Population genetic relationships in the roseate tern: globally, regionally and locally

Thesis submitted by

Anna Lashko

B. Sc. (Hons) ANU

in May 2004

for the degree of Doctor of Philosophy

in the School of Tropical Environment Studies and Geography

James Cook University

STATEMENT OF ACCESS

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

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and;

I do not wish to place any further restriction on access to this work.

…… ……..

Anna Lashko Date

Abstract

The roseate tern, *Sterna dougallii*, is a colonially nesting seabird which breeds on tropical and temperate islands in the Atlantic, Indian and Pacific Oceans. The roseate tern is considered threatened in a number of regions, including the United States, Europe and South Africa. There are currently four or five recognised subspecies of roseate tern: one in the Atlantic Ocean; and three or four in the Indo-Pacific, with many authors questioning the validity of currently circumscribed subspecies.

In this study, two kinds of genetic marker, ND2 and ND6 mitochondrial DNA sequences and microsatellite genotyping, were used to investigate the historical and contemporary relationships among roseate tern breeding colonies. I used samples from breeding colonies in Australia, Japan, Seychelles, South Africa, Azores, Ireland and the United States to investigate population structure in the roseate tern at a global scale, a regional scale and at a local scale within Australia. In addition, samples obtained from the non-breeding aggregations in Australia were included to determine whether mitochondrial DNA or microsatellite markers can be used to determine the origins of individuals in non-breeding aggregations.

Sequencing of a combined total of 1381 base pairs of the ND2 and ND6 genes, and genotyping with five polymorphic microsatellite markers, revealed two distinct lineages of roseate tern: an Atlantic lineage; and an Indo-Pacific lineage. The genetic divergence between the two lineages (3.9% corrected mitochondrial DNA sequence divergence) represents a significant historic separation, with continental Africa implicated as the major long term barrier to gene flow between the two oceanic basins. The star-shaped phylogeny observed for both lineages suggests a recent range expansion from a

common source in each oceanic basin, which is likely to have coincided with the stabilising of sea level following the last glacial period, approximately 20 000 years ago. The lack of population structure in mitochondrial DNA haplotypes within each oceanic basin suggests that cycles of range expansion and contraction to a glacial refuge have occurred during interglacial and glacial periods, respectively. This periodic range contraction, to a common glacial refuge within each oceanic basin, has allowed regular gene flow to occur among individuals within each oceanic basin, ameliorating any genetic divergence that may develop during interglacial periods. The most recent range expansion has resulted in secondary contact and introgression of the two lineages, identified in the breeding colony sampled in South Africa.

The Atlantic subspecies, *S. d. dougallii*, is well supported by this study, but the currently recognised subspecies in the Indo-Pacific are not warranted based on genetic data. I recommend a formal revision of the subspecies designations, merging the three or four currently recognised subspecies in the Indo-Pacific.

Microsatellite markers reveal that population structuring has developed among some breeding regions since the last range expansion. In the Atlantic Ocean, microsatellite analyses suggest that breeding colonies in Ireland and the United States have diverged from the Azores ($R_{ST} = 0.28 – 0.36$, $P < 0.05$); however, due to the potential influence of population declines on the genetic variation in these colonies, I cannot reject the alternate hypothesis that individuals from the north Atlantic breeding colonies are regularly interbreeding.

In the Indo-Pacific, the breeding colony in the Seychelles is divergent from the colonies in Australia and Japan ($R_{ST} = 0.14 – 0.22$, $P < 0.05$), suggesting that the Indian Ocean serves as a current barrier to gene flow. There is no evidence for genetic divergence

among colonies in Australia and Japan (R_{ST} = 0.03, P = 0.22), indicating that gene flow between the two regions, and among colonies in Australia, is occurring regularly. This may be facilitated via the mixing of individuals from different breeding colonies in the non-breeding aggregations. Due to the lack of genetic structure among breeding colonies in Australia and Japan, neither mitochondrial DNA or microsatellites could be used to determine the origins of individuals in non-breeding aggregations in Australia.

During this study, I obtained evidence of hybridisation between roseate and black-naped terns. Further investigation of this revealed that paternal leakage, the transfer of paternal mitochondrial DNA along with the usual maternal transfer, had occurred as a result of the hybridisation event. This is the first evidence of hybridisation between the roseate and black-naped tern, and one of only a handful of instances of paternal leakage of mitochondrial DNA in animals.

Acknowledgments

First and foremost I must thank my supervisors, Emma Gyuris and Michelle Waycott. But for their backing, support and guidance I would not be writing this today. They gave me flexibility and freedom, but were always readily available for discussions or draft-reading. Thanks Michelle for the level-headedness and all your help in the lab - the optimism when I had lost mine and for ideas when all seemed to have been tried. Thanks to Emma for your honesty, encouragement and enthusiasm, particularly when things weren't looking so good for a while.

Many people have contributed to my memories of unique and amazing field trips: I shared the Great Barrier Reef with many Queensland Parks and Wildlife staff – for their essential roles I would like to mention Bruce Knuckey, Kete Pelling and particularly Paul O'Neill who not only assisted with sample collection and logistics but with whom I shared many important discussions and of course an amazing excursion to Japan. In Japan, arigatõ gozaimasu to Takao Baba, Shigemoto Komeda and Kiyoaki Ozaki for sharing their field trip with me, and for dealing with my primary school Japanese. In Western Australia, Nic Dunlop made me the captain of the boat (and made me swim to get it!) and Chris Surman was the perfect host during countless days of howling winds and the occasional piratical expedition (arrrrr). I also couldn't have answered these questions without the people that have helped with collection of samples in places I was unable to get to: Clive Minton, Veronica Neves, Stephen Newton, Carlos Pacheco, Jaime Ramos, Patty Szczys, Tony Tree, Dick Veitch and Rebecca White.

Special thanks to Patty Szczys for her trust in sharing her microsatellites with me, and to Julie Feinstein for the initial ND6 sequence that got everything started. In the lab here many people have made the days pass more quickly: Ainsley Calladine (A), Mark 'the dish' Disher, Jennifer Kunzelman, Brenda McDonald, Jessica Moss (my colony picking and creative dancing queen) and Grant Whiteman (who got me up early for tennis). Thanks to Glenn Dunshea for joining me in the wrestle with the roseate tern. Scott Smithers helped me with the world of geomorphology, David Blair and Jon Luly provided valuable comments when some outside perspective was required, and Tim Harvey's fine tooth comb picked up on things at the end that I was never going to see.

Shirley Bruce, Clive Grant, Sue Knight and Rob Scott answered countless queries about paperwork or computers, and Adella Edwards helped with maps.

Financial assistance for this project was provided by the Cooperative Research Centre for the Great Barrier Reef World Heritage Area, the Raine Island Corporation and the School of Tropical Environment Studies and Geography.

To the collection of other students who have shared this crazy time with me, I can't imagine what it would have been like feeling as though I was the only one and am thankful for the constant reminder that it wasn't so. My regular uni 'housemates' Ameer, Ramona and Nadine made the office an easier place to be on weekends and at night with their regular passings by. Thanks to Ramona and Liz for breaking up each day (with food!) and especially to Liz for surviving the last few months in the office with me, helping me to find the right word. Rohan and Ben gallantly put up with being my entertainers in the last few months. Most especially to Vimoksalehi and Amanda, for providing me with an office to retreat to when everything got too insane and escape was the only option, and dealing with every up and down in the last few years. My housemates in recent years, Guido and Monica, have dealt with the traumas with good humour and forgiven me my outbursts without question – thanks guys. And last but not least of my fellow students, to Amanda for being in the same place, for understanding and sympathising and for food, laughs and a heap of trashy television – thanks maaate!

Finally, thanks to all those at home and further away, especially my mum and the girls – Anita, Faye, Imogen, Jane, Lara and Zoe - for their endless support, for helping me to forget (what's a PhD?) when I wasn't here and for always believing in me.

Table of Contents

List of Figures

List of Tables

STATEMENT OF SOURCES DECLARATION

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

……. …….

Signature Date

Chapter 1: Introduction

1.1 Study species – the roseate tern

The roseate tern, *Sterna dougallii*, is a medium sized tern (Gochfeld *et al*. 1998), named for the rosy hue on its breast when in breeding plumage. When in breeding plumage the roseate tern also has a black cap, and red bill and legs (Figure 1.1). The species nests colonially on islands in tropical and temperate regions of the Atlantic, Indian and Pacific Oceans, including Europe, the United States, Canada, the Caribbean, South Africa, the Seychelles, Kenya, India, Japan, Australia and the Pacific Islands (Figure 1.2). Nisbet (1980) estimated that at least 25 000 pairs of roseate terns bred in the Indian Ocean, and cited reports of several thousand pairs breeding in Australia. A few years later, Gochfeld (1983) estimated the world population to be between 20 000 and 55 500 pairs, with a maximum estimate of 8 000 breeding individuals in Australia. While accurate numbers were available for colonies in North America and Europe, where regular studies were in progress at the time of Gochfeld's review, many colonies in the Indo-Pacific, particularly in Australia, were yet to be discovered or documented.

In the past 20 years, surveys have revealed that Australia has the world's largest population of breeding roseate terns. Many large breeding colonies have been reported in Australia (King & Limpus 1989; Burbidge & Fuller 1996; Chatto 2001), with a recent review estimating that 30 000 pairs breed in colonies around the Australian coast (O'Neill *et al*. in press). In addition, a combined total of 50 000 - 60 000 non-breeding individuals aggregate in two regions of Australia each year: the Swain Reefs in the

Figure 1.1: A roseate tern, *Sterna dougallii*, in breeding plumage. The red bill and legs are obvious features, but note also the subtle rosy hue on the breast, for which the species is named. Photo: Kiyoaki Ozaki.

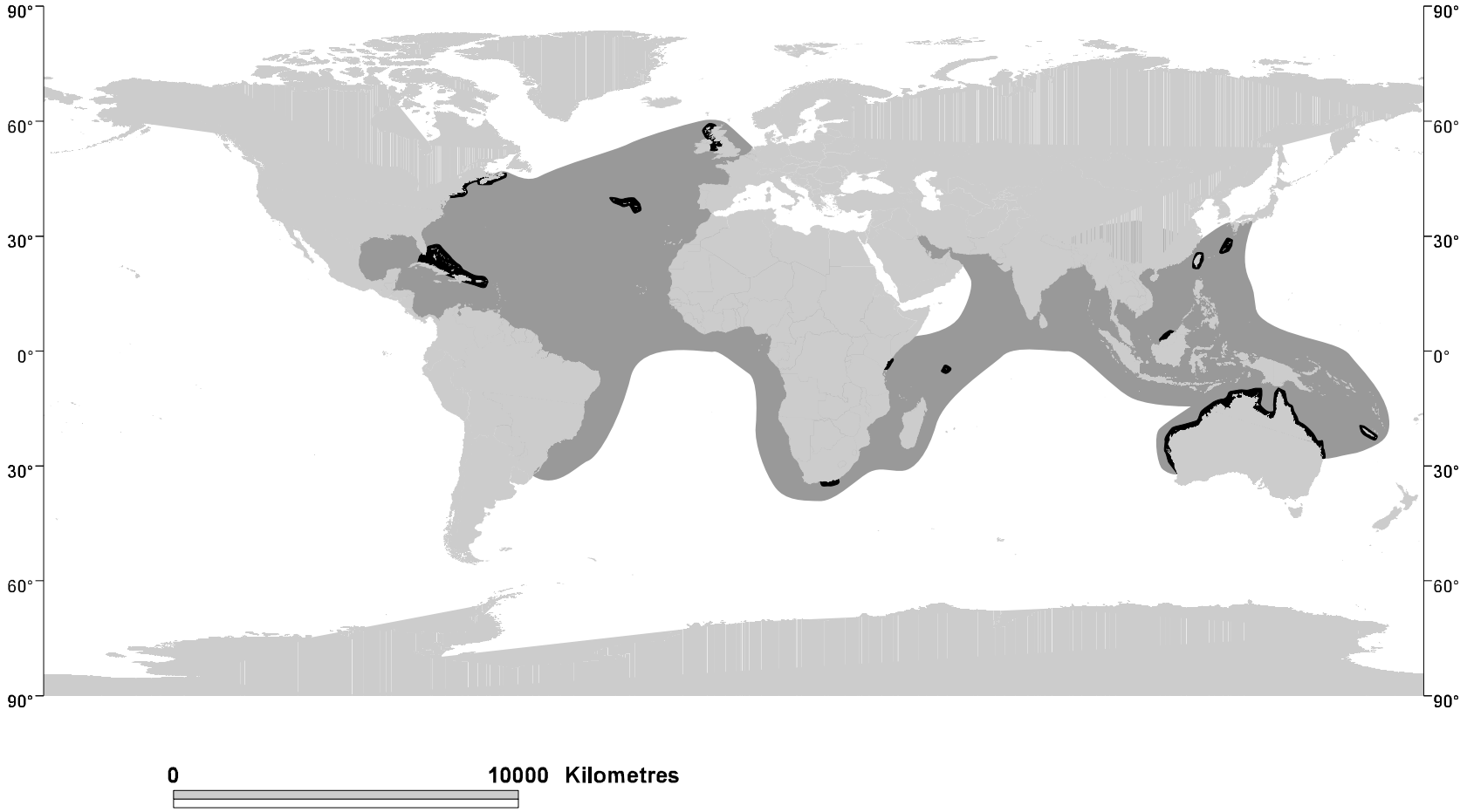


Figure 1.2: Approximate range (darker grey) and regularly recorded breeding areas (black) of the roseate tern, *Sterna dougallii*, based on a compliation of published records.

southern Great Barrier Reef (O'Neill *et al.* in press) and the Lacepede Islands in northern Western Australia (C. Minton, pers. comm.).

In contrast to the species' secure status in Australia, the roseate tern is listed as threatened in Canada (Kirkham & Nettleship 1987), and endangered in the U. S. (U. S. Fish and Wildlife Service 1987), Europe (Tucker & Heath 1994) and South Africa (Brooke 1984), with population declines observed in each locality (Witherby & Ticehurst 1908; Nisbet 1973; Cramp *et al.* 1974; Nisbet 1980; Randall & Randall 1980; Kirkham & Nettleship 1987). Particularly in Europe and North America, concern over the species' decline has lead to regular monitoring and research occurring in most major breeding colonies, which has been aided by the accessibility of colonies in these regions. The wide geographical range over which roseate tern colonies occur in Australia, in conjunction with their often remote location and non-threatened status, has resulted in a comparative dearth of information on roseate terns within Australia.

Current banding research on roseate terns in Australia is centred on the non-breeding aggregations in the Swain Reefs, Great Barrier Reef. Recent results have revealed that at least some of the roseate terns breeding in Japan and Taiwan migrate to the Swain Reefs during the northern hemisphere winter, and mix with non-breeding Australian roseate terns (O'Neill *et al.* in press). Despite this important finding, this banding research cannot, nor is it intended to, provide information about the movements of birds between breeding colonies within Australia, which is needed to better understand, manage and conserve the world's largest population of roseate terns.

While colony sizes and the timing of breeding of roseate terns are often reported in surveys which encompass a number of seabird species (Serventy & White 1951; Hulsman 1984; Burbidge & Fuller 1996; Blaber *et al.* 1998; Surman 1998; Chatto 2001), only a small number of studies have examined the breeding biology or behaviour of roseate terns in any detail (Hulsman 1976; Hulsman 1977a; Hulsman 1977b; Smith 1991; Milton *et al.* 1996; Surman & Wooller 2003). To date, no study has investigated movements of roseate terns between breeding colonies in Australia.

1.2 Philopatry in seabirds

Most species of seabird are thought to exhibit philopatry, returning to their natal colony to breed. For example, a review of band recoveries of 11 species of Antarctic seabirds, including albatrosses, petrels and skuas, found rates of philopatry close to 100% (Weimerskirch *et al.* 1985). Banding studies, however, are limited by the difficulty of banding entire populations, low band recoveries and the difficulty of surveying large areas in a single year. In particular, long distance dispersal is often underestimated as surveys for banded birds are usually restricted to areas close to the colonies where banding has occurred (Crochet 1996; Koenig *et al.* 1996). An additional bias arises because the density of dispersing birds decreases as dispersal distance increases (ie. birds dispersing one kilometre can settle on a 3.14 km circumference, while those dispersing 10 km can settle on a 31.4 km circumference). Therefore, even if an equal number of individuals disperse varying distances, there is a higher probability of encountering individuals that have dispersed the shortest distance. The limitations of banding studies are particularly pertinent in large, geographically widespread populations, as is the case for the roseate tern in Australia where colonies, in addition to

being widespread, are often extremely difficult to access, prohibiting regular surveys for banded individuals.

In areas where roseate terns have been the subject of banding studies, there is evidence for philopatry. A metapopulation analysis of 11 years of data from three major roseate tern breeding colonies in the United States found that the average probability of an individual returning to its natal colony to breed was 0.8 (range 0.53, in a declining colony, to 0.91; Lebreton *et al*. 2003). Once individuals commenced breeding, fidelity to a breeding colony was strong, with the probability of adults returning to breed at the same colony in consecutive years ranging from 0.89, in the declining colony, to 0.98 (average 0.95; Lebreton *et al*. 2003). These results confirmed those from an earlier analysis of a smaller data set (five years and four breeding colonies) which found that over 90% of surviving adults demonstrated fidelity to their breeding colony (Spendelow *et al*. 1995). The same level of philopatry or adult colony fidelity is not expected in Australia, as roseate tern colonies do not always occur on the same islands, often moving between islands in a group or region in consecutive years (Hulsman 1984; Fuller *et al*. 1994; Milton *et al*. 1996; Surman 1998).

Another method of examining philopatry is using genetic methods. In species with strong philopatry there will be limited migration between breeding colonies, and hence little gene flow. This provides a scenario that promotes genetic divergence through random genetic drift and local adaptation (Slatkin 1987). Genetic divergence among breeding colonies is thus suggestive of high rates of philopatry in a species. In Australia, because of the tendency for roseate tern breeding colonies to move between islands in consecutive years, individuals may exhibit philopatry to a region or island

group rather than to a specific island. Genetic structure cannot, therefore, exist on the scale of individual islands. However, if roseate terns display group fidelity, with individuals moving together when colony sites move, genetic structure may occur at the scale of island groups or regions.

1.3 Genetic markers as tools for identifying population structure

Genetic methods are now commonly used to investigate population structure, with different genetic markers employed depending on the focus of the study. In this thesis I use two of the most commonly used genetic markers: mitochondrial DNA sequences and multilocus genotypes based on microsatellites. The different characteristics of these two types of genetic markers make them appropriate for examining population structure at varying scales, and they are thus often used in tandem to provide a combination of broad and fine scale information about population structuring within a species.

1.3.1 Mitochondrial DNA

The mitochondrial genome is a circular, independently replicating genome. It is maternally inherited (Hutchison III *et al.* 1974; Giles *et al.* 1980; Birky 1995) and non-recombining (Eyre-Walker & Awadalla 2001, but see Eyre-Walker *et al.* 1999a; Eyre-Walker *et al.* 1999b; Macaulay *et al.* 1999; Ladoukakis & Zouros 2001; Rokas *et al.* 2003 for current debate on the subject). As a result, mitochondrial DNA is generally considered as a haploid, clonally inherited genome which reflects female-mediated processes. The clonal mode of inheritance makes understanding evolutionary processes relatively straightforward in comparison to studies using the nuclear genome.

Mitochondrial DNA has a mutation rate approximately 5 - 10 times greater than generally recognised in nuclear DNA, with mutations estimated to occur an average of 0.02 times per nucleotide per million years (Brown *et al*. 1979), though this varies in different regions of the mitochondrial genome (Cann *et al*. 1984). This makes mitochondrial DNA a valuable tool for studying the relationships among closely related taxonomic groups as, at the estimated rate of mutation, multiple mutations at nucleotide sites begin to obscure relationships beyond 10 - 20 million years previous (Brown *et al*. 1979).

Mitochondrial DNA is commonly used for studies of phylogeny, phylogeography and population genetics in animals. Abbott & Double (2003b) used the mitochondrial control region for a phylogeographic study of shy, *Thalassarche cauta*, and white-capped, *T. steadi*, albatrosses, with the aim of clarifying the evolutionary relationship between the two species, and their taxonomic status. The two species could be differentiated by a single fixed nucleotide difference, which demonstrated a lack of female-mediated gene flow between the two species, but reflected low sequence divergence in comparison with other interspecies studies. Haplotype diversity was lower in shy albatrosses, and the observed haplotypes were more closely related to one another, than in white-capped albatrosses, suggesting that the shy albatross originated more recently than the white-capped albatross. A nested clade analysis (Templeton 2001) supported this hypothesis, predicting that shy albatrosses had been founded via range expansion of white-capped albatrosses. The combined evidence for lack of gene flow between the two species and low divergence suggests that speciation of the shy and white-capped albatrosses is an ongoing or very recent occurrence (Abbott & Double 2003b).

Because of its clonal mode of inheritance, mitochondrial DNA is useful for identifying introgression, the incorporation of genes from one distinct taxon into another (Rieseberg & Wendel 1993). Unlike nuclear DNA, which undergoes recombination during meiosis, preventing the identification of either parent's specific contribution to their offspring, the mitochondrial haplotype of an individual represents the maternal contribution. A study using mitochondrial DNA identified the complete introgression of arctic charr mitochondrial DNA into a population of lake trout in Lac des Chasseurs, Québec (Wilson & Bernatchez 1998). Arctic charr, *Salvelinus alpinus*, and lake trout, *S. namaycush*, are morphologically and genetically distinct species, with 3% mitochondrial DNA divergence between the two species (Grewe *et al.* 1990). Forty seven fish which had typical lake trout morphology were analysed using restriction fragment length polymorphism (RFLP) of mitochondrial DNA, which was known to yield diagnostic DNA fragment patterns for arctic charr and lake trout. All sampled lake trout had haplotypes differing from typical lake trout haplotypes, and corresponding with the most common haplotype observed in arctic charr in nearby Lac Chaudière; there are no arctic charr in Lac des Chasseurs. Wilson & Bernatchez (1998) suggested that, following the last glacial retreat, arctic charr was one of the early colonisers of the area. The subsequent arrival of lake trout, which readily displaces arctic charr, provided a period of time in which hybridisation could occur between the two species prior to the complete displacement of arctic charr from Lac des Chasseurs. The arctic charr haplotypes in the lake trout then became fixed in the Lac des Chasseurs population via the founder effect and/or genetic drift in the small founding lake trout population, or through selection favouring arctic charr haplotypes (Wilson & Bernatchez 1998).

These two examples illustrate the utility of mitochondrial DNA for investigating historical relationships among taxa and the processes responsible for the observed relationships. In the first example, mitochondrial DNA revealed that one albatross species was derived from another through range expansion; in the second example mitochondrial DNA provided evidence of an historic interaction between two genetically distinct species following a climatic change. However, in neither example could mitochondrial DNA be used to infer modern-day processes; for this purpose a genetic marker with a higher mutation rate is required.

1.3.2 Microsatellites

Microsatellites, or simple sequence repeats (SSRs), are short, tandemly repeated DNA sequences (e.g. CACACACACA, denoted as $(CA)_5$) occurring throughout the nuclear genome (Tautz & Renz 1984). As biparentally inherited codominant markers, microsatellites provide information about genetic processes in both genders. They are non-coding regions of DNA, and consequently not subject to direct selection, although there appears to be a limit to their size (Garza *et al.* 1995; Wierdl *et al.* 1997; Huang *et al.* 2002; Calabrese & Durrett 2003). Because of their repetitive nature, mutations occur in the form of addition or loss of repeat units (in the example above, CA) via slippage of the DNA strands during DNA replication or repair (Levinson & Gutman 1987). Microsatellites have a high mutation rate, commonly estimated at between 10^{-3} and 10^{-4} (but can be as high as 10^{-2}) mutations per locus per generation (Hancock 1999), producing high levels of polymorphism within populations. However, due to their repetitive structure and high mutation rate, homoplasy, the generation of identical genotypes via multiple pathways, can be an issue. For example, an allele $(CA)_6$ which has arisen via the addition and subsequent loss of a single repeat unit $(CA)_6 \rightarrow (CA)_7 \rightarrow (CA)_6$ appears the same as a $(CA)_6$ allele which has not undergone any addition or loss

of repeat units. When examining processes occuring over longer time frames, the increasing occurrence of homoplasy obscures information about the evolutionary history of microsatellite alleles. As a result, microsatellites are best used to examine processes occurring over shorter time frames.

Microsatellite markers are thus most commonly used for intraspecies studies: to determine population structure, mating systems, kinship and parentage. Microsatellite loci were used to examine parentage in the Seychelles warbler, a cooperative breeder in which previous offspring, usually female, often remain in the natal territory and assist with raising subsequent young (Richardson *et al.* 2001). Fourteen variable microsatellites allowed parents to be identified with a high degree of confidence. Microsatellite genotyping revealed that helper females often gain maternity in their natal territory, with parentage assigned to 44% of helper females. A single male helper gained paternity, but a large number of extra-group fathers were identified, with 38% of young fathered by a male outside the natal territory; the extra-group male was always the primary male in his territory (Richardson *et al.* 2001). In the superb fairy-wren, also a cooperatively breeder, a similar study to that of Richardson *et* al (2001) revealed that 95% of broods contained young sired by an extra-pair male, with 76% of young sired by extra-group males (Mulder *et al.* 1994). These levels of extra-group paternity are some of the highest reported in birds, and demonstrates the potential for polymorphic microsatellites to differentiate even closely related individuals.

Scribner *et al.* (2003) used microsatellites to investigate population structure in the Canada goose, *Branta canadensis*, a species which is harvested by hunters in North America. Microsatellites revealed significant genetic differences between the two

subspecies, *B. c. interior* and *B. c. maxima*, and significant, but weaker, genetic structuring within either subspecies. These differences were subsequently used to determine whether harvested individuals could be identified, firstly to subspecies level, and secondarily to population. Individuals could be assigned to one of the two subspecies with a high degree of certainty (between 88.8 and 99.3% correct), although assignment to a population was less accurate. Microsatellite markers will thus be a valuable tool for assessing the relative contribution of each subspecies of Canada goose to the annual harvest, allowing the effects of the harvests on each subspecies to be addressed.

1.3.3 Combining genetic information from mitochondrial DNA and microsatellites

Mitochondrial DNA and microsatellite markers, used in combination, are a powerful genetic tool for understanding both historical and contemporary population genetic processes. Mitochondrial DNA provides information on historical processes, such as isolation of populations in glacial refugia, introgression and speciation. The more variable microsatellite markers provide information on contemporary processes, such as migration of individuals between populations (inferred from measures of gene flow), kinship and parentage. The comparison of a maternally inherited marker with a biparentally inherited marker allows any gender biases to be identified, such as sex-biased dispersal. Together the two types of genetic markers allow a greater understanding of evolutionary processes within a species than either can provide alone.

Both mitochondrial DNA and microsatellites showed evidence of a reduction in genetic variation in small populations of the greater prairie-chicken, *Tympanuchus cupido pinnatus*, a species which has suffered habitat fragmentation in North America (Johnson *et al*. 2003). However, the two kinds of genetic markers revealed varying degrees of genetic structure: mitochondrial DNA indicated minor differentiation of populations, with the sample populations divided into two significantly different groups; and microsatellite analyses identified five significantly different groups. Additionally, mitochondrial DNA revealed an isolation by distance effect, while microsatellites did not. The authors concluded that the results of mitochondrial DNA analyses reflected the period prior to habitat fragmentation, at which time populations of the greater prairie-chicken existed as a large metapopulation, hence the pattern of isolation by distance with little genetic structuring. The occurrence of the habitat fragmentation was too recent for mitochondrial DNA to reveal the current relationships among, or processes within, the now isolated populations. Microsatellites, however, provided evidence that genetic drift was the predominant force in the isolated populations, obscuring the historic pattern of isolation by distance. The combination of information from mitochondrial DNA and microsatellites demonstrated the importance of maintaining larger population sizes for each isolated population and/or maintaining gene flow between isolated populations to maximise genetic diversity in the greater prairie-chicken (Johnson *et al*. 2003).

The value of using a combination of mitochondrial and nuclear markers, such as microsatellites, was recognised by Moritz (1994) in his genetic criteria for defining units of conservation priority. Moritz proposed that the highest level of conservation priority, the evolutionarily significant unit (ESU; Ryder 1986), should be designated for

populations or groups of populations that had been historically isolated, and hence contained unique evolutionary genetic potential. Moritz (1994) proposed that to be designated ESUs, populations should display reciprocal monophyly at mitochondrial DNA regions, a measure of their historical separation. Additionally, the same populations should be significantly divergent at nuclear loci, to demonstrate that the distinct evolutionary potential developed over time was maintained today. For populations that demonstrated significant divergence at mitochondrial or nuclear loci, but did not meet the criteria for ESU designation, Moritz (1994) suggested the term management units (MUs). Management units reflect a more recent isolation of populations than ESUs, but are still recognised as important for conservation.

Stemming from Moritz's (1994) paper have been numerous others debating the definition, criteria, and application of ESUs, with many arguing that conservation priorities should not be decided based solely on genetic data (Pennock & Dimmick 1997; Bowen 1998; Waples 1998; Paetkau 1999; Taylor & Dizon 1999; Crandall *et al.* 2000; Fraser & Bernatchez 2001). While many would agree that, ideally, genetic and ecological information should be combined to identify distinct population units, often genetic data is available on a scale at which ecological data are not. As a result, Moritz's criteria have appealed to many researchers, providing a workable framework for identifying distinct units in population genetic studies.

For example, Brown Gladden *et al.* (1997; 1999) used mitochondrial DNA and microsatellite markers to examine population structure and identify conservation units in beluga whales in the North American Arctic. Beluga whales consisted of two mitochondrial lineages: an eastern North American lineage; and a western North

American lineage. Brown Gladden *et al.* (1997) proposed that the two lineages had developed during the Wisconsian glaciation, when beluga whales would have been isolated in glacial refugia in the Pacific and Atlantic, before recolonising their current distribution. Subsequent microsatellite analyses also revealed an east - west separation of beluga whales, but the separation was geographically different from that identified with mitochondrial DNA (Brown Gladden *et al.* 1999). The separation evident in the microsatellite markers corresponded with the current area of heavy ice in the Canadian Arctic, a potential barrier to gene flow between eastern and western groups of beluga whales. Two ESUs were identified for beluga whales, corresponding with the areas where the historical separation of beluga populations is maintained by current barriers to gene flow ie. where both mitochondrial DNA and microsatellites indicated an east - west separation (Brown Gladden *et al.* 1999). Eight management units were identified, based on genetic divergence at either mitochondrial or microsatellite markers.

In the example of the beluga whale and the earlier example of the greater prairie-chicken, mitochondrial DNA and microsatellites revealed different patterns, due to current processes differing from historical ones. The use of either of these markers alone would have resulted in different explanations of the processes affecting populations, and thus in different designations of genetically distinct population units. These examples highlight the importance of using both mitochondrial and nuclear markers to understand population structure, both past and present.

1.4 Aims of the thesis

This thesis aims to investigate varying scales of population structure in the roseate tern, using a combination of mitochondrial DNA and microsatellites to gain a greater understanding of the relationships among roseate tern breeding colonies, both historical and recent, and at varying geographic scales.

Specifically, the three main aims of this thesis are to:

1. Examine the genetic relationships among the currently recognised subspecies of roseate tern.

2. Determine the extent of genetic divergence among breeding colonies of the roseate tern within Australia.

3. Investigate the genetic relationships among non-breeding aggregations and breeding colonies of roseate terns within Australia and the Pacific region.

1.5 Thesis outline

In this thesis I have used two types of genetic marker, mitochondrial DNA and microsatellites, to investigate population structuring in the roseate tern on a number of spatial scales. I used mitochondrial DNA to investigate population structure within Australia and to examine the relationships between breeding colonies in both Australia and Japan and non-breeding aggregations in Australia (Chapter 2). The results of this

work led me to a global phylogeographic study using mitochondrial DNA to assess the relationships among the currently recognised subspecies of roseate tern (Chapter 3). I then used microsatellites to examine the recent relationships among the major roseate tern breeding regions of the world and, in greater detail, among breeding colonies within Australia (Chapter 4). During my work with mitochondrial DNA, I obtained evidence of a hybridisation event involving the roseate tern, and explored this using a combination of mitochondrial DNA and microsatellites (Chapter 5). In my final chapter (Chapter 6) I combined the information that each of the genetic markers have provided to explore the relationships, both historical and contemporary, among roseate tern colonies throughout the world, with a focus on those in Australia.

Chapter 2: Lack of phylogeographic structure between Australian and east-Asian subspecies of the roseate tern revealed using mitochondrial DNA markers

2.1 Introduction

In Australia, an estimated 60 000 roseate terns breed in colonies distributed along more than 6 000 km of coastline, from Lady Musgrave Island in the southern Great Barrier Reef, Queensland, around the northern coastline and as far south as Fremantle, Western Australia (Dunlop & Wooller 1986; King 1993; Figures 1.2 and 2.1). Roseate terns often breed on different islands in consecutive years, showing fidelity to a particular island group or region, rather than to a specific island (Higgins & Davies 1996). In addition, an estimated 40 000 - 60 000 non-breeding individuals occur in two regions of Australia: the Swain Reefs in the southern Great Barrier Reef, Queensland, and the Lacepede Islands in northern Western Australia. Only a few records of breeding roseate terns exist for these two locations, and none are of more than 20 breeding pairs (Limpus & Lyon 1981; Heatwole *et al.* 1996). In both locations, the non-breeding aggregations are likely to be composed of individuals from more than one breeding colony or region.

During the Austral summer (January-February), an estimated 12 000 – 25 000 non-breeding roseate terns are present on the cays of the Swain Reefs (Walker 1988; O'Neill *et al.* in press). In 2002, roseate terns banded in Japan and Taiwan, belonging to the east-Asian subspecies *S. dougallii bangsi*, were captured in Australia in the summer non-breeding aggregation in the Swain Reefs (O'Neill *et al.* in press). As a result the roseate tern was proposed for listing on the Japan-Australia Migratory Bird Agreement (JAMBA) and China-Australia Migratory Bird Agreement (CAMBA). Early estimates

based on moult data suggest up to 60% of the Swain Reefs summer aggregation could be birds of east-Asian origin (O'Neill *et al.* in press).

A smaller non-breeding aggregation, of up to 5 000 individuals (Heatwole *et al.* 1996), occurs in the Swain Reefs during the Austral winter (June-July). It was originally presumed that the Swain Reefs winter aggregation consisted of individuals that bred on cays of the Great Barrier Reef during the Austral summer (January-February), and this idea was confirmed in part with birds colour-flagged in the Swain Reefs observed breeding during the summer on Lady Elliot Island and Wilson Island in the southern Great Barrier Reef (A. Congram, B. Knuckey, pers. comm.). One individual colour-flagged in the Swain Reefs in the Austral winter (July) was seen in Okinawa in the northern hemisphere summer (July-August) of a subsequent year, indicating that some east-Asian birds remain in Australian waters rather than making the return migration each year. In the Austral spring (October-November), the Lacepede Islands host 25 000 – 30 000 non-breeding roseate terns. The entire breeding population of Western Australia is an estimated 20 000 individuals (Burbidge & Fuller 1996), hence the non-breeding aggregation must comprise individuals from multiple breeding colonies, likely from a wide geographic area.

Mitochondrial DNA has been investigated as a tool for identifying the geographic origins of individuals away from their breeding location: on non-breeding grounds (Bowen *et al.* 1996; Wenink & Baker 1996; Baker *et al.* 1999; Glenn *et al.* 2002; Kimura *et al.* 2002), and during dispersal or migration events (Bolten *et al.* 1998; Laurent *et al.* 1998; Wennerberg 2001; Wennerberg *et al.* 2002). The most successful use of mitochondrial markers for identification of birds on wintering grounds has been

in the dunlin, a migratory shorebird. Mitochondrial DNA sequencing identified five major lineages of dunlin corresponding with recognised subspecies (Wenink *et al.* 1993; Wenink *et al.* 1996), with the variation and level of divergence among the lineages suggesting that divergence occurred due to fragmentation in different Pleistocene refugia (Wenink *et al.* 1994). Subsequent work has used these mitochondrial markers, both alone and in combination with morphological characters, to identify the origins of dunlin on wintering grounds and migration routes (Wenink & Baker 1996; Wennerberg 2001). Similar studies in other species have identified migration routes (Ruegg & Smith 2002), assessed the mixing of geographically isolated breeding groups on feeding or wintering grounds (Bowen *et al.* 1996), and identified appropriate conservation units for a species or subspecies (Bolten *et al.* 1998; Laurent *et al.* 1998). The inclusion of the roseate tern on the JAMBA and CAMBA would require the governments of Australia, Japan and China to protect the species, including preservation of habitat (Department of Foreign Affairs 1995). However, the significance of these agreements to the conservation of the roseate tern, and potentially the extent of government actions in relation to the agreements, will depend on the degree of overlap of birds from the three countries.

The mitochondrial NADH dehydrogenase subunit 6 (ND6) gene was identified by Moum & Johansen (1992) as an attractive choice for phylogenetic studies because of its higher than average variation when compared to other mitochondrial genes. A comparison of the mitochondrial genomes of common and roseate terns revealed ND6 as the most variable gene (J. Feinstein, pers. comm.). The second gene used in this study, NADH dehydrogenase subunit 2 (ND2), is now increasingly used for

phylogenetic work, because of its length and ease of amplification with degenerate avian primers (Johnson & Sorenson 1998; Johnson *et al*. 2000; Zink *et al*. 2001).

In this study I use two mitochondrial genes, ND6 and ND2, to investigate population structure among Australian and Japanese breeding colonies and to determine whether mitochondrial genes can be used to identify the origins of roseate terns in non-breeding aggregations in Australia.

2.2 Methods

Blood or tissue samples were collected from a total of 110 individuals. Five breeding colonies in Australia were sampled: two in Western Australia, and three in Queensland (Figure 2.1). Samples were also collected from several small offshore rocky outcrops in Okinawa, Japan. The non-breeding aggregation in the Lacepede Islands was sampled in October, and the summer and winter non-breeding aggregations in the Swain Reefs were sampled in January and July, respectively. Samples were stored in NaCl-saturated DMSO, or in Queens Lysis Buffer (Seutin *et al*. 1991) without *n*-lauroylsarcosine. DNA was extracted using the ammonium acetate method described by Nicholls *et al*. (2000). DNA was gel-quantified in comparison with known concentration lambda standards.

The mitochondrial ND2 gene was amplified in two halves using primers L5216, H5766, L5758 and H6313 (Sorenson *et al*. 1999). After initial sequencing, primer ND2H-1 (5'-AGAAGGTTAGTAGTGTTAG-3') was designed in place of H5766. The ND6 gene was amplified using primers tRNA Pro (5'-CTTAAACCTCTATCTCCAAC-3') and Glu-2 (5'-AACAACAGTGGCTTTTCAG-3'), designed from a roseate tern sequence

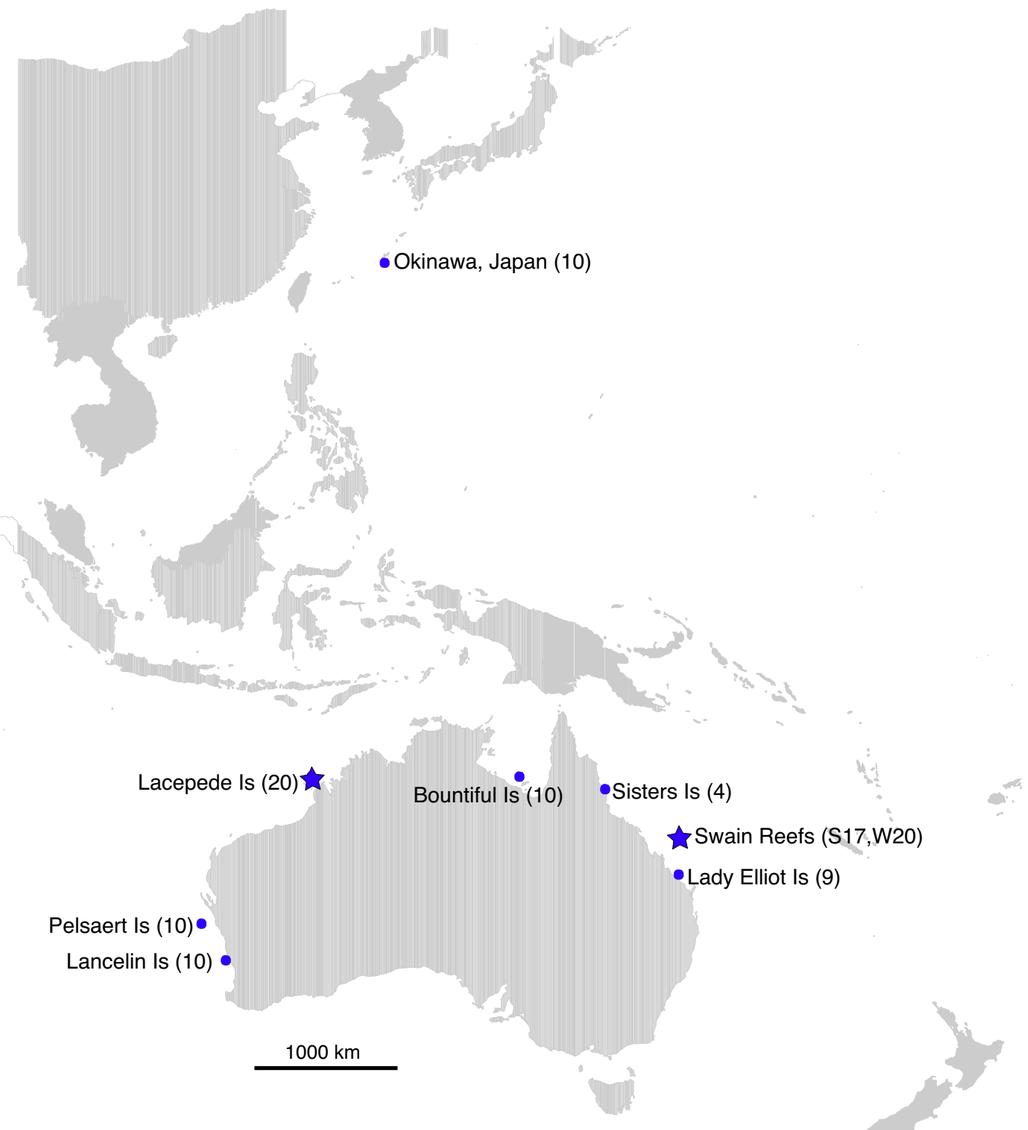


Figure 2.1: Sampling locations of roseate tern, *Sterna dougallii*, breeding colonies (circles) and non-breeding aggregations (stars). Sample numbers are in brackets. For the Swain Reefs, S = summer, W = winter.

provided by Julie Feinstein (American Museum of Natural History). The 50 μL PCR reaction mix contained 20 ng DNA, 1 x PCR buffer (containing Tris-HCl, KCl and $(NH_4)_2SO_4$), 0.16 mM of each dNTP, 2 mM $MgCl_2$, 0.48 μM of each primer and 1 unit of Taq DNA polymerase (QIAGEN). To amplify ND2, the PCR included 30 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 60 s, with a final 10 minute extension at 72°C. ND6 was amplified using the same profile, but with an annealing temperature of 55°C. PCR products were gel-purifed using a QIAquick Gel Extraction Kit according to the manufacturer's instructions (QIAGEN). Sequencing reactions were performed using Big Dye v3.1 (ABI) half reactions according to the manufacturer's instructions. Both strands of the ND2 gene, and the H strand of ND6 were sequenced at the Griffith University Molecular Analysis Facility. Sequences were edited in Sequencher v3.1.1 (Gene Codes Corporation), manually aligned using Se-Al v1.0 (Rambaut 1996), and any variable sites verified on the chromatograms. Unusual, highly divergent haplotypes were obtained for one individual, as detailed in the results. This individual was excluded from analyses.

Preliminary genetic analyses were conducted separately for each gene. A partition homogeneity test (Farris *et al.* 1995), using 100 random partitions, was performed in PAUP* (Swofford 1998) to confirm that both genes reflected similar molecular evolutionary processes. Subsequent analyses were performed with the two genes combined. Nucleotide diversity and pairwise F_{ST} (Slatkin 1995) were calculated in Arlequin v2.0 (Schneider *et al.* 1997). Tajima's D statistic was used to test the selective neutrality of each gene under the infinite site model, with significance determined using the 95% confidence intervals of a beta distribution (Tajima 1989). An analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) was conducted on two hierarchical

levels: between Australia and Japan; and between each of the states (Western Australia and Queensland) and Japan. A statistical parsimony tree, illustrating the relationships among the haplotypes, was constructed in TCS v1.13 (Clement *et al.* 2000). The population pairwise F_{ST} values, calculated in Arlequin, were used for an analysis of isolation by distance using a Mantel test in IBD v1.4 (Bohonak 2002). Geographic distances were measured as the shortest distances between breeding colonies via the coastline, as roseate terns are unlikely to fly significant distances over mainland Australia. Samples from the non-breeding aggregations were excluded from calculations of pairwise F_{ST}, AMOVA and isolation by distance due to the high likelihood, and in some cases evidence, that the aggregations contain individuals from the sampled breeding colonies.

2.3 Results

As mentioned previously, unusual haplotypes were obtained for one individual. In individual RTS4, identical amplification conditions produced two ND6 haplotypes on separate occasions: the most common haplotype and another differing by 28 base changes. RTS4 also had a divergent ND2 haplotype, differing from the common haplotype at 26 nucleotide positions. This result was confirmed by repeat amplification and sequencing. Several paths were explored to determine whether either the 'normal' or 'abnormal' haplotypes represented nuclear pseudogenes, which have been reported in a number of avian taxa (Arctander 1995; Sorenson & Fleischer 1996; Kidd & Friesen 1998; Allende *et al.* 2001; Nielsen & Arctander 2001). Neither the 'normal' or 'abnormal' haplotypes for RTS4 contained evidence of base ambiguities in the sequence chromatograms. The ratio of transitions to transversions (ts:tv) for the 'normal'

haplotypes was 5:0 at ND6, and 4:1 at ND2; inclusion of the 'abnormal' haplotype did not change the ts:tv for ND2, but for ND6 it increased to 12:1. The observed bias towards transitions is indicative of a mitochondrial gene (Arctander 1995). 'Normal' and 'abnormal' haplotypes were translated and compared with the appropriate chicken (*Gallus gallus*) genes (Desjardins & Morais 1990). No premature stop codons were present, and nucleotide substitutions matched the expected distribution for a coding gene, predominating at the third position. Relative to chicken, 67% and 65% of nucleotide substitutions in the 'normal' and 'abnormal' ND6 haplotypes, respectively, occurred at the third position. Third position substitutions made up 62% and 61% of substitutions in the 'normal' and 'abnormal' haplotypes, respectively, at ND2. Additionally, both haplotypes at both genes showed similar nucleotide composition biases to those seen in birds (Kocher *et al.* 1989; Desjardins & Morais 1990). A paucity of G is observed both at the third position (4.6 – 5.5%) and overall (10.3 – 10.6%) in ND2, with C and A most commonly occurring at the third position. ND6 is coded on the opposite (L) strand and hence shows a preference for codons ending in T or G, with C as the low frequency nucleotide (1.8% of third position nucleotides; 8.6% of total). Neither 'normal' or 'abnormal' haplotypes could be clearly identified as nuclear pseudogenes, as all conformed to expectations for mitochondrial coding genes. While it was not possible to confirm the identity of the 'abnormal' haplotypes, and investigation of this continues, I was confident that the 'normal' haplotypes in RTS4 represented true mitochondrial genes and were comparable to those obtained for the remaining samples. However, as mentioned previously, sample RTS4 was excluded from analyses as a precaution.

A total of 1029 base pairs (bp) of ND2 and 477 bp of ND6 were sequenced, revealing 15 and 13 variable sites respectively, each describing 14 haplotypes among 110 individuals. The most common ND2 haplotype was represented in 84 individuals, with 77 individuals sharing the most common ND6 haplotype; 67 had the most common haplotype at both genes. Nucleotide diversity was low at both loci: 0 - 0.001751 at ND6; and 0 - 0.002268 at ND2.

For ND6, significant pairwise F_{ST} values among breeding colonies occurred between Lady Elliot Island and both Pelsaert Island ($F_{ST} = 0.24$, $p < 0.001$) and Lancelin Island ($F_{ST} = 0.3$, $P = 0.009$). Despite its lower level of variation, ND2 revealed the same two significant pairs: Lady Elliot Island – Pelsaert Island, $F_{ST} = 0.3$, $P = 0.009$; Lady Elliot Island – Lancelin Island, $F_{ST} = 0.3$, $P = 0.045$. For ND6, Tajima's $D = -1.9$, which falls outside the 95% confidence interval (-1.78 – 2.08), and is therefore significant. For ND2 $D = -1.3$, which falls within the 95% confidence interval and is non-significant. The partition homogeneity test yielded a *P* value of 1, indicating that both genes reflected similar processes. All subsequent results are the product of analyses with the two genes combined.

Combined, 1506 base pairs of the ND6 and ND2 genes had 28 variable sites, describing 23 haplotypes. Nucleotide diversity ranged from 0 (for Sisters Island, where all four samples had the same haplotype) to 0.002398. The statistical parsimony tree showed a star pattern radiating from the most common haplotype (Figure 2.2) with only one branch longer than two base changes. Only four (of 23) haplotypes occurred more than two base changes from the common haplotype. There was a lack of notable pattern with regard to sampling location in the statistical parsimony tree, although the significant

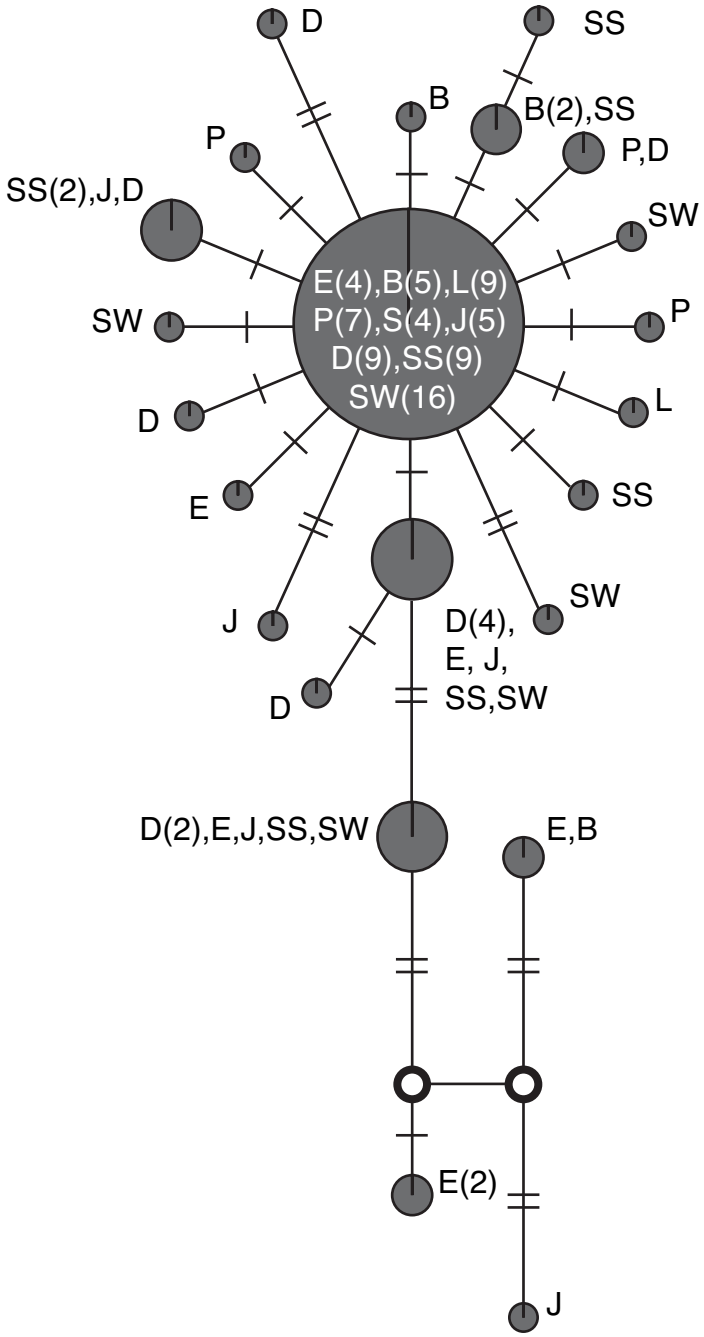


Figure 2.2: Minimum spanning tree of roseate tern, *Sterna dougallii*, ND2 and ND6 mitochondrial haplotypes. Circle sizes are proportional to the number of individuals sharing the haplotype. Letters refer to the sampling location of individuals, and the number in brackets indicates the number of individuals from each location with the haplotype, if greater than one; E = Lady Elliot, S = Sisters, B = Bountiful, P = Pelsaert, L = Lancelin, J = Japan, D = Lacepedes, SS = Swain Reefs summer, SW = Swain Reefs winter. Number of crosshatches denotes the number of nucleotide changes between haplotypes. Open circles represent haplotypes at inferred branching points.

pairwise F_{ST} values were reflected in the occurrence of four Lady Elliot Island samples on the long branch and the corresponding absence of any samples from Lancelin Island or Pelsaert Island. Particularly evident was the lack of any clustering of samples from Japan. Samples from non-breeding aggregations all occur within three base changes of the common haplotype and, as a result of the lack of structure among breeding colonies, their distribution in the statistical parsimony tree cannot be used to infer the composition of non-breeding aggregations. When an AMOVA was performed with the Australia breeding colonies grouped into two regions, Western Australia and Queensland, it revealed differentiation between the two regions, with 82.4% of the variation distributed within the regions, and 8.2% between them (F_{ST} = 0.18, P = 0.02). Japanese and Australian breeding colonies were weakly differentiated when all five Australian breeding colonies were grouped (F_{ST} = 0.04, P = 0.041), but there was no significant difference between breeding individuals in Japan and colonies grouped within Western Australia or Queensland (Japan - Western Australia F_{ST} = 0.11, P = 0.12; Japan - Queensland F_{ST} = -0.01, P = 0.24; see Appendix 1 for AMOVA tables for this section). Despite the divergence between breeding colonies in Western Australia and Queensland, there was no support for isolation by distance among the breeding colonies in Australia (r = 0.23, P = 0.2) or when the samples from Japan were included (r = 0.04, P = 0.41). Further exploratory analyses using AMOVA revealed that the three samples from Lady Elliot Island most different from the common haplotype (two samples six changes away, and one eight changes away) were responsible for all significant differences, as their exclusion resulted in non-significant results for all comparisons.

2.4 Discussion

While no evidence of genetic structure was found among the Australian and Japanese breeding colonies, weak differentiation was observed between breeding colonies in Queensland and Western Australia, and between Australia and Japan. The significant differences seen in this study stem from Lady Elliot Island, which had the highest observed level of nucleotide diversity of the breeding colonies sampled for this study. This was a result of three samples having haplotypes six or more changes from the common haplotype. By contrast, the haplotypes of individuals sampled at Lancelin Island and Pelsaert Island all occur within one base change of the common haplotype, resulting in the significant pairwise F_{ST} values observed. There are three possible explanations for the higher diversity and apparent differentiation of the Lady Elliot Island colony from the colonies in Western Australia. Firstly, increased genetic diversity is anticipated in a colony with a larger effective population size. Lancelin Island does have a smaller number of breeding pairs than Lady Elliot Island and nearby islands (< 50 cf. 100-200); however, the reverse is true for Pelsaert Island and its surrounding islands, where colonies of more than 500 breeding individuals regularly occur (King 1993; Burbidge & Fuller 1996). Secondly, early colonisation and increased time in which to accumulate mutations cannot be invoked as an explanation for the higher diversity observed at Lady Elliot Island. The formation of Lady Elliot Island began approximately 3 000 years ago (Flood *et al*. 1979), so it is unlikely to represent one of the islands first colonised by roseate terns. Additionally, even if Lady Elliot Island was one of the first islands colonised, the mutation rate of mitochondrial DNA is insufficient for higher comparative diversity to have developed in the short evolutionary

time since the formation of Lady Elliot Island. The third explanation for the higher level of diversity observed at Lady Elliot Island is that, by chance, sampling of this colony encompassed a greater proportion of the total existing haplotypic diversity than sampling at other colonies. Because of the low overall diversity seen in this study, it is not possible to differentiate between these hypotheses. Increased sample size and a more variable genetic marker should allow a higher resolution analysis.

A global phylogeographic study of the sooty tern revealed genetic divergence between sooty terns from the Indo-Pacific and the Atlantic (Avise *et al*. 2000). However, despite levels of nucleotide diversity an order of magnitude higher than those found in this study (0.018 – 0.029 cf. 0 – 0.002398), minimal structure was detected within each oceanic basin. Reviews of phylogenetic studies of seabirds (Friesen 1997; Avise *et al*. 2000) note that mitochondrial genetic structure is seldom detected within oceanic basins, and conclude that this is likely a combination of lower than expected philopatry and recent colonisation of contemporary breeding sites from glacial refugia.

This study revealed a star-shaped phylogeny for the roseate tern, which is indicative of an expanding population (Slatkin & Hudson 1991). The negative significant value of Tajima's D for ND6 supports this, as population expansion results in negative, and often significant, D values (Aris-Brosou & Excoffier 1996). The star pattern in the minimum spanning tree for the roseate tern, combined with low nucleotide diversity and negative values of Tajima's D for both genes, suggests a recent colonisation of the Australasian region by a small number of roseate terns. In contrast to the dunlin, in which subspecies are thought to have diverged due to isolation in separate Pleistocene refugia (Wenink *et al*. 1994), roseate terns in Australasia appear to have colonised the region from a

common glacial refuge. The current locations of breeding colonies in Australia and Japan were landlocked during the last glacial maximum, and would have become suitable breeding sites within the last 7 000 years, and in some instances much more recently (Pirazzoli 1991). For example, the formation of Lady Elliot Island is estimated to have started only 3 000 years ago (Flood *et al.* 1979). Colonisation of these sites by a small number of individuals, and little evolutionary time in which to accumulate mutations, has resulted in low mitochondrial DNA diversity and hence a lack of detectable genetic structure between the Australian and east-Asian subspecies of roseate tern. Recent colonisation of islands south of the Houtman Abrolhos in Western Australia (Dunlop & Wooller 1986) provide evidence that, at least within Australia, roseate terns are still undergoing range expansion.

Postglacial range expansion in North America has been given as the most likely explanation for similar combinations of low diversity and star phylogenies found in Swainson's thrush, *Catharus ustulatus,* (Ruegg & Smith 2002), MacGillivray's warbler, *Oporornis tolmiei*, (Mila *et al.* 2000), greenfinch, *Carduelis chloris,* (Merilä *et al.* 1997) and mountain sheep (Ramey II 1995). As in this study, Ramey II's (1995) study of mountain sheep in south-western North America found a lack of mitochondrial divergence among currently recognised subspecies. With 94% of genetic variation found within, rather than between, the five subspecies it was suggested that four of the mountain sheep subspecies be combined.

Mitochondrial haplotypes do not support the separate subspecific status of Australian, *Sterna dougalli gracilis*, and east Asian, *S. d. bangsi*, roseate terns. The short time since colonisation of their current breeding range, reflected in the low diversity of the

mitochondrial regions used in this study, is insufficient for mutations in mitochondrial genes to reflect limited gene flow between the two subspecies. Therefore, despite the lack of divergence, I cannot conclude that gene flow is unrestricted today. A global genetic assessment of roseate tern subspecies' is needed to determine large scale phylogenetic relationships in the species, and to incorporate multiple oceanic basins, at which scale divergence is usually detected for seabirds (Avise *et al*. 2000, and references therein).

It is clear that mitochondrial genes are inappropriate for differentiating between roseate terns from east-Asia and Australia within non-breeding aggregations. The shared history of individuals from the two subspecies means that a higher resolution marker is needed to reflect recent processes. Haig *et al*. (1997) explored the use of random amplified polymorphic DNA (RAPDs) for identification of several species of migratory shorebird. RAPD bands could be used to detect east vs. west origins, and to determine whether migrant groups came from a similar location. Additionally, population specific markers were identified for three of the nine species assayed. Kimura *et al*. (2002) used mitochondrial control region sequences to examine phylogenetic relationships among breeding populations of migratory Wilson's warbler in North America. Eastern and western lineages could be identified, but a more variable marker was required for identifying the origin of wintering migrants over a finer geographic scale. A combination of microsatellites and isotope analysis was subsequently used in an attempt to identify finer scale structure in Wilson's warbler (Clegg *et al*. 2003). Microsatellites identified the same east-west separation as mitochondrial DNA, and the isotope analyses revealed a north-south division, in combination proving more informative than mitochondrial DNA alone.

The colonisation of the current Australasian breeding range, from a common glacial refuge, is too recent for mitochondrial DNA to provide evidence of population structure in the roseate tern. As a result, mitochondrial DNA markers cannot be used to assess the composition of non-breeding aggregations. More variable genetic markers such as microsatellites and other fingerprinting methods are worth exploring for examining genetic structure and identifying the origins of non-breeding roseate terns.

Chapter 3: Global phylogeography of the roseate tern: are the current subspecies supported?

3.1 Introduction

There are four or five currently recognised subspecies of roseate tern, *Sterna dougallii*, based on differences in wing length, bill length and bill colouration (Gochfeld 1983; Cramp 1985; Higgins & Davies 1996). Peters (1934) and Gochfeld (1983) list the five subspecies as: *Sterna dougallii dougallii*, which breeds in North America, the Caribbean, Ireland, Europe, the Azores and South Africa; *S. d. arideensis*, which breeds in the Seychelles, and possibly Madagascar; *S. d. korustes*, which breeds in India, Sri Lanka and as far east as the western Malay Peninsula; *S. d. bangsi*, which breeds in Asia east of the Malay Peninsula, and on some of the Pacific islands; and *S. d. gracilis*, endemic to Australia (Figure 1.2). However, neither Higgins & Davies (1996) or Cramp (1985) recognise *S. d. arideensis*, and both consider birds breeding in New Caledonia part of the Australian subspecies *S. d. gracilis* rather than *S. d. bangsi*. Cramp (1985) observes that in the nominate subspecies *S. d. dougallii*, variation of morphological characters is nearly as large as in the whole species. Both Cramp (1985) and Higgins & Davies (1996) comment on the similarities between some of the current subspecies, the resulting difficulty in identifying individuals to subspecies level, and the need for a detailed assessment of the current subspecies distributions.

As with species concepts, there is much discussion over the definition and use of taxonomic units below the species level (eg. Phillips 1982; Zusi 1982; Ryder 1986; Avise & Ball 1990; O'Brien & Mayr 1991). A common theme, however, is that

subspecies should be largely distinguishable from one another in some way (morphologically or genetically), while still maintaining the ability to interbreed (otherwise they would likey be classified as species). For example, O'Brien & Mayr (1991) propose that members of a subspecies share a unique range, natural history and phylogenetically concordant characters in comparison with other subspecific groups, but are still reproductively compatible with other groups. From a phylogenetic perspective, such definitions align with the concept of an evolutionarily significant unit (ESUs; Ryder 1986; Moritz 1994), which is considered to be a group of phylogenetically distinguishable individuals (this does not preclude reproductive incompatibility with other ESUs), and is a definition commonly used when identifying and classifying genetically distinct populations.

The first genetic study to include individuals from more than one subspecies of roseate tern failed to find differences between the Australian subspecies, *S. d. gracilis*, and the east-Asian subspecies, *S. d. bangsi* using two mitochondrial genes (Chapter 2). I concluded that there had been insufficient time for mitochondrial DNA to show evidence of divergence between the Australian and east-Asian roseate terns, which are likely to have colonised their current breeding locations in Australasia from a common glacial refuge within the last 7 000 years.

Results from genetic studies of other marine species also imply that the recognition of four or five roseate tern subspecies across the species' distribution may be inappropriate. While divergence of mitochondrial DNA between oceanic basins is often observed, strong differentiation within an oceanic basin is less common. For example, Avise *et al.* (2000) found divergence between sooty tern breeding colonies in the

Atlantic and Indo-Pacific, but a lack of differentiation within either oceanic basin. Similarly, clear Atlantic – Indo-Pacific separations were seen in studies of the common murre, *Uria aalge* (Friesen *et al.* 1996) and thick-billed murre, *U. lomvia* (Birt-Friesen *et al.* 1992). Within the Atlantic, however, the thick-billed murre showed no evidence of genetic structure, and the common murre showed clinal structuring among colonies. Although not universal, the same pattern of minimal intra-ocean divergence is seen in a diverse range of marine organisms, including turtles (Bowen *et al.* 1992), fish (Alvarado Bremer *et al.* 1998; Graves 1998; Ovenden & Street 2003) and echinoderms (Lessios *et al.* 1998). The green turtle has a similar distribution to the roseate tern, and also demonstrates philopatry, a high dispersal capability and migration away from breeding areas outside the breeding season. Restricted fragment length polymorphism (RFLP) analysis of the green turtle revealed a phylogenetic split differentiating all Indo-Pacific rookeries from those in the Atlantic and Mediterranean by a minimum of five restriction site differences (Bowen *et al.* 1992). Genetic structuring within both oceanic basins was evident, but less pronounced, with differences mainly attributable to haplotype frequencies rather than fixed differences. Bowen *et al.* (1992) interpreted the shallow intra-ocean genetic structure as evidence for more recent gene flow between rookeries, possibly due to cycles of colony extinction and recolonisation driven by climatic or environmental factors. Considering the results of previous studies of marine organisms, many of which share distributions and/or life history traits with the roseate tern, significant differentiation between roseate tern colonies in the Atlantic and Indo-Pacific is expected. However, significant genetic differentiation among the colonies of the three or four roseate tern subspecies occurring within the Indo-Pacific is less likely and, if so, the current number of subspecies may be unwarranted.

This study uses two mitochondrial genes to examine the phylogenetic relationships within a global sample of roseate tern colonies. These relationships will be examined in light of the current four or five recognised subspecies, to assess whether genetic relationships are congruent with the currently defined subspecies.

3.2 Methods

Blood, solid tissue or feather samples for this study were collected from breeding colonies in each of five countries: the Seychelles (*Sterna dougallii arideensis*), and South Africa, the United States, the Republic of Ireland, and the Azores, Portugal (*S. d. dougallii*), in addition to those collected previously from Australia (*S. d. gracilis*) and Japan (*S. d. bangsi*) (see section 2.2 and Table 3.1). All samples were stored in salt-saturated DMSO or Queens Lysis Buffer (Seutin *et al.* 1991) without *n*-lauroylsarcosine.

DNA was extracted from blood samples as described previously (see section 2.2). DNA extractions from the additional feather samples were carried out using a protocol similar to those described by Kocher *et al.* (1989) and Taberlet & Bouvet (1991). The base of the feather was digested overnight in 250 μL extraction buffer (10 mM Tris, 20 mM EDTA, 200 mM NaCl, 2% SDS, 10 mg/mL DTT and 50 μg/mL Proteinase K) at 37°C with agitation. DNA was extracted using a standard phenol-chloroform extraction (Sambrook *et al.* 1989) followed by precipitation with 2.5 volumes of 100% ethanol and 0.1 volumes of 2.5 M sodium acetate. DNA concentration was gel-quantified in comparison with known concentration lambda standards. The mitochondrial ND2 and

Table 3.1: Roseate tern sample locations, numbers of samples used in this study, sample type (B = blood, F = feather, S = solid tissue), and subspecies designations. Sample groups from Japan, Queensland and Western Australia were created from previous sampling and data (see section 2.2 for specific sampling location details of Australian colonies)

Sampling Location	Co-ordinates	Current Subspecies	Tissue Type	N
Rockabill Is, Republic of Ireland	53° 37′ N, 6° W	*S. d. dougallii*	B	6
Terceira Is, Azores	38° 41′ N, 27° 3′ W	"	B	6
Bird Is, Massachusetts, United States	41° 40′ N, 70° 43′ W	"	B	6
Bird Is, South Africa	33° 50′ S, 25° 50′ E	"	F	6
Aride Is, Seychelles	4° 10′ S, 55° 40′ E	*S. d. arideensis*	F, S	6
Okinawa, Japan	26° N, 127° E	*S. d. bangsi*	B	6
Queensland, Australia	N/A	*S. d. gracilis*	B, S	8
Western Australia	N/A	"	B	6

ND6 genes were amplified and sequenced in six individuals from each geographic region sampled (described by country of origin), as described in Chapter 2 but both strands of each gene were sequenced. Sequences were edited in Sequencher v3.1.1 (Gene Codes Corporation), manually aligned using Se-Al (Rambaut 1996), and any variable sites verified on the chromatograms. Additional ND2 and ND6 sequence data from a previous study of roseate terns in Japan (*S. d. bangsi*) and Australia (*S. d. gracilis*) were included (Chapter 2), bringing the number of sampled subspecies to four of the five currently recognised. Individuals from Japan (N = 6) and Australia (N = 14) were specifically chosen to encompass the range of haplotypes observed in breeding colonies of the two countries in the previous study. To avoid large differences in sample

sizes, Australia was divided into two regions: Queensland (N = 8) and Western Australia (N = 6).

Tests for partition homogeneity and selective neutrality of each gene were performed, and nucleotide diversity was calculated, as described in section 2.2. Genetic distance, using the method of Tamura (1992) which takes uneven base frequencies and transition: transversion ratio into account, was calculated in Arlequin, and used for an analysis of isolation by distance in IBD v1.4 (Bohonak 2002). Geographic distance was measured as the shortest distance between sampling locations without crossing major land masses. For the pooled Queensland and Western Australia groups, distances were measured from Lady Elliot Island and Lancelin Island, respectively. Phylogenetic trees were constructed in PAUP* using maximum likelihood and maximum parsimony methods, with two bridled terns (*S. anaethetus*) and a black-naped tern (*S. sumatrana*) from Australia as the outgroups. The black-naped tern and roseate tern are often considered to be a species pair, while the bridled tern is more distantly related (Moynihan 1959; Schnell 1970a; Schnell 1970b). Robustness of phylogenetic trees was tested using 500 bootstrap replicates (Felsenstein 1985). The best base substitution model and parameters for the maximum likelihood analysis were determined using MODELTEST v3.06 (Posada & Crandall 1998). Uncorrected percentage sequence divergence between phylogenetically distinct groups was calculated using the average number of nucleotide site differences between individuals in the compared groups according to Tamura's (1992) method, and corrected using the average number of differences within each group. Once phylogenetically distinct groups were identified, analyses of nucleotide diversity, Tajima's D and isolation by distance were repeated for each group.

3.3 Results

The partition homogeneity test confirmed that the ND2 and ND6 genes contained similar phylogenetic signal (*P* value = 0.58), so all analyses were performed with the two genes combined, unless specified otherwise. A total of 1381 base pairs (bp) of the ND2 (1029 bp) and ND6 (352 bp) genes were sequenced in 50 individuals, revealing 70 variable sites which described 23 haplotypes (Tables 3.2 and 3.3; see Appendix 2 for full haplotype alignment). The four most common haplotypes were present in 30 individuals. There were eighteen unique haplotypes, of which only three differed from one of the common haplotypes by more than one nucleotide substitution. Nucleotide diversity ranged from 0, at Bird Island, Massachusetts, where all samples had the same haplotype, to 0.2065 at Bird Island, South Africa. Tajima's D value was positive for both genes: for ND2, D was outside the 95% confidence interval (D = 2.97, C. I. = -1.8 – 2.04), but for ND6 was within (D = 1.44). With all the sampled colonies included, there was a strong correlation between genetic and geographical distance (r = 0.88, P = 0.004).

Phylogenetic analyses

Including the outgroup individuals, 172 sites were variable, with 58 parsimony-informative. The most appropriate model for the maximum likelihood analysis was HKY85 with a gamma correction (Hasegawa *et al.* 1985). The overall transition:transversion ratio was 11.52 and base frequencies were uneven: 34.4% A, 32.5% C, 9.9% G, 23.1% T. The maximum likelihood and maximum parsimony analyses revealed the same tree topology, although the maximum likelihood analysis collapsed two minor branches in comparison to the maximum parsimony analysis; a

Table 3.2: Variable sites in the mitochondrial DNA sequence obtained from 1029bp of the ND2 gene (variable sites 45 to 1029) and 352bp of the ND6 gene (variable sites 1038 to 1362) of the roseate tern, *Sterna dougallii*. Sequence identity to haplotype 1 is indicated by a '.'.

	Alignment of variable sites																																			
	4	5	6	8	10	11	12	13	14	18	19	21	21	22	26	30	30	35	36	37	40	41	43	44	46	46	48	48	55	56	57	58	63	63	70
Haplotype	5	7	9	4	5	7	9	2	1	9	6	3	5	9	1	7	9	7	8	3	2	8	0	8	2	5	3	9	2	4	0	1	1	6	5
1	T	G	G	C	T	C	G	G	C	C	T	C	T	C	C	T	G	A	T	C	C	T	C	C	G	T	G	G	A	T	C	A	G	T	T
2	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
3	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
5	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
6	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
7	.	.	.	.	C	A	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
8	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
9	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
11	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
14	.	.	.	.	C	A	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
15	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.
16	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
17	.	.	.	.	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
18	C	.	A	T	.	.	A	A	.	T	C	T	.	T	T	C	.	.	C	T	T	G	.	.	A	C	.	A	C	A	T	G	A	A	C
19	C	.	A	T	.	.	A	A	.	T	C	T	.	T	T	C	.	.	C	T	T	G	.	.	A	C	.	A	C	A	T	G	A	A	C
20	C	.	A	T	.	.	A	A	.	T	C	T	.	T	T	C	.	G	C	T	T	G	.	.	A	C	.	A	C	A	T	G	A	A	C
21	C	.	A	T	.	.	A	A	.	T	C	T	.	T	T	C	T	G	C	T	T	G	.	.	A	C	.	A	C	A	T	G	A	A	C
22	C	.	A	T	.	.	A	A	.	T	C	T	.	T	T	C	.	.	C	T	T	G	.	.	A	C	.	A	C	A	T	G	A	A	C
23	C	.	A	T	.	.	A	A	.	T	C	T	.	T	T	C	.	.	C	T	T	G	.	.	A	C	.	A	C	A	T	G	A	A	C

Table 3.2 continued: Variable sites in the mitochondrial DNA sequence obtained from 1029bp of the ND2 gene (variable sites 45 to 1029) and 352bp of the ND6 gene (variable sites 1038 to 1362) of the roseate tern, *Sterna dougallii*. Sequence identity to haplotype 1 is indicated by a '.'.

	Alignment of variable sites																																		
	7	7	7	7	7	8	8	8	8	8	8	9	9	9	9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	0	1	2	3	8	0	2	4	4	4	5	0	5	5	8	0	0	0	0	1	1	1	1	1	1	1	1	2	2	2	2	2	3	3	6
Haplotype	8	8	6	2	8	4	3	4	6	9	8	0	5	7	4	2	3	6	9	0	1	1	2	3	3	7	7	0	2	3	4	6	4	7	2
																9	8	2	2	5	8	1	6	5	1	2	8	9	0	1	1	3	4	3	7	2
1	A	G	G	G	T	T	A	C	A	T	A	G	A	C	A	T	C	A	A	C	G	A	C	T	T	A	G	G	A	C	T	A	A	T	T
2	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.
3	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.
5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.
7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
9	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	.
13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
18	G	A	A	C	C	.	G	T	G	C	G	A	G	T	.	C	.	G	G	.	A	G	T	A	C	.	.	A	G	T	.	.	G	.	C
19	G	A	A	C	C	.	G	T	G	C	G	A	G	T	.	C	.	G	G	.	A	G	T	A	C	.	.	A	G	T	C	.	G	.	C
20	G	A	A	C	C	.	G	T	G	C	G	A	G	T	.	C	.	G	G	.	A	G	T	A	C	.	.	A	G	T	.	.	G	.	C
21	G	A	A	C	Y	.	G	T	G	C	G	A	G	T	.	C	.	G	G	.	A	G	T	A	C	.	.	A	G	T	.	.	G	.	C
22	G	A	A	C	C	.	G	T	G	C	G	A	G	T	.	C	.	G	G	.	A	G	T	A	C	.	.	A	G	T	.	G	G	.	C
23	G	A	A	C	C	.	G	T	G	C	G	A	G	T	.	C	T	G	G	.	A	G	T	A	C	..	.	A	G	T	.	.	G	.	C

Table 3.3: Frequency of roseate tern, *Sterna dougallii*, mitochondrial DNA haplotypes, nucleotide diversity (π) and Tajima's D, based on the ND2 and ND6 genes, according to sampling region. Haplotype numbers correspond to those in Table 3.2. QLD = Queensland, WA = Western Australia, Seych = Seychelles, S. Africa = South Africa, Ire = Ireland, U. S. = United States

Haplotype	QLD	W A	Japan	Seych	S. Africa	Azores	Ire	U. S.	Total
1	1	-	1	-	-	-	-	-	2
2	1	-	-	-	-	-	-	-	1
3	1	-	-	-	-	-	-	-	1
4	-	-	1	-	-	-	-	-	1
5	1	-	-	-	-	-	-	-	1
6	1	-	-	-	-	-	-	-	1
7	1	-	-	-	-	-	-	-	1
8	1	3	1	1	-	-	-	-	6
9	-	1	-	-	-	-	-	-	1
10	-	1	-	-	-	-	-	-	1
11	-	1	-	-	-	-	-	-	1
12	-	-	1	-	-	-	-	-	1
13	-	-	1	-	-	-	-	-	1
14	-	-	-	1	-	-	-	-	1
15	-	-	-	1	-	-	-	-	1
16	-	-	-	-	1	-	-	-	1
17	1	-	1	3	2	-	-	-	7
18	-	-	-	-	-	5	4	-	9
19	-	-	-	-	-	1	-	-	1
20	-	-	-	-	2	-	-	6	8
21	-	-	-	-	1	-	-	-	1
22	-	-	-	-	-	-	1	-	1
23	-	-	-	-	-	-	1	-	1
Total	8	6	6	6	6	6	6	6	50
π	0.02	0.007	0.02	0.005	0.21	0.002	0.005	0	
Tajima's D	0.04	-1.23	-1.3	-1.23	2.21*	-0.93	-1.13	0	

*significant at $p < 0.05$

a single tree is presented with bootstrap values from both analyses included (Figure 3.1). Phylogenetic analyses revealed two strongly supported clades within the roseate tern, separated by 49 fixed nucleotide differences; 43 transitions and six transversions. Adjusting for unequal base frequencies and the transition:transversion ratio, this translates to 4% uncorrected inter-oceanic sequence divergence or 3.9% corrected sequence divergence. By comparison, the outgroup species, black-naped tern and bridled tern are, respectively, 4.4% and 13.1% divergent from the roseate tern. One clade contains six haplotypes, all from Atlantic Ocean breeding colonies; and the other 17 haplotypes, all from Indian and Pacific Ocean colonies (Figure 3.2). There is minimal structure within either clade. South Africa, located between the two oceanic basins, has three individuals in each clade. Within the Atlantic clade, the samples from North America and the three from South Africa differ from the samples from Ireland and the Azores by one fixed nucleotide transition (G → A). There were no fixed differences between breeding colonies in the Indo-Pacific clade.

The separation of the two clades was evident in the results of within-clade analyses, which often revealed different patterns than those from the initial overall analyses (Table 3.4). When homogeneity tests were repeated for the Atlantic and Indo-Pacific clades separately, all values of Tajima's D were negative and within 95% confidence intervals. Isolation by distance analyses of Atlantic and Indo-Pacific colonies, performed twice for each oceanic basin with South Africa excluded and included, produced contrasting results for the two oceanic basins. While there was no support for isolation by distance in the Indo-Pacific ($P = 0.22 – 0.498$), strong evidence for isolation by distance among Atlantic colonies was maintained ($P = 0.001$).

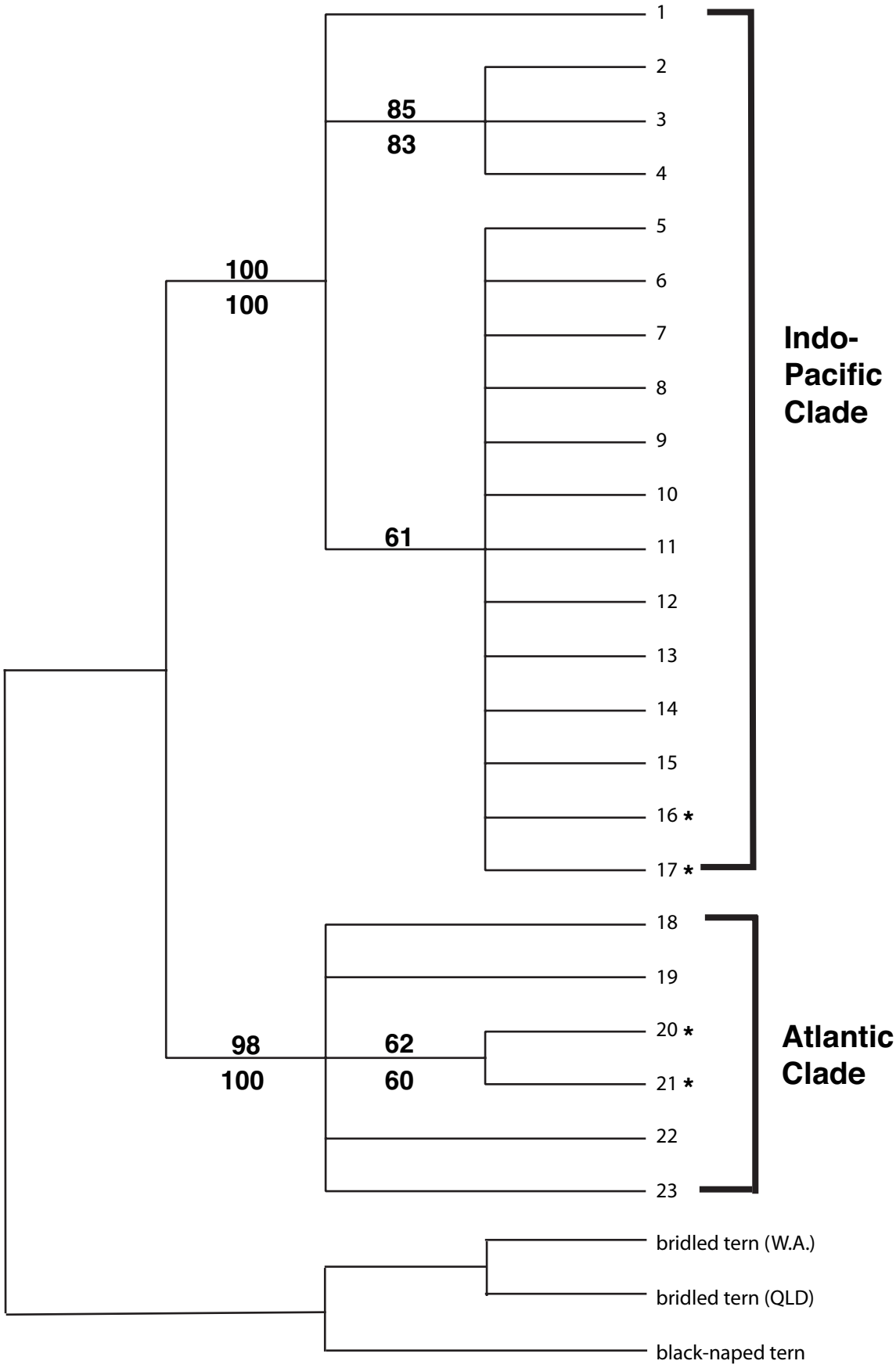


Figure 3.1: Maximum likelihood cladogram of mitochondrial haplotypes based on the ND2 and ND6 genes of the roseate tern, *Sterna dougallii*. Values above branches are bootstrap values from the maximum likelihood analysis, and those below branches from maximum parsimony analysis. Numbers at terminal branches correspond to haplotype numbers in Table 3.2. Stars indicate haplotypes observed in individuals from South Africa.

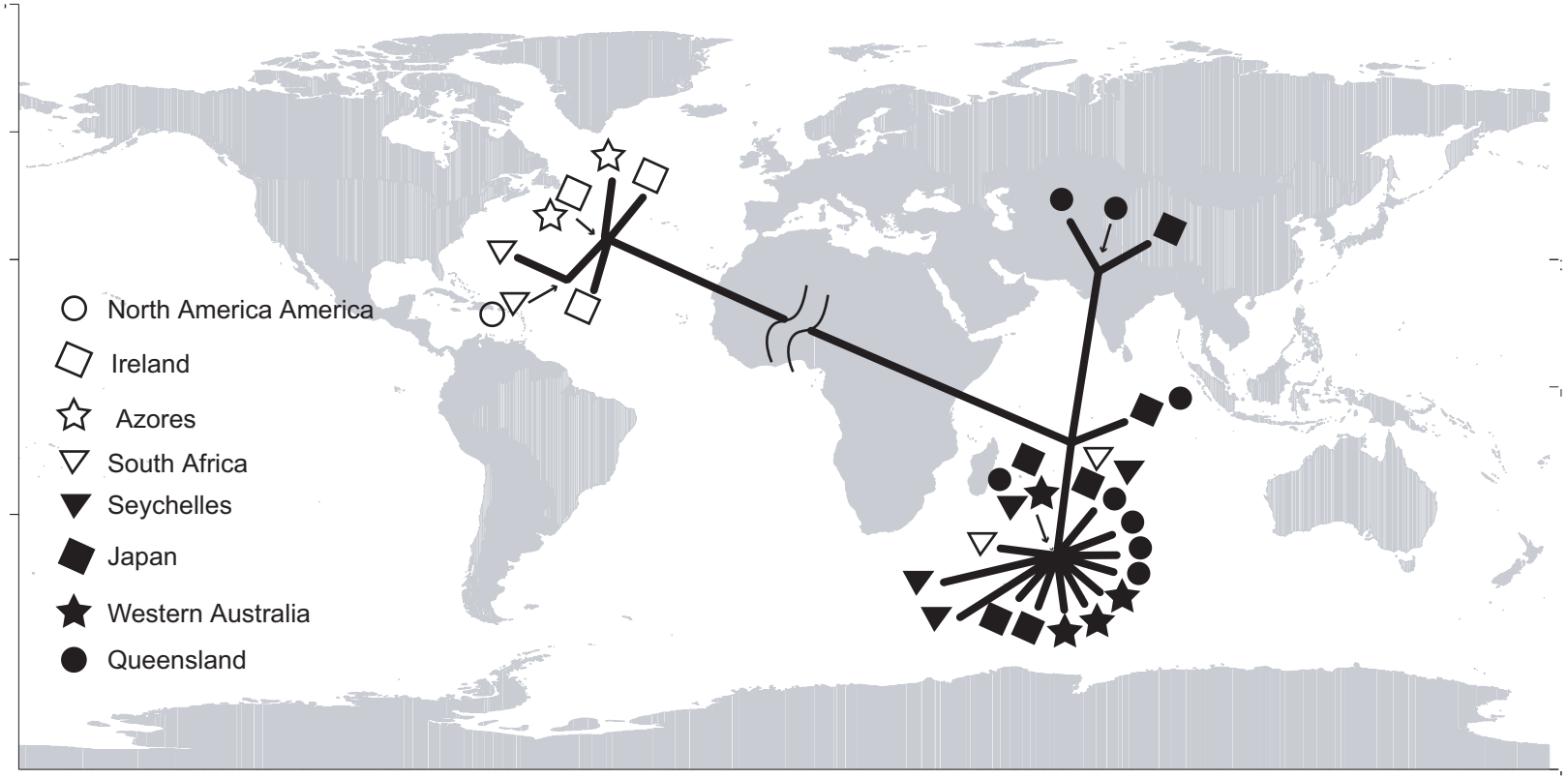


Figure 3.2: The unrooted phylogram of haplotypes based on maximum likelihood analysis of the ND2 and ND6 mitochondrial genes of the roseate tern, *Sterna dougallii*, illustrates the concurrent geographical and genetic separation of the two clades. Each node represents a haplotype, with the symbols at each node corresponding with the sampling location of individuals with that haplotype. The major branch of the phylogram is abbreviated; all other branches are proportional. Arrows indicate the internal nodes with which haplotypes are associated.

Table 3.4: Numbers of samples, haplotypes, transition: transversion ratio (Ts:Tv), nucleotide diversity (π), Tajima's D and support for isolation by distance (r^) for the two identified clades of roseate tern (including/excluding South Africa), and for all samples combined.

Clade	N	No. haplotypes	Ts:Tv	π	Tajima's D ND2	Tajima's D ND6	r^
Atlantic	21	6	4	0.00065	-0.59	-0.96	0.96*/0.85
Indo-Pacific	29	17	8	0.0017	-1.19	-1.04	0.33/0.048
Overall	50	23	6.9	0.02	2.97*	1.44	0.88*

* indicates $P < 0.05$

^ r is the correlation coefficient for genetic versus geographic distance

3.4 Discussion

This study has identified two lineages of roseate tern: one represented by the Atlantic clade, and one by the Indo-Pacific clade, with the high inter-oceanic divergence (4%) contrasting with the lack of phylogeographic structure within either oceanic basin. Based on their data from direct sequencing of the mitochondrial cytochrome *b* gene, along with Sibley & Alquist's (1990) DNA-DNA hybridisation data and the estimated divergence date of *Sterna* (terns) and *Larus* (gulls), Crochet *et al.* (2000) estimated a divergence rate of 0.47% per million years for transversions only. The divergence between the two roseate tern lineages for transversions only is 0.43%, which indicates, if the divergence rate calculated by Crochet *et al.* (2000) is relatively constant within *Sterna* and between genes, that the two lineages of roseate tern have been separated for almost one million years. Despite this being a rough estimate, it indicates that the

lineage separation in the roseate tern has been maintained through several glacial maxima. By comparison, Atlantic and Indo-Pacific lineages of sooty tern are estimated to have diverged less than 200 000 years ago (Avise *et al*. 2000). This difference may be due, in part, to the rough estimation of divergence rates, and different methods used in each instance. However, the sooty tern has a broader Pacific Ocean distribution than the roseate tern, breeding throughout the Pacific (Higgins & Davies 1996), whereas the roseate tern is confined to the west Pacific. Avise *et al*. (2000) postulate that the rise of the Isthmus of Panama has been the primary reason for the Atlantic – Indo-Pacific divergence in the sooty tern, with the occasional crossing of this barrier resulting in relatively shallow divergence. The primary barrier to gene flow in the roseate tern (and probably also the sooty tern) is likely to be continental Africa, as the species' distribution does not include the west coast of South America. It may be, therefore, that Africa has historically been a more impenetrable barrier for seabirds than South America and the Isthmus of Panama, with movements around Cape Horn less frequent than those across the Isthmus of Panama.

In contrast to the overall pattern, there is no evidence for restricted gene flow among roseate tern colonies within each oceanic basin. This could be due to 1) gene flow among colonies/regions since the separation of the two lineages, or 2) periods where gene flow is absent or limited, but subsequently masked by periodic gene flow, as may have occurred in glacial periods. The low nucleotide diversity and close relationship of haplotypes to one another within each lineage resemble the pattern seen in the previous chapter (see section 2.3), and similarly suggest either: that the recent colonisation of the current breeding locations renders mitochondrial DNA inappropriate for resolving any divergence; or the occurrence of recent gene flow between individuals from different

breeding colonies. If the latter is true, the most recent occurrence of gene flow is likely to have been during the last glaciation, approximately 20 000 years ago (Yokoyama *et al*. 2000), when current breeding locations would have been unsuitable, and range contraction may have confined roseate terns to a common refuge in each oceanic basin.

The Atlantic clade

Within the Atlantic clade, a single fixed nucleotide difference between east and west Atlantic breeding individuals implies that gene flow across the Atlantic is restricted. This fixed difference is the likely cause of the isolation by distance effect observed in the Atlantic clade, with the two closest colonies (Azores and Ireland) sharing the fixed difference that differentiates them from the more distant North American and South African colonies. However, this apparently fixed difference may be a result of low sample sizes, as trans-Atlantic movements of breeding roseate terns have been recorded in both directions (Hays *et al*. 2002a; Hays *et al*. 2002b). Additionally, birds from colonies in the Azores mix with west-Atlantic breeders on the South American coast (Hays *et al*. 2002a), a scenario which increases the chances of a return migration to a non-natal breeding colony. Alternatively, the colonies in Ireland and the Azores may exchange individuals with one another more often than with more distant Atlantic colonies, and the fixed nucleotide difference may be an accurate reflection of an isolation by distance effect in the Atlantic.

A negative, but non-significant, value of D was obtained for the Atlantic clade, suggestive of a population bottleneck and subsequent range expansion (Rand 1996). The low number of observed haplotypes, and their close relationship to one another suggest that the bottleneck and range expansion were relatively recent occurrences,

consistent with the idea that roseate terns occupied a common Atlantic refuge during the last glacial period.

The Indo-Pacific clade

As observed in the Atlantic clade and the previous study of Australian and Japanese roseate terns (Chapter 2), negative values of Tajima's D and close relationships among the observed haplotypes in the Indo-Pacific clade suggest a recent range expansion to current breeding locations from a common source. Even if individuals from different colonies are not currently interbreeding, the time since their range expansion and subsequent geographic separation is too recent for divergence to be detected using mitochondrial DNA. Therefore, despite the lack of divergence of the Indo-Pacific colonies, it cannot be concluded that gene flow is high today, or has been in the recent past. Unlike roseate tern colonies in the Atlantic, those in the Indo-Pacific have been the subject of few studies, and there is a lack of band recovery data to provide information on inter-colony movements. Particularly within the Indo-Pacific lineage there is a need for genetic studies using a higher resolution marker to elucidate contemporary processes among roseate tern colonies.

Secondary contact

Recent range expansion to higher latitudes has resulted in secondary contact, and likely introgression, of the two roseate tern lineages in South Africa. Secondary contact can cause positive values of Tajima's D (Rand 1996), as seen for both genes in this study when the two lineages were analysed in combination. Given that values of Tajima's D

were negative in separate analyses of the two lineages, it appears that in the roseate tern, the positive overall D values for both genes are a result of secondary contact between the two lineages, rather than selective processes. A similar geographic pattern of lineage overlap is seen in bigeye tuna, *Thunnus obesus*, which also have an Atlantic – Indo-Pacific separation (Chow *et al.* 2000). Individuals from each bigeye tuna lineage mix in South African waters, however despite the geographic similarity, the range overlap does not translate into gene flow because bigeye tuna breed at lower latitudes. Another marine species, *Lates calcarifer*, an estuarine teleost, shows clinal introgression around the north-east of Australia, resulting from secondary contact of divergent lineages since the last reopening of the Torres Strait (Chenoweth & Hughes 2003). Clinal introgression is less likely to occur in roseate terns because of the discontinuous distribution of breeding colonies. Increased sampling of the breeding colonies closest to South Africa would reveal the geographic extent of the secondary contact between the two roseate tern lineages.

Subspecies recommendation

The nominate subspecies, *Sterna dougallii dougallii,* is well supported by this study, with all breeding colonies sampled occurring solely within the Atlantic lineage, with the exception of South Africa. By contrast, the three or four subspecies currently recognised in the Indo-Pacific are unwarranted on a genetic basis, with no mitochondrial differentiation evident among *S. d. arideensis*, *S. d. bangsi* and *S. d. gracilis*. The subspecies not sampled in this study, *S. d. korustes*, has been described as the most morphologically distinct of the roseate tern subspecies (Cramp 1985; Higgins & Davies 1996). However, the distribution of *S. d. korustes* falls within the Indo-Pacific sampling range of this study, and it would be unexpected if it did not also fall within the Indo-

Pacific clade. I therefore highlight the need for a taxonomic revision, with the nominate subspecies, *Sterna dougallii dougallii,* maintained, representing the Atlantic lineage, and the remaining subspecies merged to create a single Indo-Pacific subspecies. The South African breeding colonies, given the equal representation of individuals from each lineage, cannot be assigned to either subspecies, but rather exist as a zone of secondary contact between the two, as discussed previously.

The Indo-Pacific lineage consists of at least 80 000 – 100 000 breeding pairs (Gochfeld 1983; O'Neill *et al.* in press), the majority of which spend some part of each year in Australia, which hosts an estimated 30 000 breeding pairs and up to 15 000 migrants from east-Asia (O'Neill *et al.* in press). By comparison, the Atlantic lineage, which corresponds with the currently recognised *S. d. dougallii*, is far less numerous, consisting of an estimated 10 000 breeding pairs, with the greatest breeding concentration occurring in the west Atlantic colonies of North America and the Caribbean (Monteiro *et al.* 1997; Gochfeld *et al.* 1998; Nisbet & Spendelow 1999; Newton & Crowe 2000). The two lineages show reciprocal monophyly for mitochondrial DNA, meeting the mitochondrial criterion for consideration as evolutionarily significant units (Moritz 1994), which recognises the evolutionary potential resulting from historical separation. However, studies of nuclear DNA markers in the two lineages are needed to confirm their status as ESUs. As mentioned previously, higher resolution nuclear markers, such as microsatellites, would provide information on contemporary processes within the roseate tern, elucidating the interactions among breeding colonies since the recent range expansion of each subspecies. If nuclear markers reveal significant recent gene flow between the two subspecies, they will meet the criteria for management units (MUs; Moritz 1994), a

designation which reflects current relationships among groups. Alternatively, if nuclear markers reveal a lack of gene flow among colonies since range expansion, recognition of MUs within each lineage may be warranted.

Chapter 4: Microsatellites reveal contemporary relationships among roseate tern populations

4.1 Introduction

Previous mitochondrial DNA analysis of the roseate tern revealed historical divergence between breeding colonies in the Indo-Pacific and the Atlantic (Chapter 3). Within each oceanic basin, roseate tern haplotypes showed a pattern consistent with a recent range expansion, and no genetic structure was detected within either oceanic basin. The most recent colonisation of current locations of roseate tern colonies is likely to have followed the last glaciation, occurring once sea levels had stabilised at their present levels approximately 7 000 years ago (Pirazzoli 1991). The mutation rate of mitochondrial DNA is insufficient for detecting population structure which may have developed over this short evolutionary timescale.

Microsatellites are commonly used to examine population structure in recently diverged populations. Many of the species distributions seen today are a result of range shift or expansion since the most recent glacial maximum. As in the roseate tern, mitochondrial DNA may provide evidence of range expansion, but the short evolutionary time since the colonisation of current ranges often means that mitochondrial DNA is unable to provide information about the current relationships among populations. In such instances, microsatellites, with their higher rate of mutation, are commonly employed to elucidate contemporary population structure (Andersen *et al.* 1998; Bérubé *et al.* 1998; Small *et al.* 2003). Mitochondrial DNA analyses of coastal and island populations of the American marten, *Martes americana*, in Alaska found evidence of two clades,

corresponding to the continental subspecies *M. a. americana* and the coastal subspecies *M. a. caurina*, both of which showed evidence of recent range expansion (Stone *et al.* 2002). The authors proposed that the two clades had diverged due to separation in different glacial refugia. However, there was little evidence for genetic structure within either clade, and microsatellite analyses were subsequently undertaken to investigate this. Microsatellites did reveal divergence among populations within each clade, with stronger divergence among populations of *M. a. caurina* (Small *et al.* 2003). To explain this pattern, Small *et al.* (2003) suggested an earlier range expansion of *M. a. caurina*, with island populations subsequently created and isolated due to Holocene sea level rise. The lesser divergence among populations of *M. a. americana* was explained by a more recent range expansion coupled with occasional gene flow among continental populations (Small *et al.* 2003).

The few microsatellite studies of genetic structure in seabirds have produced contrasting results, even within genera. Abbott & Double's (2003a) research on shy (*Thalassarche cauta*) and white-capped albatross (*T. steadi*) revealed strong population structuring among the shy albatross colonies, but not among colonies of the white-capped albatross. This difference was attributed to distance between colonies, with the three white-capped albatross colonies closer to one another than the three colonies of shy albatross (3 – 25 km between colonies cf. 50 – 500 km, respectively). A study of black-browed (*T. melanophris* and *T. impavida*) and grey-headed albatross (*T. chrysostoma*) across their sub-Antarctic distribution found three divergent groups of black-browed albatross, but no structure among colonies of grey-headed albatross (Burg & Croxall 2001). This result was explained by differences in the foraging and dispersal patterns of the two species: while the three divergent groups of black-browed albatross correspond with

differences in foraging grounds, grey-headed albatross roam more widely and individuals from different colonies are believed to mix at sea, potentially allowing inter-colony matings to occur. Roeder *et al*. (2001) found no evidence for genetic divergence of Adelie penguin colonies around Antarctica, despite the results of banding studies inferring high rates of philopatry in the species. A microsatellite study focussing on the two main colonies of Audouin's gull in the Mediterranean found no evidence of differentiation between the two colonies, suggesting gene flow between the two was a regular occurrence (Genovart *et al*. 2003). These examples illustrate that, while seabirds are generally considered to exhibit philopatry, the strength of that philopatry can differ greatly, even when comparing closely related species. In the roseate tern, levels of philopatry appear to differ in different areas, with banding studies suggesting that roseate terns in North America display philopatry at an island level, while in Australia, the tendency for breeding colonies to occur on different islands in different years precludes philopatry below a regional level (see Chapter 1).

In this chapter I describe my use of microsatellites to genotype roseate tern samples from most of the species' range. While microsatellites have not commonly been used for population studies on a global scale, this allows me to explore contemporary relationships in different regions and at a number of spatial scales, including between the two oceanic basins (Atlantic and Indo-Pacific), within each oceanic basin, and among colonies in Japan and Australia. Finding genetic structure among colonies in Australia and Japan may also allow the origin of individuals in the non-breeding aggregations in Australia to be identified.

4.2 Methods

Samples for microsatellite analysis were collected from five breeding colonies and three non-breeding aggregations in Australia, and from breeding colonies in Japan, the Seychelles, South Africa, the Azores, Ireland and the United States, as described previously (see Figure 2.1 and Table 3.1). The number of samples from breeding colonies ranged from four, from Sisters Island, to 22, from Lady Elliot Island. Larger sample sizes, of at least 35 individuals, were obtained from the non-breeding aggregations to allow assessment of their composition if genetic structure was detected among the breeding colonies. DNA was extracted as described in Chapters 2 and 3.

Samples were genotyped at seven microsatellite loci: three dinucleotide loci isolated from the red-billed gull (RBG18, RBG27 and RBG29; Given *et al*. 2002), and four from the roseate tern (Sdaat20, Sdaat27, Sdaat46 and Scaac20; P. Szczys, unpublished data). The roseate tern microsatellite motifs contained interrupted repeats; as a result alleles were anticipated to differ by as little as a single nucleotide, whereas the alleles at the red-billed gull loci should occur an even number of nucleotides apart. Each 12.5 μL PCR reaction mix contained 8 - 10 ng DNA, 1 x PCR buffer (containing Tris-HCl, KCl and $(NH_4)_2SO_4$), 0.08 mM of each dNTP, 2.5 mM $MgCl_2$, 0.24 μM of each primer, one of which was fluorescently labelled with HEX, FAM or TET, and 0.5 units of Taq DNA polymerase (QIAGEN). A touchdown PCR was used, which included a series of steps, each consisting of two cycles of primer annealing for 30 s, polymerisation at 72°C for 45 s and denaturation at 94°C for 10 s. For the red-billed gull loci, the initial annealing temperature was 65°C and was reduced by 1°C each step to a final annealing temperature of 60°C, at which 28 cycles of annealing, polymerisation and denaturation

occurred, followed by a final two minute extension at 72°C. The roseate tern loci were amplified using a similar profile, but the initial annealing temperature was 62°C, decreased by 1°C per step to 60°C, and then by 2°C per step to a final annealing temperature of 56°C. Products were run on a MegaBace Genetic Analyser (Amersham Biosciences) using ET 400-R (Amersham Biosciences) internal size standard. Allele sizes were determined using Fragment Profiler (Amersham Biosciences).

Allele frequencies and observed and expected heterozygosity were calculated in Arlequin v2.0 (Schneider *et al.* 1997). Loci were tested for Hardy-Weinberg equilibrium and linkage disequilibrium, which tests for the independence of the loci, for each breeding colony and non-breeding aggregation. Mitochondrial DNA analyses identified two lineages of roseate tern: an Atlantic lineage and an Indo-Pacific lineage (Chapter 3). To examine the contemporary relationships among the breeding colonies in each lineage, pairwise F_{ST} (Weir & Cockerham 1984) and R_{ST} (Slatkin 1995) were calculated separately for the Atlantic and the Indo-Pacific colonies. F_{ST} is calculated from measures of heterozygosity, and assumes that microsatellites conform to the infinite allele model (IAM), in which every mutation produces a unique allele. R_{ST} assumes stepwise mutation, where mutation results in the loss or gain of a single repeat unit, and takes into account allele size. South Africa has been identified as an area of secondary contact between the Atlantic and Indo-Pacific mitochondrial lineages (Chapter 3), and as it is likely to be subject to gene flow from both oceanic basins, it was included in comparisons of pairwise F_{ST} and R_{ST} for both oceanic basins. The population pairwise F_{ST} values were used for an analysis of isolation by distance using a Mantel test in IBD v1.4 (Bohonak 2002), for all breeding colonies, for each oceanic basin (including South Africa in both because it may be involved in gene flow in both

oceanic basins), for the Australian and Japanese colonies, and for the Australian colonies only. Geographic distance was measured as the shortest distance between sampling locations without crossing major land masses. An analysis of molecular variance (Excoffier *et al*. 1992) of breeding colonies was performed at four hierarchical levels: between mitochondrial lineages (Atlantic and Indo-Pacific), excluding South Africa as it contains both mitochondrial lineages; among colonies within the Indo-Pacific; between Australian and Japanese colonies; and between the colonies in the two Australian states, Queensland and Western Australia. Samples from the non-breeding aggregations were excluded from calculations of pairwise F_{ST}, R_{ST}, AMOVA and isolation by distance due to the high likelihood, and in some instances evidence, that the aggregations contain individuals from the sampled breeding colonies.

4.3 Results

A total of 308 individuals were genotyped at seven microsatellite loci. Due to interruptions in the microsatellite repeats at loci Sdaat20, Sdaat27 and Scaac20, alleles commonly occurred one nucleotide apart. The chromatograms for Sdaat20 allowed alleles to be scored consistently, however variable chromatogram profiles for Sdaat27 and Scaac20 prevented consistent scoring for these two loci and they were excluded from final analyses. Significant linkage disequilibrium of loci was detected in seven of the 70 population-locus comparisons, however these did not consistently involve the same population or pair of loci. The five loci were thus considered to be independent.

A total of 33 alleles were scored from the five microsatellite loci used for analysis: RBG18, RBG27, RBG29, Sdaat20 and Sdaat46 (Table 4.1; see Appendix 3 for individual genotypes). Loci had few gaps in expected allele sizes, conforming predominantly to the stepwise mutation model (SMM). The number of alleles observed at each locus ranged from three, at Sdaat46, to 11, at Sdaat20. Japan was the only breeding colony with a private allele (locus RBG18, allele 173), with a further four private alleles occurring in non-breeding aggregations. Private alleles always occurred at low frequency. Values for observed heterozygosity for each sampled group at each locus varied from 0 to 0.818 (Table 4.1). The absence of heterozygosity was most commonly due to allele fixation, with four breeding colonies displaying fixed alleles for at least one locus: Sisters Island at RBG18; Ireland and the United States at RBG29; and all three north Atlantic breeding colonies at Sdaat46. Although not a result of allele fixation, all individuals from Lancelin Island and Pelsaert Island were homozygous at locus Sdaat20; this was also the case for Lancelin Island at RBG18.

Eight of the fourteen sampled populations had significant heterozygote deficiencies at one or more microsatellite loci (Table 4.1), however none showed a consistent departure from Hardy-Weinberg equilibrium across all loci. The non-breeding aggregations most commonly deviated from expectation, with the Swain Reefs non-breeding aggregations showing significant heterozygote deficiencies at two and three loci, for the summer and winter aggregations, respectively. Heterozygote deficiencies were anticipated for the non-breeding aggregations, as the combination of individuals from multiple breeding colonies increases the number of alleles observed, without increasing the number of combinations in which they occur (Wahlund effect). The expected heterozygosity is thus elevated without any increase in the observed heterozygosity, resulting in a

Table 4.1: Sample numbers, allele frequencies, observed (H_O) and expected (H_E) heterozygosity for each sampled roseate tern group at the five microsatellite loci used for analyses. Pels = Pelsaert, Lanc = Lancelin, Boun = Bountiful, Lacep = Lacepede, Swains summ = Swains summer, Seych = Seychelles, Ire = Ireland. * indicates a significant departure from Hardy-Weinberg equilibrium at $P < 0.05$

Locus	*Allele*	Pels	Lanc	Lady Elliot	Boun	Sisters	Lacep	Swains summ	Swains winter	Japan	Azores	South Africa	Seych	Ire	United States
RBG18		N = 19	N = 10	N = 21	N = 18	N = 4	N = 35	N = 79	N = 48	N = 19	N = 14	N = 6	N = 11	N = 10	N = 10
	173	0	0	0	0	0	0	0	0	0.026	0	0	0	0	0
	175	0.026	0	0	0	0	0.043	0.006	0	0	0.25	0.583	0.227	0.1	0.4
	177	0	0	0	0	0	0.043	0.006	0.021	0.026	0	0.083	0	0	0
	179	0.842	0.9	0.976	0.861	1	0.657	0.785	0.823	0.684	0	0.083	0.5	0	0
	181	0.132	0.1	0	0.083	0	0.2	0.165	0.156	0.211	0.214	0.25	0.227	0.45	0.45
	183	0	0	0	0.56	0	0.029	0.019	0	0.026	0.536	0	0.045	0.45	0.15
	185	0	0	0.024	0	0	0.029	0.013	0	0.026	0	0	0	0	0
	187	0	0	0	0	0	0	0.006	0	0	0	0	0	0	0
	H_O	0.211	0	0.048	0.278	0	0.486	0.304	0.354	0.421	0.571	0.667	0.636	0.3	0.6
	H_E	0.324	0.284	0.094	0.303	0	0.531	0.359	0.318	0.508	0.677	0.636	0.693	0.705	0.684
RBG29		N = 19	N = 10	N = 22	N = 18	N = 4	N = 35	N = 79	N = 51	N = 19	N = 14	N = 6	N = 11	N = 10	N = 10
	134	0	0	0.136	0	0.125	0.029	0.044	0.098	0.053	0.107	0.167	0	0	0
	136	0.421	0.25	0.341	0.306	0.25	0.4	0.418	0.392	0.526	0.893	0.75	0.591	1	1
	138	0	0	0.023	0.028	0	0.043	0.044	0.01	0.026	0	0	0	0	0
	144	0.579	0.75	0.5	0.667	0.625	0.529	0.494	0.5	0.395	0	0.083	0.409	0	0
	H_O	0.526	0.3	0.545	0.333	0.5	0.6	0.582	0.725	0.474	0.071	0.5	0.818	0	0
	H_E	0.531	0.395	0.651	0.527	0.607	0.581	0.582	0.6	0.579	0.267	0.439	0.506	0	0

Table 4.1 continued: Allele frequencies, observed (H_O) and expected (H_E) heterozygosity for each sampled roseate tern group at the five microsatellite loci used for analyses.

Locus	*Allele*	Pels	Lanc	Lady Elliot	Boun	Sisters	Lacep	Swains summ	Swains winter	Japan	Azores	South Africa	Seych	Ire	United States
RBG27		N = 19	N = 10	N = 22	N = 18	N = 4	N = 35	N = 79	N = 51	N = 19	N = 14	N = 6	N = 11	N = 10	N = 10
	195	0.132	0.25	0.295	0.056	0.125	0.1	0.114	0.294	0.053	0	0.083	0.091	0	0
	196	0	0	0	0	0	0	0	0.01	0	0	0	0	0	0
	197	0.263	0.05	0.023	0.222	0.25	0.114	0.152	0.049	0.184	0.536	0.667	0.455	0.5	0.6
	199	0	0	0	0	0	0	0.006	0	0	0.107	0	0	0.05	0
	201	0.605	0.7	0.545	0.639	0.625	0.7	0.608	0.48	0.658	0.357	0.167	0.409	0.45	0.4
	203	0	0	0.091	0.083	0	0.043	0.114	0.088	0.026	0	0	0	0	0
	205	0	0	0.045	0	0	0.043	0.006	0.078	0.079	0	0.083	0.045	0	0
	H_O	0.579	0.2	0.409*	0.278*	0.5	0.457	0.342*	0.49*	0.421	0.5	0.333	0.545	0.4	0.4
	H_E	0.562	0.468	0.618	0.597	0.75	0.491	0.594	0.684	0.573	0.632	0.682	0.645	0.616	0.505
Sdaat46		N = 19	N = 10	N = 21	N = 18	N = 4	N = 35	N = 78	N = 49	N = 18	N = 14	N = 6	N = 11	N = 10	N = 10
	243	0.658	0.6	0.333	0.528	0.625	0.5	0.333	0.306	0.472	0	0.167	0.227	0	0
	246	0	0	0.119	0.083	0	0.043	0.096	0.051	0.139	0	0	0.045	0	0
	250	0.342	0.4	0.548	0.389	0.375	0.457	0.571	0.643	0.389	1	0.833	0.727	1	1
	H_O	0.579	0.4	0.476	0.5	0.75	0.6	0.577	0.265*	0.5	0	0.333	0.545	0	0
	H_E	0.462	0.563	0.63	0.579	0.536	0.547	0.558	0.495	0.651	0	0.439	0.437	0	0

* indicates a significant departure from Hardy-Weinberg equilibrium at $P < 0.05$

Table 4.1 continued: Allele frequencies, observed (H_O) and expected (H_E) heterozygosity for each sampled roseate tern group at the five microsatellite loci used for analyses.

Locus	*Allele*	Pels	Lanc	Lady Elliot	Boun	Sisters	Lacep	Swains summ	Swains winter	Japan	Azores	South Africa	Seych	Ire	United States
Sdaat20		N = 19	N = 10	N = 22	N = 18	N = 4	N = 35	N = 79	N = 51	N = 19	N = 14	N = 6	N = 11	N = 10	N = 7
	187	0.158	0.2	0.068	0.222	0.25	0.114	0.203	0.039	0.105	0	0	0.136	0	0
	189	0	0	0	0	0	0.014	0	0	0	0	0	0	0	0
	191	0	0	0	0	0	0.014	0	0	0	0	0	0.045	0	0
	192	0.053	0.2	0	0.056	0.25	0.114	0.082	0.069	0.079	0	0.083	0.091	0	0
	193	0.737	0.5	0.932	0.694	0.5	0.7	0.627	0.892	0.684	0	0	0.318	0	0
	194	0.053	0.1	0	0	0	0	0.032	0	0.079	0	0	0	0	0
	196	0	0	0	0	0	0	0	0	0	0.25	0	0	0.85	0.857
	198	0	0	0	0	0	0	0.006	0	0	0	0	0	0	0
	202	0	0	0	0	0	0.043	0.051	0	0	0	0	0	0	0
	203	0	0	0	0.028	0	0	0	0	0.053	0	0	0.045	0	0
	206	0	0	0	0	0	0	0	0	0	0.75	0.917	0.364	0.15	0.143
	H_O	0*	0*	0.045	0.5	0.5	0.143*	0.253*	0.098*	0.263*	0.357	0.167	0.545	0.3	0
	H_E	0.522	0.779	0.172	0.478	0.75	0.514	0.569	0.217	0.565	0.442	0.167	0.801	0.353	0.396

* indicates a significant departure from Hardy-Weinberg equilibrium at $P < 0.05$

significant departure from Hardy-Weinberg equilibrium. None of the breeding colonies showed significant departures from Hardy-Weinberg equilibrium at more than one locus. Loci RBG27 and Sdaat20 showed heterozygote deficiencies for multiple breeding colonies: two and three, respectively. Significant heterozygote deficiencies can indicate the presence of null alleles. In this instance, however, as they do not occur consistently across all colonies it seems unlikely to be the reason for the colony specific heterozygote deficiencies.

Atlantic vs. Indo-Pacific

Atlantic breeding colonies had, on average, fewer alleles than Indo-Pacific colonies ($t = -8.5$, $df = 4$, $P < 0.05$), but did not have significantly lower average heterozygosity ($t = -2.53$, $df = 4$, $P = 0.065$) (Table 4.2). Strong differentiation was evident between the Atlantic and Indo-Pacific colonies, with 38.7% of the genetic variation distributed between the two oceanic basins ($F_{ST} = 0.43$, $R_{ST} = 0.52$, $P < 0.001$). Of the 33 alleles scored, 22 were found only in individuals sampled in the Indo-Pacific, with a single allele (locus Sdaat20, allele 196) unique to the three Atlantic breeding colonies (Table 4.1). There was a strong correlation between genetic differentiation and geographic distance when all roseate tern breeding colonies were included in the Mantel test ($r = 0.85$, $P < 0.001$).

Atlantic

Within the Atlantic Ocean, all pairwise F_{ST} and R_{ST} values were significant, with the exception of the Ireland – United States comparison ($F_{ST} = 0.01$, $P = 0.4$; $R_{ST} = 0.05$, $P = 0.27$; Table 4.3). The Mantel test did not support an isolation by distance effect among the Atlantic Ocean breeding colonies ($r = 0.57$, $P = 0.095$).

Table 4.2: Number of samples from breeding colonies in each region, and genetic diversity, measured as the average number of alleles (A) and average observed heterozygosity (H_{AV})

	Ireland	Azores	United States	South Africa	Seychelles	Japan	Australia
N	10	14	10	6	11	19	73
A	2	2.2	1.8	3	3.8	4.6	4.4
H_{AV}	0.2	0.3	0.2	0.4	0.62	0.42	0.34

Table 4.3: Pairwise F_{ST} (above diagonal) and R_{ST} values (below diagonal) between colonies of the Atlantic lineage of roseate tern, and the South African colony, which contains individuals with the haplotype typical of the Atlantic lineage.

	Azores	United States	Ireland	South Africa
Azores	-	0.16**	0.17**	0.11**
United States	0.28*	-	0.01	0.25**
Ireland	0.36**	0.05	-	0.35**
South Africa	0.14**	0.38*	0.53**	-

* indicates $P < 0.05$,** indicates $P < 0.001$

Indo-Pacific

Within the Indo-Pacific, the AMOVA revealed significant divergence among breeding colonies, grouped by country of origin ($F_{ST} = 0.08$, $R_{ST} = 0.12$, $P < 0.001$). Pairwise F_{ST} and R_{ST} values were significant for all comparisons involving South Africa, and the comparison with Sisters Island yielded the only non-significant pairwise value for the Seychelles ($R_{ST} = 0.16$, $P = 0.13$; Table 4.4). Most comparisons between the remaining Indo-Pacific breeding colonies, those in Australia and Japan, were non-significant; the only exception involved Lady Elliot Island, for which significant pairwise F_{ST} values

Table 4.4: Pairwise F_{ST} (above diagonal) and R_{ST} values (below diagonal) between the Indo-Pacific lineage of roseate tern, and the South African colony, which contains individuals with the mitochondrial haplotype typical of the Indo-Pacific lineage. L. E. = Lady Elliot, Sist = Sisters, Bount = Bountiful, Pels = Pelsaert, Lanc = Lancelin, Seych = Seychelles, S. A. = South Africa

	L. E.	Sist	Bount	Pels	Lanc	Japan	Seych	S. A.
L. E.	-	0.06	0.05*	0.06**	0.08*	0.06*	0.19**	0.48**
Sist	-0.01	-	-0.05	-0.03	-0.06	0.007	0.1**	0.43**
Bount	0.03	-0.06	-	-0.01	-0.0007	0.02	0.13**	0.43**
Pels	0.03	-0.07	-0.007	-	0.01	0.009	0.13**	0.43**
Lanc	0.05	-0.07	-0.03	-0.01	-	0.04	0.15**	0.46**
Japan	0.001	0.002	0.04	0.03	0.08	-	0.08**	0.35**
Seych	0.14**	0.16	0.22**	0.22**	0.22**	0.14**	-	0.11**
S. A.	0.66**	0.7**	0.71**	0.71**	0.72**	0.65**	0.28**	-

* indicates $P < 0.05$, ** indicates $P < 0.001$

were obtained for all but one comparison (with Sisters Island). The AMOVA of the Australian and Japanese colonies revealed weak, but significant, differentiation when colonies were grouped into each of the two regions for F_{ST} , but not R_{ST} ($F_{ST} = 0.03$, $P = 0.03$, $R_{ST} = 0.03$, $P = 0.22$). Within Australia, colonies in Western Australia and Queensland were not significantly differentiated ($F_{ST} = 0.03$, $P = 0.07$, $R_{ST} = 0.01$, $P = 0.49$). Overall, there was evidence for isolation by distance among breeding colonies in the Indo-Pacific ($r = 0.74$, $P = 0.013$). However, there was no support for isolation by distance among the breeding colonies in Australia and Japan ($r = 0.15$, $P = 0.32$), or among the breeding colonies in Australia ($r = 0.04$, $P = 0.47$).

4.4 Discussion

Microsatellite analyses reveal a global pattern of isolation by distance for breeding colonies of the roseate tern. This is largely attributable to a lack of gene flow between the Atlantic and Indo-Pacific oceanic basins, which corresponds with a geographic gap in the distribution of roseate tern breeding colonies. The large number of alleles unique to the Indo-Pacific highlights the lack of gene flow between the two oceanic basins. Even with the range expansion to higher latitudes following the last sea level rise (see Chapter 3), and the establishment of breeding colonies in South Africa, between the two oceanic basins, there appears to have been little transfer of microsatellite alleles around continental Africa. However, despite their historical separation, individuals from the two lineages are interbreeding in South Africa, as evidenced by the occurrence of microsatellite alleles unique to the Indo-Pacific lineage in individuals with mitochondrial haplotypes from the Atlantic lineage.

Africa has also restricted gene flow in the blue marlin, though not to the same extent as in the roseate tern. Using microsatellites, Buonaccorsi *et al.* (2001) found significant divergence between Atlantic and Indo-Pacific blue marlin. They identified two groups of alleles: one group unique to blue marlin in the Atlantic; and one which occurred in individuals in both the Indo-Pacific and the Atlantic. This pattern of allele distribution, in combination with relatively low divergence (14.7% of variation distributed between oceans), alluded to unidirectional gene flow in blue marlin from the Indo-Pacific to the Atlantic (Buonaccorsi *et al.* 2001). In the roseate tern, the large number of unique alleles in the Indo-Pacific indicates that gene flow between the two oceanic basins, if it occurs, is most likely to take the opposite direction, Atlantic to Indo-Pacific.

Atlantic Ocean

Microsatellites reveal that, in the Atlantic, breeding colonies in Ireland and North America are not significantly divergent from one another. This may be a result of low sample sizes, with the samples collected not an accurate representation of the genetic diversity in the Atlantic breeding colonies. Another possible explanation for this non-significant comparison is that the colonies in the United States, Ireland and the Azores all interchange individuals, and are effectively a genetically homogeneous population. However, the population declines recorded in the United States and Ireland in the last century (Witherby & Ticehurst 1908; Nisbet 1973; Cramp *et al*. 1974; Nisbet 1980) are likely to have reduced genetic diversity in these colonies relative to the Azores, decreasing the value of F_{ST} between Ireland and the United States, and increasing the value of F_{ST} for each of the comparisons with the Azores. Several studies have demonstrated a reduction in genetic diversity as a result of population declines, using a combination of pre- and post-decline samples (Nielsen *et al*. 1997; Bouzat *et al*. 1998; Rajora *et al*. 2000; Whitehouse & Harley 2001; Larson *et al*. 2002; England *et al*. 2003; Hutchinson *et al*. 2003). For example, in the white pine in Canada, an approximate 75% reduction in the white pine population reduced the number of microsatellite alleles by 26%, but heterozygosity by less than 5% (Rajora *et al*. 2000). The alleles lost, however, were all low frequency alleles (less than 0.25), with 50% of the alleles lost occurring in the pre-harvest population at frequencies less than 0.01. Roseate tern colonies in North America underwent an estimated 40-50% decline in the 1970s (Nisbet 1980), and those in Ireland a 40% decline in the 1960s (Cramp *et al*. 1974), which, while not as severe as the white pine example, are likely to have also caused the loss of rare alleles. The rare alleles observed in the current breeding colonies in the Azores (in each instance at a

frequency of 0.11) may therefore have been present in the colonies in the United States and Ireland prior to their decline. Due to low sample sizes, however, these interpretations are preliminary and should be considered with caution.

An alternate explanation for the non-significant relationship between the colonies in Ireland and the United States is that gene flow is more common between colonies in Ireland and the United States than between Ireland and the geographically closer colonies in the Azores. The limited band recovery data from these three areas supports the greater movement of birds between Ireland and the United States than between either of these regions and the Azores, although the lower banding intensity in the Azores in comparison to the two former regions may explain this imbalance (Hays *et al.* 2002a). Four roseate terns banded in the main breeding colony in Ireland, Rockabill Island, have been reported in breeding colonies in the United States (Nisbet & Cabot 1995; Hays *et al.* 2002b). There are also unconfirmed reports of movement in the opposite direction, with two individuals banded in the United States observed at Rockabill Island, Ireland (Newton & Crowe 2000). By contrast, there is a single report of a roseate tern moving from the Azores to a colony in north-west Europe, with one individual banded in the Azores sighted at a colony in Wales (Newton & Crowe 2000). There are no reports of movements from north-west Europe to the Azores, and a single record of a bird carrying a plastic band of the kind used in roseate tern colonies in North America and the Caribbean (but with no numbered metal band with which to confirm the origin) sighted in a colony in the Azores (Hays *et al.* 2002a).

The transatlantic movement of roseate terns from one colony to another may occur via mixing of individuals on wintering grounds. Adult roseate terns from North America

and the Caribbean winter in South America, and young birds spend their first two summers on the wintering grounds in South America before returning north for their first breeding attempt (Hamilton 1981; Nisbet 1984; Hays *et al*. 1997; Hays *et al*. 1999). Band recoveries of birds banded in Great Britain and the Azores have occurred in West Africa, suggesting that this is the wintering area for east Atlantic roseate terns (Langham 1971; Hays *et al*. 2002a). On occasions, however, birds have been banded on one side of the Atlantic and recovered on another. Hays *et al*. (2002a) reported the recovery of two roseate terns in the Azores which had been banded on wintering grounds elsewhere: one in Ghana, the presumed wintering ground for Azores birds; and one at Mangue Seco, Brazil, the wintering ground for west Atlantic birds. Previous recoveries at Mangue Seco included a roseate tern banded at Rockabill Island, Ireland (Hays *et al*. 2000). The latter two recoveries demonstrate that, at least occasionally, birds from breeding colonies in the east Atlantic cross the Atlantic Ocean to the South American coast, where roseate terns from the west Atlantic spend their winter. There is thus potential for birds to accompany other migrating individuals back to a breeding area other than the one from which they originated. However, as birds from both Ireland and the Azores have been captured in South America, this does not explain why Ireland and the United States show evidence for a greater exchange of individuals than either nation has with the Azores. With only a single band recovery in South America from each of the Azores and Ireland, it is not possible to draw any conclusions about the frequency of transatlantic migrations and mixing of roseate terns from opposite sides of the Atlantic. Mixing of east and west Atlantic birds in South America may be only an occasional occurrence, and particularly rare for birds from the Azores. More band recovery data is required to better understand the interactions among individuals from the Atlantic Ocean colonies.

Indo-Pacific

Roseate tern colonies in the Indo-Pacific show an overall pattern of isolation by distance, with the breeding colonies in the Seychelles and South Africa highly divergent from other Indo-Pacific colonies. This was anticipated for the colony in South Africa as it comprises individuals from both mitochondrial lineages (see Chapter 3). The divergence of breeding colonies in the Seychelles from those in Japan and Australia suggests that, since the most recent range expansion, there has been little movement of birds across the Indian Ocean. Roseate terns are not present in the Seychelles outside the breeding period (Warman 1979), and while their non-breeding distribution is unknown, it appears unlikely to overlap with individuals from breeding colonies in Australia or Japan.

Australia and Japan

The results of AMOVA and pairwise comparisons demonstrate the occurrence of gene flow between breeding colonies in Australia and Japan. Although sample sizes at some colonies, particularly Sisters Island, were too low to draw any conclusions about the processes specifically affecting those colonies, the overall pattern is clear. While the recent colonisation of the region means that genetic equilibrium of microsatellite markers may not yet have been reached, the divergence of the colony in the Seychelles indicates that the weak divergence between colonies in Australia and Japan may be due to gene flow, rather than insufficient time for divergence to occur at microsatellite loci. Within Australia, there is no consistent evidence of divergence among breeding colonies. Some pairwise comparisons involving Lady Elliot Island were significant, but only for F_{ST}. As the microsatellites used in this study appear to follow a mutation model closer to that of the stepwise mutation model (SMM) than the infinite allele model

(IAM; see section 4.2 for assumptions of each model), R_{ST} should be a more appropriate measure of divergence than F_{ST}. As pairwise comparisons were not consistently significant for both F_{ST} and R_{ST}, there is not clear evidence for the divergence of the Lady Elliot Island colony from other Australian breeding colonies. The overall trend of a lack of differentiation among breeding colonies in Australia and Japan means that microsatellites cannot be used to identify the origins of individuals in non-breeding aggregations.

The lack of genetic structure among Australian breeding colonies of roseate tern contrasts with the structure observed among green turtle rookeries in the same area, individuals from which display philopatry as well as long distance dispersal to non-breeding grounds. A microsatellite study of rookeries in the southern and northern regions of the Great Barrier Reef, the Gulf of Carpentaria, and north and central regions of Western Australia revealed three distinct rookery groups for the green turtle (FitzSimmons *et al.* 1997). These were Western Australia; Gulf of Carpentaria; and Great Barrier Reef (FitzSimmons *et al.* 1997), demonstrating that philopatry is stronger in the green turtle than I have observed in the roseate tern. However, as in the roseate tern, there was no significant microsatellite differentiation between green turtle rookeries in the northern and southern regions of the Great Barrier Reef. The primary explanation for the absence of nuclear genetic structure in green turtles on the Great Barrier Reef was the overlap of the migration path of turtles from the southern region with the courtship and mating grounds of turtles breeding in the northern region. Gene flow between the two regions was postulated to occur in the mating grounds of the northern region, where northern region turtles could mate with individuals migrating south (FitzSimmons *et al.* 1997).

Similar interactions to those proposed for green turtles may be responsible for gene flow among breeding colonies of the roseate tern. Gene flow among breeding colonies in Australia and Japan may be facilitated in the non-breeding aggregations in Australia. In the non-breeding aggregations in the Swain Reefs, band recoveries have demonstrated that birds from at least two regions mix in both summer and winter: those from Japan and from colonies in the southern Great Barrier Reef (O'Neill *et al*. in press). In the non-breeding aggregation in the Lacepede Islands, Western Australia, the size of the aggregation suggests that birds from multiple colonies are mixing. Overlap of non-breeding individuals provides the potential for gene flow to occur via pairings in the non-breeding area, or through individuals returning to breeding colonies other than their natal colony. Alternatively, there may be regular movements of birds between breeding colonies regardless of mixing in non-breeding aggregations, as demonstrated in the only other microsatellite study of a larid (Genovart *et al*. 2003). Genovart *et al*. (2003) examined the two main breeding colonies of Audouin's gull in the Mediterranean and showed that, despite being separated by 650 km, the two colonies were not significantly divergent, implying that birds regularly move from one colony to the other. Few, if any, roseate tern breeding colonies in Australia would be more than 500 km from another colony, so gene flow could occur via intercolony movements between neighbouring colonies on a similar scale to those in Audouin's gull.

In summary, microsatellites show strong divergence between Atlantic and Indo-Pacific colonies of roseate tern, concordant with the mitochondrial evidence for historical separation of the two ocean basins (Chapter 3). The distribution of microsatellite alleles provides evidence that, despite range expansion in both oceanic basins, contemporary gene flow between the two has been restricted. Within the Atlantic, the colonies in

Ireland and the United States are the only pair to show non-significant divergence. This could be due to mixing of individuals on non-breeding grounds in Brazil, because of the reduction in diversity that both colonies are likely to have suffered as a result of the population declines or due to recent colonisation from a common ancestral population. Within the Indo-Pacific, gene flow across the Indian Ocean has been limited since the most recent range expansion, with the Seychelles colony divergent from colonies in Australia and Japan. However, there is evidence for continued gene flow among colonies in Australia and Japan, which may be facilitated by the mixing of individuals in non-breeding aggregations.

Chapter 5: Paternal leakage of mitochondrial DNA associated with hybridisation of two tern species, the roseate tern and black-naped tern

5.1 Introduction

It is estimated that 9.2% of all bird species have naturally bred with another species to produce hybrid offspring (Grant & Grant 1992). In their review, Grant & Grant (1992) observed that hybridisation occurs more frequently in terrestrial birds than in seabirds. While this may be true of seabirds generally, hybridisation has been commonly observed in the family Laridae (skuas, gulls and terns). Reports of hybridisation between gull (*Larus*) species have been based on observations of morphological intermediates (Gurr 1967; Ingolfsson 1970; Andrle 1972; Andrle 1973; Godfrey 1973; Hoffman *et al*. 1978; Andrle 1980; Panov & Monzikov 2000) or mixed species pairs (Gurr 1967). For example, Hoffman *et al.* (1978) observed large numbers of hybrid gulls in colonies along the Washington coast, where glaucous-winged gulls, *Larus glaucescens*, and western gulls, *L. occidentalis*, interbreed. Individuals were considered hybrids if they were intermediate in the colour of one or more of the characters discriminating the two species: eye rings (magenta/purple – orange/red), iris (dark brown – yellow with brown flecks) and wing tip (pale grey – black), where the bracketed colours refer to western gulls and glaucous-winged gulls, respectively. Using these criteria for identifying hybrids, Hoffman *et al.* (1978) estimated that hybrids constituted half of the breeding population at the focal island of their study, Destruction Island. Panov & Monzikov (2000) compared morphological and behavioural characters of *L. cachinnans cachinnans, L. heuglini*, and the taxonomically uncertain *L. barabensis*. Morphological characters of *L. barabensis*, including weight, bill length

and tarsus length, were intermediate to those of *L. c. cachinnans* and *L. heuglini*, although usually closer to those of *L. heuglini*. A similar pattern was observed for colouration, however many behavioural characters of *L. barabensis* resembled those of *L. c. cachinnans*. Panov & Monzikov (2000) conclude that *L. barabensis* arose via introgression, the incorporation of genes from one distinct taxon into another (Rieseberg & Wendel 1993), of *L. cachinnans* genes into *L. heuglini* during range expansion, and suggest that *L. barabensis* be considered a subspecies of *L. heuglini*, and renamed *L. heuglini barabensis*.

Recently, genetic studies have confirmed the widespread occurrence of hybridisation and introgression in the genus *Larus* (Bell 1996; Liebers *et al*. 2001; Crochet *et al*. 2002). Crochet *et al*. (2002) used cytochrome *b* mitochondrial DNA sequences to examine the relationships among 21 gull taxa. They found that *L. argentatus*, *L. fuscus* and *L. hyperboreus* have all interbred in the past, and that introgression was likely to have occurred in most of the sampled taxa. Liebers *et al*. (2001) found haplotypes resulting from introgression in four of nine gull taxa studied. Within the Stercorariidae (skuas), mixed species pairs are commonly observed between south polar skuas, *Catharacta maccormicki*, and brown skuas, *C. lonnbergi* (Pietz 1987; Parmelee 1988). Based on mitochondrial DNA, random amplified polymorphic DNA (RAPDs) and allozymes, hybridisation and introgression of congeneric species has been proposed as one explanation for the unusual phylogeny of the two skua genera, *Catharacta* and *Stercorarius* (Cohen *et al*. 1997).

Hybridisation has been inferred between a number of different tern species via observations either of mixed species pairs (Brichetti & Foschi 1987; Ross *et al*. 1999) or

individuals that are morphological intermediates (Vinicombe 1980; Davis 1982). Roseate terns, *Sterna dougallii*, have been observed mating with arctic terns, *S. paradisaea*, (Whittam 1998) and paired with common terns, *S. hirundo* (Perry 1972; Robbins 1974; Hays 1975; Zingo *et al.* 1994). The frequency of recognisable roseate x common tern hybrids is approximately 0.2% in roseate tern colonies and 0.0002% in common tern colonies in north-east America (Gochfeld *et al.* 1998; Nisbet 2002). Roseate x common tern hybrids can breed successfully with other hybrid individuals, or with individuals of either parental species (Hays 1975; Gochfeld *et al.* 1998; Nisbet 2002). Within a colony of roseate and common terns on Great Gull Island, Hays (1975) observed two pairs in which both adults were morphological intermediates. Both pairs successfully hatched two chicks, with the chicks of one pair surviving to fledging (Hays 1975). At another colony in the United States, an F2 hybrid individual, resulting from a backcross, was banded as a chick and subsequently observed successfully raising young with a common tern (Gochfeld *et al.* 1998). Observations of hybrid individuals, either as chicks or adults, have occurred in colonies that are visited frequently and observed for long periods of time, predominantly those in North America. In Australia, where hybridisation of roseate terns has not previously been reported, breeding colonies are usually observed only for the time required to obtain an estimate of the number of breeding pairs or to scan for bands or colour-flags. This time restriction, coupled with the apparently low frequency of hybridisation, means that genetic methods, rather than direct observations, have a greater likelihood of identifying hybridisation events involving the roseate tern in Australia.

One of the fundamental assumptions about mitochondrial DNA is that it is maternally and clonally inherited (Birky 1995) ie. only the mother contributes an exact copy of her

mitochondrial DNA to her offspring. In the last 20 years, however, evidence for occasional paternal inheritance of mitochondrial DNA has been observed in a small number of animal taxa: mussels (Zouros *et al*. 1992; Quesada *et al*. 1996), mice (Gyllensten *et al*. 1991), humans (Bromham *et al*. 2003), *Drosophila* (Kondo *et al*. 1990), anchovies (Magoulas & Zouros 1993) and most recently in a bird, the great tit (Kvist *et al*. 2003b). In bees, paternal mitochondrial DNA was observed in the fertilised egg, but the amount steadily decreased until it was no longer detectable in larval stage L5 (Meusel & Moritz 1993). Several of the occurrences of paternal inheritance of mtDNA, commonly termed paternal leakage, involved hybridisation events between species (Kondo *et al*. 1990; Zouros *et al*. 1992), subspecies (Meusel & Moritz 1993; Kvist *et al*. 2003b), or strains (Gyllensten *et al*. 1991). In intraspecies matings within cattle and mice, the mitochondria in sperm are destroyed within the fertilised egg, preventing their transfer to offspring (Kaneda *et al*. 1995; Sutovsky *et al*. 2000). Recognition of species-specific molecules on the surface of the sperm mitochondria triggers the destruction of the mitochondria (Sutovsky *et al*. 2000). However, in inter-specific hybridisations the surface molecules on the mitochondria are not always recognised by the species-specific recognition system. This prevents the destruction of sperm mitochondria and allows the transfer of paternal mitochondrial DNA to the hybrid offspring (Kaneda *et al*. 1995; Shitara *et al*. 1998; Sutovsky *et al*. 2000).

Paternal inheritance usually occurs along with the normal maternal inheritance of the mitochondrial genome, and thus creates mitochondrial heteroplasmy, the existence of two mitochondrial haplotypes, in the hybrid individual and potentially its descendants. Kondo *et al*. (1990), however, found that after repeated backcrossing of inter-specific *Drosophila* hybrids, paternal mtDNA completely replaced maternal mtDNA in some

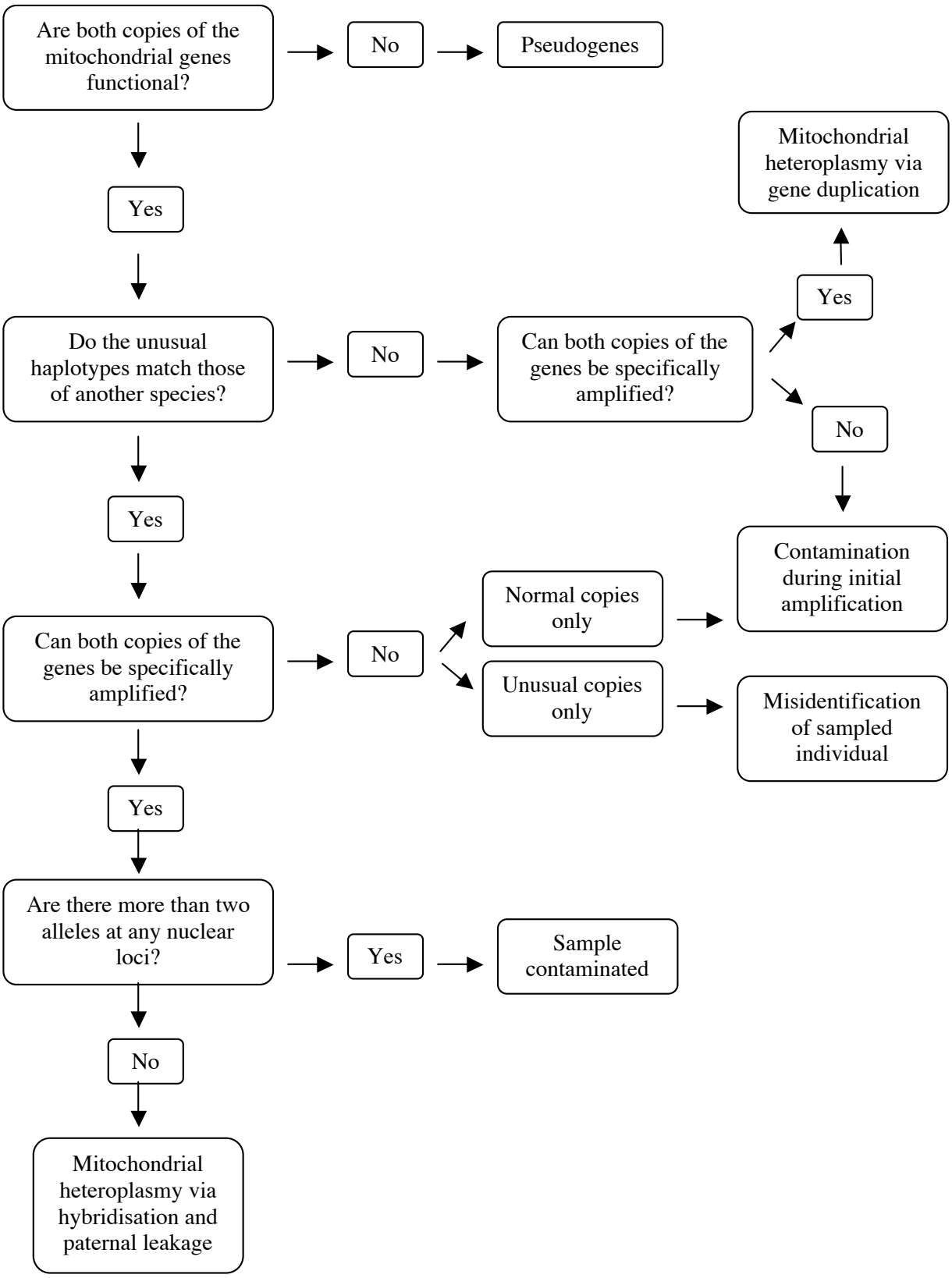
individuals. Early estimates from mice and *Drosophila* suggest that paternal leakage can occur as often as once in 10^3 to 10^4 fertilisations in some lineages (Kondo *et al*. 1990; Gyllensten *et al*. 1991).

This study presents the first genetic evidence of hybridisation and introgression between two species of tern, and the first evidence of hybridisation and paternal leakage of mitochondrial DNA between the roseate tern, *Sterna dougallii*, and the black-naped tern, *S. sumatrana*.

5.2 Methods and Results

During a study of mitochondrial differentiation in roseate terns from Japan and Australia, unusual, divergent haplotypes were observed in one individual (see section 2.3; see Appendix 4 for unusual haplotype alignment). In this individual, RTS4, two ND6 haplotypes were obtained: one typical roseate tern haplotype; and another differing at 28 nucleotide positions. The same individual had a single divergent ND2 haplotype, differing from the common roseate tern haplotype at 26 nucleotide positions. Both divergent haplotypes appeared to be functional mitochondrial copies of the two genes (see section 2.3). Subsequent investigations, reported here, involved DNA sequencing of other tern species, specific amplification of each of the haplotypes and microsatellite genotyping of individual RTS4 to determine whether the heteroplasmy observed at ND6 was evident at ND2, and whether such heteroplasmy was a result of paternal leakage (Figure 5.1).

Figure 5.1: Determining the origin of the unusual haplotypes amplified at the ND2 and ND6 mitochondrial genes, along with normal haplotypes at each gene, in the roseate tern, *Sterna dougallii*.



Sequencing of other tern species

The roseate tern sample with unusual haplotypes, RTS4, was collected on the Great Barrier Reef, Queensland, from a chick in a breeding colony. The ND2 and ND6 genes of three tern species commonly found breeding on the Great Barrier Reef, with which the roseate tern could therefore potentially hybridise, were amplified and sequenced as described in Chapter 2. These three species were the black-naped tern (*Sterna sumatrana*, N = 1), bridled tern (*S. anaethetus*, N = 2) and crested tern (*S. bergii*, N = 2). For the crested tern, different individuals were used for each of the two genes, so for each gene N = 1.

For the four species, 993 base pairs of the ND2 gene, and 352 base pairs of the ND6 gene were aligned and compared. The number of pairwise differences between each species pair were calculated (Table 5.1); ambiguities were treated conservatively (ie. assumed not to represent differences); no sequence contained more than one ambiguity. The unusual roseate tern haplotypes were identified as black-naped tern haplotypes, with the ND6 haplotype identical to that of the black-naped tern, and the ND2 haplotype differing at one nucleotide position. By comparison, the unusual roseate tern haplotypes differed from the ND2 and ND6 genes in crested and bridled terns by 7 - 8% and 14.6 - 14.9%, respectively. Partial sequencing of an additional two black-naped tern individuals at ND2 and one at ND6 confirmed that the unusual roseate tern haplotypes were black-naped tern haplotypes.

Table 5.1: Average uncorrected percentage sequence divergence of 993 base pairs of ND2 (above diagonal) and 352 base pairs of ND6 (below diagonal) genes between four species of tern commonly found on the Great Barrier Reef. RT = roseate tern, BNT = black-naped tern, BT = bridled tern, CT = crested tern. Common and unusual, respectively, refer to the most common roseate tern haplotype and the unusual haplotype obtained in individual RTS4

	RT (common)	RT (unusual)	BNT	BT	CT
RT (common)	-	4.3%	4.4%	14.5%	7.4%
RT (unusual)	6.3%	-	0.1%	14.6%	7%
BNT	6.3%	0%	-	14.4%	6.9%
BT	14.1%	14.9%	15.2%	-	12.7%
CT	6.8%	8%	8%	13.5%	-

Specific amplification of roseate and black-naped tern haplotypes from the same individual

Having identified the unusual haplotypes in RTS4 as black-naped tern haplotypes, new primers were designed, in locations with three or four fixed differences between the two species, to specifically amplify the black-naped tern haplotype or the roseate tern haplotype at each gene. For ND6, primers BNT ND6 (5'-ATGGGGTGGGGTTTGTTTC-3') and tRNA Pro (see section 2.2) were used to amplify approximately 400 bp of the black-naped tern haplotype, and primers Norm ND6 (5'-AGCGATGTTTTATTCGTGTGGTGTA-3') and tRNA Pro to amplify approximately 250 bp of the roseate tern haplotype. Primer L5216 (Sorenson *et al.* 1999) was used with H5503 (5'-AGATCTCACTGTCCTGTGTG-3') to amplify approximately 300 bp of the black-naped tern ND2 haplotype, and with H5889 (5'-TAAGGTAAGAAAAACGGCAG-3') to amplify approximately 700 bp of the roseate tern ND2 haplotype. Both RTS4 and a control individual (displaying only normal

roseate tern haplotypes) were amplified and sequenced. The PCR reaction conditions were the same as in Chapter 2, but with an annealing temperature of 60°C to ensure specific amplification of the target haplotype. Both strands of each PCR product were sequenced.

At ND2, sequencing of RTS4 produced the predicted haplotypes: roseate tern haplotype when targetting roseate tern haplotype, and black-naped tern haplotype when targetting black-naped tern (Table 5.2). Both amplifications of ND6 in RTS4 produced the same haplotype, that of the black-naped tern. The sequence obtained from the control individual, when using the ND6 primers specific for the black-naped tern haplotype, could not be identified by comparison with sequences available in GenBank.

Table 5.2: Haplotypes obtained from two roseate tern individuals by specific amplification targetting either the roseate tern or black-naped tern haplotype at two mitochondrial genes.

Haplotype	ND2		ND6	
specifically targetted	RTS4	control	RTS4	control
roseate	roseate	roseate	black-naped	roseate
black-naped	black-naped	roseate	black-naped	?

Microsatellite genotyping of RTS4

To determine whether the sample had been contaminated with the blood or DNA of another individual, RTS4 was genotyped at five microsatellite loci (for microsatellite primers and methods see section 4.2). As microsatellites are biparentally inherited markers (see section 1.1), the presence of more than two alleles at any locus would

indicate that the sample had been contaminated. The microsatellite primers were also used for amplification of three black-naped terns, to confirm that the black-naped tern alleles would have been detected if present.

No more than two alleles were observed at any of the five microsatellite loci in individual RTS4: RTS4 was homozygous at two loci and heterozygous at the remaining three loci. The microsatellite primers used for genotyping of roseate terns successfully amplified a product of similar size in the three black-naped terns. Contamination could thus be eliminated as a source of the black-naped tern haplotypes at the two mitochondrial genes.

5.3 Discussion

This study has demonstrated the co-existence of two distinct mitochondrial haplotypes in a single individual. Using specific primers, I was able to amplify a black-naped tern ND2 haplotype and a roseate tern ND2 haplotype from a single individual. The primer designed to specifically amplify the roseate tern ND6 haplotype consistently amplified the black-naped tern haplotype. This may be to the limited choice of regions in which to design the specific primer, resulting in the design of a suboptimal primer which preferentially amplifies the black-naped tern haplotype. Despite the failure of primer Norm ND6 to amplify the targetted roseate tern haplotype, the previous amplification of the ND6 haplotypes of the two species and the specific amplification of the ND2 haplotypes of the two species clearly demonstrate the presence of two mitochondrial haplotypes in individual RTS4.

This study is the first to find evidence of hybridisation between roseate and black-naped terns. In Australia, the two species are sympatric throughout the Northern Territory and Queensland, often nesting on the same islands, and occasionally forming mixed colonies (Smith & Buckley 1986; Walker 1989a; Smith 1991; Chatto 2001). On the Great Barrier Reef, roseate and black-naped terns also mix in non-breeding aggregations (O'Neill *et al.* in press). This mixing potentially allows mixed-species matings to occur at either breeding or non-breeding locations. The pair formation and courtship behaviour of the two species follows a very similar progression: males fly over the colony with fish to advertise availability (termed fish flights); both species undertake occasional courtship flights; and feeding of the female by the male occurs, first in the courtship and foraging areas, and later in the nesting area (Higgins & Davies 1996). Behaviour is therefore unlikely to prevent occasional pairing and interbreeding of the two species.

Pierotti (1987) surveyed a broad range of sympatric seabirds and found that hybridisation rarely occurs between species with different bill or foot colour. He proposed that as a result of the lack of variation in plumage and song, which act as isolating mechanisms in most groups of birds, the primary isolating mechanism in seabirds is the colour of the bill and/or feet. Breeding roseate terns have red bills and feet, while both the bills and feet of black-naped terns are black, seemingly providing an exception to Pierotti's rule (Figure 5.2). However, Donaldson (1968) observed that the red on the bill of roseate terns only appears once they have eggs, and Prendergast (1984) observed paired individuals, some with eggs, that did not yet have any red on the bill. The red colouration of the bill may, therefore, develop too late to act as an effective isolating mechanism between roseate and black-naped terns.

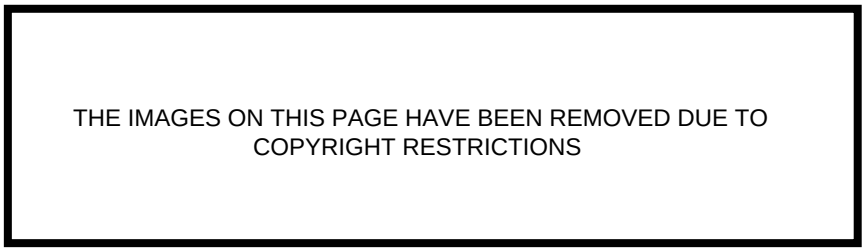


Figure 5.2: Pierotti (1987) proposed that the primary isolating mechanism in seabirds is bill or foot colour. While the roseate tern, *Sterna dougallii*, has a red bill and legs when in breeding plumage (top), in non-breeding plumage it has a black bill and legs (centre), as does the black-naped tern , *S. sumatrana*, year round (bottom). Occasional interbreeding of the two species may therfore occur prior to the development of breeding plumage in the roseate tern, bypassing the supposed isolating mechanism. Photos: Paul O'Neill and Kiyoaki Ozaki.

The hybridisation between the roseate and black-naped tern may have occurred in a breeding colony in which one of the species was present in low numbers. Tinbergen (1960) observed mixed species pairs of *Larus argentatus* and *L. fuscus* among *L. fuscus* colonies on several islands in Western Europe; in each instance there was only one mixed species pair, and it involved the only *L. argentatus* individual in the colony. Tinbergen (1960) hypothesised that hybridisation may be associated with pioneering species, with range expansion resulting in the colonisers occasionally finding themselves without conspecifics with which to mate. This may be the case for roseate and black-naped terns in Australia as, within their shared range, there is mitochondrial DNA evidence of recent range expansion in the roseate tern (Chapter 2). Gurr (1967) observed interbreeding between *L. novaehollandiae scopulinus* and *L. bulleri* in New Zealand. These two gull species breed sympatrically in only one colony on a regular basis, with *L. n. scopulinus* commencing breeding two to four weeks earlier than *L. bulleri*. The mixed species pairs reported by Gurr (1967) involved a *L. bulleri* individual paired with a *L. n. scopulinus* individual nesting within the *L. n. scopulinus* colony, on the edge of which six pairs of *L. bulleri* were nesting. Modifying Tinbergen's (1960) hypothesis to suit the two previous examples, hybridisation may thus occur when individuals are either spatially or temporally isolated from breeding conspecifics. This hypothesis fits observations of hybridisation in the roseate tern in North America, where roseate x common tern hybrids occur disproportionately often in small colonies in which roseate terns are the minority (Gochfeld *et al*. 1998).

Reports of mixed species colonies of roseate and black-naped terns in Australia include records where one of the species, usually the roseate tern, occurs in low numbers.

Between one and three pairs of roseate terns have been recorded breeding among tens to hundreds of black-naped tern pairs (Walker 1986; Walker 1989b; Chatto 2001). If a roseate tern at a predominantly black-naped tern colony was unable to find a conspecific with which to mate, for instance if there were extra male or female roseate terns associated with the colony, a roseate tern may pair with a black-naped tern and produce a hybrid offspring. As discussed previously, neither courtship behaviour nor colouration of feet and bill is likely to act as isolating mechanisms to this interspecific hybridisation.

In cattle and mice, interspecific hybridisation can allow sperm mitochondria to bypass the mechanism which, in an intraspecific mating, is designed to eliminate them, thus creating a heteroplasmic embryo (Kaneda *et al.* 1995; Sutovsky *et al.* 2000). Rokas *et al.* (2003) suggest that the degree of divergence of the two mitochondrial genomes is extremely important in determining whether paternal leakage can occur. The more divergent two species are, the greater the likelihood of paternal mtDNA bypassing the elimination mechanism. However, the more divergent two species are, the less likely they are to produce viable offspring. The window of divergence in which hybridisation can result in paternal leakage and transmission to subsequent generations is therefore narrow (Rokas *et al.* 2003), hence the low number of reports of paternal leakage. As there has not yet been a similar investigation into the biochemical processes of sperm recognition in a bird species, I can only presume that mechanisms operating in the roseate and black-naped tern parallel those described for cattle and mice.

In the case of RTS4, it would seem that the divergence between the roseate tern and the black-naped tern falls within this narrow window. I propose that either RTS4, or an ancestor of RTS4, was produced through a roseate tern x black-naped tern mating. This

hybridisation resulted in paternal leakage of mitochondrial DNA due to the mitochondrial genomes of the two species being sufficiently divergent to avoid the elimination of sperm mitochondria. The black-naped tern and roseate tern were sufficiently related to produce a hybrid offspring. Individual RTS4 could be the original hybrid offspring, in which case it could be either gender. Alternatively, RTS4 could be a descendant of the hybrid offspring (a backcross), in which case the original hybrid must have been a female, which subsequently bred with a male roseate tern, passing on her heteroplasmic mitochondrial genome.

The roseate tern and black-naped tern are 4.4% and 6.3% divergent at the ND2 and ND6 genes, respectively. The mitochondrial DNA of common and roseate terns shows the greatest divergence at the ND6 gene (J. Feinstein, pers. comm.). Assuming that ND6 is the most divergent mitochondrial region for roseate and black-naped terns, 6.3% is the maximum divergence for the mitochondrial genomes of the two species. At or below 6.3% divergence is thus sufficient to allow paternal leakage (by avoiding the elimination of sperm mitochondria), but still produce viable offspring. This is similar to the divergence of the heteroplasmic haplotypes of the great tit (6.24%; Kvist *et al.* 2003b). Paternal leakage has been observed with genomes as much as 20% divergent (Table 5.3). However, this level of divergence is associated with a unique mode of mitochondrial inheritance seen in mussels, whereby males are heteroplasmic and females are homoplasmic, and heteroplasmy is common (Skibinski *et al.* 1994). Excluding mussels due to their different inheritance mechanism, the roseate and black-naped tern show one of the higher divergence percentages of the taxa in which paternal leakage has been observed.

Table 5.3: Sequence divergence of heteroplasmic mitochondrial DNA in taxa in which paternal leakage has been reported

Taxon (reference)	heteroplasmic mtDNA sequence divergence
Human (Schwartz & Vissing 2002)	< 1%
Drosophila (Kondo *et al*. 1990)	2%
Anchovy (Magoulas & Zouros 1993)	2.82%
Bee (Meusel & Moritz 1993)	Not reported
Great tit (Kvist *et al*. 2003b)	6.24%
Tern (this study)	4.4 – 6.3%
Mouse (Gyllensten *et al*. 1991)	8.4% (Ferris *et al*. 1983)
Mussel (Skibinski *et al*. 1994)	Up to 20%+

This study provides the first evidence of hybridisation between the roseate and black-naped tern, and the first evidence for hybridisation involving either species in Australia. The common association of the two species, similar to that of the roseate and common tern in North America, suggests that the hybridisation reported in this study is unlikely to be an isolated event. In addition, this study has demonstrated that the two species are sufficiently related for paternal leakage of mitochondrial DNA to occur. I anticipate further evidence of hybridisation of the roseate and black-naped tern, potentially including paternal leakage, to be found in future genetic studies of the species.

Chapter 6: General discussion

In this thesis I have used a combination of mitochondrial DNA and microsatellite markers to examine population structure in the roseate tern at a number of geographic and temporal scales. In accordance with the majority of published studies to date, mitochondrial DNA has been of most value in investigating historic and large-scale relationships among populations, while microsatellites have proven useful for more recent, and finer-scale, relationships. Together, the two kinds of genetic markers have provided information on the relationships among roseate tern breeding colonies throughout the species range, and evidence for processes affecting these relationships in different regions.

This study has identified two distinct lineages of roseate tern, Atlantic and Indo-Pacific, using both mitochondrial DNA and microsatellite markers. The division between the two lineages concurs with Avise & Ball's (1990) definition of subspecies based on principles of genealogical concordance, which draws from both the biological (Mayr 1940) and phylogenetic species concepts (Cracraft 1983). Avise & Ball (1990) suggest that populations displaying historical, phylogenetic separation from one another for multiple, independent, genetically based traits, and which have an extrinsic reproductive barrier, rather than an intrinsic reproductive barrier, should be considered subspecies. This definition incorporates the historical separation of roseate tern lineages (see Chapter 3), while also recognising the likely ability for the two lineages to assimilate if the extrinsic reproductive barrier, continental Africa, was removed. The secondary contact and introgression of the two roseate tern lineages in South Africa illustrates the importance of differentiating between intrinsic and extrinsic reproductive barriers for

this definition. The genetic evidence obtained in this study suggests the need for a formal revision of roseate tern taxonomy, reducing the number of recognised subspecies from the current maximum of five to two: one in the Atlantic; and one in the Indo-Pacific.

In providing a genetic framework for identifying conservation priorities, Moritz (1994) proposed two criteria for a group of individuals to be considered an evolutionarily significant unit (ESU): 1) reciprocal monophyly of mitochondrial DNA sequences; and 2) significant divergence at nuclear loci. Mitochondrial DNA sequences identified two reciprocally monophyletic lineages of roseate tern, corresponding to the Atlantic and Indo-Pacific ocean basins (Chapter 3). Microsatellite analysis revealed strong divergence between individuals in the two mitochondrial lineages (Chapter 4), meeting Moritz's second criteria for designation as an ESU. Therefore the highest level of conservation priority in the roseate tern corresponds with the two lineages, at the geographic level of oceanic basins.

The divergence of roseate terns in the two oceanic basins highlights the role geographical barriers play in intraspecific population structuring. This is the first study to implicate continental Africa as the dominant long term barrier to gene flow in a seabird, although numerous studies have identified continental Africa as a barrier to gene flow in vagile fish species (Graves & McDowell 1995; Alvarado Bremer *et al.* 1998; Graves 1998; Scoles *et al.* 1998; Chow *et al.* 2000). Studies of temperate northern hemisphere seabirds have found evidence for restricted gene flow between the North Atlantic and the Pacific Ocean (Birt-Friesen *et al.* 1992; Friesen *et al.* 1996). However, only two previous studies of tropical seabirds have sampled across a range sufficient to

examine barriers to gene flow between oceanic basins in the southern hemisphere: Avise *et al*'s (2000) study of the sooty tern, which sampled from colonies in the Indian, Pacific and Atlantic Oceans; and the study of Steeves *et al*. (2003) investigating the diversification of the genus *Sula* (boobies), which sampled from the Atlantic and Pacific Oceans. In both studies, the Isthmus of Panama was implicated as a barrier to female-mediated gene flow, although in the sooty tern and masked booby, there was evidence of occasional gene flow across the Isthmus of Panama. The sampling range of the sooty tern study spanned both sides of Africa as well as the Isthmus of Panama therefore, though not mentioned, Africa is expected to have also served as a barrier to gene flow in the sooty tern.

In the roseate tern, Africa has served as a fixed barrier to gene flow between the two oceanic basins, with the strong divergence (3.9%) equating to a separation of up to one million years (see section 3.4). However, gene flow and population structure within each oceanic basin are likely to have been shaped by climatic changes (Figure 6.1). During glacial periods, roseate terns would have undergone range contraction to a common glacial refuge in each oceanic basin as a result of limited habitat availability. During this time, gene flow could have occurred among all individuals within an oceanic basin. During interglacial periods, suitable habitat would have been more widely available, allowing roseate terns to expand their range. A broadened distribution, coupled with some degree of philopatry, may have allowed genetically divergent colonies or regions to develop. During the next glacial period, however, range contraction to the common glacial refugia would have facilitated gene flow among any divergent groups, ameliorating evidence of past divergence. For the roseate tern, the current interglacial period has allowed the largest range expansion since the two

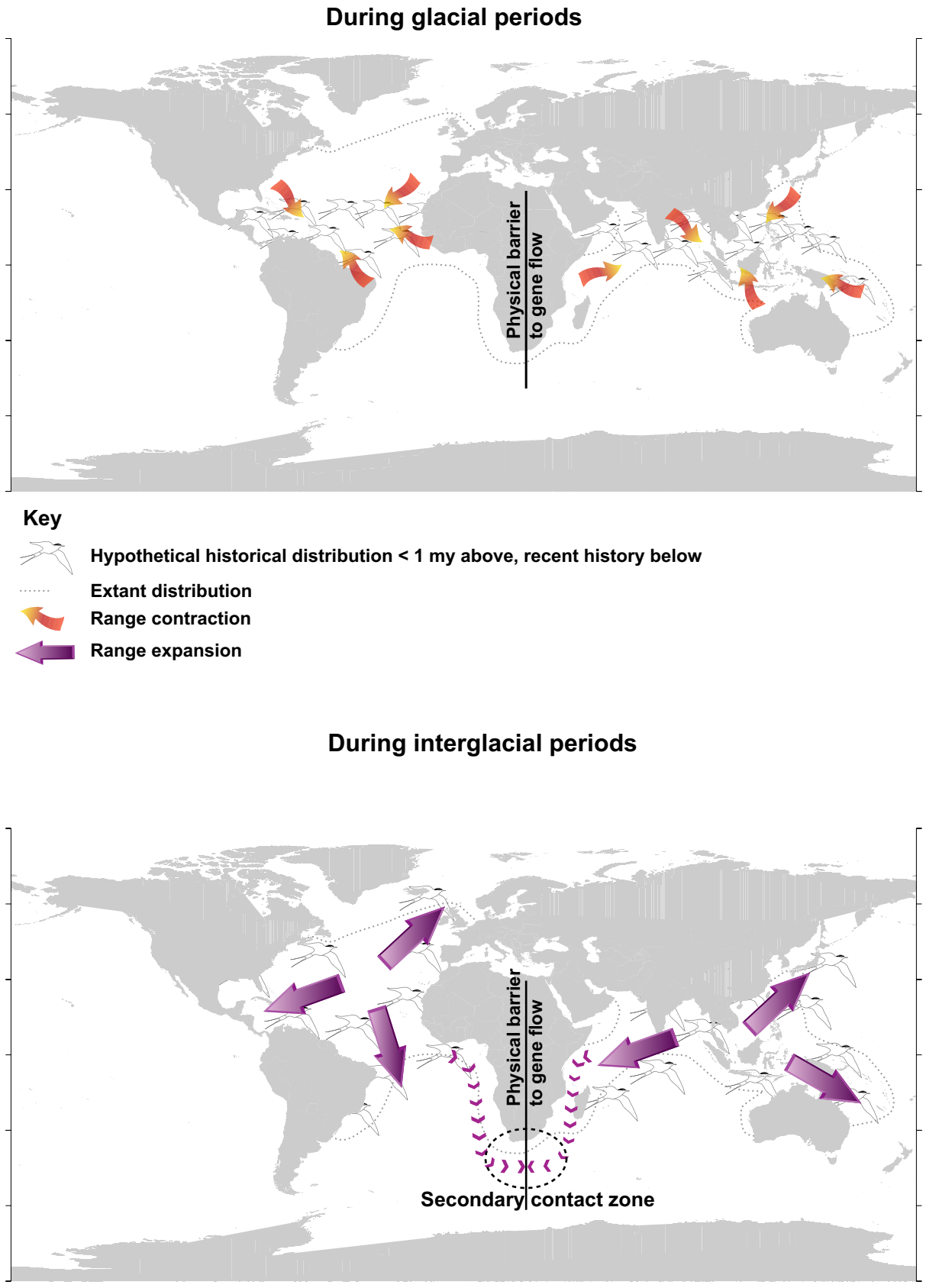


Figure 6.1: Model depicting processes affecting population structuring in the roseate tern, *Sterna dougallii*. The major long term barrier to gene flow is continental Africa, which has prevented gene flow between the Atlantic and Indo-Pacific ocean basins for up to one million years (see section 3.4), resulting in the two divergent lineages observed in this study. Within each oceanic basin, cycles of range contraction during glacial periods, and range expansion during interglacial periods have also contributed to the population structure seen today. During glacial periods, roseate terns are likely to have been confined to a single glacial refuge in either oceanic basin, allowing gene flow between all individuals within an oceanic basin. During interglacial periods, range expansion may have allowed geographically distant colonies to diverge genetically through philopatry or other barriers to gene flow. However, during the subsequent glacial period any genetic divergence would have been negated via gene flow among individuals in the glacial refuge. The current distribution of roseate terns is the broadest since the two lineages diverged, with secondary contact between individuals from each lineage occurring in South Africa.

lineages diverged, evidenced by the occurrence of both lineages in South Africa. If the ranges of the two lineages had overlapped previously, haplotypes of both lineages should be seen in both oceanic basins. Instead, the two lineages of roseate tern are interbreeding for the first time in as long as one million years. As a result of their separation, the two lineages show evidence of the different processes occurring in each oceanic basin. The Atlantic lineage appears to have reduced genetic diversity in comparison to the Indo-Pacific lineage, with fewer mtDNA haplotypes and microsatellite alleles, and lower nucleotide diversity observed in roseate terns in the Atlantic. This can be explained in part by smaller sample sizes for the Atlantic colonies, particularly for the microsatellite analyses. The Atlantic population is also smaller than the Indo-Pacific population, approximately 10 000 breeding pairs compared with over 80 000 breeding pairs in the Indo-Pacific, with smaller populations expected to contain less genetic diversity. There were also five instances of microsatellite allele fixation in Atlantic breeding colonies: at two loci each for the colonies in the United States and Ireland; and at one locus for the Azores. This reduced genetic diversity and fixation of alleles is likely to be a result of the population declines recorded in breeding colonies in the United States and Ireland in the last century (Witherby & Ticehurst 1908; Nisbet 1973; Cramp *et al*. 1974; Nisbet 1980). Roseate tern colonies in North America underwent an estimated 40-50% decline in the 1970s (Nisbet 1980), and those in Ireland a 40% decline in the 1960s (Cramp *et al*. 1974), which are likely to have caused the loss of rare alleles.

Ironically, the population declines in the nominate subspecies, *S. d. dougallii*, are the reason that the prioritising of conservation units would be most pertinent, and also the reason that identifying conservation units is difficult. Because of the population

declines, the data presented in this thesis give rise to two hypotheses to explain the unexpected results of the pairwise comparisons between colonies. If the non-significant relationship between Ireland and the United States is due to greater interchange of individuals on the non-breeding grounds, the Azores would constitute one management unit (MU; Moritz 1994), and North America and Europe another. Alternatively, if the non-significant relationship is the result of a concomitant loss of genetic diversity in North America and Europe, which would otherwise display allele frequencies similar to those observed in the Azores, the north Atlantic colonies would constitute a single MU. More band recaptures would help assess the first hypothesis, and ongoing work in this area should, in time, clarify the interactions of birds on the wintering grounds. To assess the second hypothesis would require genetic analysis of samples from North America and Europe obtained before the population declines to get an estimate of genetic diversity in colonies before the population declines.

While genetic data do not support the existence of three or four subspecies in the Indo-Pacific because of the lack of mitochondrial DNA divergence, microsatellites reveal that recent gene flow across the Indian Ocean has been restricted (Chapter 4). Since the most recent range expansion, the colony in the Seychelles has diverged from those in Australia and Japan. This contrasts with results from two studies in the Indian Ocean of marine species with high dispersal capabilities, with microsatellite analyses of shortfin mako, *Isurus oxyrinchus,* and bigeye tuna, *Thunnus obesus,* finding no evidence to suggest that gene flow across the Indian Ocean was restricted (Appleyard *et al.* 2002; Schrey & Heist 2003). With one exception, the only species in which genetic structure has been detected in the Indian Ocean have relatively restricted dispersal capabilities:

sea cucumber, *Holothuria nobilis*, (Uthicke & Benzie 2003), sea urchin, *Diadema sp.*, (Lessios *et al*. 2001) and giant tiger prawn, *Penaeus monodon*, (Benzie *et al*. 2002). The exception is the green turtle, *Chelonia mydas*, with mitochondrial analyses revealing divergence between nesting beaches in Oman and Australia, indicating strong female philopatry (Karl *et al*. 1992). Restriction fragment length polymorphism (RFLP) analyses of nuclear DNA, however, revealed significant divergence between Oman and Australia at only one of five loci, suggesting that gene flow in the green turtle is occurring through male-biased dispersal. Male-biased dispersal was invoked to explain a similar pattern in great white shark populations in South Africa, Australia and New Zealand, with mitochondrial DNA identifying two divergent groups which were not supported by microsatellite analyses (Pardini *et al*. 2001). Unlike the pattern of sex-biased dispersal observed in the green turtle and great white shark, microsatellite evidence for divergence of roseate tern colonies on either side of the Indian Ocean demonstrates that, since the last range expansion, neither sex is undergoing transoceanic dispersal. While microsatellites demonstrated the divergence of the colony in the Seychelles from those in Australia and Japan, colonies in the latter two regions did not show consistent evidence of divergence. The Indo-Pacific lineage therefore consists of two MUs under Moritz's (1994) criteria: a west Indian Ocean MU, represented by the Seychelles colony; and an Australasian MU, represented by Japan and Australia (Figure 6.2). The governments of Japan and Australia have agreed to list the roseate tern on the Japan-Australia Migratory Bird Agreement, in recognition of the ecological overlap of individuals from the two regions, providing a framework for conservation of the species beyond national borders.

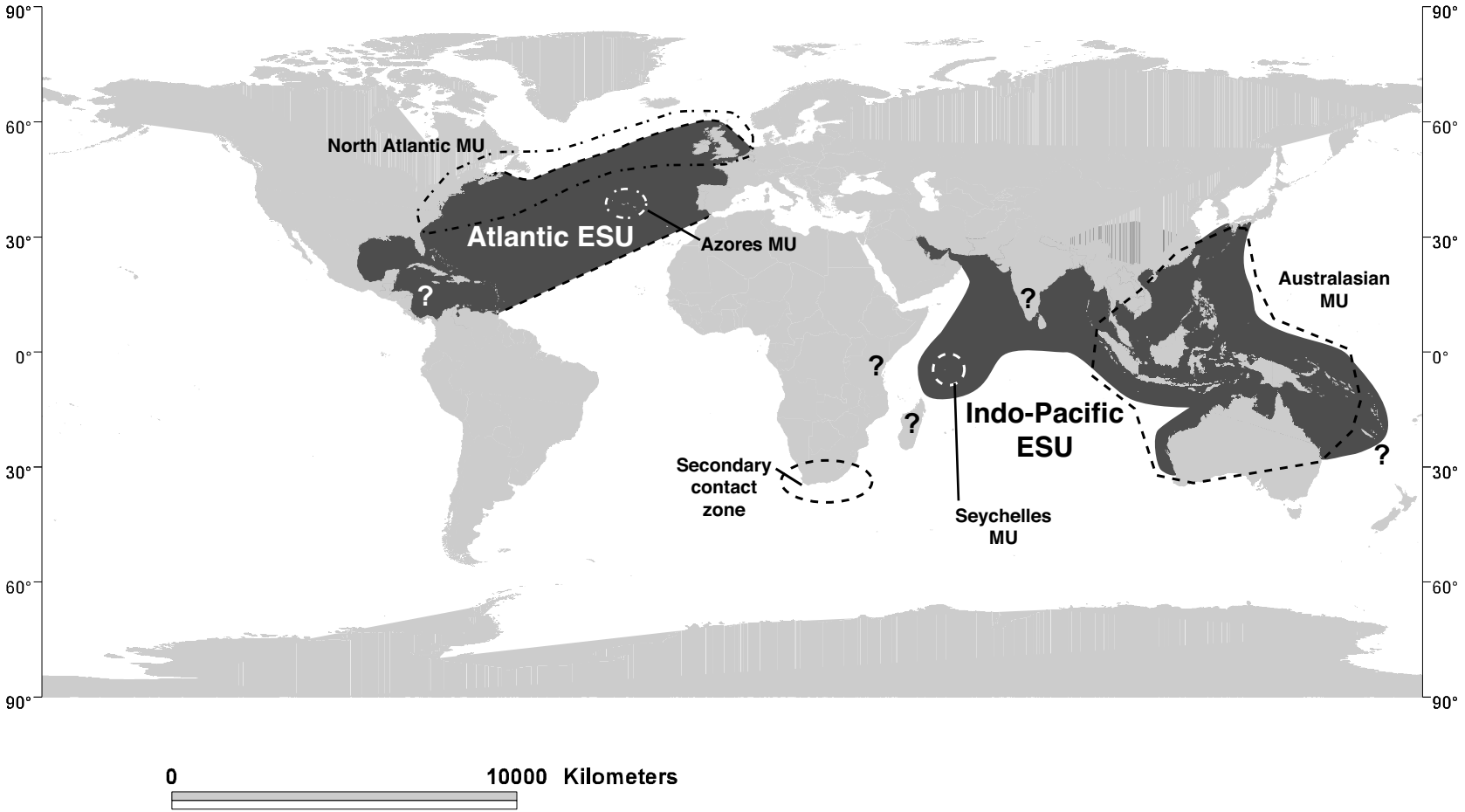


Figure 6.2: Genetically distinct breeding units in the roseate tern, *Sterna dougallii*, determined with ND2 and ND6 mitochondrial DNA sequences and microsatellite markers. Each mitochondrial lineage constitutes an evolutionarily significant unit (ESU; sensu Moritz 1994): an Atlantic ESU and an Indo-Pacific ESU (dark grey). Within the Indo-Pacific, there are two management units identified based on the data in this thesis, indicated by the areas enclosed by broken lines: an Australasian MU, combining individuals from Australia and Japan; and a Seychelles MU. In the Atlantic, the genetic data could not be used to unequivocally designate MUs; rather there were two possible designations. In one, the MU corresponded with the Atlantic ESU (broken line). In the other, the Azores constituted an MU, and the North American and Irish colonies combined to form a North Atlantic MU (enclosed by the combined dashed and dotted line). As the South African colony was a zone of secondary contact between the two ESUs it was not designated either ESU or MU status, but requires further investigation as a unique region. Question marks indicate regions for which data were unavailable for this study.

The lack of divergence among roseate tern breeding colonies in Australia and Japan indicates that philopatry is not as strong as anticipated. Breeding colonies in Australia and Japan are exchanging individuals often enough to counter the effects of genetic drift and prevent genetic divergence, both between the two regions and within Australia. Despite expectations of philopatry, numerous studies of seabirds have failed to find genetic evidence of population structure (Moen 1991; Moum *et al*. 1991; Birt-Friesen *et al*. 1992; Austin *et al*. 1994; Burg & Croxall 2001; Moum & Arnason 2001; Abbott & Double 2003a; Genovart *et al*. 2003;). Most of these studies were based on analyses of mitochondrial DNA, and hence may have encountered the same issue as this study, with recent colonisation events preventing mitochondrial DNA from providing insights into population structure. However, even those studies using microsatellites do not consistently reveal strong population structure, and therefore strong philopatry (see section 4.4).

The lack of strong philopatry in roseate terns in Australia may be a result of the frequent movement of colonies from one island to another (Hulsman 1984; Fuller *et al*. 1994; Milton *et al*. 1996; Surman 1998), in contrast to the high levels of philopatry and consistent locations of colonies observed in roseate terns in North America (Spendelow *et al*. 1995). Several studies of seabirds have found that individuals often move breeding colonies between years if reproductive success in their original colony was low and/or if the reproductive success in their subsequent colony was high in the previous year (Burger 1982; Danchin *et al*. 1998; Suryan & Irons 2001). This phenomenon has been termed 'performance-based conspecific attraction' by Danchin *et al*. (1998). In the northern Great Barrier Reef (GBR), mass mortalities of chicks have been reported at a number of breeding colonies; one example involved over 3 000 chicks dying of

starvation at an advanced stage of development (King *et al.* 1989; Milton *et al.* 1996). Milton *et al.* (1996) also reported low reproductive success in roseate terns breeding in the northern GBR, finding that the growth rate of roseate tern chicks was slow in comparison to black-naped tern chicks raised on the same island and roseate tern chicks in North America. If low reproductive success is common to roseate terns around Australia, it may induce the regular abandonment and relocation of breeding colonies. In the absence of group adherence (ie. individuals breeding together in a colony one year all moving to a new colony the following year), individuals that breed together one year may disperse to breed in multiple colonies the following year, facilitating regular gene flow between individuals in different colonies, and preventing the development of genetic structure among colonies or regions.

While this study has found evidence for the regular movement of individuals between breeding colonies in Australia, some colonies still carry signatures of their unique history. The breeding colony on Lancelin Island showed evidence of reduced genetic diversity for mitochondrial DNA and microsatellites; individuals from Lancelin Island were all homozygous at two of the five microsatellite loci (Chapter 4) and, of the ten individuals sampled, nine carried the same common roseate tern mitochondrial haplotype (Chapter 2). While the sample size for Lancelin Island was smaller than that of most other Australian breeding colonies and may have influenced this result, it seems no coincidence that Lancelin Island is one of the most recently colonised breeding islands of the roseate tern. Roseate terns have undergone a southward range extension in Western Australia in the past 50 years, with the colony on Lancelin Island one of the most recently initiated (Storr & Ford 1962; Dunlop 1979; Dunlop & Wooller 1986). The founder effect at Lancelin Island is likely to have caused the observed low level of

genetic diversity at both mitochondrial and nuclear loci through the bottleneck effect and ensuing genetic drift in the small founding population (Chakraborty & Nei 1977).

Future research

The lack of genetic differentiation among breeding colonies of the roseate tern in Australia and Japan precludes the use of genetic markers for identifying the origins of individuals in non-breeding aggregations in Australia. O'Neill *et al.* (in press) used moult to identify three different groups of roseate terns in the summer non-breeding aggregation in the Swain Reefs (see section 2.1); they identified one group as birds from Asia, and one as local Queensland birds, but could not identify an origin for the final group which constituted 38% of the non-breeding aggregation.

One potential method for identifying the origin of non-breeding roseate terns, which does not require genetically divergent colonies, is stable isotope analysis. Stable isotopes, such as carbon-13 ($\delta^{13}C$) and nitrogen-15 ($\delta^{15}N$), derived from food items, leave signatures in the tissue of an individual that are indicative of diet (Mizutani *et al.* 1990; Hobson & Clark 1992). Stable isotope signatures are often specific to a particular region, providing a method of linking individuals, via diet, with a geographic location (Alisauskas & Hobson 1993; Chamberlain *et al.* 1997; Minami & Ogi 1997; Cherel *et al.* 2000). Because feathers are metabolically inert once they have grown, they provide a measure of the stable isotopes of the region in which they were grown. Cherel *et al.* (2000) used $\delta^{13}C$ and $\delta^{15}N$ in feathers to investigate the wintering grounds of black-browed albatross, *Diomedea melanophrys*, breeding on the Kerguelen Islands. Black-browed albatross adults had significantly higher values for both $\delta^{13}C$ and $\delta^{15}N$ in comparison to chicks, indicating that adults feed in different areas during breeding and

non-breeding periods. Cherel *et al*. (2000) suggest that the higher value of both isotopes in adults is a result of their wintering north of the subtropical front, where $\delta^{13}C$ and $\delta^{15}N$ values are raised relative to waters south of the subtropical front. This correlated with band recovery data, which had indicated that black-browed albatross from the Kerguelen Islands winter in waters south of Australia. If different geographic regions in Australia have specific stable isotope signatures, it may enable finer scale identification of the breeding origins of roseate terns in non-breeding aggregations. This may allow identification of the 'unknown' component in the study of O'Neill *et al*. (in press), and the composition of the non-breeding aggregation in the Lacepede Islands.

While this study has clarified the intra-species relationships of *Sterna dougallii*, there is a need for a molecular phylogenetic analysis focussing on both generic and species-level relationships in the terns. Moynihan (1959) listed only three genera of terns, *Sterna, Anous* and *Larosterna,* based on plumage and behaviour; Sibley & Monroe (1990) describe seven genera based on DNA-DNA hybridisation work; and Gochfeld & Burger (1996) recognise a total of ten genera. As a result, different authors place the same species in different genera, depending on the taxonomy they accept. A phylogeny of the group based on direct sequencing of mitochondrial DNA would assist in designating genera and the relationships of the species therein, providing a greater understanding of inter-species relationships and providing uniformity in nomenclature for future studies. The primers used for DNA sequencing of the mitochondrial ND2 and ND6 genes in this study (see section 2.1) have been used successfully in a number of tern species (see section 5.2), and hence should prove useful for phylogenetic analyses.

This study is the first to find evidence of introgression resulting from inter-species hybridisation in terns (Chapter 5). Crochet *et al.* (2002) found evidence for introgression between several species of gull, and Cohen *et al.* (1997) proposed introgression as an explanation for the unusual phylogeny of the skuas (see section 5.1). Given the ability of several tern species pairs to interbreed (see section 5.1), a phylogeny of the terns may reveal historical instances of introgression, as in the gulls and skuas. Any phylogenetic evidence for introgression between species could be combined with information on behaviour, morphology and ecology, to potentially identify traits that act as reproductive isolating mechanisms in terns.

The area of secondary contact between the two lineages of roseate tern provides a unique opportunity for studying introgression. Most instances of secondary contact occur in species in which individuals are continuously distributed either side of the secondary contact zone (Chow *et al.* 2000; Tarkhnishvili *et al.* 2001; Kvist *et al.* 2003a) or which show evidence of introgression in multiple discrete populations (Chenoweth *et al.* 1998; Gilles *et al.* 1998; Esa *et al.* 2000; Lu *et al.* 2001). In the roseate tern, however, the secondary contact has been identified within a single breeding colony (the only currently active breeding colony) in South Africa (see Chapter 3), with the nearest breeding colony likely to be in Madagascar, over 2 000 km away. Sampling of the colonies in Madagascar would be valuable to determine whether the area of secondary contact extends beyond South Africa. Meanwhile, the secondary contact between the two lineages of roseate tern can be studied in a discrete geographic area. An increased sample size would allow testing of whether each lineage is represented by a similar number of individuals, as is suggested in this study from the six samples from South Africa. Ongoing genetic analyses could reveal whether introgression of the lineages was

directional, resulting in one lineage predominating in the colony, or whether the two lineages exist in equilibrium.

This study has provided strong genetic evidence for the existence of two lineages of roseate terns, between which gene flow has been limited, both historically and currently, by continental Africa. I have proposed that each of the two roseate tern lineages warrant status as an evolutionarily significant unit. Crandall *et al.* (2000) argue that a combination of ecological and genetic data are required to assign status as an evolutionarily significant unit. While I recognise the validity of Crandall *et al.*'s (2000) argument, the genetic data in this study clearly demonstrate the existence of two groups of roseate terns with different evolutionary histories, and thus different evolutionary potential. Unfortunately, the only ecological data available across the range of the roseate tern is for the morphological traits of wing length, bill length and bill colouration, and the variation of these traits in Atlantic roseate terns encompasses the variation seen in Indo-Pacific roseate terns. These morphological traits are unlikely to act as measures of adaptive differentiation, and the behavioural characteristics which may reveal adaptive differentiation are poorly known beyond the colonies in North America. For the roseate tern, ecological studies will be of most value where genetic data have been unable to provide answers. For example, more band-recapture data are required to understand the interactions of birds on non-breeding grounds, in both the Atlantic and in Australia. There is an ongoing need for ecological studies across the breadth of the species range to complement and allowing testing of the findings of this thesis.

References

Abbott, C. L. & Double, M. C. (2003a). Genetic structure, conservation genetics and evidence of speciation by range expansion in shy and white-capped albatrosses. *Molecular Ecology* **12**, 2953-2962.

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

Alisauskas, R. T. & Hobson, K. A. (1993). Determination of lesser snow goose diets and winter distribution using stable isotope analysis. *Journal of Wildlife Management* **57**, 49-54.

Allende, L. M., Rubio, I., Ruiz-del-Valle, V., Guillen, J., Martinez-Laso, J., Lowy, E., Varela, P., Zamora, J. & Arnaiz-Villena, A. (2001). The old world sparrows (Genus Passer) phylogeography and their relative abundance of nuclear mtDNA pseudogenes. *Journal of Molecular Evolution* **53**, 144-154.

Alvarado Bremer, J. R., Stequert, B., Robertson, N. W. & Ely, B. (1998). Genetic evidence for inter-oceanic subdivision of bigeye tuna (*Thunnus obesus*) populations. *Marine Biology* **132**, 547-557.

Andersen, L. W., Born, E. W., Gjertz, I., Wiig, O., Holm, L.-E. & Bendixen, C. (1998). Population structure and gene flow of the Atlantic walrus (*Odobenus rosmarus rosmarus*) in the eastern Atlantic Arctic based on mitochondrial DNA and microsatellite variation. *Molecular Ecology* **7**, 1323-1336.

Andrle, R. F. (1972). Another probable hybrid of *Larus marinus* and *L. argentatus*. *Auk* **89**, 669-671.

Andrle, R. F. (1973). A second probable hybrid of *Larus marinus* and *L. argentatus* on the Niagara River. *Canadian Field Naturalist* **87**, 170-171.

Andrle, R. F. (1980). Three more probable hybrids of *Larus hyperboreus* and *L. argentatus*. *Wilson Bulletin* **92**, 389-393.

Appleyard, S. A., Ward, R. D. & Grewe, P. M. (2002). Genetic stock structure of bigeye tuna in the Indian Ocean using mitochondrial DNA and microsatellites. *Journal of Fish Biology* **60**, 767-770.

Arctander, P. (1995). Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proceedings of the Royal Society of London Series B* **262**, 13-19.

Aris-Brosou, S. & Excoffier, L. (1996). The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology and Evolution* **13**, 494-504.

Austin, J. J., White, R. W. G. & Ovenden, J. R. (1994). Population-genetic structure of a philopatric, colonially nesting seabird, the short-tailed shearwater (*Puffinus tenuirostris*). *Auk* **111**, 70-79.

Avise, J. C. & Ball, R. M. J. (1990). Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology* **7**, 45-67.

Avise, J. C., Nelson, W. S., Bowen, B. W. & Walker, D. (2000). Phylogeography of colonially nesting seabirds, with special reference to global matrilineal patterns in the sooty tern (*Sterna fuscata*). *Molecular Ecology* **9**, 1783-1792.

Baker, C. S., Patenaude, N. J., Bannister, J. L., Robins, J. & Kato, H. (1999). Distribution and diversity of mtDNA lineages among southern right whales (*Eubalaena australis*) from Australia and New Zealand. *Marine Biology* **134**, 1-7.

Bell, D. A. (1996). Genetic differentiation, geographic variation and hybridization in gulls of the *Larus glaucescens - occidentalis* complex. *Condor* **98**, 527-546.

Benzie, J. A. H., Ballment, E., Forbes, A. T., Demetriades, N. T., Sugama, K., Haryanti & Moria, S. (2002). Mitochondrial DNA variation in Indo-Pacific populations of the giant tiger prawn, *Penaeus monodon*. *Molecular Ecology* **11**, 2553-2569.

Bérubé, M., Aguilar, A., Dendanto, D., Larsen, F., Notarbartolo di Sciara, G., Sears, R., Sigurjónsson, J., Urban-R., J. & Palsbøll, P. J. (1998). Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus 1758): analysis of mitochondrial and nuclear loci. *Molecular Ecology* **7**, 585-599.

Birky, C. W. J. (1995). Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the USA* **92**, 11331-11338.

Birt-Friesen, V. L., Montevecchi, W. A., Gaston, A. J. & Davidson, W. S. (1992). Genetic structure of thick-billed murre (*Uria lomvia*) populations examined using direct sequence analysis of amplified DNA. *Evolution* **46**, 267-272.

Blaber, S. J. M., Milton, D. A., Farmer, M. J. & Smith, G. C. (1998). Seabird breeding populations on the far northern Great Barrier Reef, Australia: trends and influences. *Emu* **98**, 44-57.

Bohonak, A. J. (2002). IBD (Isolation By Distance): a program for analyses of isolation by distance. *Journal of Heredity* **93**, 153-154.

Bolten, A. B., Bjorndal, K. A., Martins, H. R., Dellinger, T., Biscoito, M. J., Encalada, S. E. & Bowen, B. W. (1998). Transatlantic developmental migrations of loggerhead sea turtles demonstrated by mtDNA sequence analysis. *Ecological Applications* **8**, 1-7.

Bouzat, J. L., Lewin, H. A. & Paige, K. N. (1998). The ghost of genetic diversity past: historical DNA analysis of the greater prairie chicken. *American Naturalist* **152**, 1-6.

Bowen, B. W. (1998). What is wrong with ESUs?: the gap between evolutionary theory and conservation principles. *Journal of Shellfish Research* **17**, 1355-1358.

Bowen, B. W., Bass, A. L., Garcia-Rodriguez, A., Diez, C. E., Vandam, R., Bolten, A., Bjorndal, K. A., Miyamoto, M. M. & Ferl, R. J. (1996). Origin of hawksbill turtles in a Caribbean feeding area as indicated by genetic markers. *Ecological Applications* **6**, 566-572.

Bowen, B. W., Meylan, A. B., Ross, J. P., Limpus, C. J., Balazs, G. H. & Avise, J. C. (1992). Global population structure and natural history of the green turtle (*Chelonia mydas*) in terms of matriarchal phylogeny. *Evolution* **46**, 865-881.

Brichetti, P. & Foschi, U. F. (1987). The lesser crested tern in the Western Mediterranean and Europe. *British Birds* **80**, 276-280.

Bromham, L., Eyre-Walker, A., Smith, N. H. & Smith, J. M. (2003). Mitochondrial Steve: paternal inheritance of mitochondria in humans. *Trends in Ecology and Evolution* **18**, 2-4.

Brooke, R. K. (1984). South African Red Data Book - Birds. *South African National Scientific Programmes* **97**, 1-213.

Brown Gladden, J. G., Ferguson, M. M. & Clayton, J. W. (1997). Matriarchal genetic population structure of North American beluga whales *Delphinapterus leucas* (Cetacea: Monodontidae). *Molecular Ecology* **6**, 1033-1046.

Brown Gladden, J. G., Ferguson, M. M., Friesen, M. K. & Clayton, J. W. (1999). Population structure of North American beluga whales (*Delphinapterus leucas*) based on nuclear DNA microsatellite variation and contrasted with the

population structure revealed by mitochondrial DNA variation. *Molecular Ecology* **8**, 347-363.

Brown, W. M., George, M. J. & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA* **76**, 1967-1971.

Buonaccorsi, V. P., McDowell, J. R. & Graves, J. E. (2001). Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). *Molecular Ecology* **10**, 1179-1196.

Burbidge, A. A. & Fuller, P. J. (1996). The Western Australian Department of Conservation and Land Management Seabird Breeding Island Database. In *The status of Australia's seabirds: Proceedings of the National Seabird Workshop, Canberra, 1-2 November 1993* (Ross, G., Weaver, K. & Greig, J., eds.), pp. 73-137. Biodiversity Group, Environment Australia, Canberra.

Burg, T. M. & Croxall, J. P. (2001). Global relationships amongst black-browed and grey-headed albatrosses: analysis of population structure using mitochondrial DNA and microsatellites. *Molecular Ecology* **10**, 2647-2660.

Burger, J. (1982). The role of reproductive success in colony-site selection and abandonment in black skimmers (*Rynchops niger*). *Auk* **99**, 109-115.

Calabrese, P. & Durrett, R. (2003). Dinucleotide repeats in the drosophila and human genomes have complex, length-dependent mutation processes. *Molecular Biology and Evolution* **20**, 715-725.

Cann, R. L., Brown, W. M. & Wilson, A. C. (1984). Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* **106**, 479-499.

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

Chamberlain, C. P., Blum, J. D., Holmes, R. T., Feng, X., Sherry, T. W. & Graves, G. R. (1997). The use of isotope tracers for identifying populations of migratory birds. *Oecologia* **109**, 132-141.

Chatto, R. (2001). The distribution and status of colonial breeding seabirds in the Northern Territory. Parks and Wildlife Commission of the Northern Territory Technical Report No. 70.

Chenoweth, S. F. & Hughes, J. M. (2003). Oceanic interchange and nonequilibrium population structure in the estuarine dependent Indo-Pacific tasselfish, *Polynemus sheridani*. *Molecular Ecology* **12**, 2387-2397.

Chenoweth, S. F., Hughes, J. M., Keenan, C. P. & Lavery, S. (1998). When oceans meet: a teleost shows secondary intergradation at an Indian-Pacific interface. *Proceedings of the Royal Society of London Series B* **265**, 415-420.

Cherel, Y., Hobson, K. A. & Weimerskirch, H. (2000). Using stable-isotope analysis of feathers to distinguish moulting and breeding origins of seabirds. *Oecologia* **122**, 155-162.

Chow, S., Okamoto, H., Miyabe, N., Hiramatsu, K. & Barut, N. (2000). Genetic divergence between Atlantic and Indo-Pacific stocks of bigeye tuna (*Thunnus obesus*) and admixture around South Africa. *Molecular Ecology* **9**, 221-227.

Clegg, S. M., Kelly, J. F., Kimura, M. & Smith, T. B. (2003). Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant, Wilson's warbler (*Wilsonia pusilla*). *Molecular Ecology* **12**, 819-830.

Clement, M., Posada, D. & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1660.

Cohen, B. L., Baker, A. J., Bleschschmidt, K., Dittmann, D. L., Furness, R. W., Gerwin, J. A., Helbig, A. J., Dekorte, J., Marshall, H. D., Palma, R. L., Peter, H. U., Ramli, R., Siebold, I., Willcox, M. S., Wilson, R. H. & Zink, R. M. (1997). Enigmatic phylogeny of skuas (Aves: Stercorariidae). *Proceedings of the Royal Society of London Series B* **264**, 181-190.

Cracraft, J. (1983). Species concepts and speciation analysis. In *Current Ornithology* (Johnston, R. F., ed.), Vol. 1, pp. 159-187. Plenum Press, London.

Cramp, S., Ed. (1985). Handbook of the Birds of Europe, the Middle East and North Africa. Volume 4: Terns to Woodpeckers. Oxford University Press, New York.

Cramp, S., Bourne, W. R. P. & Saunders, D. (1974). *The seabirds of Britain and Ireland*, Collins, London.

Crandall, K. A., Bininda-Edmonds, O. R. P., Mace, G. M. & Wayne, R. K. (2000). Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution* **15**, 290-295.

Crochet, P.-A. (1996). Can measures of gene flow help to evaluate bird dispersal? *Acta Oecoloogica* **17**, 459-474.

Crochet, P.-A., Bonhomme, F. & Lebreton, J.-D. (2000). Molecular phylogeny and plumage evolution in gulls (Larini). *Journal of Evolutionary Biology* **13**, 47-57.

Crochet, P.-A., Lebreton, J. D. & Bonhomme, F. (2002). Systematics of large white-headed gulls: Patterns of mitochondrial DNA variation in western European taxa. *Auk* **119**, 603-620.

Danchin, E., Boulinier, T. & Massot, M. (1998). Conspecific reproductive success and breeding habitat selection: implications for the study of coloniality. *Ecology* **79**, 2415-2428.

Davis, A. H. (1982). Terns showing mixed characters of black and white-winged black terns. *British Birds* **75**, 579-580.

Department of Foreign Affairs. (1995). Agreement between the Government of Australia and the Government of Japan for the protection of migratory birds in danger of extinction and their environment. Department of Foreign Affairs, Canberra.

Desjardins, P. & Morais, R. (1990). Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. *Journal of Molecular Biology* **212**, 599-634.

Donaldson, G. (1968). Bill color changes in adult roseate terns. *Auk* **85**, 662-668.

Dunlop, J. N. (1979). The occurrence of breeding roseate terns at Lancelin Island, Western Australia. *Western Australian Naturalist* **14**, 118-119.

Dunlop, J. N. & Wooller, R. D. (1986). Range extensions and the breeding seasons of seabirds in south-western Australia. *Records of the Western Australian Museum* **12**, 389-394.

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

Esa, Y. B., Waters, J. M. & Wallis, G. P. (2000). Introgressive hybridization between *Galaxias depressiceps* and *Galaxias sp D* (Teleostei: Galaxiidae) in Otago, New Zealand: secondary contact mediated by water races. *Conservation Genetics* **1**, 329-339.

Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Eyre-Walker, A., Smith, N. H. & Maynard Smith, J. (1999a). How clonal are human mitochondria? *Proceedings of the Royal Society of London Series B* **266**, 477-483.

Eyre-Walker, A., Smith, N. H. & Maynard Smith, J. (1999b). Reply to Macaulay *et al.* (1999): mitochondrial DNA recombination - reasons to panic. *Proceedings of the Royal Society of London Series B* **266**, 2041-2042.

Eyre-Walker, A. & Awadalla, P. (2001). Does human mtDNA recombine? *Journal of Molecular Evolution* **53**, 430-435.

Farris, J. S., Kallerjso, M., Kluge, A. G. & Bult, C. (1995). Constructing a significance test for incongruence. *Systematic Biology* **44**, 570-572.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783-791.

Ferris, S. D., Sage, R. D., Prager, E. M., Ritte, U. & Wilson, A. C. (1983). Mitochondrial DNA evolution in mice. *Genetics* **105**, 681-721.

FitzSimmons, N. N., Moritz, C., Limpus, C. J., Pope, L. & Prince, R. (1997). Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. *Genetics* **147**, 1843-1854.

Flood, P. G., Harjanto, S. & Orme, G. R. (1979). Carbon-14 dates, Lady Elliot Reef, Great Barrier Reef. *Queensland Government Mining Journal*, 444-447.

Fraser, D. J. & Bernatchez, L. (2001). Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology* **10**, 2741-2752.

Friesen, V. L., Montevecchi, W. A., Baker, A. J., Barrett, R. T. & Davidson, W. S. (1996). Population differentiation and evolution in the common guillemot *Uria aalge*. *Molecular Ecology* **5**, 793-805.

Friesen, V. L. (1997). Population genetics and the spatial scale of conservation of colonial waterbirds. *Colonial Waterbirds* **20**, 353-368.

Fuller, P. J., Burbidge, A. A. & Owens, R. (1994). Breeding seabirds of the Houtman Abrolhos, Western Australia: 1991-1993. *Corella* **18**, 97-113.

Garza, J. C., Slatkin, M. & Freimer, N. B. (1995). Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Molecular Biology and Evolution* **12**, 594-603.

Genovart, M., Oro, D. & Bonhomme, F. (2003). Genetic and morphological differentiation between the two largest breeding colonies of Audouin's gull *Larus audouinii*. *Ibis* **145**, 448-456.

Giles, R. E., Blanc, H., Cann, H. M. & Wallace, D. C. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA* **77**, 6715-6719.

Gilles, A., Chappaz, R., Cavalli, L., Lortscher, M. & Faure, E. (1998). Genetic differentiation and introgression between putative subspecies of *Leuciscus soufia* (Teleostei: Cyprinidae) of the region of the Mediterranean Alps. *Canadian Journal of Fisheries & Aquatic Sciences* **55**, 2341-2354.

Given, A. D., Mills, J. A. & Baker, A. J. (2002). Isolation of polymorphic microsatellite loci from the red-billed gull (*Larus novaehollandiae scopulinus*) and amplification in related species. *Molecular Ecology Notes* **2**, 416-418.

Glenn, T. C., Thompson, J. E., Ballard, B. M., Roberson, J. A. & French, J. O. (2002). Mitochondrial DNA variation among wintering midcontinent gulf coast sandhill cranes. *Journal of Wildlife Management* **66**, 339-348.

Gochfeld, M. (1983). The roseate tern: world distribution and status of a threatened species. *Biological Conservation* **25**, 103-125.

Gochfeld, M. & Burger, J. (1996). Family Sternidae (Terns). In *Handbook of birds of the world,* Vol. 3 (del Hoyo, J., Elliott, A. & Sargatal, J., eds.). Lynx Edicions, Barcelona.

Gochfeld, M., Burger, J. & Nisbet, I. C. T. (1998). Roseate tern (*Sterna dougallii*). In *The birds of North America, No. 370* (Poole, A. & Gill, F., eds.). The Birds of North America, Inc., Philadelphia, PA.

Godfrey, W. E. (1973). More presumed hybrid gulls: *Larus argentatus* x *L. marinus*. *Canadian Field Naturalist* **87**, 171-172.

Grant, P. R. & Grant, B. R. (1992). Hybridization of bird species. *Science* **256**, 193-197.

Graves, J. E. (1998). Molecular insights into the population structures of cosmopolitan marine fishes. *Journal of Heredity* **89**, 427-437.

Graves, J. E. & McDowell, J. R. (1995). Inter-ocean divergence of istiophorid billfishes. *Marine Biology* **122**, 193-203.

Grewe, P. M., Billington, N. & Hebert, P. D. N. (1990). Phylogenetic relationships among members of *Salvelinus* inferred from mitochondrial DNA divergence. *Canadian Journal of Fisheries and Aquatic Science* **47**, 984-991.

Gurr, L. (1967). Interbreeding of *Larus novaehollandiae scopulinus* and *Larus bulleri* in the wild in New Zealand. *Ibis* **109**, 552-555.

Gyllensten, U., Wharton, D., Josefsson, A. & Wilson, A. C. (1991). Paternal inheritance of mitochondrial DNA in mice. *Nature* **352**, 255-257.

Haig, S. M., Gratto-Trevor, C. L., Mullins, T. D. & Colwell, M. A. (1997). Population identification of western hemisphere shorebirds throughout the annual cycle. *Molecular Ecology* **6**, 413-427.

Hamilton, J. (1981). Recoveries of wintering roseate terns. *Journal of Field Ornithology* **52**, 36-42.

Hancock, J. M. (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms. In *Microsatellites: evolution and applications* (Goldstein, D. B. & Schlotterer, C., eds.), pp. 1-9. Oxford University Press, New York.

Hasegawa, M., Kishino, H. & Yano, T.-a. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**, 160-174.

Hays, H. (1975). Probable common x roseate tern hybrids. *Auk* **92**, 219-234.

Hays, H., DiCostanzo, J., Cormons, G., de Tarso Zuquim Antas, P., do Nascimento, J., de Lima Serrano do Nascimento, I. & Bremer, R. E. (1997). Recoveries of roseate and common terns in South America. *Journal of Field Ornithology* **68**, 79-90.

Hays, H., Neves, V. & Lima, P. (2002a). Banded roseate terns from different continents trapped in the Azores. *Journal of Field Ornithology* **73**, 180-184.

Hays, H., Newton, S., Lima, P. & Crowe, O. (2000). Rockabill roseate tern recaptured in Brazil. *Irish Birds* **6**, 585-586.

Hays, H., Newton, S. F. & Cormons, G. (2002b). Rockabill roseate terns *Sterna dougallii* sighted in west Atlantic colony. *Irish Birds* **7**, 133-134.

Hays, H., Spendelow, J. A., Burger, J., Pierce, J. & Gochfeld, M. (1999). A non-breeding concentration of *Sterna dougallii* and *S. hirundo* in Bahia, Brazil. *Journal of Field Ornithology* **70**, 456-464.

Heatwole, H., O'Neill, P., Jones, M. & Preker, M. (1996). *Long-term population trends on the Swains Reefs, Great Barrier Reef, Technical Report No. 12*, CRC Reef Research Centre, Townsville.

Higgins, P. J. & Davies, S. J. J. F. E., Eds. (1996). Handbook of Australian, New Zealand and Antarctic Birds. Volume 3: Snipe to Pigeons. Oxford University Press, Melbourne.

Hobson, K. A. & Clark, R. G. (1992). Assessing avian diets using stable isotopes I: turnover of ^{13}C in tissues. *Condor* **94**, 181-188.

Hoffman, W., Wiens, J. A. & Michael Scott, J. (1978). Hybridization between gulls (*Larus glaucescens* and *L. occidentalis*) in the Pacific Northwest. *Auk* **95**, 441-458.

Huang, Q.-Y., Xu, F.-H., Shen, H., Deng, H.-Y., Liu, Y.-J., Liu, Y.-Z., Li, J.-L., Recker, R. R. & Deng, H.-W. (2002). Mutation patterns as dinucleotide microsatellite loci in humans. *American Journal of Human Genetics* **70**, 625-634.

Hulsman, K. (1976). The robbing behaviour of terns and gulls. *Emu* **76**, 143-149.

Hulsman, K. (1977a). Breeding success and mortality of terns at One Tree Island, Great Barrier Reef. *Emu* **77**, 49-60.

Hulsman, K. (1977b). Daily attendance of terns and gulls at One Tree Island. *Sunbird* **8**, 9-19.

Hulsman, K. (1984). Seabirds of the Capricornia section, Great Barrier Reef Marine Park. *Royal Society of Queensland Symposium*, 53-60.

Hutchinson, W. F., van Oosterhout, C. v., Rogers, S. I. & Carvalho, G. R. (2003). Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*). *Proceedings of the Royal Society of London Series B* **270**, 2125-2132.

Hutchison III, C. A., Newbold, J. E., Steven Potter, S. & Edgell, M. H. (1974). Maternal inheritance of mammalian mitochondrial DNA. *Nature* **251**, 536-538.

Ingolfsson, A. (1970). Hybridization of glaucous gulls *Larus hyperboreus* and herring gulls *L. argentatus* in Iceland. *Ibis* **112**, 340-362.

Johnson, J. A., Toepfer, J. E. & Dunn, P. O. (2003). Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. *Molecular Ecology* **12**, 3335-3347.

Johnson, K. P., Goodman, S. M. & Lanyon, S. M. (2000). A phylogenetic study of the Malagasy Couas with insights into cuckoo relationships. *Molecular Phylogenetics and Evolution* **14**, 436-444.

Johnson, K. P. & Sorenson, M. D. (1998). Comparing molecular evolution in two mitochondrial protein coding genes (cytochrome *b* and ND2) in the dabbling ducks (Tribe: Anatini). *Molecular Phylogenetics and Evolution* **10**, 82-94.

Kaneda, H., Hayashi, J.-I., Takahama, S., Taya, C., Lindahl, K. F. & Yonekawa, H. (1995). Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. *Proceedings of the National Academy of Sciences of the USA*. **92**, 4542-4546.

Karl, S. A., Bowen, B. W. & Avise, J. C. (1992). Global population genetic structure and male-mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analyses of anonymous nuclear loci. *Genetics* **131**, 163-173.

Kidd, M. G. & Friesen, V. L. (1998). Sequence variation in the guillemot (Alcidae: *Cepphus*) mitochondrial control region and its nuclear homolog. *Molecular Biology and Evolution* **15**, 61-70.

Kimura, M., Clegg, S. M., Lovette, I. J., Holder, K. R., Girman, D. J., Mila, B., Wade, P. & Smith, T. B. (2002). Phylogeographical approaches to assessing demographic connectivity between breeding and overwintering regions in a Neartic-Neotropical warbler (*Wilsonia pusilla*). *Molecular Ecology* **11**, 1605-1616.

King, B. R. (1993). The status of Queensland seabirds. *Corella* **17**, 65-93.

King, B. R., Hicks, J. & Oldroyd, A. (1989). Seabird Islands No. 190: Cholmondeley Island, Great Barrier Reef, Queensland. *Corella* **13**, 41-42.

King, B. R. & Limpus, C. J. (1989). Seabird Islands No. 191: Wallace Island, Great Barrier Reef, Queensland. *Corella* **13**, 43-44.

Kirkham, I. R. & Nettleship, D. N. (1987). Status of the roseate tern in Canada. *Journal of Field Ornithology* **58**, 505-515.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X. & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the USA* **86**, 6196-6200.

Koenig, W. D., van Vuren, D. & Hooge, P. H. (1996). Detectability, philopatry, and the distribution of dispersal distances in vertebrates. *Trends in Ecology and Evolution* **11**, 514-517.

Kondo, R., Satta, Y., Matsuura, E. T., Ishiwa, H., Takahata, N. & Chigusa, S. I. (1990). Incomplete maternal transmission of mitochondrial DNA in drosophila. *Genetics* **126**, 657-663.

Kvist, L., Martens, J., Higuchi, H., Nazarenko, A. A., Valchuk, O. P. & Orell, M. (2003a). Evolution and genetic structure of the great tit (*Parus major*) complex. *Proceedings of the Royal Society of London Series B* **270**, 1447-1454.

Kvist, L., Martens, J., Nazarenko, A. A. & Orell, M. (2003b). Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Molecular Biology and Evolution* **20**, 243-247.

Ladoukakis, E. D. & Zouros, E. (2001). Recombination in animal mitochondrial DNA: evidence from published sequences. *Molecular Biology and Evolution* **18**, 2127-2131.

Langham, N. P. E. (1971). Seasonal movements of British terns in the Atlantic Ocean. *Bird Study* **18**, 155-175.

Larson, S., Jameson, R., Etnier, M., Fleming, M. & Bentzen, P. (2002). Loss of genetic diversity in sea otters (*Enhydra lutris*) associated with the fur trade of the 18th and 19th centuries. *Molecular Ecology* **11**, 1899-1903.

Laurent, L., Casale, P., Bradai, M. N., Godley, B. J., Gerosa, G., Broderick, A. C., Schroth, W., Schierwater, B., Levy, A. M., Freggi, D., Abd El-Mawla, E. M., Hadoud, D. A., Gomati, H. E., Domingo, M., Hadjichristophorou, M., Kornaraky, L., Demirayak, F. & Gautier, C. (1998). Molecular resolution of marine turtle stock composition in fishery bycatch: a case study in the Mediterranean. *Molecular Ecology* **7**, 1529-1542.

Lebreton, J. D., Hines, J. E., Pradel, R., Nichols, J. D. & Spendelow, J. A. (2003). Estimation by capture-recapture of recruitment and dispersal over several sites. *Oikos* **101**, 253-264.

Lessios, H. A., Kessing, B. D. & Pearse, J. S. (2001). Population structure and speciation in tropical seas: Global phylogeography of the sea urchin *Diadema*. *Evolution* **55**, 955-975.

Lessios, H. A., Kessing, B. D. & Robertson, D. R. (1998). Massive gene flow across the world's most potent marine biogeographic barrier. *Proceedings of the Royal Society of London Series B* **265**, 583-588.

Levinson, G. & Gutman, G. A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* **4**, 203-221.

Liebers, D., Helbig, A. J. & de Knijff, P. (2001). Genetic differentiation and phylogeography of gulls in the *Larua cachinnans - fuscus* group (Aves: Charadriiformes). *Molecular Ecology* **10**, 2447-2462.

Limpus, C. J. & Lyon, B. J. (1981). Seabirds breeding on the Swain Reefs, Queensland. *Corella* **5**, 101-105.

Lu, G., Basley, D. J. & Bernatchez, L. (2001). Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. *Molecular Ecology* **10**, 965-985.

Macaulay, V., Richards, M. & Sykes, B. (1999). Mitochondrial DNA recombination - no need to panic. *Proceedings of the Royal Society of London Series B* **266**, 2037-2039.

Magoulas, A. & Zouros, E. (1993). Restriction-site heteroplasmy in anchovy (*Engraulis encrasicolus*) indicates incidental biparental inheritance of mitochondrial DNA. *Molecular Biology and Evolution* **10**, 319-325.

Mayr, E. (1940). Speciation phenomena in birds. *American Naturalist* **74**, 249-278.

Merilä, J., Björklund, M. & Baker, A. J. (1997). Historical demography and present day population structure of the greenfinch, *Carduelis chloris* - an analysis of mtDNA control-region sequences. *Evolution* **51**, 946-956.

Meusel, M. S. & Moritz, R. F. A. (1993). Transfer of paternal mitochondrial DNA during fertilization of honeybee (*Apis mellifera* L.) eggs. *Current Genetics* **24**, 539-543.

Mila, B., Girman, D. J., Kimura, M. & Smith, T. B. (2000). Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *Proceedings of the Royal Society of London Series B* **267**, 1033-1040.

Milton, D. A., Smith, G. C. & Blaber, S. J. M. (1996). Variable success in breeding of the roseate tern *Sterna dougallii* on the Northern Great Barrier Reef. *Emu* **96**, 123-131.

Minami, H. & Ogi, H. (1997). Determination of migratory dynamics of the sooty shearwater in the Pacific using stable carbon and nitrogen isotope analysis. *Marine Ecology Progress Series* **158**, 249-256.

Mizutani, H., Fukuda, M., Kabaya, Y. & Wada, E. (1990). Carbon isotope ratio of feathers reveals feeding behavior of cormorants. *Auk* **107**, 400-403.

Moen, S. M. (1991). Morphologic and genetic variation among breeding colonies of the Atlantic puffin (*Fratercula arctica*). *Auk* **108**, 755-763.

Monteiro, L., Pereira, J. & Feio, R. (1997). Monitoring Report - Azores. *Roseate Tern Newsletter* **10**, 3-4.

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

Moum, T. & Arnason, E. (2001). Genetic diversity and population history of two related seabird species based on mitochondrial DNA control region sequences. *Molecular Ecology* **10**, 2463-2478.

Moum, T., Erikstad, K. E. & Bjorklid, E. (1991). Restriction fragment analysis of mitochondrial DNA in Common Murres, *Uria aalge*, from four Norwegian seabird colonies. *Candian Journal of Zoology* **69**, 1577-1584.

Moum, T. & Johansen, S. (1992). The mitochondrial NADH dehydrogenase subunit 6 (ND6) gene in murres: relevance to phylogenetic and population studies among birds. *Genome* **35**, 903-906.

Moynihan, M. (1959). A revision of the Family Laridae (Aves). *American Museum Novitates* **1928**, 1-42.

Mulder, R. A., Dunn, P. O., Cockburn, A., Lazenby-Cohen, K. A. & Howell, M. J. (1994). Helpers liberate female fairy-wrens from constraints on extra-pair mate choice. *Proceedings of the Royal Society of London Series B* **255**, 223-229.

Newton, S. & Crowe, O. (2000). Roseate terns - the natural connection. Maritime INTERREG Rpt. No 2. Marine Institute, Dublin.

Nicholls, J. A., Double, M. C., Rowell, D. M. & Magrath, R. D. (2000). The evolution of cooperative and pair breeding in thornbills *Acanthiza* (Pardalotidae). *Journal of Avian Biology* **31**, 165-176.

Nielsen, E. E., Hansen, M. M. & Loeschcke, V. (1997). Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years. *Molecular Ecology* **6**, 487-492.

Nielsen, K. K. & Arctander, P. (2001). Recombination among multiple mitochondrial pseudogenes from a Passerine genus. *Molecular Phylogenetics and Evolution* **18**, 362-369.

Nisbet, I. C. T. (1973). Terns in Massachusetts: present numbers and historical changes. *Bird Banding* **44**, 27-55.

Nisbet, I. C. T. (1980). Status and trends of the roseate tern *Sterna dougallii* in North America and the Caribbean. U. S. Department of the Interior.

Nisbet, I. C. T. (1984). Migration and winter quarters of North American roseate terns as shown by banding recoveries. *Journal of Field Ornithology* **55**, 1-17.

Nisbet, I. C. T. (2002). Common tern (*Sterna hirundo*). In *The Birds of North America* (Poole, A. & Gill, F., eds.), No. 618. The Birds of North America, Inc., Philadelphia, PA.

Nisbet, I. C. T. & Cabot, D. (1995). Transatlantic recovery of a ringed roseate tern *Sterna dougallii*. *Ringing and Migration* **16**, 14-15.

Nisbet, I. C. T. & Spendelow, J. A. (1999). Contribution of research to management and recovery of the roseate tern: review of a twelve-year project. *Waterbirds* **22**, 239-252.

O'Brien, S. J. & Mayr, E. (1991). Bureaucratic mischief: recognizing endangered species and subspecies. *Science* **251**, 1187-1188.

O'Neill, P., Minton, C., Ozaki, K. & White, R. (in press). Three populations of non-breeding roseate terns *Sterna dougallii* in the Swain Reefs, southern Great Barrier Reef. *Emu*.

Ovenden, J. R. & Street, R. (2003). Genetic population structure of mangrove jack, *Lutjanus argentimaculatus* (Forsskål). *Marine & Freshwater Research* **54**, 127-137.

Paetkau, D. (1999). Using genetics to identify intraspecific conservation units: a critique of current methods. *Conservation Biology* **13**, 1507-1509.

Panov, E. N. & Monzikov, D. G. (2000). Status of the form *barabensis* within the '*Larus argentatus - cachinnans - fuscus* complex'. *British Birds* **93**, 227-241.

Pardini, A. T., Jones, C. S., Noble, L. R., Kreiser, B., Malcolm, H., Bruce, B. D., Stevens, J. D., Cliff, G., Scholl, M. C., Francis, M., Duffy, C. A. J. & Martin, A. P. (2001). Sex-biased dispersal of great white sharks. *Nature* **412**, 139-140.

Parmelee, D. F. (1988). The hybrid skua: a southern ocean enigma. *Wilson Bulletin* **100**, 345-356.

Pennock, D. S. & Dimmick, W. W. (1997). Critique of the evolutionarily significant unit as a definition for "Distinct Population Segments" under the U. S. Endangered Species Act. *Conservation Biology* **11**, 611-619.

Perry, R. (1972). *At the turn of the tide*, Taplinger Publishing, London.

Peters, J. L. (1934). *Check-list of birds of the world*, Vol 2, Harvard University Press, Cambridge.

Phillips, A. R. (1982). Subspecies and species: fundamentals, needs, and obstacles. *Auk* **99**, 612-615.

Pierotti, R. (1987). Isolating mechanisms in seabirds. *Evolution* **41**, 559-570.

Pietz, P. J. (1987). Feeding and nesting ecology of sympatric south polar and brown skuas. *Auk* **104**, 617-627.

Pirazzoli, P. A., Ed. (1991). World atlas of Holocene sea-level changes. Elsevier Science, Amsterdam.

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

Prendergast, H. D. V. (1984). Bill colour of breeding roseate terns. *Sunbird* **14**, 69-76.

Quesada, H., Skibinski, D. A. G. & Skibinski, D. O. F. (1996). Sex biased heteroplasmy and mitochondrial DNA inheritance in the mussel *Mytilus galloprovincialis* Lmk. *Current Genetics* **29**, 423-426.

Rajora, O. P., Rahman, M. H., Buchert, G. P. & Dancik, B. P. (2000). Microsatellite analysis of genetic effects of harvesting in old-growth eastern white pine (*Pinus strobus*) in Ontario, Canada. *Molecular Ecology* **9**, 339-348.

Rambaut, A. (1996). Sequence Alignment Editor 1.0 edit.

Ramey II, R. R. (1995). Mitochondrial DNA variation, population structure, and evolution of mountain sheep in the south-western United States and Mexico. *Molecular Ecology* **4**, 429-439.

Rand, D. M. (1996). Neutrality tests of molecular markers and the connection between DNA polymorphism, demography, and conservation biology. *Conservation Biology* **10**, 665-671.

Randall, R. M. & Randall, B. M. (1980). Status and distribution of the roseate tern in South Africa. *Ostrich* **51**, 14-20.

Richardson, D. S., Jury, F. L., Blaakmeer, K., Komdeur, J. & Burke, T. (2001). Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (*Acrocephalus sechellensis). Molecular Ecology* **10**, 2263-2273.

Rieseberg, L. H. & Wendel, J. F. (1993). Introgression and its consequences in plants. In *Hybrid zones and the evolutionary process* (Harrison, R. G., ed.). Oxford University Press, New York.

Robbins, C. S. (1974). Probable interbreeding of common and roseate terns. *British Birds* **67**, 168-170.

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.

Rokas, A., Ladoukakis, E. & Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. *Trends in Ecology and Evolution* **18**, 411-417.

Ross, G. A., Egan, K. & Priddel, D. (1999). Hybridization between little tern *Sterna albifrons* and fairy tern *Sterna nereis* in Botany Bay, New South Wales. *Corella* **23**, 33-36.

Ruegg, K. C. & Smith, T. B. (2002). Not as the crow flies: a historical explanation for circuitous migration in Swainson's thrush (*Catharus ustulatus*). *Proceedings of the Royal Society of London Series B* **269**, 1375-1381.

Ryder, O. A. (1986). Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution* **1**, 9-10.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a laboratory manual*. 2nd edit, Cold Spring Harbour Laboratory Press, U. S. A.

Schneider, S., Kueffer, J.-M., Roessli, D. & Excoffier, L. (1997). Arlequin: a software for population genetic data analysis 1.1 edit. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schnell, G. D. (1970a). A phenetic study of the suborder Lari (Aves) I. Methods and results of principal components analyses. *Systematic Zoology* **19**, 35-57.

Schnell, G. D. (1970b). A phenetic study of the suborder Lari (Aves) II. Phenograms, discussion, and conclusions. *Systematic Zoology* **19**, 264-302.

Schrey, A. W. & Heist, E. J. (2003). Microsatellite analysis of population structure in the shortfin mako (*Isurus oxyrinchus*). *Canadian Journal of Fisheries and Aquatic Science* **60**, 670-675.

Schwartz, M. & Vissing, J. (2002). Paternal inheritance of mitochondrial DNA. *New England Journal of Medicine* **347**, 576-580.

Scoles, D. R., Collette, B. B. & Graves, J. E. (1998). Global phylogeography of mackerels of the genus *Scomber*. *Fishery Bulletin* **96**, 823-842.

Scribner, K. T., Warrillow, J. A., Leafloor, J. O., Prince, H. H., Inman, R. L., Luukkonen, D. R. & Flegel, C. S. (2003). Genetic methods for determining racial composition of Canada goose harvests. *Journal of Wildlife Management* **67**, 122-135.

Serventy, V. N. & White, S. R. (1951). The roseate tern: notes on nesting behaviour. *Emu* **50**, 145-151.

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

Shitara, H., Hayashi, J.-I., Takahama, S., Kaneda, H. & Yonekawa, H. (1998). Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by prevention of subsequent paternal leakage. *Genetics* **148**, 851-858.

Sibley, C. G. & Ahlquist, J. E. (1990). *Phylogeny and classification of birds: a study in molecular evolution*, Yale University Press, New Haven.

Sibley, C. G. & Monroe, B. L. J. (1990). *Distribution and taxonomy of birds of the world*, Yale University Press, New Haven.

Skibinski, D. O. F., Gallagher, C. & Beynon, C. M. (1994). Sex-limited mitochondrial DNA transmission in the marine mussel *Mytilus edulis*. *Genetics* **138**, 801-809.

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

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Slatkin, M. & Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**, 555-562.

Small, M. P., Stone, K. D. & Cook, J. A. (2003). American marten (*Martes americana*) in the Pacific Northwest: population differentiation across a landscape fragmented in time and space. *Molecular Ecology* **12**, 89-103.

Smith, G. C. (1991). The roseate tern *Sterna dougallii gracilis* breeding on the northern Great Barrier Reef Queensland. *Corella* **15**, 33-36.

Smith, G. C. & Buckley, R. C. (1986). Seabird Islands No. 161: Eagle Island, Great Barrier Reef, Queensland. *Corella* **10**, 81-83.

Sorenson, M. D., Ast, J. C., Dimcheff, D. E., Yuri, T. & Mindell, D. P. (1999). Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular Phylogenetics and Evolution* **12**, 105-114.

Sorenson, M. D. & Fleischer, R. C. (1996). Multiple independent transpositions of mitochondrial control region sequences to the nucleus. *Proceedings of the National Academy of Sciences of the USA* **93**, 15239-15243.

Spendelow, J. A., Nichols, J. D., Nisbet, I. C. T., Hays, H., Cormons, G. D., Burger, J., Safina, C., Hines, J. E. & Gochfeld, M. (1995). Estimating annual survival and

movement rates of adults within a metapopulation of roseate terns. *Ecology* **76**, 2415-2428.

Steeves, T. E., Anderson, D. J., McNally, H., Kim, M. H. & Friesen, V. L. (2003). Phylogeography of *Sula*: the role of physical barriers to gene flow in the diversification of tropical seabirds. *Journal of Avian Biology* **34**, 217-223.

Stone, K. D., Flynn, R. W. & Cook, J. A. (2002). Post-glacial colonization of northwestern North America by the forest-associated American marten (*Martes americana*, Mammalia: Carnivora: Mustelidae). *Molecular Ecology* **11**, 2049-2063.

Storr, G. M. & Ford, J. R. (1962). The recent expansion of the roseate tern into south-western Australia. *Emu* **61**, 293-294.

Surman, C. A. (1998). Seabird breeding schedules at the Pelsaert Group of islands, Houtman Abrolhos, Western Australia between 1993 and 1998. *Records of the Western Australian Museum* **19**, 209-215.

Surman, C. A. & Wooller, R. D. (2003). Comparative foraging ecology of five sympatric terns at a sub-tropical island in the eastern Indian Ocean. *Journal of Zoology* **259**, 219-230.

Suryan, R. M. & Irons, D. B. (2001). Colony and population dynamics of black-legged kittiwakes in a heterogeneous environment. *Auk* **118**, 636-649.

Sutovsky, P., Moreno, R. D., Ramalho-Santos, J., Dominko, T., Simerly, C. & Schatten, G. (2000). Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos. *Biology of Reproduction* **63**, 582-590.

Swofford, D. L. (1998). *PAUP*: Phylogenetic analysis using parsimony (*and other methods), Version 4*, Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P. & Bouvet, J. (1991). A single plucked feather as a source of DNA for bird genetic studies. *Auk* **108**, 959-960.

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

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Molecular Biology and Evolution* **9**, 678-687.

Tarkhnishvili, D., Hille, A. & Bohme, W. (2001). Humid forest refugia, speciation and secondary introgression between evolutionary lineages: differentiation in a Near Eastern brown frog, *Rana macrocnemis*. *Biological Journal of the Linnean Society* **74**, 141-156.

Tautz, D. & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* **12**, 4127-4138.

Taylor, B. L. & Dizon, A. E. (1999). First policy then science: why a management unit based solely on genetic criteria cannot work. *Molecular Ecology* **8**, S11-S16.

Templeton, A. R. (2001). Using phylogeographic analyses of gene trees to test species status and processes. *Molecular Ecology* **10**, 779-791.

Tinbergen, N. (1960). *The herring gull's world*, Basic Books, United States.

Tucker, G. M. & Heath, M. F. (1994). *Birds in Europe: their conservation status*, Birdlife International (Birdlife Conservation Series No. 3), Cambridge, U. K.

U. S. Fish and Wildlife Service (1987). Endangered and threatened wildlife and plants: determination of endangered and threatened status for two populations of the roseate tern. *Federal Register* **52**, 42064-42071.

Uthicke, S. & Benzie, J. A. H. (2003). Gene flow and population history in high dispersal marine invertebrates: mitochondrial DNA analysis of *Holothuria*

nobilis (Echinodermata: Holothuroidea) populations from the Indo-Pacific. *Molecular Ecology* **12**, 2635-2648.

Vinicombe, K. (1980). Tern showing mixed characters of black tern and white-winged black tern. *British Birds* **73**, 223-225.

Walker, T. (1986). The birds of Lady Elliot Island. *Sunbird* **16**, 73-82.

Walker, T. A. (1988). Roseate terns *Sterna dougallii* on southern Great Barrier Reef islands. *Corella* **12**, 56-58.

Walker, T. A. (1989a). Seabird Islands No. 200: North Reef Island, Great Barrier Reef, Queensland. *Corella* **13**, 115-117.

Walker, T. A. (1989b). Seabird Islands No. 201: Lady Elliot Island, Great Barrier Reef, Queensland. *Corella* **13**, 118-121.

Waples, R. S. (1998). Evolutionarily significant units, distinct population segments, and the Endangered Species Act: reply to Pennock and Dimmick. *Conservation Biology* **12**, 718-721.

Warman, S. R. (1979). The roseate tern *Sterna dougallii arideensis* on Aride Island, Seychelles. *Bulletin of the British Ornithologists' Club* **99**, 124-128.

Weimerskirch, H., Jouventin, P., Mougin, J. L., Stahl, J. C. & Van Beveren, M. (1985). Banding recoveries and the dispersal of seabirds breeding in French Austral and Antarctic territories. *Emu* **85**, 22-33.

Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38(6)**, 1358-1370.

Wenink, P. W. & Baker, A. J. (1996). Mitochondrial DNA lineages in composite flocks of migratory and wintering dunlins (*Calidris alpina*). *Auk* **113**, 744-756.

Wenink, P. W., Baker, A. J., Rosner, H.-U. & Tilanus, M. G. J. (1996). Global mitochondrial DNA phylogeography of holarctic breeding dunlins (*Calidris alpina*). *Evolution* **50**, 318-330.

Wenink, P. W., Baker, A. J. & Tilanus, M. G. J. (1993). Hypervariable-control-region sequences reveal global population structuring in a long-distance migrant shorebird, the dunlin (*Calidris alpina*). *Proceedings of the National Academy of Sciences of the USA* **90**, 94-98.

Wenink, P. W., Baker, A. J. & Tilanus, M. G. J. (1994). Mitochondrial control-region sequences in two shorebird species, the turnstone and the dunlin, and their utility in population genetic studies. *Molecular Biology and Evolution* **11**, 22-31.

Wennerberg, L. (2001). Breeding origin and migration pattern of dunlin (*Calidris alpina*) revealed by mitochondrial DNA analysis. *Molecular Ecology* **10**, 1111-1120.

Wennerberg, L., Klaassen, M. & Lindstrom, A. (2002). Geographical variation and population structure in the white-rumped sandpiper *Calidris fuscicollis* as shown by morphology, mitochondrial DNA and carbon isotope ratios. *Oecologia* **131**, 380-390.

Whitehouse, A. M. & Harley, E. H. (2001). Post-bottleneck genetic diversity of elephant populations in South Africa, revealed using microsatellite analysis. *Molecular Ecology* **10**, 2139-2149.

Whittam, R. M. (1998). Interbreeding of roseate and arctic terns. *Wilson Bulletin* **110**, 65-70.

Wierdl, M., Dominska, M. & Petes, T. D. (1997). Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics* **146**, 769-779.

Wilson, C. C. & Bernatchez, L. (1998). The ghost of hybrids past - fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Molecular Ecology* **7**, 127-132.

Witherby, H. F. & Ticehurst, M. F. (1908). On the more important additions to our knowledge of British birds since 1899. *British Birds* **2**, 307.

Yokoyama, Y., Lambeck, K., De Dekkar, P., Johnston, P. & Fifield, L. K. (2000). Timing of the last glacial maximum from observed sea level minima. *Nature* **406**, 713-716.

Zingo, J. M., Church, C. A. & Spendelow, J. A. (1994). Two hybrid common x roseate terns fledge at Falkner Island, Connecticut. *The Connecticut Warbler* **14**, 50-55.

Zink, R. M., Kessen, A. E., Line, T. V. & Blackwell-Rago, R. C. (2001). Comparative phylogeography of some aridland bird species. *Condor* **103**, 1-10.

Zouros, E., Freeman, K. R., Oberhauser Ball, A. & Pogson, G. H. (1992). Direct evidence for extensive paternal mitochondrial inheritance in the marine mussel *Mytilus*. *Nature* **359**, 412-414.

Zusi, R. L. (1982). Infraspecific geographic variation and the subspecies concept. *Auk* **99**, 606-608.

Appendix 1: AMOVA tables from Chapter 2, copied from Arlequin results files.

Table 1: AMOVA results when breeding colonies were grouped into Australia and Japan

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	0.692	-0.07497	-8.82
Among populations within groups	4	7.007	0.11111	13.07
Within populations	47	38.244	0.8137	95.75
Total	52	45.943	0.84985	

FST = 0.04253
P = 0.04106

Table 2: AMOVA results when breeding colonies were grouped into Japan and Queensland

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	0.517	-0.11892	-10.09
Among populations within groups	2	3.869	0.10248	8.70
Within populations	29	34.644	1.19464	101.39
Total	32	39.030	1.17820	

FST = -0.01395
P = 0.24633

Table 3: AMOVA results when breeding colonies were grouped into Japan and Western Australia

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	1.633	0.10750	16.95
Among populations within groups	1	0.200	-0.03630	-5.72
Within populations	27	15.200	0.56296	88.77
Total	29	17.033	0.63417	

FST = 0.11228
P = 0.06061

Table 4: AMOVA results when breeding colonies were grouped into Queensland and Western Australia

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	2.938	0.06947	8.16
Among populations within groups	3	4.069	0.08042	9.45
Within populations	38	26.644	0.7011	82.39
Total	42	33.651	0.85106	

FST = 0.17613
P = 0.01760

Appendix 2: Global roseate tern haplotype alignment. Haplotype numbers correspond to those in Table 3.2.

#NEXUS

[Created by Se-Al]

BEGIN DATA;

DIMENSIONS NTAX=26 NCHAR=1381;

FORMAT MISSING=? GAP=- DATATYPE=DNA;

MATRIX

Haplotype 1

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCGCAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 2

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCAAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT

CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACCAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCGCAACAAGCACCAAAACTAACCCTGTCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 3

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCAAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACCAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT

AACAACTACCCCCGCAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGYACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 4

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCAAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACCAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCGCAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGACGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 5

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA

GCCTTAGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 6

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGAGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT

AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 7

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCACAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 8

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT

GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 9

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCACTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 10

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT

CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTGTCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 11

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA

CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCTCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 12

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCACAACAAGCACCAAAACTAACCCCATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 13

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA

TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
TACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 14

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCACG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCGCAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA

CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 15

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTACTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCGCAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 16

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACATTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA

AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCCACAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 17

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACTATC
ACAATTTCGAGCAACCATTGGGCGATAGCTTGGACCGGACTAGAAAT
CAACACTCTCGCCATCATCCCACTTATCTCAAAGTCGCATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCCT
CAGCCTTAGTTCTCTTCTCAAGCATAACCAACGCATGACATACAGGA
CAATGAGATCTCACTCAATTAGTCCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAATTGGGATTAGTTCCATTTCATTTTTGA
TTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTATCGCCCTACTACTA
TCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATGGCTCTTGCCTCAGCA
GCCTTGGGTGGGTGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTTATTATCAT
TATTTATAACCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTGCCGTTTTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCTATACTTAAT
GCCGCTCTTATGTTAACGCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCATTATCCAAGAATTAACTAAACA
AGAAATAACTACAACAGCTACAATCATTGCCATACTATCTCTACTAG
GATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACGCTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATTACAGAAGACATGCCCCCA
CAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCCCACT
AACAACTACCCCCCGCAACAAGCACCAAAACTAACCCTATCCCATACC
CAGCAACCCGCCAATCCCCCCCACGCCTCCGGAAAAGGGTCTGCTGCC
AAAGACACAGAATATACAAAAACCACCAATATCCCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACAAA
ACACAACC

Haplotype 18

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACCATC

ACAATTTCGAGCAACCATTGAGCGATAGCTTGGACTGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAATCACATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCTT
CAGCCCTAGTTCTCTTCTCAAGTATAACCAACGCATGATATACAGGA
CAATGAGATCTCACTCAATTAGTTCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAACTGGGATTAGTTCCATTTCATTTTTG
ATTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTACCGCCTTACTACT
ATCAACAGTAATAAAATTCCCTCCAATCACTATCCTAGTCATAACTTC
ACACTCCCTCAACCCAACACTGCTTACCTCCATAGCCCTTGCCTCAGC
AGCCTTGGGTGGATGAATAGGATTAAACCAAACACAAGTCCGAAAA
ATCTTAGCCTTCTCCTCCATCTCTCACCTCGGTTGAATAGTAATTATT
ATTATTTATAGCCCAAAACTAACACTACTAACCTTCTACCTATATTCC
CTAATAACAACTACCGTATTTCTTACCTTAAACACAATCAAAGCTGT
AAAGCTATCAACAATAATAACCTCATGAACAAAAACTCCCATGCTTA
ATGCCACTCTTATATTAACCCTACTTTCACTAGCAGGACTTCCCCCAC
TCACAGGATTCCTACCTAAGTGACTCACTATCCAAGAATTAACTAAA
CAAGAAATAACTACAGCAGCTACAATCATTGCCATATTGTCCCTACT
AGGGTTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACACT
TCCACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAAT
CAACAGGTACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTAC
TACCTCTCTCCCCCATAATCCTATCTATCATCACAGAAGACATGCCCC
CACAATCCACCGTAACGACCCCAAGATCCCAACATTCAACAACCCCG
CTAACAACTACCCCCACAGCAAGTACCAAAACAAACCCCATCCCATA
CCCAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGATCTGCTG
CCAAAGACACAGAGTATACAAAAACCACCAATATTCCCCCCAAATAC
ACCATAAACAGTACCAATGACACAAAAGAGACCCCCAAACTCAACA
ACCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCA
TAATAAGGAGAAGGATTAGATGCAACTGCTAACCCCCCCCAACACAA
AACACAACC

Haplotype 19

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACCATC
ACAATTTCGAGCAACCATTGAGCGATAGCTTGGACTGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAATCACATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCTT
CAGCCCTAGTTCTCTTCTCAAGTATAACCAACGCATGATATACAGGA
CAATGAGATCTCACTCAATTAGTTCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAACTGGGATTAGTTCCATTTCATTTTTG
ATTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTACCGCCTTACTACT
ATCAACAGTAATAAAATTCCCTCCAATCACTATCCTAGTCATAACTTC
ACACTCCCTCAACCCAACACTGCTTACCTCCATAGCCCTTGCCTCAGC
AGCCTTGGGTGGATGAATAGGATTAAACCAAACACAAGTCCGAAAA
ATCTTAGCCTTCTCCTCCATCTCTCACCTCGGTTGAATAGTAATTATT
ATTATTTATAGCCCAAAACTAACACTACTAACCTTCTACCTATATTCC
CTAATAACAACTACCGTATTTCTTACCTTAAACACAATCAAAGCTGT
AAAGCTATCAACAATAATAACCTCATGAACAAAAACTCCCATGCTTA
ATGCCACTCTTATATTAACCCTACTTTCACTAGCAGGACTTCCCCCAC
TCACAGGATTCCTACCTAAGTGACTCACTATCCAAGAATTAACTAAA
CAAGAAATAACTACAGCAGCTACAATCATTGCCATATTGTCCCTACT
AGGGTTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACACT
TCCACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAAT

CAACAGGTACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTAC
TACCTCTCTCCCCCATAATCCTATCTATCATCACAGAAGACATGCCCC
CACAATCCACCGTAACGACCCCAAGATCCCAACATTCAACAACCCCG
CTAACAACTACCCCCACAGCAAGTACCAAAACAAACCCCATCCCATA
CCCAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGATCTGCTG
CCAAAGACACAGAGTATACAAAAACCACCAATATTCCCCCCAAACA
CACCATAAACAGTACCAATGACACAAAAGAGACCCCCAAACTCAAC
AACCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCC
ATAATAAGGAGAAGGATTAGATGCAACTGCTAACCCCCCCAACACA
AAACACAACC

Haplotype 20

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACCATC
ACAATTTCGAGCAACCATTGAGCGATAGCTTGGACTGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAATCACATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCTT
CAGCCCTAGTTCTCTTCTCAAGTATAACCAACGCATGATATACAGGA
CAATGAGATCTCACTCAATTAGTTCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAACTGGGATTAGTTCCATTTCATTTTTG
ATTCCCAGAAGTTCTCCAAGGTTCGTCCCTAACTACCGCCTTACTACT
ATCAACAGTAATAAAATTCCCTCCAATCACTATCCTAGTCATAACTTC
ACACTCCCTCAACCCAACACTGCTTACCTCCATAGCCCTTGCCTCAGC
AGCCTTGGGTGGATGAATAGGATTAAACCAAACACAAGTCCGAAAA
ATCTTAGCCTTCTCCTCCATCTCTCACCTCGGTTGAATAGTAATTATT
ATTATTTATAGCCCAAAACTAACACTACTAACCTTCTACCTATATTCC
CTAATAACAACTACCGTATTTCTTACCTTAAACACAATCAAAGCTGT
AAAGCTATCAACAATAATAACCTCATGAACAAAAACTCCCATGCTTA
ATGCCACTCTTATATTAACCCTACTTTCACTAGCAGGACTTCCCCCAC
TCACAGGATTCCTACCTAAGTGACTCACTATCCAAGAATTAACTAAA
CAAGAAATAACTACAGCAGCTACAATCATTGCCATATTGTCCCTACT
AGGGTTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACACT
TCCACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAAT
CAACAGGTACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTAC
TACCTCTCTCCCCCATAATCCTATCTATCATCACAGAAGACATGCCCC
CACAATCCACCGTAACGACCCCAAGATCCCAACATTCAACAACCCCG
CTAACAACTACCCCCACAGCAAGTACCAAAACAAACCCCATCCCATA
CCCAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGATCTGCTG
CCAAAGACACAGAGTATACAAAAACCACCAATATTCCCCCCAAATAC
ACCATAAACAGTACCAATGACACAAAAGAGACCCCCAAACTCAACA
ACCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCA
TAATAAGGAGAAGGATTAGATGCAACTGCTAACCCCCCCAACACAA
AACACAACC

Haplotype 21

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACCATC
ACAATTTCGAGCAACCATTGAGCGATAGCTTGGACTGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAATCACATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCTT
CAGCCCTAGTTCTCTTCTCAAGTATAACCAACGCATGATATACAGGA
CAATGAGATCTCACTCAATTAGTTCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAACTTGGATTAGTTCCATTTCATTTTTGA

TTCCCAGAAGTTCTCCAAGGTTCGTCCCTAACTACCGCCTTACTACTA
TCAACAGTAATAAAATTCCCTCCAATCACTATCCTAGTCATAACTTCA
CACTCCCTCAACCCAACACTGCTTACCTCCATAGCCCTTGCCTCAGCA
GCCTTGGGTGGATGAATAGGATTAAACCAAACACAAGTCCGAAAAA
TCTTAGCCTTCTCCTCCATCTCTCACCTCGGTTGAATAGTAATTATTAT
TATTTATAGCCCAAAACTAACACTACTAACCTTCTACCTATATTCCCT
AATAACAACTACCGTATTTCTTACCTTAAACACAATCAAAGCTGTAA
AGCTATCAACAATAATAACCTCATGAACAAAAACTCCCATGCTTAAT
GCCACTCTTATATTAACCCTACTTTCACTAGCAGGACTTCCCCCACTC
ACAGGATTCCTACCTAAGTGACTCAYTATCCAAGAATTAACTAAACA
AGAAATAACTACAGCAGCTACAATCATTGCCATATTGTCCCTACTAG
GGTTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACACTTC
CACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAGGTACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTACTA
CCTCTCTCCCCCATAATCCTATCTATCATCACAGAAGACATGCCCCCA
CAATCCACCGTAACGACCCCAAGATCCCAACATTCAACAACCCCGCT
AACAACTACCCCCACAGCAAGTACCAAAACAAACCCCATCCCATACC
CAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGATCTGCTGCC
AAAGACACAGAGTATACAAAAACCACCAATATTCCCCCCAAATACA
CCATAAACAGTACCAATGACACAAAAGAGACCCCCAAACTCAACAA
CCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCAT
AATAAGGAGAAGGATTAGATGCAACTGCTAACCCCCCCCAACACAAA
ACACAACC

Haplotype 22

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACCATC
ACAATTTCGAGCAACCATTGAGCGATAGCTTGGACTGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAATCACATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCTT
CAGCCCTAGTTCTCTTCTCAAGTATAACCAACGCATGATATACAGGA
CAATGAGATCTCACTCAATTAGTTCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAACTGGGATTAGTTCCATTTCATTTTTG
ATTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTACCGCCTTACTACT
ATCAACAGTAATAAAATTCCCTCCAATCACTATCCTAGTCATAACTTC
ACACTCCCTCAACCCAACACTGCTTACCTCCATAGCCCTTGCCTCAGC
AGCCTTGGGTGGATGAATAGGATTAAACCAAACACAAGTCCGAAAA
ATCTTAGCCTTCTCCTCCATCTCTCACCTCGGTTGAATAGTAATTATT
ATTATTTATAGCCCAAAACTAACACTACTAACCTTCTACCTATATTCC
CTAATAACAACTACCGTATTTCTTACCTTAAACACAATCAAAGCTGT
AAAGCTATCAACAATAATAACCTCATGAACAAAAACTCCCATGCTTA
ATGCCACTCTTATATTAACCCTACTTTCACTAGCAGGACTTCCCCCAC
TCACAGGATTCCTACCTAAGTGACTCACTATCCAAGAATTAACTAAA
CAAGAAATAACTACAGCAGCTACAATCATTGCCATATTGTCCCTACT
AGGGTTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACACT
TCCACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAAT
CAACAGGTACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTAC
TACCTCTCTCCCCCATAATCCTATCTATCATCACAGAAGACATGCCCC
CACAATCCACCGTAACGACCCCAAGATCCCAACATTCAACAACCCCG
CTAACAACTACCCCCACAGCAAGTACCAAAACAAACCCCATCCCATA
CCCAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGATCTGCTG
CCAAAGACACAGAGTATACAAAAACCACCAATATTCCCCCCAAATAC

GCCATAAACAGTACCAATGACACAAAAGAGACCCCCAAACTCAACA
ACCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCA
TAATAAGGAGAAGGATTAGATGCAACTGCTAACCCCCCCAACACAA
AACACAACC

Haplotype 23

CACGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTACTACCATC
ACAATTTCGAGCAACCATTGAGCGATAGCTTGGACTGGACTAGAAAT
CAACACTCTTGCCATCATCCCCCTTATCTCAAAATCACATCACCCCCG
GGCTATCGAAGCTGCAATCAAATACTTCCTAGTCCAAGCAGCCGCTT
CAGCCCTAGTTCTCTTCTCAAGTATAACCAACGCATGATATACAGGA
CAATGAGATCTCACTCAATTAGTTCACCCAACTTCATGCCTACTCCTA
ACAGCCGCAATTGGAATAAAACTGGGATTAGTTCCATTTCATTTTTG
ATTCCCAGAAGTTCTCCAAGGTTCATCCCTAACTACCGCCTTACTACT
ATCAACAGTAATAAAATTCCCTCCAATCACTATCCTAGTCATAACTTC
ACACTCCCTCAACCCAACACTGCTTACCTCCATAGCCCTTGCCTCAGC
AGCCTTGGGTGGATGAATAGGATTAAACCAAACACAAGTCCGAAAA
ATCTTAGCCTTCTCCTCCATCTCTCACCTCGGTTGAATAGTAATTATT
ATTATTTATAGCCCAAAACTAACACTACTAACCTTCTACCTATATTCC
CTAATAACAACTACCGTATTTCTTACCTTAAACACAATCAAAGCTGT
AAAGCTATCAACAATAATAACCTCATGAACAAAAACTCCCATGCTTA
ATGCCACTCTTATATTAACCCTACTTTCACTAGCAGGACTTCCCCCAC
TCACAGGATTCCTACCTAAGTGACTCACTATCCAAGAATTAACTAAA
CAAGAAATAACTACAGCAGCTACAATCATTGCCATATTGTCCCTACT
AGGGTTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACACT
TCCACCAAATTCCACAAACCATATAAAACAATGACACATCAATAAAT
CAACAGGTACCCCTATCGCCGTTGTTGCCTCCCTATCAATCCTGCTAC
TACCTCTCTCCCCCATAATCCTATCTATCATCACAGAAGATATGCCCC
CACAATCCACCGTAACGACCCCAAGATCCCAACATTCAACAACCCCG
CTAACAACTACCCCCACAGCAAGTACCAAAACAAACCCCATCCCATA
CCCAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGATCTGCTG
CCAAAGACACAGAGTATACAAAAACCACCAATATTCCCCCCAAATAC
ACCATAAACAGTACCAATGACACAAAAGAGACCCCCAAACTCAACA
ACCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCCCA
TAATAAGGAGAAGGATTAGATGCAACTGCTAACCCCCCCAACACAA
AACACAACC

bridled tern (WA) -------

AACTAATCACATCCCTAAGCCTACTCTTAGGCACAACCATTACAATTT
CAAGCAACCATTGAGTGATAGCTTGAACTGGATTAGAAATCAACACC
CTTGCTATTATCCCCCTTATCTCCAAATCGCACCACCCTCGGGCTATT
GAAGCTGTAATCAAATACTTCTTAGTTCAAGCAGCTGCTTCAACTTTA
ATTCTCTTCTCAAGCATAATCAATGCATCATATACAGGACAATGAGA
TATCACCCAACTAACCCACCCAACTTCATGCTTACTTCTAACAGCTGC
AATCGGAATAAAACTAGGACTAGTCCCATTCCACTTCTGATTTCCAG
AAGTACTTCAAGGTTCACCCCTAACTACTGCCCTACTACTATCAACA
ATGATAAAATTCCCCCCAATCACCATTCTATTTATAACCTCACACTCC
CTTGACCCAACACTACTAACCTCTATGGCCATCGCCTCAACAGCTCTG
GGTGGATGAATGGGGTTAAACCAAACACAAATTCGAAAAATCCTAG
CCTTCTCCTCTATCTCTCACCTGGGCTGAATAGTAATCATCATTATCT
ACAGCCCAAAACTCACACTATTAACCTTTTACCTATACTCTTTAATAA

CTACCACTGTATTTCTAACCCTTAATACAATCAAAGTTATAAAACTAT
CAACAATAATAACCTCATGAACAAAAACCCCCATACTCAACGCTACT
CTTATACTAACCCTACTTTCACTAGCAGGGCTCCCCCCACTCACAGGA
TTCCTACCTAAATGACTCATCATTCAAGAATTAACTAAACAAGAAAT
AACTACAACAGCTACAATTATCGCTATATTATCCTTACTAGGATTATT
CTTCTACCTTCGTCTTGCATACTACTCAACAATTACCCTCCCACCAAA
CTCCACAAACCACATAAAACAGTGACACATTAACAAATCAACAAAC
ACCTTCATCGCCATTGTCACCTCTCTGTCAATCCTACTATTACCCCTCT
CCCCCATAATCCTTTCCATCATTGCAGAAGATATACCCCCACTATCCA
CAGTAACAATCCCAGAATTTCAACACTCAATAACTCCACTAACAACC
ACCCCCACAATAAGTACCAAAACGAACCCTCCCCCATACCCCATAAC
CCGCCAATCCCCCCAAGTTTCCGGAAAAGGATCCGCCGCCAAGGACA
CAGAGTACACAAAAACTACCAGTATCCCCCCCAAATATACCATGAAC
AGTACCAACGATACAAAAGACACCCCCAAACTCAACAGCCACCCAC
ACCCTGCAACAGAGGCTAACACTAAACCAACCACCCCATAATAAGG
AGAAGGATTAGATGCAACCGCCAAACCCCCTAGCACAAAACACAAT
C

bridled tern (QLD) ----------

TCACATCCCTAAGCCTACTCTTAGGCACAACCATTACAATTTCAAGC
AACCATTGAGTGATAGCTTGAACTGGATTAGAAATCAACACCCTTGC
TATTATCCCCCTTATCTCCAAATCGCACCACCCTCGGGCTATTGAAGC
TGTAATCAAATACTTCTTAGTTCAAGCAGCTGCTTCAACTTTAATTCT
CTTCTCAAGCATAATCAATGCATCATATACAGGACAATGAGATATCA
CCCAACTAACCCACCCAACTTCATGCTTACTTCTAACAGCTGCAATCG
GAATAAMACTAGGACTAGTCCCATTCCACTTCTGATTTCCAGAAGTA
CTTCAAGGTTCACCCCTAACTACTGCCCTACTACTATCAACAATGATA
AAATTCCCCCCAATCACCATTCTATTTATAACCTCACACTCCCTTGAC
CCAACACTACTAACCTCTATGGCCATCGCCTCAACAGCTCTGGGTGG
ATGAATGGGGTTAAACCAAACACAAATTCGAAAAATCCTAGCCTTCT
CCTCTATCTCTCACCTGGGCTGAATAGTAATCATCATTATCTACAACC
CAAAACTCACACTATTAACCTTTTACCTATACTCTTTAATAACTACCA
CTGTATTTCTAACCCTTAATACAATCAAAGTTATAAAACTATCAACA
ATAATAACCTCATGAACAAAAACCCCCATACTCAACGCTACTCTTAT
ACTAACCCTACTTTCACTAGCAGGGCTCCCCCCACTCACAGGATTCCT
ACCTAAATGACTCATCATTCAAGAATTAACTAAACAAGAAATAACTA
CAACAGCTACAATTATCGCTATATTATCCTTACTAGGATTATTCTTCT
ACCTTCGTCTTGCATACTACTCAACAATTACCCTCCCACCAAACTCCA
CAAACCACATAAAACAGTGACACATTAACAAATCAACAAACACCTTC
ATCGCCATTGTCACCTCTCTGTCAATCCTACTATTACCCCTCT----------

GCAGAAGATATACCCCCACTATCCACAGTAACAATCCCAGAATTTCA
ACACTCAATAACTCCACTAACAACCACCCCCACAATAAGTACCAAAA
CGAACCCTCCCCCATACCCCATAACCCGCCAATCCCCCCAAGTTTCC
GGAAAAGGATCCGCCGCCAAGGACACAGAGTACACAAAAACTACCA
GTATCCCCCCCCAAATATACCATGAACAGTACCAACGATACAAAAGAC
ACCCCCAAACTCAACAGCCACCCACACCCTGCAACAGAGGCTAACAC
TAAACCAACCACCCCATAATAAGGAGAAGGATTAGATGCAACTGCC
AAACCCCCTAGCACAAAACACAATC

black-naped tern ----------

TCACTTCCCTAAGCCTACTCCTGGGTACTACTATCACAATTTCAAGCA
ACCATTGAGCGATAGCTTGGACCGGACTAGAAATCAATACTCTTGCC
ATCATCCCCCTTATCTCAAAATCACATCACCCCCGGGCTATCGAAGC
CGCAATCAAATACTTCCTAGTCCAAGCAGCTGCTTCAGCCCTAGTTCT
CTTCTCAAGCATAACCAACGCATGACACACAGGACAGTGAGATCTCA
CTCAGTNAATCCACCCAACTTCTTGCCTACTCCTAACAGCCGCAATTG
GAATAAAATTGGGATTAGTCCCATTTCATTTTTGATTCCCAGAAGTCC
TCCAAGGTTCATCCCTAACTACCGCCCTACTACTATCAACAGTAATA
AAATTCCCCCCAATCACTATCCTATTCATAACTTCACACTCCCTTAAC
CCAACACTGCTCACCTCCATAGCCCTTGCCTCAGCAGCCTTGGGCGG
GTGAATAGGATTAAACCAAACACAAGTCCGAAAAATCCTAGCCTTCT
CCTCCATCTCTCACCTAGGTTGAATAGTAATTATCATTATTTATAACC
CAAAACTAACACTACTAACCTTCTACCTATATTCCCTAATAACAATTA
CCGTATTTCTTACCTTAAACACAATCAAAGCTGTAAAGCTATCAACA
ATAATAACCTCATGAACAAAAACTCCTATACTTAATGCCACTCTTAT
ATTAACTCTGCTTTCACTGGCAGGACTTCCCCCACTCACAGGATTCCT
ACCCAAGTGACTCATTATCCAAGAATTAACTAAACAAGAAATAACTA
CAGCAGCTACAGTCATTGCCATACTATCCCTACTAGGATTATTCTTCT
ACCTTCGCCTTGCATACTACTCAACAATCACACTTCCACCAAACTCCA
CAAACCATATAAAACAATGACACATCAATAAATCAACAGGCACTCCT
ATCGCCATTGTTGCCTCCCTATCAATCCTGCTACTACCTCTCTCCCCC
ATAATCCTAT--------
ACAGAAGACATACCCCCACAATCCACCGTAACAACCCCAAGATCCCA
ACATCCAACAACCCCACTAACAGCTACCCCCACAACAAGCACCGAA
ACAAACCCCACCCCATATCCAACAACCCGCCAATCTCCCCAAGCCTC
AGGAAAAGGATCTGCTGCCAAAGACACAGAGTATACAAAAACCACC
AACATCCCCCCCAAATACACCATAAACAACACCAATGACACAAAAG
AAACCCCCAAACTCAACAACCACCCACACCCTGCAATAGAAGCCGA
CACCAAACCAACCACCCCATAATAAGGAGAAGGATTAGATGCAACT
GCTAACCCCCCTAACACAAAGCACAACC

;

END;

Appendix3: Individual microsatellite genotypes. Numbers indicate the allele size in base pairs. Missing data is coded as 0

		Microsatellite loci									
Sample	Sampling location	RBG27		RBG29		Sdaat20		RBG18		Sdaat46	
RTP1	Pelsaert	197	201	144	144	194	194	179	179	243	250
RTP2	Pelsaert	201	201	144	144	187	187	179	179	243	250
RTP3	Pelsaert	197	201	136	144	193	193	179	179	243	250
RTP4	Pelsaert	201	201	136	136	193	193	179	179	243	250
RTP5	Pelsaert	195	201	136	144	187	187	179	179	243	250
RTP7	Pelsaert	197	201	136	144	193	193	179	179	243	243
RTP8	Pelsaert	197	197	144	144	193	193	179	181	250	250
RTP9	Pelsaert	197	201	136	144	193	193	179	179	243	250
RTP10	Pelsaert	197	201	136	144	193	193	179	179	243	243
RTP11	Pelsaert	195	201	136	144	193	193	179	179	243	250
RTP12	Pelsaert	201	201	136	144	187	187	179	179	243	250
RTP13	Pelsaert	201	201	136	144	193	193	179	179	243	250
RTP14	Pelsaert	195	195	136	144	193	193	179	179	243	243
RTP15	Pelsaert	197	201	136	136	193	193	179	181	243	243
RTP18	Pelsaert	197	201	144	144	193	193	179	181	243	250
RTP19	Pelsaert	195	201	144	144	192	192	179	179	243	243
RTP20	Pelsaert	197	201	136	144	193	193	175	179	243	250
RTP21	Pelsaert	201	201	136	136	193	193	181	181	243	243
RTP22	Pelsaert	201	201	144	144	193	193	179	179	243	243
RTL1	Lancelin	195	201	136	144	192	192	179	179	243	243
RTL2	Lancelin	201	201	144	144	194	194	179	179	243	250
RTL3	Lancelin	195	195	144	144	193	193	179	179	250	250
RTL4	Lancelin	195	195	144	144	187	187	179	179	243	250
RTL5	Lancelin	201	201	144	144	193	193	179	179	243	243
RTL6	Lancelin	201	201	136	136	193	193	181	181	250	250
RTL7	Lancelin	201	201	144	144	193	193	179	179	243	250
RTL8	Lancelin	201	201	144	144	193	193	179	179	243	250
RTL9	Lancelin	201	201	136	144	192	192	179	179	243	243
RTL10	Lancelin	197	201	136	144	187	187	179	179	243	243
RTE1	Lady Elliot	195	201	144	144	193	193	179	179	250	250
RTE2	Lady Elliot	195	195	136	144	193	193	179	179	243	243
RTE3	Lady Elliot	201	201	134	144	193	193	179	179	250	250
RTE4	Lady Elliot	195	201	136	144	193	193	179	179	243	250
RTE5	Lady Elliot	201	201	144	144	187	193	179	179	250	250
RTE6	Lady Elliot	201	201	136	136	193	193	179	179	250	250
RTE7	Lady Elliot	201	201	144	144	193	193	179	179	243	250
RTE8	Lady Elliot	201	201	144	144	193	193	179	179	250	250
RTE9	Lady Elliot	201	203	144	144	193	193	179	179	243	243
RTE10	Lady Elliot	201	201	134	136	193	193	179	179	250	250
RTE12	Lady Elliot	201	201	136	144	193	193	0	0	0	0
RTE13	Lady Elliot	203	205	144	144	193	193	179	179	243	250
RTE14	Lady Elliot	203	203	136	136	193	193	179	179	250	250
RTE15	Lady Elliot	195	201	144	144	193	193	179	179	243	250
RTE16	Lady Elliot	195	195	136	144	193	193	179	179	243	250
RTE17	Lady Elliot	197	201	136	144	187	187	179	179	246	250
RTE18	Lady Elliot	201	205	134	144	193	193	179	179	243	250
RTE19	Lady Elliot	195	201	134	136	193	193	179	185	246	250
RTE20	Lady Elliot	201	201	138	144	193	193	179	179	243	243
RTE21	Lady Elliot	195	201	134	136	193	193	179	179	246	246
RTE22	Lady Elliot	195	195	136	136	193	193	179	179	243	250
RTE23	Lady Elliot	195	195	134	136	193	193	179	179	243	246
RTN1	Bountiful	197	197	136	136	187	193	179	179	243	250
RTN2	Bountiful	201	201	144	144	193	193	179	179	250	250
RTN3	Bountiful	201	201	136	144	193	193	179	183	243	243
RTN4	Bountiful	195	201	144	144	187	193	179	179	243	246
RTN5	Bountiful	201	201	136	144	193	193	179	179	246	250

RTN6	Bountiful	197	197	144	144	187	193	179	181	243	243
RTN7	Bountiful	195	201	144	144	187	193	179	179	243	250
RTN8	Bountiful	203	203	144	144	193	193	179	179	243	250
RTN9	Bountiful	197	197	136	136	193	193	179	181	243	243
RTN10	Bountiful	201	201	144	144	193	193	179	179	243	250
RTN13	Bountiful	201	201	136	136	193	203	179	179	243	243
RTN15	Bountiful	201	201	136	144	187	193	179	179	243	250
RTN16	Bountiful	201	201	144	144	193	193	179	179	243	250
RTN17	Bountiful	197	201	144	144	187	192	179	179	246	250
RTN19	Bountiful	197	201	136	144	193	193	179	179	250	250
RTN20	Bountiful	201	203	138	144	187	193	179	183	250	250
RTN21	Bountiful	201	201	136	144	187	192	179	181	243	243
RTN22	Bountiful	201	201	144	144	193	193	179	179	243	243
RTS3	Sisters	201	201	136	144	187	193	179	179	243	250
RTS5	Sisters	195	197	144	144	193	193	179	179	243	250
RTS6	Sisters	197	201	144	144	192	192	179	179	243	243
RTS7	Sisters	201	201	134	136	187	193	179	179	243	250
RTW1	Lacepede	197	201	136	138	187	187	177	179	243	250
RTW2	Lacepede	201	201	144	144	193	193	177	179	250	250
RTW3	Lacepede	197	201	136	144	187	187	177	179	250	250
RTW4	Lacepede	201	201	144	144	193	193	179	179	243	250
RTW5	Lacepede	201	201	136	144	187	193	179	181	243	243
RTW6	Lacepede	201	201	144	144	193	193	179	179	243	250
RTW7	Lacepede	201	201	144	144	193	193	179	179	250	250
RTW8	Lacepede	201	201	136	144	192	192	179	179	250	250
RTW9	Lacepede	201	201	134	144	193	193	175	175	243	246
RTW10	Lacepede	197	201	136	144	193	193	179	183	243	243
RTW11	Lacepede	197	201	144	144	193	193	179	179	250	250
RTW12	Lacepede	201	203	136	144	193	193	179	179	243	243
RTW13	Lacepede	195	201	136	144	191	193	183	185	243	250
RTW14	Lacepede	201	201	144	144	192	192	179	179	250	250
RTW15	Lacepede	197	201	136	144	193	193	179	179	243	246
RTW16	Lacepede	201	203	136	144	192	192	179	181	243	250
RTW17	Lacepede	197	201	136	144	193	193	179	179	243	250
RTW18	Lacepede	201	201	136	144	193	193	179	181	243	250
RTW19	Lacepede	195	201	136	144	187	193	179	179	243	250
RTW20	Lacepede	195	201	136	144	193	193	179	181	243	246
RTW21	Lacepede	201	201	136	144	193	193	179	181	243	250
RTW22	Lacepede	201	201	136	136	193	193	179	179	243	243
RTW23	Lacepede	195	203	136	136	192	192	179	181	243	243
RTW24	Lacepede	195	201	138	144	193	193	175	179	243	250
RTW25	Lacepede	205	205	136	144	193	193	179	181	250	250
RTW26	Lacepede	197	201	136	136	189	193	179	181	243	250
RTW27	Lacepede	197	201	136	144	193	193	181	181	243	250
RTW28	Lacepede	195	195	144	144	193	193	179	181	243	250
RTW29	Lacepede	201	201	144	144	193	193	179	179	243	250
RTW30	Lacepede	201	201	144	144	193	202	179	179	243	243
RTW31	Lacepede	201	201	134	144	193	193	181	181	243	250
RTW32	Lacepede	201	201	136	138	187	187	179	179	243	250
RTW33	Lacepede	201	201	136	136	193	193	179	185	243	243
RTW34	Lacepede	201	205	136	144	202	202	179	179	243	250
RTW35	Lacepede	201	201	136	136	193	193	179	181	243	250
RTG1	Swains summer	201	201	136	144	193	193	179	181	243	250
RTG2	Swains summer	201	201	136	144	193	193	179	179	243	250
RTG3	Swains summer	201	203	134	144	198	202	179	185	250	250
RTG4	Swains summer	201	201	136	144	202	202	179	179	243	250
RTG5	Swains summer	201	203	136	144	194	194	181	183	250	250
RTG6	Swains summer	201	203	136	144	193	193	179	181	243	250
RTG7	Swains summer	201	201	136	144	187	193	179	179	250	250

RTG8	Swains summer	195	195	136	136	193	193	179	181	250	250
RTG9	Swains summer	199	201	144	144	193	193	179	179	243	243
RTG10	Swains summer	201	201	144	144	192	192	179	179	250	250
RTG11	Swains summer	201	201	144	144	193	193	177	181	243	250
RTG12	Swains summer	201	203	144	144	192	192	179	185	246	250
RTG13	Swains summer	201	201	136	144	192	192	179	179	246	250
RTG14	Swains summer	201	201	136	144	193	202	179	181	243	250
RTG15	Swains summer	197	201	138	144	193	193	179	179	243	243
RTG16	Swains summer	201	201	144	144	193	193	179	181	250	250
RTG17	Swains summer	201	201	136	138	193	202	179	179	243	243
RTG18	Swains summer	195	201	136	144	193	193	175	179	243	250
RTG19	Swains summer	201	201	136	144	187	194	179	179	250	250
RTG20	Swains summer	203	203	144	144	193	193	179	179	250	250
RTG21	Swains summer	195	195	136	138	193	193	179	179	243	250
RTG22	Swains summer	197	201	138	144	193	193	179	181	246	250
RTG23	Swains summer	195	203	136	144	193	193	179	179	246	246
RTG24	Swains summer	201	201	134	136	193	193	181	181	250	250
RTG26	Swains summer	201	201	136	144	187	187	179	181	250	250
RTG27	Swains summer	201	203	144	144	192	192	179	179	243	250
RTG29	Swains summer	195	195	144	144	192	192	179	179	246	250
RTG30	Swains summer	201	201	144	144	193	202	179	179	243	250
RTG31	Swains summer	195	195	136	144	187	193	179	179	243	250
RTG32	Swains summer	201	201	144	144	187	187	179	181	243	250
RTG33	Swains summer	201	203	136	144	193	193	179	179	243	250
RTG34	Swains summer	197	201	136	144	193	193	179	179	250	250
RTG35	Swains summer	203	203	136	144	187	193	179	179	243	246
RTG36	Swains summer	201	201	136	144	187	193	179	179	243	243
RTG37	Swains summer	201	203	136	144	193	193	179	179	243	250
RTG38	Swains summer	201	201	136	144	193	193	179	187	246	250
RTG39	Swains summer	201	201	144	144	187	187	179	181	243	246
RTG40	Swains summer	201	205	136	136	187	187	179	179	246	250
RTG41	Swains summer	195	195	136	144	193	193	179	181	243	243
RTG42	Swains summer	197	201	144	144	187	187	179	181	243	243
RTG43	Swains summer	201	201	136	136	193	202	179	179	0	0
RTG44	Swains summer	201	201	138	144	193	193	179	179	243	250
RTG45	Swains summer	195	201	136	144	193	193	179	179	243	250
RTG46	Swains summer	201	201	136	136	187	193	179	179	243	250
RTG47	Swains summer	201	201	134	134	193	193	179	179	250	250
RTG48	Swains summer	197	201	136	136	192	202	179	179	250	250
RTG49	Swains summer	197	201	136	144	187	193	179	179	243	250
RTG50	Swains summer	197	197	134	136	187	187	181	181	243	243
RTG51	Swains summer	195	201	136	144	194	194	179	181	243	246
RTG52	Swains summer	201	203	136	144	187	193	179	179	243	250
RTG53	Swains summer	201	201	136	144	193	193	179	179	243	250
RTG54	Swains summer	197	201	144	144	187	187	181	181	243	250
RTB55	Swains summer	201	201	136	136	193	193	179	179	243	250
RTB56	Swains summer	201	201	136	144	187	193	179	181	243	250
RTB57	Swains summer	201	201	144	144	187	193	179	179	250	250
RTB58	Swains summer	195	201	144	144	192	192	179	179	243	250
RTB59	Swains summer	201	201	136	136	193	193	179	179	243	250
RTB60	Swains summer	201	201	136	144	187	193	179	179	243	250
RTB61	Swains summer	195	201	136	144	193	193	179	179	243	250
RTB62	Swains summer	197	197	136	136	193	193	179	179	246	250
RTB63	Swains summer	201	203	136	144	187	193	179	179	250	250
RTB64	Swains summer	201	201	136	144	193	193	179	179	250	250
RTB65	Swains summer	195	195	136	144	193	193	179	181	250	250
RTB66	Swains summer	201	201	136	136	193	193	179	179	250	250
RTB68	Swains summer	203	203	138	144	193	193	179	179	243	243
RTB69	Swains summer	201	201	144	144	187	193	179	179	246	250

RTB70	Swains summer	197	201	136	144	193	193	181	181	243	250
RTB71	Swains summer	197	197	136	144	193	193	179	179	250	250
RTB72	Swains summer	201	201	136	136	187	187	179	179	243	246
RTB73	Swains summer	197	201	144	144	193	193	179	183	243	250
RTB74	Swains summer	197	197	134	138	187	193	179	179	250	250
RTB75	Swains summer	201	201	136	144	193	193	179	179	250	250
RTB76	Swains summer	203	203	136	144	193	193	179	179	243	250
RTB77	Swains summer	197	197	144	144	187	187	179	179	250	250
RTB78	Swains summer	197	197	136	136	193	193	179	179	246	250
RTB79	Swains summer	201	201	144	144	193	193	179	179	243	250
RTB80	Swains summer	197	201	136	136	193	193	179	181	243	250
RTB81	Swains summer	201	201	136	136	193	193	179	181	250	250
RTB82	Swains summer	197	197	134	136	193	193	179	183	250	250
RT48	Swains winter	201	201	134	144	187	192	179	179	243	250
RT49	Swains winter	195	201	136	136	193	193	179	179	250	250
RT50	Swains winter	197	201	134	136	193	193	179	179	243	250
RT51	Swains winter	201	201	144	144	187	193	179	179	243	243
RT52	Swains winter	195	201	136	144	193	193	179	179	243	250
RT53	Swains winter	205	205	144	144	193	193	179	181	250	250
RT54	Swains winter	197	201	136	144	193	193	179	179	243	243
RT55	Swains winter	197	201	136	144	193	193	179	179	243	246
RT56	Swains winter	195	195	136	144	193	193	179	181	250	250
RT57	Swains winter	195	197	136	136	193	193	179	179	250	250
RT58	Swains winter	201	201	136	136	193	193	179	181	250	250
RT59	Swains winter	201	201	136	144	193	193	179	181	246	250
RT60	Swains winter	201	203	144	144	192	192	179	181	243	250
RT61	Swains winter	195	201	136	144	193	193	179	181	250	250
RT62	Swains winter	201	201	134	136	193	193	179	179	243	243
RT63	Swains winter	196	201	136	136	192	192	179	179	250	250
RT64	Swains winter	203	205	134	144	193	193	179	179	243	243
RT66	Swains winter	195	201	136	144	193	193	179	179	243	243
RT67	Swains winter	195	201	136	144	193	193	179	179	243	246
RT69	Swains winter	195	195	136	144	193	193	179	181	250	250
RT70	Swains winter	201	201	134	144	193	193	179	179	243	243
RT71	Swains winter	201	201	136	144	193	193	0	0	250	250
RT72	Swains winter	201	201	136	144	193	193	179	181	250	250
RT73	Swains winter	195	195	138	144	193	193	179	181	250	250
RT76	Swains winter	197	205	144	144	193	193	0	0	0	0
RT77	Swains winter	203	203	134	144	193	193	179	179	250	250
RT78	Swains winter	205	205	136	144	193	193	179	179	243	250
RT79	Swains winter	201	201	136	144	193	193	179	179	250	250
RT80	Swains winter	201	201	136	144	193	193	179	179	250	250
RT81	Swains winter	195	195	136	144	193	193	0	0	0	0
RT82	Swains winter	195	195	136	144	193	193	179	179	250	250
RT83	Swains winter	201	201	144	144	187	192	179	179	243	243
RT84	Swains winter	195	201	136	144	193	193	179	179	243	243
RT85	Swains winter	201	201	144	144	187	193	177	179	243	243
RT86	Swains winter	201	205	134	144	193	193	179	181	250	250
RT87	Swains winter	195	205	134	144	193	193	179	181	243	250
RT88	Swains winter	195	195	136	136	193	193	179	179	246	250
RT89	Swains winter	195	201	136	144	192	193	179	179	250	250
RT90	Swains winter	201	201	134	144	193	193	179	181	250	250
RT91	Swains winter	195	201	134	144	193	193	179	179	250	250
RT92	Swains winter	195	201	136	144	193	193	179	179	246	250
RT94	Swains winter	201	203	136	144	193	193	179	179	250	250
RT95	Swains winter	201	203	136	136	193	193	179	179	250	250
RT96	Swains winter	201	203	136	144	193	193	179	179	243	250
RT97	Swains winter	195	203	136	144	193	193	179	179	250	250
RT98	Swains winter	195	195	144	144	193	193	179	181	250	250

RT99	Swains winter	201	201	136	144	193	193	179	181	243	250
RT100	Swains winter	201	201	136	144	193	193	179	179	250	250
RT101	Swains winter	195	203	136	144	193	193	179	179	250	250
RT102	Swains winter	195	195	136	144	193	193	179	181	250	250
RT103	Swains winter	195	201	144	144	193	193	177	179	243	243
Sdb1	Japan	201	201	136	144	193	193	179	183	250	250
Sdb2	Japan	197	205	136	136	194	194	181	181	0	0
Sdb3	Japan	201	201	144	144	193	194	179	181	250	250
Sdb4	Japan	197	201	136	136	192	192	179	179	246	246
Sdb5	Japan	195	195	136	136	193	193	179	179	243	250
Sdb6	Japan	201	201	136	136	187	187	179	179	246	250
Sdb7	Japan	197	201	136	138	193	193	179	179	246	250
Sdb8	Japan	201	205	136	144	193	193	179	179	243	243
Sdb9	Japan	201	203	136	144	193	203	179	181	243	250
Sdb10	Japan	201	201	136	144	193	193	179	181	250	250
Sdb11	Japan	197	201	136	136	187	193	179	181	243	246
Sdb12	Japan	201	201	144	144	193	193	179	179	243	243
Sdb13	Japan	201	201	136	144	193	193	179	179	243	243
Sdb14	Japan	197	197	136	144	193	193	173	177	243	250
Sdb15	Japan	201	201	134	144	192	203	179	179	243	243
Sdb16	Japan	197	201	136	136	193	193	181	185	243	250
Sdb17	Japan	201	205	134	136	193	193	179	179	243	250
Sdb18	Japan	201	201	144	144	187	193	179	179	243	250
Sdb19	Japan	201	201	144	144	193	193	179	181	243	243
Sdd1	Azores	197	197	136	136	206	206	183	183	250	250
Sdd2	Azores	197	201	136	136	206	206	181	181	250	250
Sdd3	Azores	197	197	136	136	206	206	175	183	250	250
Sdd4	Azores	197	201	136	136	196	206	183	183	250	250
Sdd5	Azores	197	199	136	136	206	206	181	183	250	250
Sdd6	Azores	201	201	136	136	196	206	183	183	250	250
Sdd7	Azores	197	197	136	136	196	206	181	181	250	250
Sdd8	Azores	197	199	136	136	196	206	175	183	250	250
Sdd9	Azores	197	201	136	136	206	206	175	183	250	250
Sdd10	Azores	201	201	136	136	206	206	175	183	250	250
Sdd11	Azores	197	199	134	134	196	206	183	183	250	250
Sdd12	Azores	197	201	136	136	196	196	175	181	250	250
Sdd13	Azores	201	201	136	136	206	206	175	183	250	250
Sdd14	Azores	197	197	134	136	206	206	175	183	250	250
Sdd15	South Africa	201	201	136	144	192	206	175	181	250	250
Sdd16	South Africa	197	197	134	136	206	206	175	175	250	250
Sdd17	South Africa	195	197	136	136	206	206	175	177	243	250
Sdd18	South Africa	197	205	134	136	206	206	175	175	250	250
Sdd19	South Africa	197	197	136	136	206	206	175	181	243	250
Sdd20	South Africa	197	197	136	136	206	206	179	181	250	250
Sda1	Seychelles	197	201	136	144	206	206	179	181	243	250
Sda2	Seychelles	197	197	136	136	191	203	179	179	250	250
Sda3	Seychelles	201	201	136	144	193	206	175	175	243	250
Sda4	Seychelles	201	201	136	136	187	193	175	183	246	250
Sda6	Seychelles	197	201	136	144	193	206	179	181	250	250
Sda7	Seychelles	195	197	136	144	193	193	179	179	243	250
Sda8	Seychelles	197	197	136	144	206	206	175	181	243	250
Sda9	Seychelles	197	201	136	144	193	193	179	181	250	250
Sda10	Seychelles	195	201	136	144	187	187	179	179	250	250
Sda11	Seychelles	197	197	136	144	192	206	175	179	250	250
Sda12	Seychelles	201	205	136	144	192	206	179	181	243	250
IR1	Ireland	201	201	136	136	196	196	181	181	250	250
IR2	Ireland	197	201	136	136	196	206	181	181	250	250
IR3	Ireland	201	201	136	136	196	196	183	183	250	250
IR4	Ireland	197	199	136	136	196	206	183	183	250	250

IR5	Ireland	197	197	136	136	196	196	181	181	250	250
IR6	Ireland	201	201	136	136	196	206	181	183	250	250
IR7	Ireland	197	201	136	136	196	196	181	183	250	250
IR8	Ireland	197	197	136	136	196	196	175	175	250	250
IR9	Ireland	197	197	136	136	196	196	181	183	250	250
IR10	Ireland	197	201	136	136	196	196	183	183	250	250
PS1	United States	197	201	136	136	0	0	175	175	250	250
PS2	United States	197	201	136	136	196	196	175	175	250	250
PS3	United States	197	201	136	136	196	196	175	181	250	250
PS4	United States	197	197	136	136	196	196	181	181	250	250
PS5	United States	197	197	136	136	0	0	175	183	250	250
PS6	United States	197	201	136	136	206	206	181	181	250	250
PS7	United States	197	197	136	136	196	196	175	181	250	250
PS8	United States	201	201	136	136	196	196	181	183	250	250
PS9	United States	201	201	136	136	0	0	181	183	250	250
PS10	United States	197	197	136	136	196	196	175	181	250	250

Appendix 4: Alignment of individual RTS4 with the typical roseate tern (RT) ND2 and ND6 haplotypes, as well as those of the black-naped tern (BNT), and bridled and crested tern.

#NEXUS

[Created by Se-Al]

BEGIN DATA;

DIMENSIONS NTAX=13 NCHAR=1078;

FORMAT MISSING=? GAP=- DATATYPE=DNA;

MATRIX

RT normal ND2

TGACTCCTTCCCTTACTAATGAACCCACACGCTAAACTAATCACTTCC
CTAAGCCTACTCCTAGGTACTACTATCACAATTTCGAGCAACCATTG
GGCGATAGCTTGGACCGGACTAGAAATCAACACTCTCGCCATCATCC
CACTTATCTCAAAGTCGCATCACCCCCGGGCTATCGAAGCTGCAATC
AAATACTTCCTAGTCCAAGCAGCCGCCTCAGCCTTAGTTCTCTTCTCA
AGCATAACCAACGCATGACATACAGGACAATGAGATCTCACTCAATT
AGTCCACCCAACTTCATGCCTACTCCTAACAGCCGCAATTGGAATAA
AATTGGGATTAGTTCCATTTCATTTTTGATTCCCAGAAGTTCTCCAAG
GTTCATCCCTAACTATCGCCCTACTACTATCAACAGTAATAAAATTCC
CCCCAATCACTATCCTATTCATAACTTCACACTCCCTCAACCCAACAC
TGCTTACCTCCATGGCTCTTGCCTCAGCAGCCTTGGGTGGGTGAATAG
GATTAAACCAAACACAAGTCCGAAAAATCTTAGCCTTCTCCTCCATC
TCTCACCTAGGTTGAATAGTTATTATCATTATTTATAACCCAAAACTA
ACACTACTAACCTTCTACCTATATTCCCTAATAACAACTGCCGTTTTT
CTTACCTTAAACACAATCAAAGCTGTAAAGCTATCAACAATAATAAC
CTCATGAACAAAAACTCCTATACTTAATGCCGCTCTTATGTTAACGCT
ACTTTCACTAGCAGGACTTCCCCCACTCACAGGATTCCTACCTAAGTG
ACTCATTATCCAAGAATTAACTAAACAAGAAATAACTACAACAGCTA
CAATCATTGCCATACTATCTCTACTAGGATTATTCTTCTACCTTCGCC
TTGCATACTACTCAACAATCACGCTTCCACCAAATTCCACAAACCAT
ATAAAACAATGACACATCAATAAATCAACAAGCACCCCTATCGCCGT
TGTTGCCTCCCTATCAATCCTGCTACTACCTCTCTCCCCCATAATCCT
ATCTATCATTTAGAAACTTAGGATAGCCCTAA

RTS4 normal ND2

TGACTCCTTCCCTTACTAATGAACCCACACGCTAAACTAATCACTTCC
CTAAGCCTACTCCTAGGTACTACTATCACAATTTCGAGCAACCATTG
GGCGATAGCTTGGACCGGACTAGAAATCAACACTCTCGCCATCATCC
CACTTATCTCAAAGTCGCATCACCCCCGGGCTATCGAAGCTGCAATC
AAATACTTCCTAGTCCAAGCAGCCGCCTCAGCCTTAGTTCTCTTCTCA
AGCATAACCAACGCATGACATACAGGACAATGAGATCTCACTCAATT
AGTCCACCCAACTTCATGCCTACTCCTAACAGCCGCAATTGGAATAA
AATTGGGATTAGTTCCATTTCATTTTTGATTCCCAGAAGTTCTCCAAG

GTTCATCCCTAACTATCGCCCTACTACTATCAACAGTAATAAAATTCC
CCCCAATCACTATCCTATTCATAACTTCACACTCCCTCAACCCAACAC
TGCTTACCTCCATGGCTCTTGCCTCAGCAGCCTTGGGTGGGTGAATAG
GATTAAACCAAACACAAGTCCGAAAAATCTTAGCCTTCTCCTCCATC
TCTCACCTAGGTTGAATAGTTATTATCATTATTTATAACCCAAAACTA
ACACTACTAA----------

RTS4 unusual ND2 ----------

AACCCACACGCTAAACTAATCACTTCCCTAAGCCTACTCCTGGGTAC
TACTATCACAATTTCAAGCAACCATTGAGCGATAGCTTGGACCGGAC
TAGAAATCAATACTCTTGCCATCATCCCCCTTATCTCAAAATCACATC
ACCCCCGGGCTATCGAAGCCGCAATCAAATACTTCCTAGTCCAAGCA
GCTGCTTCAGCCCTAGTTCTCTTCTCAAGCATAACCAACGCATGACAC
ACAGGACAGTGAGATCTCACTCAGTTAATCCACCCAACTTCTTGCCT
ACTCCTAACAGCCGCAATTGGAATAAAATTGGGATTAGTCCCATTTC
ATTTTTGATTCCCAGAAGTCCTCCAAGGTTCATCCCTAACTACCGCCC
TACTACTATCAACAGTAATAAAATTCCCCCCAATCACTATCCTATTCA
TAACTTCACACTCCCTTAACCCAACACTGCTCACCTCCATAGCCCTTG
CCTCAGCAGCCTTGGGCGGGTGAATAGGATTAAACCAAACACAAGTC
CGAAAAATCTTAGCCTTCTCCTCCATCTCTCACCTAGGTTGAATAGTA
ATTATCATTATTTATAACCC----------

BNT ND2 ----------

TCACTTCCCTAAGCCTACTCCTGGGTACTACTATCACAATTTCAAGCA
ACCATTGAGCGATAGCTTGGACCGGACTAGAAATCAATACTCTTGCC
ATCATCCCCCTTATCTCAAAATCACATCACCCCCGGGCTATCGAAGC
CGCAATCAAATACTTCCTAGTCCAAGCAGCTGCTTCAGCCCTAGTTCT
CTTCTCAAGCATAACCAACGCATGACACACAGGACAGTGAGATCTCA
CTCAGTNAATCCACCCAACTTCTTGCCTACTCCTAACAGCCGCAATTG
GAATAAAATTGGGATTAGTCCCATTTCATTTTTGATTCCCAGAAGTCC
TCCAAGGTTCATCCCTAACTACCGCCCTACTACTATCAACAGTAATA
AAATTCCCCCCAATCACTATCCTATTCATAACTTCACACTCCCTTAAC
CCAACACTGCTCACCTCCATAGCCCTTGCCTCAGCAGCCTTGGGCGG
GTGAATAGGATTAAACCAAACACAAGTCCGAAAAATCTTAGCCTTCT
CCTCCATCTCTCACCTAGGTTGAATAGTAATTATCATTATTTATAACC
CAAAACTAACACTACTAACCTTCTACCTATATTCCCTAATAACAATTA
CCGTATTTCTTACCTTAAACACAATCAAAGCTGTAAAGCTATCAACA
ATAATAACCTCATGAACAAAAACTCCTATACTTAATGCCACTCTTAT
ATTAACTCTGCTTTCACTGGCAGGACTTCCCCCACTCACAGGATTCCT
ACCCAAGTGACTCATTATCCAAGAATTAACTAAACAAGAAATAACTA
CAGCAGCTACAGTCATTGCCATACTATCCCTACTAGGATTATTCTTCT
ACCTTCGCCTTGCATACTACTCAACAATCACACTTCCACCAAACTCCA

CAAACCATATAAAACAATGACACATCAATAAATCAACAGGCACTCCT
ATCGCCATTGTTGCCTCCCTATCAATCCTGCTACTACCTCTCTCCCCC
ATAATCCTAT----------

bridled tern ND2 ----------

TCACATCCCTAAGCCTACTCTTAGGCACAACCATTACAATTTCAAGC
AACCATTGAGTGATAGCTTGAACTGGATTAGAAATCAACACCCTTGC
TATTATCCCCCTTATCTCCAAATCGCACCACCCTCGGGCTATTGAAGC
TGTAATCAAATACTTCTTAGTTCAAGCAGCTGCTTCAACTTTAATTCT
CTTCTCAAGCATAATCAATGCATCATATACAGGACAATGAGATATCA
CCCAACTAACCCACCCAACTTCATGCTTACTTCTAACAGCTGCAATCG
GAATAAMACTAGGACTAGTCCCATTCCACTTCTGATTTCCAGAAGTA
CTTCAAGGTTCACCCCTAACTACTGCCCTACTACTATCAACAATGATA
AAATTCCCCCCAATCACCATTCTATTTATAACCTCACACTCCCTTGAC
CCAACACTACTAACCTCTATGGCCATCGCCTCAACAGCTCTGGGTGG
ATGAATGGGGTTAAACCAAACACAAATTCGAAAAATCCTAGCCTTCT
CCTCTATCTCTCACCTGGGCTGAATAGTAATCATCATTATCTACAACC
CAAAACTCACACTATTAACCTTTTACCTATACTCTTTAATAACTACCA
CTGTATTTCTAACCCTTAATACAATCAAAGTTATAAAACTATCAACA
ATAATAACCTCATGAACAAAAACCCCCATACTCAACGCTACTCTTAT
ACTAACCCTACTTTCACTAGCAGGGCTCCCCCCACTCACAGGATTCCT
ACCTAAATGACTCATCATTCAAGAATTAACTAAACAAGAAATAACTA
CAACAGCTACAATTATCGCTATATTATCCTTACTAGGATTATTCTTCT
ACCTTCGTCTTGCATACTACTCAACAATTACCCTCCCACCAAACTCCA
CAAACCACATAAAACAGTGACACATTAACAAATCAACAAACACCTTC
ATCGCCATTGTCACCTCTCTGTCAATCCTACTATTACCCCTCT----------
----

crested tern ND2 ----------

TGCTAAACTAATCACTTCCCTAAGCCTACTCCTAGGTTCTACCATCAC
AATTTCAAGCAACCATTGAGTAATAGCTTGAACCGGACTAGAGATCA
ATACTCTTGCCATCATCCCTCTTATCTCAAAATCACATCATCCTCGAG
CTATCGAAGCCGCAATCAAATACTTCCTAGTCCAAGCAGCCGCTTCA
ACCCTAGTTCTTTTCTCAAGTATAACTAACGCATGATATACAGGACA
ATGGGATCTCACCCAATTGGTCCACCCAACTTCATGTCTACTCCTAAC
AGCTGCAATTGGAATAAAATTAGGATTAGTTCCATTCCATTTTTGATT
CCCAGAAGTACTTCAGGGCTCATCCCTAGCTACCGCCCTACTACTATC
AACAGTAATGAAATTCCCCCCAATCACCATTCTATTCATAACTTCACA
CTCCCTCAACCCAACACTGCTTACCTCCATAGCCCTCGCCTCAGCAGC
CCTAGGTGGATGAATAGGATTAAACCAAACACAAGTCCGAAAAATC
CTAGCCTTCTCCTCCATCTCTCACCTGGGCTGAATAACAATCATCATT
ATTTATAGCCCAAAACTAACACTATTAACCTTCTACCTATATTCCCTA
ATAACAACTACCGTATTTCTTACCTTAAACACAATCAAAGCTGTAAA
GCTATCAACAATAATAACCTCATGAACAAAAACTCCCATACTTAATG
CTACTCTTATATTAACTCTACTTTCATTAGCAGGACTCCCCCCACTCA
CAGGATTCCTACCCAAGTGACTCATTATCCAAGAATTAACTAAACAA
GAAATAACTACAACAGCTACAATCATTGCTATACTATCCCTACTAGG
ATTATTCTTCTACCTTCGCCTTGCATACTACTCAACAATCACACTTCC
ACCAAACTCCACAAACCATATAAAACAATGACACATCAATAAATCA
ACAAGCACCCCTATCGCCGTTGTTACCTCCCTGTCAATCCTACTACTA
CCCCTCTCCCCCATAATCCT----------

RT normal ND6

CCCACGAGATAACCCCCGTACTAACTCTAACACAACAAACAAAGTCA
ACAACAATCCCCAACCCGCAACCAAAAACATCCCTACACCACACGA
ATAAAACATCGCTACCCCACCAAAATCTAACCGAACAGAAGACATG
CCCCCACAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAAC
CCCACTAACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCC
CATACCCAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCT
GCTGCCAAAGACACAGAATATACAAAAACCACCAATATCCCCCCCA
AATACACCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACT
CAACAACCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACT
ACCCCATAATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAA
CACAAAACACAACCCTAAAAAAAGCACAAAGTAAGTCATAGCAGTT
CCTGCTTGGCTTCTCTCCAAGACCCGCGGCCTGA----------

RTS4 normal ND6 -

CCACGAGATAACCCCCGTACTAACTCTAACACAACAAACAAAGTCAA
CAACAATCCCCAACCCGCAACCAAAAACATCCCTACACCACACGAAT
AAAACATCGCTACCCCACCAAAATCTAACCGAACAGAAGACATGCC
CCCACAATCCACCGTAACAACCCCAAGATCCCAACATTCAACAACCC
CACTAACAACTACCCCCACAACAAGCACCAAAACTAACCCTATCCCA
TACCCAGCAACCCGCCAATCCCCCCACGCCTCCGGAAAAGGGTCTGC
TGCCAAAGACACAGAATATACAAAAACCACCAATATCCCCCCCAAAT
ACAYCATAAACAGTACCAATGACACAAAAGAAACCCCCAAACTCAA
CAACCACCCACACCCTGCAATAGAAGCCAACACCAAACCAACTACCC
CATAATAAGGAGAAGGATTAGATGCAACTGCTAACCCTCCCAACACA
AAACACAACCCTAAAAAAAGCACAAAGTAAGTCATAGCAGTTCCTG
CTTGGCTTCTCTCCAAGACCCGCGGCCTGA----------

RTS4 unusual ND6

CCCACGAGATAACCCTCGTACTAATTCCAACACAACAAACAAAGTTA
ACAACAACCCCCAACCCGCAACCAAAAACATCCCCACACCATACGA
ATAAAACATCGCCACCCCACCAAAATCTAACCGAACAGAAGACATA
CCCCCACAATCCACCGTAACAACCCCAAGATCCCAACATCCAACAAC
CCCACTAACAGCTACCCCCACAACAAGCACCGAAACAAACCCC--------

BNT ND6 ------

AGATAACCCTCGTACTAATTCCAACACAACAAACAAAGTTAACAACA
ACCCCCAACCCGCAACCAAAAACATCCCCACACCATACGAATAAAA
CATCGCCACCCCACCAAAATCTAACCGAACAGAAGACATACCCCCAC
AATCCACCGTAACAACCCCAAGATCCCAACATCCAACAACCCCACTA
ACAGCTACCCCCACAACAAGCACCGAAACAAACCCCACCCCATATCC
AACAACCCGCCAATCTCCCCAAGCCTCAGGAAAAGGATCTGCTGCCA
AAGACACAGAGTATACAAAAACCACCAACATCCCCCCCAAATACAC
CATAAACAACACCAATGACACAAAAGAAACCCCCAAACTCAACAAC
CACCCACACCCTGCAATAGAAGCCGACACCAAACCAACCACCCCATA
ATAAGGAGAAGGATTAGATGCAACTGCTAACCCCCCTAACACAAAG
CACAACCCTAAAAAAAGCACAAAGTAAGTCATAACAGTTCCTGCTTG
GCTTCTCTCCAAGACCCGTGGCCTGA----------

---

bridled tern ND6 ----------

GCAGAAGATATACCCCCACTATCCACAGTAACAATCCCAGAATTTCA
ACACTCAATAACTCCACTAACAACCACCCCCACAATAAGTACCAAAA
CGAACCCTCCCCCATACCCCATAACCCGCCAATCCCCCCAAGTTTCC
GGAAAAGGATCCGCCGCCAAGGACACAGAGTACACAAAAACTACCA
GTATCCCCCCCCAAATATACCATGAACAGTACCAACGATACAAAAGAC
ACCCCCAAACTCAACAGCCACCCACACCCTGCAACAGAGGCTAACAC
TAAACCAACCACCCCATAATAAGGAGAAGGATTAGATGCAACTGCC
AAACCCCCTAGCACAAAACACAATC----------

---

crested tern ND6 ---

ACAGAAGACATACCCCCACAATCCACCGNAACAACCCCAAGATCCC
AACATTCAACAACCCCACTAACAACTACCCCCACAACAAGCACCAAA
ACAAACCCCATCCCATACCCAACAACCCGCCAATCCCCTCAAGCCTC
CGGAAAAGGGTCCGCCGCCAAAGACACAGAGTATACAAAAACCACT
AATATCCCCCCCAAGTACACCATAAATAGCACCAACGACACAAAAG
AAACCCCCAAACTCAACAACCACCCACATCCTGCAACAGAAGCTAAC
ACCAAACCAACCACCCCATAGTAGGGAGAAGGATTAGATGCAACTG
CCAAACCCCCCAACACAAAACATAACC---

;

END;