PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE AGRONOMÍA E INGENIERÍA FORESTAL
DIRECCIÓN DE INVESTIGACIÓN Y POSTGRADO
MAGÍSTER EN RECURSOS NATURALES

# HIGH GENE FLOW AND LIMITED POPULATION STRUCTURE IN CHINSTRAP PENGUIN (*Pygoscelis antarcticus*) POPULATIONS

Thesis presented as requisite to obtain the degree of

Master in Natural Resources

by:

Isidora Mura Jornet

Thesis Committee:
Advisor:
Dr. Juliana de Abreu Vianna
Examiners:
Dr. Elie Poulin
Dr. Cristián Bonacic

October 2017

Santiago – Chile

**Acknowledgements**

Financial support for this study was provided by Chilean Antarctic Institute (INACH T_27-10, G_06-11, RT_12-14), FONDECYT projects (N°110060 and 1150517), National Council for Scientific and Technological Development – CNPq (n 482501/2013-8; INCT-APA 574018/2008-5) and the Spanish Ministry of Economy and Competitiveness (CGL2004-01348, CGL2007-60369, CTM2011-24427). I would like to express my gratitude to the Department of Ecosystems and Environment of the Pontifical Catholic University of Chile and the scholarship for master awarded by the National Commission for Scientific and Technological Research (CONICYT-PCHA/MagísterNacional/2016 – 22162312), which allowed me to finance the master program.

Many people accompanied me throughout this process, and I would like to thank each one of them:

My advisor Dr Juliana A. Vianna, who has been there for me for almost 5 years, teaching me and guiding me not only in the molecular ecology field, but also in life itself.

My colleagues and friends that I made while working in the laboratory, for their intellectual support and for so many bursts of laughter through all of these years: Bernardita Julio, Camila Barría, Felipe Córdova, Marcelo Fuentes, Máximo González, Ángela Guajardo, María José Frugone, Daly Noll, Bárbara Ramos and Nicole Sallaberry. I would like to give special thanks to Paulina Valenzuela and Rocío Álvarez, who were my mentors during my first steps in science; thanks for believing in me from the beginning.

My family, for their constant love, support and patience throughout the completion of this thesis. I would like to especially thank my mom, who without her support I would not have been able to finish this work.

All of my friends, who without them my life would not be the same. All of them have contributed in some way to starting and/or finishing this investigation: Natalia Jordan, Daniela Baeza, Alejandro Oñate, M. Beatriz Ugalde, Josefa Hojman, Hernán Briones,

Soledad Kremer, Rodrigo Díaz, Jacqueline Calvillo, Ruperto Abarca and Trinidad Quito (who designed the map of the locations sampled in this study).

Thank you so much, to all of you.

*This thesis is dedicated to the tiny humans that I love the most in the world:*

*Sofía Mura,*

*Tomás Mura*

*and the most beautiful penguin of all: Mila Borquez.*

# Index

# High gene flow and limited population structure in chinstrap penguin (*Pygoscelis antarcticus*) populations

Mura-Jornet, I.[1], Dantas, G.P.M.[2], Petry, M. V[3], González-Acuña, D.[4], Barbosa, A[5], Lowther, A. D[6], Kovacs, K. M[6], Poulin, E. [7], Vianna, J.A.[1]

[1]Laboratorio Biodiversidad Molecular, Facultad de Agronomía e Ingeniería Forestal, Pontificia Universidad Católica de Chile. [2]PPG in Biology of Vertebrate, Pontificia Universidade Católica de Minas Gerais. [3]Laboratório de Ornitologia e Animais Marinhos, Universidade do Valle Rio Sinos.[4]Facultad de Ciencias Veterinarias, Universidad de Concepción. [5]Museo Nacional de Ciencias Naturales, Departamento de Ecología Evolutiva, Madrid, España. [6]Norwegian Polar Institute, Tromsø, Norway. [7]Instituto de Ecología y Biodiversidad (IEB), Facultad de Ciencias, Universidad de Chile.

## Abstract

**Mura-Jornet I, Dantas GPM, Petry MV, González-Acuña D, Barbosa A, Lowther AD, Kovacs, KM, Poulin E and Vianna JA**. Thesis, Master in Natural Resources, Faculty of Agronomy and Forestry, Pontifical Catholic University. Santiago, Chile. 36 pp. Historical and contemporary climate change has impacted Antarctic biota affecting species distribution, demography and connectivity. In Antarctic marine environments, seabirds are important bioindicators of ecological change and can respond to climate change in two main ways: local adaptation or dispersal. To characterise connectivity, sex-biased dispersal, diversity and genetic structure 12 microsatellite *loci* and a fragment of CHD1 gene on 12 breeding colonies in the South Shetland Islands and the West Antarctic Peninsula (WAP), and one previously unstudied sub-Antarctic island 3,600 km away from the WAP (Bouvetøya). High genetic diversity and evidence of female bias-dispersal were detected. Limited population genetic structure and lack of isolation by distance throughout the region were found, along with no differentiation between the WAP and Bouvetøya (overall microsatellite $F_{ST}$ = 0.0016*,* p = 0.273), indicating large distance dispersal. Therefore, genetic assignment tests could not assign individuals to their population(s) of origin. The most differentiated location was Georges Point, one of the southernmost breeding colonies of this species in the WAP. The subtle differentiation found may be explained by some combination of low natal philopatric behavior, high rates of dispersal and/or generally high mobility among colonies of chinstrap penguins compared to other *Pygoscelis* species.

**Key words**: Seabirds, Antarctica, genetic diversity, sex-biased, dispersal, population genetics.

# Introduction

Population genetic structure is typically investigated using indirect methods, such as inferring gene flow levels among colonies (Bohonak, 1999). As levels of gene flow increase towards panmixia, the power to statistically detect distinct populations using clustering algorithms decreases (Waples and Gaggiotti, 2006; Latch *et al*., 2006). Additionally, population size has implications for genetic differentiation, as larger populations are more robust to the effects of genetic drift than smaller ones (Taylor and Dizon, 1996). Thus, considering these processes, three broadly different patterns of population genetic structure can be observed: 1) absence of genetic structure and differentiation among populations, 2) significant genetic structure, but no geographic pattern to explain it, or 3) significant genetic and geographic structured populations (Bossart and Prowell, 1998).

In the Southern Ocean, penguins are a major component of avian biomass (Woehler *et al*., 2001) and are dominant predators in the marine ecosystem (Brooke, 2004). Since they are not a directly human target, penguins are considered to be good bioindicators of ecosystem changes (Scheifler *et al*., 2005; Carravieri *et al*., 2013), which makes them an important subject of study. In the South Shetland Islands and the Western Antarctic Peninsula (WAP), three species of *Pygoscelis* penguins breed sympatrically: Adélie (*Pygoscelis adelie*), gentoo (*P. papua*) and chinstrap (*P. antarcticus*) (Borboroglu and Boersma, 2015). Population genetic structure of these penguins has recently been well documented. Microsatellite and mitochondrial data on Adélie penguins have revealed a lack of genetic differentiation between colonies around the Antarctic continent, and a signature of population expansion after the LGM (Roeder *et al*., 2001; Clucas, *et al*., 2014; Younger *et al*., 2016). In contrast, the same genetic markers employed on gentoo penguins have revealed significant population genetic structure in Antarctica, and evidenced of divergent lineages between Antarctica and each sub-Antarctic colony studied (Peña *et al*., 2014; Vianna *et al*., 2016).

Unlike these two species, almost the entire breeding distribution of chinstrap penguins is restricted to the Antarctic Peninsula (to approximately 64° south) and the South Shetland, South Orkney, and South Sandwich Islands in the Scotia Sea region (Woehler, 1993; Borboroglu and Boersma, 2015; Petry *et al*., 2016). Additionally, small breeding populations are described on South Georgia, Bouvetøya, Heard and the Balleny

Islands (Croxall and Kirkwood, 1979; Borboroglu and Boersma, 2015). The non-breeding range of the chinstrap penguin is extensive, with large dispersal being reported. For example, Trivelpiece *et al*., (2007), using satellite telemetry of adult birds, showed that penguins can migrate distances of 800 and 1,300 km from the South Shetland Islands to the South Orkney and South Sandwich Islands, respectively. Biuw *et al*., (2010) described a long-range migration of 3,600 km for a single pre-moulting adult chinstrap penguin from Bouvetøya to the South Sandwich Islands. Although all three *Pygoscelis* species show some degree of natal philopatry, chinstrap penguins are the least philopatric of the genus (Ainley *et al*., 1995; Macdonald *et al*., 2002). At fine geographical scales, this species appears to show weak or even no significant population structure, with no isolation by distance (Korczak-Abshire *et al*., 2012; Clucas *et al*., 2014; Freer *et al*., 2015). No sex-bias has been detected for these birds using microsatellite *loci* (Freer *et al*., 2015), although the authors reported test values consistent with female bias dispersal. Currently, chinstrap penguins are listed as being of Least Concern on the IUCN's Red List of Threatened Species (BirdLife International, 2016). However, there have been reports of continuous declines at nearly all breeding sites of this species (Ciaputa and Sierakowski, 1999; Macdonald *et al*., 2002; Forcada *et al*., 2006; Hinke *et al*., 2007; Sander *et al*., 2007; Trivelpiece *et al*., 2011; Barbosa *et al*., 2012; Korczak-Abshire *et al*., 2012; Lynch *et al*., 2012; Naveen *et al*., 2012; Dunn *et al*., 2016; Niemandt *et al*., 2016).

Even though population genetic structure of chinstrap penguins has been studied in some parts of its range (Korczak-Abshire *et al*., 2012; Clucas *et al*., 2014; Freer *et al*., 2015), this species remains the least studied of the *Pygoscelis* penguins, and the connectivity between breeding colonies in the WAP, or between the WAP and the eastern-most breeding colony in the species distribution (Bouvetøya), is unknown. The population declines reported in numerous colonies highlight the importance of investigating the connectivity of breeding colonies in terms of source and sink population dynamics and other genetic effects these reductions might have. Indeed, in the context of conservation biology, the proper identification of population genetic structure is crucial (Palsbøll *et al*., 2006). Therefore, we used 12 nuclear markers specifically designed for *Pygoscelis* penguins and a fragment of CHD1 gene on adult breeding chinstrap penguins from 13 different colonies to: (1) describe patterns of distribution of genetic diversity and population structure, (2) quantify levels of connectivity among colonies in the WAP and

the easternmost limit of the species distribution at Bouvetøya and (3) evaluate levels of sex-biased dispersal.

## Materials and methods

### Field sampling and DNA extraction

Between 2009 and 2016 (exception: Miers Bluff from 2003), a total of 251 blood samples from chinstrap penguins were collected at 13 sites, including 10 locations in the South Shetland Islands (n = 183), two in the Antarctic Peninsula (n = 45) and Bouvetøya (n = 23) (Fig.1 and Table 1). To avoid disturbance within the breeding colonies, adult penguins were captured using hand-held nets when they were walking into the water, and all procedures were done following a method for restraining penguins (Wilson, 1997). Up to 1 mL of blood samples were obtained from the brachial or foot vein using a 23G needle and stored in 96% ethanol.

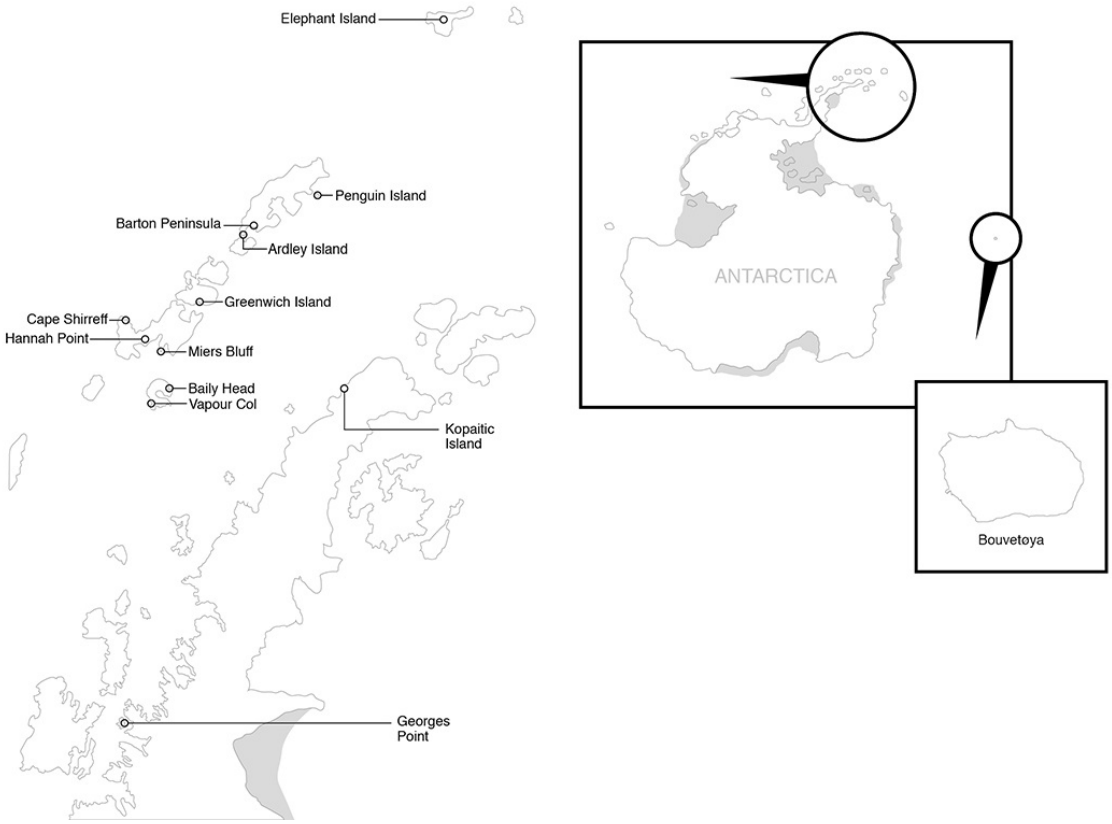


**Fig. 1** Chinstrap penguin sampled sites during this study (total n = 251).

All samples from the WAP were collected with permits in accordance to Annex II, Article 3 of the Protocol on Environmental Protection to the Antarctic Treaty, and under the regulation of the Scientific Committee on Antarctic Research (SCAR) provided by the Chilean Antarctic Authorities (INACH), the Brazilian Antarctic Authorities through the PROANTAR and Environmental Ministry, and the Spanish Polar Committee. The permits incorporated authorization for sample collection for all locations including special authorization to the studied protected areas. Bioethics permits were provided by Universidad de Concepción and Pontificia Universidad Católica de Chile, which were required to obtain INACH permits. Samples from Bouvetøya were collected under Animal Ethics Permit #7001 issued by the Norwegian Food Safety Authority.

**Table 1.** Summary of the chinstrap penguin samples used in the analyses. Information of localities (with abbreviation), geographic regions (SAI: Sub-Antarctic Island; WAP: West Antarctic Peninsula; SSI: South Shetland Island; AP: Antarctic Peninsula), coordinates and the number of samples used from each locality.

| Colony name (abbreviation) | Area | Coordinates | | Sample size |
|---|---|---|---|---|
| | | | | Microsatellite (n=251) |
| Elephant Island (EI) | WAP - SSI | 61°10’S | 55°00'W | 17 |
| Penguin Island (PI) | WAP - SSI | 62°06’S | 57°56’W | 19 |
| Barton Peninsula (BP) | WAP- SSI | 62°14’S | 58°46’W | 29 |
| Ardley Island (AI) | WAP-SSI | 62°13’S | 58°56’W | 14 |
| Greenwich Island (GI) | WAP-SSI | 62°31’S | 59°47’W | 14 |
| Miers Bluff (MB) | WAP- SSI | 62°43’S | 60°26’W | 11 |
| Hannah Point (HP) | WAP- SSI | 62°39’S | 60°52’W | 25 |
| Cape Shirreff (CS) | WAP- SSI | 62°28’S | 60°48’W | 30 |
| Baily Head (BH) | WAP- SSI | 62°58’S | 60°30’W | 9 |
| Vapour Col (VC) | WAP- SSI | 63°00’S | 60°44’W | 15 |
| Kopaitic Island (KI) | WAP - AP | 63°19’S | 57°55'W | 30 |
| Georges Point (GP) | WAP - AP | 64°40’S | 62°39’W | 15 |
| Bouvetøya (BI) | SAI | 54°26’S | 3°23’E | 23 |

Total genomic DNA was extracted using a salt protocol (Aljanabi and Martinez, 1997) modified as follows: a buffer based on TNE 1X, Tris-HCl and SDS 25% in place of Tris-HCl, EDTA and SDAD. Additionally, 10 M ammonium acetate was used instead of NaCl and tubes spun down for 20 min at 14, 000 rpm. After the extraction, DNA samples were stored in Low TE buffer at -20°C or -80°C.

**Amplification and genotyping protocols**

Genetic diversity and population differentiation were examined at 12 tetranucleotide microsatellite *loci* (AP-3, AP-19, AP-26, AP-61, AP-78, AP-85, AP-90, CP-6, CP-25, GP-6, GP-15 and GP-36) isolated from the genome of three species of *Pygoscelis* penguin sequenced by NGS (Next Generation Sequencing) as part of another study (Vianna *et al*., 2017). Forward primers were synthesized using 5'-end-M13 tail-labelled fluorophores with one of three dyes (6-FAM, HEX, or NED; Applied Biosystems) to suit simultaneous genotyping at multiple *loci* with overlapping size ranges. The primer's sequences, PCR conditions and amplification cycles for microsatellite *loci* were followed by Vianna *et al*., (2017). DNA samples were separated by electrophoresis through a 2% agarose gel, run for 0.5 h at 300 V. The resultant PCR products were sent for genotyping at Macrogen Inc. (Seoul, South Korea). All PCR reactions were conducted on an Applied Biosystem machine and the mixtures contained 10-100 ng of genomic DNA. The microsatellite genotypes were assigned using the software program GeneMarker® v.1.75 (Softgenetics LLC™) for allele size identification.

**Gene diversity**

For all microsatellite data, PGDSpider v.2.1.0.1 software was used as an automated data conversion tool (Lischer and Excoffier, 2012). For these data set, the presence of null alleles or potential genotyping errors were evaluated using Micro-Checker v.2.2.3 (Van Oosterhout *et al*., 2004). Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010) was used to study genetic diversity within samples from each chinstrap penguin colony, calculate the mean number of alleles per *locus*, and to evaluate the observed (Ho) and expected heterozygosities (He). Expectations for Hardy-Weinberg equilibrium

(HWE) were estimated as deviation of the Wright's $F_{IS}$ index and these were tested for each *locus*, for all *loci*, and for each population using randomization procedures using 10,000 permutations with GENETIX v.4.05.2 (Belkhir *et al*., 2004). To test the presence of linkage disequilibrium, the same program was used with a likelihood-ratio test and the empirical distribution generated by 10,000 permutations. Corrections for multiple testing were made using the False Discovery Rate (FDR, Benjamini and Hochberg, 1995; Pike, 2011).

**Population genetic structure and isolation by distance**

Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010) was used to calculate $F_{ST}$ between pairwise populations on microsatellite data using 10,000 permutations. P values were corrected with the FDR method for multiple tests (Benjamini and Hochberg, 1995; Pike, 2011). Employing microsatellite *loci*, isolation by distance (IBD) was evaluated using the adegenet package in R (Jombart, 2008). For this, adegenet uses a Mantel test between a matrix of genetic distances and a matrix of geographic distances (Jombart, 2015). Google Earth (Google, v.7.1.8.3036) was used to calculate the shortest geographical distance by sea between locations.

To determine the most likely number of clusters (*K*), multilocus genotypes were analyzed using Bayesian clustering methods implemented in STRUCTURE v2.3.4 (Pritchard *et al*., 2000), BAPS v.6.0 (Corander *et al*., 2008) and GENELAND v.3.1 (Guillot *et al*., 2005). First, the software STRUCTURE v2.3.4 was run using different models assuming (ad)mixture, (un)correlated allele frequencies both with and without *a priori* specification of sample locations (Falush *et al*., 2003; Pritchard *et al*., 2000). The models were run with the likely number of populations (*K*) set from 1 to 13. For each *K*, the model was run 10 times with a burn-in length of 100,000 iterations followed by 1,000,000 Markov Chain Monte Carlo (MCMC) subsequent iterations. The optimum number of clusters was inferred by deriving the posterior probability of *K* (LnP(D)) from each independent run. As the Δ*K* method of Evanno's does not allow $K < 2$ to be tested (Evanno *et al*., 2005), this method was employed when *K* was higher than one for log-likelihood using STRUCTURE HARVESTER (Earl and VonHoldt, 2012). To align multiple replicates of STRUCTURE result files, CLUMPP v1.1.2 (CLUster Matching and Permutation Program; Jakobsson

and Rosenberg, 2007) was used. The results generated by the genetic clustering program were visualized using DISTRUCT v1.1 (Rosenberg, 2004).

Second, a Bayesian Analysis of Population Structure (BAPS v.6.0) was performed using a combination of analytical and stochastic methods, based on molecular markers and geographical sampling (Corander *et al*., 2008). Calculations were performed over 10,000 iterations with both spatial and non-spatial, and both a mixture and an admixture model, with the maximum number of populations possible set to 13.

Third, an analysis of spatial structure using the R package GENELAND v3.1 was carried out to determine the most likely number of populations and to assign individuals to population clusters. This program is based on an algorithm which includes not only genotypes, but also the geographic location information of all individuals to estimate the number of groups and delineate their spatial boundaries (Guillot *et al*., 2008). Analyses were performed under the spatial model assuming both correlated and uncorrelated allele frequency. The correlated frequency model might be more capable of detecting subtle differentiations, but it also could be more sensitive to departure from model assumptions (as presence of isolation-by-distances), and more prone to algorithm instabilities than the uncorrelated frequency model (Guillot *et al*., 2008). Ten independent MCMC simulations were run allowing the number of populations to vary between 1 and 13, with the following parameters: 1,000,000 MCMC iterations with a thinning of 100, a maximum rate of Poisson processes fixed to 500 and a maximum number of nuclei in the Poisson-Voroni tessellation fixed to 300. The best-supported *K* value was determined based on the highest averaged maximum likelihood score of the models.

Also, a Discriminant Analysis of Principal Components (DAPC) was carried out to determine the number of clusters of genetically related individuals, using a non-Bayesian approach. DAPC uses sequential *K*-means and model selection to identify genetic clusters (Jombart *et al.,* 2010*)*. For this, the adegenet package in R (Jombart, 2008) was used, retaining all principal components.

To assign or exclude the reference populations as being the origins of individuals based on genotype data, assignment testing of microsatellite *loci* was done using GENECLASS2 v.2.0.h (Piry *et al*., 2004). To do this, two separate analyses were performed: one employed the likelihood method based on allele frequencies (Paetkau *et al.*, 1995), and the other used the Bayesian method approach (Rannala and Mountain,

1997). The probability that each individual was assigned to a candidate population was estimated using a Monte Carlo resampling method (Paetkau *et al*., 2004; number of simulated individuals = 10,000; type I error = 0.01). The same program and parameters were also used for the detection of first-generation migrants.

**Sex determination and sex-biased dispersal**

For molecular sex identification, a region of the Chromosome-helicase-DNA binding protein (CHD1) gene was amplified using PCR, with primer pair 2550F/2718R (Fridolfsson and Ellegren, 1999). PCR reactions were carried out in 25 $\mu$L volume containing 10-100 ng genomic DNA, 1X reaction buffer, 0.5 $\mu$M of each primer, 1.5mM $MgCl_2$, 100 $\mu$M dNTPs and 0.7 U *Taq* DNA polymerase (Invitrogen Life Technologies). The reactions' conditions were as follows: an initial denaturing step at 94°C for 5 min; followed by 45 cycles at 94°C for 30 s, 46°C for 45 s, and 72°C for 25 s; and a final extension step at 72°C for 5 min. All PCR reactions were conducted on an Applied Biosystem machine. The amplification products were electrophoresed through a 2% agarose gel for approximately 1 h at 150 V and visualized with GelRed® under UV light.

Using microsatellite data, sex-biased dispersal was evaluated with FSTAT v.2.9.3.2 (Goudet, 2001). For this, a one-tailed test was done assuming males as the most philopatric group, since dispersal is female-biased in most birds (Greenwood, 1980). Differences in the inbreeding coefficient ($F_{IS}$), fixation index ($F_{ST}$), relatedness between individuals (*r*), mean Assignment Index (*mAIc*) and variance of Assignment Indices (*vAIc*) between sexes were calculated. $F_{ST}$, *r* and *mAIc* are expected to be lower in the sex that disperses most, whereas $F_{IS}$ and *vAIc* should be higher (Goudet *et al*., 2002). The p values of each test were estimated using 10,000 randomizations.

# Results

## Genetic diversity

For microsatellite data, only one *locus* was monomorphic (GP-6), so it was not used in further analyses. The remaining 11 microsatellite *loci* were polymorphic for all populations, except for AP-3 in Baily Head (BH) (Table S1). The inbreeding coefficient ($F_{IS}$) was low in all populations, and no significant p-values were found, indicating no significant heterozygote excess or deficiencies. Therefore, departures from Hardy-Weinberg equilibrium were not detected on populations at the 11 *loci* (Table 2). Linkage disequilibrium of each pair of *loci* were not detected within or among populations. For microsatellite markers, overall allele numbers *per locus* varied between three (*locus* AP-3) and 12 (*locus* GP-15), with an average of 5.69 alleles over all sample sites. The expected heterozygosity ranged from 0.62 (Elephant Island, Barton Peninsula and Miers Bluff) to 0.71 (BH), with an average of 0.65. The observed heterozygosity exhibits a similar level of variation with an average of 0.63 over all locations. Values of allelic richness ranged between 5.00 to 6.46 per sample site. *Locus*-by-*locus* allelic richness and diversity measures for each sample location are shown in the Appendix (Table S1).

**Table 2.** Genetic diversity indices from 11 microsatellite *loci* by localities. Short forms refer to: N: sample size; A: mean number of alleles per locus; *Ho*: mean observed heterozygosity; *He*: mean expected heterozygosity and $F_{IS}$: inbreeding coefficient.

| Colony | N | A | Ho±SD | He±SD | $F_{IS}$ |
|---|---|---|---|---|---|
| EI | 17 | 5.55±1.97 | 0.62±0.25 | 0.62±0.23 | 0.003 |
| PI | 19 | 5.91±1.92 | 0.70±0.24 | 0.67±0.22 | -0.053 |
| BP | 29 | 6.18±2.48 | 0.57±0.28 | 0.62±0.27 | 0.081 |
| AI | 14 | 5.27±2.05 | 0.60±0.26 | 0.66±0.22 | 0.096 |
| GI | 14 | 5.18±1.89 | 0.61±0.28 | 0.64±0.25 | 0.048 |
| MB | 11 | 5.00±2.28 | 0.61±0.29 | 0.62±0.28 | 0.019 |
| HP | 25 | 5.91±2.47 | 0.64±0.23 | 0.64±0.24 | 0.001 |
| CS | 30 | 6.46±2.73 | 0.61±0.27 | 0.64±0.25 | 0.039 |
| BH | 9 | 5.10±1.29 | 0.70±0.17 | 0.71±0.16 | 0.011 |
| VC | 15 | 5.46±2.02 | 0.67±0.24 | 0.65±0.24 | -0.025 |
| KI | 30 | 6.36±2.50 | 0.63±0.27 | 0.64±0.25 | 0.028 |
| GP | 15 | 5.82±2.27 | 0.64±0.26 | 0.64±0.26 | -0.003 |
| BI | 23 | 5.73±2.01 | 0.65±0.24 | 0.65±0.20 | -0.001 |
| Total | 251 | 5.69±2.14 | 0.63±0.24 | 0.65±0.24 | 0.003 |

**Population genetic structure and isolation by distance**

Notably, there was an absence of population genetics structure between WAP and Bouvetøya for microsatellite *loci* ($F_{ST}$ = 0.0016, p = 0.273). The $F_{ST}$ values were generally not significant (Fig. 2, Table S2). Seven of 78 pairwise $F_{ST}$ comparisons were significantly different, all corresponding to the southernmost locality of this study: Georges Point (GP). However, significant $F_{ST}$ values were small (0.031 - 0.054), indicating only weak differentiation between GP and other colonies (Table S2). GP $F_{ST}$ values differed significantly from seven of the 12 northernmost studied sites (Elephant Island, Penguin Island, Barton Peninsula, Ardley Island, Cape Shirreff, Kopaitic Island and Bouvetøya, Fig.1). Mantel's testing did not detect isolation by distance in microsatellite data (r = 0.048, p = 0.401).

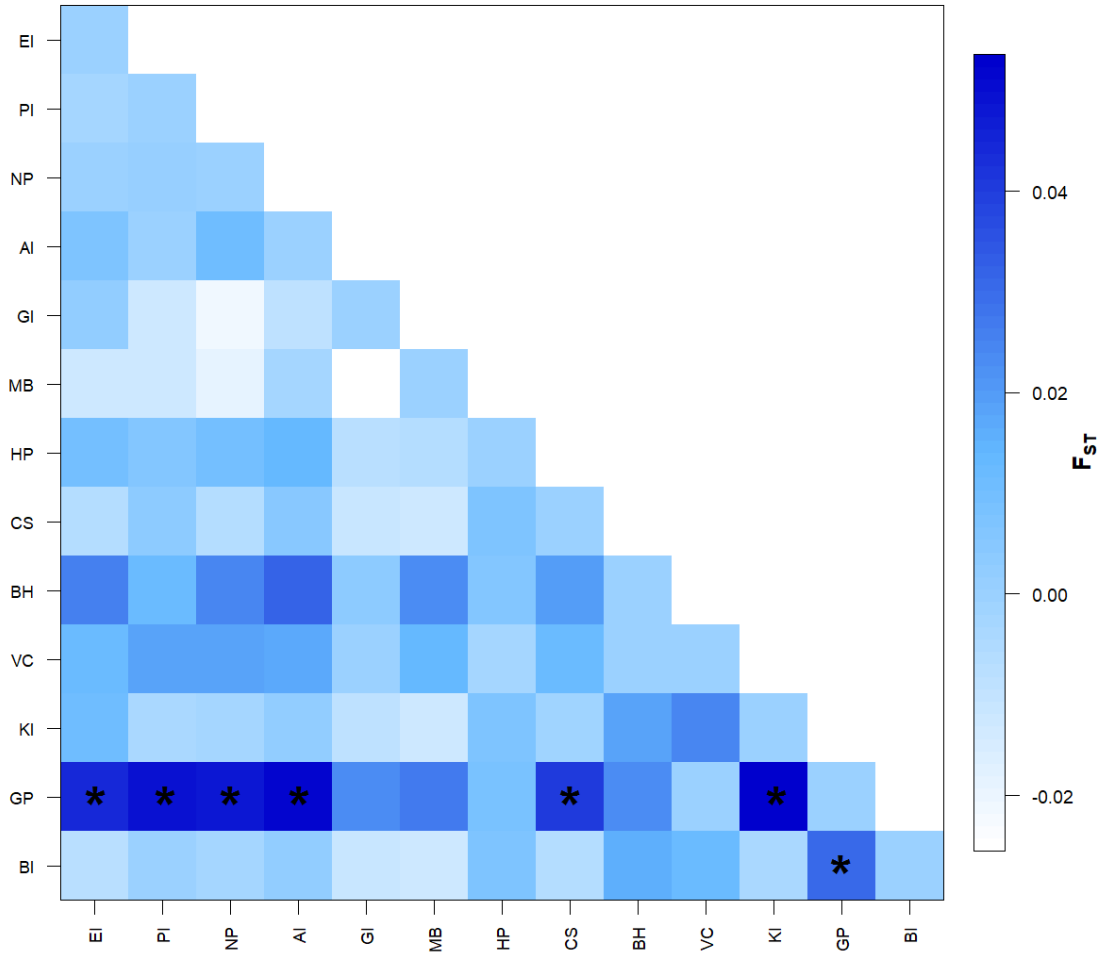


**Fig. 2.** Distance matrix showing the population differentiation. Each cell of the heat plot is color coded, illustrating relative differences. Blue shades indicate higher levels of genetic differences and white lower. The asterisks indicate significant $F_{ST}$ values. Short forms refer to EI: Elephant Island; PI: Penguin Island; BP: Barton Peninsula; AI: Ardley Island; GI: Greenwich Island; MB: Miers Bluff; HP: Hannah Point; CS: Cape Shirreff; BH: Baily Head; VC: Vapour Col; KI: Kopaitic Island; GP: Georges Point and BI: Bouvetøya.

To identify the number of populations among the 13 locations, four approaches were used, yielding different optimal numbers of clusters. First, using the mean log-likelihood in STRUCTURE, the analysis inferred that the number of populations (*K*) was one, for seven of the eight different model assumptions tested. Only one model selected *K* = 3, therefore it was also evaluated with Evanno's method, which suggested *K* = 2. Nonetheless, when looking at the individual assignment plots, no group could be identified (Fig. S1). Second, for BAPS analysis, when the spatial model was used, the inferred number of populations was *K* = 1 despite whether mixture or admixture models were performed. Conversely, when a non-spatial model was run, the optimal number of clusters was *K* = 7 but with no geographical relation (Fig. S2). Third, for GENELAND, the number of groups estimated varied depending on whether the uncorrelated or correlated frequency model was used, although in both models the 10 runs consistency converged on a single *K* value. When employing the uncorrelated model, the inferred number of populations *K* was one. Contrasting, GENELAND estimated *K* = 3 clusters for chinstrap penguins when employing the correlated allele frequency model. These clusters corresponded to three distinct populations: (1) Kopaitic Island, (2) Georges Point and (3) northern WAP locations and Bouvetøya (Fig. S3). Like pairwise $F_{ST}$ comparison, GENELAND also identified Georges Point as the most differentiated breeding colony, however, the probabilities of cluster membership were very low (<0.5). Models and estimated number of populations (*K*) for all Bayesian programs used are summarized in Table 3. The final approach, DAPC, estimated the optimal number of clusters to *K* = 6, however, they were geographically meaningless and overlapped extensively (Fig. S4). Although some analyses suggest clusters larger than one, the graphic results do not show any consistent group. Thus, only one panmictic population is estimated for chinstrap penguins.

Finally, for estimating dispersal patterns, assignment tests were only successful for assigning 13.9% of the individuals (assignment threshold of 0.05) to the proper colony, and low values were again observed for the first-migrant generation, revealing high gene flow among all sampled colonies (Table S3).

**Table 3.** Summary of the Bayesian clustering analyses and different models used to infer the optimal number of clusters (*K*).

| Bayesian clustering software | Model use | Inferred number of cluster (*K*) |
|---|---|---|
| BAPS | Spatial, with mixture model | *K* = 1 |
| | Spatial, with admixture model | *K* = 1 |
| | Non-spatial model, with admixture model | *K* = 7 |
| GENELAND | Spatial model, with uncorrelated allele frequency | *K* = 1 |
| | Spatial model, with correlated allele frequency | *K* = 3 |
| STRUCTURE | Admixture, with correlated allele frequency, using location information | *K* = 1 |
| | Admixture, with independent allele frequency, using location information | *K* = 1 |
| | Admixture, with correlated allele frequency, no location information supplied | *K* = 1 |
| | Admixture, with independent allele frequency, no location information supplied | *K* = 1 |
| | No admixture, with correlated allele frequency, using location information | *K* = 1 |
| | No admixture, with independent allele frequency, using location information | *K* =1 |
| | No admixture, with correlated allele frequency, no location information supplied | *K* =1 |
| | No admixture, with independent allele frequency, no location information supplied | *K* = 3, K =2 * |

* Inferred number of cluster using Evanno's method.

**Sex-bias dispersal**

It was possible to determine the sex of 196 individuals: 93 females and 103 males. Three test values indicated a female bias dispersal (higher $F_{ST}$, relatedness and mean assignment index values in males). Nonetheless, only the mean assignment index test

revealed a significant difference between males and females, indicating females as the dispersing sex in chinstrap penguins (Table 4).

**Table 4.** Test results for sex-biased dispersal in chinstrap penguins, their corresponding p-values, and the number (N) of females and males used for the analyses.

| | | | | Assignment indices | | |
|---|---|---|---|---|---|---|
| | N | $F_{IS}$ | $F_{ST}$ | Relatedness | Mean | Variance |
| Females | 93 | 0.0016 | 0.0016 | 0.1452 | -0.5088 | 9.5572 |
| Males | 103 | 0.0418 | 0.0969 | 0.1708 | 0.4594 | 10.428 |
| p-value | | 0.1107 | 0.2912 | 0.3966 | **0.0286** | 0.6984 |

Significant results ($p < 0.05$) are given in bold type.

## Discussion

Chinstrap penguins throughout the 13 colonies studied herein, showed high levels of genetic diversity, low levels of genetic structure between study sites with no isolation by distance and evidence of female bias-dispersal.

Although chinstrap penguin populations have decreased dramatically over the last four decades, high genetic diversity was found in all studied colonies. This could be the result of a historically large population size or simply a result of the currently large population (7,5 million pairs) (Woehler, 1993), in combination with high levels of gene flow between colonies. This study's finding of high genetic diversity is consistent with previous reports for chinstrap penguin colonies from the WAP, South Orkney and South Sandwich Islands using microsatellite and mtDNA data (Freer *et al*., 2015; Clucas *et al*., 2014). High genetic diversity also has been documented using mtDNA from gentoo penguins (Dantas *et al*., 2014; Peña *et al*., 2014; Clucas *et al*., 2014; Vianna *et al*., 2016) and for Adélie penguins using microsatellite markers (Roeder *et al*., 2001). Additionally, other penguins also have demonstrated high genetic diversity, such as rockhopper penguins (Jouventin *et al*., 2006), magellanic penguins (Bouzat *et al*., 2009) and Humboldt penguins (Schlosser *et al*., 2009).

In this current study, most clustering data analyses suggests only one genetic group for chinstrap penguin. The lack of consensus achieved for a few microsatellite analyses (Table 3) could be explained because it is common that the accuracy of Bayesian analyses diminish when levels of genetic differentiation among population decrease, performing better with $F_{ST} > 0.05$ (Latch *et al*., 2006). However, using microsatellite *loci*, it is clear that there is reduced or no population structure among chinstrap penguins breeding in the WAP, but, also and interestingly, there is an absence of structure between the WAP and Bouvetøya (Fig. S1 and S2, and Table S2). These results complement previous genetic investigations that have found little (if any) population structure between chinstrap penguins in Antarctica. For example, weak genetic differences and high level of gene flow between two colonies from South Shetland Islands were also found using amplified fragment length polymorphism (AFLP) analyses, but within a short distance of 32 km (Korczak-Abshire *et al*., 2012). Using a fragment of the hypervariable region 1 (HVR1) on four breeding sites, weak differentiation between colonies of chinstrap penguins from the WAP, South Shetland and South Orkney Islands from the South Sandwich Islands were found (Clucas *et al*., 2014).

Also, limited genetic variation was found among colonies from the WAP and archipelagos within the Scotia Arc using microsatellite *loci*. Nonetheless, limited number of breeding colonies (only two in WAP, and two in Scotia Arc) were studied (Freer *et al*., 2015). The limited genetic structure found in this study is likely a result of recurrent and long-distance migration of individuals between sample sites, supported by the inability of assignment tests to successfully place individuals to their correct population of origin. The slight levels of genetic differentiation reported between Georges Point and the northernmost locations studied coincide with one of the southernmost distribution of chinstrap penguins in the Antarctic Peninsula. Thus, the incipient differentiation may be explained by a founder effect from the northernmost colonies to the southern. Although the source colony is expected to present higher genetic diversity values than the new ones (Bensch and Hasselquist, 1999), the genetic diversity indices found here were similar at all sample sites. A similar pattern to this was observed in the trumpeter finch (*Bucanetes githagineus*) at peripheral populations (Barrientos *et al*., 2008). Another observed patter that could support the idea of chinstrap penguin colonizing new breeding habitats is that they are currently expanding their range southward in the Antarctic Peninsula (Forcada and Trathan, 2009). Numerous studies have reported the presence of small number of chinstrap penguins south of their normal breeding range (Ainley *et al*., 1978; Raymond, 1975; Spurr, 1985). During field work conducted in January 2017, also various dispersed observations of chinstraps south of their normal breeding range were made: two breeding pairs on Waterboat point (Gabriel González Videla base; 64°49'S, 62°51'W), a single individual surrounded by gentoo colonies on Doumer Island (Yelcho base; 64°65'S, 63°35'W) and another single bird on Avian Island (67°46'S, 68°54'W) surrounded by Adélie penguins (Fig. S5). These may suggest that chinstraps tend to prospect other colonies and breeding habitats far away for their colony of origin, similar to king penguins (Clucas *et al*., 2016).

Dispersal has significant effects on the population size (growth or reduction), species persistence and genetics of species (Lowe and Allendorf, 2010). The predisposition of an individual to return to its natal colony over its reproductive lifetime is known as philopatry (Greenwood, 1980). In most birds, females have a greater tendency to leave their natal groups and disperse larger distances than males (Greenwood, 1980; Pusey, 1987). The first study which compared connectivity between male and female with genetic tools in chinstrap penguins, reported several value tests (females with higher $F_{IS}$, negative *mAIc* and higher *vAIC*) pointing to a female-bias dispersal, though none of the indexes were significant (Freer

*et al*., 2015). In the current study, the data suggest that females are the dispersing sex and males are the philopatric sex. Some studies suggest that penguins are not always philopatric (Ainley *et al*. 1995; Dugger *et al*., 2010). Natal philopatry suggest that individuals are likely to have low rates of movement between colonies (Roeder *et al*., 2001). However, only a proportion of all the individuals are faithful to one locality (Greenwood, 1980), and a small number of migrants could homogenize population structure easily (Wright, 1969). Indeed, Adélie penguins, which exhibit strong natal philopatry, do not show strong genetic difference among colonies, potentially due to an interaction between large effective population size in combination with some dispersal (Roeder *et al*., 2001). Evolutionary reasonings for sex-biased dispersal are inbreeding avoidance and evasion of intersexual competition (Pusey, 1987). Although philopatry has several benefits, such as the development of antipredator strategies, social facilitation and spatial heterogeneity of breeding and foraging habitats (Ainley *et al*., 1995), stressful environmental conditions (such as extensive sea ice or obstruction to unusual migration patterns) may be driving an increase in dispersion rates, leading penguin species to have less philopatric behaviour than previously thought (Dugger *et al*., 2010).

## Conclusions

This study has revealed minimal population structure for chinstrap penguin among breeding colonies in the WAP, but also between the WAP and a sub-Antarctic island 3,600 km away (Bouvetøya). Georges Point, one of the southernmost breeding colonies of chinstrap penguin in the Antarctic Peninsula, was the most differentiated of all. The observed genetic structure in chinstrap penguins may be due to different factors, such as a historical large population size making it robust to drift, long-range gene flow leading to panmixia between breeding colonies, stressful environmental conditions forcing penguins to increase dispersion rates or post-LGM recolonization between WAP and Bouvetøya.

## References


Ainley, D.G., Nur, N. and Woehler, E.J. 1995. Factors affecting the distribution and size of Pygoscelid penguin colonies in the Antarctic. American Ornithological Society 112(1):171–182.

Ainley, D.G., Wood, R.C. and Sladen, W.J.L. 1978. Bird Life at Cape Crozier, Ross Island. Wilson Bull 90(4):492–510.

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

Barbosa, A., Benzal, J., De León, A., and Moreno, J. 2012. Population decline of chinstrap penguins (*Pygoscelis antarctica*) on Deception Island, South Shetlands, Antarctica. Polar Biology 35(9):1453-1457.

Barrientos R., Kvist, L., Barbosa, A., Valera, F., López-Iborra G.M. and Moreno, E. 2009. Colonization patterns and genetic structure of peripheral populations of the trumpeter finch (*Bucanetes githagineus*) from Northwest Africa, the Canary Islands and the Iberian Peninsula. Journal of Biogeography *36*(2):210-219.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Catch, F. 2004. GENETIX 4.0. 5.2., Software under Windows™ for the genetics of the populations. University of Montpellier, Montpellier, France.

Bensch, S. and Hasselquist, D. 1999. Phylogeographic population structure of great reed warblers: an analysis of mtDNA control region sequences. Biological Journal of the Linnean Society *66*(2):171-185

Benjamini, Y. and Hochberg, Y. 1995. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 57(1):289–300.

BirdLife International. 2016. Pygoscelis antarcticus. The IUCN Red List of Threatened Species 2016: e.T22697761A93638235. Available at: http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22697761A93638235.en.

Biuw, M., Lydersen, C., De Bruyn, P.N., Arriola, A., Hofmeyr, G.G., Kritzinger, P. and Kovacs, K. M. 2010. Long-range migration of a chinstrap penguin from Bouvetøya to Montagu Island, South Sandwich Islands. Antarctic Science 22(2):157-162.

Bohonak, A.J. 1999. Dispersal, gene flow, and population structure. The Quarterly Review of Biology 74(1): 21-45.

Bossart, J.L. and Prowell, D.P. 1998. Genetic estimates of population structure and gene flow: limitations, lessons and new directions. Trends in Ecology and Evolution 13(5):202–206.

Borboroglu, P.G. and Boersma, P.D. (Eds.). 2015. Penguins: natural history and conservation. University of Washington Press. 360 pp.

Bouzat, J.L., Walker, B.G. and Dee Boersma, P. 2009. Regional genetic structure in the Magellanic penguin (*Spheniscus magellanicus*) suggests metapopulation dynamics. The American Ornithologist's Union 126(2):326–334.

Brooke, M. de L. 2004. The food consumption of the world's seabirds. Proceedings of the Royal Society of London B: Biological Sciences 271(Suppl_4):S246–S248.

Carravieri, A., Bustamante, P., Curlaud, C. and Cherel, Y. 2013. Penguins as bioindicators of mercury contamination in the Southern Ocean: birds from the Kerguelen Islands as a case study. Science of the Total Environment 454-455:141–148.

Ciaputa, P. and Sierakowski, K. 1999. Long-term population changes of Adélie, chinstrap, and gentoo penguins in the regions of SSSI No. 8 and SSSI No. 34, King George Island, Antarctica. Polish Polar Research 20(4):355–365.

Clucas, G.V., Younger, J.L., Kao, D., Rogers, A.D., Handley, J., Miller, G.D., Jouventin, P., Nolan, P., Gharbi, K., Miller, K.J. and Hart, T. 2016. Dispersal in the sub-Antarctic: king penguins show remarkably little population genetic differentiation across their range. BMC Evolutionary Biology 16(211):1–14.

Clucas, G.V., Dunn, M.J., Dyke, G., Emslie, S.D., Levy, H., Naveen, R., Polito, M.J., Pybus, O.G., Rogers, A.D. and Hart, T. 2014. A reversal of fortunes: climate change "winners" and "losers" in Antarctic Peninsula penguins. Scientific Reports 4:1–7.

Corander, J., Marttinen, P. and Tang, J. 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9(539):1–14.

Croxall, J.P. and Kirkwood, E.D. 1979. The distribution of penguins on the Antarctic Peninsula and islands of the Scotia Sea. Balogh Scientific Books.186 pp.

Dantas, G.P.M., de Oliveira, L. R., Marasco, A.C.M., de Araujo, J., Hurtado, R., Durigon, E. L., San Fillipo, L.F. and Morgante, J.S. 2014. Demographic history of the gentoo penguin (*Pygoscelis papua*) and the Adélie penguin (*Pygoscelis adeliae*) on Admiralty Bay, King George Island, Antarctica. Waterbirds 37(4), pp.410–418.

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

Dunn, M. J., Jackson, J. A., Adlard, S., Lynnes, A. S., Briggs, D. R., Fox, D., and Waluda, C. M. 2016. Population size and decadal trends of three penguin species nesting at Signy Island, South Orkney Islands. PloS one 11(10): pp.e0164025.

Earl, D.A. and 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.

Evanno, F., Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14(8):2611–2620.

Excoffier, L. and Lischer, H.E. 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.

Falush, D., Stephens, M. and Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked *loci* and correlated allele frequencies. Genetics Society of America 164(4):1567–1587.

Forcada, J., Trathan, P.N., Reid, K., Murphy, E.J. and Croxall, J.P. 2006. Contrasting population changes in sympatric penguin species in association with climate warming. Global Change Biology 12(3):411–423.

Forcada, J. and Trathan, P.N. 2009. Penguin responses to climate change in the Southern Ocean. Global Change Biology 15(7):1618–1630.

Freer, J.J., Mable, B.K., Clucas, G., Rogers, A.D., Polito, M.J., Dunn, M., Naveen, R., Levy, H. and Hart, T. 2015. Limited genetic differentiation among chinstrap penguin (*Pygoscelis antarctica*) colonies in the Scotia Arc and Western Antarctic Peninsula. Polar Biology 38(9):1493–1502.

Fridolfsson, A.K. and Ellegren, H. 1999. A simple and universal method for molecular sexing of non-ratite birds. Journal of Avian Biology 30(1):116-121.

Goudet, J. 2001. FSTAT, version 2.9.3, A program to estimate and test gene diversities and fixation indices. Lausanne University, Lausanne, Switzerland.

Goudet, J., Perrin, N. and Waser, P. 2002. Tests for sex-biased dispersal using bi-parentally inherited. Molecular Ecology 11(6):1103–1114.

Greenwood, P.J. 1980. Mating systems, philopatry and dispersal in birds and mammals. Animal behaviour 28(4):1140–1162.

Guillot, G., Mortier, F. and Estoup, A. 2005. GENELAND: a computer package for landscape genetics. Molecular Ecology Notes 5(3):712–715.

Guillot, G., Santos, F. and Estoup, A. 2008. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. Bioinformatics 24(11):1406–1407.

Hinke, J.T., Salwicka, K., Trivelpiece, S.G., Watters, G.M. and Trivelpiece, W.Z. 2007. Divergent responses of *Pygoscelis* penguins reveal a common environmental driver. Oecologia 153(4):845–855.

Jakobsson, M. and Rosenberg, N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23(14):1801–1806.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24(11):1403–1405.

Jombart, T. 2015. An introduction to adegenet 2.0.0.

Jombart, T., Devillard, S. and Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11(94):1–15.

Jouventin, P., Cuthbert, R. and Ottvall, R. 2006. Genetic isolation and divergence in sexual traits: evidence for the northern rockhopper penguin *Eudyptes moseleyi* being a sibling species. Molecular Ecology 15(11):3413–3423.

Korczak-Abshire, M., Chwedorzewska, K. J., Wąsowicz, P. and Bednarek, P.T. 2012. Genetic structure of declining chinstrap penguin (*Pygoscelis antarcticus*) populations from South Shetland Islands (Antarctica). Polar Biology 35(11):1681–1689.

Latch, E. K., Dharmarajan, G., Glaubitz, J. C. and Rhodes, O. E. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conservation Genetics 7(2):295–302.

Lischer, H.E.L. and Excoffier, L. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28(2):298–299.

Lowe, W.H. and Allendorf, F.W. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19(15):3038–3051.

Lynch, H.J., Naveen, R., Trathan, P.N. and Fagan, W.F. 2012. Spatially integrated assessment reveals widespread changes in penguin populations on the Antarctic Peninsula. Ecology Society of America 93(6):1367–1377.

Macdonald, J.A., Barton, K.J. and Metcalf, P. 2002. Chinstrap penguins (*Pygoscelis antarctica*) nesting on Sabrina Islet, Balleney Islands, Antarctica. Polar Biology 25(6):442–447.

Naveen, R., Lynch, H. J., Forrest, S., Mueller, T. and Polito, M. 2012. First direct, site-wide penguin survey at Deception Island, Antarctica, suggests significant declines in breeding chinstrap penguins. Polar Biology 35(12):879–1888.

Niemandt, C., Kovacs, K. M., Lydersen, C., Dyer, B. M., Isaksen, K., Hofmeyr, G. G., Mehlum, F., and de Bruyn, P.N.J. 2016. Chinstrap and macaroni penguin diet and demography at Nyrøysa, Bouvetøya. Antarctic Science 28(2):91-100.

Paetkau, D., Slade, R., Burden, M. and Estoup, A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology 13(1):55–65.

Paetkau, D., Calvert, W., Stirling, I. and Strobeck, C. 1995. Microsatellite analysis of population structure in Canadian polar bears. Molecular Ecology 4(3):347–354.

Palsbøll, P.J., Bérubé, M. and Allendorf, F.W. 2006. Identification of management units using population genetic data. Trends in Ecology and Evolution 22(1):11-16.

Peña, F., Poulin, E., Dantas, G.P.M, González-Acuña, D., Petry, M.V. and Vianna, J.A. 2014. Have historical climate changes affected gentoo penguin (*Pygoscelis papua*) populations in Antarctica? PLoS one 9(4):p.e95375.

Petry MV, Valls FCL, de Souza Petersen E, Krüger L, da Cruz Piuco R, dos Santos CR. 2016. Breeding sites and population of seabirds on Admiralty Bay, King George Island. Antarctica. Polar Biology 39(7):1343-1349.

Pike, N. 2011. Using false discovery rates for multiple comparisons in ecology and evolution. Methods in Ecology and Evolution 2(3):278–282.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L. and Estoup, A. 2004. GENECLASS2: A software for genetic assignment and first-generation migrant detection. Journal of Heredity 95(6):536–539.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics Society of America 155(2):945–959.

Pusey, A.E. 1987. Sex-biased dispersal and inbreeding avoidance in birds and mammals. Trends in Ecology and Evolution 2(10):295–299.

Rannala, B. and Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. PNAS 94(17):9197–9201.

Raymond, J.A. 1975. Chinstrap penguin at McMurdo Sound. Antarctic Journal of United States 10:275.

Roeder, A.D., Marshall, R.K., Mitchelson, A.J., Visagathilagar, T., Ritchie, P.A., Love, D.R.,Pakai, T.J., McPartland, H.C., Murray, N.D., Robinson, N.A., Kerry, K.E. and Lambert, D.M. 2001. Gene flow on the ice: genetic differentiation among Adélie penguin colonies around Antarctica. Molecular Ecology 10(7):1645–1656.

Rosenberg, N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology 4(1):137–138.

Sander M, Balbão TC, Costa ES, Dos Santos CR, Petry MV. 2007. Decline of the breeding population of *Pygoscelis antarctica* and *Pygoscelis adeliae* on Penguin Island, South Shetland, Antarctica. Polar Biology 30(5):651-654.

Scheifler, R., Gauthier-Clerc, M., Bohec, C. L., Crini, N., Cœurdassier, M., Badot, P.M., Giraudoux, P. and Maho, Y.L. 2005. Mercury concentrations in king penguin (Aptenodytes patagonicus) feathers at Crozet Islands (sub-Antarctic): temporal trend between 1966-1974 and 2000-2001. Environmental Toxicology and Chemistry 24(1):125–128.

Schlosser, J.A., Dubach, J.M., Garner, T.W., Araya, B., Bernal, M., Simeone, A., Smith, K.A. and Wallace, R.S. 2009. Evidence for gene flow differs from observed dispersal patterns in the Humboldt penguin, *Spheniscus humboldti*. Conservation Genetics 10(4):839–849.

Spurr, E. 1985. Records of chinstrap penguins in the Ross Sea region, Antarctica. Notornis 32(1):1–4.

Taylor, B.L. and Dizon, A.E. 2017. The need to estimate power to link genetics and demography for conservation. Society for Conservation Biology 10(2):661–664.

Trivelpiece, W.Z., Buckelew, S., Reiss, C. and Trivelpiece, S.G. 2007. The winter distribution of chinstrap penguins from two breeding sites in the South Shetland Islands of Antarctica. Polar Biology 30(10):231–1237.

Trivelpiece, W.Z., Hinke, J.T., Miller, A.K., Reiss, C.S., Trivelpiece, S.G. ad Watters, G.M. 2011. Variability in krill biomass links harvesting and climate warming to penguin population changes in Antarctica. Proceedings of the National Academy of Sciences 108(18):7625–7628.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. and Shipley, P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Resources 4(3):535-538.

Vianna, J.A., Noll, D., Mura-Jornet, I., Valenzuela-Guerra, P., González-Acuña, D., Navarro, C., Loyola D.E., and Dantas, G.P.M. 2017. Comparative genome-wide polymorphic microsatellite markers in Antarctic penguins through next generation sequencing. Genetics and Molecular Biology 40(3):676-687.

Vianna, J.A., Noll, D., Dantas, G.P.M, Petry, M.V., Barbosa, A., González-Acuña, D., Le Bohec, C., Bonadonna, F. and Poulin, E. 2016. Marked phylogeographic structure of Gentoo penguin reveals an ongoing diversification process along the Southern Ocean. Molecular Phylogenetics and Evolution 107:486–498.

Younger, J.L., Emmerson, L.M., and Miller, K. J. 2016. The influence of historical climate changes on Southern Ocean marine predator populations: a comparative analysis. Global change biology 22(2):474-493.

Waples, R. and Gaggiotti, O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15(6):1419–1439.

Wilson, R.P. 1997. A Method for restraining penguins. Marine Ornithology 25:72–73.

Woehler, E.J., Cooper, J., Croxall, J.P., Fraser, W.R., Kooyman, G.L., Miller, G.D., Patterson, D.L., Peter, H.U., Ribic, C.A., Salwicka, K., Trivelpiece, W.Z. and Weimerskirch, H. 2001. A statistical assessment of the status and trends of a statistical assessment of the status and trends of antarctic and subantarctic seabirds. SCAR. 44 pp. Available at: http://www.scar.org/scar_media/documents/publications/Woehler_etal_Status_of_Seabirds_2001.pdf

Woehler, E.J. 1993. The distribution and abundance of Antarctic and Subantarctic penguins, Cambridge, United Kingdom: Scientific Committee on Antarctic Research (SCAR).

Wright, S. 1969. Evolution and the genetics of populations: Vol. 2. The theory of gene frequencies. Chicago: University of Chicago Press. 520 pp.

# Appendix

**Table S1.** Allelic richness (A), and expected ($H_E$) and observed heterozigosity ($H_O$) values for 11 microsatellite *loci*, for all population examined. Study site (collection location) abbreviations correspond to EI: Elephant Island, PI: Penguin Island, BP: Barton Peninsula, AI: Ardley Island, GI: Greenwich Island, MB: Miers Bluff, HP: Hannah Point, CS: Cape Shirreff, BH: Baily Head, VC: Vapour Col, KI: Kopaitic Island, GP: Georges Point and BI: Bouvetøya.

| *Locus* | EI | PI | BP | AI | GI | MB | HP | CS | BH | VC | KI | GP | BI | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **AP-19** | | | | | | | | | | | | | | |
| A | 5 | 8 | 5 | 5 | 5 | 6 | 7 | 7 | 6 | 6 | 7 | 7 | 7 | 9 |
| $H_E$ | 0.74 | 0.83 | 0.78 | 0.73 | 0.79 | 0.80 | 0.77 | 0.79 | 0.78 | 0.75 | 0.81 | 0.81 | 0.83 | |
| $H_O$ | 0.93 | 0.94 | 0.82 | 0.71 | 0.50 | 0.72 | 0.84 | 0.82 | 0.78 | 0.67 | 0.56 | 0.80 | 1.00 | |
| **AP-26** | | | | | | | | | | | | | | |
| A | 7 | 8 | 10 | 6 | 6 | 8 | 7 | 6 | 6 | 8 | 8 | 8 | 8 | 11 |
| $H_E$ | 0.83 | 0.81 | 0.81 | 0.79 | 0.80 | 0.88 | 0.82 | 0.80 | 0.84 | 0.82 | 0.78 | 0.85 | 0.81 | |
| $H_O$ | 0.63 | 084 | 0.69 | 0.71 | 0.69 | 0.73 | 0.80 | 0.80 | 1.00 | 0.87 | 0.92 | 0.73 | 0.71 | |
| **AP-3** | | | | | | | | | | | | | | |
| A | 3 | 2 | 2 | 2 | 2 | 2 | 3 | 2 | 1 | 2 | 2 | 2 | 2 | 3 |
| $H_E$ | 0.22 | 0.10 | 0.03 | 0.14 | 0.08 | 0.09 | 0.16 | 0.06 | - | 0.12 | 0.10 | 0.13 | 0.16 | |
| $H_O$ | 0.24 | 0.10 | 0.03 | 0.00 | 0.08 | 0.09 | 0.17 | 0.06 | - | 0.13 | 0.11 | 0.13 | 0.17 | |
| **AP-90** | | | | | | | | | | | | | | |
| A | 8 | 7 | 9 | 9 | 8 | 4 | 9 | 9 | 7 | 8 | 9 | 9 | 8 | 9 |
| $H_E$ | 0.83 | 0.80 | 0.85 | 0.91 | 0.87 | 0.77 | 0.85 | 0.84 | 0.88 | 0.89 | 0.87 | 0.89 | 0.86 | |
| $H_O$ | 0.93 | 0.87 | 0.79 | 0.82 | 0.91 | 1.00 | 0.84 | 0.90 | 0.89 | 0.80 | 0.85 | 0.86 | 0.87 | |
| **AP-85** | | | | | | | | | | | | | | |
| A | 5 | 6 | 6 | 4 | 6 | 6 | 6 | 7 | 5 | 5 | 7 | 6 | 6 | 8 |
| $H_E$ | 0.74 | 0.75 | 0.77 | 0.74 | 0.82 | 0.77 | 0.78 | 0.76 | 0.82 | 0.77 | 0.77 | 0.71 | 0.76 | |
| $H_O$ | 0.69 | 0.76 | 0.69 | 0.57 | 0.83 | 0.72 | 0.72 | 0.70 | 0.67 | 0.80 | 0.70 | 0.80 | 0.90 | |
| **AP-78** | | | | | | | | | | | | | | |
| A | 4 | 6 | 5 | 4 | 4 | 2 | 3 | 4 | 5 | 4 | 4 | 4 | 5 | 10 |
| $H_E$ | 0.36 | 0.53 | 0.25 | 0.38 | 0.27 | 0.09 | 0.28 | 0.38 | 0.55 | 0.49 | 0.29 | 0.25 | 0.44 | |
| $H_O$ | 0.41 | 0.44 | 0.17 | 0.29 | 0.14 | 0.09 | 0.28 | 0.23 | 0.67 | 0.53 | 0.26 | 0.20 | 0.39 | |
| **CP-6** | | | | | | | | | | | | | | |
| A | 4 | 5 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 3 | 6 |
| $H_E$ | 0.57 | 0.63 | 0.64 | 0.67 | 0.58 | 0.58 | 0.61 | 0.52 | 0.57 | 0.39 | 0.64 | 0.64 | 0.66 | |
| $H_O$ | 0.55 | 0.78 | 0.56 | 0.69 | 0.75 | 0.75 | 0.60 | 0.60 | 0.45 | 0.47 | 0.62 | 0.67 | 0.50 | |
| **AP-61** | | | | | | | | | | | | | | |
| A | 6 | 6 | 7 | 5 | 5 | 6 | 7 | 8 | 6 | 7 | 7 | 6 | 7 | 10 |
| $H_E$ | 0.69 | 0.64 | 0.68 | 0.71 | 0.72 | 0.68 | 0.75 | 0.76 | 0.78 | 0.80 | 0.69 | 0.62 | 0.73 | |
| $H_O$ | 0.71 | 0.58 | 0.63 | 0.50 | 0.80 | 0.50 | 0.64 | 0.64 | 0.78 | 0.80 | 0.73 | 0.53 | 0.74 | |
| **CP-25** | | | | | | | | | | | | | | |
| A | 8 | 7 | 8 | 7 | 8 | 7 | 8 | 10 | 5 | 6 | 10 | 8 | 6 | 11 |
| $H_E$ | 0.84 | 0.84 | 0.85 | 0.78 | 0.84 | 0.77 | 0.78 | 0.83 | 0.78 | 0.82 | 0.82 | 0.83 | 0.67 | |
| $H_O$ | 0.85 | 0.84 | 0.83 | 0.92 | 0.79 | 0.72 | 0.75 | 0.87 | 0.67 | 0.86 | 0.86 | 0.72 | 0.68 | |
| **GP-36** | | | | | | | | | | | | | | |
| A | 3 | 3 | 5 | 4 | 4 | 3 | 3 | 5 | 3 | 4 | 4 | 3 | 5 | 5 |
| $H_E$ | 0.31 | 0.54 | 0.43 | 0.57 | 0.58 | 0.54 | 0.44 | 0.47 | 0.39 | 0.48 | 0.46 | 0.48 | 0.52 | |
| $H_O$ | 0.21 | 0.73 | 0.31 | 0.58 | 0.58 | 0.50 | 0.60 | 0.33 | 0.44 | 0.47 | 0.39 | 0.67 | 0.52 | |
| **GP-15** | | | | | | | | | | | | | | |
| A | 8 | 7 | 8 | 8 | 6 | 8 | 9 | 10 | 5 | 7 | 8 | 7 | 7 | 12 |
| $H_E$ | 0.78 | 0.79 | 0.77 | 0.81 | 0.71 | 0.82 | 0.83 | 0.81 | 0.69 | 0.81 | 0.46 | 0.81 | 0.71 | |
| $H_O$ | 0.71 | 0.73 | 0.77 | 0.75 | 0.64 | 0.82 | 0.80 | 0.77 | 0.67 | 0.93 | 0.39 | 0.93 | 0.65 | |

**Table S2.** Summary of pairwise genetic differences ($F_{ST}$) between chinstrap penguin colonies calculated from the 11 microsatellite *loci*.

| | EI | PI | BP | AI | GI | MB | HP | CS | BH | VC | KI | GP | BI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| EI | - | 0.744 | 0.765 | 0.542 | 0.669 | 0.971 | 0.266 | 0.971 | 0.200 | 0.266 | 0.266 | 0.002 | 0.971 |
| PI | -0.002 | - | 0.687 | 0.740 | 0.971 | 0.971 | 0.325 | 0.619 | 0.295 | 0.159 | 0.917 | 0.002 | 0.655 |
| BP | 0.000 | 0.001 | - | 0.372 | 1.000 | 1.000 | 0.256 | 1.000 | 0.159 | 0.147 | 0.927 | 0.000 | 0.927 |
| AI | 0.008 | 0.000 | 0.010 | - | 0.971 | 0.917 | 0.266 | 0.643 | 0.232 | 0.280 | 0.669 | 0.007 | 0.656 |
| GI | 0.003 | -0.012 | -0.021 | -0.009 | - | 1.000 | 0.971 | 1.000 | 0.677 | 0.740 | 0.971 | 0.159 | 0.971 |
| MB | -0.012 | -0.012 | -0.019 | -0.002 | -0.026 | - | 0.959 | 1.000 | 0.266 | 0.331 | 0.971 | 0.130 | 0.994 |
| HP | 0.010 | 0.007 | 0.009 | 0.013 | -0.007 | -0.006 | - | 0.303 | 0.534 | 0.847 | 0.295 | 0.303 | 0.284 |
| CS | -0.006 | 0.003 | -0.006 | 0.005 | -0.011 | -0.012 | 0.007 | - | 0.245 | 0.256 | 0.847 | 0.000 | 0.971 |
| BH | 0.025 | 0.012 | 0.025 | 0.032 | 0.003 | 0.023 | 0.006 | 0.020 | - | 0.687 | 0.256 | 0.183 | 0.256 |
| VC | 0.012 | 0.019 | 0.018 | 0.017 | 0.000 | 0.013 | -0.003 | 0.013 | -0.001 | - | 0.088 | 0.677 | 0.248 |
| KI | 0.011 | -0.004 | -0.003 | 0.003 | -0.009 | -0.012 | 0.007 | -0.001 | 0.018 | 0.025 | - | 0.000 | 0.927 |
| GP | **0.045** | **0.049** | **0.047** | **0.052** | 0.023 | 0.027 | 0.008 | **0.040** | 0.023 | 0.000 | **0.054** | - | 0.007 |
| BI | -0.007 | 0.000 | -0.003 | 0.002 | -0.011 | -0.013 | 0.007 | -0.006 | 0.016 | 0.012 | -0.004 | **0.031** | - |

Bold values are significantly different from zero after the FDR correction ($p < 0.05$)

**Table S3.** GeneClass2 percentage test results using microsatellite data for chinstrap penguins from 13 colonies for (a) genetic assignment using Peatkau *et al*., (1995) criterion and (b) first-generation migrant. Rows indicate the sample site collection and columns indicate the colonies to which the individuals were assigned. Colony self-assignments are in bold.

a) Assignment test

| | **EI** | **PI** | **BP** | **AI** | **GI** | **MB** | **HP** | **CS** | **BH** | **VC** | **KI** | **GP** | **BI** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **EI** | **12** | 6 | 0 | 0 | 12 | 24 | 0 | 24 | 12 | 0 | 6 | 6 | 0 |
| **PI** | 0 | **26** | 5 | 0 | 5 | 16 | 5 | 0 | 11 | 21 | 5 | 5 | 0 |
| **BP** | 10 | 3 | **14** | 0 | 17 | 3 | 3 | 17 | 3 | 0 | 17 | 3 | 7 |
| **AI** | 14 | 0 | 14 | **14** | 14 | 7 | 0 | 14 | 0 | 7 | 0 | 7 | 7 |
| **GI** | 0 | 7 | 36 | 0 | **0** | 14 | 7 | 7 | 0 | 0 | 21 | 0 | 7 |
| **MB** | 0 | 9 | 18 | 0 | 18 | **0** | 9 | 9 | 0 | 9 | 9 | 0 | 18 |
| **HP** | 8 | 12 | 16 | 0 | 12 | 8 | **16** | 8 | 0 | 8 | 4 | 8 | 0 |
| **CS** | 3 | 3 | 13 | 0 | 7 | 0 | 7 | **20** | 3 | 10 | 20 | 3 | 10 |
| **BH** | 11 | 11 | 0 | 0 | 0 | 0 | 22 | 11 | **11** | 22 | 0 | 0 | 11 |
| **VC** | 7 | 13 | 7 | 0 | 7 | 0 | 20 | 7 | 7 | **13** | 0 | 20 | 0 |
| **KI** | 7 | 10 | 10 | 7 | 17 | 10 | 3 | 3 | 7 | 3 | **20** | 0 | 3 |
| **GP** | 0 | 7 | 7 | 13 | 13 | 0 | 7 | 7 | 0 | 33 | 0 | **13** | 0 |
| **BI** | 4 | 13 | 4 | 4 | 4 | 9 | 17 | 4 | 0 | 9 | 9 | 4 | **17** |

b) First-generation migrant test

| | **EI** | **PI** | **BP** | **AI** | **GI** | **MB** | **HP** | **CS** | **BH** | **VC** | **KI** | **GP** | **BI** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **EI** | **12** | 6 | 0 | 0 | 12 | 24 | 0 | 24 | 12 | 0 | 6 | 6 | 0 |
| **PI** | 0 | **21** | 5 | 0 | 5 | 16 | 5 | 0 | 11 | 21 | 11 | 0 | 5 |
| **BP** | 10 | 3 | **14** | 0 | 17 | 3 | 3 | 17 | 3 | 0 | 17 | 3 | 7 |
| **AI** | 14 | 0 | 14 | **14** | 14 | 7 | 0 | 14 | 0 | 7 | 7 | 0 | 7 |
| **GI** | 0 | 7 | 36 | 0 | **0** | 14 | 7 | 7 | 0 | 7 | 14 | 0 | 7 |
| **MB** | 0 | 9 | 18 | 0 | 18 | **0** | 9 | 9 | 0 | 9 | 9 | 0 | 18 |
| **HP** | 4 | 12 | 16 | 0 | 12 | 12 | **16** | 8 | 0 | 8 | 8 | 4 | 0 |
| **CS** | 3 | 3 | 13 | 0 | 7 | 0 | 7 | **20** | 3 | 10 | 20 | 3 | 10 |
| **BH** | 11 | 11 | 0 | 0 | 0 | 0 | 22 | 11 | **11** | 22 | 0 | 0 | 11 |
| **VC** | 7 | 13 | 7 | 0 | 7 | 0 | 20 | 7 | 7 | **13** | 0 | 20 | 0 |
| **KI** | 7 | 10 | 10 | 7 | 17 | 10 | 3 | 3 | 10 | 0 | **20** | 0 | 3 |
| **GP** | 0 | 7 | 7 | 13 | 13 | 0 | 0 | 7 | 0 | 40 | 0 | **13** | 0 |
| **BI** | 4 | 13 | 4 | 4 | 4 | 9 | 17 | 4 | 0 | 9 | 9 | 4 | **17** |

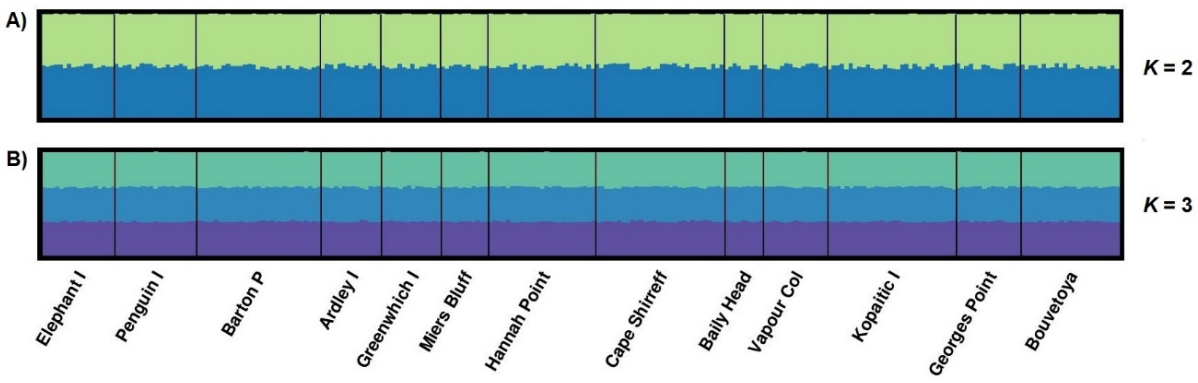


**Fig. S1**. Plot of assignment probabilities from STRUCTURE. A vertical bar represents an individual and the colors represent the different clusters found. All plots were generated via running 10 replicates. Figures show the optimal number of clusters for no admixture model, with independent allele frequency and no location information supplied using A) Posterior probability of K (LnP(D)) and B) Evanno's method.

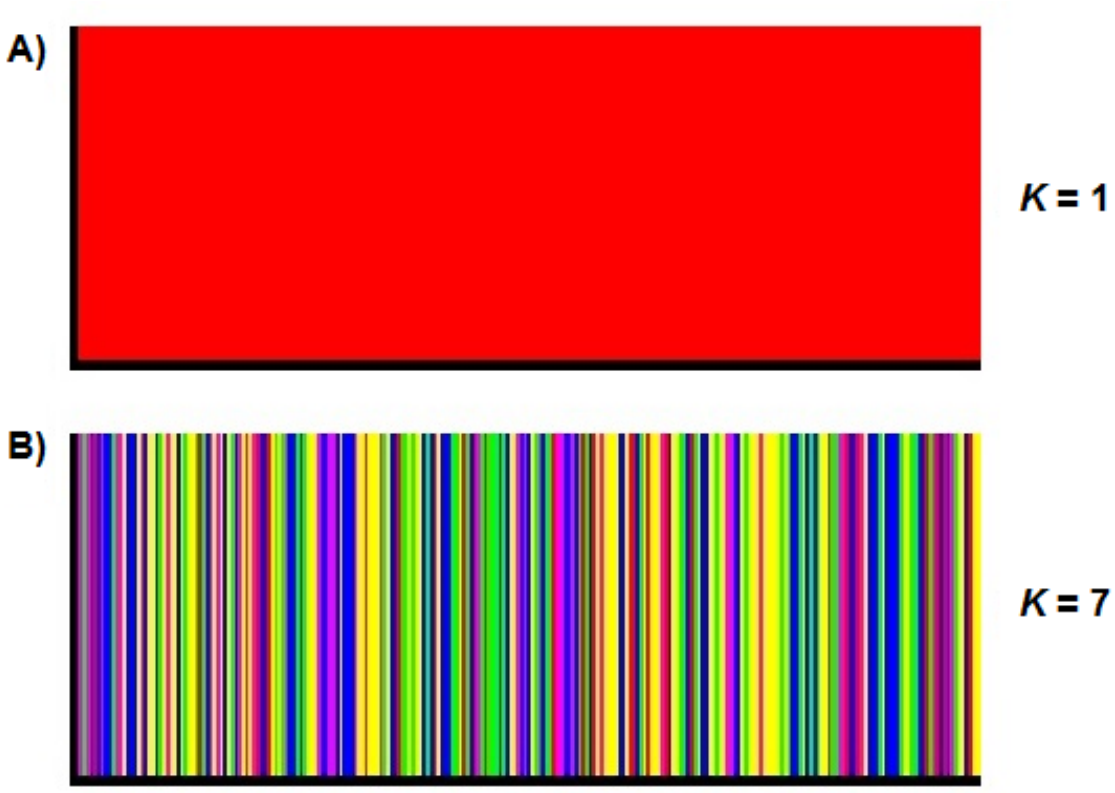


**Fig. S2**. Plot of assignment probabilities from BAPS. Vertical lines represent each individual and the color refers to clusters found by this analysis. A) spatial with both, mixture and admixture models ($K$ = 1) and B) non-spatial admixture model ($K$ = 7).

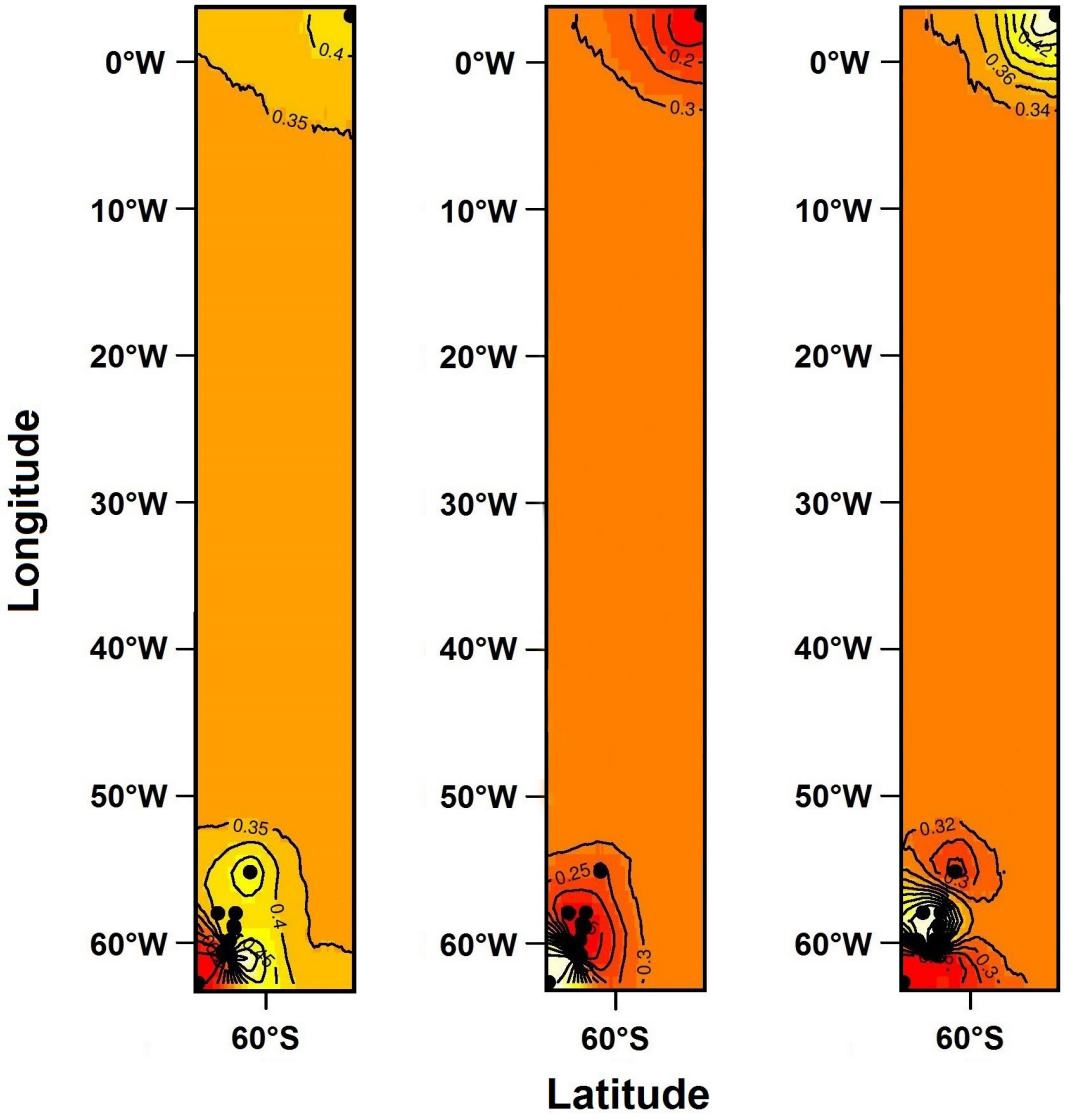


**Fig. S3**. Posterior probabilities of population membership from the spatial model with correlated allele frequencies model using GENELAND. Lighter colors indicate higher probabilities of population membership. Three genetic clusters were identified. Left: Kopaitic Island, middle: Georges Point, and right: northern WAP locations and Bouvetøya.

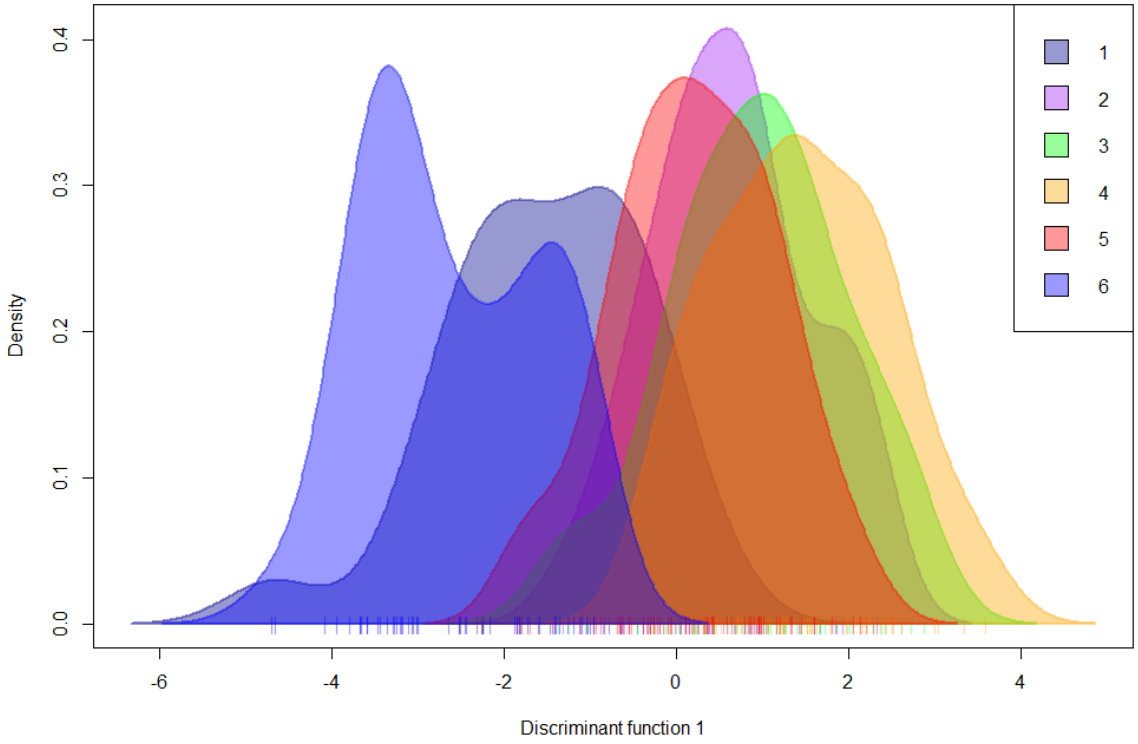


**Fig. S4.** Discriminant Analysis of Principal Components (DAPC). The six genetic clusters identified by adegenet are shown in different colors. All six groups overlapped extensively, and none of them represent a specific colony.

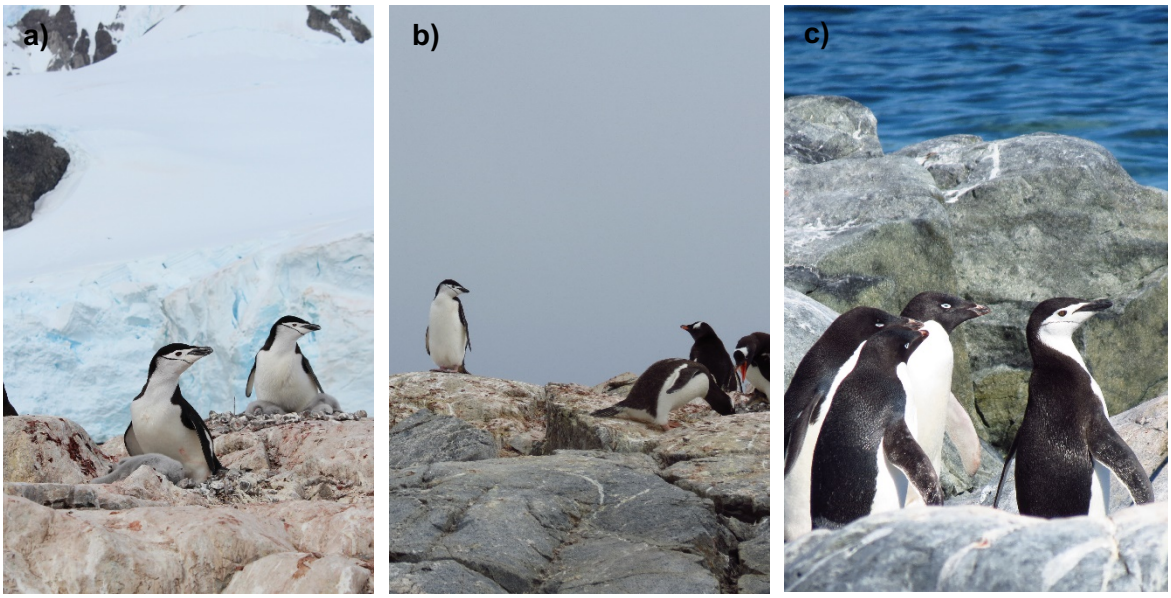


**Fig. S5.** Chinstrap penguins south of their normal breeding range. A) Breeding pairs on Waterboat point (Gabriel González Videla base; 64°49'S, 62°51'W), B) a single individual surrounded by gentoo colonies on Doumer Island (Yelcho base; 64°65'S, 63°35'W) and c) another single bird on Avian Island (67°46'S, 68°54'W) surrounded by Adélie penguins.