summer of 1990. *New Zealand Veterinary Journal* **41**, 39-42.

Glenn, T. C. & Schable, N. A. (2005) Isolating microsatellite DNA loci. In: *Molecular Evolution: Producing the Biochemical Data, Part B* (eds. Zimmer EA, Roalson EH), pp. 202-222. Elsevier Academic Press Inc, San Diego.

Goudet, J. (2002) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Insitute of Ecology, Lausanne. Available from.

Graham, C. H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A. T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution* **19**, 497-503.

Haig, S. M. (1998) Molecular contributions to conservation. *Ecology* **79**, 413-425.

Hansen, M. M., Ruzzante, D. E., Nielsen, E. E., Bekkevold, D. & Mensberg, K. L. D. (2002) Long-term effective population sizes, temporal stability of genetic composition and potential for local adaptation in anadromous brown trout (*Salmo trutta*) populations. *Molecular Ecology* **11**, 2523-2535.

Hanski, I. (1998) Metapopulation dynamics. *Nature* **396**, 41-49.

Hastings, A. (1993) Complex interactions between dispersal and dynamics - lessons from coupled logistic equations. *Ecology* **74**, 1362-1372.

Headland, R. K. (1989) *Chronological List of Antarctic Expeditions and Related Historical Events*. Cambridge University Press, Cambridge.

Heath, R. A. (1985) A review of the physical oceanography of the seas around New Zealand -1982. *New Zealand Journal of Marine and Freshwater Research* **19**, 79-124.

Hedrick, P. W., Lacy, R. C., Allendorf, F. W. & Soule, M. E. (1996) Directions in conservation biology: Comments on Caughley. *Conservation Biology* **10**, 1312-1320.

Hedrick, P. W. (2004) Recent developments in conservation genetics. *Forest Ecology and Management* **197**, 3-19.

Helbig, A. J., Knox, A. G., Parkin, D. T., Sangster, G. & Collinson, M. (2002) Guidelines for assigning species rank. *IBIS* **144**, 518-525.

Hennache, A., Rasmussen, P., Lucchini, V., Rimondi, S. & Randi, E. (2003) Hybrid origin of the imperial pheasant *Lophura imperialis* (Delacour and Jabouille, 1924) demonstrated by morphology, hybrid experiments, and DNA analyses. *Biological Journal of the Linnean Society* **80**, 573-600.

Hewitt, G. M. (1996) Some genetic consequences of ice ages,and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247-276.

Higham, T., Anderson, A. & Jacomb, C. (1999) Dating the first New Zealanders: The chronology of Wairau Bar. *Antiquity* **73**, 420-427.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O. A. & Wilson, A. C. (1984) DNA sequences from the Quagga, an extinct member of the horse family. *Nature* **312**, 282-284.

Hofreiter, M., Serre, D., Poinar, H. N., Kuch, M. & Pääbo, S. (2001) Ancient DNA. *Nature Reviews Genetics* **2**, 353-359.

Hofreiter, M. (2007) Pleistocene extinctions: Haunting the survivors. *Current Biology* **17**, 609-611.

Holdaway, R. N. & Jacomb, C. (2000) Rapid extinction of the moas (Aves : Dinornithiformes): Model, test, and implications. *Science* **287**, 2250-2254.

Hombron, J. B. & Jacquinot, H. (1841) Description de plusieurs oiseaux nouveaux ou peu connus, provenant de l'expédition autour du monde faite sur les corvettes l'Astolabe et la Zélée. *Annales Des Sciences Naturelles, Zoologie* **16**, 312-320.

Howell, N., Kubacka, I. & Mackey, D. A. (1996) How rapidly does the human mitochondrial genome evolve? *American Journal of Human Genetics* **59**, 501-509.

Hurles, M. E., Matisoo-Smith, E., Gray, R. D. & Penny, D. (2003) Untangling Oceanic settlement: the edge of the knowable. *Trends in Ecology & Evolution* **18**, 531-540.

Jamieson, I. G. (2007) Has the debate over genetics and extinction of island endemics truly been resolved? *Animal Conservation* **10**, 139-144.

Jamieson, I. G., Grueber, C. E., Waters, J. M. & Gleeson, D. M. (2008) Managing genetic diversity in threatened populations: a New Zealand perspective. *New Zealand Journal of Ecology* **32**, 130-137.

Johnson, J. A., Bellinger, M. R., Toepfer, J. E. & Dunn, P. (2004) Temporal changes in allele frequencies and low effective population size in greater prairie-chickens. *Molecular Ecology* **13**, 2617-2630.

Jones, T. L., Porcasi, J. F., Erlandson, J. M., Dallas, H., Jr., Wake, T. A. & Schwaderer, R. (2008) The protracted Holocene extinction of California's flightless sea duck (*Chendytes lawi*) and its implications for the Pleistocene overkill hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4105-4108.

Jouventin, P., Cuthbert, R. J. & Ottvall, R. (2006) Genetic isolation and divergence in sexual traits: evidence for the northern rockhopper penguin *Eudyptes moseleyi* being a sibling species. *Molecular Ecology* **15**, 3413-3423.

Keller, L. F. & Waller, D. M. (2002) Inbreeding effects in wild populations. *Trends in Ecology and Evolution* **17**, 230-241.

Kennedy, M. & Spencer, H. G. (2000) Phylogeny, biogeography, and taxonomy of Australasian teals. *Auk* **117**, 154-163.

King, M. (2003) *Penguin History of New Zealand*. Penguin Books, Auckland.

Knox, A. G. (1993) Richard Meinertzhagen - a case of fraud examined. *Ibis* **135**, 320-325.

Ksepka, D. T., Bertelli, S. & Giannini, N. P. (2006) The phylogeny of the living and fossil Sphenisciformes (penguins). *Cladistics* **22**, 412-441.

Lacy, R. C. (1987) Loss of geetic diversity from managad populations; interacting effecs of drift, mutation, immigration, selection and population subdivision. *Conservation Biology* **1**, 143-158.

Lalas, C., Ratz, H., McEwan, K. & McConkey, S. D. (2007) Predation by New Zealand sea lions (*Phocarctos hookeri*) as a threat to the viability of yellow-eyed penguins (*Megadyptes antipodes*) at Otago Peninsula, New Zealand. *Biological Conservation* **135**, 235-246.

Lambert, D. M., Ritchie, P. A., Millar, C. D., Holland, B., Drummond, A. J. & Baroni, C. (2002) Rates of evolution in ancient DNA from Adelie penguins. *Science* **295**, 2270-2273.

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

Lande, R. & Shannon, S. (1996) The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* **50**, 434-437.

Larsson, J. K., Jansman, H. A. H., Segelbacher, G., Hoglund, J. & Koelewijn, H. P. (2008) Genetic impoverishment of the last black grouse (*Tetrao tetrix*) population in the Netherlands: detectable only with a reference from the past. *Molecular Ecology* **17**, 1897-1904.

Leberg, P. (2005) Genetic approaches for estimating the effective size of populations. *Journal of Wildlife Management* **69**, 1385-1399.

Lecis, R., Pierpaoli, M., Biro, Z. S., Szemethy, L., Ragni, B., Vercillo, F. & Randi, E. (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology* **15**, 119-131.

Lee, P. L. M. & Griffiths, R. (2003) Sexing errors among museum skins of a sexually monomorphic bird, the Moorhen *Gallinula chloropus*. *Ibis* **145**, 695-698.

Lee, P. L. M. & Prŷs-Jones, R. P. (2008) Extracting DNA from museum bird eggs, and whole genome amplification of archive DNA. *Molecular Ecology Resources* **8**, 551-560.

Leonard, J. A., Wayne, R. K. & Cooper, A. (2000) Population genetics of Ice age brown bears. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 1651-1654.

Leonard, J. A., Vila, C. & Wayne, R. K. (2005) Legacy lost: genetic variability and population size of extirpated US grey wolves (*Canis lupus*). *Molecular Ecology* **14**, 9-17.

Leonard, J. A., Vila, C., Fox-Dobbs, K., Koch, P. L., Wayne, R. K. & Van Valkenburgh, B. (2007) Megafaunal extinctions and the disappearance of a specialized wolf ecomorph. *Current Biology* **17**, 1146-1150.

Leonard, J. A. (2008) Ancient DNA applications for wildlife conservation. *Molecular Ecology* **17**, 4186-4196.

Leonard, J. A. & Wayne, R. K. (2008) Native Great Lakes wolves were not restored. *Biology Letters* **4**, 95-98.

Lynch, M. & Lande, R. (1998) The critical effective size for a genetically secure population. *Animal Conservation* **1**, 70-72.

Mackintosh, N. A. (1960) The pattern of distribution of the antarctic fauna. *Proceedings of the Royal Society of London Series B-Biological Sciences* **152**, 624-631.

Madsen, T., Shine, R., Olsson, M. & Wittzell, H. (1999) Conservation biology - restoration of an inbred adder population. *Nature* **402**, 34-35.

Manel, S., Berthier, P. & Luikart, G. (2002) Detecting wildlife poaching: Identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. *Conservation Biology* **16**, 650-659.

Manel, S., Bellemain, E., Swenson, J. E. & Francois, O. (2004) Assumed and inferred spatial structure of populations: the Scandinavian brown bears revisited. *Molecular Ecology* **13**, 1327-1331.

Manel, S., Gaggiotti, O. E. & Waples, R. S. (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology & Evolution* **20**, 136-142.

Marchant, S. & Higgins, P. J. (1990) *Handbook of Australian, New Zealand and Antarctic Birds*. Oxford University Press, Melbourne.

McGlone, M. S. (1983) Polynesian deforestation of New Zealand: a preliminary synthesis. *Archaeology in Oceania* **18**, 11-23.

McKinlay, B. (2001) Hoiho (*Megadyptes antipodes*) recovery plan 2000-2025. Department of Conservation, Wellington.

Meiri, S. & Mace, G. M. (2007) New taxonomy and the origin of species. *PLOS Biology* **5**, e194.

Miller, C. R. & Waits, L. P. (2003) The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): Implications for conservation. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 4334-4339.

Milne-Edwards, A. (1880) Expéditions scientifiques du Travailleur. *Annales des Sciences Naturelle* **6**, 1-56.

Milot, E., Weimerskirch, H. & Bernatchez, L. (2008) The seabird paradox: dispersal, genetic structure and population dynamics in a highly mobile, but philopatric albatross species. *Molecular Ecology* **17**, 1658-1673.

Moore, P. J. (1992) Population estimates of Yellow-eyed penguin (*Megadyptes antipodes*) on Campbell and Auckland Islands 1987-90. *Notornis* **39**, 1-15.

Moore, P. J. (2001) Historical records of yellow-eyed penguin (*Megadyptes antipodes*) in southern New Zealand. *Notornis* **48**, 145-156.

Moore, P. J., Fletcher, D. & Amey, J. (2001) Population estimates of yellow-eyed Penguins, *Megadyptes antipodes*, on Campbell Island, 1987-98. *Emu* **101**, 225-236.

Moors, P. J. (1983) Predation by mustelids and rodents on the eggs and chicks of native and introduced birds in Kowhai Bush, New Zealand. *Ibis* **125**, 137-154.

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

Nagaoka, L. (2001) Using diversity indices to measure changes in prey choice at the Shag River Mouth Site, Southern New Zealand. *International Journal of Osteoarchaeology* **11**, 101-111.

Nei, M., Maruyama, T. & Chakraborty, R. (1975) The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1-10.

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583-590.

Neigel, J. E. (2002) Is F_{ST} obsolete? *Conservation Genetics* **3**, 167-173.

NIWA (2008) Climate change: projections for New Zealand. National Institute of Water & Atmospheric Research, Wellington, New Zealand.

Nylander, J. A. A., Wilgenbusch, J. C., Warren, D. L. & Swofford, D. L. (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24**, 581-583.

O'Donnell, C. F. J. (1996) Predators and the decline of New Zealand forest birds: An introduction to the hole-nesting bird and predator programme. *New Zealand Journal of Zoology* **23**, 213-219.

Overeem, R. L., Peucker, A. J., Austin, C. M., Dann, P. & Burridge, C. P. (2008) Contrasting genetic structuring between colonies of the World's smallest penguin, *Eudyptula minor* (Aves: Spheniscidae). *Conservation Genetics* **9**, 893-905.

Paetkau, D., Slade, R., Burden, M. & Estoup, A. (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**, 55-65.

Palsbøll, P. J., Berube, M. & Allendorf, F. W. (2007) Identification of management units using population genetic data. *Trends in Ecology & Evolution* **22**, 11-16.

Palstra, F. P. & Ruzzante, D. E. (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology* **17**, 3428-3447.

Paquette, S. R., Behncke, S. M., O'Brien, S. H., Brenneman, R. A., Louis, E. E. & Lapointe, F. J. (2007) Riverbeds demarcate distinct conservation units of the radiated tortoise (*Geochelone radiata*) in southern Madagascar. *Conservation Genetics* **8**, 797-807.

Paxinos, E. E., James, H. F., Olson, S. L., Ballou, J. D., Leonard, J. A. & Fleischer, R. C. (2002) Prehistoric decline of genetic diversity in the nene. *Science* **296**, 1827-1827.

Peacock, L., Paulin, M. & Darby, J. T. (2000) Investigations into climate change influence on population dynamics of yellow-eyed penguins *Megadyptes antipodes*. *New Zealand Journal of Zoology* **27**, 317-325.

Pearse, D. E. & Crandall, K. A. (2004) Beyond F_{ST}: Analysis of population genetic data for conservation. *Conservation Genetics* **5**, 585-602.

Peel, D., Ovenden, J. R. & Peel, S. L. (2004) NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries.

Penn, D. J., Damjanovich, K. & Potts, W. K. (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 11260-11264.

Pennisi, E. (2000) Taxonomic revival. *Science* **289**, 2306-2308.

Perrin, C. & Roy, M. S. (2000) Rapid and efficient identification of microsatellite loci from the sea urchin, *Evechinus chloroticus*. *Molecular Ecology* **9**, 2221-2223.

Peterson, A. T., Navarro-Sigüenza, A. G. & Pereira, R. S. (2004) Detecting errors in biodiversity data based on collector's itineraries. *Bulletin of the British Ornithologists' Club* **124**, 143-151.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L. & Estoup, A. (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536-539.

Posada, D. & Crandall, K. A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Pritchard, J. K., Wen, X. & Falush, D. (2007) Documentation for structure software: Version 2.2. Department of Human Genetics, University of Chicago.

Pruett, C. L. & Winker, K. (2005) Northwestern song sparrow populations show genetic effects of sequential colonization. *Molecular Ecology* **14**, 1421-1434.

Pruvost, M., Schwarz, R., Correia, V. B., Champlot, S., Braguier, S., Morel, N., Fernandez-Jalvo, Y., Grange, T. & Geigl, E. M. (2007) Freshly excavated fossil bones are best for amplification of ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 739-744.

Pääbo, S. (1985) Molecular cloning of ancient Egyptian mummy DNA. *Nature* **314**, 644-645.

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L. & Hofreiter, M. (2004) Genetic analyses from ancient DNA. *Annual Review of Genetics* **38**, 645-679.

Rambaut, A. & Drummond, A. J. (2007) Tracer v1.4. Available from http://beast.bio.ed.uk/Tracer.

Rannala, B. & Mountain, J. L. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 9197-9201.

Rasmussen, P. C. & Collar, N. J. (1999) Major specimen fraud in the Forest Owlet *Heteroglaux* (*Athene* auct.) *blewitti*. *Ibis* **141**, 11-21.

Rasmussen, P. C. & Prŷs-Jones, R. P. (2003) History vs. mystery: the reliability of museum specimen data. In: *Why Museums Matter: Avian Archives in the Age of Extinction. Bulletin of the British Ornithologists' Club 123A* (eds. Collar NJ, Fisher CT, Feare CJ), pp. 66-94.

Ratz, H., Darby, J. T., Edge, K. A. & Thompson, C. (2004) Survival and breeding of yellow-eyed penguins (*Megadyptes antipodes*), at two locations on Otago Peninsula, South Island, New Zealand, 1991-96. *New Zealand Journal of Zoology* **31**, 133-147.

Rice, W. R. (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Richards, C. M. (2000) Inbreeding depression and genetic rescue in a plant metapopulation. *American Naturalist* **155**, 383-394.

Richdale, L. E. (1957) *A population study of penguins*. Clarendon Press, Oxford.

Roeder, A. D., Marshall, R. K., Mitchelson, A. J., Visagathilagar, T., Ritchie, P. A., Love, D. R., Pakai, T. J., McPartlan, H. C., Murray, N. D., Robinson, N. A., Kerry, K. R. & Lambert, D. M. (2001) Gene flow on the ice: genetic differentiation among Adelie penguin colonies around Antarctica. *Molecular Ecology* **10**, 1645-1656.

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

Rotschild, M. L. (1983) *Dear Lord Rothschild: birds, butterflies and history*. Hutchinson, London.

Rousset, F. (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

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

Schlosser, J. A., Dubach, J. M., Garner, T. W. J., Araya, B., Bernal, M., Simeone, A., Smith, K. A. & Wallace, R. S. (2009) Evidence for gene flow differs from observed dispersal patterns in the Humboldt penguin, *Spheniscus humboldti*. *Conservation Genetics* **In press**.

Schlötterer, C. (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**, 365-371.

Schneider, I. (1998) Gene codes sequencher. *Genetic Engineering News* **18**, 24-24.

Seddon, P. J. & Davis, L. S. (1989) Nest-site selection by yellow-eyed penguins. *The Condor* **91**, 653-659.

Sefc, K. M., Payne, R. B. & Sorenson, M. D. (2003) Microsatellite amplification from museum feather samples: Effects of fragment size and template concentration on genotyping errors. *Auk* **120**, 982-989.

Sefc, K. M., Payne, R. B. & Sorenson, M. D. (2007) Single base errors in PCR products from avian museum specimens and their effect on estimates of historical genetic diversity. *Conservation Genetics* **8**, 879-884.

Seutin, G., White, B. N. & Boag, P. T. (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* **69**, 82-90.

Shaffer, H. B., Fisher, R. N. & Davidson, C. (1998) The role of natural history collections in documenting species declines. *Trends in Ecology & Evolution* **13**, 27-30.

Shaffer, M. L. (1981) Minimum population sizes for species conservation. *Bioscience* **31**, 131-134.

Shaffer, M. L. (1987) Minimum viable populations: coping with uncertainty. In: *Viable Populations for Conservation* (ed. Soulé ME), pp. 69-86. Cambridge University Press, Cambridge.

Shapiro, B., Drummond, A. J., Rambaut, A., Wilson, M. C., Matheus, P. E., Sher, A. V., Pybus, O. G., Gilbert, M. T. P., Barnes, I., Binladen, J., Willerslev, E., Hansen, A. J., Baryshnikov, G. F., Burns, J. A., Davydov, S., Driver, J. C., Froese, D. G., Harington, C. R., Keddie, G., Kosintsev, P., Kunz, M. L., Martin, L. D., Stephenson, R. O., Storer, J., Tedford, R., Zimov, S. & Cooper, A. (2004) Rise and fall of the Beringian steppe bison. *Science* **306**, 1561-1565.

Sharpe, R. B. (1891) A review of recent attempts to classify birds **90**.

Shivji, M., Clarke, S., Pank, M., Natanson, L., Kohler, N. & Stanhope, M. (2002) Genetic identification of pelagic shark body parts for conservation and trade monitoring. *Conservation Biology* **16**, 1036-1047.

Sites, J. W. & Marshall, J. C. (2003) Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology & Evolution* **18**, 462-470.

Slatkin, M. (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics* **16**, 393-430.

Smith, I. & Anderson, A. (2007) Codfish Island/Whenua Hou Archaeological Project: Preliminary Report. In: *Otago Archaeological Laboratory Report*. University of Otago, Dunedin.

Soulé, M. E. (1987) *Viable populations for conservation*. Cambridge University Press, Cambridge.

Steadman, D. W. (1995) Prehistoric extinctions of Pacific island birds - biodiversity meets zooarchaeology. *Science* **267**, 1123-1131.

Steadman, D. W. & Martin, P. S. (2003) The late Quaternary extinction and future resurrection of birds on Pacific islands. *Earth-Science Reviews* **61**, 133-147.

Suarez, A. V. & Tsutsui, N. D. (2004) The value of museum collections for research and society. *Bioscience* **54**, 66-74.

Swofford, D. L. (2003) Paup* Phylogenetic Analysis Using Parsimony (* and Other Methods). Sinauer & Associates, Massachusetts. Available from http://paup.csit.fsu.edu/.

Taberlet, P., Swenson, J. E., Sandegren, F. & Bjarvall, A. (1995) Localization of a contact zone between 2 highly divergent mitochondrial-DNA

lineages of the brown bear *Ursus arctos* in Scandinavia. *Conservation Biology* **9**, 1255-1261.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L. P. & Bouvet, J. (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Tallmon, D. A., Luikart, G. & Beaumont, M. A. (2004) Comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. *Genetics* **167**, 977-988.

Tarr, C. L., Conant, S. & Fleischer, R. C. (1998) Founder events and variation at microsatellite loci in an insular passerine bird, the Laysan finch (*Telespiza cantans*). *Molecular Ecology* **7**, 719-731.

Tautz, D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463-6471.

Taylor, S. S., Jamieson, I. G. & Wallis, G. P. (2008) Ancestral and contemporary levels of genetic variation in two New Zealand passerines with different histories of decline. *Journal of Evolutionary Biology* **20**, 2035-2047.

Triggs, S. & Darby, J. T. (1989) Genetics and conservation of yellow-eyed penguin: an interim report. Science and Research Directorate, Department of Conservation, Wellington.

UNEP (2008) The New Zealand Subantarctic Islands. United Nations Environment Program, World Conservation Monitoring Centre.

Urquhart, A., Oldroyd, N. J., Kimpton, C. P. & Gill, P. (1995) Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. *Biotechniques* **18**, 116-121.

Valdiosera, C. E., Garcia-Garitagoitia, J. L., Garcia, N., Doadrio, I., Thomas, M. G., Hänni, C., Arsuaga, J. L., Barnes, I., Hofreiter, M., Orlando, L. & Götherström, A. (2008) Surprising migration and population size dynamics in ancient Iberian brown bears (*Ursus arctos*). *Proceedings of the National Academy of Sciences of the United States of America* **105**, 5123-5128.

van Heezik, Y. & Davis, L. (1990) Effects of food variability on growth rates, fledging sizes and reproductive success in the yellow-eyed penguin *Megadyptes antipodes*. *Ibis* **132**, 354-365.

Vilà, C., Sundqvist, A. K., Flagstad, O., Seddon, J., Björnerfeldt, S., Kojola, I., Casulli, A., Sand, H., Wabakken, P. & Ellegren, H. (2003) Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 91-97.

Waits, L., Taberlet, P., Swenson, J. E., Sandegren, F. & Franzen, R. (2000) Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (*Ursus arctos*). *Molecular Ecology* **9**, 421-431.

Walsh, P. S., Metzger, D. A. & Higuchi, R. (1991) Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**, 506-513.

Walter, R., Smith, I. & Jacomb, C. (2006) Sedentism, subsistence and socio-political organization in prehistoric New Zealand. *World Archaeology* **38**, 274-290.

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

Wang, J. L. (2001) A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genetical Research* **78**, 243-257.

Wang, J. L. & Whitlock, M. C. (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* **163**, 429-446.

Waples, R. S. (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**, 379-391.

Waples, R. S. (2002) Definition and estimation of effective population size in the conservation of endangered species. In: *Population Viability*

Analysis (eds. Beissinger SR, McCullough DR). The University of Chicago Press, Chicago.

Waples, R. S. & Gaggiotti, O. (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* **15**, 1419-1439.

Waples, R. S. & Yokota, M. (2007) Temporal estimates of effective population size in species with overlapping generations. *Genetics* **175**, 219-233.

Waser, P. M. & Strobeck, C. (1998) Genetic signatures of interpopulation dispersal. *Trends in Ecology & Evolution* **13**, 43-44.

Wasser, S. K., Shedlock, A. M., Comstock, K., Ostrander, E. A., Mutayoba, B. & Stephens, M. (2004) Assigning African elephant DNA to geographic region of origin: Applications to the ivory trade. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14847-14852.

Wayne, R. K., Leonard, J. A. & Cooper, A. (1999) Full of sound and fury: The recent history of ancient DNA. *Annual Review of Ecology and Systematics* **30**, 457-477.

Weir, B. S. & Cockerham, C. C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Whyte, A. L. H., Marshall, S. J. & Chambers, G. K. (2005) Human evolution in Polynesia. *Human Biology* **77**, 157-177.

Willerslev, E. & Cooper, A. (2005) Ancient DNA. *Proceedings of the Royal Society of London Series B-Biological Sciences* **272**, 3-16.

Willi, Y., Van Buskirk, J. & Hoffmann, A. A. (2006) Limits to the adaptive potential of small populations. *Annual Review of Ecology Evolution and Systematics* **37**, 433-458.

Wilmshurst, J. M., Anderson, A. J., Higham, T. F. G. & Worthy, T. H. (2008) Dating the late prehistoric dispersal of polynesians to New Zealand using the commensal Pacific rat. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 7676-7680.

Winker, K. (2000) Obtaining, preserving, and preparing bird specimens. *Journal of Field Ornithology* **71**, 250-297.

Worthy, T. H. (1997) The identification of fossil *Eudyptes* and *Megadyptes* bones at Marfells Beach, Marlborough, South Island. *New Zealand Natural Sciences* **23**, 71-85.

Worthy, T. H. (1998) A remarkable fossil and archaeological avifauna from Marfells Beach, Lake Grassmere, South Island, New Zealand. *Records of the Canterbury Museum* **12**, 79-176.

Worthy, T. H. (1999) What was on the menu? Avian extinction in New Zealand. *New Zealand Journal of Archaeology* **19**, 125-160.

Worthy, T. H. & Holdaway, R. N. (2002) *The lost world of the moa: prehistoric life of New Zealand.* Indiana University Press, Bloomington.

Worthy, T. H., Tennyson, A. J. D., Archer, M., Musser, A. M., Hand, S. J., Jones, C., Douglas, B. J., McNamara, J. A. & Beck, R. M. D. (2006) Miocene mammal reveals a Mesozoic ghost lineage on insular New Zealand, southwest Pacific. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19419-19423.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.

Appendices

Appendix 9.1 Historic *Megadyptes* samples used for DNA analyses presented in Chapter 2

Museum abbreviation and accession number, laboratory code, collection location, collection year, sampling region as indicated in figure 2.1 and control region haplotype (H). The following museum abbreviations are used: AM = Auckland Museum, AMNH = American Natural History Museum, AUM = Australian Museum, CM = Canterbury Museum, MCZ = Museum of Comparative Zoology, MHNG = Natural History Museum Geneva, MNZ = Museum of New Zealand Te Papa Tongarewa, NHMT = Natural History Museum Tring, NMNHP = Natural History Museum Paris, NMW = Natural History Museum Vienna, NRM = Swedish Museum of Natural History, OM = Otago Museum, SAMA = South Australian Museum, USNM = Smithsonian Institution, ZMB = Museum für Naturkunde Berlin. Location abbreviations are as follows: AI = Auckland Islands, CI = Campbell Island, SI = South Island, STI = Stewart Island.

Museum accession no.	Lab code	Collection location	Collection year	Sampling region	H
AM LB5045[b]	YM12	Stewart Island	1932	M9	A8
AM LB5046[a,b]	YM13	Stewart Island	1935	M9	
AM LB5047[a,b]	YM14	Stewart Island	1935	M9	
AMNH 525843	YM01	Otago Coast, SI	1895	M5/6	A8
AMNH 525844	YM02	Otago Heads, SI	1895	M6	A8
AMNH 525845	YM03	Otago Heads, SI	1895	M6	A8
AMNH 525849	YM04	Campbell Island	1894	M11	A8
AMNH 525850	YM05	Campbell Island	1893	M11	A2
AMNH 525851	YM06	Campbell Island	1894	M11	A2
AMNH 525852	YM07	Campbell Island	1893	M11	A2
AMNH 525853	YM08	Auckland Islands	1893	M10	A8
AMNH 525854	YM09	Auckland Islands	1893	M10	A8
AMNH 525855	YM10	Auckland Islands	1893	M10	A8
AMNH 525856	YM11	Auckland Islands	1894	M10	A2
AUM O.23935[a]	YM49	Otago, SI	1915	M5/6/7	
AUM O.37154[a]	YM50	Catlins River Mouth, SI	1938	M7	

Appendix 9.1 continued

Museum accession no.	Lab code	Collection location	Collection year	Sampling region	H
AUM O.37155[a]	YM51	Catlins River Mouth, SI	1938	M7	
AUM O.37156	YM52	Catlins River Mouth, SI	1938	M7	A8
AUM O.37157	YM53	Catlins River Mouth, SI	1938	M7	A2
AUM O.37158	YM54	Catlins River Mouth, SI	1938	M7	A8
AUM O.37159	YM55	Catlins River Mouth, SI	1938	M7	A2
CM AV801	YM15	Otago Coast, SI	1895	M5/6/7	A8
CM AV802	YM16	Otago, SI	1895	M5/6/7	A2
CM AV862	YM17	Dunedin, SI	1937	M6	A8
MCZ 86732	YM20	South Island	1895	M4-7	A2
MHNG 754.46[a]	YM26	Otago Peninsula, SI	1911	M6	
MNZ OR.12968	YM40	Perseverance Harbour, CI	1943	M11	A2
MNZ OR.12969	YM41	Penguin Bay, AI	1942	M10	A2
MNZ OR.12970	YM42	Auckland Islands	1942	M10	A10
MNZ OR.12971	YM43	Ocean Island, AI	1943	M10	A2
MNZ OR.5378	YM44	Campbell Island	1944	M11	A8
MNZ OR.5379	YM45	Otago Heads, SI	1895	M6	A8
MNZ OR.5383	YM46	Otago Peninsula, SI	1895	M6	A8
MNZ OR.5385	YM47	Otago Heads, SI	1895	M6	A2
MNZ OR.5386	YM48	Stewart Island	1888	M9	A8
NHMT 1842.12.16.165[a]	YM21	Auckland Islands	1840	M10	A2
NHMT 1897.12.6.40[a]	YM22	Otago, SI	1895	M5/6/7	
NHMT 1901.1.7.15	YM23	Campbell Island	1899	M11	
NHMT 1901.1.7.16	YM24	Campbell Island	1899	M11	A1
NHMT 1905.12.30.233	YM25	Auckland Islands	1904	M10	A8

Appendix 9.1 continued

Museum accession no.	Lab code	Collection location	Collection year	Sampling region	H
NMNHP CG: 1875-522[a]	YM18	Campbell Island	1875	M11	
NMW 4385	YM27	Stewart Island	1888	M9	A2
NMW 4386	YM28	Stewart Island	1888	M9	A2
NMW 4402	YM29	Stewart Island	1888	M9	A8
NMW 4403	YM30	Stewart Island	1888	M9	A8
NMW 4404	YM31	Stewart Island	1888	M9	A8
NMW 4405	YM32	Stewart Island	1888	M9	A8
NMW 4406	YM33	Stewart Island	1888	M9	A8
NMW 4408	YM34	Stewart Island	1888	M9	A2
NRM 569465[a]	YM39	Campbell Island	1924	M11	
OM AV034	YM35	Otago Heads, SI	1895	M6	A2
SAMA B13913	YM38	Otago, SI	1911	M5/6/7	A8
USNM 124683[a]	YM36	Stewart Island	1891	M9	
USNM 15655[a]	YM37	Auckland Islands	1840	M10	
ZMB2000.8244	YM19	Stewart Island	1888	M11	A2

[a] No or only sporadic amplification, not included in genetic analyses.
[b] DNA extracts kindly provided by C. Millar.

Appendix 9.2 Prehistoric *Megadyptes* samples used for DNA analyses presented in Chapter 2.

Museum abbreviation and accession number, collection location, element, sampling region as indicated in figure 2.1 and control region haplotype (H). Museum accession numbers were used as laboratory codes. Specimens CM AV34157, UO SMC/BB360-1 and UO Map1 are the prehistoric bones from South Island locations identified as *M. antipodes*. The following museum abbreviations are used: CM = Canterbury Museum, NMNZ = Museum of New Zealand Te Papa Tongarewa, OM = Otago Museum, UO = University of Otago (Department of Archaeology). Location abbreviations are as follows: AI = Auckland Islands, SI = South Island, STI = Stewart Island.

Museum accession no.	Location	Element	Sampling region	H
CM AV10456[a]	Wairau Bar, SI[e]	coracoid	P2	
CM AV10458[a]	Wairau Bar, SI[e]	coracoid	P2	
CM AV10459[a]	Wairau Bar, SI[e]	femur	P2	
CM AV11540[a]	Marfells Beach, SI[f]	humerus	P3	
CM AV11995[d]	Marfells Beach, SI[f]	femur	P3	W1
CM AV12535[a]	Marfells Beach, SI[f]	femur	P3	
CM AV13269[c]	Marfells Beach, SI[f]	femur	P3	W2
CM AV13641	Old Neck, STI[e]	femur	P9	W10
CM AV15787A(.2)1[a]	Redcliffs, SI[e]	femur	P4	
CM AV15787B(.1)[a]	Redcliffs, SI[e]	femur	P4	
CM AV16046A(.1)	Redcliffs, SI[e]	femur	P4	W10
CM AV16046B(.2)[a]	Redcliffs, SI[e]	femur	P4	
CM AV16200	Redcliffs, SI[e]	femur	P4	W10
CM AV16256	Redcliffs, SI[e]	femur	P4	W10
CM AV16258X(.1)	Redcliffs, SI[e]	femur	P4	W10
CM AV16258Z(.2)[d]	Redcliffs, SI[e]	femur	P4	W13
CM AV32860	Pounawea, SI[e]	femur	P7	W6
CM AV32877	Pounawea, SI[e]	femur	P7	W5

Appendix 9.2 continued

Museum accession no.	Location	Element	Sampling region	H
CM AV34157	Old Neck, STI[e]	femur	P9	A8
CM AV34198	Pounawea, SI[e]	femur	P7	W10
CM AV34367	Pounawea, SI[e]	femur	P7	W10
CM AV34373	Pounawea, SI[e]	femur	P7	W4
CM AV34566	Old Neck, STI[e]	femur	P9	W12
CM AV34941[d]	Pounawea, SI[e]	femur	P7	W4
CM AV35004[a]	Tumbledown Bay, SI[e]	humerus	P4	
CM AV37359.1	Old Neck, STI[e]	femur	P9	W8
CM AV37359.2	Old Neck, STI[e]	femur	P9	W10
CM AV37359.3	Old Neck, STI[e]	femur	P9	W9
NMNZ S.41937.1	Enderby Island, AI[e]	femur	P10	A2
NMNZ S.41937.2	Enderby Island, AI[e]	femur	P10	A9
NMNZ S.41937.3	Enderby Island, AI[e]	femur	P10	A8
NMNZ S.41975.1	Enderby Island, AI[e]	cranial fragment	P10	A5
NMNZ S.41975.2	Enderby Island, AI[e]	cranial fragment	P10	A6
NMNZ S.41984.1	Enderby Island, AI[e]	femur	P10	A8
NMNZ S.41984.2	Enderby Island, AI[e]	femur	P10	A2
NMNZ S.42029	Enderby Island, AI[e]	furcula	P10	A2
NMNZ S.42156.1	Delaware Bay, SI	femur	P1	W1
NMNZ S.42156.2	Delaware Bay, SI	femur	P1	W3
OM PN/J17/L2	Pounawea, SI[e]	femur	P7	W5
OM PN/J18/L1/2[a]	Pounawea, SI[e]	femur	P7	
UO 50BB2	Pleasant River, SI[e]	coracoid	P5	W4
UO 65BB1[a]	Pleasant River, SI[e]	coracoid	P5	
UO 277BB2[a]	Pleasant River, SI[e]	pelvis	P5	
UO 392BB1	Pleasant River, SI[e]	ulna	P5	W10

Appendix 9.2 continued

Museum accession no.	Location	Element	Sampling region	H
UO SMA/7BB1[a]	Shag River Mouth, SI[e]	humerus	P5	
UO SMA/17BB1[a]	Shag River Mouth, SI[e]	humerus	P5	
UO SMB/9BB1[a]	Shag River Mouth, SI[e]	tibiotarsus	P5	
UO SMC/BB292-2[a]	Shag River Mouth, SI[e]	ulna	P5	
UO SMC/BB298-1	Shag River Mouth, SI[e]	femur	P5	W10
UO SMC/BB360-1	Shag River Mouth, SI[e]	ulna	P5	A4
UO SMC/BB366-6[a]	Shag River Mouth, SI[e]	femur	P5	
UO SMC/BB382-3	Shag River Mouth, SI[e]	ulna	P5	W12
UO SMC/BB494-1	Shag River Mouth, SI[e]	ulna	P5	W11
UO SMC/BB520-20	Shag River Mouth, SI[e]	femur	P5	W12
UO SMC/BB641-1[a]	Shag River Mouth, SI[e]	ulna	P5	
UO SMC/BB699-1	Shag River Mouth, SI[e]	ulna	P5	W10
UO SMC/BB714-3	Shag River Mouth, SI[e]	ulna	P5	W10
UO SMC/BB733-3[a]	Shag River Mouth, SI[e]	tibiotarsus	P5	
UO SMD/625BB3[a]	Shag River Mouth, SI[e]	coracoid	P5	
UO SMD/633BB4[a]	Shag River Mouth, SI[e]	coracoid	P5	
UO CF218-B4[a]	Sealers Bay, Codfish Island[g]	coracoid	P8	
UO CF228-B1[a]	Sealers Bay, Codfish Island[g]	femur	P8	
UO CF231-B1-1[a]	Sealers Bay, Codfish Island[g]	coracoid	P8	
UO CF231-B1-2[a]	Sealers Bay, Codfish Island[g]	coracoid	P8	
UO CF311-B1	Sealers Bay, Codfish Island[g]	coracoid	P8	W4
UO CF343-B2[a]	Sealers Bay, Codfish Island[g]	coracoid	P8	
UO Map1	Mapoutahi Pa, SI[h]	femur	P6	A8
Harw 1[a,b]	Harwood, SI[i]	tibiotarsus	P6	
Harw 2[b]	Harwood, SI[i]	tibiotarsus	P6	W7

Appendix 9.2 continued

[a] No or only sporadic amplification, not included in genetic analyses.
[b] Personal collection Chris Lalas
[c] *Megadyptes waitaha* holotype
[d] *Megadyptes waitaha* paratype
[e] Moahunter site dated 1250-1450 AD (Anderson 1989; Anderson 1991; Higham *et al.* 1999)
[f] Natural and moahunter sites dated 600-1500 AD (Worthy 1998)
[g] Moahunter site dated 1250-1450 AD (Smith & Anderson 2007)
[h] Early fortified village (Pa) site dated ~1700 AD (Smith & James-Lee, in prep)
[i] Natural site dated 1300-1400 (C. Lalas pers. comm.)

Appendix 9.3 Specimens used for morphometric measurements presented in Chapter 2.

Specimens are listed by museum abbreviation and accession number. Modern *Megadyptes antipodes* specimens consisted of whole skeletons for which all four bone types were measured. Prehistoric specimens are listed per bone type. Museum abbreviations are as follows: CM = Canterbury Museum, NMNZ = Te Papa Museum of Natural History New Zealand, OM = Otago Museum, UO = University of Otago (Department of Archaeology)

M. antipodes modern skeletons

OM AV831, OM AV832, OM AV954, OM AV962, OM AV986, OM AV1001, OM AV1003, OM AV1005, OM AV1009, OM AV1010, OM AV1012, OM AV1014, OM AV1015, OM AV1319, OM AV1906, OM AV1908, OM AV4173, OM AV4174, OM AV7420, OM AV7843, OM AV7844, OM AV7845, OM AV7860, OM AV7903, OM AV7904, OM AV7905

M. antipodes prehistoric specimens

Femur

NMNZ S.41937.1, NMNZ S.41937.2, NMNZ S.41937.3, NMNZ S.41984.1, NMNZ S.41984.2, CM AV34157, UO Map1

M. waitaha prehistoric specimens

Femur

CM AV10459, CM AV11995, CM AV12535, CM AV13269, CM AV13641AA, CM AV15787A(.2), CM AV15787B(.1), CM AV16046A(.1), CM AV16046B(.2), CM AV16200C, CM AV16256A, CM AV16258X(.1), CM AV16258Z(.2), CM AV32877, CM AV34198, CM AV34367, CM AV34373, CM AV34566, CM AV34941, CM AV36190, CM AV37358, CM AV37359.1, CM AV37359.2, CM AV37359.3, NMNZ S.42156.1, UO BB298-1, UO BB366-6, UO CF228-1

Appendix 9.3 continued

Humerus

CM AV9654, CM AV11987, CM AV12083, CM AV12447, CM AV13448, CM AV13641N, CM AV15782B, CM AV15782C, CM AV15782D, CM AV15782E, CM AV16046F, CM AV16255G, CM AV16258N, CM AV19926, CM , CM AV25723A, CM AV25723B, CM AV32860, CM AV32861, CM AV34091, CM AV34208, CM AV34373, CM AV34374, CM AV34647, CM AV35004, UO CF77B2

Coracoid

CM AV10456, CM AV10943, CM AV13641K, CM AV13653, CM AV15784C, CM AV15784D, CM AV16256C, CM AV16258I, CM AV36190, UO 50BB3, UO CF311B1

Tarsometatarsus

CM AV9654, CM AV10464, CM AV10944, CM AV11083, CM AV11718, CM AV13973, CM AV16258I, CM AV16258T, CM AV22791, CM AV33133, CM AV34219

Appendix 9.4 Sequence alignment for 402 bp of the control region from *Megadyptes antipodes* and *M. waitaha*

A1-W13 refer to haplotypes, number in brackets reflects basepairs.

```
A1  AGCACATTATACTGATATTAAGCAAGTACAGTTAAATGTATGTGCTATAACCATATTAATGGTGGGTTGGGTGGAAATGGTTTCTTCACTGCTATGTTTCA [101]
A2  ................G.................................................................................... [101]
A3  ................G.................................................................................... [101]
A4  ................G.....................................................................T.............. [101]
A5  ................G.................................................................................... [101]
A6  ................G.................................................................................... [101]
A7  ................G.....................................G.....................................C.............. [101]
A8  ................G.....................................................................C.............. [101]
A9  ................G.............................C.....................................C.............. [101]
A10 ...........T....G.............................C.....................................C.............. [101]
W1  ................G.............................C..T.....................A.............T.C.............. [101]
W2  ................G.............................C..T.....................A.............T.C.............. [101]
W3  ................G.............................C..T.....................A.............T.C.............. [101]
W4  ...........T....G.............C...........T..C..T.....................................T.............. [101]
W5  ...........T....G.............C.............C..T.....................................T.............. [101]
W6  ...........T....G.............C.............C..T.....................................T.............. [101]
W7  ...........T....G.............C.............C..T.....................................T.............. [101]
W8  ...........T..G.G.............................C..T.....................................T.............. [101]
W9  ...........T..G.G.............................C.....................................T.............. [101]
W10 ...........T....G.............................C..T.....................................T.............. [101]
W11 ...........T....G.............................C..T....C................................T.............. [101]
W12 ...........T....G.............................C..T.....................................T.............. [101]
W13 ...........T....G.............................C..T.....................................T..T............. [101]
```

```
A1  GGGATAATTGGAGTAATACTTTCAAGGATTAACTGAGTAATGGTATGAGGATTAGCTCCTATTACATTGACTAAACGAGTTTAATGTGATGGTTTAAGGAA [202]
A2  ........................................................................................................ [202]
A3  ........................................................................................................ [202]
A4  ........................................................................................................ [202]
A5  ..............................................A......................................................... [202]
A6  ........................................................................................................ [202]
A7  ..............................................................................C..................... [202]
A8  ..............................................................................C..................... [202]
A9  ..............................................................................C..................... [202]
A10 ..............................................................................C..................... [202]
W1  .........A.....................................G.A..............C....G...................C..................... [202]
W2  .........A.....................................G.A..............C....G...................C..................... [202]
W3  ........GA.......................A............A..............C....G...................C..................... [202]
W4  .........A.....................................................G...................C..................... [202]
W5  .........A.....................................................G...................C..................... [202]
W6  .........A.....................................................G...................C..................... [202]
W7  .........A.....................................A...............G...................C..................... [202]
W8  .........A.....................................................G...................C..................... [202]
W9  .........A.....................................................G...................C..................... [202]
W10 .........A.....................................................G...................C..................... [202]
W11 .........A.....................................................G...................C..................... [202]
W12 .........A.....................................................G.......................................... [202]
W13 .........A.....................................................G...................C..................... [202]
```

```
A1  TTGGTATGTTATGTGCTAAATGATTATCTTCTGTTCCGTTGGAAAGGACTAAACCCATAGGTTGGTATCGTTTGTGTATCTCTGCTGTGGATAGTACGGCT [303]
A2  ..............................................................................................................  [303]
A3  ..............................................................................G...............  [303]
A4  ..............................................................................................................  [303]
A5  ..............................................................................................................  [303]
A6  ..............................................................................................................  [303]
A7  ..............................................................................................................  [303]
A8  ..............................................................................................................  [303]
A9  ..............................................................................................................  [303]
A10 ..............................................................................................................  [303]
W1  ..........................................................T...................................  [303]
W2  .........C................................................T...................................  [303]
W3  ..........................................................T...................................  [303]
W4  ..........................................................T.....G.............................  [303]
W5  ..........................................................T.....G.............................  [303]
W6  ..........................................................T.....G.............................  [303]
W7  ..........................................................T.....G.............................  [303]
W8  ..........................................................T.....G.............................  [303]
W9  ..........................................................T.....G.............................  [303]
W10 ..........................................................T.....G.............................  [303]
W11 ..........................................................T.....G.............................  [303]
W12 ..........................................................T.....G.............................  [303]
W13 ..........................................................T.....G.............................  [303]
```

```
A1  GTGCTTGGTTACCTGAACTCAATGGTGGAAGCGCTTAATATCCAATCATCTCTTGGAGTGCCAGTTTCAGGAACTGGGTTATCTATTAATTTGACTTCT [402]
A2  ........................................................................................................ [402]
A3  ........................................................................................................ [402]
A4  ........................................................................................................ [402]
A5  ........................................................................................................ [402]
A6  ..................................................C..................................................... [402]
A7  ..................................................C..................................................... [402]
A8  ..................................................C..................................................... [402]
A9  ..................................................C..................................................... [402]
A10 ..................................................C..................................................... [402]
W1  .......A..................A.....................C..................................................... [402]
W2  .......A..................A.....................C..................................................... [402]
W3  .......A..................A.....................C..................................................... [402]
W4  ........................A....A.................C..................................................... [402]
W5  ........................A....A.................C..................................................... [402]
W6  ........................A....A.................C....C................................................. [402]
W7  ........................A....A.................C..................................................... [402]
W8  ........................A....A.................C..................................................... [402]
W9  ........................A....A.................C..................................................... [402]
W10 ........................A....A.................C..................................................... [402]
W11 ........................A....A.................C..................................................... [402]
W12 ........................A....A.................C..................................................... [402]
W13 ........................A....A.................C..................................................... [402]
```

Appendix 9.5 Allele frequencies for contemporary yellow-eyed penguin breeding areas at 12 microsatellite loci

Allele frequencies, sample size/ locus (N), expected heterozygosity (H_e) and observed heterozygosity (H_o).

Locus	North Otago	Otago P.	Catlins	Stewart Island	Codfish Island	Auckl. Islands	Campb. Island
Mano3							
N	35	86	38	40	50	52	49
112	0.000	0.000	0.000	0.000	0.000	0.048	0.000
116	1.000	1.000	1.000	1.000	1.000	0.914	0.888
118	0.000	0.000	0.000	0.000	0.000	0.010	0.000
120	0.000	0.000	0.000	0.000	0.000	0.029	0.112
H_e	0.000	0.000	0.000	0.000	0.000	0.164	0.201
H_o	0.000	0.000	0.000	0.000	0.000	0.173	0.225
Mano8							
N	35	86	38	40	50	52	49
116	0.757	0.895	0.868	0.950	0.890	0.673	0.796
122	0.243	0.087	0.079	0.050	0.110	0.077	0.020
126	0.000	0.017	0.053	0.000	0.000	0.250	0.184
H_e	0.373	0.192	0.240	0.096	0.198	0.483	0.336
H_o	0.429	0.209	0.158	0.100	0.140	0.442	0.306

Appendix 9.5 continued

Locus	North Otago	Otago P.	Catlins	Stewart Island	Codfish Island	Auck. Islands	Camp. Island
Man13							
N	35	86	38	40	50	52	49
122	0.300	0.279	0.197	0.250	0.350	0.250	0.133
128	0.400	0.483	0.592	0.463	0.520	0.740	0.806
130	0.300	0.238	0.210	0.288	0.130	0.010	0.061
H_e	0.670	0.636	0.574	0.649	0.596	0.393	0.332
H_o	0.829	0.593	0.579	0.625	0.560	0.481	0.388
Man21							
N	35	86	38	40	50	52	49
128	0.000	0.000	0.000	0.000	0.000	0.048	0.102
130	0.800	0.814	0.750	0.538	0.700	0.471	0.622
138	0.200	0.186	0.250	0.463	0.300	0.481	0.276
H_e	0.325	0.305	.0380	0.504	0.424	0.550	0.532
H_o	0.286	0.209	0.290	0.425	0.400	0.539	0.531
Man22							
N	35	86	38	40	50	52	49
126	0.000	0.000	0.013	0.075	0.010	0.212	0.184
134	0.986	0.994	0.987	0.913	0.990	0.702	0.725
136	0.014	0.000	0.000	0.000	0.000	0.010	0.000
138	0.000	0.006	0.000	0.013	0.000	0.077	0.092
H_e	0.029	0.012	0.263	0.164	0.020	0.461	0.437
H_o	0.029	0.012	0.263	0.176	0.020	0.462	0.449

Appendix 9.5 continued

Locus	North Otago	Otago P.	Catlins	Stewart Island	Codfish Island	Auck. Islands	Camp. Island
Man27							
N	35	86	38	40	50	52	49
136	0.000	0.000	0.000	0.000	0.000	0.144	0.245
150	1.000	1.000	1.000	1.000	1.000	0.856	0.755
H_e	0.000	0.000	0.000	0.000	0.000	0.249	0.374
H_o	0.000	0.000	0.000	0.000	0.000	0.289	0.367
Man39							
N	35	86	38	39	50	52	49
137	0.186	0.233	0.355	0.269	0.260	0.587	0.622
139	0.000	0.000	0.013	0.000	0.000	0.221	0.031
143	0.514	0.436	0.500	0.436	0.390	0.164	0.316
145	0.300	0.331	0.132	0.295	0.350	0.010	0.020
147	0.000	0.000	0.000	0.000	0.000	0.019	0.010
H_e	0.620	0.650	0.614	0.659	0.664	0.586	0.516
H_o	0.714	0.698	0.679	0.462	0.740	0.404	0.571

Appendix 9.5 continued

Locus	North Otago	Otago P.	Catlins	Stewart Island	Codfish Island	Auck. Islands	Camp. Island
Man47							
N	35	86	38	39	50	52	49
123	0.257	0.215	0.211	0.115	0.130	0.240	0.133
126	0.000	0.000	0.000	0.000	0.000	0.010	0.000
129	0.000	0.012	0.000	0.000	0.010	0.164	0.347
132	0.000	0.000	0.000	0.000	0.000	0.067	0.122
135	0.486	0.552	0.632	0.756	0.550	0.356	0.398
138	0.000	0.000	0.000	0.000	0.000	0.010	0.000
141	0.257	0.209	0.145	0.128	0.310	0.154	0.000
144	0.000	0.012	0.013	0.000	0.000	0.000	0.000
H_e	0.641	0.608	0.543	0.403	0.590	0.768	0.696
H_o	0.600	0.535	0.526	0.359	0.640	0.692	0.653
Man50							
N	35	86	38	36	50	52	49
103	0.900	0.855	0.855	0.750	0.950	0.721	0.735
112	0.100	0.145	0.145	0.250	0.050	0.279	0.265
H_e	0.183	0.250	0.251	0.380	0.096	0.406	0.394
H_o	0.200	0.244	0.237	0.333	0.100	0.481	0.490

Appendix 9.5 continued

Locus	North Otago	Otago P.	Catlins	Stewart Island	Codfish Island	Auck. Islands	Camp. Island
Man51							
N	35	86	38	37	50	52	49
131	0.000	0.000	0.000	0.000	0.000	0.019	0.000
134	0.600	0.622	0.526	0.568	0.540	0.490	0.378
137	0.114	0.041	0.040	0.135	0.120	0.183	0.102
140	0.100	0.041	0.026	0.081	0.080	0.077	0.153
143	0.186	0.262	0.316	0.135	0.250	0.067	0.010
146	0.000	0.006	0.000	0.000	0.000	0.067	0.000
152	0.000	0.029	0.079	0.081	0.010	0.029	0.020
155	0.000	0.000	0.000	0.000	0.000	0.029	0.000
158	0.000	0.000	0.000	0.000	0.000	0.029	0.214
164	0.000	0.000	0.000	0.000	0.000	0.010	0.102
167	0.000	0.000	0.000	0.000	0.000	0.000	0.020
170	0.000	0.000	0.013	0.000	0.000	0.000	0.000
H_e	0.591	0.544	0.623	0.637	0.631	0.715	0.774
H_o	0.743	0.605	0.605	0.595	0.520	0.596	0.776
Man54							
N	35	86	37	39	50	52	49
133	0.000	0.000	0.000	0.000	0.000	0.096	0.000
136	0.686	0.587	0.581	0.692	0.670	0.702	0.786
139	0.314	0.413	0.419	0.308	0.330	0.173	0.214
142	0.000	0.000	0.000	0.000	0.000	0.129	0.000
H_e	0.437	0.488	0.494	0.432	0.447	0.472	0.340
H_o	0.400	0.477	0.568	0.564	0.540	0.442	0.306

Appendix 9.5 continued

Locus	North Otago	Otago P.	Catlins	Stewart Island	Codfish Island	Auck. Islands	Camp. Island
Man55							
N	35	86	38	37	50	52	49
146	0.000	0.000	0.000	0.000	0.000	0.010	0.000
152	0.000	0.000	0.000	0.000	0.000	0.019	0.357
158	1.000	1.000	1.000	1.000	1.000	0.971	0.643
H_e	0.000	0.000	0.000	0.000	0.000	0.057	0.464
H_o	0.000	0.000	0.000	0.000	0.000	0.058	0.388

Appendix 9.6 Historic *M. antipodes* samples used for DNA analyses presented in Chapter 5.

Museum abbreviation and accession number, laboratory code, collection location and collection year. The following museum abbreviations are used: AM = Auckland Museum, AMNH = American Natural History Museum, AUM = Australian Museum, CM = Canterbury Museum, MCZ = Museum of Comparative Zoology, MHNG = Natural History Museum Geneva, MNZ = Museum of New Zealand Te Papa Tongarewa, NHMT = Natural History Museum Tring, NMNHP = Natural History Museum Paris, NMW = Natural History Museum Vienna, NRM = Swedish Museum of Natural History, OM = Otago Museum, SAMA = South Australian Museum, USNM = Smithsonian Institution, ZMB = Museum für Naturkunde Berlin. Location abbreviation SI = South Island.

Museum accession no.	Lab code	Collection location	Collection year
AM LB5045 [a,b]	YM12	Stewart Island	1932
AM LB5046[a,b]	YM13	Stewart Island	1935
AM LB5047[a,b]	YM14	Stewart Island	1935
AMNH 525843	YM01	Otago Coast, SI	1895
AMNH 525844	YM02	Otago Heads, SI	1895
AMNH 525845	YM03	Otago Heads, SI	1895
AUM O.23935[a]	YM49	Otago, SI	1915
AUM O.37154[a]	YM50	Catlins River Mouth, SI	1938
AUM O.37155	YM51	Catlins River Mouth, SI	1938
AUM O.37156	YM52	Catlins River Mouth, SI	1938
AUM O.37157 [a]	YM53	Catlins River Mouth, SI	1938
AUM O.37158	YM54	Catlins River Mouth, SI	1938
AUM O.37159	YM55	Catlins River Mouth, SI	1938
CM AV801	YM15	Otago Coast, SI	1895
CM AV802	YM16	Otago, SI	1895
CM AV862	YM17	Dunedin, SI	1937
MCZ 86732	YM20	South Island	1895
MHNG 754.46[a]	YM26	Otago Peninsula, SI	1911

Appendix 9.6 continued

Museum accession no.	Lab code	Collection location	Collection year
MNZ OR.5379	YM45	Otago Heads, SI	1895
MNZ OR.5383	YM46	Otago Peninsula, SI	1895
MNZ OR.5385	YM47	Otago Heads, SI	1895
MNZ OR.5386	YM48	Stewart Island	1888
NHMT 1897.12.6.40[a]	YM22	Otago, SI	1895
NMW 4385	YM27	Stewart Island	1888
NMW 4386	YM28	Stewart Island	1888
NMW 4402	YM29	Stewart Island	1888
NMW 4403	YM30	Stewart Island	1888
NMW 4404	YM31	Stewart Island	1888
NMW 4405	YM32	Stewart Island	1888
NMW 4406	YM33	Stewart Island	1888
NMW 4408	YM34	Stewart Island	1888
OM AV034	YM35	Otago Heads, SI	1895
SAMA B13913	YM38	Otago, SI	1911
USNM 124683[a]	YM36	Stewart Island	1891
ZMB2000.8244	YM19	Stewart Island	1888

[a] No or only sporadic amplification, not included in genetic analyses.
[b] DNA extracts kindly provided by C. Millar.

Appendix 9.7 Allele frequencies for contemporary and historic South Island and contemporary subantarctic yellow-eyed penguins at 10 microsatellite loci

Allele frequencies, sample size/ locus (N), expected heterozygosity (H_e) and observed heterozygosity (H_o).

Locus	Historic South Island	Contemporary South Island	Contemporary subantarctic
Mano3			
N	27	249	101
112	0.000	0.000	0.025
116	1.000	1.000	0.901
118	0.000	0.000	0.005
120	0.000	0.000	0.069
H_e	0.000	0.000	0.184
H_o	0.000	0.000	0.198
Mano8			
N	27	249	101
116	0.889	0.880	0.733
122	0.111	0.106	0.050
126	0.000	0.014	0.218
H_e	0.201	0.215	0.415
H_o	0.148	0.201	0.376

Appendix 9.7 continued

Locus	Historic South Island	Contemporary South Island	Contemporary subantarctic
Man13			
N	24	249	101
122	0.167	0.279	0.193
128	0.688	0.492	0.772
130	0.146	0.229	0.035
H_e	0.489	0.629	0.367
H_o	0.458	0.623	0.436
Man21			
N	22	249	101
128	0.000	0.000	0.074
130	0.636	0.735	0.545
138	0.364	0.265	0.381
H_e	0.474	0.390	0.555
H_o	0.364	0.305	0.535
Man39			
N	23	248	101
137	0.500	0.256	0.604
139	0.000	0.002	0.129
143	0.304	0.448	0.238
145	0.196	0.294	0.015
147	0.000	0.000	0.015
H_e	0.633	0.649	0.565
H_o	0.652	0.653	0.485

Appendix 9.7 continued

Locus	Historic South Island	Contemporary South Island	Contemporary subantarctic
Man47			
N	24	248	101
123	0.396	0.188	0.188
126	0.000	0.000	0.005
129	0.000	0.006	0.253
132	0.000	0.000	0.094
135	0.458	0.587	0.376
138	0.000	0.000	0.005
141	0.146	0.214	0.079
144	0.000	0.006	0.000
H_e	0.625	0.576	0.748
H_o	0.542	0.536	0.673
Man50			
N	27	245	101
103	0.833	0.865	0.728
112	0.167	0.135	0.272
H_e	0.283	0.234	0.398
H_o	0.333	0.220	0.485

Appendix 9.7 continued

Locus	Historic South Island	Contemporary South Island	Contemporary subantarctic
Man51			
N	27	246	101
131	0.000	0.000	0.010
134	0.741	0.579	0.436
137	0.037	0.081	0.144
140	0.000	0.061	0.114
143	0.222	0.238	0.040
146	0.000	0.002	0.035
152	0.000	0.037	0.025
155	0.000	0.000	0.015
158	0.000	0.000	0.119
164	0.000	0.000	0.055
167	0.000	0.000	0.010
170	0.000	0.002	0.000
H_e	0.408	0.597	0.760
H_o	0.444	0.606	0.683
Man54			
N	26	247	101
133	0.000	0.000	0.050
136	0.615	0.634	0.743
139	0.385	0.366	0.193
142	0.000	0.000	0.015
H_e	0.483	0.465	0.411
H_o	0.385	0.506	0.376

Appendix 9.7 continued

Locus	Historic South Island	Contemporary South Island	Contemporary subantarctic
Man55			
N	22	246	101
146	0.000	0.000	0.005
152	0.000	0.000	0.183
158	1.000	1.000	0.812
H_e	0.000	0.000	0.309
H_o	0.000	0.000	0.218

Appendix 9.8 Historic *M. antipodes* samples used for DNA analyses presented in Chapter 6.

Museum abbreviation and accession number, laboratory code, purported collection location and collection year. The following museum abbreviations are used: AM = Auckland Museum, AMNH = American Natural History Museum, AUM = Australian Museum, CM = Canterbury Museum, MCZ = Museum of Comparative Zoology, MHNG = Natural History Museum Geneva, MNZ = Museum of New Zealand Te Papa Tongarewa, NHMT = Natural History Museum Tring, NMNHP = Natural History Museum Paris, NMW = Natural History Museum Vienna, NRM = Swedish Museum of Natural History, OM = Otago Museum, SAMA = South Australian Museum, USNM = Smithsonian Institution, ZMB = Museum für Naturkunde Berlin. Location abbreviations are as follows: AI = Auckland Islands, CI = Campbell Island, SI = South Island, STI = Stewart Island.

Museum accession no.	Lab code	Collection location	Collection year
AM LB5045 [a,b]	YM12	Stewart Island	1932
AM LB5046[a,b]	YM13	Stewart Island	1935
AM LB5047[a,b]	YM14	Stewart Island	1935
AMNH 525843	YM01	Otago Coast, SI	1895
AMNH 525844	YM02	Otago Heads, SI	1895
AMNH 525845	YM03	Otago Heads, SI	1895
AMNH 525849	YM04	Campbell Island	1894
AMNH 525850	YM05	Campbell Island	1893
AMNH 525851	YM06	Campbell Island	1894
AMNH 525852	YM07	Campbell Island	1893
AMNH 525853	YM08	Auckland Islands	1893
AMNH 525854	YM09	Auckland Islands	1893
AMNH 525855	YM10	Auckland Islands	1893
AMNH 525856	YM11	Auckland Islands	1894
AUM O.23935[a]	YM49	Otago, SI	1915
AUM O.37154[a]	YM50	Catlins River Mouth, SI	1938
AUM O.37155	YM51	Catlins River Mouth, SI	1938

Appendix 9.8 continued

Museum accession no.	Lab code	Collection location	Collection year
AUM O.37156	YM52	Catlins River Mouth, SI	1938
AUM O.37157[a]	YM53	Catlins River Mouth, SI	1938
AUM O.37158	YM54	Catlins River Mouth, SI	1938
AUM O.37159	YM55	Catlins River Mouth, SI	1938
CM AV801	YM15	Otago Coast, SI	1895
CM AV802	YM16	Otago, SI	1895
CM AV862	YM17	Dunedin, SI	1937
MCZ 86732	YM20	South Island	1895
MHNG 754.46[a]	YM26	Otago Peninsula, SI	1911
MNZ OR.12968	YM40	Perseverance Harbour, CI	1943
MNZ OR.12969	YM41	Penguin Bay, AI	1942
MNZ OR.12970	YM42	Auckland Islands	1942
MNZ OR.12971	YM43	Ocean Island, AI	1943
MNZ OR.5378	YM44	Campbell Island	1944
MNZ OR.5379	YM45	Otago Heads, SI	1895
MNZ OR.5383	YM46	Otago Peninsula, SI	1895
MNZ OR.5385	YM47	Otago Heads, SI	1895
MNZ OR.5386	YM48	Stewart Island	1888
NHMT 1842.12.16.165[a]	YM21	Auckland Islands	1840
NHMT 1897.12.6.40	YM22	Otago, SI	1895
NHMT 1901.1.7.15	YM23	Campbell Island	1899
NHMT 1901.1.7.16	YM24	Campbell Island	1899
NHMT 1905.12.30.233	YM25	Auckland Islands	1904
NMNHP CG: 1875-522[a]	YM18	Campbell Island	1875
NMW 4385	YM27	Stewart Island	1888
NMW 4386	YM28	Stewart Island	1888
NMW 4402	YM29	Stewart Island	1888
NMW 4403	YM30	Stewart Island	1888
NMW 4404	YM31	Stewart Island	1888

Appendix 9.8 continued

Museum accession no.	Lab code	Collection location	Collection year
NMW 4405	YM32	Stewart Island	1888
NMW 4406	YM33	Stewart Island	1888
NMW 4408	YM34	Stewart Island	1888
NRM 569465[a]	YM39	Campbell Island	1924
OM AV034	YM35	Otago Heads, SI	1895
SAMA B13913	YM38	Otago, SI	1911
USNM 124683[a]	YM36	Stewart Island	1891
USNM 15655[a]	YM37	Auckland Islands	1840
ZMB2000.8244	YM19	Stewart Island	1888

[a] No or only sporadic amplification, not included in genetic analyses.
[b] DNA extracts kindly provided by C. Millar.

Appendix 9.9 Presented conference papers and invited talks

Boessenkool S, Seddon PJ & Waters JM. Rare migration events despite range expansion: spatiotemporal genetic analyses reveal the dynamic history of yellow-eyed penguins. European Science Foundation, ConGen 2008, Trondheim, Norway.

Boessenkool S. Identifying management units of yellow-eyed penguin. Yellow-eyed penguin Recovery Group Meeting 2008, Department of Conservation. Dunedin, New Zealand.

Boessenkool S, Austin JJ, Worthy TH, Cooper A, Seddon PJ & Water JM. Penguin on the menu: prehistoric DNA reveals cryptic extinction and colonisation within 500 yrs of human settlement in New Zealand. Joint meeting of the Society for the Study of Evolution, Society of Systematic Biologists and American Society of Naturalists 2008, Minneapolis, United States of America.

Boessenkool S. Genetics and conservation of yellow-eyed penguins. Ornithological Society of New Zealand 2008. Dunedin, New Zealand.

Boessenkool S, Seddon PJ & Waters JM. The SLOSS of the penguin: do yellow-eyed penguins exhibit single large or several small populations? Joint meeting of the Society for the Study of Evolution, Society of Systematic Biologists and American Society of Naturalists 2007, Christchurch, New Zealand.

Boessenkool S. Genetics as a tool in yellow-eyed penguin conservation. Threatened Bird Research Workshop II 2006, Waianakarua, New Zealand.

Appendix 9.10 Coverage of research in popular press (selection)

New York Times (21 November 2008) Studying rare penguin, scientist finds new penguin. Henry Fountain.

Otago Daily Times (20 November 2008) Otago Researchers find ancient penguin species. Rebecca Fox.

New Zealand Herald (20 November 2008) 500 years later, rare penguin still new.

Reuters (20 November 2008) Researchers stumble upon new penguin species. Pauline Askin.

BBC News (19 November 2008) Rare penguin took over from rival.

ABC News (19 November 2008) Scientists find new penguin, extinct for 500 years. Ray Lilley.

Otago Daily Times (19 January 2008) Diving deep into penguin's past. Rebecca Fox.