POPULATION DIFFERENTIATION OF SOUTHERN OCEAN SEABIRDS

DILINI K. ABEYRAMA

BSc. Honours, University of Colombo, 2017

A thesis submitted in partial fulfilment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

in

EVOLUTION AND BEHAVIOUR

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

POPULATION DIFFERENTIATION OF SOUTHERN OCEAN SEABIRDS

DILINI K. ABEYRAMA

Date of thesis defence: 11th April 2023

Dr Theresa Burg Thesis supervisor	Professor	Ph.D
Dr Cameron Goater Thesis examination committee member	Professor	Ph.D
Dr Robert Laird Thesis examination committee member	Professor	Ph.D
Dr Andrew Iwaniuk Internal examiner	Associate Professor	Ph.D
Dr Cathryn Abbott External Examiner Pacific Biological Station Fisheries and Oceans Canada Canada	Research Scientist	Ph.D
Dr Steve Wiseman Chair, Thesis Examination Committee	Associate Professor	Ph.D

GENERAL ABSTRACT

The Southern Ocean is a remote but unique ecosystem with high winds, strong currents, and a handful of islands surrounding the Antarctic continent. Reduced gene flow due to these physical and non-physical barriers supports rapid evolution and endemism within the Southern Ocean. Seabirds are a good model to study barrier-mediated speciation as they face a limited number of physical barriers, yet they are a highly diversified group. In my thesis, I used molecular markers to study population differentiation in five Southern Ocean seabird species at the three levels: among ocean basins, within oceans and within a single island. Sooty albatross (*Phoebetria fusca*) and yellow-nosed albatross showed population differentiation between Atlantic and Indian Ocean basins. Two sister species of yellow-nosed albatross, Atlantic (*Thalassarche chlororhynchos*) and Indian (*Thalassarche carteri*), both showed population genetic structure within Atlantic and Indian Oceans, respectively. The other two study species, Kerguelen shags (*Phalacrocorax verrucosus*) and gentoo penguins (*Pygoscelis papua*) breeding on Kerguelen Island, showed genetic structure among different breeding colonies of each species on the same island. Non-physical barriers such as natal philopatry and at-sea distribution, are limiting gene flow in the Southern Ocean at different geographic scales.

THESIS ACKNOWLEDGEMNTS

First and foremost, I want to thank my supervisor Dr Theresa Burg for everything she has done for me. Her guidance, from day one when I wrote to her looking for a graduate student position up until now, is enormous. If it wasn't for her countless emails and random stopovers at our grad office, I would have missed several deadlines and would still be staring at my computer not knowing that I've been enjoying my grad life for too long. Thank you so much for believing in me and encouraging me to do my best. During the times my stupidity came to play, I survived thanks to your patience. Thank you so much for being a person with whom I can share not only my research and academic interests, but also my personal struggles. Also, I am thankful for all the sweet treats you brought us frequently, they were very helpful to get through long and tiring days. Thank you so much for taking me to where I am today and shaping me to who I am today.

I am also very grateful to my supervisory committee, Dr Robert Laird and Dr Cameron Goater for their great support throughout my project. Thank you very much for your feedback, guidance, questions and encouragement during every committee meeting as well as when I see you in the building. I am also thankful to Dr Andrew Iwaniuk, Dr Steve Wiseman, Dr Matt Bogard and Dr Greg Pyle for their valuable feedback and great support during my transfer and comprehensive exams. I am thankful to Dr Alice Hontela and Dr Roy Golsteyn for their feedback and advice during my seminar course at the University of Lethbridge in my first year. Special thanks to Dr Brendan Graham, a former postdoc fellow in Burg Lab for answering all my stupid questions, saving me countless times, teaching me and encouraging me all these years. Without the guidance and help of

my undergrad supervisor, I would not be here today. Dr Sampath Seneviratne, I am indebted you and thank you so much for showing me the right path.

A big thank you to Danika Schramm and Kanishka Senevirathna, my former labmates and partners in crime for rooting for me and supporting me though my thick and thin. Thank you very much my current labmates Thilini, Aaron, and Bryan for your support and making my grad life happier. I also want to thank current postdoc fellows in our lab, Dr Linda Lait and Dr Juan Enciso for their valuable feedback and guidance. Without my former labmates, Bukola, Ian and Amanda, it wouldn't be easy. Thank you, guys, for all your support. Without a good teaching coordinator, my teaching assistant experience at the University of Lethbridge would not have been easy. Thank you very much Quintin Steynen and all my colleague TAs for your help and making my TA experience interesting.

Moving to the other side of the world to an unknown country all by myself and settling here in Canada after finishing my undergrad in Sri Lanka was not easy. If it is not for Himani aunty, Densil uncle, Soba akka and Chinthaka ayya, I would not have survived these 5 years. Their love and support gave me strength to adapt to this new life. Also thank you so much Dilshadi, Supun, Dilumi akka Asha akka and Akalanka ayya for your support. Dion my little boy, thank you for being my bundle of joy. Thank you all for being my second family.

My heartfelt gratitude goes to all our collaborators, Dr Peter Ryan, Dr Henri Weimerskirch, Dr Thierry Boulinier, Dr Yves Cherel, Dr Michelle Risi, Dr Maelle Connan, Dr Charles-André Bost, Belinda Bauer and Tasmanian Museum and Art Gallery for providing us samples. Thank you very much Dr Brian Boyle and Dr Kyle Wellband

for your technical support and valuable feedback. Without the financial support for this project as well as for my stipend, this project would not have gone smooth. Thank you very much Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship, Alberta Society of Professional Biologists, Mitacs Research Training Award, and the Seabird Group for funding my project and my stipend. I am also very grateful to the University of Lethbridge for accepting me as a graduate student and giving me a tuition fee scholarship, without which I would not have been able to complete my degree. I also want to thank Imaru Baquero, Karis Dykstra and Claire Brodrick from the International Centre, University of Lethbridge for their great support throughout these years and for always answering my questions. Araba Pee, my mentor, very first friend in Lethbridge and my sister, thank you so much for your advice, guidance, love and especially for your friendship.

Last not least my family, to whom I am forever indebt. I have no words to thank my mom and dad for all the struggles they went through to make my life successful. I am forever grateful to all the love, strength, guidance and literally everything you have given and are giving me. Thank you so much for raising me to be strong, independent and human. My siblings Ruwanthika, Vindyani, Gayan and their partners Ravin, Tharaka and Piyumi, thank you so much for loving me, making me strong and always being there for me. Thank you for raising me up. I am also thankful to my niece and nephews Kuwi, Nuha and Buwi for making my life happier and lighting my days up. Namith Gunasekara my husband, thank you so much for being in my life. Thank you so much for encouraging me, cooking for me, taking care of me, loving me and especially for bearing up with my

stress and drama. Also, I want to thank my relatives and all friends for their love and support. Thank you, my coworkers and those random customers I met during my parttime job at Subway for your encouraging words and support. I am utterly grateful to all of you and you all will forever be in my heart.

PREFACE

The thesis present original research work completed by Dilini Abeyrama in the Department of Biological Sciences of the University of Lethbridge, Alberta, Canada under the supervision of Dr Theresa M. Burg.

This thesis is based on four chapters arranged in the form of manuscripts accompanied by general introduction and discussion chapters. The first data chapter focuses on the population differentiation of sooty albatrosses breeding in Atlantic and Indian Ocean basins. This chapter is a combination of analyses of molecular, morphological, and biochemical data. I collected and analysed all molecular data. I wrote the manuscript with input from others. I am one of the co-first authors of this manuscript. In the second data chapter I looked at the genetic structure between Atlantic and Indian yellow-nosed albatrosses using two different types of molecular markers and was published on Conservation Genetics in 2021. I completed all data collections and analyses. I wrote the manuscript along with Dr Burg and input from other coauthors. In the third data chapter, I looked at the population genetic structure within the two yellow-nosed albatross species using next-generation sequencing data. For this chapter also, I collected and analysed all the data. Dr Burg and I wrote the manuscript with input from other two coauthors. In my last data chapter, I looked at the intraspecific genetic structuring in the two seabird species: Kerguelen shags and gentoo penguins breeding on the Kerguelen Island. For this chapter, I collected and analysed all data. Similar to other data chapters, I wrote the manuscript with Dr Burg and input from other coauthors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS, ACRONYMS AND SYMBOLS

A	Adenine
ACAP	Agreement on the Conservation of Albatrosses and Petrels
ACC	Antarctic Circumpolar Current
AMOVA	Analysis of molecular variance
APF	Antarctic polar front
Arp	Arlequin file format (file type)
B-H	Benjamini-Hochberg correction
C	Cytosine
CNES	National Centre for Space Studies
COI	Cytochrome oxidase I
ddRAD	Double digest restriction-site associated DNA sequencing
Df	Degree of freedom
DNA	Deoxyribonucleic acid
DnaSP	DNA sequence polymorphism (software)
dNTP	Deoxyribonucleotide triphosphate
ESU	Evolutionary significant unit
F	Forward primer
F_{ST}	Fixation index; a measure of population differentiation
G	Guanine
GEBCO	General Bathymetric Chart of the Oceans
GEPE	Gentoo penguin
Gff	General feature format (file type)
GLMM	Generalized linear mixed model
GLS	Global location sensing (geolocators)
H_E	Expected heterozygosity
H_O	Observed heterozygosity
HWE	Hardy-Weinberg equilibrium
IBIS	Institute of Integrative Biology and Systems
IUCN	International Union for Conservation of Nature
K	Number of clusters
KESH	Kerguelen shag
km	Kilometers
KRFo	Kerguelen Ratmanoff open-sea colony
LD	Linkage disequilibrium
LDEO	Lamont-Doherty Earth Observatory
M	Million
M13	Universal primer sequence tag
MEGA	Molecular Evolutionary Genetics Analysis (software)

$MgCl_2$	Magnesium chloride
mM	Millimolar
mtCR	Mitochondrial control region
mtDNA	Mitochondrial DNA
MU	Management unit
N	Number of samples
N_2	Atmospheric nitrogen
NA	Not applicable
Ng	Nanograms
NGA	National Geospatial-Intelligence Agency
NGS	Next-generation sequencing
Nm	Nanometers
NOAA	National Oceanic and Atmospheric Administration
NSERC	Natural Science and Engineering Research Council
NSF	National Science Foundation
P	Probability value
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
R	Reverse primer
R^2	Coefficient of determination
RADSeq	Restriction-site associated DNA sequencing
ROI	Region of interest
RRS	Reduced representation sequencing
SAF	Sub-Antarctic front
SEA	Standard ellipse area
SEA_B	Bayesian standard ellipses area
SIBER	Stable isotope Bayesian ellipses
SIO	Scripps Institution of Oceanography
SNP	Single nucleotide polymorphism
SOAL	Sooty albatross
STF	Subtropical front
T	Thymine
T_1	Primer annealing temperature one
T_2	Primer annealing temperature two
TCS	Templeton Crandall and Sing method (software)
TCSBU	TCS Beautifier (software)
U.S.	United States
v	Version
Vcf	Variant call format (file type)
WGS	Whole genome sequencing

YNAL	Yellow-nosed albatross
X	Number of times
$\delta^{13}C$	Carbon 13 stable isotope
$\delta^{15}N$	Nitrogen 15 stable isotope
ΔK	Delta K (Structure harvester value)
°C	Degree Celsius
°E	Degrees east of the meridian
°S	Degrees south of the equator
°W	Degrees west of the meridian
µl	Microliters
µM	Micromolar
^{12}C	Carbon 12 isotope
^{14}N	Nitrogen 14 isotope

Abbreviations for populations

AI	Amsterdam Island
CI	Crozet Island
GI	Gough Island
II	Inaccessible Island
KI	Kerguelen Island
NI	Nightingale Island
NZB	New Zealand bycatch
PEI	Prince Edward Islands
SAfB	South African bycatch
SI	Stoltenhoff Island
SPI	Saint Paul Island
TDC	Tristan da Cunha Island

Chapter 1: General introduction

1.1 Speciation and seabirds

There are many definitions of what a species is, but it is essentially a unique evolutionary unit. New species can arise because of isolation due to physical and non-physical barriers, natural selection or genetic drift and sometimes secondary contact or hybridization. Speciation can be allopatric (two or more populations become reproductively isolated due to geographical barriers), parapatric (individuals in a large continuous population start inbreeding locally eventually forming two reproductively isolated groups), peripatric (a subset of individuals separate and colonize a new habitat at the edge of their original population, and evolve independently) or sympatric (two populations of the same species inhabiting the same geographic area evolve independently and eventually become reproductively isolated) (Mayr, 1942, 1963; Coyne & Orr, 2004). Although allopatric speciation is believed to be the most common type of speciation, it does not apply to groups like seabirds. Seabirds are highly vagile and face only limited number of physical barriers, yet they are a very diversified group.

Seabirds spend more than half of their lives at sea and visit land only for breeding. According to Friesen et al. (2007), parapatric and sympatric speciation are more common in seabirds. Physical isolation, philopatry, differences in ocean regime, non-breeding distribution, foraging distribution during the breeding season and breeding phenology act as barriers to gene flow in seabirds.

Natal philopatry or returning to their birth site to breed is one of the most common genetic barriers in many seabird species (Coulson, 2016). Seabirds are capable of dispersing thousands of kilometres, yet show genetic structure between islands less

than 200 km apart, and this is mainly due to natal philopatry (Inchausti & Weimerskirch, 2002; Burg & Croxall, 2004; Danckwerts et al., 2021; Herman et al., 2022). Even when natal philopatry is high, it is still possible to see genetic divergence within a site due to individuals breeding at different times of the year (Power et al., 1986; Taylor et al., 2019). For example, Taylor et al. (2019) found genetic differentiation between hot and cold season breeding populations of band-rumped storm petrels (*Hydrobates* spp.) on the Azores. At-sea distribution is also an important genetic barrier in seabirds (Burg & Croxall, 2001; Pertierra et al., 2020). Burg and Croxall (2001) found there is a higher chance of gene flow between breeding populations if the birds are utilizing a common foraging ground. Therefore, in seabirds, the gene flow is mostly limited by non-physical barriers.

According to IUCN red list, about 28% of 346 seabird species are globally threatened (Croxall et al., 2012). To provide an effective conservation for this amazing group of animals, it is required to understand their actual diversity. The study of evolution and speciation of seabirds is critical not only for the betterment of these animals themselves, but also for the ecosystem as they are apex-predators and can be used as indicators of status of the marine health. Therefore, as a first step to protect the global oceans which cover more than 70% of the earth's surface, these birds need to be conserved and managed properly. The Southern Ocean is a region with high seabird diversity compared to other oceans of the world and it provides a very good system to study seabird speciation (Munro and Burg, 2017).

1.2 Southern Ocean

The Southern Ocean is made up of southern parts of Pacific, Atlantic and Indian Oceans. In the Southern Ocean, there are few large islands of continental origin and many smaller islands of volcanic origin (Bergstrom & Chown, 1999). The Southern

Ocean islands provide breeding grounds for many species including albatrosses, petrels, cormorants and seals. Biodiversity of the Southern Ocean is mainly affected by the Antarctic Circumpolar Current (ACC) which is one of the largest currents. ACC flows eastwards around Antarctica and isolates the Southern Ocean from the warmer northern waters (Bergstrom & Chown, 1999). Distinct hydrological regions create lateral temperature and salinity gradients within the Southern Ocean (Dong et al., 2006). Unique environmental conditions and geographical features of the region have resulted in high levels of endemism within the Southern Ocean (Griffiths, 2010). Among these, some groups are endemic to specific ocean basins or islands (Harrison et al., 2013; Jackson et al., 2014; Abeyrama et al., 2021). While most of the marine species in this region face a limited number of physical barriers to gene flow or are capable of easily overcoming these barriers, the gene flow between different breeding populations is restricted mostly by non-physical barriers resulting in high levels of speciation and endemism (Friesen et al., 2007; Jackson et al., 2014; Munro & Burg, 2017). Sometimes, population divergence leading to speciation is not readily detectable through phenotypic markers. Therefore, these populations need to be analysed with high resolution molecular markers to study ongoing divergence.

1.3 Molecular markers

Speciation is a continuous process and differences accumulating in populations may not always be readily detectable. Therefore, we can use genetic markers to determine if populations are genetically isolated and identify the drivers of this genetic divergence (Karl et al., 2012). Molecular markers include DNA and can be either nuclear or organellar (Al-Samarai & Al-Kazaz, 2015). Molecular markers such as single nucleotide polymorphisms (SNPs) and microsatellites are used to study evolution (Vignal et al., 2002).

1.3.1 Mitochondrial DNA

Mitochondrial DNA (mtDNA) is a common molecular marker used to study populations and speciation. The popularity of mtDNA is due to its rapid evolutionary rate, maternal inheritance, lack of recombination, ease of use and amplification (Galtier et al., 2009; Ballard & Pichaud, 2013). The rapid evolutionary rate in mtDNA makes it a good marker to detect recent genetic divergence or differences among and within populations (Harrison, 1989). For example, a high-resolution DNA barcoding database has been constructed for European butterflies using 22,306 cytochrome oxidase I (COI) mtDNA sequences. This database covers 62% of total haplotype diversity from 97% of European butterfly species (Dincă et al., 2021). Although mtDNA is a good molecular marker, it needs to be used with caution as mtDNA does not represent the nuclear genome and features such as differential introgression, incomplete lineage sorting or selection may lead to discrepancies between patterns shown by mtDNA and nuclear markers.

1.3.2 Microsatellite markers

Microsatellites are simple repeating DNA motifs with a 1- 6 nucleotide core sequence. The variation in the number of repeats at a given locus results in microsatellite polymorphism (Liu & Cordes, 2004). These polymorphic regions of the genome are amplified using polymerase chain reactions (PCR) and the size of the amplified DNA fragments is used to differentiate individuals. Microsatellite markers are codominant, highly variable, and have a higher mutation rate than other nuclear markers (Abdul-Muneer, 2014). Microsatellite markers can be used not only to study population structure, but also for pedigree analyses, functional genomics, forensic studies, genome mapping and disease diagnostics. Although microsatellites are good markers to identify population differentiation, studies using them often examine < 20 loci, and

therefore do not provide the same level of genome coverage to detect loci under selection.

1.3.3 Single nucleotide polymorphism

A single nucleotide polymorphism (SNP) arises due to a point mutation on a specific locus (Yang et al., 2012). SNPs are the most abundant type of DNA variation found within organisms and can occur both in coding and non-coding regions of the genome. Using advanced high-throughput sequencing technologies like next-generation sequencing (NGS), it is possible to genotype thousands of SNPs. The number of SNPs genotyped can be adjusted by selecting different sequencing protocols such as reduced representation sequencing (RRS; Good, 2011) to genotype SNPs only within selected regions of the genome or whole genome sequencing (WGS; Ng & Kirkness, 2010) to look at SNPs across the complete genome. In restriction-site associated DNA sequencing (RADSeq), genome coverage and the number of SNPs genotyped can be adjusted using different restriction enzyme combinations or sequencing technologies (Miller et al., 2007; Abeyrama et al., 2022). The type of the sequencing protocol used depends on the study question and the budget available. By looking at the location of SNPs, especially within the coding region, it is possible to detect processes such as selection and local adaptation to identify the drivers of genetic variation.

1.4 Study species

Seabirds are a good model system to study non-physical barrier-mediated speciation as they face a limited number of physical barriers, yet they are a highly diversified group. The family Diomedeidae is no exception. This group of Procellariiformes, tube-nosed birds, has 21-24 currently identified species (Robertson & Nunn, 1998). This family is composed of four genera: *Diomedea* (great albatrosses), *Phoebastria* (Northern Pacific albatrosses), *Phoebetria* (sooty albatrosses) and *Thalassarche*

(mollymawks) (Onley & Scofield, 2013) and are currently found throughout all oceans except the North Atlantic.

The focus of the first three data chapters of my thesis are three species of the two genera, *Phoebetria* and *Thalassarche* which show intraspecific morphological and behavioural differences at macrogeographic scales. The three species, sooty albatross (*Phoebetria fusca*), Atlantic yellow-nosed albatross (*Thalassarche chlororhynchos*), and Indian yellow-nosed albatross (*Thalassarche carteri*) are listed as endangered by International Union for Conservation of Nature (IUCN) and the global population sizes are continuing to decline (ACAP, 2009a; ACAP, 2009b; ACAP, 2010). Conservation actions such as eradication of invasive mammals, managing the longline fishing and establishing systematic monitoring schemes have been implemented for all three species. But the scarcity of genetic data for these species limits the effectiveness of the conservation and management plans.

The fourth data chapter of my thesis focuses on genetic differentiation of seabirds at microgeographic levels. In this chapter I examine the genetics of two seabirds, Kerguelen shag (*Leucocarbo verrucosus*) and gentoo penguin (*Pygoscelis papua*), breeding on Kerguelen Island in the Southern Ocean. Both species breeding on the Kerguelen Island have two ecotypes; open-sea and closed-sea based on their foraging range at-sea. IUCN has listed both species as least concern, but if genetic data show that birds from each ecotype on the Kerguelen Island are distinct, new conservation plans will need to be implemented to protect these evolutionarily significant populations.

1.4.1 Sooty albatross

Sooty albatross is a medium sized albatross with a sooty brown complexion. Its broken white eye-ring, yellowish-pink bill stripe, long but slender wings and tapered

tail distinguish it from other Procellariiformes. Sooty albatrosses breed only on a few islands in south Atlantic and south Indian Oceans (Figure 1.1). Recent observations suggest that the bill stripe colour of birds breeding on south Atlantic Islands is bright yellow while it is cream/white in birds breeding on south Indian Islands (Figure 1.2). Although populations in the two ocean basins show some morphological differences, they are considered a single species (ACAP, 2010).

1.4.2 Yellow-nosed albatross

Atlantic and Indian yellow-nosed albatross look similar in many aspects, but they can be differentiated based on the head and neck plumage and the bill colouration (Figures 1.3 and 1.4). In adult Atlantic yellow-nosed albatross, the hood covering the head and hind-neck is light grey coloured while the forehead and the fore-crown are much paler (Figure 1.4A). In the Indian yellow-nosed albatross, the head and neck are white while there is a light grey shade on face and cheeks (Figure 1.4B). The eye patch in the Atlantic yellow-nosed adult is more triangular whereas in the Indian yellow-nosed adult, the eye patch is rounded and less clear (Figure 1.4) (Flood, 2016). The two species also show some differences in behaviour, breeding range and at-sea distribution (ACAP, 2009a, 2009b). The taxonomy of the two species is controversial; Brooke (2004) lists them as two species while Penhallurick & Wink (2004) list them as a single species. Little is known about the genetics of either species to support either claim (Nunn et al., 1996; Nunn & Stanley, 1998).

The breeding range of Atlantic yellow-nosed albatross is highly restricted to the central south Atlantic Ocean and they breed on the five islands; Tristan da Cunha Island (TDC), Nightingale Island (NI), Inaccessible Island (II), Stoltenhoff Island (SI), and Gough Island (GI) (Figure 1.5). Indian yellow-nosed albatross breed on the four

French subantarctic islands; Crozet (CI), Kerguelen (KI), Amsterdam (AI), Saint Paul (SPI), and one South African island, Prince Edward (PEI) Islands (Figure 1.5).

1.4.3 Kerguelen shags and gentoo penguins

While seabirds such as albatrosses show genetic differentiation across their breeding range (macrogeographic), some other species such as shags and penguins show microgeographic differentiation across breeding sites which are only a few kilometers apart on a single island (Bost et al., 1992; Cook et al., 2013).

The Kerguelen archipelago is composed of a main island, Kerguelen Island, and about 300 small satellite islands. The main island and all the small islands cover a combined area of ~7,000 km^2 (Blain et al., 2001). On Kerguelen Island, the complex landscape with jagged open coasts and sheltered bays creates different oceanographic conditions. Lescroël & Bost (2005) found individuals from colonies on Kerguelen form two different types of colonies: open-sea and closed-sea. Individuals breeding in colonies facing the closed-sea forage in shallow water while individuals breeding in colonies facing open-sea are foraging offshore in deeper waters.

The Kerguelen shag is endemic to Kerguelen Island while gentoo penguin breeds on islands across Southern Ocean including Kerguelen Island. Both Kerguelen shags and gentoo penguins breeding on the island have both open-sea and closed-sea colonies. Both species show similar and stable morphological and behavioural differences corresponding to the two ecotypes regardless of their colonies being relatively close to each other and located on the same island (Bost et al., 1992; Cook et al., 2013). Shags and gentoo penguins from closed-sea colonies are ~60% and ~36% lighter in weight respectively than individuals of the same species from open-sea colonies.

1.5 Thesis overview

This thesis is organized into six chapters. Chapter one provides a general introduction to speciation, genetic barriers in seabirds, molecular markers and introduces the study species. Chapter two examines the population genetic structuring of sooty albatross breeding in southern Atlantic and Indian Oceans. The genetic analyses for the species in this chapter are combined with morphological and biochemical analyses assessing the taxonomy of sooty albatross breeding in two ocean basins. The third chapter examines the population differentiation among and within the two species of yellow-nosed albatross and has been published in Conservation Genetics (Abeyrama et al., 2021). In this chapter, more conventional molecular markers such as mtDNA and nuclear markers including microsatellites are used. Chapter four investigates the population genetic structuring among and within Atlantic and Indian yellow-nosed albatross using high throughput sequencing data from NGS. The two main aims of this chapter are to detect the fine-scale genetic differentiation especially within the two species and to identify the drivers of genetic divergence in the two species. The final data chapter, chapter five, investigates the microgeographic genetic differentiation in two species of seabirds breeding on the Kerguelen Island in the southern Indian Ocean. RADSeq data are used in this chapter to identify the drivers of microgeographic genetic differentiation on the Kerguelen Island. The final chapter summarizes the findings of the four data chapters (Chapters 2-5), discusses the role of non-physical barriers in population genetic differentiation in Southern Ocean seabirds at macro and microgeographic scales and suggests future work on these amazing study systems.

1.6 References

Abdul-Muneer, P. M. (2014). Application of microsatellite markers in conservation genetics and fisheries management: Recent advances in population structure analysis and conservation strategies. *Genetics Research International*, *2014*, 1–11. https://doi.org/10.1155/2014/691759

Abeyrama, D. K., Boyle, B., & Burg, T. M. (2022). Comparison of genotyping by sequencing procedures to determine population genetic structure. *Functional & Integrative Genomics*, *23*, 9. https://doi.org/10.1007/S10142-022-00929-6

Abeyrama, D. K., Dempsey, Z. W., Ryan, P. G., & Burg, T. M. (2021). Cryptic speciation and population differentiation in the yellow-nosed albatross species complex. *Conservation Genetics*, *22*, 757–766. https://doi.org/10.1007/s10592-021-01358-x

Al-Samarai, F. R., & Al-Kazaz, A. A. (2015). Molecular Markers: an Introduction and Applications. *European Journal of Molecular Biotechnology*, *9*, 118–130. https://doi.org/10.13187/ejmb.2015.9.118

Agreement on the Conservation of Albatrosses and Petrels (2009a). ACAP Species assessment: Indian yellow-nosed albatross *Thalassarche carteri.*

Agreement on the Conservation of Albatrosses and Petrels (2009b). ACAP Species assessment: Atlantic yellow-nosed albatross *Thalassarche chlororhynchos*.

Agreement on the Conservation of Albatrosses and Petrels (2010). ACAP Species assessments: sooty albatross *Phoebetria fusca*.

Ballard, J. W. O., & Pichaud, N. (2014). Mitochondrial DNA: more than an evolutionary bystander. *Functional Ecology*, *28*, 218–231. https://doi.org/10.1111/1365-2435.12177

Bergstrom, D. M., & Chown, S. L. (1999). Life at the front: history, ecology and change on Southern Ocean islands. *Trends in Ecology & Evolution*, *14*, 472–477.

Blain, S., Tréguer, P., Belviso, S., Bucciarelli, E., Denis, M., Desabre, S., Fiala, M., Martin Jézéquel, V., Le Fèvre, J., Mayzaud, P., Marty, J. C., & Razouls, S. (2001). A biogeochemical study of the island mass effect in the context of the iron hypothesis: Kerguelen Islands, Southern Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, *48*, 163–187. https://doi.org/10.1016/S0967-0637(00)00047-9

Bost, C. A., Jouventin, P., & Sel, N. P. du. (1992). Morphometric variability on a microgeographical scale in two inshore seabirds. *Journal of Zoology*, *226*, 135–149. https://doi.org/10.1111/J.1469-7998.1992.TB06132.X

Brooke, de L. M. (2004) Albatrosses and petrels across the world. Oxford University Press, Oxford.

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

Burg, T. M., & Croxall, J. P. (2004). Global population structure and taxonomy of the wandering albatross species complex. *Molecular Ecology*, *13*, 2345–2355.

https://doi.org/10.1111/j.1365-294X.2004.02232.x

Cook, T. R., Lescroël, A., Cherel, Y., Kato, A., & Bost, C. A. (2013). Can foraging ecology drive the evolution of body size in a diving endotherm? *PLoS One*, *8*, e56297. https://doi.org/10.1371/JOURNAL.PONE.0056297

Coulson, J. C. (2016). A review of philopatry in seabirds and comparisons with other waterbird species. *Waterbirds*, *39*, 229–240. https://doi.org/10.1675/063.039.0302

Coyne, J. A. Orr H. A. 2004. Speciation. Sunderland, MA: Sinauer. 545 pp.

Croxall, J. P., Butchart, S. H., Lascelles, B. E. N., Stattersfield, A. J., Sullivan, B. E. N., Symes, A., & Taylor, P. H. I. L. (2012). Seabird conservation status, threats and priority actions: a global assessment. *Bird Conservation International*, *22*(1), 1–34.

Danckwerts, D. K., Humeau, L., Pinet, P., McQuaid, C. D., & Le Corre, M. (2021). Extreme philopatry and genetic diversification at unprecedented scales in a seabird. *Scientific Reports*, *11*, 1–12. https://doi.org/10.1038/s41598-021-86406-9

Dincă, V., Dapporto, L., Somervuo, P., Vodă, R., Cuvelier, S., Gascoigne-Pees, M., Huemer, P., Mutanen, M., Hebert, P. D. N., & Vila, R. (2021). High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity. *Communications Biology*, *4*, 1–11. https://doi.org/10.1038/s42003-021-01834-7

Dong, S., Sprintall, J., & Gille, S. T. (2006). Location of the Antarctic polar front from AMSR-E satellite sea surface temperature measurements. *Journal of Physical Oceanography*, *36*, 2075–2089.

Friesen, V. L., Burg, T. M., & McCoy, K. D. (2007). Mechanisms of population differentiation in seabirds: Invited review. *Molecular Ecology*, *16*, 1765–1785. https://doi.org/10.1111/j.1365-294X.2006.03197.x

Galtier, N., Nabholz, B., GlÉmin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology*, *18*, 4541–4550. https://doi.org/10.1111/J.1365-294X.2009.04380.X

Good, J. M. (2011). Reduced representation methods for subgenomic enrichment and next-generation sequencing. *Methods in Molecular Biology*, *772*, 85–103. https://doi.org/10.1007/978-1-61779-228-1_5/COVER

Griffiths, H. J. (2010). Antarctic marine biodiversity – What do we know about the distribution of life in the Southern Ocean? *PLoS One*, 5, e11683. https://doi.org/10.1371/JOURNAL.PONE.0011683

Harrison, R. G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends in Ecology and Evolution*, *4*, 6–11. https://doi.org/10.1016/0169-5347(89)90006-2

Harrison, P., Sallaberry, M., Gaskin, C. P., Baird, K. A., Jaramillo, A., Metz, S. M., Pearman, M., Michael O'Keeffe, Dowdall, J., Enright, S., Fahy, K., Gilligan, J., & Lillie, G. (2013). A new storm-petrel species from Chile. *The Auk*, *130*, 180–191. https://doi.org/10.1525/AUK.2012.12071

Herman, R. W., Winger, B. M., Dittmann, D. L., & Harvey, M. G. (2022). Fine-scale population genetic structure and barriers to gene flow in a widespread seabird (*Ardenna pacifica*). *Biological Journal of the Linnean Society*, *137*, 125–136. https://doi.org/10.1093/BIOLINNEAN/BLAC091

Inchausti, P., & Weimerskirch, H. (2002). Dispersal and metapopulation dynamics of an oceanic seabird, the wandering albatross, and its consequences for its response to long-line fisheries. *Journal of Animal Ecology*, *71*, 765–770. https://doi.org/10.1046/J.1365-2656.2002.00638.X

Jackson, J. A., Steel, D. J., Beerli, P., Congdon, B. C., Olavarría, C., Leslie, M. S., Pomilla, C., Rosenbaum, H., & Scott Baker, C. (2014). Global diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20133222. https://doi.org/10.1098/RSPB.2013.3222

Karl, S. A., Toonen, R. J., Grant, W. S., & Bowen, B. W. (2012). Common misconceptions in molecular ecology: Echoes of the modern synthesis. *Molecular Ecology*, *21*, 4171–4189. https://doi.org/10.1111/j.1365-294X.2012.05576.x

Lescroël, A., & Bost, C. A. (2005). Foraging under contrasting oceanographic conditions: The gentoo penguin at Kerguelen Archipelago. *Marine Ecology Progress Series*, *302*, 245–261. https://doi.org/10.3354/MEPS302245

Liu, Z. J., & Cordes, J. F. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*, *238*, 1–37. https://doi.org/10.1016/j.aquaculture.2004.05.027

Mayr, E. 1942. Systematics and the Origin of Species. New York: Columbia Univ. Press. 382 pp.

Mayr, E. 1963. Animal Species and Evolution. Cambridge, MA: Belknap. 811 pp.

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, *17*, 240–248. https://doi.org/10.1101/GR.5681207

Munro, K. J., & Burg, T. M. (2017). A review of historical and contemporary processes affecting population genetic structure of Southern Ocean seabirds. *Emu - Austral Ornithology*, 11, 4–18. https://doi.org/10.1080/01584197.2016.1271988

Ng, P. C., & Kirkness, E. F. (2010). Whole genome sequencing. *Methods in Molecular Biology*, *628*, 215–226. https://doi.org/10.1007/978-1-60327-367-1_12/COVER

Nunn, G. B., Cooper, J., Jouventin, P., Robertson, C. J. R., & Robertson, G. G. (1996). Evolutionary relationships among extant albatrosses (Procellariiformes: Diomedeidae) established from complete cytochrome-b gene sequences. *The Auk*, *113*, 784–801. https://doi.org/10.2307/4088857

Onley, D., & Scofield, P. (2013). *Albatrosses, Petrels and Shearwaters of the World - Derek Onley, Paul Scofield - Google Books*. Bloomsbury Publishing, 2013.

Penhallurick, J., & Wink, M. (2004). Analysis of the taxonomy and nomenclature of

the Procellariiformes based on complete nucleotide sequences of the mitochondrial cytochrome b gene. *Emu*, *104*, 125–147. https://doi.org/10.1071/MU01060

Pertierra, L. R., Segovia, N. I., Noll, D., Martinez, P. A., Pliscoff, P., Barbosa, A., Aragón, P., Raya Rey, A., Pistorius, P., Trathan, P., Polanowski, A., Bonadonna, F., Le Bohec, C., Bi, K. E., Wang-Claypool, C. Y., González-Acuña, D., Dantas, G. P. M., Bowie, R. C. K., Poulin, E., & Vianna, J. A. (2020). Cryptic speciation in gentoo penguins is driven by geographic isolation and regional marine conditions: Unforeseen vulnerabilities to global change. *Diversity and Distributions*, *26*, 958–975.

Power, D. M., & Ainley, D. G. (1986). Seabird geographic variation: Similarity among populations of Leach's storm-petrel. *The Auk*, *103*, 575–585. https://doi.org/10.1093/AUK/103.3.575

Robertson, C. J. R., & Nunn, G. B. (1998). Towards a new taxonomy for albatrosses. In Albatross Biology and Conservation, 13–19.

Taylor, R. S., Bolton, M., Beard, A., Birt, T., Deane-Coe, P., Raine, A. F., González-Solís, J., Lougheed, S. C., & Friesen, V. L. (2019). Cryptic species and independent origins of allochronic populations within a seabird species complex (*Hydrobates* spp.). *Molecular Phylogenetics and Evolution*, *139*, 106552. https://doi.org/10.1016/J.YMPEV.2019.106552

Vignal, A., Milan, D., SanCristobal, M., & Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, *34*, 1–31. https://doi.org/10.1186/1297-9686-34-3-275/METRICS

Yang, H., Tao, Y., Zheng, Z., Li, C., Sweetingham, M. W., & Howieson, J. G. (2012). Application of next-generation sequencing for rapid marker development in molecular plant breeding: a case study on anthracnose disease resistance in *Lupinus angustifolius L*. *BMC Genomics*, *13*, 1. https://doi.org/10.1186/1471-2164-13-318

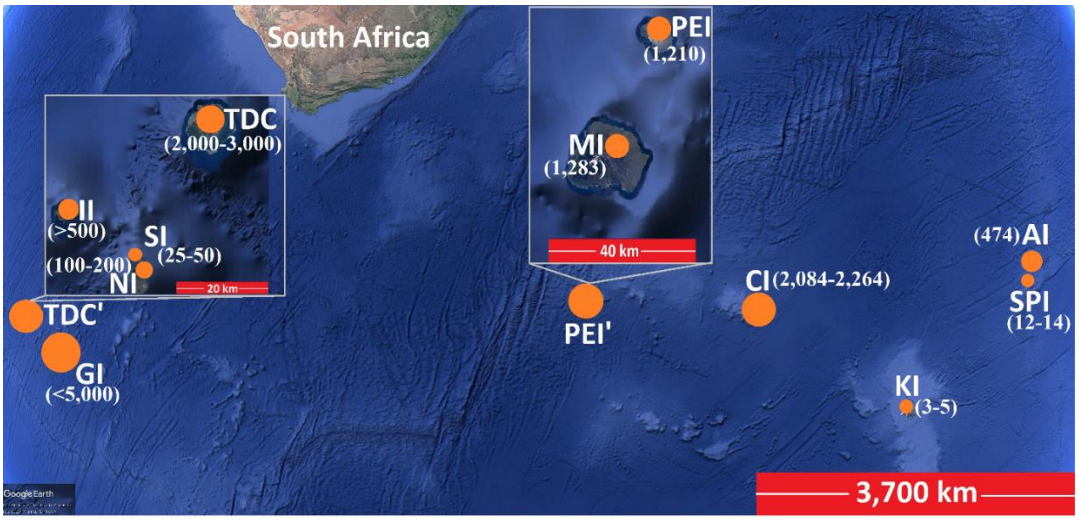

Figure 1.1: Breeding sites of sooty albatross *(P. fusca)*. Inset on left shows four breeding islands within the Tristan archipelago (TDC') and inset on right shows two islands in the Prince Edward Island group (PEI'). Tristan da Cunha (TDC), Nightingale (NI), Stoltenhoff (SI), Inaccessible (II), Gough (GI), Marion (MI), Prince Edward (PEI), Crozet (CI), Kerguelen (KI), Amsterdam (AI), and St. Paul (SPI). Numbers within brackets represent approximate number of breeding pairs on each island. Map data: Google Earth, Maxar Technologies, National Centre for Space Studies, Scripps Institution of Oceanography, National Oceanic and Atmospheric Administration, U.S. Navy, National Geospatial-Intelligence Agency, General Bathymetric Chart of the Oceans, Lamont-Doherty Earth Observatory Columbia University and National Science Foundation.

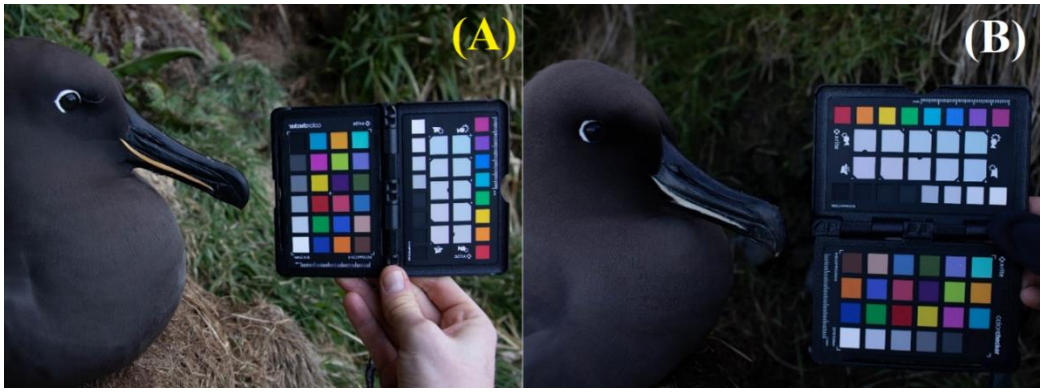

Figure 1.2: Sooty albatross with X-Rite colour-checker passport 2 on (A) Atlantic Ocean bird from Gough Island © Danielle Keys and (B) Indian Ocean bird from Marion Island ©Christopher Jones.

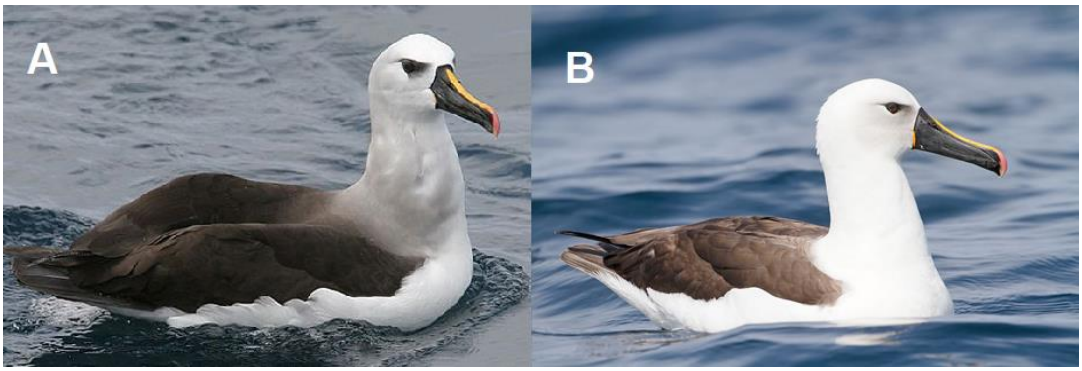

Figure 1.3: Photo of adult (A) Atlantic (*T. chlororhynchos*) and (B) Indian (*T. carteri*) yellow-nosed albatrosses. © J. J. Harrison and www.worldlifeexpectancy.com.

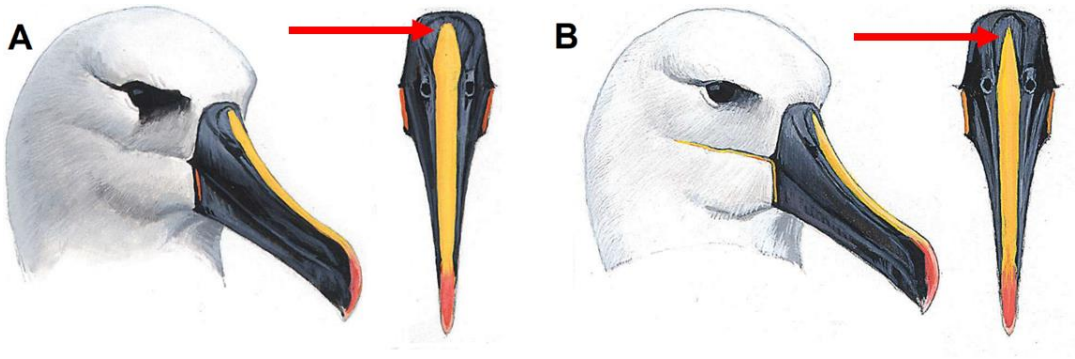

Figure 1.4: Comparison of shape of culminicorn (labelled with red arrow) and eye patch in adult (A) Atlantic (*T. chlororhynchos*) and (B) Indian (*T. carteri*) yellow-nosed albatross (Flood, 2015; used with permission from author).

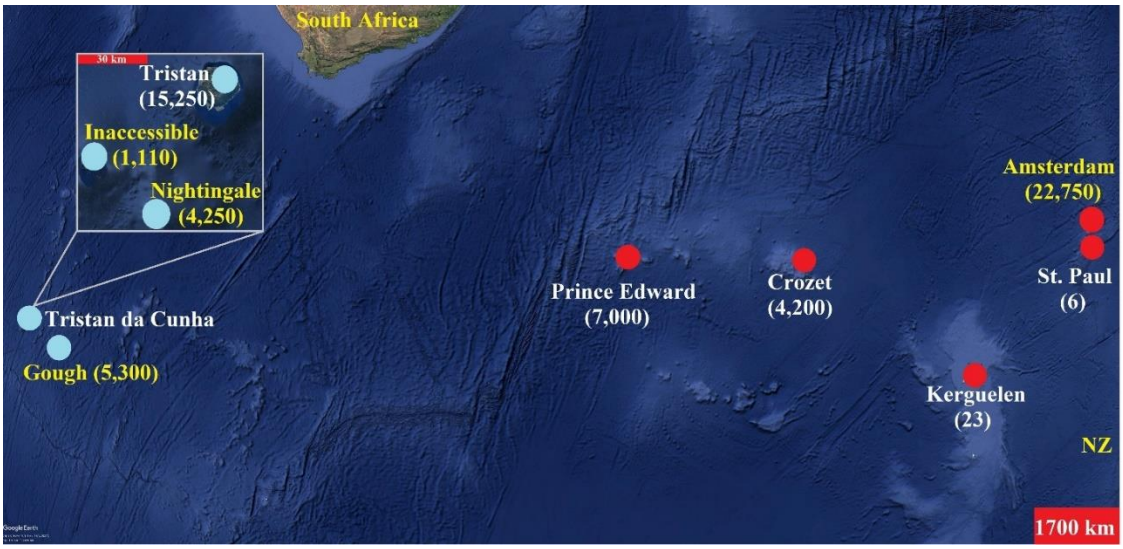

Figure 1.5: Breeding sites of Atlantic (*T. chlororhynchos*, blue) and Indian (*T. carteri*, red) yellow-nosed albatrosses. Inset shows three breeding islands within the Tristan archipelago (TDC'). Tristan da Cunha (TDC), Nightingale (NI), Inaccessible (II), Gough (GI), Crozet (CI), Kerguelen (KI), Amsterdam (AI), St. Paul (SPI) and, Prince Edward (PEI). Numbers within brackets represent approximate number of breeding pairs on each island. Map data: Google Earth, Maxar Technologies, National Centre for Space Studies, Scripps Institution of Oceanography, National Oceanic and Atmospheric Administration, U.S. Navy, National Geospatial-Intelligence Agency, General Bathymetric Chart of the Oceans, Lamont-Doherty Earth Observatory Columbia University and National Science Foundation.

Chapter 2: Another cryptic albatross. The case of the sooty albatross *Phoebetria fusca* species complex

Dilini K. Abeyrama[1], Michelle M. Risi[2,3], Maelle Connan[4], Peter G. Ryan[3], Christopher W. Jones[2,3], Trevor Glass[5], Alexis M. Osborne[2,3], Antje Steinfurth[2], Andy Schofield[2], Pierre Pistorius[4], Dannielle Keys[4], Yves Cherel[6], Theresa M. Burg[1]

1 University of Lethbridge, Lethbridge, Alberta, Canada

2 Royal Society for the Protection of Birds, The Lodge, Sandy, Bedfordshire SG19 2DL, United Kingdom

3 FitzPatrick Institute of African Ornithology, University of Cape Town, Rondebosch 7701, South Africa

4 Department of Zoology, Institute for Coastal and Marine Research, Marine Apex Predator Research Unit, Nelson Mandela University, Port Elizabeth, South Africa.

5 Tristan Conservation Department, Edinburgh of the Seven Seas, Tristan da Cunha, TDCU 1ZZ, South Atlantic

6 Centre d'Etudes Biologiques de Chizé, UPR 1934 du Centre National de la Recherche Scientifique, BP 14, 79360 Villiers-en-Bois, France

*This chapter has been prepared as a manuscript for submission and author contributions are listed under section 2.8. Michelle M. Risi and I are co-first authors of the manuscript. I collected and analysed all the molecular data. Michelle M. Risi and Maelle Connan collected and analysed isotope and sulcus colour data. The manuscript has been reviewed thoroughly by coauthors.

2.1 Abstract

Regardless of being remote and featuring harsh environmental conditions, the Southern Ocean is home to a large variety of plants and animals, including many endemic species. High levels of endemism are due in part to non-physical barriers limiting gene flow among oceanic populations. The sooty albatross, *Phoebetria fusca*, is an endangered seabird species breeding on a few islands in the Atlantic and Indian Oceans. There is some morphological and isotopic variation among individuals of the two oceanic populations. To compare the population genetics of sooty albatross breeding within Atlantic and Indian Ocean basins, we genotyped 44 birds with mitochondrial control region and 77 birds with 11 microsatellite markers. Birds were sampled from Tristan da Cunha, Gough, Marion and Amsterdam Islands. We analysed the richness of yellow colour in the sulcus strip and carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) stable isotope compositions in body feathers of breeding adults from Gough and Marion Islands. Both mitochondrial and microsatellite markers identified two genetic clusters separating sooty albatross breeding in the Atlantic and Indian Ocean basins, though differences between the Tristan da Cunha population and the Indian populations were not significant with microsatellite data. The blue-yellow hue of sooty albatrosses sampled on Gough Island was significantly higher than individuals sampled on Marion Island. Feathers from Marion sooty albatrosses exhibited significantly higher $\delta^{13}C$ values than feathers collected from Gough birds. Sooty albatrosses breeding on islands in two ocean basins are different from each other in their genetics, morphology and feather chemical composition. The taxonomy of sooty albatross needs to be reassessed, and new management and conservation plans must be implemented to prevent the loss of distinct populations of the species.

Keywords: molecular markers, feather stable isotopes, genetic structure, sulcus colour

2.2 Introduction

Although it is very remote and hostile, the Southern Ocean is home to many flora and fauna. Undersea habitats and most of the islands in the Southern Ocean are not easily accessible and are therefore poorly studied. As a result, the biodiversity of the Southern Ocean is highly underestimated (Brooks & Christian, 2023). The region's unique environmental conditions and geographical features have resulted in high endemism levels within the Southern Ocean (Griffiths, 2010). Some species, or different populations of species, are endemic not only to the Southern Ocean but to specific ocean basins or islands (Harrison et al., 2013; Jackson et al., 2014; Abeyrama et al., 2021).

Most of the marine top predators in this region face a limited number of physical barriers and gene flow between different breeding populations of these groups is restricted mostly by non-physical barriers (Friesen et al., 2007; Jackson et al., 2014; Munro & Burg, 2017). In seabirds, non-physical barriers such as at-sea distribution, natal philopatry and breeding phenology play a more prominent role than physical barriers (Friesen et al., 2007; Munro & Burg, 2017). Despite being good flyers or swimmers with the potential to overcome most physical barriers, seabirds are a highly diversified group. A review by Friesen et al. (2007) revealed that both parapatric and sympatric speciation are common in seabirds. Physical isolation, philopatry, non-breeding distribution, narrow foraging range during the breeding season and breeding phenology are all known to restrict the gene flow in seabirds, including albatrosses (Abbott & Double, 2003a; Friesen et al., 2007; Rains et al., 2011; Coulson, 2016; Munro & Burg, 2017; Danckwerts et al., 2021).

The sooty albatross (*Phoebetria fusca*) is an endangered Procellariiform breeding on 11 islands in the Atlantic and Indian Oceans though a few of these islands have <50

breeding pairs (Figure 2.1; ACAP, 2010; BirdLife International, 2023). The populations of both ocean basins are ecologically separated; birds mostly segregate in their preferred foraging areas during breeding and moulting (Delord et al., 2013, Schoombie et al., 2022). The global population size of sooty albatross is relatively small (<14,400 breeding pairs; Figure 2.1; BirdLife International, 2023), and birds breeding on most islands face several threats. The populations declined by ~60% over the last three generations (90 years: BirdLife International, 2023). The most significant threat sooty albatross face is being caught as bycatch in longline fisheries (Schoombie et al., 2022). The south Atlantic sooty albatross population (~62% of the global population; ACAP, 2010) is also greatly affected by climate change (Weimerskirch, 2007; Schoombie et al., 2017). Ocean warming has resulted in longer foraging trips to reach feeding grounds. Atlantic individuals now forage farther from the Exclusive Economic Zone compared to individuals breeding in the south Indian Ocean (Weimerskirch, 2007; Schoombie et al., 2017). Moreover, Gough and Tristan da Cunha Islands, which together contain ~91% of breeding pairs in the south Atlantic Ocean, currently both have invasive mammals which prey on sooty albatross chicks (Cuthbert et al., 2013). Sooty albatrosses breeding in the south Indian Ocean are also threatened by fisheries, invasive mice, and diseases like avian cholera (Dilley et al., 2016; Jaeger et al., 2018; Schoombie et al., 2022).

Both at-sea and on-land threats continue to affect this endangered species resulting in declines for some populations although the populations on a few islands are stable. Therefore, effective conservation and management plans must be implemented. Effective planning requires understanding the connectivity of populations. Since Atlantic and Indian populations of sooty albatross segregate in their preferred foraging areas during breeding (Schoombie et al., 2017) and moulting (Cherel et al., 2013), our

study used mitochondrial and nuclear markers to determine (i) if sooty albatross populations breeding in the south Atlantic and south Indian Ocean basins are genetically different and (ii) if there is any population structuring within each ocean basin. In addition, sulcus colour (fleshy skin on either side of the lower mandible) and feather stable isotopes were used to determine whether differences could be observed between Gough (Atlantic Ocean) and Marion (west Indian Ocean) populations. Altogether, this information will determine whether the individuals breeding in the two ocean basins are distinct species. If that is the case, a reassessment of the taxonomy and conservation status of the (potentially two different) species must be done.

2.3 Methods

2.3.1 Study sites

The Tristan da Cunha archipelago (37° 06′ S, 12° 17′ W) and Gough Island (40° 20′ S, 9° 50′ W, 380 km south-southeast of Tristan) are located in the middle of the south Atlantic Ocean approximately 2,800 km west of South Africa and 3,500 km east of South America, north of Antarctic polar front (APF) and sub-Antarctic front (SAF). The two are separated by the subtropical front (STF) which is north of Gough. Marion Island (46° 50′ S, 37° 50′ E) forms part of the Prince Edward Islands group, a South African territory situated 1,750 km southeast of South Africa. Like Gough Island, Marion Island is also located south of STF and north of APF and SAF. Amsterdam Island (37° 50′ S, 77° 33′ E) is a part of the French southern and Antarctic lands in the southern Indian Ocean and situated in the central southern Indian Ocean, roughly equidistant to Africa, Australia, and Antarctica. Like Tristan da Cunha, Amsterdam Island is located north of APF, SAF and STF (Figure 2.1). These four sites represent ~70% of the global sooty albatross population (BirdLife International, 2023).

2.3.2 Sampling

Blood samples were obtained from adult sooty albatrosses breeding at Tristan da Cunha (n=8), Gough (n=30), Marion (n=30) and Amsterdam Islands (n=9). Blood samples (<0.5 ml) were collected from the tarsus vein of adult birds using 25-gauge hypodermic needles and stored in 70% ethanol. Samples were stored at -20°C as soon as possible after collection.

In addition, six body feathers were collected from each individual sampled on Gough and Marion Islands and stored frozen in sealed plastic bags. Photographs of breeding adults at Gough (n=40) and Marion (n=40) Islands were taken, clearly showing the bill and sulcus alongside an X-Rite colour-checker passport 2 (X-Rite, Incorporated, 2022; Figure 2.2). Gough birds were photographed with a Canon 800D in shutter priority mode, and Marion birds were photographed using the same procedure and a Canon 750D. An attempt was made to keep the lighting on the bird and colour chart uniform, and to have both be equidistant from the lens. All images were captured in a RAW file format. Birds were photographed in late incubation/brood guard phase and not all birds photographed were sampled for blood or feathers.

2.3.3 Molecular analyses

2.3.3.1 DNA extraction

A volume of ~5 µl of blood was used for each DNA extraction. DNA was extracted using either modified chelex (Walsh et al., 1991; Burg & Croxall, 2001) or salt-out (Aljanabi & Martinez, 1997) extraction protocols and was stored at -20°C.

2.3.3.2 Microsatellite genotyping

A total of 77 samples was genotyped at 11 microsatellite loci (Tc65, Dc26, 11H7, 11H1, 12C8, 12E1, Cd11, De35, De11, Dc27 and Parm03). These loci were originally

isolated for other Procellariiformes including albatrosses (Burg, 1999; Burg & Croxall, 2001; Abbott & Double, 2003b; Dubois et al., 2005; Bried et al., 2007; Brown & Jordan, 2009; Sun et al., 2009) and were selected after checking their variability in sooty albatrosses. DNA was amplified in 10 µl reactions containing FroggaBio buffer, 0.25 mM dNTP, 0.5 µM forward primer, 1 µM reverse primer, 0.05 µM fluorescent labelled M13 primer and 0.5 units of FroggaBio taq polymerase. PCR consisted of one cycle of 120 s at 94°C, seven cycles of 60 s at 94°C, 45 s at T_1, 30 s at 72°C, 25 cycles of 30 s at 89°C, 30 s at T_2, 45 s at 72°C and one final cycle of 300 s at 72°C and 20 s at 4°C. For four loci (Dc26, Dc27, 11H1 and Cd11), T_1 and T_2 temperatures were 45°C and 48°C, 50°C and 52°C for another five loci (Tc65, De35, 11H7, 12C8 and 12E1), and 52°C and 54°C for other two loci (De11 and Parm03). PCR with the two higher annealing temperatures had 28 cycles at step three instead of 25 cycles. Denatured PCR products were run on a 6% polyacrylamide gel in a LI-COR 4300 DNA Analyzer. Three positive and one negative control were included in each load, and a calibration gel was done to ensure consistent scoring. Alleles were scored manually, with a second person checking them for consistent and accurate scoring. Genepop v. 4.7.5 was used to check for deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) (Raymond & Rousset, 1995; Rousset, 2008). The false discovery rate (0.05) in P values for HWE and LD were corrected using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). Microsatellite data were analysed using STRUCTURE v. 2.3.4. Analyses were done with 100,000 burnin followed by 200,000 runs incorporating the admixture model with locpriors and correlated allele frequencies to identify the optimal number of genetic clusters (K) (Pritchard et al., 2000). Initially, ten replicates were performed for each K (K=1-5) value. Average likelihoods for each K value were calculated using

Structure Harvester v. 0.6.94 (Earl & vonHoldt, 2012). The optimal number of genetic clusters within the data was selected using ΔK and Evanno method. Genetic variation within and among the two ocean basins were calculated using F_{ST} statistics. The F_{ST} values were obtained by performing an analysis of molecular variance (AMOVA, Excoffier et al., 1992) in Arlequin v. 3.5.2.2 (Excoffier & Lischer, 2010).

2.3.3.3 Mitochondrial DNA analysis

A 365 bp fragment of the mitochondrial control region (mtCR) was amplified using the primers H505 (5′-GAAAGAATGGTCCTGAAGC-3′, Burg & Croxall, 2001) and SpecF1 (5′-CAGCTTATGTATAAATGCG-3′, Abbott & Double, 2003a). DNA amplification was done in 25 µl reactions with GoTaq® Flexi buffer, 1.5 mM $MgCl_2$, 0.1 mM dNTP, 0.4 µM forward and reverse primers and 0.1 unit of GoTaq® Flexi DNA polymerase. Amplification consisted of one cycle of 120 s at 94°C, 45 s at 54°C, 60 s at 72°C, 37 cycles of 30 s at 94°C, 45 s at 54°C, 60 s at 72°C and one final cycle of 300 s at 72°C. The quality of PCR products was assessed using a 0.8% agarose gel stained with ethidium bromide. The PCR products were sent to Génome Québec, Canada, for sequencing. DNA sequences were edited manually and aligned in MEGA v. 5.2 (Tamura et al., 2011). A haplotype network was constructed using TCS v.1.21 (Clement et al., 2000) and edited using tcsBU (TCS Beautifier; dos Santos et al., 2016). Diversity parameters such as nucleotide diversity and haplotype diversity were calculated in DnaSP v. 6.12.03 (Rozas et al., 2017).

2.3.4 Sulcus colour analysis

The micaToolbox plugin (Troscianko & Stevens, 2015) was used in ImageJ v. 1.51 to process and extract colour values for all photos (Schneider et al., 2012). Colour values were then normalised and linearised by converting each raw photo into a multispectral image using the 36% reflectance patch of a colour chart as a reference reflectance

value (Troscianko, 2022). The colour chart was used to develop a cone-catch model based on human vision in the 400–700 nm spectral range under D65 illuminant. Successful linearisation was confirmed by model R^2 values >0.98 for each photoreceptor in the cone-catch model (Troscianko & Stevens, 2015). The cone-catch model was used to convert each multispectral image into XYZ colour space. A region of interest (ROI) was created on each image by selecting the area of the sulcus stripe, and a mean value of X (LW), Y (MW) and Z (SW) were calculated for all the pixels inside the ROI using the Measure ROIs tool. Then the richness of yellow in the sulcus stripe was measured by calculating the blue-yellow hue using the equation: yellow = (LW+MW)/2, blue-yellow = (yellow-SW)/(yellow+SW).

2.3.5 Stable isotope analysis

Carbon and nitrogen stable isotope composition were determined using four body feathers per bird from 30 breeding adults from Gough and Marion Islands. These same birds were included in the genetics analyses. Feathers were washed individually in chloroform:methanol (2:1), and rinsed in successive baths of methanol and distilled water. Feathers were then dried at 50°C before being finely cut and homogenised. The relative isotopic abundance of carbon ($^{13}C/^{12}C$; $\delta^{13}C$) and nitrogen ($^{15}N/^{14}N$; $\delta^{15}N$) were determined on Thermo Delta V Plus mass spectrometer interfaced with a Conflo IV device to a Thermo Flash EA 2000 elemental analyzer (Stable Light Isotope Unit, University of Cape Town, South Africa). Results were presented relative to Vienna PeeDee Belemnite and atmospheric N_2 for $\delta^{13}C$ and $\delta^{15}N$, respectively. Replicates of laboratory standards (Merck Gel, Sucrose, Valine, Chocolate) indicated measurement errors < 0.22‰ and < 0.16‰ for carbon and nitrogen stable isotope measurements, respectively.

The effect of island on $\delta^{13}C$ and $\delta^{15}N$ composition was tested using a generalized linear mixed model using the package lme4 (Bates et al., 2015). Standard (SEA) and Bayesian standard ellipses areas (SEA_B) were then estimated for each island and their overlaps using the Stable Isotope Bayesian Ellipses (*SIBER*) package in R (Jackson et al., 2011). Stable isotope data failed the multivariate normality assumption (Mardia test $P < 0.05$; package *MVN*; Korkmaz et al. 2014), but Syväranta et al. (2013) determined that the reliability of SEA increased for $n \geq 30$ in such case. Finally, to ease comparisons with published values (Cherel et al., 2013), feathers were attributed to three water masses according to their carbon stable isotope composition: 3 $\delta^{13}C$ values < -21.2‰, between -21.2 and -18.3‰, and > -18.3‰ corresponded to Antarctic waters, sub-Antarctic waters, and sub-tropical waters, respectively (Jaeger et al., 2010).

2.4 Results

2.4.1 Nuclear DNA analysis

The number of alleles ranged from two to 10 per locus at the 11 polymorphic microsatellite loci. Five loci deviated significantly from HWE for some populations, but the pattern was inconsistent. Following Benjamini-Hochberg correction, 12C8, 12E1, and Dc26 deviated significantly from HWE at Marion Island and 11H1 and De11 at Gough Island. Following Benjamini-Hochberg correction, none of the loci were in linkage disequilibrium.

STRUCTURE clustering analysis with all 11 microsatellite loci resulted in K=2 (mean LnP(K) = -1305 and $\Delta K = 1.188$), and the two genetic clusters corresponded to the two ocean basins: Atlantic and Indian (Figure 2.3). All sooty albatrosses sampled in the Atlantic Ocean showed admix of the two ancestries. For the Indian Ocean, 29 of

the 30 sooty albatrosses sampled on Marion Island and all Amsterdam Island birds possessed 100% and >95% of the Indian ancestry, respectively (Figure 2.3). According to the hierarchical population structure analysis in AMOVA, Atlantic and Indian Ocean populations of sooty albatrosses are genetically different from each other (F_{ST} 0.0954 and $P < 0.001$). All population pairs of sooty albatrosses except Tristan da Cunha and the two Indian Oceanic populations were also significantly different from each other (Table 2.1).

2.4.2 mtDNA

Within the 365 bp mtCR sequences of 44 sooty albatrosses, 39 haplotypes were identified (Figure 2.4). A total of 43 variable sites were found, 34 in the Atlantic and 22 in the Indian Ocean populations. The haplotype network showed two distinct groups, mostly corresponding to two ocean basins; the Atlantic and Indian Oceans, separated by a minimum of 11 mutations. While no haplotypes were shared between the two ocean basins or among populations within the Atlantic Ocean, one haplotype was shared between the two populations in the Indian Ocean. Two sooty albatrosses from Tristan da Cunha grouped with the Indian Ocean cluster. The two individuals from Tristan da Cunha that placed within the Indian Ocean group of haplotypes had more than 70% Indian Ocean ancestry in the structure analysis using microsatellite markers and no missing data. The one bird from Marion Island, which had more than 60% Atlantic ancestry for the microsatellite data, possessed an Indian Ocean-specific mtCR haplotype, which was shared with a bird from Amsterdam Island. Nucleotide diversity was highest for the Tristan da Cunha population (0.0307) and lowest for Amsterdam Island (0.0095) (Table 2.2). Overall, nucleotide diversity was lower in birds from the Indian Ocean basin (0.0095-0.0150) than in the Atlantic Ocean

(0.0174-0.0307). Haplotype diversity for all four populations were comparable (0.972-1.000).

2.4.3 Sulcus colour

The blue-yellow hue of sooty albatrosses sampled on Gough Island (mean = 0.471±0.002) was significantly higher than individuals sampled on Marion Island (mean = 0.233±0.004; $t(66) = -19.01$, $P < 0.001$), with no overlap in their ranges (Figure 2.5). Despite slightly more variation within the Marion individuals, results indicate a consistent pattern of richer yellow within the sulcus stripes of Gough Island birds (Figure 2.5).

2.4.4 Stable isotope analysis

Carbon and nitrogen stable isotope data can be used as an indicator of the food consumed by the birds. Carbon stable isotope values ranged from -25.0‰ to -16.5 ‰ and from -20.3‰ to -16.1 ‰ for Gough and Marion Island sooty albatrosses respectively (Figure 2.6). Nitrogen stable isotope values ranged from 9.8‰ to 15.5 ‰ and from 7.6‰ to 16.2‰ for Gough and Marion Island birds respectively (Figure 2.6). Overall, feathers from Marion Island sooty albatrosses exhibited significantly higher $\delta^{13}C$ (GLMM: df = 58, t-value = 7.487, $P < 0.001$) and $\delta^{15}N$ (GLMM: df = 58, t-value = 5.080, $P < 0.001$) values than feathers collected from Gough Island birds. SEAs were of similar size for both islands, 2.16‰ versus 2.33‰ for Gough and Marion Islands respectively (Figure S2.1) and the isotopic overlap between the two ellipses reached 0.74‰. $\delta^{13}C$ indicated that the bulk (61%) of feathers collected from Gough Island birds had grown in sub-Antarctic waters, while almost 80% of feathers collected from Marion Island birds had grown in sub-tropical waters. Three feathers from three Gough Island birds (one each) had grown in Antarctic waters (2.5% of

feathers), and no feathers collected from Marion Island birds had grown in Antarctic waters.

2.5 Discussion

Using a combination of genetic, morphological and isotope data, we found population structuring in sooty albatrosses. All three datasets confirmed populations of sooty albatrosses in the Atlantic and Indian Oceans are significantly different from each other. Interestingly, when analysed with nuclear DNA, the sooty albatrosses from Tristan da Cunha seem to be genetically closer to Indian Ocean birds than to birds from Gough Island which is only ~415 km away, but when analysed with mtDNA, most Tristan da Cunha birds grouped with birds from Gough Island.

2.5.1 Divergence of Atlantic and Indian Ocean populations

As all datasets identify significant differences between Atlantic and Indian Ocean populations of sooty albatrosses, the two ocean basin populations have experienced a more prolonged split. The higher nucleotide and haplotype diversities in both Atlantic Ocean populations than the two Indian Ocean populations indicates a possible genetic drift from Atlantic Ocean to Indian Ocean. The sooty albatross population on Tristan da Cunha Island was not significantly different from either of the Indian Ocean populations based on microsatellite data, and two Tristan birds also contained mtDNA from the Indian Ocean populations. Therefore, sooty albatrosses from Tristan da Cunha Island are genetically closer to Indian Ocean populations than to the other Atlantic Ocean population although they are expressing Atlantic Ocean morphology. Based on this, it is possible that the Indian Ocean population was founded by sooty albatrosses from the Tristan da Cunha Island. But the fact that sooty albatrosses from two Indian Ocean populations are significantly genetically different from each other although they are not from the Tristan birds and significant morphological differences

between the two oceanic populations also suggests the possibility of a more recent secondary contact of the Tristan da Cunha population with Indian Ocean populations despite its large breeding population size.

The higher amounts of yellow in the sulcus stripes of adult sooty albatrosses breeding on Gough Island than on Marion Island are seemingly indicative of a wider trend. Time and resources did not allow for colour analysis of images on other islands. Photographs captured opportunistically at other breeding sites in the Tristan da Cunha archipelago suggest higher yellow trend holds for breeding sites in the Atlantic Ocean. Similarly, a bird opportunistically photographed on Prince Edward Island also had paler sulcus stripes like Marion Island birds, confirming consistent paler sulcus colour across the Indian Ocean.

In terms of isotope data, the pattern observed with our samples is same as the one seen in Cherel et al. (2013) where Gough Island birds moulted in more southern waters than Marion Island birds. Global location sensing (GLS or geolocator) data of Marion Island birds in non-breeding and moulting zones found that most of birds travel north in the Indian Ocean and moult (Schoombie et al., 2022). GLS data for Gough Island birds show that they mostly travel between 40° and 55°S within the central Atlantic Ocean (ACAP, 2010): i.e., both populations are mostly separated during moulting. According to Delord et al. (2013) sooty albatrosses from Crozet Island foraged in the Indian Ocean (mostly sub-tropical waters) during non-breeding season while Amsterdam birds foraged mostly east of Amsterdam in Australian waters. This differential behaviour could keep the populations separated and continue driving a separation.

2.5.2 Genetic barriers to sooty albatrosses

Except for the geographic distance, there is no other distinct physical barrier for gene flow between the sooty albatrosses breeding in the Atlantic and Indian Oceans. These birds can fly long distances; therefore, it is very unlikely that physical distance is acting as a barrier to gene flow between the two ocean basins. Other than physical barriers, non-physical barriers such as natal philopatry, breeding phenology and foraging distribution, which are more common genetic barriers for other seabird species, might be limiting gene flow between the two groups of sooty albatrosses (Steeves et al., 2005; Calderón et al., 2014; Taylor et al., 2018). It is also possible a genetic drift followed by reproductive isolation due to non-physical barriers results in the contemporary genetic structure observed in sooty albatrosses. A population genetic study on *Pachyptila* observed genetic structure representing founder effect (Masello et al., 2022).

Breeding phenology or temporal differences in breeding acts as a barrier in some seabird species such as band-rumped storm petrels, *Oceanodroma castro* (Taylor et al., 2019). Sooty albatrosses on Gough and Tristan da Cunha Islands start laying eggs mid-September although Tristan birds arrive the site about two weeks earlier than Gough Island birds (Elliott, 1957; Swales, 1965). On Marion Island, they start laying eggs early-October (Berruti, 1979). Although there are differences in breeding time among different islands for sooty albatrosses, it is not a large enough temporal difference to maintain the observed genetic divergence in sooty albatrosses as variation in breeding phenology within populations is much greater than the average differences between islands.

Seabirds as a group show high levels of natal philopatry, and many return close to their natal site to breed limiting the gene flow among populations (Harris, 1973;

Prince et al., 1994; Weimerskirch et al., 1997; Friesen et al., 2006; Coulson, 2016; Herman et al., 2022). Similarly, sooty albatross is also a species that shows high site fidelity. On Marion Island, 88% of breeders returned to the same colony to breed in following years (Berruti, 1979). Although there are no published data on natal philopatry of sooty albatrosses from all breeding populations, it is possible they show a similar trend to Marion Island birds. Therefore, natal philopatry is a potential genetic barrier for sooty albatrosses.

Furthermore, according to the BirdLife seabird tracking database, the foraging ranges of sooty albatrosses from the Atlantic and Indian Oceans are highly distinct with minimal overlap during both breeding and non-breeding seasons. In black-browed albatross (*Thalassarche melanophris*), two genetic groups were found in the southwest Atlantic, each with distinct at-sea foraging grounds (Burg & Croxall, 2001). Moreover, significant differences in isotopic composition in feathers of two oceanic groups also confirm differences in sooty albatross at-sea distributions during both breeding and non-breeding seasons (Delord et al., 2013; Schoombie et al., 2017; Carpenter-Kling et al., 2020; Schoombie et al., 2022). Similarly, differences in sulcus colour can be a result of differential food resources that birds ingest within their respective foraging grounds. Also, it is possible that the foraging range of non-breeders from Tristan da Cunha overlaps with foraging ranges of Indian Ocean sooty albatrosses to some extent allowing secondary contact so making them genetically closer. Unfortunately, there are no tracking or stable isotope data for non-breeding sooty albatrosses from Tristan da Cunha to confirm their at-sea distribution. It is possible that at-sea distribution acts as a major genetic barrier and drives the morphological and isotopic variation between the two oceanic groups of sooty albatrosses.

2.5.3 Taxonomic implications

There is no one specific definition for species. But in the biological species concept, two groups which are reproductively isolated and do not interbreed are known as two species (Mayr 1996). In the morphological species concept, groups that are morphologically different are identified as separate species while phylogenetic species concept identifies two groups that have descended from the same ancestor but show different derived characters as two species (Wheeler & Platnick, 2000). As the sooty albatrosses breeding in Atlantic and Indian Ocean basins show differences in their genetics, morphology and biochemical composition, they fit the criteria of biological, morphological and phylogenetic species concepts. Moreover, when populations of the same species are geographically isolated and show diverging characters, they can be identified as subspecies of the same species (Mayr, 1942, 1982b; Mallet, 2011). Therefore, the taxonomy of sooty albatross needs to be reassessed and need to be identified as Atlantic and Indian Ocean subspecies of the sooty albatross species.

2.6 Conclusions

In sooty albatrosses, genetic drift, natal philopatry and different at-sea distributions seem to be driving the population divergence. Differential at-sea distribution results in different nutritional intakes between the two oceanic populations, thereby making their stable isotope compositions significantly different between Gough and Marion Islands. The sooty albatross breeding in Atlantic and Indian Ocean basins are genetically, morphologically and biochemically different, and therefore, a reassessment of their taxonomy is required. The sooty albatross is already considered endangered by IUCN. Since Atlantic and Indian Ocean populations are two distinct groups, their geographical range is halved, and new conservation and management plans are required.

2.7 Acknowledgements

Logistical support was provided by the South African Department of Environment, Forestry and Fisheries, the FitzPatrick Institute of African Ornithology at the University of Cape Town and the Royal Society for the Protection of Birds. The Tristan da Cunha Administrator, Island Council and Conservation Department provided permission to work on Tristan da Cunha and Gough Islands. The South African Department of Environment, Forestry and Fisheries provided permission to work on Marion Island. Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship, Alberta Society of Professional Biologists and the Seabird Group provided funding. Thanks to fieldworkers on Marion (Tegan Carpenter-Kling) and Amsterdam who collected blood and feather samples. Thanks to Claire Spottiswoode for advice on photographing for colour analysis, Jolyon Troscianko for the guidance in analysing the sulcus photos, Julie Luyt and Patricia Groenewald for conducting the stable isotope analyses at the University of Cape Town. Special thanks to members of Burg lab, University of Lethbridge, for their enormous support.

2.8 Author contributions

DKA, MMR, MC, DK, PGR and TMB designed the study and MMR, CJ, AO, TG, DK, ASc and ASt performed and facilitated the fieldwork. DKA and MC completed the lab work. DKA, MMR and MC led the analyses. DKA, TMB, MMR and MC wrote the manuscript with input from the others.

2.9 Funding

Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship, Alberta Society of Professional Biologists, and the Seabird Group.

2.10 References

Abbott, C. L., & Double, M. C. (2003a). Phylogeography of shy and white-capped albatrosses inferred from mitochondrial DNA sequences: Implications for population history and taxonomy. *Molecular Ecology*, *12*(10), 2747–2758. https://doi.org/10.1046/j.1365-294X.2003.01944.x

Abbott, C. L., & Double, M. C. (2003b). Genetic structure, conservation genetics and evidence of speciation by range expansion in shy and white-capped albatrosses. *Molecular Ecology*, *12*(11), 2953–2962. https://doi.org/10.1046/j.1365-294X.2003.01980.x

Abeyrama, D. K., Dempsey, Z. W., Ryan, P. G., & Burg, T. M. (2021). Cryptic speciation and population differentiation in the yellow-nosed albatross species complex. *Conservation Genetics*, *22*, 757–766. https://doi.org/10.1007/s10592-021-01358-x

Al-Samarai, F. R., & Al-Kazaz, A. A. (2015). Molecular markers: an Introduction and applications. *European Journal of Molecular Biotechnology*, *9*, 118–130. https://doi.org/10.13187/ejmb.2015.9.118

Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, *25*(22), 4692–4693. https://doi.org/10.1093/NAR/25.22.4692

Agreement on the Conservation of Albatrosses and Petrels (2010). ACAP Species assessments: Sooty albatross *Phoebetria fusca*.

Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/J.2517-6161.1995.TB02031.X

Berruti, A. (1979). The breeding biologies of the sooty albatrosses *Phoebetria fusca* and *P. palpebrata*. *Emu*, *79*(4), 161–175.

BirdLife International (2023). Species factsheet: *Phoebetria fusca*.

Bried, J., Nicolaus, M., Jarne, P., Dubois, M. P., & Jouventin, P. (2007). Population biology of the wandering albatross (*Diomedea exulans*) in the Crozet and Kerguelen archipelagos, southern Indian Ocean, approached through genetic and demographic methods. *Journal of Zoology*, *272*(1), 20–29. https://doi.org/10.1111/j.1469-7998.2006.00242.x

Brooks, C., Christian, C. (2023). Antarctica and the Southern Ocean: Our Last Great Wilderness. In: Obaidullah, F. (eds) The Ocean and Us. Springer, Cham. https://doi.org/10.1007/978-3-031-10812-9_17

Brown, R. M., & Jordan, W. C. (2009). Characterization of polymorphic microsatellite loci from Round Island Petrels (*Pterodroma arminjoniana*) and their utility in other seabird species. *Journal of Ornithology*, *150*, 925–929.

Burg, T. M. (1999). Isolation and characterization of microsatellites in albatrosses. *Molecular Ecology*, *8*(2), 335–346. https://doi.org/10.1046/j.1365-294X.1999.00534.x

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.

Calderón, L., Quintana, F., Cabanne, G. S., Lougheed, S. C., & Tubaro, P. L. (2014). Phylogeography and genetic structure of two Patagonian shag species (Aves: Phalacrocoracidae). *Molecular Phylogenetics and Evolution*, *72*, 42–53. https://doi.org/10.1016/J.YMPEV.2013.12.011

Carpenter-Kling, T., Reisinger, R.R., Orgeret, F., Connan, M., Stevens, K.L., Ryan, P.G., Makhado, A. and Pistorius, P.A. (2020). Foraging in a dynamic environment: Response of four sympatric sub-Antarctic albatross species to interannual environmental variability. *Ecology and Evolution*, *10*(20), 11277–11295.

Cherel, Y., Jaeger, A., Alderman, R., Jaquemet, S., Richard, P., Wanless, R. M., Phillips, R. A., & Thompson, D. R. (2013). A comprehensive isotopic investigation of habitat preferences in non-breeding albatrosses from the Southern Ocean. *Ecography*, *36*(3), 277–286. https://doi.org/10.1111/j.1600-0587.2012.07466.x

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, *9*(10), 1657–1659. https://doi.org/10.1046/J.1365-294X.2000.01020.X

Coulson, J. C. (2016). A review of philopatry in seabirds and comparisons with other waterbird species. *Waterbirds*, *39*(3), 229–240. https://doi.org/10.1675/063.039.0302

Croxall, J. P., Butchart, S. H. M., Lascelles, B., Stattersfield, A. J., Sullivan, B., Symes, A., & Taylor, P. (2012). Seabird conservation status, threats and priority actions: a global assessment. *Bird Conservation International*, *22*(1), 1–34. https://doi.org/10.1017/S0959270912000020

Cuthbert, R. J., Louw, H., Parker, G., Rexer-Huber, K., & Visser, P. (2013). Observations of mice predation on dark-mantled sooty albatross and Atlantic yellow-nosed albatross chicks at Gough Island. *Antarctic Science*, *25*(6), 763-766.

Danckwerts, D. K., Humeau, L., Pinet, P., McQuaid, C. D., & Le Corre, M. (2021). Extreme philopatry and genetic diversification at unprecedented scales in a seabird. *Scientific Reports*, 11, 1–12. https://doi.org/10.1038/s41598-021-86406-9

Delord, K., Barbraud, C., Bost, C. A., Cherel, Y., Guinet, C., & Weimerskirch, H. (2013). Atlas of Top Predators from French Southern Territories in the Southern Indian Ocean. CEBC-CNRS. pp. 252.

Dilley, B. J., Schoombie, S., Schoombie, J., & Ryan, P. G. (2016). 'Scalping' of albatross fledglings by introduced mice spreads rapidly at Marion Island. *Antarctic Science*, *28*(2), 73–80. https://doi.org/10.1017/S0954102015000486

dos Santos Múrias, A., Cabezas, M. P., Tavares, A. I., Xavier, R., & Branco, M. (2016). tcsBU: a tool to extend TCS network layout and visualization. *Bioinformatics*, *32*(4), 627–628. https://doi.org/10.1093/BIOINFORMATICS/BTV636

Dubois, M. P., Jarne, P., & Jouventin, P. (2005). Ten polymorphic microsatellite markers in the wandering albatross *Diomedea exulans*. *Molecular Ecology Notes*, *5*(4), 905–907. https://doi.org/10.1111/j.1471-8286.2005.01108.x

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/S12686-011-9548-7

Elliott, H. F. I. (1957). The birds of Tristan da Cunha. *Ibis*, *99*, 546–585.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/J.1755-0998.2010.02847.X

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*(2), 479–491. https://doi.org/10.1093/GENETICS/131.2.479

Friesen, V. L., Burg, T. M., & McCoy, K. D. (2007). Mechanisms of population differentiation in seabirds: Invited review. *Molecular Ecology*, *16*(9), 1765–1785. https://doi.org/10.1111/j.1365-294X.2006.03197.x

Friesen, V. L., González, J. A., & Cruz-Delgado, F. (2006). Population genetic structure and conservation of the Galápagos petrel (*Pterodroma phaeopygia*). *Conservation Genetics 2006 7:1*, *7*(1), 105–115. https://doi.org/10.1007/S10592-005-8704-6

Griffiths, H. J. (2010). Antarctic marine biodiversity – What do we know about the distribution of life in the Southern Ocean? *PLoS One*, *5*(8), e11683. https://doi.org/10.1371/JOURNAL.PONE.0011683

Harris, M. P. (1973). The biology of the waved albatross *Diomedea irrorata* of Hood Island, Galapagos. *Ibis*, *115*(4), 483–510. https://doi.org/10.1111/J.1474-919X.1973.TB01988.X

Harrison, P., Sallaberry, M., Gaskin, C. P., Baird, K. A., Jaramillo, A., Metz, S. M., Pearman, M., Michael O'Keeffe, Dowdall, J., Enright, S., Fahy, K., Gilligan, J., & Lillie, G. (2013). A new storm-petrel species from Chile. *Auk*, *130*(1), 180–191. https://doi.org/10.1525/AUK.2012.12071

Jackson, A. L., Inger, R., Parnell, A. C., & Bearhop, S. (2011). Comparing isotopic niche widths among and within communities: SIBER–Stable Isotope Bayesian Ellipses in R. J*ournal of Animal Ecology*, *80*(3), 595–602.

Jackson, J. A., Steel, D. J., Beerli, P., Congdon, B. C., Olavarría, C., Leslie, M. S., Pomilla, C., Rosenbaum, H., & Scott, B. C. (2014). Global diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). *Proceedings of the*

Royal Society B: Biological Sciences, *281*(1786), 20133222. https://doi.org/10.1098/RSPB.2013.3222

Jaeger, A., Lecomte, V. J., Weimerskirch, H., Richard, P., & Cherel, Y. (2010). Seabird satellite tracking validates the use of latitudinal isoscapes to depict predators' foraging areas in the Southern Ocean. *Rapid Communications in Mass Spectrometry*, *24*(23), 3456–3460.

Jaeger, A., Lebarbenchon, C., Bourret, V., Bastien, M., Lagadec, E., Thiebot, J. B., Boulinier, T., Delord, K., Barbraud, C., Marteau, C., Dellagi, K., Tortosa, P., & Weimerskirch, H. (2018). Avian cholera outbreaks threaten seabird species on Amsterdam Island. *PLoS One*, *13*(5), e0197291. https://doi.org/10.1371/JOURNAL.PONE.0197291

Karl, S. A., Toonen, R. J., Grant, W. S., & Bowen, B. W. (2012). Common misconceptions in molecular ecology: Echoes of the modern synthesis. *Molecular Ecology*, *21*, 4171–4189. https://doi.org/10.1111/j.1365-294X.2012.05576.x

Korkmaz, S., Göksülük, D., & Zararsiz, G. Ö. K. M. E. N. (2014). MVN: An R package for assessing multivariate normality. *The R Journal*, 6(2), 151–162.

Mallet, J. (2001). Subspecies, semispecies, superspecies. In: Levin S ed. Encyclopedia of biodiversity, Vol. 5. London: Academic Press, 523–526.

Masello, J. F., Ryan, P. G., Shepherd, L. D., Quillfeldt, P., Cherel, Y., Tennyson, A. J., Alderman, R., Calderón, L., Cole, T .L., Cuthbert, R. J. & Dilley, B. J., 2022. Independent evolution of intermediate bill widths in a seabird clade. *Molecular Genetics and Genomics*, 1–16.

Mayr, E. (1996). What is a species, and what is not? *Philosophy of Science*, *63*(2), 262–277.

Mayr, E. (1982b). Of what use are subspecies? *Auk*, 99, 593–595.

Munro, K. & Burg, T. M. (2017) Southern Ocean seabirds: patterns and processes. *Emu*. 117, 4–18.

Phillips, R. A., Croxall, J. P., Silk, J. R. D., & Briggs, D. R. (2007). Foraging ecology of albatrosses and petrels from South Georgia: two decades of insights from tracking technologies. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *17*(S1), S6–S21. https://doi.org/10.1002/AQC.906

Prince, P. A., Rothery, P., Croxall, J. P., & Wood, A. G. (1994). Population dynamics of black-browed and grey-headed albatrosses *Diomedea melanophris* and *D. chrysostoma* at Bird Island, South Georgia. *Ibis*, *136*(1), 50–71. https://doi.org/10.1111/J.1474-919X.1994.TB08131.X

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/GENETICS/155.2.945

Rains, D., Weimerskirch, H., & Burg, T. M. (2011). Piecing together the global population puzzle of wandering albatrosses: Genetic analysis of the Amsterdam albatross *Diomedea amsterdamensis*. *Journal of Avian Biology*, *42*(1), 69–79. https://doi.org/10.1111/j.1600-048X.2010.05295.x

Raymond M., & Rousset F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86:248–249.

Rousset, F. (2008). Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sanchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, *34*(12), 3299–3302. https://doi.org/10.1093/MOLBEV/MSX248

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671-675. https://doi.org/10.1038/nmeth.2089

Schoombie, S., Dilley, B. J., Davies, D., Glass, T., & Ryan, P. G. (2017). The distribution of breeding sooty albatrosses from the three most important breeding sites: Gough, Tristan and the Prince Edward Islands. *Emu-Austral Ornithology, 117*(2), 160–169. https://doi.org/10.1080/01584197.2017.1289804

Schoombie, S., Connan, M., Dilley, B. J., Davies, D., Makhado, A. B., & Ryan, P. G. (2022). Non-breeding distribution, activity patterns and moulting areas of sooty albatrosses *(Phoebetria fusca*) inferred from geolocators, satellite trackers and biochemical markers. *Polar Biology*, *45*, 1–14.

Steeves, T. E., Anderson, D. J., & Friesen, V. L. (2005). A role for non-physical barriers to gene flow in the diversification of a highly vagile seabird, the masked booby (*Sula dactylatra*). *Molecular Ecology*, 14, 3877–3887. https://doi.org/10.1111/J.1365-294X.2005.02713.X

Sun, Z., Gómez-Díaz, E., Bailie, A., & Friesen, V. (2009). Isolation and characterization of microsatellite loci for storm-petrels. *Molecular Ecology Resources*, *9*(3), 913–915. https://doi.org/10.1111/J.1755-0998.2008.02487.X

Swales. M. K. (1965). The seabirds of Gough Island. *Ibis*, *107*, 17–42.

Syväranta, J., Lensu, A., Marjomäki, T. J., Oksanen, S., & Jones, R. I. (2013). An empirical evaluation of the utility of convex hull and standard ellipse areas for assessing population niche widths from stable isotope data. *PLoS One*, *8*(2), e56094.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, *28*(10), 2731–2739. https://doi.org/10.1093/MOLBEV/MSR121

Taylor, R. S., Bailie, A., Gulavita, P., Birt, T., Aarvak, T., Anker-Nilssen, T., Barton, D. C., Lindquist, K., Bedolla-Guzmán, Y., Quillfeldt, P., & Friesen, V. L. (2018). Sympatric population divergence within a highly pelagic seabird species complex (*Hydrobates* spp.). *Journal of Avian Biology*, *49*, e01515. https://doi.org/10.1111/JAV.01515

Taylor, R. S., Bolton, M., Beard, A., Birt, T., Deane-Coe, P., Raine, A. F., González-Solís, J., Lougheed, S. C., & Friesen, V. L. (2019). Cryptic species and independent origins of allochronic populations within a seabird species complex

(*Hydrobates* spp.). *Molecular Phylogenetics and Evolution*, *139*, 106552. https://doi.org/10.1016/J.YMPEV.2019.106552

Troscianko, J. (2022). Chart-Based Cone-Catch Model. http://empiricalimaging.com/knowledge-base/chart-based-cone-catch-model (6 March 2022, date last accessed).

Troscianko, J., & Stevens, M. (2015). Image calibration and analysis toolbox–a free software suite for objectively measuring reflectance, colour and pattern. *Methods in Ecology and Evolution*, 6(11), 1320-1331. https://doi.org/10.1111/2041-210X.12439

Vignal, A., Milan, D., SanCristobal, M., & Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, *34*, 1–31. https://doi.org/10.1186/1297-9686-34-3-275/METRICS

Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques 10,* 506–513

Weimerskirch, H. (2007). Are seabirds foraging for unpredictable resources? *Deep-Sea Research Part II: Topical Studies in Oceanography*, *54*(3–4), 211–223. https://doi.org/10.1016/J.DSR2.2006.11.013

Weimerskirch, H., Brothers, N., & Jouventin, P. (1997). Population dynamics of wandering albatross *Diomedea exulans* and Amsterdam albatross *D. amsterdamensis* in the Indian Ocean and their relationships with longline fisheries: Conservation implications. *Biological Conservation*, *79*(2–3), 257–270. https://doi.org/10.1016/S0006-3207(96)00084-5

Wheeler, Q. D., and Platnick, N. I. (2000). The phylogenetic species concept. Species Concepts and Phylogenetic Theory: A Debate. Columbia University Press, NY, Pp. 55-69

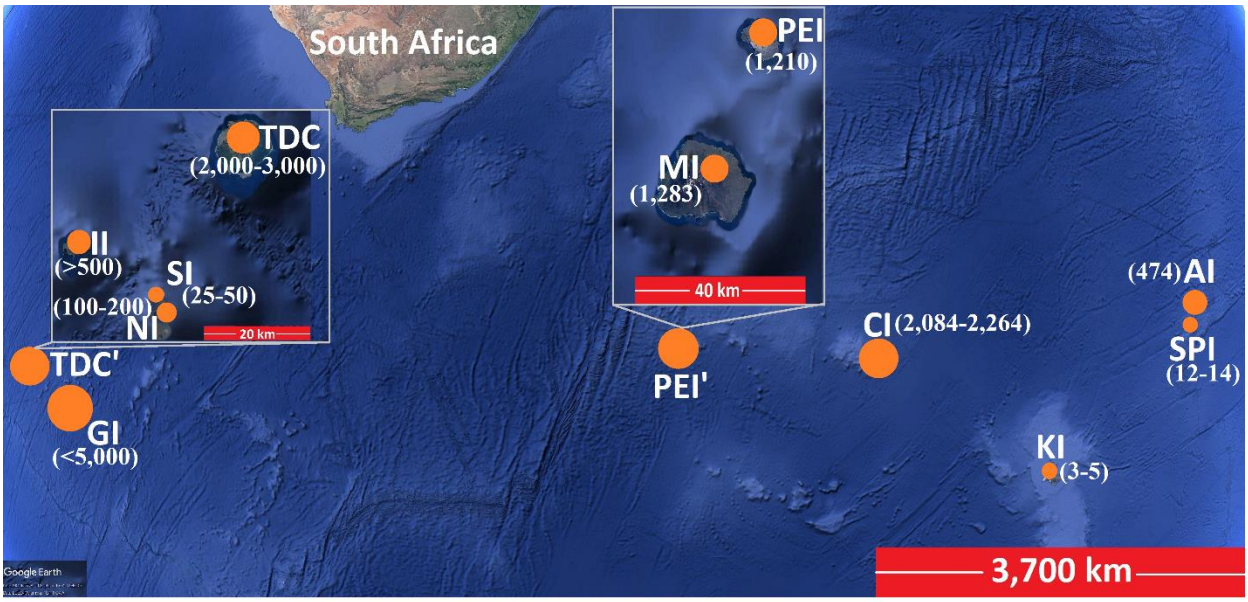

Figure 2.1: Breeding sites of sooty albatross (*Phoebetria fusca*). Insets show four breeding islands within the Tristan archipelago (TDC') and two breeding sites within the Prince Edward Islands group (PEI'). Tristan da Cunha (TDC), Nightingale (NI), Inaccessible (II), Stoltenhoff (SI) Gough (GI), Marion (MI), Prince Edward (PEI), Crozet (CI), Kerguelen (KI), Amsterdam (AI), and St. Paul (SPI). Samples were obtained from AI, MI, GI and TDC. The numbers within brackets represent the number of annual breeding pairs on each island, and the size of the circles is proportional to the number of breeding pairs (ACAP 2010). Map data: Google Earth, Maxar Technologies, National Centre for Space Studies (CNES), Scripps Institution of Oceanography (SIO), National Oceanic and Atmospheric Administration (NOAA), U.S. Navy, National Geospatial-Intelligence Agency (NGA), General Bathymetric Chart of the Oceans (GEBCO), Lamont-Doherty Earth Observatory (LDEO) Columbia University and National Science Foundation (NSF).

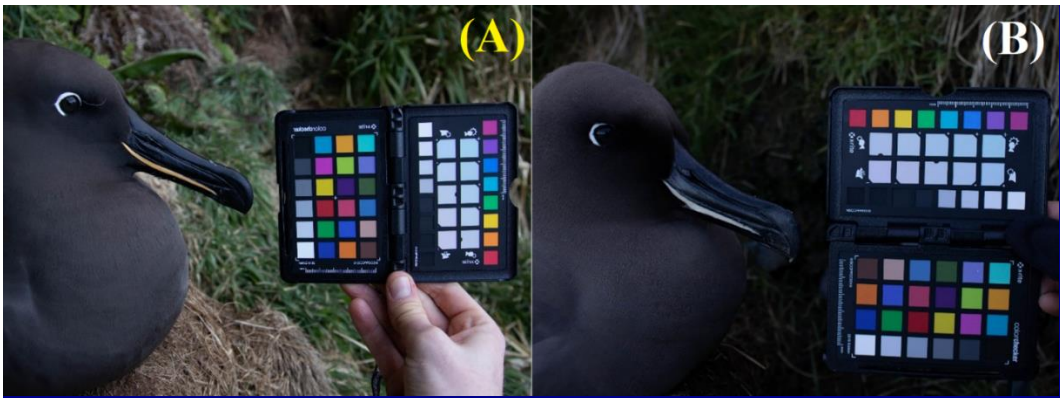

Figure 2.2: Sooty albatross with X-Rite colour-checker passport 2 on (A) Gough Island ©Christopher Jones and (B) Marion Island © Danielle Keys.

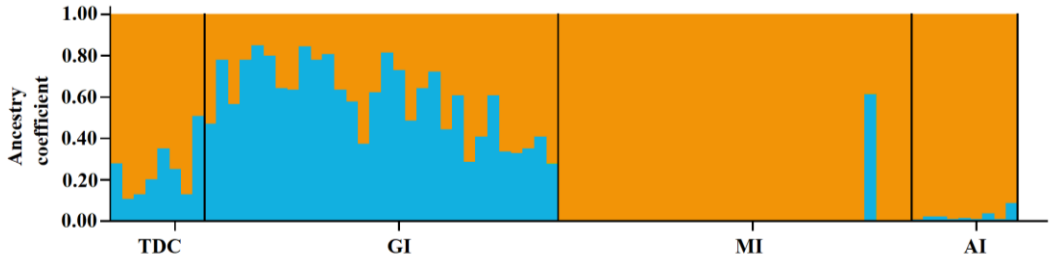

Figure 2.3: STRUCTURE plot for four populations of sooty albatross at K=2. (TDC-Tristan da Cunha: n=8, GI-Gough: n=30, MI-Marion: n=30 and AI-Amsterdam: n=9).

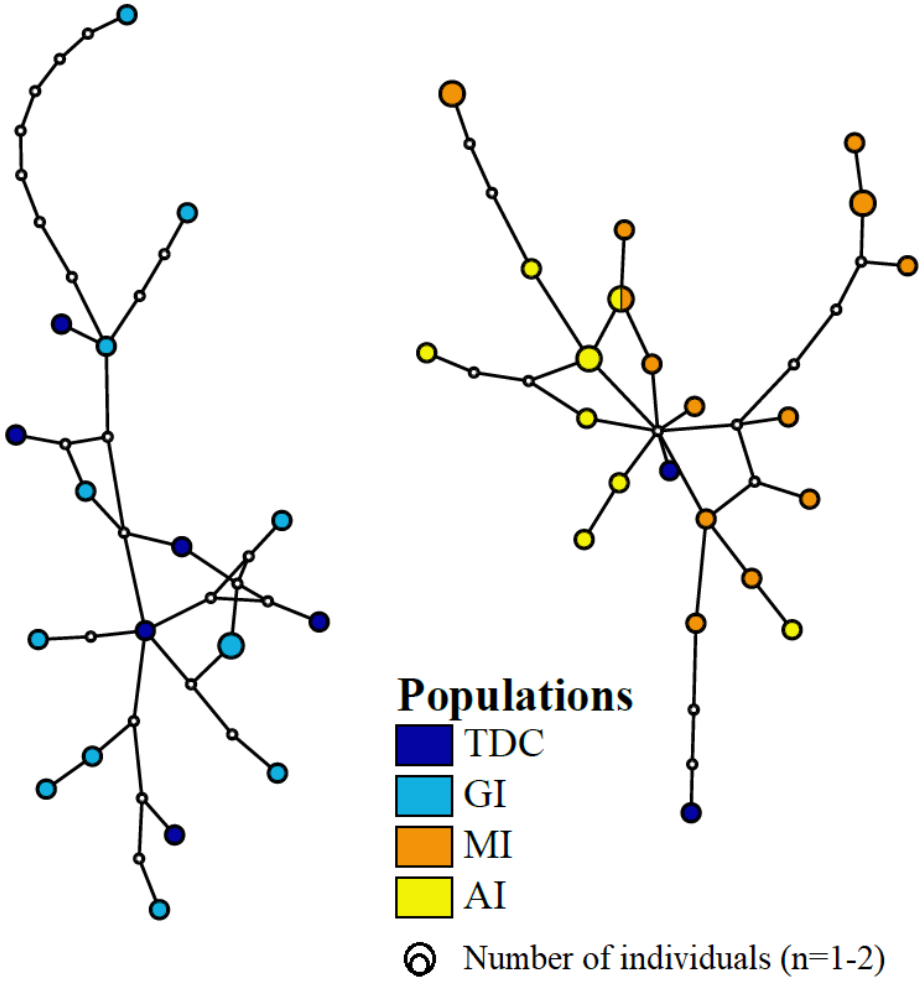

Figure 2.4: Haplotype network from mtCR sequences of 44 sooty albatrosses. Coloured circles represent sampled haplotypes, and the size of the circle is proportional to the number of individuals sharing that haplotype. Small white circles are inferred haplotypes.

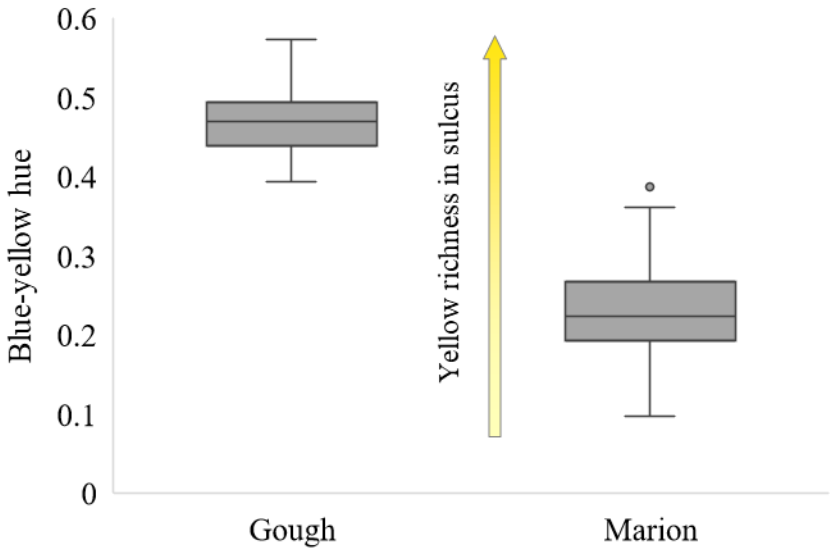

Figure 2.5: Plot summarising blue-yellow hue values for sooty albatross sulcus stripes on Gough and Marion Islands. Higher values indicate richer yellow within the sulcus.

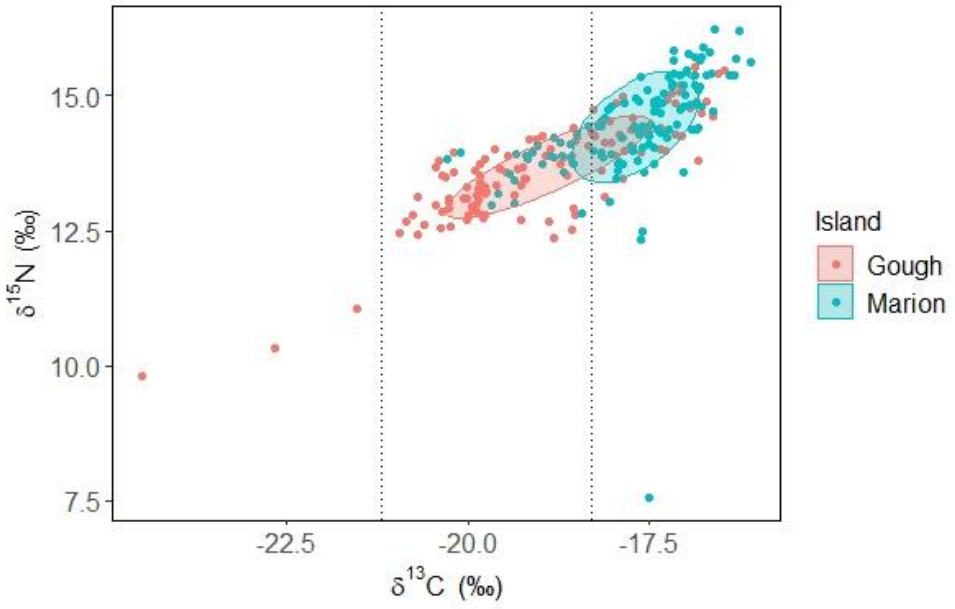

Figure 2.6: Carbon and nitrogen stable isotope values (‰) of individual feathers of sooty albatrosses (n = 30) sampled on breeding adults at Gough and Marion Islands. Standard ellipse areas were estimated using the Stable Isotope Bayesian Ellipses in R (Jackson et al., 2011).

Table 2.1: Matrix of pairwise comparison of F_{ST} (below the diagonal) and P-values (above the diagonal) between the four sampled breeding populations of sooty albatross based on microsatellite data.

	TDC	**GI**	**MI**	**AI**
TDC	-----	<0.001	0.073	0.209
GI	0.088	-----	<0.001	<0.001
MI	0.031	0.120	-----	<0.001
AI	0.038	0.155	0.116	-----

Table 2.2: Number of individuals sequenced, nucleotide diversity, number of haplotypes and haplotype diversity observed in the mtCR sequences for sooty albatrosses from four sampled breeding populations.

Population	**Number of individuals**	**Nucleotide diversity**	**Number of haplotypes**	**Haplotype diversity**
TDC	8	0.0307	8	1.000
GI	12	0.0174	11	0.985
MI	15	0.0150	13	0.981
AI	9	0.0095	8	0.972

2.11 Supplementary materials

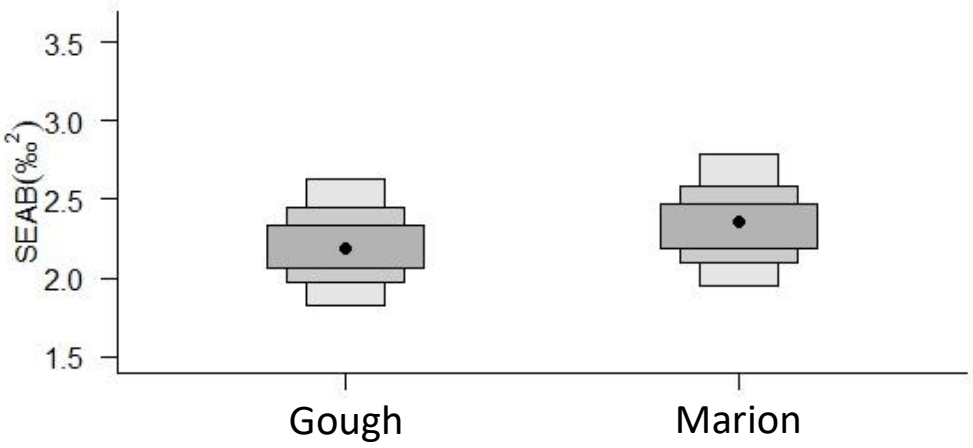

Figure S2.1: Bayesian standard ellipse areas estimated from 120 feathers (30 individuals) collected from sooty albatrosses breeding on Gough and Marion Islands.

Chapter 3: Cryptic speciation and population differentiation in the yellow-nosed albatross species complex

Dilini K. Abeyrama[1], Zach W. Dempsey[1], Peter G. Ryan[2] and Theresa M. Burg[1]

1 University of Lethbridge, Lethbridge, Alberta, Canada

2 FitzPatrick Institute of African Ornithology, DST-NRF Centre of Excellence, University of Cape Town, Rondebosch 7701, South Africa

Corresponding author

Dilini K. Abeyrama: dilini.abeyrama@uleth.ca

ORCID ID:

DKA 0000-0002-9724-3708

ZWD

PGR 0000-0002-3356-2056

TMB 0000-0001-5096-3479

*This manuscript was published in Conservation Genetics on 3rd of April 2021. The author contributions are listed under section 3.7.

Citation: Abeyrama, D. K., Dempsey, Z. W., Ryan, P. G., & Burg, T. M. (2021). Cryptic speciation and population differentiation in the yellow-nosed albatross species complex. *Conservation Genetics*, *22*(5), 757–766. The author contributions are listed under section 3.7.

3.1 Abstract

The two species of yellow-nosed albatross, Atlantic (*Thalassarche chlororhynchos*) and Indian (*T. carteri*), are morphologically similar, but they differ in breeding behaviour and distribution. Both species are listed as endangered by the IUCN due to the limited number of breeding sites, threats from introduced predators and diseases, and impact of commercial fishing. We quantified genetic variation between and within the two species. Using nuclear (microsatellites and two nuclear sequences) and mitochondrial (control region) markers, we analysed 354 samples from four breeding islands (Atlantic: Nightingale, Inaccessible, and Gough; Indian: Amsterdam) and bycatch samples from South Africa and New Zealand. In addition to all markers separating the two species, nuclear markers showed Atlantic yellow-nosed albatrosses from Gough Island are genetically distinct from those breeding at Nightingale and Inaccessible Islands in the Tristan da Cunha archipelago. Nuclear markers confirmed that all bycatch samples were Indian yellow-nosed albatrosses, however, the bycatch birds from South Africa and New Zealand were distinct from each other and from birds breeding on Amsterdam Island, suggesting colony specific dispersal at sea. Our study supports the current recognition of two yellow-nosed albatross species and recognises genetically distinct groups of both Atlantic and Indian yellow-nosed albatross breeding on different islands, which is important for their conservation and management.

Keywords: Thalassarche, genetic diversity, microsatellite, conservation, threatened species

3.2 Introduction

Reproductively isolated populations evolve independently and can give rise to new species (Coyne & Orr, 2004). Reproductive isolation can occur as a result of physical and non-physical barriers restricting gene flow between populations (Friesen et al., 2007). In homogenous environments, the level of intraspecific genetic differentiation is believed to be low and driven by sympatric speciation. Recent studies have shown allopatric speciation can also occur in these homogenous environments as a result of cryptic barriers (Bergek & Björklund, 2007). Isolation due to physical barriers is more readily discerned, whereas non-physical barriers, such as behavioural segregation, are more difficult to identify. Ecological and behavioural differences including song, diet, breeding phenology and migration patterns can act as non-physical barriers to gene flow (Friesen et al., 2007; Amaral et al., 2012; Viricel & Rosel, 2014). Such barriers are species-specific and depend on factors such as dispersal ability, climate, habitat preference and food availability (Burg & Croxall, 2001; Geffen et al., 2004).

Oceanic islands provide an ideal environment to study speciation. Some islands are separated by large distances and others are part of archipelagos. For marine species, various factors act as barriers to gene flow regardless of the organism's dispersal potential or location. A number of studies have identified population structure in marine species with a diverse range of dispersion modes. Viricel & Rosel (2014) found population differentiation in a highly mobile marine species, the Atlantic spotted dolphin (*Stenella frontalis*), however, genetic differences and geographic distance were independent in the absence of physical barriers. Some populations separated by larger distances (~4,500 km) were genetically similar while other populations in close proximity (~1,800 km) were genetically distinct. Other studies found a combination of factors such as geographical distance, geography and habitat

use affect population genetic structure (Riginos & Nachman, 2001) or population structure can occur without any obvious barrier (Knutsen et al., 2003).

Oceanic seabirds are good examples of behavioural barrier-mediated speciation (Friesen et al., 2007). For oceanic seabirds, including albatross, philopatry, non-overlapping distributions, and breeding phenology are important factors in restricting gene flow (Friesen et al., 2007). Strong natal philopatry, combined with non-overlapping at-sea distributions correspond to patterns of population genetic structure in a number of seabirds (Abbott & Double, 2003a; Rains et al., 2011; Silva et al., 2015). However, even species with extreme philopatry, occasionally disperse to other breeding colonies such as wandering albatrosses (*Diomedea exulans*) which have been reported breeding ~1,450 km away from their natal site (Inchausti & Weimerskirch, 2002). Non-physical barriers including at-sea distribution play a prominent role in speciation, especially in the Southern Ocean where land barriers are rare. For example, Burg and Croxall (2001) detected genetically distinct groups corresponding to non-breeding distribution not proximity of breeding sites in black-browed albatross (*Thalassarche melanophris*) while Taylor et al. (2019) found temporal isolation with two breeding periods in the same region for band-rumped storm petrels (*Hydrobates castro*).

The yellow-nosed albatross (*Thalassarche chlororhynchos*) was split into the nominate Atlantic form and Indian species (*T. carteri*) in the 1990s (Nunn et al., 1996; Nunn & Stanley, 1998). This taxonomic revision is not universally accepted. Penhallurick and Wink (2004) suggested that they are better recognised as two subspecies of *T. chlororhynchos*; however, Rheindt & Austin (2005) questioned Penhallurick and Wink's recommendation, leaving the taxonomic status of two yellow-nosed albatross species in controversy. The two species have similar

morphology, but show differences in behaviour, plumage and at-sea distribution (ACAP, 2009a, b; Flood, 2015; BirdLife International, 2018a, b). Their breeding ranges are distinct, but their at-sea ranges overlap around southern Africa (Figure 3.1; Figure S3.1; BirdLife International Seabird Tracking Database). Quantifying genetic variation within and among the two species is important because both are listed as endangered (BirdLife International, 2018). The main reasons for population declines are low rates of reproductive success due to diseases, threats from introduced predators and fishing-related mortality (Gales et al., 1998; Weimerskirch 2004; Angel & Cooper, 2006; Wanless et al., 2007; Bugoni et al., 2008). Global population estimates of mature individuals are 35,000-73,500 Atlantic and ~82,000 Indian yellow-nosed albatrosses. Most individuals from both species breed on a single island: ~60% of Atlantic yellow-nosed albatrosses breed on Tristan da Cunha and ~65% of Indian yellow-nosed albatrosses on Amsterdam Island. From 1982 to 2006, the population on Amsterdam Island decreased from 37,000 to 27,000 due to infectious diseases (avian cholera introduced from domestic poultry) and bycatch in commercial fisheries (Rolland et al., 2009, Jaeger et al., 2018). The Atlantic yellow-nosed albatross populations on Gough and Inaccessible Islands in the Tristan archipelago have become stable (Cuthbert et al., 2014; Ryan et al., 2019). According to ACAP (2009b) estimates, the population on Nightingale is increasing while the population on Tristan da Cunha is decreasing (Wolfaardt et al., 2009), but data are scant.

This study uses nuclear and mitochondrial markers to determine if: i) Atlantic and Indian yellow-nosed albatross are genetically distinct, ii) there is genetic structure within the two species corresponding to different breeding islands and iii) genetic markers can be used to determine the origin of bycatch birds killed in fisheries. The

results will provide a better foundation for the conservation and management of the two species as well as their constituent populations.

3.3 Methods

3.3.1 Sampling and DNA extraction

Blood samples were obtained from Atlantic yellow-nosed albatrosses breeding on two islands in the Tristan archipelago (Nightingale, n=42 and Inaccessible, n=23) and Gough Island (n=136). Indian yellow-nosed albatrosses were only sampled on Amsterdam Island (n=33) because other islands where they breed (Prince Edward, Pingouins and Apôtres in the Crozet archipelago, and Île de Croy, Nuageuses, Kerguelen) are very seldom visited (no landings since 2008 at Prince Edward, and 1980s for French islands; Weimerskirch et al. 2018). Bycatch samples of Indian yellow-nosed albatrosses from South Africa (n=38) and New Zealand (n=82) were compared with birds from the sampled breeding colonies to determine their provenance. DNA was extracted from all the samples using a modified chelex extraction protocol (Walsh et al., 1991; Burg and Croxall, 2001) and stored at -20°C.

3.3.2 Nuclear DNA analyses

3.3.2.1 Occa9 and Pema7 genotyping

Primers from Silva et al. (2011, 2012) were screened for variation. Occa9 and Pema7 showed the highest variation and eight to ten samples per ocean basin (Atlantic and Indian) were sequenced at both loci using previously published primers (Silva et al., 2011; Silva et al., 2012). Within Occa9 a single substitution, G/A, was identified. New primers (Occa9F:5′-ATGGTGGCAGAAGGGACA-3′ and Occa9R Nla: 5′-TGCAGCTAGCTATGGTAATACAGTCA-3′) were developed to target a single nucleotide polymorphic (SNP) site using polymerase chain reaction (PCR) and

genotype with restriction digests (Abbott & Double, 2003a). DNA amplification with Occa9F/ Occa9R Nla generated a 410 bp fragment. The restriction enzyme *NlaIII* cut the fragment into two (109+301 bp) fragments in the presence an A at the targeted SNP. Within Pema7, two SNP sites T/G (Hpy) and G/C (Alu) were identified. Two forward primers (Pema7F157 Alu: 5′-AGACAGTTCTTTGCTCTCTATGGAG-3′ and Pema7F154 Hpy: 5′-TACAGACAGTTCTTTGCTCTCTACG-3′) and a common reverse primer (Pema7Rint dig: 5′-TCTGCAAATCACCAGAGCAC-3′) were developed targeting these SNPs. PCR with Pema7F154 Hpy and the common reverse primer generated a 198 bp product and when a G was present, remained uncut or produced 23+175 bp fragments when a T was present at the polymorphic site and the PCR product was digested with *HpyCH4IV*. Pema7F157 Alu primer generated a 195 bp product. *AluI* digests produced 159+36 bp fragments with a C at the SNP site or 195 bp uncut fragment with a G at the polymorphic site. For all restriction digests, approximately 100-160 ng of amplified DNA was incubated with smartcut buffer and 0.1 units of restriction enzyme at 37°C for 2-4 hours. Digested DNA was visualized on a 3% agarose gel stained with ethidium bromide. Out of 354 samples, 302 samples were screened with Occa9 and 288 with Pema7. SNP genotypes at Occa9 and Pema7 loci were numerically coded as two loci and included in our STRUCTURE analysis (see below) along with microsatellite data.

3.3.2.2 Microsatellite genotyping

A total of 354 samples were screened at 13 microsatellite loci (Cd11, 11H7, Dc22, Dc9, Dc27, De18, Tc65, Oc51, De35, De11, 12E1, Dc26 and 11H1) originally isolated from other Procellariiformes including albatrosses (Burg, 1999; Burg & Croxall, 2001; Abbott & Double, 2003b; Dubois et al., 2005; Bried et al., 2008; Sun et al., 2009) after checking their variability in yellow-nosed albatrosses. DNA was

amplified in 10 µl reactions containing FroggaBio buffer, 0.25 mM dNTP, 0.5 µM forward primer, 1 µM reverse primer, 0.05 µM fluorescent labelled M13 primer and 0.5 units of FroggaBio taq polymerase. PCR consisted of one cycle of 120 s at 94°C, 7 cycles of 60 s at 94°C, 45 s at T_1, 30 s at 72°C, 25 cycles of 30 s at 89°C, 30 s at T_2, 45 s at 72°C and one final cycle of 300 s at 72°C and 20 s at 4°C. For six loci (Cd11, Dc22, Dc27, Oc51, Dc26 and 11H1), T_1=45°C and T_2= 48°C and for other seven microsatellite loci (11H7, Dc9, De18, Tc65, De35, De11 and 12E1) T_1=50°C and T_2=52°C. PCR with T_1=50°C and T_2=52°C was amplified with 28 cycles. Denatured PCR products were run on a 6% polyacrylamide gel in a LI- COR 4300 DNA Analyzer. Three positive and one negative controls were included in each load to calibrate samples across runs and to ensure no contamination. Alleles were scored manually with a second person checking for consistent and accurate scoring. Genepop 4.7 was used to check for deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) (Raymond & Rousset, 2002). The false discovery rate in p-values for HWE and LD were corrected using Benjamini–Hochberg (B-H) procedure (Benjamini & Hochberg, 1995). Microsatellite data were analysed using STRUCTURE v. 2.3.4 with 100K burnin followed by 200K runs, admixture model with locpriors and correlated allele frequencies to identify the optimal number of genetic clusters (K) (Pritchard et al., 2000). Initially, ten replicates were performed for each K value (K=1-8) on all 354 samples from all six sites. Average likelihoods for each K value were calculated using Structure Harvester v. 0.6.94 (Earl & vonHoldt, 2012). Optimum number of genetic clusters within data was selected using ΔK and Evanno method (Earl & vonHoldt, 2012). Hierarchical STRUCTURE analyses were performed for the Atlantic and Indian yellow-nosed albatrosses separately. For both runs 10 replicates were performed for each K (K=1-3) using the same settings used in

the species analysis to examine genetic clustering within each species. Finally, two additional STRUCTURE analyses were performed: one with individuals from Amsterdam Island and South African bycatch, and another with Amsterdam Island and New Zealand bycatch using the same settings as above.

Genetic differences within and among two species were estimated using F_{ST} statistics. An analysis of molecular variance (AMOVA) was used to calculate F_{ST} (Excoffier et al., 1992). AMOVA was performed in Arlequin 3.5 (Excoffier & Lischer, 2010) and significance was determined using 1000 permutations.

3.3.3 Species assignment tests

Species assignment tests were performed using GenAlEx 6.51b2 (Peakall and Smouse, 2012) to check the assignment of bycatch samples and the ability of the genetic markers to differentiate between the two species. First assignment tests were done using samples from the breeding populations and a second set of tests were done on all of the samples.

3.3.4 Mitochondrial DNA analysis

A 365 bp fragment of the mitochondrial control region (mtCR) was amplified using the primers H505 (5′-GAAAGAATGGTCCTGAAGC-3′; Burg & Croxall, 2001) and SpecF1 (5′-CAGCTTATGTATAAATGCG-3′; Abbott & Double, 2003a). DNA was amplified in 25 µl reactions with GoTaq® Flexi buffer, 1.5 mM $MgCl_2$, 0.1 mM dNTP, 0.4 µM forward and reverse primers and 0.1 units of GoTaq® Flexi DNA polymerase. Amplification consisted of one cycle of 120 s at 94°C, 45 s at 54°C, 60 s at 72°C; 37 cycles of 30 s at 94°C, 45 s at 54°C, 60 s at 72°C and one final cycle of 300 s at 72°C. Quality of PCR products was assessed using a 0.8% agarose gel stained with ethidium bromide. The PCR products were sent to Génome Québec, McGill

University, Quebec, Canada for sequencing. DNA sequences were aligned manually in MEGA 5.2 (Tamura et al., 2011), and a haplotype network was constructed using TCS v.1.21 (Clement et al., 2000) and edited using tcsBU (TCS Beautifier; Múrias dos Santos et al., 2016). Nucleotide diversity, number of haplotypes and haplotype diversity for species and populations were calculated using the program DnaSP 6 (Rozas et al., 2017).

Fixed differences between the two species were identified by sequencing 50 individuals from both species, and species-specific forward primers were developed. H289Atl (5′-GGTTTAATCCGTTTAAGAGG-3′) exclusively amplified Atlantic yellow-nosed albatross and H289Ind (5′-GGTTTAATCCGTTTGAGAGT-3′) exclusively identified Indian yellow-nosed albatross. Both primers were used in conjunction with the SpecF1 primer. Using the species-specific primers, a total of 255 samples (including 50 individuals that were sequenced) from Nightingale (n=41), Gough (n=63), Amsterdam (n=33), South African bycatch (n=36), and New Zealand bycatch (n=82) were screened. Genomic DNA was amplified in 10 µl reactions with GoTaq® Flexi buffer, 2 mM $MgCl_2$, 0.08 mM dNTP, 0.4 µM of each SpecF1 and H289Atl/ H289Ind primers and 0.5 units of GoTaq® Flexi DNA polymerase. PCR conditions followed those outlined above but used annealing temperatures of 50°C for H289Atl and 54°C for H289Ind.

3.4 Results

3.4.1 Nuclear DNA analysis

Based on the percentage of individuals with different genotypes at the locus Occa9 (Figure S3.2), around 80% of Indian yellow-nosed albatrosses from all three groups, Amsterdam Island (25/31), South African bycatch (28/35) and New Zealand bycatch

(28/37), were homozygous for the A allele compared to 20% of Atlantic birds from all three islands. Within the Atlantic yellow-nosed albatross, more than 50% of individuals from both Nightingale (20/39) and Inaccessible (13/23) Islands were G/A heterozygotes at the Occa9 SNP whereas about 50% of individuals from Gough Island (67/134) were homozygous for G (Figure S3.2). When analysed with AMOVA, 30.4% and 1.4% ($P<0.05$) of the variation was observed between and within the two species, respectively for Occa9. Occa9 differentiated the two species as well as breeding populations of Atlantic yellow-nosed albatross (Figure S3.2). For Pema7 (Figure S3.3), variation between the two species was 3.9% and within species was 16.7% ($P<0.05$). Pema7 showed highest variability for yellow-nosed albatross among all the nuclear markers.

At the 13 polymorphic microsatellite loci, the number of alleles ranged from 3-16 per locus (Table S3.1). Seven loci deviated significantly from HWE for some populations. Following B-H correction, Dc26 deviated significantly from HWE at four populations, 11H1 and Dc22 at three populations, De11 at two populations and De35, Dc9 and Cd11 at one population. With the exception of Dc26, all loci were in HWE for three or more populations therefore they were included in further analyses. Nightingale and Inaccessible Islands showed significant heterozygote deficiencies. Following B-H correction, none of the loci were in linkage disequilibrium.

STRUCTURE clustering analysis with all 13 microsatellite loci, Occa9 and Pema7 showed an optimal K value of two (mean LnP(K)= -10503 and $\Delta K= 162$) and the two genetic clusters corresponded to the two species (Figure 3.2a). However, four Indian yellow-nosed albatrosses (three from Amsterdam and one South African bycatch) showed genetic admixture (Figure 3.2a). Our analysis of hierarchical structure for Atlantic yellow-nosed albatross showed an optimal K value of two (LnP(K)= -6348

and ΔK=8), separating Atlantic yellow-nosed albatrosses on Gough Island from those on Nightingale and Inaccessible Islands (Figure 3.2b). Our analysis of structure for Indian yellow-nosed albatross showed an optimal K value of two, separating New Zealand bycatch from the Amsterdam population and South African bycatch (Figure 3.2c). The STRUCTURE runs with Amsterdam Island birds and South African bycatch, and Amsterdam Island birds and New Zealand bycatch both showed K=2 separating each of the groups (Figure 3.2d and e).

According to the hierarchical population structure analysis in AMOVA, all populations of Atlantic yellow-nosed albatross were significantly different from all Indian yellow-nosed albatross populations (Table 3.1). Supporting the optimal K value detected by STRUCTURE, the Atlantic yellow-nosed albatross population on Gough Island was significantly different from both Nightingale and Inaccessible Island populations while Nightingale and Inaccessible Islands were not significantly different from each other (Table 3.1). Within the Indian yellow-nosed albatross, the Amsterdam Island population was significantly different from the New Zealand bycatch population while South African bycatch population was not significantly different from Amsterdam or New Zealand bycatch populations (Table 3.1).

3.4.2 Species assignment

Using only known origin birds, 192/201 of the Atlantic yellow-nosed albatrosses and 30/33 of the Indian yellow-nosed albatrosses were correctly assigned to each species using nuclear markers. Six of the nine misassigned Atlantic birds had similar values between the two oceans. When the bycatch samples were included, 33/38 of the South African bycatch and 63/82 of the New Zealand bycatch were assigned to Indian yellow-nosed albatross. Two of the five misassigned South African bycatch and nine

of the nineteen misassigned New Zealand bycatch birds had similar values between the two oceans.

3.4.3 mtDNA

Within the 365 bp mtCR sequences of 111 yellow-nosed albatrosses, 75 haplotypes were identified (Figure 3.3). A total of 48 variable sites were found, 39 in the Atlantic and 31 in the Indian. The haplotype network showed two distinct groups, mostly corresponding to Indian and Atlantic yellow-nosed albatrosses, separated by a minimum of six mutations. While none of the haplotypes were shared between the two species, one Atlantic yellow-nosed albatross from Gough Island clustered with Indian yellow-nosed albatross. Nucleotide and haplotype diversities were similar between the Indian and Atlantic yellow-nosed albatrosses, while Atlantic birds possessed more haplotypes than Indian Ocean birds (Table 3.2). Nucleotide diversity was lower in birds from Inaccessible Island and New Zealand bycatch (0.0117 and 0.0126 respectively vs 0.0161-0.0197 at other sites). Among Atlantic yellow-nosed albatrosses, four haplotypes were shared between Gough and Inaccessible Islands, while none of the Nightingale Island haplotypes were shared with Gough or Inaccessible Islands (Figure 3.3). When samples were screened with species-specific primers, all 33 samples from Amsterdam Island possessed the Indian Ocean mtCR genotype along with two from Nightingale Island; the remaining birds from Gough (n= 62) and Nightingale (n=39) had the Atlantic Ocean genotype. All the samples from Inaccessible Island were sequenced and contained the Atlantic mtCR genotype. All South African and 79 out of 82 New Zealand bycatch samples possessed the Indian Ocean mtCR genotype. The three remaining New Zealand bycatch samples had the Atlantic Ocean mtCR genotype.

3.5 Discussion

Using a combination of nuclear and mitochondrial markers, we found population structuring in yellow-nosed albatross. Both nuclear and mitochondrial markers confirmed Atlantic and Indian yellow-nosed albatrosses as two genetically distinct groups with additional genetic differentiation within both Atlantic and Indian yellow-nosed albatrosses. These results will allow for assignment of at-sea birds not only to species, but to determine ancestry to breeding population based on their unique genetic structuring.

3.5.1 Assignment of breeding sites for bycatch groups of Indian yellow-nosed albatross

All individuals from the Indo-Pacific Ocean (one breeding site and two sets of bycatch samples) were identified as Indian yellow-nosed albatrosses using nuclear markers and all except three New Zealand bycatch samples with mitochondrial markers. The two sets of bycatch samples did not originate from the largest breeding population on Amsterdam Island. Adults from Amsterdam tend to disperse east, however, little is known about the at-sea distribution of immature birds which typically account for ~50% of albatross populations. The fact that we find differences between Amsterdam and the two sets of bycatch samples and no mixing of birds from the sampled populations suggests each population might have different at-sea ranges. The South African bycatch samples are too far west to have originated from Amsterdam Island and the STRUCTURE analysis shows the bycatch samples are genetically distinct from Amsterdam birds. This leaves two main breeding colonies on Prince Edward Island and Îles Crozet. The at-sea distribution of birds from Îles Crozet is unknown; however, Makhado et al. (2018) found breeding adults from Prince Edward Island forage off the south and east coasts of South Africa, in the same area where the

bycatch birds were caught. Three of the South African bycatch birds killed in October and November had a brood patch suggesting they were nesting; Prince Edward Island is the closest breeding site (~1,800 km vs Crozet ~2,500 km). The 35 other bycatch samples from the same area were genetically similar to the three with brood patches and therefore likely came from the same population. All 82 of the New Zealand bycatch samples formed a third genetic group based on nuclear data. Other than Îles Crozet, the only two other breeding sites are St. Paul (~90 km south of Amsterdam Island, < 20 breeding birds; ACAP, 2009a) and Kerguelen (~1,250 km east of Îles Crozet, ~1,500 km southwest of Amsterdam Island, ~50 breeding birds; Weimerskirch et al., 2018). Both St. Paul and Kerguelen populations are too small to survive a loss of 82 individuals. Thus, the New Zealand bycatch is most likely comprised of birds predominantly from Îles Crozet. This supports our assumption and Makhado et al. (2018) of substantial partitioning of foraging grounds of Indian yellow-nosed albatrosses, limiting the extent of at-sea mixing of individuals from different breeding populations. Other studies found that seabirds with non-overlapping at-sea distributions are genetically distinct (Burg and Croxall, 2001; Abbott and Double, 2003a; Rains et al., 2011; Silva et al., 2015).

3.5.2 Interspecific and intraspecific genetic variation

Except for the distinct breeding distribution, which is an obvious genetic barrier for the two yellow-nosed albatross species, there are no large land masses separating their breeding ranges. Therefore, the physical distance between the breeding islands might act as a barrier for gene flow (Wright, 1943). A similar pattern is found in masked boobies (*Sula dactylatra*) where genetic structure corresponds to their distribution in Indo-Pacific and Atlantic Oceans (Steeves et al., 2005).

Population differentiation is observed in both Atlantic and Indian yellow-nosed albatrosses. For Atlantic yellow-nosed albatross, Nightingale and Inaccessible Islands are located about 20 km apart while Gough Island is about 380 km southeast of Nightingale Island, the nearest island in the Tristan da Cunha archipelago. Physical distance is the only known barrier between these populations. In the closely related black-browed albatross, two genetic groups were found in the southwest Atlantic, each with distinct at-sea foraging grounds (Burg and Croxall, 2001). Foraging ranges of Atlantic yellow-nosed albatross from different islands, span the entire South Atlantic Ocean and show overlap between islands with some evidence of differential dispersal within Tristan da Cunha archipelago (ACAP, 2009b; BirdLife International Seabird Tracking database). All recoveries of Atlantic yellow-nosed albatross from Nightingale have been from Africa, whereas most Tristan and Inaccessible recoveries have been from South America. All tracking of Nightingale breeding birds (N=19) show they go to the Benguela, whereas some Inaccessible breeders go to South America and birds from Gough disperse in both directions (ACAP, 2009b; BirdLife International Seabird Tracking database). Therefore, the reasons for genetic differentiation among the Atlantic yellow-nosed albatrosses from Nightingale, Inaccessible and Gough Islands are not clear.

Based on our earlier hypothesis regarding the origin of the bycatch samples, birds from Amsterdam, Prince Edward Island (South African bycatch) and Îles Crozet (New Zealand bycatch) are genetically distinct. There appears to be some gene flow from Prince Edward Island to both Amsterdam and the two Crozet colonies as some individuals showed admixture at the nuclear loci (Figure 3.2c) while there does not appear to be gene flow between Amsterdam and Îles Crozet. While limited morphological and behavioural data are available from Prince Edward and Crozet,

there is no evidence for morphological or behavioural differences among Indian yellow-nosed albatrosses from different breeding populations except for the substantial partitioning in foraging ranges (Makhado et al., 2018). Therefore, the possible reason for population differentiation in Indian yellow-nosed albatross is not clear. But as mentioned above, natal philopatry or physical distance may act as genetic barriers.

The three Atlantic yellow-nosed albatrosses from Gough and Nightingale Islands with Indian Ocean mtCR genotypes were not identified as Indian yellow-nosed albatross using nuclear markers. Similarly, three Indian yellow-nosed albatross bycatch samples which possessed Atlantic Ocean mtCR genotype were not identified as Atlantic yellow-nosed albatross using nuclear markers. Possible reasons why a small number of individuals have mtDNA haplotypes similar to their sister species include incomplete lineage sorting, migration or hybridization. None of these birds had more than 35% of missing data for 15 nuclear loci or showed evidence of admixture. This suggests that these six individuals are not recent migrants, visitors or hybrids. Although the breeding sites of the two species are restricted to different ocean basins with minimal overlap in their foraging ranges, a few cases of individuals visiting breeding sites of other species have been recorded. In early 2019, an Indian yellow-nosed albatross was spotted on Gough Island for the first time (Jones et al., 2020). Similarly, on Chatham Island, New Zealand, a pair of Indian yellow-nosed albatross has been recorded, suggesting a possible range expansion of the species (Gales, 1998). Both cases show a pattern of vagrancy in Indian yellow-nosed albatross. Although hybridization is rare in many species of seabirds, a few hybrid chicks have been identified in other *Thalassarche* albatross species (Moore et al., 1997; Phalan et al.,

2004; Phillips et al., 2018; Jones et al., 2020), but there are no records of hybridization between the two yellow-nosed albatross species.

3.5.3 Conservation and management

The patterns of genetic variation in both the Atlantic and Indian yellow-nosed albatrosses suggest they should be treated as separate units for conservation and management. Populations on different islands facing different problems may require individual management. For Atlantic yellow-nosed albatross, we did not have samples from the largest breeding colony on Tristan (~60% of the population), but samples from the three other large breeding populations show fine scale genetic structure over short distances. Since Nightingale and Inaccessible populations are genetically similar, it is more likely birds from Tristan da Cunha Island in the same archipelago belong to the same genetic cluster. Gough contains ~20% of the species and is genetically distinct from Nightingale and Inaccessible Islands which contain ~4% and 16%, respectively, of the Atlantic yellow-nosed albatross. The largest population of Indian yellow-nosed albatross on Amsterdam Island is genetically distinct from the other populations and the bycatch data suggest two other genetically distinct groups of Indian yellow-nosed albatross populations are also present; likely from Îles Crozet and Prince Edward Island.

The species-specific primers developed in our study allowed us to differentiate Atlantic and Indian yellow-nosed albatrosses with a high accuracy and low cost. Using this knowledge, we can develop a rapid molecular test to identify the bycatch yellow-nosed albatross samples which cannot be identified from their morphology.

Using genetic analyses combined with genetic data from breeding sites and bycatch birds, along with at-sea locations of bycatch birds, we were able to determine not only

the at-sea distribution of birds from colonies which are monitored regularly but also colonies which are seldom visited further assisting in conservation efforts and their genetic differentiation. The at-sea distribution data can be used to help manage the commercial fishing by minimizing the interaction of birds at-sea with fishing vessels. Further studies including samples from the two large Indian Ocean breeding sites of Prince Edward and Crozet as well as using a genomic approach may help to uncover additional genetic differences among yellow-nosed albatross populations.

3.6 Acknowledgements

We thank Henri Weimerskirch and Thierry Boulinier for providing samples from Amsterdam Island. Ben Dilley assisted with sample collection from Tristan da Cunha archipelago and Gough Island. Belinda Bauer, A Hewer and Tasmanian Museum and Art Gallery generously provided us samples. Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship and the Seabird Group provided funding. Special thanks to members of Burg lab, University of Lethbridge for their enormous support.

3.7 Author contributions

PGR and TMB designed the study and PGR performed the fieldwork. ZWD and DKA completed the lab work and DKA led the analyses. DKA and TMB wrote the manuscript with input from the others.

3.8 Funding

Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship, and the Seabird Group.

3.9 Data availability

Sequencing and genotype data supporting the findings of this study are available at dryad and GenBank.

3.10 Declarations

Conflict of interest: The authors declare that they have no conflict of interest.

3.11 References

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

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

Agreement on the Conservation of Albatrosses and Petrels (2009a). ACAP Species assessment: Indian yellow-nosed albatross *Thalassarche carteri*.

Agreement on the Conservation of Albatrosses and Petrels (2009b). ACAP Species assessment: Atlantic yellow-nosed albatross *Thalassarche chlororhynchos*.

Amaral, A. R., Beheregaray, L. B., Bilgmann, K., Boutov, D., Freitas, L., Robertson, K. M., Sequeira, M., Stockin, K. A., Coelho, M. M., & Möller, L. M. (2012). Seascape genetics of a globally distributed, highly mobile marine mammal: the short-beaked common dolphin (genus *Delphinus*). *PLoS One*, 7, e31482.

Angel, A., & Cooper, J. (2006). A review of the impacts of introduced rodents on the islands of Tristan da Cunha and Gough. Royal Society for the Protection of Birds, Sandy, UK.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57, 289–300.

Bergek, S., & Björklund, M. (2007). Cryptic barriers to dispersal within a lake allow genetic differentiation of Eurasian perch. *Evolution*, 61, 2035–2041.

BirdLife International (2018a). *Thalassarche carteri*. The IUCN Red List of Threatened Species 2018: e. T22728372A132657962. Downloaded on 20 September 2019.

BirdLife International (2018b). *Thalassarche chlororhynchos*. The IUCN Red List of Threatened Species 2018: e.T22698425A132645225. Downloaded on 20 September 2019.

Bried, J., Dubois, M. P., Jouventin, P., & Santos, R. S. (2008). Eleven polymorphic microsatellite markers in Cory's shearwater, *Calonectris diomedea*, and cross-species amplification on threatened Procellariiformes. *Molecular Ecology Resources*, 8, 602–604.

Bugoni, L., Mancini, P. L., Monteiro, D. S., Nascimiento, L., & Neves, T. (2008). Seabird bycatch in the Brazilian pelagic longline fishery and a review of capture rates in the southwestern Atlantic Ocean. *Endangered Species Research*, 5, 137–147.

Burg, T. M. (1999). Isolation and characterization of microsatellites in albatrosses. *Molecular Ecology*, 8, 338–341.

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.

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

Coyne, A. E., & Orr, H. A. (2004). Speciation. Sinauer Associates, Sunderland.

Cuthbert, R. J., Cooper, J., & Ryan, P. G. (2014). Population trends and breeding success of albatrosses and giant petrels at Gough Island in the face of at-sea and on-land threats. *Antarctic Science*, 26, 163–171.

Dubois, M. P., Jarne, P., & Jouventin, P. (2005). Ten polymorphic microsatellite markers in the wandering albatross *Diomedea exulans*. *Molecular Ecology Notes*, 5, 905–907.

Earl, D. A., & von Holdt, B. M. (2012). Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.

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.

Flood, R. L. (2015). The two yellow-nosed albatrosses: Molt, age, and identification. *Birding*, 47, 36–47.

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

Gales, R. (1998). Albatross populations: status and threats, in Albatross Biology and Conservation, G. Robertson and R. Gales (Eds). Surrey Beatty & Sons, Chipping Norton.

Gales, R., Brothers, N., & Reid, T. (1998). Seabird mortality in the Japanese tuna longline fishery around Australia, 1988-1995. *Biological Conservation,* 86, 37–56.

Geffen, E. L. I., Anderson, M. J., & Wayne, R. K. (2004). Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology*, 13, 2481–2490.

Inchausti, P., & Weimerskirch, H. (2002). Dispersal and metapopulation dynamics of an oceanic seabird, the wandering albatross, and its consequences for its response to long-line fisheries. *Journal of Animal Ecology*, 71, 765–770.

Jaeger, A., Lebarbenchon, C., Bourret, V., Bastien, M., Lagadec, E., Thiebot, J. B., Boulinier, T., Delord, K., Barbraud, C., Marteau, C., Dellagi, K., Tortosa, P., &

Weimerskirch, H. (2018). Avian cholera outbreaks threaten seabird species on Amsterdam Island. *PLoS One*, 13, e0197291.

Jones, M. G. W., Techow, N. M. S., Risi, M. M., Jones, C. W., Hagens, Q. A., Taylor, F., & Ryan, P. G. (2020). Hybridisation and cuckoldry between black-browed and grey-headed albatrosses. *Antarctic Science*, 32, 10–14. doi:10.1017/S0954102019000506

Knutsen, H., Jorde, P. E., André, C., & Stenseth, N. C. (2003). Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. *Molecular Ecology*, 12, 385–394.

Makhado, A. B., Crawford, R. J. M., Dias, M. P., Dyer, B. M., Lamont, T., Pistorius, P., Ryan, P. G., Upfold, L., Weimerskirch, H., & Reisinger, R. R. (2018). Foraging behaviour and habitat use by Indian yellow-nosed albatrosses (*Thalassarche carteri*) breeding at Prince Edward Island. *Emu*, 118, 353–362.

Moore, P. J., Taylor, G. A., & Amey, J. M. (1997). Interbreeding of black-browed albatross *Diomedea m. melanophris* and New Zealand black-browed albatross *D. m. impavida* on Campbell Island. *Emu*, 97, 322–324.

Múrias dos Santos, A., Cabezas, M. P., Tavares, A. I., Xavier, R., & Branco, M. (2016). tcsBU: a tool to extend TCS network layout and visualization. *Bioinformatics*, 32, 627–628.

Nunn, G. B., Cooper, J., Jouventin, P., Robertson, C. J., & Robertson, G. G. (1996). Evolutionary relationships among extant albatrosses (Procellariiformes: Diomedeidae) established from complete cytochrome-b gene sequences. *Auk*, 113, 784–801.

Nunn, G. B., & Stanley, S. E. (1998). Body size effects and rates of cytochrome b evolution in tube-nosed seabirds. *Molecular Biology Evolution*, 15, 1360–1371.

Peakall, R., & Smouse, P. E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28, 2537–2539.

Penhallurick, J., & Wink, M. (2004). Analysis of the taxonomy and nomenclature of the Procellariiformes based on complete nucleotide sequences of the mitochondrial cytochrome b gene. *Emu*, 104, 125–147.

Phalan, B., Phillips, R. A., & Double, M. C. (2004). A white-capped albatross, *Thalassarche* [*cauta*] *steadi*, at South Georgia: first confirmed record in the south-western Atlantic. *Emu*, 104, 359–361.

Phillips, R. A., Cooper, J., & Burg, T. M. (2018). Breeding-site vagrancy and hybridization in albatrosses. *Ibis*, 160, 907–913.

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

Rains, D., Weimerskirch, H., & Burg, T. M. (2011). Piecing together the global population puzzle of wandering albatrosses: genetic analysis of the Amsterdam albatross *Diomedea amsterdamensis*. *Journal of Avian Biology*, 42, 69–79.

Raymond, M., & Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248–249.

Rheindt, F. E., & Austin, J. J. (2005). Major analytical and conceptual shortcomings in a recent taxonomic revision of the Procellariiformes—a reply to Penhallurick and Wink (2004). *Emu-Austral Ornithology*, 105, 181–186.

Riginos, C., & Nachman, M. W. (2001). Population subdivision in marine environments: the contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. *Molecular Ecology*, 10, 1439–1453.

Rolland, V., Barbraud, C., & Weimerskirch, H. (2009). Assessing the impact of fisheries, climate and disease on the dynamics of the Indian yellow-nosed albatross. *Biological Conservation*, 142, 1084–1095.

Ryan, P. G., Dilley, B. J., & Ronconi, R. A. (2019). Population trends of spectacled petrels *Procellaria conspicillata* and other seabirds at Inaccessible Island. *Marine Ornithology*, 47, 257–265.

Silva, M. C., Duarte, M. A., & Coelho, M. M. (2011). Anonymous nuclear loci in the white-faced storm-petrel *Pelagodroma marina* and their applicability to other procellariiform seabirds. *Journal of Heredity*, 102, 362–365.

Silva, M. C., Silva, M. F., & Coelho, M. M. (2012) Anonymous nuclear loci in Madeiran storm-petrel *Oceanodroma castro* (Procellariiformes: Hydrobatidae) and applicability across other procellariiform families. *Conservation Genetics Resources*, 4, 1093–1095.

Silva, M. C., Matias, R., Wanless, R. M., Ryan, P. G., Stephenson, B. M., Bolton, M., Ferrand, N., & Coelho, M. M. (2015). Understanding the mechanisms of antitropical divergence in the seabird White-faced storm-petrel (Procellariiformes: *Pelagodroma marina*) using a multilocus approach. *Molecular Ecology*, 24, 3122–3137.

Steeves, T. E., Anderson, D. J., & Friesen, V. L. (2005). A role for nonphysical barriers to gene flow in the diversification of a highly vagile seabird, the masked booby (*Sula dactylatra*). *Molecular Ecology*, 14, 3877–3887.

Sun, Z., Gómez-Diaz, E., Bailie, A., & Friesen, V. (2009). Isolation and characterization of microsatellite loci for storm-petrels. *Molecular Ecology Resources*, 9, 913–915.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. *Molecular Biology and Evolution*, 34, 3299–3302. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum

likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28, 2731–2739.

Taylor, R. S., Bolton, M., Beard, A., Birt, T., Deane-Coe, P., Raine, A. F., González-Solís, J., Lougheed, S. C., & Friesen, V. L. (2019). Cryptic species and independent origins of allochronic populations within a seabird species complex (*Hydrobates* spp.). *Molecular Phylogenetics and Evolution*, 139, 106552.

Viricel, A., & Rosel, P. E. (2014). Hierarchical population structure and habitat differences in a highly mobile marine species: the Atlantic spotted dolphin. *Molecular Ecology*, 23, 5018–5035.

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

Wanless, R. M., Angel, A., Cuthbert, R. J., Hilton, G. M., & Ryan, P. G. (2007). Can predation by invasive mice drive seabird extinctions? *Biology Letters*, 3, 241–244.

Weimerskirch, H. (2004). Diseases threaten Southern Ocean albatrosses. *Polar Biology*, 27, 374–379.

Weimerskirch, H., Delord, K., Fretwell, P. T., Marteau, C., Ryan, P. G., & Barbraud, C. (2018) Status and trends of albatrosses in the French Southern Territories, western Indian Ocean. *Polar Biology*, 41, 1963–1972.

Wolfaardt, A. C., Glass, J., & Glass, T. (2009). Tristan da Cunha implementation plan for the Agreement on the Conservation of Albatrosses and Petrels (ACAP): review of current work and a prioritised work programme for the future. Tristan da Cunha Government, Tristan da Cunha.

Wright, S. (1943). Isolation by distance. *Genetics*, 28, 114–138.

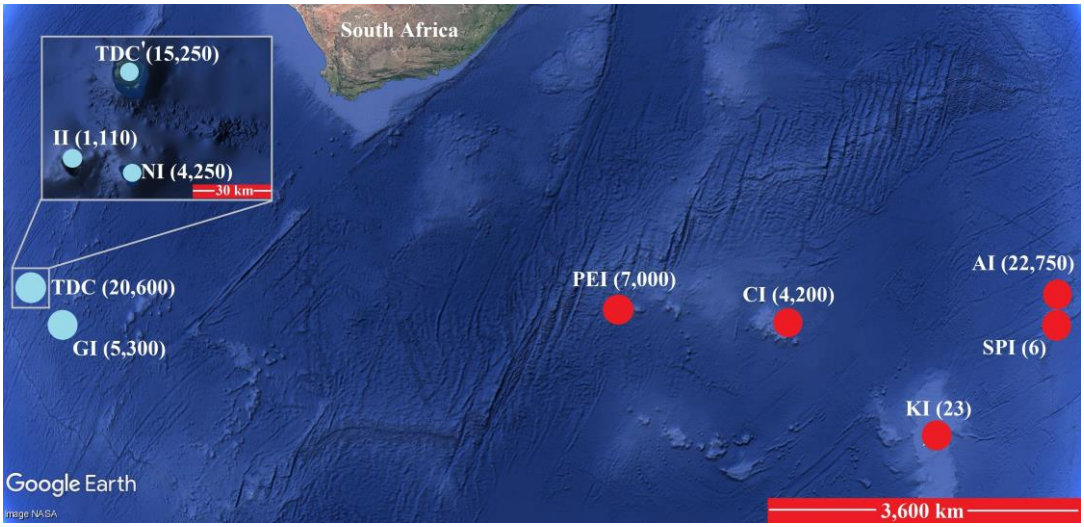

Figure 3.1: Breeding sites of Atlantic (*T. chlororhynchos*, blue) and Indian (*T. carteri*, red) yellow-nosed albatrosses with approximate numbers of annual breeding pairs on each island. Inset shows three breeding islands within the Tristan archipelago (TDC). Tristan da Cunha (TDC'), Nightingale (NI), Inaccessible (II), Gough (GI), Crozet (CI), Kerguelen (KI), Amsterdam (AI), St. Paul (SPI) and, Prince Edward (PEI). Map data: Google Earth, Maxar Technologies, National Centre for Space Studies (CNES), Scripps Institution of Oceanography (SIO), National Oceanic and Atmospheric Administration (NOAA), U.S. Navy, National Geospatial-Intelligence Agency (NGA), General Bathymetric Chart of the Oceans (GEBCO), Lamont-Doherty Earth Observatory (LDEO) Columbia and National Science Foundation (NSF).

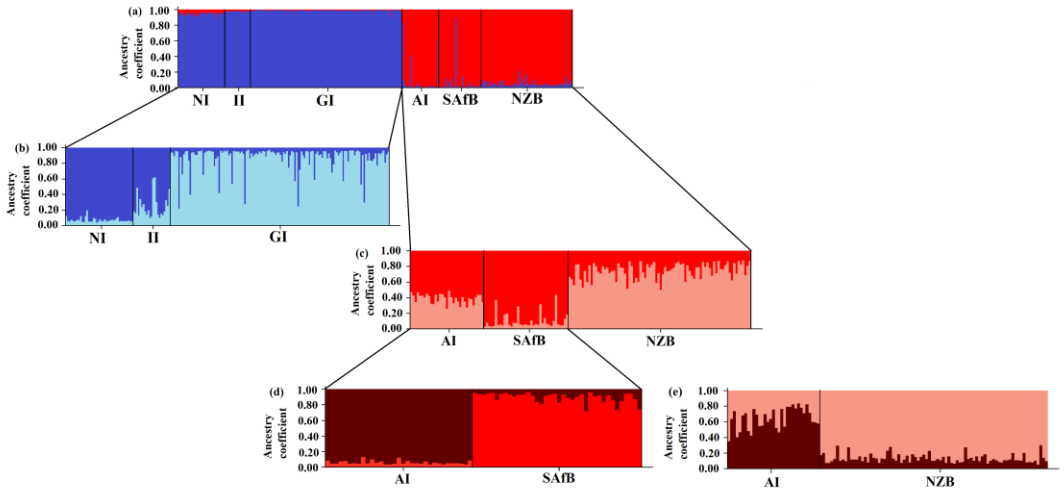

Figure 3.2: (a) STRUCTURE plot for six populations of yellow-nosed albatross at K=2. (NI-Nightingale, II-Inaccessible, GI-Gough, AI-Amsterdam, SAfB-South African bycatch, and NZB-New Zealand bycatch). (b) STRUCTURE plot for Atlantic yellow-nosed albatross shown at K=2. (c) STRUCTURE plot for Indian yellow-nosed albatross shown at K=2. (d) STRUCTURE plot for Amsterdam and South African bycatch at K=2. (e) STRUCTURE plot for Amsterdam and New Zealand bycatch at K=2.

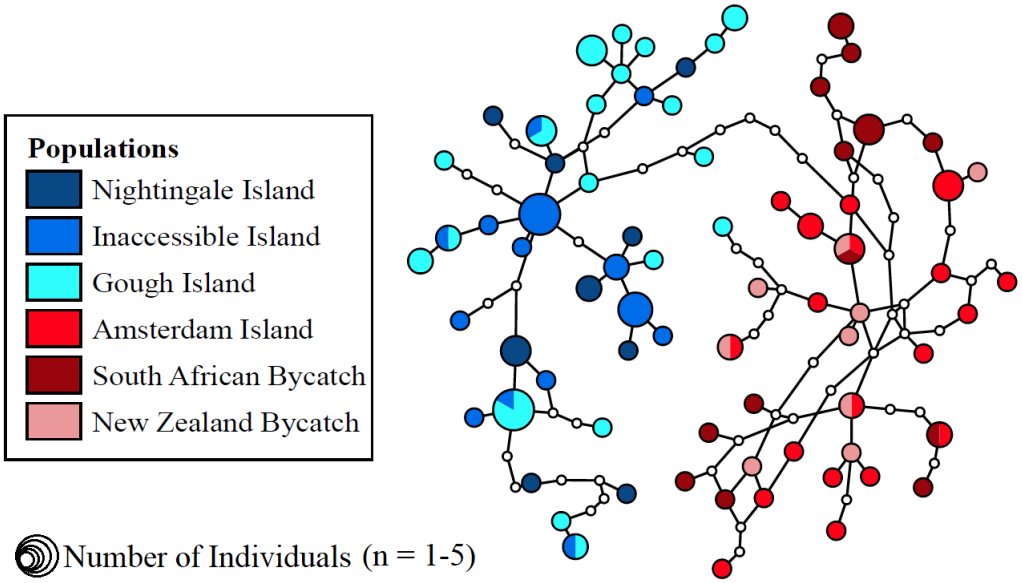

Figure 3.3: Haplotype network from mtCR sequences of 111 Atlantic and Indian yellow-nosed albatrosses. Coloured circles represent sampled haplotypes and the size of the cirle is proportional to the number of individuals sharing that haplotype. Small white circles are inferred haplotypes.

Table 3.1: Matrix of pairwise comparison of F_{ST} (below the diagonal) and P values (above the diagonal) between the four breeding populations and two bycatch groups of yellow-nosed albatross.

	Atlantic			**Indian**		
	NI	**II**	**GI**	**AI**	**SAfB**	**NZB**
NI		0.08	<0.001	<0.001	<0.001	<0.001
II	0.01		0.04	<0.001	<0.001	<0.001
GI	0.04	0.02		<0.001	<0.001	<0.001
AI	0.07	0.07	0.07		0.19	0.01
SAfB	0.05	0.06	0.06	0.01		0.08
NZB	0.07	0.08	0.05	0.02	0.01	

Table 3.2: Nucleotide diversity, number of haplotypes and haplotype diversity observed in the mtCR sequences for two species of yellow-nosed albatross from four breeding populations and two groups of bycatch.

	Nucleotide diversity	**Number of haplotypes**	**Haplotype diversity**
Atlantic	0.0171	39	0.970
NI	0.0161	9	0.909
II	0.0117	14	0.901
GI	0.0197	20	0.944
Indian	0.0173	36	0.978
AI	0.0167	19	0.961
SAfB	0.0174	13	0.935
NZB	0.0126	9	0.948

3.12 Supplementary materials

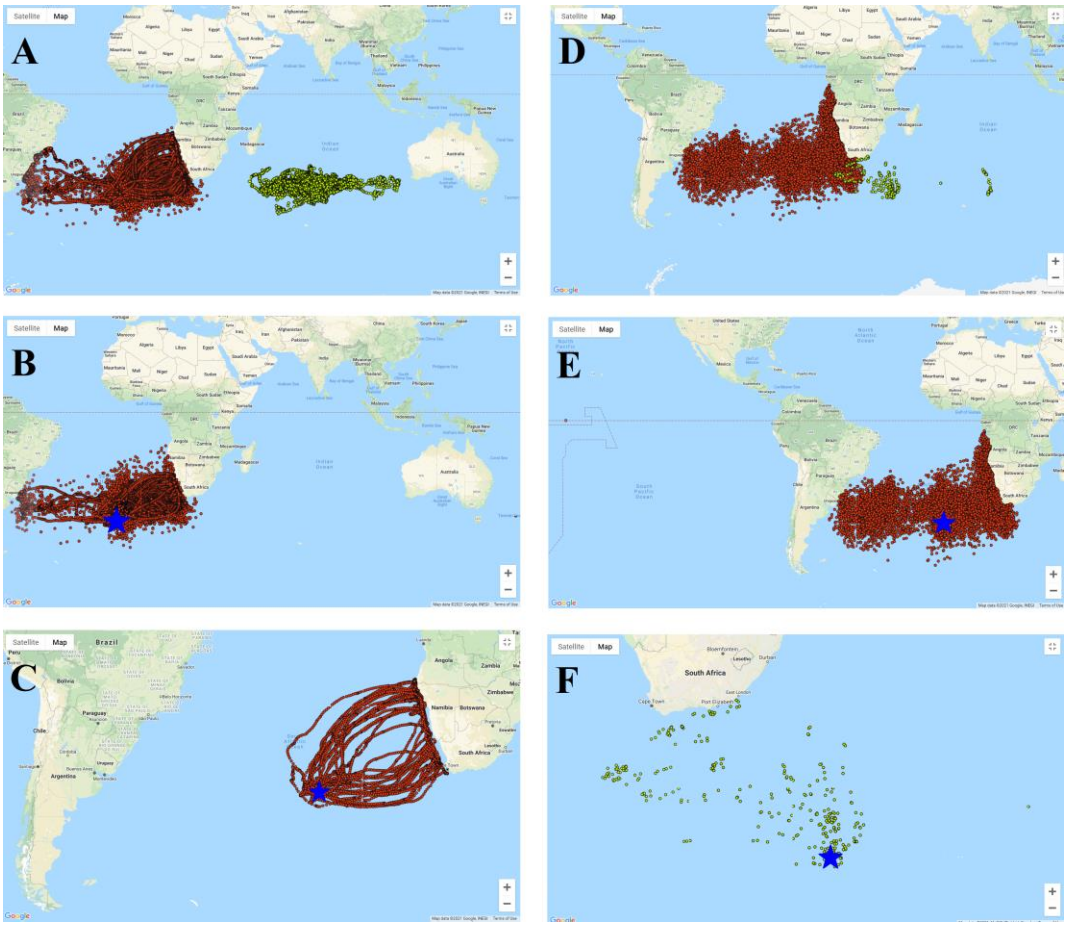

Figure S3.1: At-sea distribution of Atlantic (red) and Indian (green) yellow-nosed albatross. (A) Distribution of breeding adults, n=409, (B) Distribution of breeding Atlantic yellow-nosed albatross from Gough Island, n=148, (C) Distribution of breeding Atlantic yellow-nosed albatross from Nightingale Island, n=19, (D) Distribution of non-breeding adults of two species, n=50, (E) Distribution of non-breeding Atlantic yellow-nosed albatross adults from Gough Island, n=39, and (F) Distribution of non-breeding Indian yellow-nosed albatross adults from Prince Edward Island, n=10. Blue stars represent the approximate location of the breeding island for each population. Map data: BirdLife International Seabird Tracking Database.

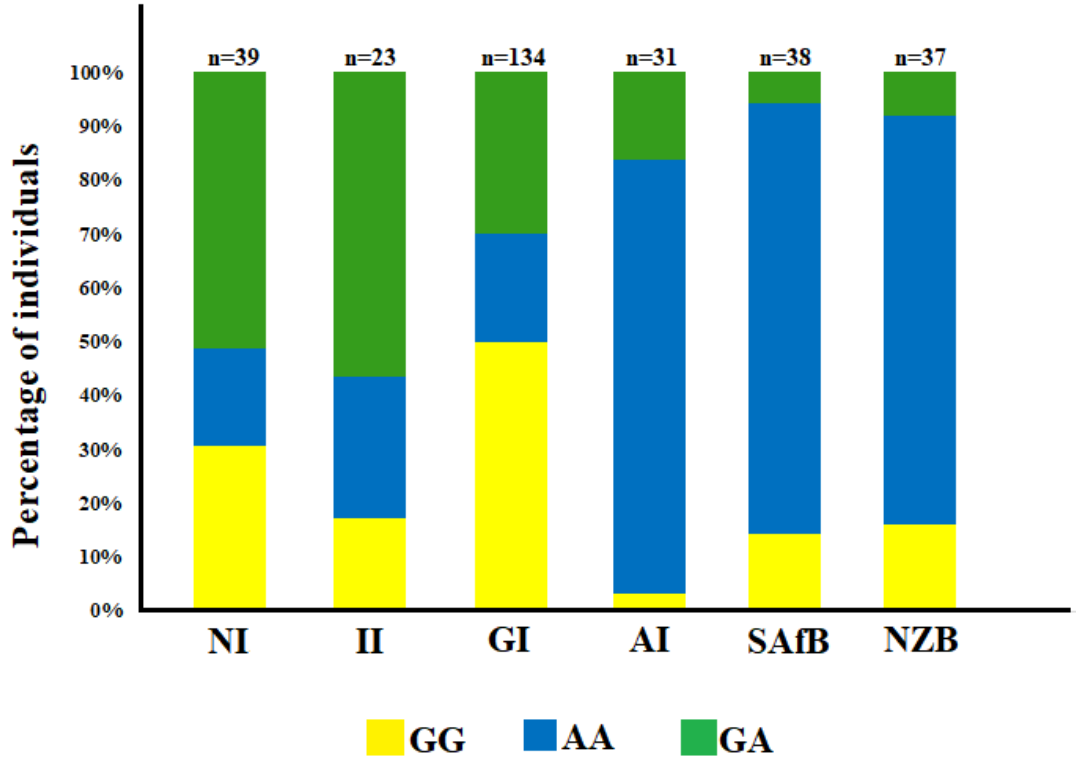

Figure S3.2: Percentage of individuals with genotypes; homozygous for G (yellow) or A (blue) alleles or heterozygous for G/A (green) for the SNP at locus Occa9 in six populations. NI-Nightingale, II-Inaccessible, GI-Gough, AI-Amsterdam, SAfB-South African bycatch, and NZB-New Zealand bycatch. Sample sizes (n) are included for each population.

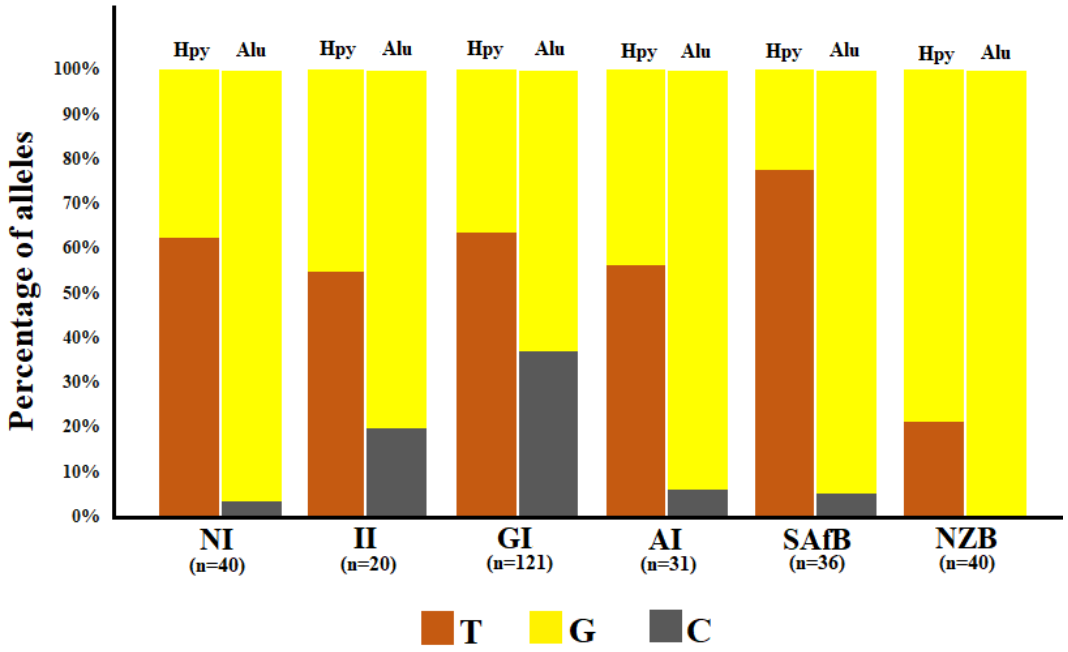

Figure S3.3: Percentage of different alleles found in each population for Hpy and Alu SNPs at locus Pema7. NI-Nightingale, II-Inaccessible, GI-Gough, AI-Amsterdam, SAfB-South African bycatch, and NZB-New Zealand bycatch. Sample sizes are included below the population abbreviation on the x-axis.

Table S3.1: Number of alleles (N), observed heterozygosity (Ho) and expected heterozygosity (He) at 13 microsatellite loci among four yellow-nosed albatross populations and bycatch samples from South Africa and New Zealand.

Species	Population		Cd11	11H7	Dc22	Dc9	Dc27	De18	Tc65	Oc51	De35	De11	12E1	Dc26	11H1	Average
Atlantic	NI	*N*	3	2	4	10	4	4	3	6	11	2	4	6	12	5.46
		Ho	0.33	0.34	0.47	0.76	0.26	0.20	0.27	0.53	0.65	0.26	0.63	0.44	0.78	0.46
		He	0.45	0.42	0.68	0.75	0.24	0.19	0.37	0.53	0.82	0.33	0.60	0.71	0.82	0.53
	II	*N*	2	2	5	10	3	3	3	6	9	2	4	5	9	4.85
		Ho	0.35	0.17	0.74	0.91	0.26	0.22	0.29	0.59	0.74	0.52	0.62	0.27	0.83	0.5
		He	0.39	0.16	0.75	0.80	0.24	0.20	0.26	0.66	0.87	0.39	0.64	0.67	0.87	0.53
	GI	*N*	4	2	6	11	4	4	3	6	11	6	7	9	13	6.62
		Ho	0.27	0.18	0.58	0.75	0.29	0.19	0.35	0.54	0.78	0.46	0.57	0.63	0.77	0.49
		He	0.33	0.19	0.67	0.83	0.37	0.20	0.38	0.57	0.83	0.49	0.62	0.78	0.86	0.55
		A	6	2	10	20	8	10	4	12	20	21	12	18	32	13.46
Indian	AI	*N*	3	2	4	6	1	3	4	4	10	3	6	7	7	4.61
		Ho	0.22	0.13	0.47	0.64	0.00	0.48	0.21	0.58	0.77	0.22	0.78	0.76	0.51	0.44
		He	0.26	0.12	0.69	0.64	0.00	0.51	0.22	0.58	0.86	0.37	0.80	0.80	0.80	0.51
	SAfB	*N*	3	2	4	8	2	6	4	6	11	3	6	6	9	5.38
		Ho	0.29	0.24	0.61	0.76	0.03	0.53	0.29	0.72	0.79	0.42	0.69	0.71	0.51	0.51
		He	0.30	0.29	0.47	0.77	0.03	0.62	0.26	0.66	0.82	0.50	0.76	0.75	0.78	0.54
	NZB	*N*	3	3	5	11	4	4	3	5	16	3	6	6	10	6.08
		Ho	0.13	0.12	0.47	0.60	0.10	0.45	0.20	0.50	0.78	0.27	0.70	0.67	0.50	0.42
		He	0.20	0.15	0.60	0.74	0.09	0.47	0.18	0.53	0.91	0.32	0.79	0.78	0.75	0.50

Chapter 4: Discovery of population genetic differentiation in two pelagic seabird species using ddRAD sequencing

Dilini K. Abeyrama[1], Kyle Wellband[2], Peter G. Ryan[3] and Theresa M. Burg[1]

1 University of Lethbridge, Lethbridge, Alberta, Canada

2 Fisheries and Oceans Canada, West Vancouver, British Columbia, Canada

3 FitzPatrick Institute of African Ornithology, University of Cape Town, Rondebosch 7701, South Africa

Corresponding author

Dilini K. Abeyrama: dilini.abeyrama@uleth.ca

ORCID ID:

DKA 0000-0002-9724-3708

KW

PGR 0000-0002-3356-2056

TMB 0000-0001-5096-3479

*This chapter has been prepared as a manuscript and submitted to Journal of Marine Biology. The manuscript has been reviewed thoroughly by all coauthors. Author contributions are listed under section 4.8.

4.1 Abstract

Population differentiation can result from physical or non-physical barriers such as isolation by distance or adaptation. Despite being highly mobile, population differentiation is not uncommon in seabirds; the driving force of this genetic differentiation in pelagic seabirds is not clear. We sequenced 48 Atlantic yellow-nosed albatrosses (*Thalassarche chlororhynchos*) from Gough, Inaccessible and Nightingale Islands and 48 Indian yellow-nosed albatrosses (*T. carteri*) breeding on Amsterdam Island or caught as bycatch around South Africa (likely from Prince Edward Island) and New Zealand (likely from Iles Crozet) using ddRAD sequencing. Three genetic clusters were identified within both species, separating all breeding populations and the two sets of bycatches from each other. Limited gene flow between breeding populations due to high philopatry is likely responsible for the genetic divergence of these highly vagile organisms.

Key words: philopatry, genetic divergence, gene flow, management units

4.2 Introduction

Speciation results from the independent evolution of lineages following reproductive isolation (Coyne & Orr, 2004). The early stages of divergence among lineages can result from physical barriers (e.g., mountain ranges, water bodies, deserts, and even roads) or non-physical barriers (e.g., natal philopatry, phenology, diet and song) that reduce gene flow among populations (Pilot et al., 2010; Gabirot et al., 2012; Roy et al., 2016). Interacting with reduced gene flow are the evolutionary forces of natural selection and genetic drift that can drive speciation (Kelly & Ward, 2018; Thorogood et al., 2018; Bergmann & Morinaga, 2019). Understanding the relative importance of these forces for speciation is both a central focus of evolutionary biology and a key component of decision-making for the conservation and management of endangered species.

Seabirds are a good model system to study non-physical barrier-mediated speciation as they face a limited number of physical barriers, yet they are a highly diversified group (Friesen et al., 2007). Physical isolation, philopatry, difference in ocean regime, non-breeding distributions, foraging distribution during the breeding season, and breeding phenology are all examples of barriers to gene flow in seabirds (Burg & Croxall, 2001; Ritchie et al., 2004; Friesen et al., 2007; Taylor et al., 2019). The initial isolation of some seabird populations is due to oceanographic changes during Pleistocene and genetic and phylogeographic structuring corresponding to historical separation can still be seen today (Friesen et al., 2007; Munro & Burg 2017).

Natal philopatry is one of the most common barriers in many seabird species, which results in reduced gene flow despite high vagility (Coulson, 2016). Seabirds are capable of dispersing thousands of kilometres, yet show genetic structure between islands that are

less than 200 km apart (Inchausti & Weimerskirch, 2002; Burg & Croxall, 2004; Danckwerts et al., 2021; Herman et al., 2022). Differences in breeding phenology can also lead to genetic divergence in seabird populations (Power et al., 1986; Taylor et al., 2019). For example, Taylor et al. (2019) found genetic differentiation between hot- and cold-season breeding populations of band-rumped storm petrels (*Hydrobates castro sensu lato*) breeding at different times of the year at a site in the Azores.

Seabirds are an important group because they are apex-predators as well as key indicators of the marine health (Croxall et al., 2012). Furthermore, 29% of seabird species are threatened with extinction (Spatz et al., 2014), so the identification of population structure in this group is particularly important for effective conservation and management.

Atlantic (*Thalassarche chlororhynchos*) and Indian (*T. carteri*) yellow-nosed albatrosses are two sister species that breed on islands in the south Atlantic and southwest Indian Oceans (Figure 4.1; ACAP, 2009a, b). The global population size of each species is ~40,000 breeding pairs with most individuals breeding on a single island (Tristan da Cunha for Atlantic yellow-nosed albatrosses and Amsterdam Island for Indian yellow-nosed albatrosses). Both are classified as endangered by the IUCN (BirdLife International, 2023a, b). Using traditional nuclear and mitochondrial markers, Abeyrama et al. (2021) detected population differentiation within these two species, but the underlying drivers of this differentiation are unclear. In this study, we use double digest restriction-site associated DNA (ddRAD) sequencing to assess the population genetic structure, identify outlier genomic regions, and compare and contrast patterns of population structure between the different marker sets. Our current study expands on our previous work and uses high resolution genomic data to detect the drivers of the

interspecific and intraspecific genetic diversity observed within these two endangered seabird species.

4.3 Methods

4.3.1 Sample collection and DNA extraction

Blood or tissue samples were collected from 48 Atlantic yellow-nosed albatrosses breeding on three of the four breeding islands (Gough, Inaccessible, and Nightingale) and 48 Indian yellow-nosed albatrosses breeding on Amsterdam Island or caught as bycatch around South Africa (likely birds breeding on Prince Edward Island; Abeyrama et al., 2021) and New Zealand (likely birds breeding on the Crozet Islands; Abeyrama et al., 2021) (Figure 4.1). These two populations could not be sampled at their breeding islands as there have been no land-based visits by people to Prince Edward Island since 2011 or Crozet colonies since the 1980s. For each sampling site (three populations per species), 16 individuals were sequenced. DNA was extracted using either phenol-chloroform (Sambrook & Russell, 2006) or modified salting-out (Aljanabi & Martinez, 1997) extraction. DNA was quantified using a Nanodrop spectrophotometer and normalized to a concentration of 20 ng/µl.

4.3.2 Sequencing and processing of data

Individually barcoded genotyping-by-sequencing libraries were prepared as described in Poland et al. (2012) using restriction enzymes *PstI* and *MspI* at the Institute of Integrative Biology and Systems (IBIS) at Université Laval, Canada. Samples were pooled and then sequenced at Genome Quebec, Canada, using Illumina TruSeq HT adaptors on an Illumina NovaSeq 6000 S4 PE100 with an expected sequencing effort of 25M reads per sample.

Forward and reverse sequencing reads were demultiplexed using the "sabre pe" option available for paired end reads in the program sabre 1.00 (https://github.com/najoshi/sabre). Quality of demultiplexed samples was checked using FastQC 0.11.9 (Andrews, 2010) and cutadapt 1.16 (Martin, 2011) was used to remove adaptors and trim all sequences to 80 bp. Sequences were then aligned to the Atlantic yellow-nosed albatross reference genome from the B10K project (GenBank assembly GCA_013400895.1, assembled to scaffold level) using Burrows-Wheeler alignment (bwa) tool 0.7.15 (Li & Durbin, 2009) and SAMtools 1.5 (Li et al., 2009; Cock et al., 2015). Gstacks and populations steps of Stacks 2.3e (Catchen et al., 2013) were used to identify single nucleotide polymorphic sites (SNPs) for all 96 individuals. To prevent the effect of linked SNPs, analyses were restricted to the first SNP at each locus. At least 50% of individuals in a population were required to be genotyped at a locus to include that locus in downstream analyses. For all other steps, default program parameters were used. Three datasets were used for downstream analyses. The first dataset compared the two species, with individuals categorised as either Atlantic or Indian yellow-nosed albatrosses. The second dataset compared only the three populations of Indian yellow-nosed albatrosses from Amsterdam Island and bycatch birds from New Zealand and South Africa. The last dataset compared Atlantic yellow-nosed albatrosses from the three sampled breeding populations: Gough, Inaccessible and Nightingale Islands.

4.3.3 Analysis of genome-wide genetic patterns

Individuals with more than 50% missing data were removed using VCFtools 0.1.16 (Danecek et al., 2011). Using the filtered variant caller format (vcf) files, Euclidean distance (Danielsson, 1980) between each pair of individuals was calculated using

Adegenet 2.1.5 (Jombart & Bateman, 2008) in RStudio. Pairwise Euclidean distance matrices were then used in GenAlEx 6.51b2 (Peakall & Smouse, 2012) for a principal coordinate analysis (PCoA). Interspecific as well as intraspecific genetic variation were determined for each of the three datasets. Observed (H_O) and expected (H_E) heterozygosities were calculated using the populations step in the Stacks pipeline.

The filtered vcf files were converted to arlequin (arp) file format using the software PGD Spider 2.1.1.5 (Lischer & Excoffier, 2012). The arlequin files were used to calculate population pairwise F_{ST} and corresponding P values in Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) and the significance of P values was determined using 110 permutations.

4.3.4 Outlier analysis

Pairwise F_{ST} values for each SNP were calculated by the populations step of Stacks pipeline (Catchen et al., 2013). Programs pcadapt (Luu et al., 2016) and OutFLANK (Whitlock & Lotterhos, 2015) were used to test for outliers between populations. These tests for outlier detection did not reveal strong outlier regions (Figure S4.1). Therefore, we defined SNPs with F_{ST} values ≥ 0.5 as outliers and a whitelist was generated including only these SNPs. The populations step of Stacks pipeline was rerun to create a vcf file with only the SNPs listed on the whitelist. This was done separately for each of the three datasets. PCoA were performed for all three datasets to detect the genetic clustering corresponding to the outlier SNPs between and within species.

4.3.5 Gene identification

Pairwise F_{ST} values of each SNP identified between each pair of populations was calculated by populations step of the Stacks pipeline. Manhattan plots were constructed for significance (p) of F_{ST} values in the form of $-\log_{10}(p)$ for each SNP in each dataset

using R package qqman (Turner, 2018) and genes found in regions with distinct peaks of the Manhattan plots were identified manually by looking at the general feature (gff) format file of the reference genome (GenBank assembly GCA_013400895.1) used in our study.

4.4 Results

4.4.1 Genome-wide genetic structure

Following removal of SNPs and individuals with more than 50% missing data, 132,900 SNPs and 70 individuals (Table S4.1) were retained in the dataset comparing the two species. Based on the variation explained by the first coordinate of the PCoA (5.14%), two distinct genetic clusters corresponding to the two species were identified (Figure 4.2). One Indian yellow-nosed albatross caught as bycatch around South Africa clustered with the Atlantic yellow-nosed albatross even though mitochondrial sequencing indicated that it is an Indian Ocean bird. This individual was excluded from further intraspecific analyses for Indian yellow-nosed albatrosses. Pairwise F_{ST} value for the two species was 0.068 and was significant ($P < 0.001$).

For the Indian yellow-nosed albatross dataset, we retained 111,609 SNPs and 35 individuals (Table S4.1) following the filtering of SNPs and individuals with more than 50% missing data. The three 'populations' of Indian yellow-nosed albatross were genetically different from each other without any overlap (Figure 4.3) and all pairwise F_{ST} values (0.004-0.017) were significant (Table 4.1). The pairwise F_{ST} value between Amsterdam Island and New Zealand bycatch was ~4x lower (0.004) than the two comparisons with South African bycatch (0.016-0.017). For all three populations, H_O was slightly higher than H_E but comparable when all sites (0.0003-0.0007 vs 0.0003-0.0007)

or only variable sites (0.0951-0.1817 vs 0.0899-0.1743 were considered (Table 4.2). Heterozygosities for New Zealand bycatch were approximately half that found in South African bycatch and Amsterdam Island samples.

In the Atlantic yellow-nosed albatross dataset, 76,993 SNPs and 38 individuals (Table S4.1) were retained for the downstream analyses. Three distinct genetic clusters corresponding to the three breeding sites were identified, although one individual from Inaccessible Island clustered with individuals from Gough Island (Figure 4.4). All pairwise F_{ST} values (0.013-0.029) between the three populations were significant (Table 4.1). The pairwise F_{ST} value for Nightingale and Gough Islands was ~2x higher than comparisons between other island pairs. H_O and H_E were either similar (when considering all SNPs) or H_O was slightly greater than H_E for two of the three populations (Table 4.2).

4.4.2 Outlier analysis and gene identification

Formal outlier tests using pcadapt and OutFLANK failed to identify any outlier SNPs. Outlier SNPs were chosen to be those with $F_{ST} \geq 0.5$. For all three datasets, population genetic structure identified when using only the outlier SNPs ($F_{ST} \geq 0.5$) was similar to the genetic structure identified with all SNPs.

No obvious peaks of divergence were observed in the Manhattan plots when comparing the two species or different populations of the two species using outlier loci (Figure S4.1). Therefore, no loci under selection were identified. Some regions of the genome containing loci with $-\log_{10}(p)$ above the cut-off level in Manhattan plots coded some random genes while others did not contain any protein coding sequences.

4.5 Discussion

The ddRAD sequencing data confirmed earlier findings based on mitochondrial and nuclear microsatellite data (Abeyrama et al., 2021) that Atlantic and Indian yellow-nosed albatrosses are genetically distinct. ddRAD sequencing also discovered fine-scale genetic differences among breeding populations within each species. This divergence among populations is probably driven by limited gene flow between breeding populations due to strong philopatric behaviour. The absence of strong outlier signals suggests that any selection driven differences are less important for explaining the genetic divergence among species and populations.

4.5.1 Fine-scale genetic structure

Bycatch samples used in this study from around New Zealand and South Africa are likely birds breeding on Crozet and Prince Edward Islands, respectively (see Abeyrama et al., 2021), suggesting that Indian yellow-nosed albatrosses breeding on Amsterdam, Crozet, and Prince Edward Islands are genetically distinct. The three sampled populations of Atlantic yellow-nosed albatross also showed significant intraspecific genetic differentiation. For both species, no loci were identified as being under selection following the gene identification analysis. Therefore, genetic divergence within each species is probably due to reduced gene flow between the islands, mutations or genetic drift. As there are no large land masses between these islands and albatrosses fly long distances, physical barriers do not appear to be involved in driving genetic divergence within species. Geographic distance seems to be irrelevant to genetic differentiation as islands 20-30 km as well as 2400 km apart are distinct. Barriers such as breeding phenology, non-breeding distribution, or philopatry could be hindering gene flow

between islands (Harris, 1973; Prince et al., 1994; Weimerskirch et al., 1997; Sagar et al., 1998; Steeves et al. 2005; Friesen et al., 2007; Calderón et al. 2014; Coulson, 2016; Munro & Burg, 2017; Taylor et al. 2018; Herman et al., 2022).

4.5.2 Breeding phenology

Based on available data, there is no evidence of differences in breeding phenology for Indian yellow-nosed albatrosses breeding on different islands (ACAP, 2009a). On average, Atlantic yellow-nosed albatrosses breed a week or so later on the more southerly Gough Island compared to birds breeding at Tristan (ACAP, 2009b), but this is unlikely to prevent potential gene flow because individual variation in breeding phenology within populations is much greater than the average difference between Tristan and Gough Islands (PGR pers. obs.).

4.5.3 Natal philopatry

Although there are no published data on philopatry of the two species of yellow-nosed albatross, researchers have observed adults of the two species returning to the same breeding area every year similar to other species of albatross (Harris, 1973; Prince et al., 1994; Weimerskirch et al., 1997). At Gough Island, adults are extremely site faithful, seldom nesting more than 100 m away from the previous year's nest, and virtually all chicks recruit within 1 km of their natal nest. Researchers have found only one chick farther than this, and it was only about 2 km from its original nest site (PGR unpubl. data). Similarly, there is no evidence of exchange of ringed birds between Tristan and Nightingale Islands, where large numbers of birds have been ringed regularly. Inaccessible Island is less well studied, but a large number of birds were ringed there in 1984 and none have been seen on any other island (PGR pers. obs.). The only evidence of

a possible inter-island movement was a ringed bird seen on Inaccessible in 2009 that probably came from Nightingale, but it was not breeding, and could not be caught to confirm its origin. Rare long-range dispersal is possible such as the case of the pair of Indian yellow-nosed albatross that had started to breed at the Chatham Islands, east of New Zealand (probably a system similar to wandering albatross described by Inchausti and Weimerskirch (2002)). Strong philopatry is very common among seabirds (Harris, 1973; Prince et al., 1994; Weimerskirch et al., 1997; Sagar et al., 1998; Friesen et al., 2006; Coulson, 2016; Herman et al., 2022). Genetic structure observed in waved albatross (*Diomedea irrorata*) endemic to Galapagos archipelago is believed to be driven by strong philopatry as there are no other barriers to restrict gene flow between populations breeding on different islands of that archipelago (Harris, 1973).

Regardless of the mechanism, reduced gene flow between populations of yellow-nosed albatross as well as other seabirds can have drastic effects. Population sizes of many Southern Ocean seabird colonies including yellow-nosed albatross are relatively small and only one or two colonies contain more than half of the global population of those species. When the gene flow between such small populations is highly restricted, their genetic diversity is reduced leading to inbreeding depression. As a result, those small populations might disappear. To prevent loss of small but critical populations, biodiversity managers are using different conservation units such as evolutionary significant units (ESUs) and management units (MUs).

When a population or a group of populations is reproductively isolated and genetically and ecologically distinct from the rest of their species, it can be identified as an evolutionary significant unit (Moritz 1994). And management units are demographically

independent populations which are not affected by migration but required for the persistence of evolutionary significant units (Allendorf et al., 2012). An evolutionary significant unit may contain several management units. As Atlantic and Indian yellow-nosed albatrosses were identified as different genetic clusters, and they are morphologically and behaviourally different, the two species should be considered as evolutionary significant units. Moreover, different populations of each species are genetically distinct with no or very rare migration between islands and therefore, all sampled populations in both species should be considered as separate management units.

4.6 Conclusions

All sampled populations of Atlantic and Indian yellow-nosed albatrosses are genetically distinct and strong natal philopatry seems to be the major driver of this intraspecific genetic diversity in both species. Based on our results, other non-physical barriers such as non-breeding distributions, foraging distribution during the breeding season, and breeding phenology do not appear to limit gene flow between breeding populations. Since both species of yellow-nosed albatross are considered endangered and each population is genetically different from one another, each of the breeding populations of yellow-nosed albatross should be considered as a separate conservation and management units.

4.7 Acknowledgements

We thank Henri Weimerskirch, Thierry Boulinier, Ben Dilley, Belinda Bauer, A Hewer and Tasmanian Museum and Art Gallery for providing samples. A big thank to Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship, Alberta Society of

Professional Biologists, the Seabird Group for funding this project. Special thanks to members of Burg lab, University of Lethbridge for their huge support.

4.8 Authors contributions

DKA, TMB and PGR designed the study. DKA completed the lab work. DKA and KW did the analyses. DKA, KW, PGR and TMB wrote the manuscript.

4.9 Funding

Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship, Alberta Society of Professional Biologists, and the Seabird Group.

4.10 Data availability

Sequences are in the process of being archived.

4.11 Declarations

4.12 Competing interest

The authors declare that they have no competing interest.

4.13 References

Abeyrama, D. K., Dempsey, Z. W., Ryan, P. G., & Burg, T. M. (2021). Cryptic speciation and population differentiation in the yellow-nosed albatross species complex. *Conservation Genetics*, *22*, 757–766. https://doi.org/10.1007/s10592-021-01358-x

Agreement on the Conservation of Albatrosses and Petrels. (2009a). ACAP Species assessments: Indian yellow-nosedaAlbatross *Thalassarche carteri.*

Agreement on the Conservation of Albatrosses and Petrels (2009b). ACAP Species assessment: Atlantic yellow-nosed albatross *Thalassarche chlororhynchos*.

Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, *25*, 4692–4693. https://doi.org/10.1093/NAR/25.22.4692

Allendorf, F. W., Luikart, G. H., & Aitken, S. N. (2012). Conservation and the genetics of populations. John Wiley & Sons.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bergmann, P. J., & Morinaga, G. (2019). The convergent evolution of snake-like forms by divergent evolutionary pathways in squamate reptiles. *Evolution*, *73*, 481–496. https://doi.org/10.1111/EVO.13651

BirdLife International (2023a). Species factsheet: *Thalassarche carteri*.

BirdLife International (2023b). Species factsheet: *Thalassarche chlororhynchos*.

Burg, T. M., & Croxall, J. P. (2004). Global population structure and taxonomy of the wandering albatross species complex. *Molecular Ecology*, *13*, 2345–2355. https://doi.org/10.1111/j.1365-294X.2004.02232.x

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

Calderón, L., Quintana, F., Cabanne, G. S., Lougheed, S. C., & Tubaro, P. L. (2014). Phylogeography and genetic structure of two Patagonian shag species (Aves: Phalacrocoracidae). *Molecular Phylogenetics and Evolution*, *72*, 42–53.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. https://doi.org/10.1111/mec.12354

Cock, P., Bonfield, J. K., Chevreux, B., & Li, H. (2015). SAM/BAM format v1.5 extensions for *de novo* assemblies. *BioRxiv*, 020024. https://doi.org/10.1101/020024

Coulson, J. C. (2016). A review of philopatry in seabirds and comparisons with other waterbird species. *Waterbirds*, *39*, 229–240. https://doi.org/10.1675/063.039.0302

Coyne, A. E., & Orr, H. A. (2004) Speciation. Sinauer Associates, Sunderland

Croxall, J. P., Butchart, S. H., Lascelles, B. E. N., Stattersfield, A. J., Sullivan, B. E. N., Symes, A., & Taylor, P. H. I. L. (2012). Seabird conservation status, threats and priority actions: a global assessment. B*ird Conservation International*, *22*, 1–34.

Danckwerts, D. K., Humeau, L., Pinet, P., McQuaid, C. D., & Le Corre, M. (2021). Extreme philopatry and genetic diversification at unprecedented scales in a seabird. *Scientific Reports*, *11*, 1–12. https://doi.org/10.1038/s41598-021-86406-9

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danielsson, P. E. (1980). Euclidean distance mapping. *Computer Graphics and Image Processing*, *14*, 227–248. https://doi.org/10.1016/0146-664X(80)90054-4

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. https://doi.org/10.1111/J.1755-0998.2010.02847.X

Friesen, V. L., Burg, T. M., & McCoy, K. D. (2007). Mechanisms of population differentiation in seabirds: Invited review. *Molecular Ecology*, *16*, 1765–1785. https://doi.org/10.1111/j.1365-294X.2006.03197.x

Gabirot, M., Lopez, P., & Martín, J. (2012). Differences in chemical sexual signals may promote reproductive isolation and cryptic speciation between Iberian wall lizard populations. *International Journal of Evolutionary Biology*, *2012*, 1–13. doi: 10.1155/2012/698520.

Harris, M. P. (1973). The biology of the waved albatross *Diomedea irrorata* of Hood Island, Galapagos. *Ibis*, *115*, 483–510. https://doi.org/10.1111/J.1474-919X.1973.TB01988.X

Herman, R. W., Winger, B. M., Dittmann, D. L., & Harvey, M. G. (2022). Fine-scale population genetic structure and barriers to gene flow in a widespread seabird (*Ardenna pacifica*). *Biological Journal of the Linnean Society*, *137*, 125–136. https://doi.org/10.1093/BIOLINNEAN/BLAC091

Inchausti, P., & Weimerskirch, H. (2002). Dispersal and metapopulation dynamics of an oceanic seabird, the wandering albatross, and its consequences for its response to long-line fisheries. *Journal of Animal Ecology*, *71*, 765–770. https://doi.org/10.1046/J.1365-2656.2002.00638.X

Jombart, T., & Bateman, A. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/BIOINFORMATICS/BTN129

Kelly, A. C., & Ward, T. J. (2018). Population genomics of Fusarium graminearum reveals signatures of divergent evolution within a major cereal pathogen. *PLoS One*, *13*, e0194616. https://doi.org/10.1371/JOURNAL.PONE.0194616

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R., (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics Applications Note*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298–299. https://doi.org/10.1093/BIOINFORMATICS/BTR642

Luu, K., Bazin, E., & Blum, M. G. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*(1), 67–77.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*, 10–12. https://doi.org/10.14806/EJ.17.1.200

Moritz, C. (1994). Defining 'evolutionarily significant units' for conservation. *Trends in Ecology & Evolution*, *9*(10), 373–375.

Munro, K. J., & Burg, T. M. (2017). A review of historical and contemporary processes affecting population genetic structure of Southern Ocean seabirds. *Emu-Austral Ornithology*, *117*, 4–18.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. *Bioinformatics*, *28*, 2537. https://doi.org/10.1093/BIOINFORMATICS/BTS460

Pilot, M., Dahlheim, M. E., & Hoelzel, A. R. (2010). Social cohesion among kin, gene flow without dispersal and the evolution of population genetic structure in the killer whale (*Orcinus orca*). *Journal of Evolutionary Biology*, *23*, 20–31. https://doi.org/10.1111/J.1420-9101.2009.01887.X

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One*, *7*, e32253. https://doi.org/10.1371/JOURNAL.PONE.0032253

Power, D. M., , & Ainley, D. G. (1986). Seabird geographic variation: Similarity among populations of Leach's storm-petrel. *The Auk*, *103*, 575–585. https://doi.org/10.1093/AUK/103.3.575

Prince, P. A., Rothery, P., Croxall, J. P., & Wood, A. G. (1994). Population dynamics of black-browed and grey-headed albatrosses *Diomedea melanophris* and *D. chrysostoma* at Bird Island, South Georgia. *Ibis*, *136*, 50–71. https://doi.org/10.1111/J.1474-919X.1994.TB08131.X

Ritchie, P. A., Millar, C. D., Gibb, G. C., Baroni, C., & Lambert, D. M. (2004). Ancient DNA enables timing of the Pleistocene origin and Holocene expansion of two Adélie

penguin lineages in Antarctica. *Molecular Biology and Evolution*, *21*, 240–248. https://doi.org/10.1093/MOLBEV/MSH012

Roy, L., Bon, M. C., Cesarini, C., Serin, J., & Bonato, O. (2016). Pinpointing the level of isolation between two cryptic species sharing the same microhabitat: a case study with a scarabaeid species complex. *Zoologica Scripta*, *45*, 407–420. https://doi.org/10.1111/ZSC.12166

Sagar, P. M., Stahl, J. C., Molloy, J. (1998) Sex determination and natal philopatry of southern Buller's mollymawks (*Diomedea bulleri bulleri*). *Notornis, 45*, 271–278.

Sambrook, J., & Russell, D. W. (2006). The condensed protocols: from molec-ular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY. *Cold Spring Harbor Protocols*, pdb.prot4455. https://doi.org/10.1101/pdb.prot4455.

Spatz, D. R., Newton, K. M., Heinz, R., Tershy, B., Holmes, N. D., Butchart, S. H., & Croll, D. A. (2014). The biogeography of globally threatened seabirds and island conservation opportunities. *Conservation Biology*, *28*, 1282–1290.

Steeves, T. E., Anderson, D. J., & Friesen, V. L. (2005). A role for nonphysical barriers to gene flow in the diversification of a highly vagile seabird, the masked booby (*Sula dactylatra*). *Molecular Ecology*, *14*, 3877–3887.

Taylor, R. S., Bailie, A., Gulavita, P., Birt, T., Aarvak, T., Anker-Nilssen, T., Barton, D.C., Lindquist, K., Bedolla-Guzmán, Y., Quillfeldt, P. and Friesen, V.L., (2018). Sympatric population divergence within a highly pelagic seabird species complex (*Hydrobates* spp.). *Journal of Avian Biology*, *49*, e01515.

Taylor, R. S., Bolton, M., Beard, A., Birt, T., Deane-Coe, P., Raine, A. F., González-Solís, J., Lougheed, S. C., & Friesen, V. L. (2019). Cryptic species and independent origins of allochronic populations within a seabird species complex (*Hydrobates* spp.). *Molecular Phylogenetics and Evolution*, *139*, 106552. https://doi.org/10.1016/J.YMPEV.2019.106552

Thorogood, C. J., Bauer, U., & Hiscock, S. J. (2018). Convergent and divergent evolution in carnivorous pitcher plant traps. *New Phytologist*, *217*, 1035–1041. https://doi.org/10.1111/NPH.14879

Turner, S. D. (2018). qqman: an R package for visualizing GWAS results using Q-Q and Manhattan plots. *Journal of Open Source Software*, *3*, 731. https://doi.org/10.21105/joss.00731

Weimerskirch, H., Brothers, N., & Jouventin, P. (1997). Population dynamics of wandering albatross *Diomedea exulans* and Amsterdam albatross *D. amsterdamensis* in the Indian Ocean and their relationships with long-line fisheries: Conservation implications. *Biological Conservation*, *79*, 257–270. https://doi.org/10.1016/S0006-3207(96)00084-5

Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F_{ST}.

The American Naturalist, *186*, S24–S36.

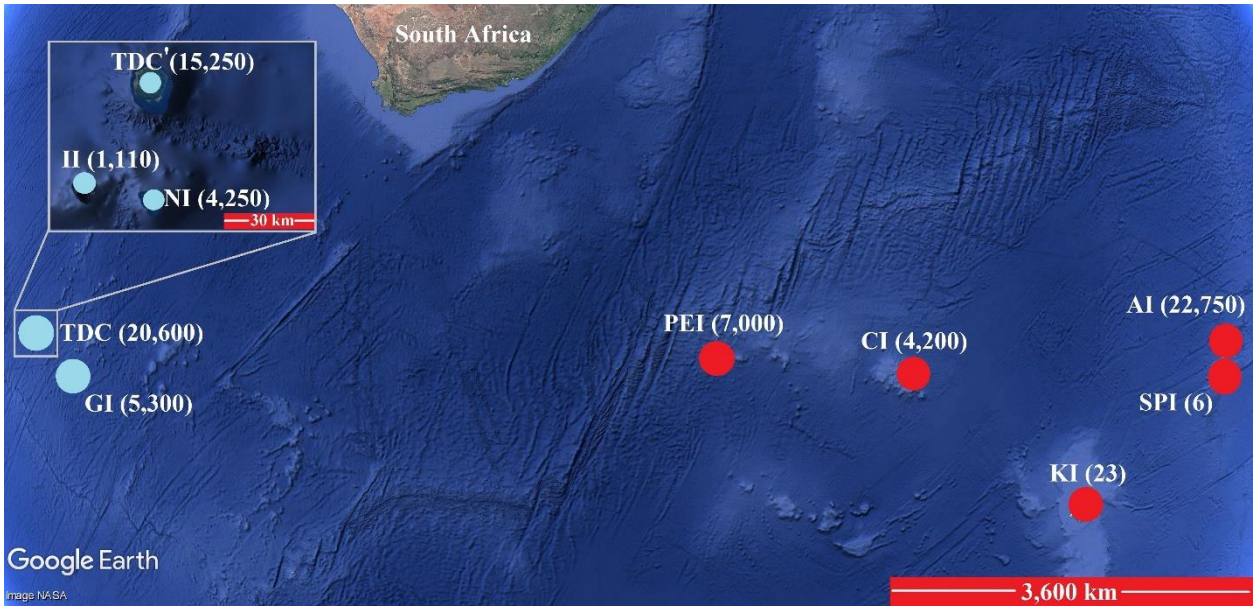

Figure 4.1: Breeding sites of Atlantic (*T. chlororhynchos*, blue) and Indian (*T. carteri*, red) yellow-nosed albatrosses with approximate numbers of annual breeding pairs on each island. Inset shows three breeding islands within the Tristan archipelago (TDC). Tristan da Cunha (TDC'), Nightingale (NI), Inaccessible (II), Gough (GI), Crozet (CI), Kerguelen (KI), Amsterdam (AI), St. Paul (SPI) and Prince Edward (PEI) Islands. Map data: Google Earth, Maxar Technologies, National Centre for Space Studies (CNES), Scripps Institution of Oceanography (SIO), National Oceanic and Atmospheric Administration (NOAA), U.S. Navy, National Geospatial-Intelligence Agency (NGA), General Bathymetric Chart of the Oceans (GEBCO), Lamont-Doherty Earth Observatory (LDEO) Columbia University and National Science Foundation (NSF).

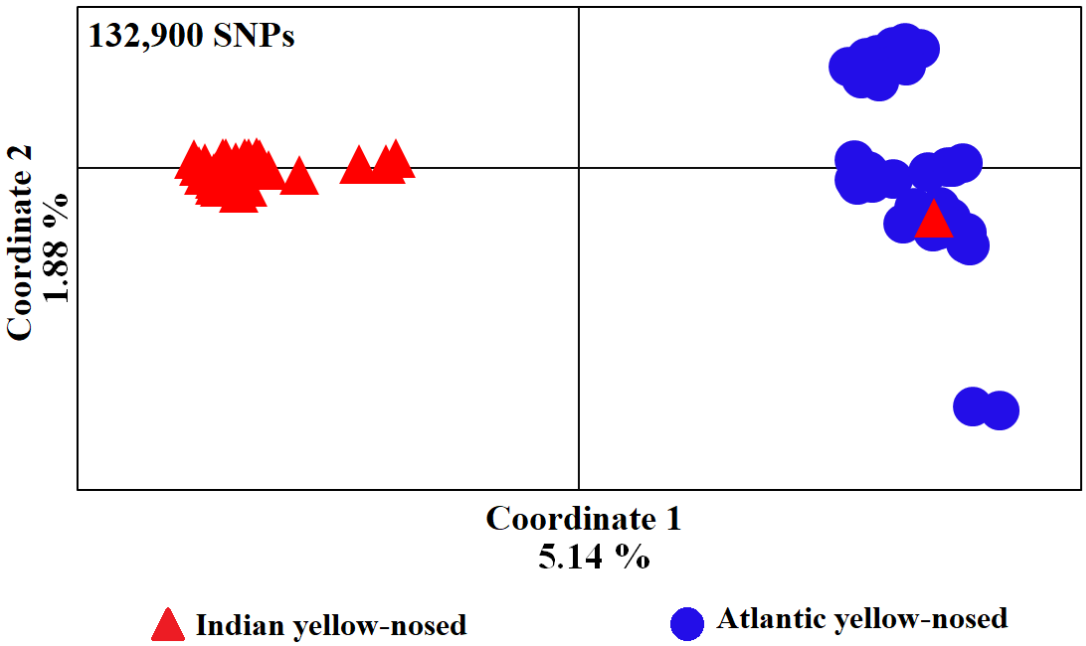

Figure 4.2: PCoA plot representing the genetic variation between Atlantic and Indian yellow-nosed albatrosses.

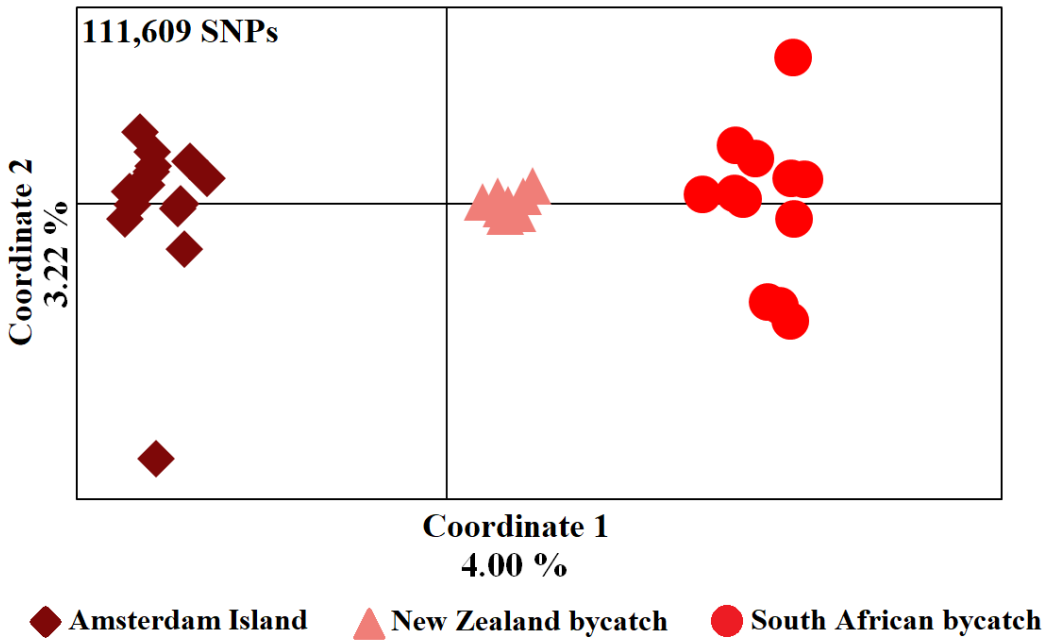

Figure 4.3: PCoA plot representing the genetic variation between the three Indian yellow-nosed albatross populations.

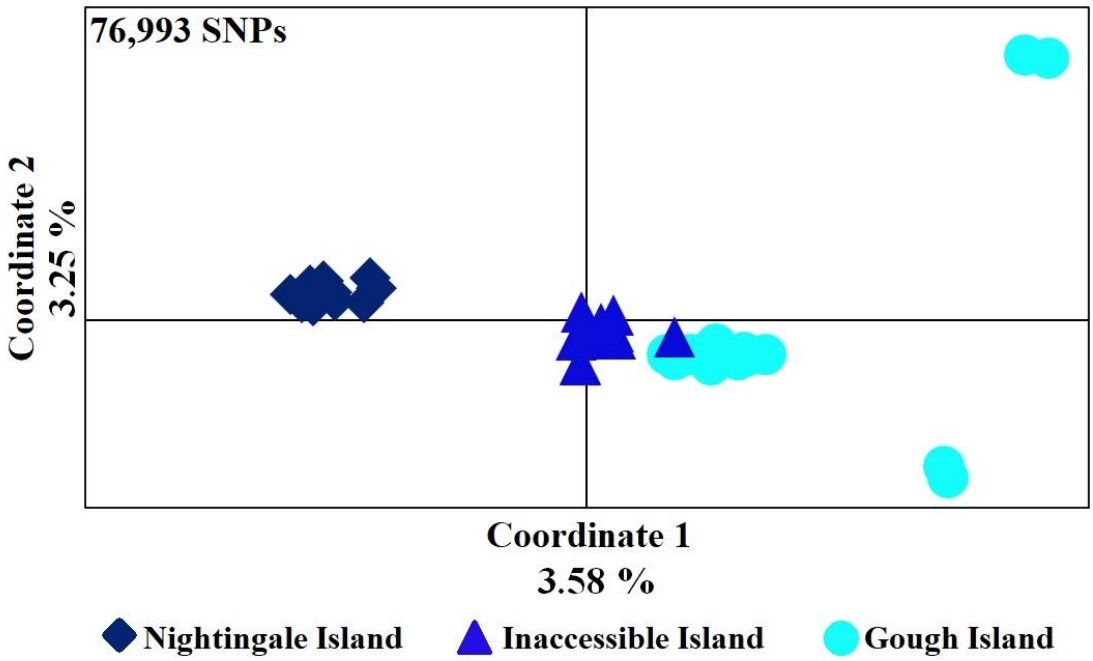

Figure 4.4: PCoA representing the genetic variation between three Atlantic yellow-nosed albatross populations.

Table 4.1: Pairwise F_{ST} (above the diagonal) and associated P values (below the diagonal) of calculated between populations within each species (Asterisk denotes significance after the Benjamini-Hochberg correction).

	AI	**NZB**	**SAfB**
AI	-----	0.004	0.016
NZB	<0.001*	-----	0.017
SAfB	<0.001*	<0.001*	-----
	GI	**II**	**NI**
GI	-----	0.013	0.029
II	<0.001*	-----	0.016
NI	<0.001*	<0.001*	-----

Table 4.2: Expected (H_E) and observed (H_O) heterozygosity calculated for six populations of yellow-nosed albatrosses.

		All sites			**Variable sites only**		
Population	**Number of individuals**	**Number of positions**	**H_E**	**H_O**	**Number of SNPs**	**H_E**	**H_O**
Indian yellow-nosed							
AI	12	39,416,982	0.0006	0.0006	151,798	0.1586	0.1640
NZB	8	22,606,690	0.0003	0.0003	77,465	0.0899	0.0951
SAfB	15	45,116,294	0.0007	0.0007	169,199	0.1743	0.1817
Atlantic yellow-nosed							
GI	14	35,840,505	0.0006	0.0006	136,195	0.1544	0.1553
II	12	21,039,441	0.0005	0.0005	77,323	0.1237	0.1252
NI	12	34,333,667	0.0005	0.0005	128,851	0.1442	0.1444

4.13 Supplementary materials

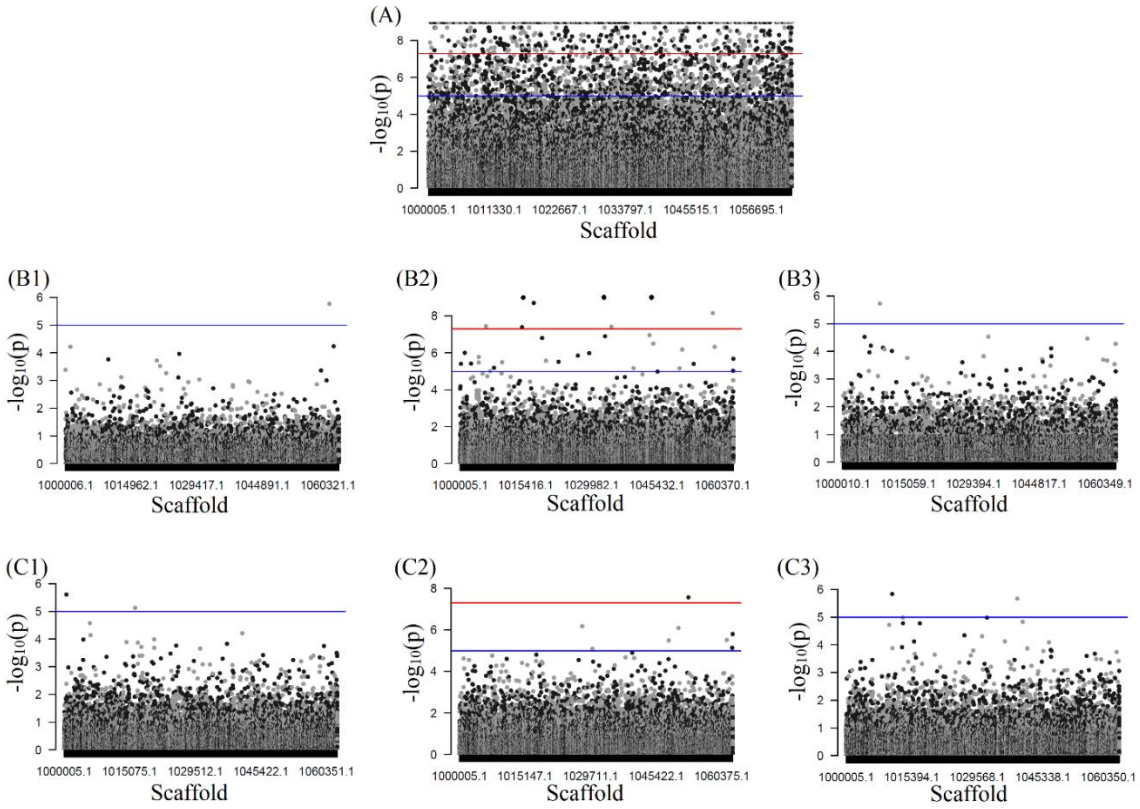

Figure S4.1: Manhattan plots (-log10(p-values)) for the datasets comparing (A) Atlantic and Indian yellow-nosed albatrosses, Indian yellow-nosed albatrosses from (B1) Amsterdam Island and New Zealand bycatch, (B2) Amsterdam Island and South African bycatch, (B3) New Zealand bycatch and South African bycatch, Atlantic yellow-nosed albatrosses from (C1) Gough and Inaccessible Islands, (C2) Gough and Nightingale Islands and (C3) Inaccessible and Nightingale Islands. Red and blue lines indicate cut-off levels of significance picked automatically by the program.

Table S4.1: Number of individuals from each population used in downstream analyses of each dataset.

Population	Initial number of individuals	Number of individuals retained after filtering in Atlantic–Indian comparison	Number of individuals retained after filtering for Indian yellow-nosed	Number of individuals retained after filtering for Atlantic yellow-nosed
Gough Island	16	13	NA	14
Inaccessible Island	16	9	NA	12
Nightingale Island	16	12	NA	12
Amsterdam Island	16	12	12	NA
New Zealand bycatch	16	8	8	NA
South African bycatch	16	16	15*	NA

* Excluding the individual that clustered with Atlantic yellow-nosed albatrosses

Chapter 5: Microgeographic genetic diversity in seabirds breeding on Kerguelen Island

Dilini K. Abeyrama[1], Charles-André Bost[2], Theresa M. Burg[1]

1 University of Lethbridge, Lethbridge, Alberta, Canada

2 CNRS-Chizé, Villiers-en-Bois, France

Corresponding author

Dilini K. Abeyrama: dilini.abeyrama@uleth.ca

ORCID ID:

DKA 0000-0002-9724-3708

CAB

TMB 0000-0001-5096-3479

*This chapter has been prepared as a manuscript for submission. The manuscript has been reviewed thoroughly by coauthors. Author contributions are listed under section 5.8.

5.1 Abstract

Population genetic structure is more common at large spatial scales than at microgeographic levels. In seabirds, gene flow is restricted mostly by non-physical barriers though in rare instances population genetic structure is present at smaller geographic ranges (<100 km). Two species, Kerguelen shag (*Leucocarbo verrucosus*) and gentoo penguin (*Pygoscelis papua*), breeding on Kerguelen Island each has two ecotypes consisting of individuals foraging in inshore shallow waters (closed-sea) or offshore deep waters (open-sea). Both species show stable morphological and behavioural variation between the two ecotypes with smaller individuals foraging inshore and larger ones offshore; colonies contain a single ecotype. To determine if the two ecotypes are genetically different, we genotyped 32 shags from one closed-sea and two open-sea colonies and 84 gentoo penguins from two closed-sea and four open-sea colonies on Kerguelen Island. In both species the two ecotypes are genetically different. The Kerguelen shags showed further differentiation among the three sampled breeding colonies. Non-physical barriers have reproductively isolated breeding colonies of Kerguelen shags and gentoo penguins on Kerguelen Island and a reassessment of the taxonomy of the two species is required.

Key words: ecotypes, Southern Ocean birds, behavioural isolation

5.2 Introduction

Restricted or reduced gene flow due to physical and non-physical barriers facilitates intraspecific genetic differentiation in both plants and animals (Friesen et al., 2007; Zhang et al., 2007; Widmer et al., 2008; Hollatz et al., 2011; Robertson & Vega, 2011; Munro & Burg, 2017; Pascual et al., 2017). This genetic differentiation is more common at macrogeographic levels than at the microgeographic level (Ruiz-Garcia & Klein, 1997; Bezault et al., 2011; Pujolar et al., 2011). Microgeographic genetic differentiation is seen mostly in sedentary species such as plants as their immobility restricts gene flow unless there is dispersal by seeds or pollinating agents. In highly mobile organisms such as most mammals, birds and fish, microgeographic genetic variation is not uncommon. It has gained limited attention as most species are capable of overcoming barriers and seem to be exchanging genes with other populations. Genetic differentiation at the microgeographic level is found in a range of species including mammals and birds and is often associated with genetic drift, changes in habitat or divergent selection (Schulte-Hostedde et al., 2001; Verdade et al., 2002; Adams et al., 2016; Mikles et al., 2020; Cheek et al., 2022; Estivals et al., 2022).

Kerguelen is an archipelago located in the southern Indian Ocean and is composed of a main island, Kerguelen Island, and about 300 small satellite islands covering an area of ~7,000 km^2 (Blain et al., 2001). Although the floral diversity on this archipelago is low due to harsh environmental conditions, volcanic eruptions and physical isolation, the faunal diversity is relatively high accommodating species of penguins, shags, whales, seals, and albatrosses. The jagged coastline of the archipelago creates different oceanographic conditions namely: open-sea and closed-sea colonies for animals breeding

on the islands (Lescroël & Bost, 2005). For some species, individuals breeding in open-sea and closed-sea colonies have behavioural and morphological differences. Both Kerguelen shags (*Leucocarbo verrucosus*) and gentoo penguins (*Pygoscelis papua*) breeding on Kerguelen Archipelago show similar and stable morphological and behavioural divergence corresponding to the two ecotypes: open-sea and closed-sea colonies (Figure 5.1). Shags and gentoo penguins from closed-sea colonies are ~60% and ~36% lighter in weight, respectively than individuals of the same species from open-sea colonies in addition to other morphological and behavioural differences such as diving depth and diet (Bost et al., 1992; Cook et al., 2013).

The Kerguelen shag is endemic to Kerguelen Island while gentoo penguin breeds on islands across Southern Ocean including the Kerguelen Island. Within the Southern Ocean, population structure has been detected in gentoo penguins breeding on different islands although gene flow was observed between islands which are ~900 km apart (Vianna et al., 2017); however, ocean currents appear to restrict gene flow among some islands. In this study we used restriction-site associated DNA sequence (RADSeq) data to determine if the two ecotypes of shags and gentoo penguins breeding on Kerguelen Island show microgeographic population genetic structure and if so, does it correspond to differences in morphology and behaviour?

5.3 Methods

5.3.1 Sample collection and DNA extraction

Blood or tissue samples were collected from 32 Kerguelen shags breeding in one closed-sea: Île Mayes (n=11) and two open-sea: Point Suzanne (n=10) and Sourcils Noirs (n=11) colonies and 86 gentoo penguins breeding in four open-sea: Ratmanoff (n=14), Point

Morne (n=19), Point Suzanne (n=20), and Cataractes (n=2; this population was used as outgroup and therefore only available two samples were used) and two closed-sea: Point Guite (n=16) and Île Longue (n=15) colonies on Kerguelen Island (Figure 5.1). DNA was extracted using modified salting-out extraction protocol (Aljanabi & Martinez, 1997), quantified using a Nanodrop spectrophotometer, and normalized to a concentration of 20 ng/µl.

5.3.2 Sequencing and processing of data

Individually barcoded genotyping-by-sequencing libraries were prepared as described in Poland et al. (2012) using three restriction enzymes *PstI*, *NsiI* and *MspI* at the Institute of Integrative Biology and Systems (IBIS) at Université Laval, Canada. Samples were pooled and sequenced at Genome Quebec, Canada, using Illumina TruSeq HT adaptors on an Illumina NovaSeq 6000 S4 PE100.

In the program sabre 1.00, the "sabre pe" option available for paired end reads was used to demultiplex forward and reverse sequencing reads. Quality of demultiplexed samples was checked using FastQC 0.11.9 (Andrews, 2017). Adaptors were removed and all sequences were trimmed to 80 bp using program cutadapt 1.16 (Martin, 2011). Only individuals with more than 10,000 reads were included in downstream analyses. Sequences were analysed *de novo* without aligning to a reference genome in Stacks 2.3e (Catchen et al., 2013; Rochette et al., 2019). To reduce the effect of linkage, analyses were restricted to the first SNP at each locus. For all other steps, default parameters in the programs were used. The Stacks pipeline was run for the two species separately focusing on two ecotypes: open-sea and closed-sea. A subset of samples from each species was also analysed aligning to reference genomes. For gentoo penguin, a reference genome

from the same species (Assembly GCA_003264615.1 in GenBank) assembled to scaffold level was used. For Kerguelen shag, a reference genome from pelagic cormorant (*Phalacrocorax pelagicus*, isolate: Assembly GCA_002173435.1 in GenBank) with scaffold level assembly was used. Gstacks and populations steps of program Stacks 2.3e (Catchen et al., 2013) were used to identify single nucleotide polymorphic sites (SNPs) for all individuals in two species separately.

5.3.3 Analysis of genome-wide genetic patterns

Individuals and SNPs with more than 50% of data were retained for downstream analyses following filtering using VCFtools 0.1.16 (Danecek et al., 2011). Filtered variant call format (vcf) files were used to calculate Euclidean distances (Danielsson, 1980) between each pair of individuals in Adegenet 2.1.5 (Jombart & Bateman, 2008) of RStudio. Principal coordinate analyses (PCoA) were done using pairwise Euclidean distance matrices in GenAlEx 6.51b2 (Peakall & Smouse, 2012). Intraspecific genetic variation such as observed (H_O) and expected (H_E) heterozygosities were calculated using the populations step in Stacks pipeline.

Using PGD Spider 2.1.1.5 software (Lischer & Excoffier, 2012), filtered vcf files were converted to arlequin (arp) file format. Population pairwise F_{ST} and corresponding P values were calculated using arlequin files in Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) and 110 permutations to determine significance.

5.4 Results

5.4.1 Gentoo penguin

Both *de novo* and reference-based analyses for RADSeq data identified two genetic clusters corresponding to two ecotypes in gentoo penguins breeding on Kerguelen Island. As reference-based analysis showed a separation based on the sequencing batch, sequences from only one batch were aligned to the reference genomes to look at the overall genetic pattern. Output files from *de novo* analysis were used in all downstream analyses.

Out of 86 sequenced individuals, 81 individuals had more than 10,000 sequences. Following filtering to remove individuals and SNPs with more than 50% missing data, 44 gentoo penguins (Table S5.1) and 49,263 SNPs were retained for the downstream analyses. The variation explained by the first axis of the PCoA (3.95%) identified two distinct genetic clusters corresponding to the two ecotypes: open-sea and closed-sea (Figure 5.2). The second axis explained 2.64% of variation but showed no further division. No genetic structure was observed between different breeding colonies within each ecotype. One open-sea individual from Ratmanoff (GEPE KRFo021) showed intermediate grouping between the open-sea and closed-sea populations (Figure 5.2). Pairwise F_{ST} value for open-sea and closed-sea colonies of gentoo penguins was 0.028 and was significant ($P<0.001$). For both open-sea and closed-sea expected heterozygosities were greater than the observed heterozygosities (0.192 vs 0.168 and 0.189 vs 0.177, respectively) (Table 5.1).

5.4.2 Kerguelen shag

In Kerguelen shags, three genetic clusters representing the three sampled breeding colonies were identified with both *de novo* and reference-based analyses. For Kerguelen shags, only output files from the *de novo* analysis were used for downstream analyses due to sequencing batch effect when aligning to the reference genome.

For Kerguelen shags, 28 out of 32 individuals had more than 10,000 sequences. Following filtering of individuals and SNPs with more than 50% missing data, 20 individuals (Table S5.1) and 66,503 SNPs were retained for downstream analyses. The variation explained by the first coordinate of the PCoA (6.78%) identified three genetic clusters corresponding to the three sampled breeding colonies. The variation within each genetic group (explained by the second axis of the PCoA: 5.61%) was comparatively high (Figure 5.3), but showed no further separation among breeding colonies. Pairwise F_{ST} value between the two ecotypes of Kerguelen shag was 0.027 and was significant ($P<0.001$) and all three colonies were significantly different (F_{ST} 0.015-0.056; Table 5.2). Interestingly, the F_{ST} value between two open-sea colonies (0.026) was greater than the F_{ST} value between the open-sea colony Point Suzanne and the closed-sea colony Île Mayes (0.015). Heterozygosity for both ecotypes as well as three separate breeding colonies were comparable (0.168-0.233; Table 5.1 and 0.186-0.230; Table 5.3).

5.5 Discussion

RADSeq data show that gentoo penguins breeding in open-sea colonies are genetically different from gentoo penguins breeding in closed-sea colonies on the Kerguelen Island. Similarly, the two Kerguelen shag ecotypes are also genetically distinct. However, further population structure is present at each of the three Kerguelen shag breeding colonies on

the island and they are genetically different from each other regardless of their ecotype. For both study species, genetic results match behavioural and morphological differences.

5.5.1 Population genetic structure in gentoo penguins

In gentoo penguins breeding across the globe, six distinct genetic clusters: Antarctic, Falklands, Crozet, Macquarie, South Georgia and Kerguelen have been identified (de Dinechin et al., 2012; Vianna et al., 2017; Pertierra et al., 2020; Tyler et al., 2020; Noll et al., 2022). Factors such as ocean currents, differences in marine habitats and behaviour (acoustics during courtship) have been identified as genetic barriers for penguins at macrogeographic scales (Jouventin, 1982; De Dinechin et al., 2012; Vianna et al., 2017; Pertierra et al., 2020; Tyler et al., 2020; Noll et al., 2022). But until now, no genetic barrier had ever been identified at a microgeographic scale for a penguin species. Furthermore, on Kerguelen Island, open-sea gentoo penguin colonies are comprised only of larger offshore deep water divers while closed-sea colonies contain only smaller inshore shallow water divers (Bost et al., 1992, Cook et al., 2013). Genetic structure identified in gentoo penguins breeding on Kerguelen Island corresponds to these morphological and behavioural differences.

5.5.2 Population genetic structure in Kerguelen shags

Although there are no previous studies on genetic structure of Kerguelen shags, a study on closely related European shags (*Phalacrocorax aristotelis*) using mitochondrial and microsatelite markers discovered genetic structure between distant populations (>500 km; Thanou et al., 2017). Another study on a more insular species of the same genus, the flightless cormorant (*P. harrisi*), also found population genetic at macrogeographic level between neighbouring Galapagose Islands, but not within each island regardless of their

sedentary status compared to other species of the genus (Duffie et al., 2009). Such genetic divergence is driven by differences in landscape, philopatry and local adaptation (Duffie et al., 2009; Thanou et al., 2017). Interestingly, none of these factors seem to be acting as genetic barriers at microgeographic levels in this genus as evidence of gene flow between adjacent populations or colonies was detected. In a previous study, it was found that Kerguelen shags from the closed-sea colony at Île Mayes are the smallest and dive in shallow waters while shags from Point Suzanne, an open-sea colony, are medium sized birds foraging in medium depth water. More, largest birds from the open-sea colony Sourcils Noirs forage in deepest waters of the three colonies (Cook et al., 2013). Among seabirds, larger animals can dive longer and hence deeper because oxygen storage scales isometrically with body mass (Halsey et al., 2006). It is possible that variation in diving distance and depth between individuals from Île Mayes and Point Suzanne might be less than that between Point Suzanne and Sourcils Noir and this resulting in a variation in the level of environmental pressure experienced by individuals from each colony. This can also explain the hight F_{ST} between two open-sea colonies than between Île Mayes and Sourcils Noir. Furthermore, the three genetic clusters identified in our study correspond to these morphological and behavioural differences observed in Kerguelen shags.

5.5.3 Genetic barriers on Kerguelen Island

Colonies of both shags and gentoo penguins breeding on Kerguelen Island are close enough that individuals from one colony can access other colonies, as such there are no obvious physical barriers restricting movement between these breeding colonies. Indeed, the distances that both shags and gentoo penguins travel for foraging and provisioning their chicks are comparable or longer than the distances between the different colonies in

the archipelago (Lescroël and Bost, 2005; Cook et al., 2013). Although differences in seascape and ocean currents are not acting as genetic barriers for shags or gentoo penguins breeding on Kerguelen Island, local adaptation and non-physical barriers such as natal philopatry and at-sea foraging distribution could be limiting gene flow within the archipelago.

At-sea foraging distribution corresponds to the genetic differentiation observed in both Kerguelen shags and gentoo penguins breeding on the Kerguelen Archipelago. Not only at microgeographic level but at-sea foraging distribution also acts as a genetic barrier at macrogeographic levels in other seabird species such as albatrosses (Burg & Croxall, 2001; Abbott & Double, 2003b; Makhado et al., 2018). The difference in foraging distribution in both Kerguelen shags and gentoo penguins on Kerguelen Island seems to be a solution to reduce competition for resources within each species, similar to Galapagos finches though at different time scales (Huber et al., 2007).

The shallow bays and continental shelves around Kerguelen Island have created different niches that species can exploit, therefore minimizing competition for resources between and within species (Fromant et al., 2022). On Kerguelen Island, both shags and gentoo penguins from the two ecotypes are occupying distinct niches: open-sea and closed-sea, so minimizing the potential competition for food. Differences in available food sources and adaptation to different environmental pressures in two niche types seem to be driving the morphological divergence between ecotypes or breeding colonies of inshore seabirds (Bost & Jouventin, 1990). A study looking at another insular seabird species, flightless cormorant (*Phalacrocorax harrisi*) endemic to Galápagos using microsatellite markers found strong population genetic structure between neighbouring islands, but not between

breeding colonies on each island (Duffie et al., 2009). Therefore, reproductive isolation following morphological and behavioural divergence in response to niche segregation likely explains the genetic structure observed in both species on Kerguelen Island. Speciation followed by niche segregation and adaptive divergence has been shown in other bird species as well as mammals and fish (Hoelzel et al., 1998; Huber et al., 2007; Musiani et al., 2007; Wolf et al., 2008; Foote et al., 2009; Louis et al., 2014). Moreover, it is believed that populations of the same species mostly avoid competetion for resources through spatial niche seggregation while different species avoid competition through niche seggregation in other dimensions than space (Taylor et al., 2019; Fromant et al., 2022).

Furthermore, natal philopatry is another common genetic barrier in seabirds (Inchausti & Weimerskirch, 2002; Burg & Croxall, 2004; Milot et al., 2008; Danckwerts et al., 2021; Herman et al., 2022), which can nevertheless be overcome under exceptional circumstances (Dugger et al., 2010). Even though, we have no data about natal philopatry in our study species, a study examining breeding philopatry at one colony of Kerguelen shags over a period of two years observed 70% of birds (not accounting for annual survival) returning to the same colony with 96% of these returning to the same nest as the previous year (Cook et al., 2013). There are no data available on philopatry of gentoo penguins from Kerguelen Island, but gentoo penguins breeding on other islands are known to show high site fidelity (Bost & Jouventin, 1990; Bost & Jouventin, 1991; Vianna et al., 2017). During the period 1987-1989, an average of 95% of gentoo penguins breeding on South Georgia Island returned to the same colony while 89-100% returned to the same nest (Williams & Rodwell, 1992). Therefore, natal philopatry seems to be one of

the main genetic barriers that facilitates the speciation followed by adaptive divergence due to niche segregation in both shags and gentoo penguins breeding on Kerguelen Island.

5.5.4 Taxonomic and conservation implications

There is no single definition for species. The biological species concept is a common concept used to define a species. Under the biological species concept, when two groups do not interbreed and are reproductively isolated, they are identified as different species. As gentoo penguins breeding on Kerguelen Island are reproductively isolated and genetically, morphologically, and behaviourally different from gentoo penguins breeding on other islands, they fit the criteria of distinct species under the biological species concept (de Dinechin et al., 2012; Vianna et al., 2017; Pertierra et al., 2020; Tyler et al., 2020; Noll et al., 2022). Therefore, as other studies on gentoo penguins also suggested, gentoo penguins breeding on Kerguelen Island needs to be identified as a separate species like Kerguelen shags. In addition, the combination of differences between two ecotypes of gentoo penguins and three breeding colonies of Kerguelen shags meets the criteria of species under the ecological species concept, in which groups sharing ecological niches showing adaptive divergence are identified as separate species. Therefore, it is also worth considering a taxonomic reassessment for both Kerguelen shag and Kerguelen gentoo penguins at the scale of this large archipelago based on the two ecotypes and breeding colonies.

Although Kerguelen shags and gentoo penguins are considered a species of least concern by IUCN (BirdLife International, 2018, 2020), more thorough conservation plans need to be implemented for both species as there are distinct genetic groups within both species

on Kerguelen Island. To provide effective management for small but distinct populations, biodiversity managers are using different conservation units such as evolutionary significant units (ESUs) and management units (MUs). An ESU is a historic evolutionary group which is substantially reproductively isolated from the rest of the species (Moritz, 1994) while a MU is a more recently diverged genetic group. As both species, Kerguelen shags and Kerguelen gentoo penguins are reproductively isolated and distinct, they need to be considered as ESUs. Moreover, on Kerguelen Island, at least two ecotypes of gentoo penguins and three sampled colonies of Kerguelen shags need to be treated as separate MUs as each group within both species are distinct from one another.

5.6 Conclusions

Regardless of the barriers limiting gene flow, there are distinct groups of Kerguelen shags and gentoo penguins breeding on Kerguelen Island. As these groups are morphologically, behaviourally, and genetically different, the taxonomy of shags and gentoo penguins breeding on the island needs to be reassessed. Moreover, further studies need to be carried out, especially for Kerguelen shags including samples from other colonies on the island to check if those colonies belong to one of the three genetic groups identified in this study or new genetic groups. Other inshore species such as common diving petrels should also be considered. Finally, it is clear that Kerguelen Island needs be given priority in conservation and management as it is a unique habitat with a very high genetic diversity despite its remoteness and harsh environmental conditions.

5.7 Acknowledgements

We thank Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship,

Alberta Society of Professional Biologists, and the Seabird Group for providing funding. Dr Kyle Wellband for technical support. Special thanks to members of Burg lab, University of Lethbridge for their huge support.

5.8 Author contributions

DKA, CAB and TMB designed the study and DKA completed the lab work. CAB led the field work and provided the samples. DKA completed the analyses. DKA and TMB wrote the manuscript with input from CAB.

5.9 Funding

French Polar Institute (Institut polaire français Paul-Émile-Victor), Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, Alberta Innovates-Technology Futures Graduate Student Scholarship, Alberta Society of Professional Biologists and the Seabird Group.

5.10 References

Abbott, C. L., & Double, M. C. (2003b). Genetic structure, conservation genetics and evidence of speciation by range expansion in shy and white-capped albatrosses. *Molecular Ecology*, *12*(11), 2953–2962. https://doi.org/10.1046/j.1365-294X.2003.01980.x

Adams, R. V., Lazerte, S. E., Otter, K. A., & Burg, T. M. (2016). Influence of landscape features on the microgeographic genetic structure of a resident songbird. *Heredity*, *117*(2), 63–72. https://doi.org/10.1038/hdy.2016.12

Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, *25*(22), 4692–4693. https://doi.org/10.1093/NAR/25.22.4692

Bezault, E., Balaresque, P., Toguyeni, A., Fermon, Y., Araki, H., Baroiller, J. F., & Rognon, X. (2011). Spatial and temporal variation in population genetic structure of wild Nile tilapia (*Oreochromis niloticus*) across Africa. *BMC Genetics*, *12*(1), 1–16. https://doi.org/10.1186/1471-2156-12-102/TABLES/6

BirdLife International (2018). *Leucocarbo verrucosus*. The IUCN Red List of Threatened Species 2018: e.T22696813A133555257.

BirdLife International (2020). *Pygoscelis papua*. The IUCN Red List of Threatened Species 2020: e.T22697755A157664581.

Blain, S., Tréguer, P., Belviso, S., Bucciarelli, E., Denis, M., Desabre, S., Fiala, M., Martin Jézéquel, V., Le Fèvre, J., Mayzaud, P., Marty, J. C., & Razouls, S. (2001). A biogeochemical study of the island mass effect in the context of the iron hypothesis: Kerguelen Islands, Southern Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, *48*(1), 163–187. https://doi.org/10.1016/S0967-0637(00)00047-9

Bost, C. A., & Jouventin, P. (1990) Evolutionary ecology of the gentoo penguin (*Pygoscelis papua*). In Penguin Biology (Darby, J. T. & L. S. Davis, Eds.), pp. 85–112, Academic Press, Orlando.

Bost, C. A., P. Jouventin (1991). The breeding performance of the gentoo penguin P*ygoscelis papua* at the northern edge of its range. *Ibis* (*33*), 14–25.

Bost, C. A., Jouventin, P., & du Sel, N. P. (1992). Morphometric variability on a microgeographical scale in two inshore seabirds. *Journal of Zoology*, *226*(1), 135–149. https://doi.org/10.1111/J.1469-7998.1992.TB06132.X

Burg, T. M., & Croxall, J. P. (2004). Global population structure and taxonomy of the wandering albatross species complex. *Molecular Ecology*, *13*(8), 2345–2355. https://doi.org/10.1111/j.1365-294X.2004.02232.x

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

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks:

An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Cheek, R. G., Forester, B. R., Salerno, P. E., Trumbo, D. R., Langin, K. M., Chen, N., Sillett, T. S., Morrison, S. A., Ghalambor, C. K., & Funk, C. W. (2022). Habitat-linked genetic variation supports microgeographic adaptive divergence in an island-endemic bird species. *Molecular Ecology*, *31*(10), 2830–2846. https://doi.org/10.1111/MEC.16438

Cook, T. R., Lescroël, A., Cherel, Y., Kato, A., & Bost, C. A. (2013). Can foraging ecology drive the evolution of body size in a diving endotherm? *PLoS One*, *8*(2), e56297. https://doi.org/10.1371/JOURNAL.PONE.0056297

Danckwerts, D. K., Humeau, L., Pinet, P., McQuaid, C. D., & Le Corre, M. (2021). Extreme philopatry and genetic diversification at unprecedented scales in a seabird. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-86406-9

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/BIOINFORMATICS/BTR330

Danielsson, P. E. (1980). Euclidean distance mapping. *Computer Graphics and Image Processing*, *14*(3), 227–248. https://doi.org/10.1016/0146-664X(80)90054-4

de Dinechin, M., Dobson, F.S., Zehtindjiev, P., Metcheva, R., Couchoux, C., Martin, A., Quillfeldt, P. and Jouventin, P. (2012). The biogeography of gentoo penguins (*Pygoscelis papua*). *Canadian Journal of Zoology*, *90(*3), 352–360.

Duffie, C. V., Glenn, T. C., Vargas, F. H., & Parker, P. G. (2009). Genetic structure within and between island populations of the flightless cormorant (*Phalacrocorax harrisi*). *Molecular Ecology*, *18*(10), 2103–2111.

Dugger, K. M., Ainley, D. G., Lyver, P. O. B., Barton, K., & Ballard, G. (2010). Survival differences and the effect of environmental instability on breeding dispersal in an Adélie penguin meta-population. *Proceedings of the National Academy of Sciences*, *107*(27), 12375–12380.

Estivals, G., Duponchelle, F., García-Dávila, C., Römer, U., Mariac, C., & Renno, J. F. (2022). Exceptional genetic differentiation at a micro-geographic scale in *Apistogramma agassizii* (Steindachner, 1875) from the Peruvian Amazon: Sympatric speciation? *Evolutionary Biology*, *1*, 1–17. https://doi.org/10.1007/S11692-022-09587-9/FIGURES/5

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/J.1755-0998.2010.02847.X

Foote, A. D., Newton, J., Piertney, S. B., Willerslev, E., & Gilbert, M. T. P. (2009). Ecological, morphological and genetic divergence of sympatric North Atlantic killer

whale populations. *Molecular Ecology*, *18*(24), 5207–5217.

Friesen, V. L., Burg, T. M., & McCoy, K. D. (2007). Mechanisms of population differentiation in seabirds: Invited review. *Molecular Ecology*, *16*(9), 1765–1785. https://doi.org/10.1111/j.1365-294X.2006.03197.x

Fromant, A., Arnould, J.P., Delord, K., Sutton, G.J., Carravieri, A., Bustamante, P., Miskelly, C.M., Kato, A., Brault-Favrou, M., Cherel, Y. and Bost, C.A., 2022. Stage-dependent niche segregation: insights from a multi-dimensional approach of two sympatric sibling seabirds. *Oecologia*, *199*(3), 537–548.

Halsey, L. G., Butler, P. J., & Blackburn, T. M. (2006). A phylogenetic analysis of the allometry of diving. *The American Naturalist*, *167*(2), 276–287.

Herman, R. W., Winger, B. M., Dittmann, D. L., & Harvey, M. G. (2022). Fine-scale population genetic structure and barriers to gene flow in a widespread seabird (*Ardenna pacifica*). *Biological Journal of the Linnean Society*, *137*(1), 125–136. https://doi.org/10.1093/BIOLINNEAN/BLAC091

Hoelzel, A. R., Dahlheim, M., & Stern, S. J. (1998). Low genetic variation among killer whales (*Orcinus orca*) in the eastern North Pacific and genetic differentiation between foraging specialists. *Journal of Heredity*, *89*(2), 121–128.

Hollatz, C., Vilaça, S. T., Redondo, R. A. F., Marmontel, M., Baker, C. S., & Santos, F. R. (2011). The Amazon River system as an ecological barrier driving genetic differentiation of the pink dolphin (*Inia geoffrensis*). *Biological Journal of the Linnean Society*, *102*(4), 812–827. https://doi.org/10.1111/J.1095-8312.2011.01616.X

Huber, S. K., León, L. F. D., Hendry, A. P., Bermingham, E., & Podos, J. (2007). Reproductive isolation of sympatric morphs in a population of Darwin's finches. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1619), 1709–1714.

Inchausti, P., & Weimerskirch, H. (2002). Dispersal and metapopulation dynamics of an oceanic seabird, the wandering albatross, and its consequences for its response to long-line fisheries. *Journal of Animal Ecology*, *71*(5), 765–770. https://doi.org/10.1046/J.1365-2656.2002.00638.X

Jombart, T., & Bateman, A. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/BIOINFORMATICS/BTN129

Jouventin, P. (1982) Visual and vocal signals in penguins, their evolution and adaptive characters. Advances in ethology, vol 24. Paul Parey Verlag, Hamburg and Berlin

Lescroël, A., & Bost, C. A. (2005). Foraging under contrasting oceanographic conditions: The gentoo penguin at Kerguelen Archipelago. *Marine Ecology Progress Series*, *302*, 245–261. https://doi.org/10.3354/MEPS302245

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299. https://doi.org/10.1093/BIOINFORMATICS/BTR642

Louis, M., Fontaine, M.C., Spitz, J., Schlund, E., Dabin, W., Deaville, R., Caurant, F., Cherel, Y., Guinet, C. and Simon-Bouhet, B., 2014. Ecological opportunities and specializations shaped genetic divergence in a highly mobile marine top predator. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1795), 20141558.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10–12. https://doi.org/10.14806/EJ.17.1.200

Mayr, E. (1942). Systematics and the Origin of Species. Columbia University Press, New York, NY, USA.

Mayr, E. (1996). What is a species, and what is not? *Philosophy of Science*, *63*(2), 262–277.

Mayr, E. (2000). The biological species concept. Species concepts and phylogenetic theory: a debate. Columbia University Press, New York, 17–29.

Mikles, C. S., Aguillon, S. M., Chan, Y. L., Arcese, P., Benham, P. M., Lovette, I. J., & Walsh, J. (2020). Genomic differentiation and local adaptation on a microgeographic scale in a resident songbird. *Molecular Ecology*, *29*(22), 4295–4307. https://doi.org/10.1111/MEC.15647

Milot, E., Weimerskirch, H., & Bernatchez, L. (2008). The seabird paradox: dispersal, genetic structure and population dynamics in a highly mobile, but philopatric albatross species. *Molecular Ecology*, 17(7), 1658–1673.

Moritz, C. (1994). Defining 'evolutionarily significant units' for conservation. *Trends in Ecology & Evolution*, *9*(10), 373–375.

Munro, K. J., & Burg, T. M. (2017). A review of historical and contemporary processes affecting population genetic structure of Southern Ocean seabirds. *Emu - Austral Ornithology*, *117*(1), 4–18. https://doi.org/10.1080/01584197.2016.1271988

Musiani, M., Leonard, J.A., Cluff, H.D., Gates, C.C., Mariani, S., Paquet, P.C., Vila, C. and Wayne, R.K., 2007. Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat colour and association with migratory caribou. *Molecular Ecology*, *16*(19), 4149–4170.

Noll, D., Leon, F., Brandt, D., Pistorius, P., Le Bohec, C., Bonadonna, F., Trathan, P. N., Barbosa, A., Rey, A. R., Dantas, G. P. M., Bowie, R. C. K., Poulin, E & Vianna, J. A. (2022). Positive selection over the mitochondrial genome and its role in the diversification of gentoo penguins in response to adaptation in isolation. *Scientific Reports*, *12*, 3767.

Pascual, M., Rives, B., Schunter, C., & Macpherson, E. (2017). Impact of life history traits on gene flow: A multispecies systematic review across oceanographic barriers in the Mediterranean Sea. *PLoS One*, *12*(5), e0176419. https://doi.org/10.1371/JOURNAL.PONE.0176419

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537.

https://doi.org/10.1093/BIOINFORMATICS/BTS460

Pertierra, L. R., Segovia, N. I., Noll, D., Martinez, P. A., Pliscoff, P., Barbosa, A., Aragón, P., Raya Rey, A., Pistorius, P., Trathan, P., Polanowski, A., Bonadonna, F., Le Bohec, C., Bi, K. E., Wang-Claypool, C. Y., González-Acuña, D., Dantas, G. P. M., Bowie, R. C. K., Poulin, E., & Vianna, J. A. (2020). Cryptic speciation in gentoo penguins is driven by geographic isolation and regional marine conditions: Unforeseen vulnerabilities to global change. *Diversity and Distributions*, *26*, 958–975.

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One*, *7*(2), e32253. https://doi.org/10.1371/JOURNAL.PONE.0032253

Pujolar, J. M., Lucarda, A. N., Simonato, M., & Patarnello, T. (2011). Restricted gene flow at the micro- and macro-geographical scale in marble trout based on mtDNA and microsatellite polymorphism. *Frontiers in Zoology*, *8*(1), 1–10. https://doi.org/10.1186/1742-9994-8-7/TABLES/5

Robertson, J. M., & Vega, A. (2011). Genetic and phenotypic variation in a colourful treefrog across five geographic barriers. *Journal of Biogeography*, *38*(11), 2122–2135. https://doi.org/10.1111/J.1365-2699.2011.02548.X

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754. https://doi.org/10.1111/mec.15253

Ruiz-Garcia, M., & Klein, K. K. (1997). Genetic structure of populations of the domestic cat in Catalonia (Spain) and upper mid western USA: A microgeographic and macrogeographic study. *Journal of Genetics*, *76*(2), 99–115. https://doi.org/10.1007/BF02923555/METRICS

Schulte-Hostedde, A. I., Lisle Gibbs, H., & Millar, J. S. (2001). Microgeographic genetic structure in the yellow-pine chipmunk (*Tamias amoenus*). *Molecular Ecology*, *10*(7), 1625–1631. https://doi.org/10.1046/J.1365-294X.2001.01298.X

Taylor, R.S., Bolton, M., Beard, A., Birt, T., Deane-Coe, P., Raine, A.F., Gonzalez-Solis, J., Lougheed, S.C. and Friesen, V.L., (2019). Cryptic species and independent origins of allochronic populations within a seabird species complex (*Hydrobates* spp.). *Molecular Phylogenetics and Evolution*, *139*, 106552.

Thanou, E., Sponza, S., Nelson, E. J., Perry, A., Wanless, S., Daunt, F., & Cavers, S. (2017). Genetic structure in the European endemic seabird, *Phalacrocorax aristotelis*, shaped by a complex interaction of historical and contemporary, physical and nonphysical drivers. *Molecular Ecology*, *26*(10), 2796–2811.

Tyler, J., Bonfitto, M. T., Clucas, G. V., Reddy, S., & Younger, J. L. (2020). Morphometric and genetic evidence for four species of gentoo penguin. *Ecology and Evolution*, *10*, 13836–13846.

Verdade, L. M., Zucoloto, R. B., & Coutinho, L. L. (2002). Microgeographic variation in *Caiman latirostris*. *Journal of Experimental Zoology*, *294*(4), 387–396. https://doi.org/10.1002/JEZ.10200

Vianna, J. A., Noll, D., Dantas, G. P. M., Petry, M. V., Barbosa, A., González-Acuña, D., Le Bohec, C., Bonadonna, F., & Poulin, E. (2017). Marked phylogeographic structure of gentoo penguin reveals an ongoing diversification process along the Southern Ocean. *Molecular Phylogenetics and Evolution*, *107*, 486–498. https://doi.org/10.1016/J.YMPEV.2016.12.003

White, M. M., & Turner, B. J. (1984). Microgeographic differentiation in a stream population of *Goodea atripinnis* (Goodeidae) from the Mexican Plateau. *Environmental Biology of Fishes*, *10*(1–2), 123–127. https://doi.org/10.1007/BF00001669

Widmer, A., Lexer, C., & Cozzolino, S. (2008). Evolution of reproductive isolation in plants. *Heredity*, *102*(1), 31–38. https://doi.org/10.1038/hdy.2008.69

Williams, T. D., & Rodwell, S. (1992). Annual variation in return rate, mate and nest-site fidelity in breeding gentoo and macaroni penguins. *The Condor*, *94*(3), 636–645.

Wolf, J. B., Harrod, C., Brunner, S., Salazar, S., Trillmich, F., & Tautz, D. (2008). Tracing early stages of species differentiation: ecological, morphological and genetic divergence of Galápagos sea lion populations. *BMC Evolutionary Biology*, *8*, 1–14.

Zhang, Z. Y., Zheng, X. M., & Ge, S. (2007). Population genetic structure of *Vitex negundo* (Verbenaceae) in Three-Gorge area of the Yangtze River: The riverine barrier to seed dispersal in plants. *Biochemical Systematics and Ecology*, *35*(8), 506–516. https://doi.org/10.1016/J.BSE.2007.01.014

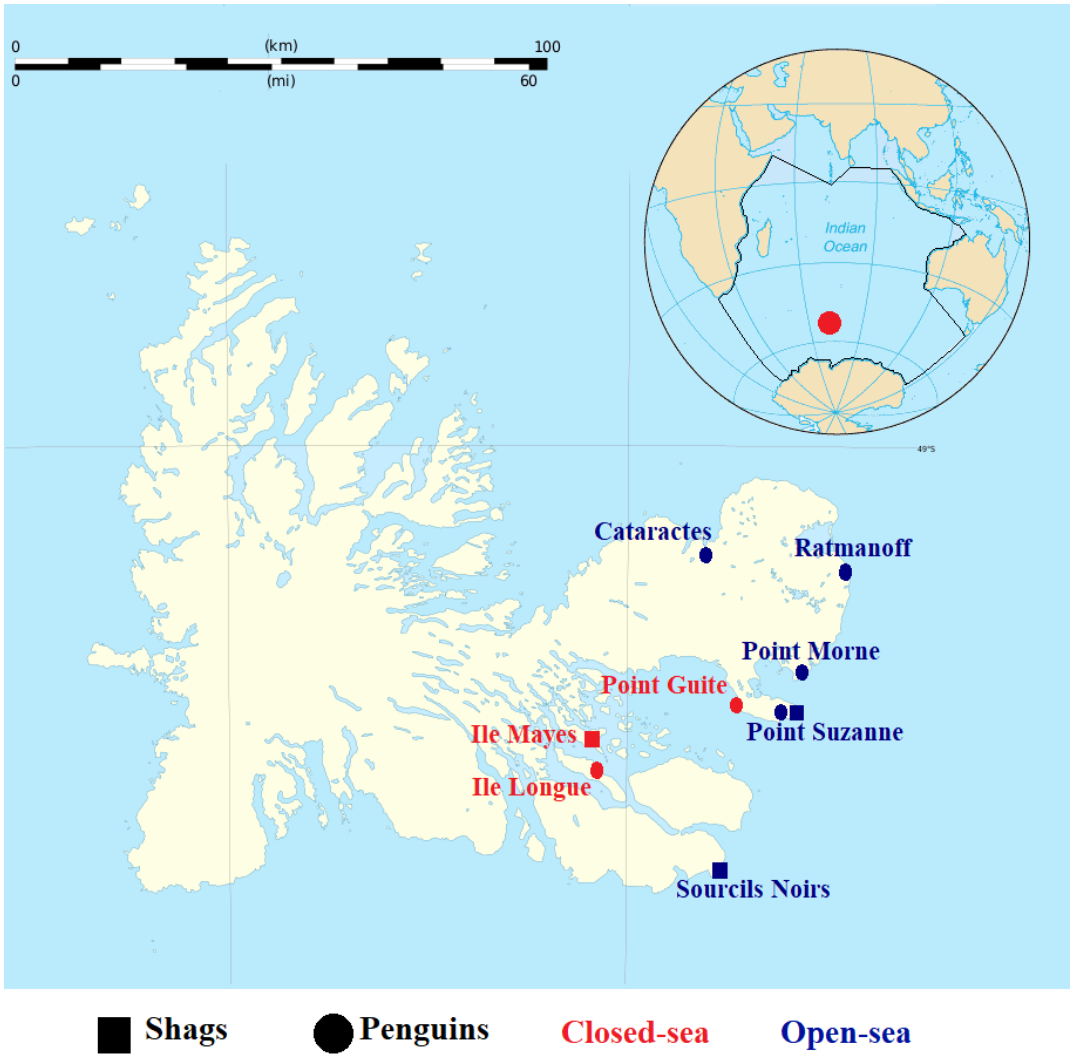

Figure 5.1: Sampled colonies of shags (square) and gentoo penguins (circle) breeding on Kerguelen Island in the Indian Ocean.

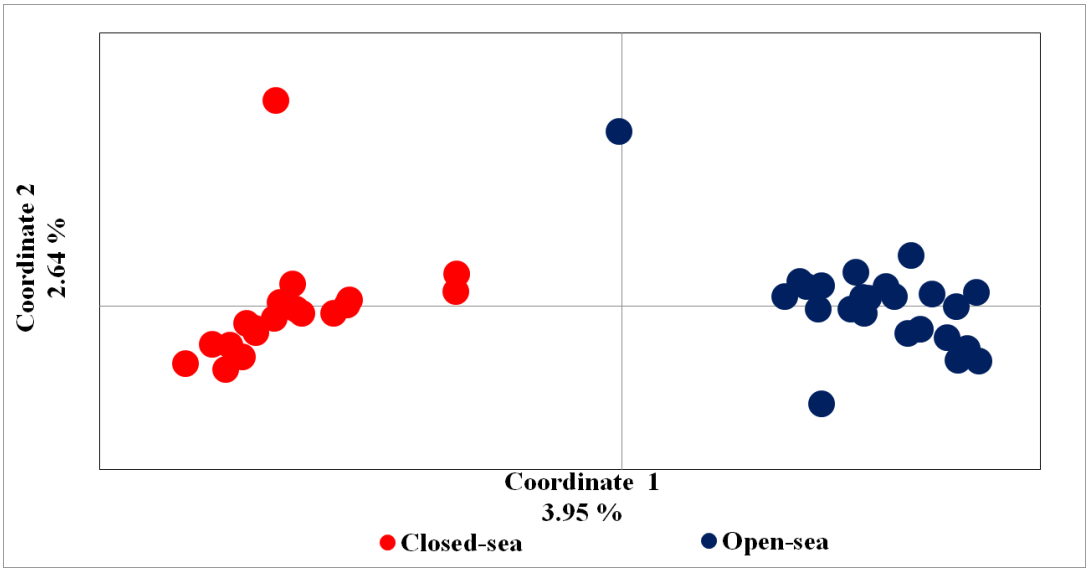

Figure 5.2: PCoA plot representing the genetic variation between closed-sea (red) and open-sea (blue) ecotypes of gentoo penguin breeding on Kerguelen Island (44 individuals and 49,263 SNPs).

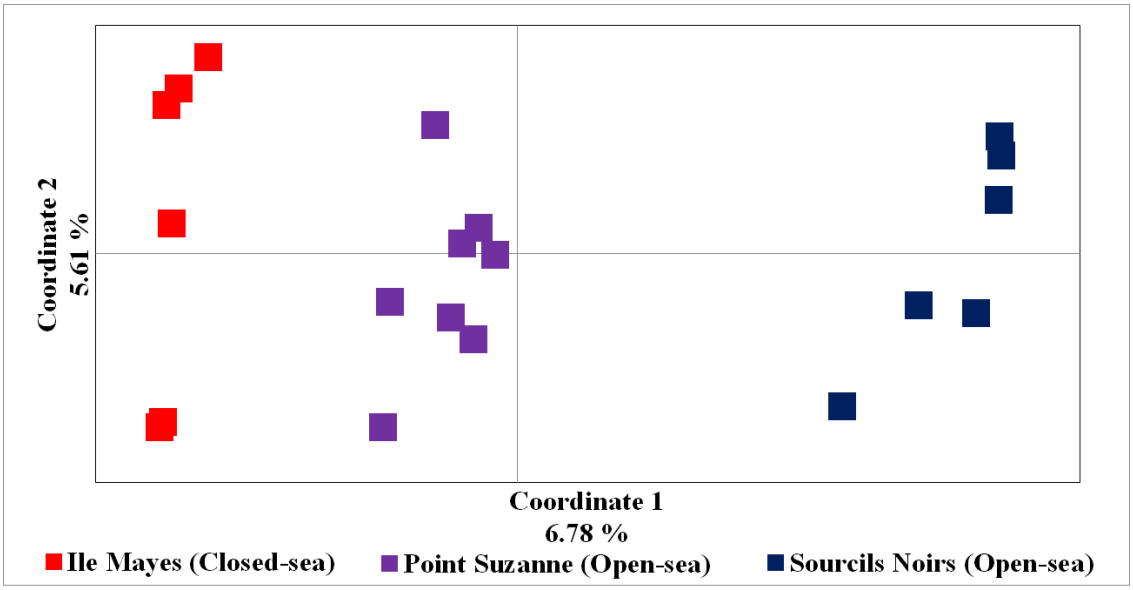

Figure 5.3: PCoA plot representing the genetic variation between three sampled populations of Kerguelen shags (20 individuals and 66,503 SNPs).

Table 5.1: Comparison of observed (H_O) and expected (H_E) heterozygosity for open and closed-sea colonies of gentoo penguins and Kerguelen shags breeding on Kerguelen Island

	Closed-sea		Open-sea	
	H_O	H_E	H_O	H_E
Gentoo penguin	0.177	0.189	0.168	0.192
Kerguelen shag	0.218	0.194	0.226	0.233

Table 5.2: Pairwise F_{ST} values (below the diagonal) and associated P values (above the diagonal) compared between sampled colonies of Kerguelen shag breeding on Kerguelen Island. All values are significant.

	Île Mayes	Point Suzanne	Sourcils Noirs
Île Mayes_Closed	-----	<0.001	<0.001
Point Suzanne_Open	0.015	-----	<0.001
Sourcils Noirs_Open	0.056	0.026	-----

Table 5.3: Comparison of observed (H_O) and expected (H_E) heterozygosity for three sampled colonies of Kerguelen shags breeding on Kerguelen Island.

Breeding Colony	H_O	H_E
Île_Mayes_Closed	0.218	0.194
Point_Suzanne_Open	0.230	0.211
Sourcils_Noirs_Open	0.212	0.186

5.11 Supplementary materials

Table S5.1: Number of individuals sequenced and used in downstream analyses for each species.

Population	Number of individuals sequenced	Number of individuals with more than 10,000 sequences	Initial number of individuals with more than 50% data
Gentoo penguin			
Open-sea	55	53	25
Closed-sea	31	28	19
Kerguelen shag			
Île Mayes (Closed-sea)	11	9	6
Point Suzanne (Open-sea)	10	10	8
Sourcils Noirs (Open-sea)	11	9	6

Chapter 6: General discussion

6.1 Population genetic structure in seabirds

All five study species from the Southern Ocean- Atlantic yellow-nosed albatross, Indian yellow-nosed albatross, sooty albatross, Kerguelen shag and Kerguelen gentoo penguin- showed distinct population genetic structure (Figure 6.1). Population differentiation observed in the three albatross species was at a macrogeographic scale while for the two species on Kerguelen Island it was within an island. Despite differences in spatial scales, common barriers are impacting gene flow in Southern Ocean seabirds (Inchausti & Weimerskirch, 2002; Burg & Croxall, 2004; Steeves et al., 2005; Friesen et al., 2007; Calderón et al., 2014; Munro & Burg, 2017; Makhado et al., 2018; Taylor et al., 2018; Danckwerts et al., 2021).

6.1.1 Population genetic structure between ocean basins

Breeding ranges of some seabird species are limited to one ocean basin while others breed on several islands in different ocean basins in the Southern Ocean. Most of these islands or island groups are isolated from one another and yet seem to be connected as seabirds are capable of flying or swimming longer distances (Milot et al., 2008). Regardless of the potential connectivity, breeding populations of some species in different ocean basins are reproductively isolated with reduced gene flow and show population genetic structure between ocean basins. Both physical and non-physical barriers such as geographic distance, ocean currents, seascapes, local adaptation, sea surface temperature, at-sea distribution, philopatry and breeding phenology limit gene flow between seabird populations breeding in different ocean basins (Friesen et al., 2007; Bicknell et al., 2012; Munro & Burg, 2017; Thanou et al., 2017; Vianna et al., 2017; Rexer-Huber et al., 2019;

Torres et al., 2021; Herman et al., 2022). Population differentiation between ocean basins is seen not only in seabirds but also in other marine organisms such as turtles, oysters and fish species (Mirams et al., 2011; Lal et al., 2017; Roberts et al., 2019). Supporting all these studies, my study also confirms that barriers such as natal philopatry and at-sea distribution can result in strong genetic structuring between populatons of seabird species like albatrosses breeding in different ocean basins.

6.1.2 Population genetic structure within ocean basins

Although some islands or island groups are located in the same ocean basin experiencing less oceanographic variation, gene flow among breeding populations on these islands can still be limited resulting in within ocean population genetic structure. Most of the time in seabirds, non-physical barriers such as natal philopatry, at-sea distribution, breeding phenology and local adaptation play a major role in shaping the population genetic structure within oceans rather than physical barriers (Abbott & Double, 2003b; Dearborn et al., 2003; Friesen et al., 2007, Munro & Burg, 2017; Vianna et al., 2017; Rexer-Huber et al., 2019). In great frigatebirds (*Fregata minor*) breeding in central Pacific Ocean, significant genetic differentiation was observed between islands (2,125 km apart) regardless of their regular movements between islands, but the force driving this population structure was not clear (Dearborn et al., 2003). Similar patterns were observed in shy albatrosses and white-chinned petrels on islands close to each other (Abbott & Double, 2003b; Rexer-Huber et al., 2019). In my study, population differentaition observed within both Atlantic and Indian yellow-nosed albatrosses confirms that a non-physical barrier like natal philopatry can drive population genetic differentiation in highly mobile seabirds breeding within the same ocean basin.

6.1.3 Population genetic structure within islands

Sympatric speciation, or speciation through reproductive isolation of populations which are not separated by physical barriers to prevent them interbreeding while they live in the same geographic range, is not as common and hard to detect (Bolnick & Fitzpatrick, 2007). Although sympatric speciation is said to be more popular among sedentary groups like plants, it is rarely seen in mobile species like fish and birds (Gíslason et al., 1999; Barluenga et al., 2006; Friesen et al., 2007; Taylor et al., 2019). For example, Taylor et al. (2019) found genetic differentiation between hot and cold season breeding populations of band-rumped storm petrels (*Hydrobates castro sensu lato*) breeding at different times of the year at the same site in Azores. The open-sea and closed-sea ecotypes of both Kerguelen shag and gentoo penguin studied in my project are good examples of sympatric speciation as two ecotypes in each species breed on the same island and there are no obvious barriers between the colonies, yet they are morphologically, behaviourally and genetically different with no evidence of admixture.

6.2 Genetic barriers in seabirds

Population differentiation seen at different levels in seabirds is mostly driven by non-physical barriers as they face limited number of physical barriers and are capable of moving longer distances overcoming physical barriers like geographic distance. For non-physical barriers, strong philopatry is commonly seen among seabirds. Even at small spatial scales, philopatry can result population genetic structure (Harris, 1973; Prince et al., 1994; Weimerskirch et al., 1997; Sagar et al., 1998; Abbott & Double, 2003b; Friesen et al., 2007; Coulson, 2016; Rexer-Huber et al., 2019; Herman et al., 2022). Genetic structure in Galapagos petrels (*Diomedea irrorate*), a species endemic to Galapagos

archipelago, is believed to be driven by strong philopatry as there are no other physical or non-physical barrier to restrict gene flow between the islands within the archipelago (Harris, 1973).

For my study species, there are no published data on philopatry. But in Atlantic yellow-nosed albatross, researchers have observed high levels of site fidelity in adults, where they seldom nest more than 100 m from the previous year's nest and virtually all chicks are recruited within 1 km of their natal nest (PGR unpubl. data). There is also no evidence of exchange of ringed birds between breeding islands (PGR pers. obs.). There are known cases where albatrosses expand their ranges and start new breeding sites. For example, a pair of Indian yellow-nosed albatross was observed breeding on the Chatham Islands, east of New Zealand, which is not a known breeding island for this species. Also, black-browed albatross *Thalassarche melanophris* recently colonized Campbell Island and limited hybridization occurs with Campbell albatross (*T. impavida*) (Moore et al., 1997). This confirms possible occasional long-range dispersal in seabirds. Moreover, natal philopatry in certain seabirds does not limit faithfulness only to their breeding island or colony, but also to their nest. A study examining philopatry only at one colony of Kerguelen shags over a period of two years observed 70% of birds returning to the same colony while 90% returned to the same nest as the previous year (Cook et al., 2013). Natal philopatry is common among species as it makes individuals knowledgeable about critical factors like availability of food, predators and competition, making them and their offspring more prepared to survive during the breeding season than they would in a new environment. But natal philopatry is not always advantageous and it can also result high population density and inbreeding (Coulson, 2016).

One way to minimize competition is breeding phenology or individuals of the same species breeding during different times of the year at the same location. Taylor et al. (2019) identified two genetically distinct groups corresponding to hot and cold season breeders in band-rumped storm petrel (*Hydrobates spp.*) breeding on Azores Archipelago. For my study species, there is no evidence for differences in breeding phenology. On average, Atlantic yellow-nosed albatrosses breed a week or so later on the more southerly Gough Island compared to birds breeding at Tristan (ACAP, 2009), but this is unlikely to prevent potential gene flow, because individual variation within populations is higher than the difference between Tristan and Gough Islands (PGR pers. obs.). In sooty albatrosses, birds from two Atlantic Ocean populations start laying eggs about two weeks before the Marion Island population. Similar to Atlantic yellow-nosed albatrosses, this temporal difference is not large enough to maintain the observed genetic differentiation in sooty albatrosses. The three other study species do not have any known temporal differences in breeding times within species. Therefore, breeding phenology does not seem to be one of the main drivers of speciation in my study species.

Differences in foraging distribution is considered as a genetic barrier for seabirds (Burg & Croxall, 2001). Seabird tracking database of BirdLife International shows foraging distributions of both Atlantic and Indian yellow-nosed albatrosses and Atlantic and Indian Ocean populations of sooty albatross are distinct with minimal overlap. Within both Atlantic and Indian yellow-nosed albatrosses, substantial partitioning in foraging ranges of different breeding populations can be seen with some overlap (Makhado et al., 2018). But in other albatross species such as black-browed albatross, genetic clusters corresponding to their distinct at-sea foraging grounds have been identified (Burg &

Croxall, 2001). Moreover, smaller spatial scale niche partitioning with respect to foraging occurring in both the Kerguelen shags and the gentoo penguins from Kerguelen Island correspond to their genetic structure. Therefore, at-sea foraging distribution is a major genetic barrier for my study species, especially for populations between oceans and even within a island like in several other seabird species (Abbott & Double, 2003a; Burg & Croxall, 2001; Friesen et al., 2007; Munro & Burg, 2017; Makhado et al., 2018).

6.3 Conservation units

When gene flow in small populations is highly restricted, their genetic diversity is reduced leading to inbreeding depression. As a result, small populations might not be able to survive events such as pandemics and climate change. Population sizes of many Southern Ocean seabird colonies are relatively small with only one or two large colonies containing more than half of the global population, and many species are globally threatened. For example, according to BirdLife International, 15/22 and 9/18 species of albatrosses (68% of family Diomedeidae) and penguins respectively are listed as threatened (BirdLife International 2012). To prevent loss of small but critical populations, biodiversity managers are using different conservation units such as evolutionary significant units (ESUs) and management units (MUs). An ESU is defined as a historically reproductively isolated population which is genetically and ecologically distinct from the rest of the species while a MU is a demographically independent, recently diverged population (Moritz 1994; Allendorf et al., 2012). An ESU may contain several MUs, where these management units can be diverging subpopulations of a metapopulation which is an ESU.

Based on these definitions of ESU and MU, my study species Atlantic yellow-nosed albatross, Indian yellow-nosed albatross, sooty albatross populations from Atlantic and Indian Ocean basins separately, Kerguelen shag and Kerguelen gentoo penguin can be identified as ESUs (Cook et al., 2013; Vianna et al., 2017). Moreover, as each breeding population of Atlantic and Indian yellow-nosed albatrosses, all sampled colonies of Kerguelen shags and two ecotypes of Kerguelen gentoo penguins are genetically and ecologically distinct, they are separate MUs.

6.4 Conclusions

Non-physical barriers: natal philopatry and foraging distribution are main genetic barriers leading to reproductive isolation and speciation in Southern Ocean seabirds. Since all of my study species along with Atlantic and Indian Ocean basin counterparts of sooty albatross are evolutionary significant units while all the sampled populations of two species of yellow-nosed albatross, colonies of Kerguelen shags and two ecotypes of gentoo penguins are separate management units, they should be considered accordingly for conservation and management regardless of the global population size and number of breeding populations or colonies. As Atlantic and Indian Ocean basin populations of sooty albatross, two ecotypes of gentoo penguins and three sampled breeding colonies of Kerguelen shags are all behaviourally, morphologically and genetically different from one another within each species, the taxonomies of the three species need to be reassessed to new species or subspecies.

6.5 Future directions

A study similar to Antoniou et al. (2023) incorporating environmental data both at breeding sites and at-sea ranges with genotypic data would reveal how local

environmental conditions affect the adaptive and neutral genomic variation in Southern Ocean seabirds as well as how these birds would response to global climate change.

Also, all three breeding colonies of Kerguelen shags studied in my study are genetically, morphologically and behaviourally different. But there are other Kerguelen shag colonies on the Kerguelen Island. A future study with samples from other shag colonies on Kerguelen Island would allow us to see if they are also distinct.

In my study, using only mtDNA and microsatellite markers, I could not detect fine genetic resolution between sooty albatross populations within Atlantic and Indian Ocean basins. Therefore, a future study with RADSeq data for sooty albatrosses to examine the genetic connectivity between populations within each ocean as well test for the secondary contact of Tristan da Cunha sooty albatrosses with Indian Ocean populations.

In my study, I used mitochondrial markers, traditional nuclear markers, microsatellite markers and RADSeq data to identify population genetic structure in five seabird species breeding in the Southern Ocean. Future research studying whole genome sequences of these species would provide more detailed insight on what parts of the genome are diverging and which genes are affected (Campo et al., 2013; Love et al., 2016; Jensen et al., 2021).

6.6 References

Abbott, C. L., & Double, M. C. (2003a). Phylogeography of shy and white-capped albatrosses inferred from mitochondrial DNA sequences: Implications for population history and taxonomy. *Molecular Ecology*, *12*(10), 2747–2758. https://doi.org/10.1046/j.1365-294X.2003.01944.x

Abbott, C. L., & Double, M. C. (2003b). Genetic structure, conservation genetics and evidence of speciation by range expansion in shy and white-capped albatrosses. *Molecular Ecology*, *12*(11), 2953–2962.

Agreement on the Conservation of Albatrosses and Petrels (2009b). ACAP Species assessment: Atlantic yellow-nosed albatross *Thalassarche chlororhynchos*.

Allendorf, F. W., Luikart, G. H., & Aitken, S. N. (2012). Conservation and the genetics of populations. John Wiley & Sons.

Antoniou, A., Manousaki, T., Ramírez, F., Cariani, A., Cannas, R., Kasapidis, P., Magoulas, A., Albo-Puigserver, M., Lloret-Lloret, E., Bellido, J.M. & Pennino, M.G. (2023). Sardines at a junction: seascape genomics reveals ecological and oceanographic drivers of variation in the NW Mediterranean Sea. *Molecular Ecology*, *32*(7), 1608-1628.

Barluenga, M., Stölting, K. N., Salzburger, W., Muschick, M., & Meyer, A. (2006). Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, *439*(7077), 719-723.

Bicknell, A. W. J., Knight, M. E., Bilton, D., Reid, J. B., Burke, T., & Votier, S. C. (2012). Population genetic structure and long-distance dispersal among seabird populations: Implications for colony persistence. *Molecular Ecology*, *21*(12), 2863–2876.

BirdLife International (2012) Spotlight on seabirds. Presented as part of the BirdLife State of the world's birds website.

Bolnick, D. I., & Fitzpatrick, B. M. (2007). Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics*, *38*, 459–487.

Burg, T. M., & Croxall, J. P. (2004). Global population structure and taxonomy of the wandering albatross species complex. *Molecular Ecology*, *13*(8), 2345–2355. https://doi.org/10.1111/j.1365-294X.2004.02232.x

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

Campo, D., Lehmann, K., Fjeldsted, C., Souaiaia, T., Kao, J., & Nuzhdin, S. V. (2013). Whole-genome sequencing of two North American *Drosophila melanogaster* populations reveals genetic differentiation and positive selection. *Molecular Ecology*, *22*(20), 5084–5097. https://doi.org/10.1111/MEC.12468

Cook, T. R., Lescroël, A., Cherel, Y., Kato, A., & Bost, C. A. (2013). Can foraging

ecology drive the evolution of body size in a diving endotherm? *PLoS One*, *8*(2), e56297. https://doi.org/10.1371/JOURNAL.PONE.0056297

Coulson, J. C. (2016). A review of philopatry in seabirds and comparisons with other waterbird species. *Waterbirds*, *39*(3), 229–240. https://doi.org/10.1675/063.039.0302

Danckwerts, D. K., Humeau, L., Pinet, P., McQuaid, C. D., & Le Corre, M. (2021). Extreme philopatry and genetic diversification at unprecedented scales in a seabird. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-86406-9

Dearborn, D. C., Anders, A. D., Schreiber, E. A., Adams, R. M., & Mueller, U. G. (2003). Inter-island movements and population differentiation in a pelagic seabird. *Molecular Ecology*, *12*(10), 2835–2843.

Friesen, V. L., Burg, T. M., & McCoy, K. D. (2007). Mechanisms of population differentiation in seabirds: Invited review. *Molecular Ecology*, *16*(9), 1765–1785. https://doi.org/10.1111/j.1365-294X.2006.03197.x

Gíslason, D., Ferguson, M. M., Skúlason, S., & Snorrason, S. S. (1999). Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (*Salvelinus alpinus*). *Canadian Journal of Fisheries and Aquatic Sciences*, 56(12), 2229–2234.

Harris, M. P. (1973). The biology of the waved albatross *Diomedea irrorata* of Hood Island, Galapagos. *Ibis*, *115*(4), 483–510. https://doi.org/10.1111/J.1474-919X.1973.TB01988.X

Herman, R. W., Winger, B. M., Dittmann, D. L., & Harvey, M. G. (2022). Fine-scale population genetic structure and barriers to gene flow in a widespread seabird (*Ardenna pacifica*). *Biological Journal of the Linnean Society*, *137*(1), 125–136. https://doi.org/10.1093/BIOLINNEAN/BLAC091

Inchausti, P., & Weimerskirch, H. (2002). Dispersal and metapopulation dynamics of an oceanic seabird, the wandering albatross, and its consequences for its response to long-line fisheries. *Journal of Animal Ecology*, *71*(5), 765–770. https://doi.org/10.1046/J.1365-2656.2002.00638.X

Jensen, A., Lillie, M., Bergström, K., Larsson, P., & Höglund, J. (2021). Whole genome sequencing reveals high differentiation, low levels of genetic diversity and short runs of homozygosity among Swedish wels catfish. *Heredity*, *127*(1), 79–91. https://doi.org/10.1038/s41437-021-00438-5

Lal, M. M., Southgate, P. C., Jerry, D. R., Bosserelle, C., & Zenger, K. R. (2017). Swept away: ocean currents and seascape features influence genetic structure across the 18,000 km Indo-Pacific distribution of a marine invertebrate, the black-lip pearl oyster P*inctada margaritifera*. *BMC Genomics*, *18*, 1–21.

Love, R. R., Steele, A. M., Coulibaly, M. B., Traore, S. F., Emrich, S. J., Fontaine, M. C., & Besansky, N. J. (2016). Chromosomal inversions and ecotypic differentiation in *Anopheles gambiae*: the perspective from whole-genome sequencing. *Molecular Ecology*, *25*(23), 5889–5906. https://doi.org/10.1111/MEC.13888

Makhado, A. B., Crawford, R. J. M., Dias, M. P., Dyer, B. M., Lamont, T., Pistorius, P.,

Ryan, P. G., Upfold, L., Weimerskirch, H., & Reisinger, R. R. (2018). Foraging behaviour and habitat use by Indian yellow-nosed albatrosses (*Thalassarche carteri*) breeding at Prince Edward Island. *Emu - Austral Ornithology*, *118*(4), 353–362. https://doi.org/10.1080/01584197.2018.1469959

Milot, E., Weimerskirch, H., & Bernatchez, L. (2008). The seabird paradox: dispersal, genetic structure and population dynamics in a highly mobile, but philopatric albatross species. *Molecular Ecology*, *17*(7), 1658–1673.

Mirams, A. G. K., Treml, E. A., Shields, J. L., Liggins, L., & Riginos, C. (2011). Vicariance and dispersal across an intermittent barrier: population genetic structure of marine animals across the Torres Strait land bridge. *Coral Reefs*, *30*, 937–949.

Moore, P. J., Taylor, G. A., & Amey, J. M. (1997). Interbreeding of black-browed albatr*oss Diomedea m. melanophris* and New Zealand black-browed albatross D. *m. impavida* on Campbell Island. *Emu-Austral Ornithology*, 97(4), 322–324.

Moritz, C. (1994). Defining "Evolutionarily Significant Units" for conservation. *Trends in Ecology & Evolution*, *9*(10), 373–375. https://doi.org/10.1016/0169-5347(94)90057-4

Munro, K. J., & Burg, T. M. (2017). A review of historical and contemporary processes affecting population genetic structure of Southern Ocean seabirds. *Emu - Austral Ornithology*, *117*(1), 4–18. https://doi.org/10.1080/01584197.2016.1271988

Prince, P. A., Rothery, P., Croxall, J. P., & Wood, A. G. (1994). Population dynamics of black-browed and grey-headed albatrosses *Diomedea melanophris* and *D. chrysostoma* at Bird Island, South Georgia. *Ibis*, *136*(1), 50–71. https://doi.org/10.1111/J.1474-919X.1994.TB08131.X

Rexer-Huber, K., Veale, A. J., Catry, P., Cherel, Y., Dutoit, L., Foster, Y., McEwan, J. C., Parker, G. C., Phillips, R. A., Ryan, P. G., Stanworth, A. J., van Stijn, T., Thompson, D. A., Waters, J., & Robertson, B. C. (2019). Genomics detects population structure within and between ocean basins in a circumpolar seabird: The white-chinned petrel. *Molecular Ecology*, *28*, 4552–4572.

Roberts, M. A., Schwartz, T. S., & Karl, S. A. (2004). Global population genetic structure and male-mediated gene flow in the green sea turtle (*Chelonia mydas*): analysis of microsatellite loci. *Genetics*, *166*(4), 1857–1870.

Taylor, R. S., Bolton, M., Beard, A., Birt, T., Deane-Coe, P., Raine, A. F., González-Solís, J., Lougheed, S. C., & Friesen, V. L. (2019). Cryptic species and independent origins of allochronic populations within a seabird species complex (*Hydrobates* spp.). *Molecular Phylogenetics and Evolution*, *139*, 106552. https://doi.org/10.1016/J.YMPEV.2019.106552

Thanou, E., Sponza, S., Nelson, E. J., Perry, A., Wanless, S., Daunt, F., & Cavers, S. (2017). Genetic structure in the European endemic seabird, *Phalacrocorax aristotelis*, shaped by a complex interaction of historical and contemporary, physical and nonphysical drivers. *Molecular Ecology*, *26*(10), 2796–2811.

Torres, L., Pante, E., González-Solís, J., Viricel, A., Ribout, C., Zino, F., MacKin, W., Precheur, C., Tourmetz, J., Calabrese, L., Militão, T., Zango, L., Shirahai, H., Bretagnolle, V. (2021). Sea surface temperature, rather than land mass or geographic distance, may drive genetic differentiation in a species complex of highly dispersive seabirds. *Ecology and Evolution* 11, 14960–14976.

Vianna, J. A., Noll, D., Dantas, G. P. M., Petry, M. V., Barbosa, A., González-Acuña, D., Le Bohec, C., Bonadonna, F., & Poulin, E. (2017). Marked phylogeographic structure of gentoo penguin reveals an ongoing diversification process along the Southern Ocean. *Molecular Phylogenetics and Evolution*, *107*, 486–498. https://doi.org/10.1016/J.YMPEV.2016.12.003

Weimerskirch, H., Brothers, N., & Jouventin, P. (1997). Population dynamics of wandering albatross *Diomedea exulans* and Amsterdam albatross *D. amsterdamensis* in the Indian Ocean and their relationships with long-line fisheries: Conservation implications. *Biological Conservation*, *79*(2–3), 257–270. https://doi.org/10.1016/S0006-3207(96)00084-5

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114. https://doi.org/10.1093/GENETICS/28.2.114

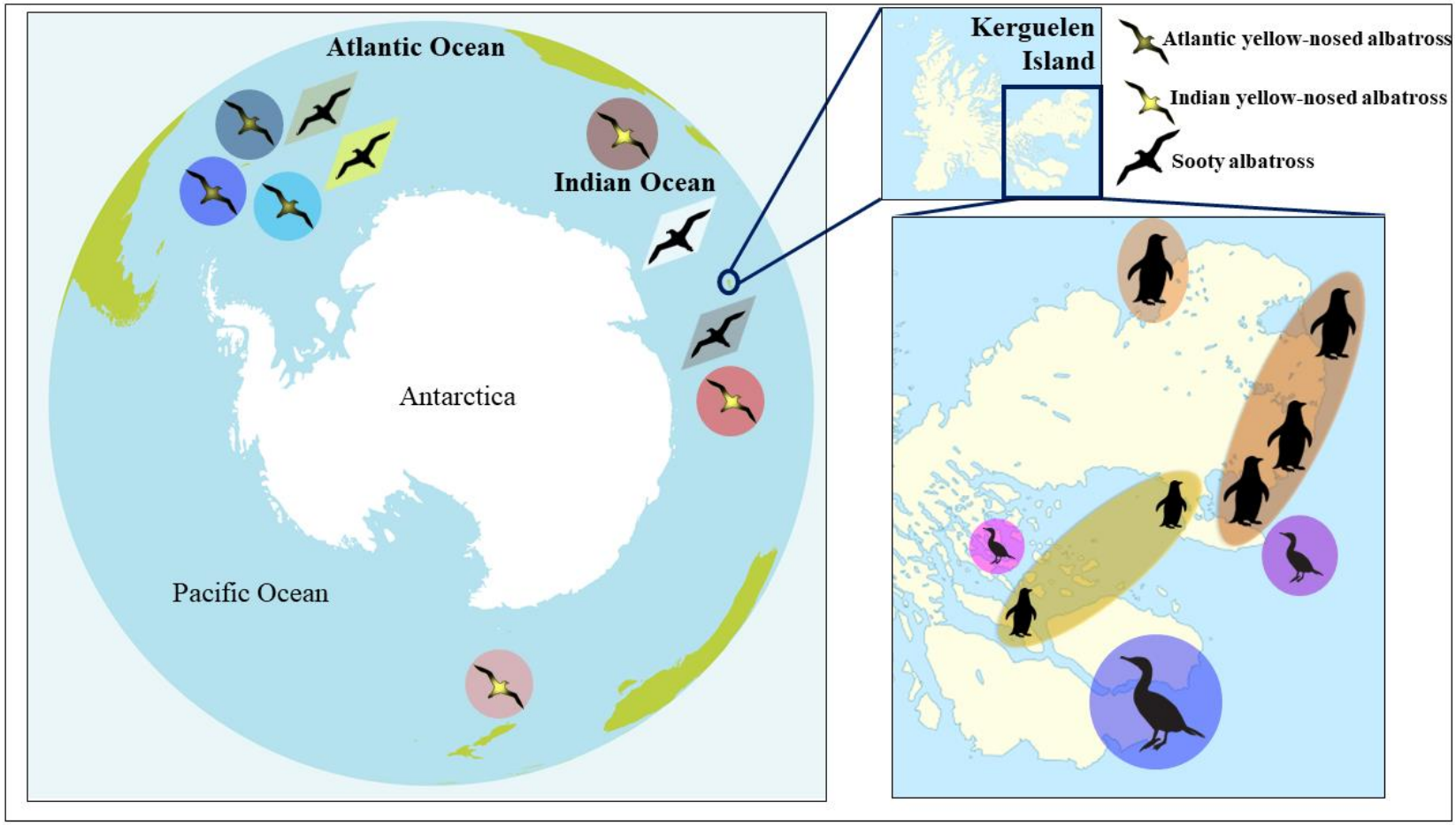

Figure 6.1: Figure summarizing the population genetic structure identified for all five study species. Coloured shapes represent distinct genetic clusters.