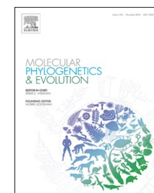
Molecular Phylogenetics and Evolution 139 (2019) 106563

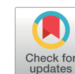
Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev



More than the eye can see: Genomic insights into the drivers of genetic differentiation in Royal/Macaroni penguins across the Southern Ocean



María José Frugone[a,b,c], María Eugenia López[d,e], Nicolás I. Segovia[b,f], Theresa L. Cole[g,h], Andrew Lowther[i], Pierre Pistorius[j], Gisele P.M. Dantas[k], Maria Virginia Petry[l], Francesco Bonadonna[m], Phil Trathan[n], Andrea Polanowski[o], Barbara Wienecke[o], Ke Bi[p], Cynthia Y. Wang-Claypool[p], Jonathan M. Waters[g], Rauri C.K. Bowie[p], Elie Poulin[a,b], Juliana A. Vianna[c,*]

[a] *Laboratorio de Ecología Molecular, Departamento de Ciencias Ecológicas, Facultad de Ciencias, Universidad de Chile, Las Palmeras # 3425, Ñuñoa, Santiago, Chile*
[b] *Instituto de Ecología y Biodiversidad (IEB), Las Palmeras # 3425, Ñuñoa, Santiago, Chile*
[c] *Pontificia Universidad Católica de Chile, Facultad de Agronomía e Ingeniería Forestal, Departamento de Ecosistemas y Medio Ambiente, Vicuña Mackenna 4860, Macul, Santiago, Chile*
[d] *Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden*
[e] *Facultad de Ciencias Veterinarias y Pecuarias, Universidad de Chile, Santiago 8820808, Chile*
[f] *Universidad Católica del Norte, Facultad de Ciencias del Mar, Departamento de Biología Marina, Coquimbo, Chile*
[g] *Department of Zoology, University of Otago, P.O. Box 56, Dunedin 9054, New Zealand*
[h] *Manaaki Whenua Landcare Research, PO Box 69040, Lincoln, Canterbury 7640, New Zealand*
[i] *Norwegian Polar Institute, Tromsø N-9297, Norway*
[j] *DST/NRF Centre of Excellence at the Percy FitzPatrick Institute for African Ornithology, Department of Zoology, Nelson Mandela University, Port Elizabeth 6031, South Africa*
[k] *Pontificia Universidade Católica de Minas Gerais, PPG in Vertebrate Biology, Belo Horizonte, Brazil*
[l] *Universidade do Vale do Rio dos Sinos, Laboratório de Ornitologia e Animais Marinhos, Av. Unisinos, 950, São Leopoldo, RS, Brazil*
[m] *CEFE UMR 5175, CNRS, Université de Montpellier, Université Paul-Valéry Montpellier, EPHE, 1919 route de Mende, 34293 Montpellier cedex 5, France*
[n] *British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK*
[o] *Australian Antarctic Division, 203 Channel Highway, Kingston, Tasmania 7050, Australia*
[p] *Museum of Vertebrate Zoology and Department of Integrative Biology, 3101 Valley Life Science Building, University of California, Berkeley, CA 94720-3160, USA*

ARTICLE INFO

Keywords:
ddRAD
Single nucleotide polymorphisms
Speciation
Population structure
Eudyptes chrysolophus
Eudyptes schlegeli

ABSTRACT

The study of systematics in wide-ranging seabirds can be challenging due to the vast geographic scales involved, as well as the possible discordance between molecular, morphological and behavioral data. In the Southern Ocean, macaroni penguins (*Eudyptes chrysolophus*) are distributed over a circumpolar range including populations in Antarctic and sub-Antarctic areas. Macquarie Island, in its relative isolation, is home to a closely related endemic taxon — the royal penguin (*Eudyptes schlegeli*), which is distinguishable from *E. chrysolophus* mainly by facial coloration. Although these sister taxa are widely accepted as representing distinct species based on morphological grounds, the extent of their genome-wide differentiation remains uncertain. In this study, we use genome-wide Single Nucleotide Polymorphisms to test genetic differentiation between these geographically isolated taxa and evaluate the main drivers of population structure among breeding colonies of macaroni/royal penguins. Genetic similarity observed between macaroni and royal penguins suggests they constitute a single evolutionary unit. Nevertheless, royal penguins exhibited a tendency to cluster only with macaroni individuals from Kerguelen Island, suggesting that dispersal occurs mainly between these neighboring colonies. A stepping stone model of differentiation of macaroni/royal populations was further supported by a strong pattern of isolation by distance detected across its whole distribution range, possibly driven by large geographic distances between colonies as well as natal philopatry. However, we also detected intraspecific genomic differentiation between Antarctic and sub-Antarctic populations of macaroni penguins, highlighting the role of environmental factors together with geographic distance in the processes of genetic differentiation between Antarctic and sub-Antarctic waters.

* Corresponding author.
E-mail address: jvianna@uc.cl (J.A. Vianna).

https://doi.org/10.1016/j.ympev.2019.106563
Received 18 April 2019; Received in revised form 12 July 2019; Accepted 15 July 2019
Available online 16 July 2019
1055-7903/ © 2019 Elsevier Inc. All rights reserved.

1. Introduction

Understanding the systematics and taxonomy of widespread seabirds can be a challenging task. Limited morphological variation can mask ongoing diversification processes in recently-differentiated lineages (Vianna et al., 2017). This could potentially result in recently-diverged lineages being difficult to distinguish from each other in the absence of obvious differentiation in external morphology (Grosser et al., 2015), even when the lineages are effectively reproductively isolated (Garg et al., 2016). On the other hand, phenotypic plasticity or phenotypic variation can lead to the description of different taxa within single genetic units (Cole et al., 2019; Grosser et al., 2015; Mason and Taylor, 2015).

Seabirds in general exhibit high dispersal capabilities. Although they can be distributed over an apparently uniform marine environment, populations are often isolated by cryptic geographic barriers (Friesen et al., 2007). In some cases, population genetic structure within widespread taxa may by influenced by processes occurring across both the breeding and nonbreeding distribution as well as by fidelity to natal colonies (philopatry) (Friesen et al., 2007). Particularly in penguins, the main drivers of genetic differentiation and within species lineage diversification appear to be associated with the extent of natal philopatry, the distribution range occupied during the inter-breeding period, and possibly, the presence of oceanic fronts (Clucas et al., 2018; Frugone et al., 2018; Moon et al., 2017).

Macaroni penguins (*Eudyptes chrysolophus*) are the most widely distributed *Eudyptes* species and the only member of this genus that occupies the Antarctic region. Macaroni breeding sites include colonies located in Antarctic waters like Elephant Island and South Georgia, as well as colonies located in sub-Antarctic waters in areas such as Crozet and Kerguelen Islands. Around 5000 km east of the Kerguelen Islands, Macquarie Island (and nearby Bishop and Clerk Islands) provides the only nesting grounds for the endemic royal penguin (*Eudyptes schlegeli*). While royal penguins are commonly considered a separate species to macaroni penguins (Bertelli and Giannini, 2005; Ksepka et al., 2006), the taxonomic classification remains controversial, with some recent studies considering them a sub-species (Cole et al., 2019), and others considering them conspecific (Christidis and Boles, 2008). Differences between royal and macaroni penguins include: (1) distribution (royal are endemic to Macquarie Island, whereas macaroni, though distributed over a wide range, but do not reach Macquarie; Fig. 1) and (2) their morphology. In particular, royal penguins have white to grey faces (in contrast to black faces of macaroni penguins), are slightly larger, and have longer and deeper beaks compared to macaroni penguins (Hart et al., 2009; Hull, 1996). Additionally, there are some differences in breeding phenology but differences among colonies of macaroni penguins tend to be in a similar range of variation (days or a few weeks) than those observed between macaroni and royal penguins (Hindell, 1988; Hull, 1999). The same is true for differences in diet composition between royal and macaroni penguins since similar variations have been observed among macaroni penguins colonies, suggesting that prey availability could account for these differences (Hindell, 1988; Hull, 1999; Kooyman, 2002).

The hypothesis that royal and macaroni penguins may be a single

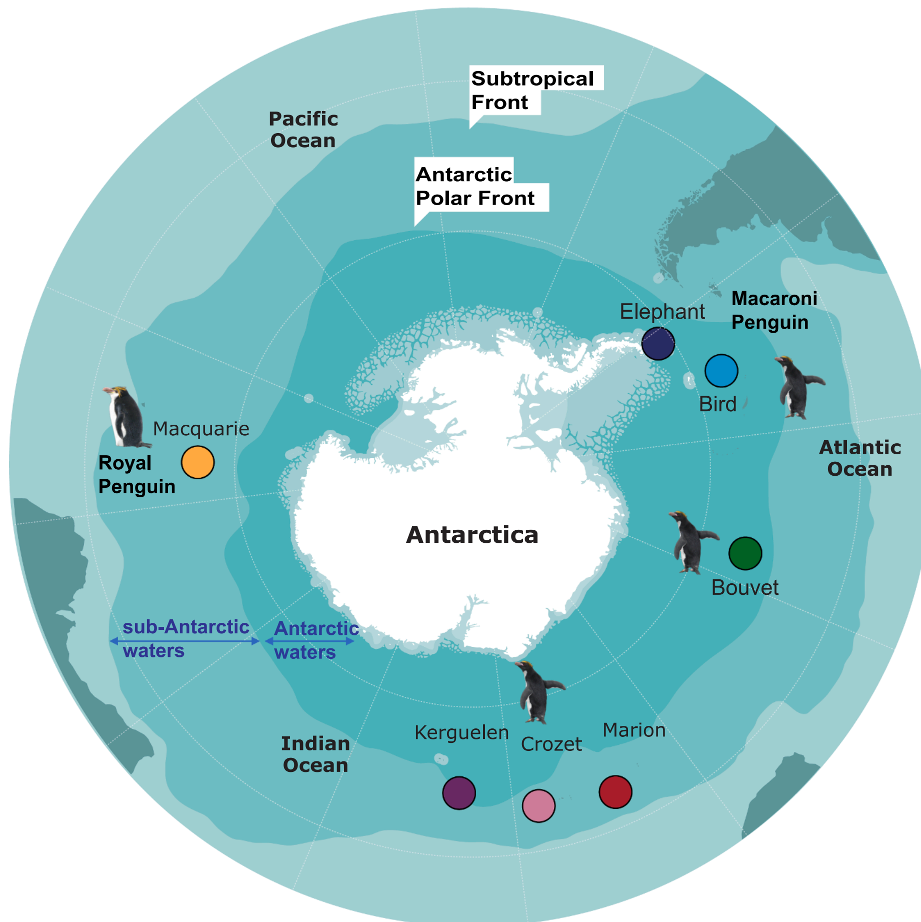


Fig. 1. Sampling locations of macaroni and royal (Macquarie) penguins.

evolutionary unit has been hinted at by several recent molecular studies, mainly using mtDNA (Cole et al., 2019; Cole et al., 2018b; Frugone et al., 2018). However, it remains unclear whether these taxa represent recently diverged lineages in the process of speciation (Cole et al., 2018b), or a single taxon. On the other hand, a study using mtDNA and nuclear introns across the distributional range of macaroni penguins suggested there may be some differentiation between Antarctic and sub-Antarctic populations (Frugone et al., 2018). Here, we use genome-wide Single Nucleotide Polymorphisms (SNPs) to test whether royal penguins are genetically isolated from macaroni penguins and evaluate the main drivers of population structure among breeding colonies of macaroni penguins.

2. Methods

2.1. Genomic data collection

Tissue samples were collected from royal penguins at Macquarie Island (n = 7), and blood samples from macaroni penguins were collected from six Southern Ocean locations (Fig. 1; Table 1) representing the circumpolar range of the species (total n = 69). In the capture and handling of penguins, we followed established procedures described in Wilson (1997), to reduce stress for both sampled penguins and surrounding colony members. Approximately 0.5 mL of blood was taken from the brachial or external metatarsal vein, using 23 G needles and were stored in > 95% ethanol. Permission for sampling, access to the penguin colonies, and animal ethics approval were granted by the responsible authorities for each sample location (Supplementary Table 1).

Genomic DNA was isolated from each sample following the salting out protocol from Aljanabi and Martinez (1997) with the modifications described in Vianna et al. (2017). We evaluated the quality of DNA through electrophoresis, in a 1% agarose gel. The concentration of each sample was determined by a Qubit fluorometer (Life Technologies). We prepared double digest Restriction-site Associated DNA (ddRAD seq) libraries, following the protocol described in Peterson et al. (2012). Genomic DNA of each individual (500 ng) was digested using 0.5 µL of EcoRI (0.1 U/µL) and 0.5 µL of SphI-HF (0.1 U/µL), at 37 °C for 3 h. Each sample was then ligated to one of 24 unique barcodes (P1). We used a Pippin Prep (Sage Science) to select fragments of 300–400 bp and assessed size selection, integrity, and quantification of the samples using the Agilent 2100 Bioanalyzer system. Illumina adapters (P5 and P7) were ligated for each sample, and each pool was amplified using 8–10 PCR cycles with an Applied Biosystems 2720 Thermal Cycler, using the following protocol: initial denaturation was at 98 °C for 30 s, then 8–10 cycles of 98 °C for 30 s, 60 °C for 30 s, and 72 °C for one minute, ending with a final elongation step of 72 °C for five minutes. We purified samples using the Thermo Scientific MagJET Separation Rack and performed a final quantification using a Qubit fluorometer (Life Technologies). Libraries were sequenced across three lanes on the Illumina HiSeq 4000 platform at the Vincent J. Coates Genomics Sequencing Laboratory, California Institute for Quantitative Biosciences (http://qb3.berkeley.edu/gsl/Home.html).

2.2. Construction of species-specific reference genome

In order to improve the efficiency and accuracy of short-read mapping, and to reduce alignment bias to a divergent genome (Shafer et al., 2017), we sequenced the genome of a macaroni penguin, which, based on past genetic studies (Frugone et al., 2018; Cole et al., 2019), would be effectively used as a reference genome for both royal and macaroni penguin sequences. We prepared libraries for the genome sequencing of macaroni penguins using an Illumina TruSeq Nano kit following the manufacturer's instructions. In brief, 100 ng of genomic DNA was fragmented to 350 bp segments using an ultrasonicator. After cleaning with beads, fragmented DNA was treated with end repair mix and then with A-tailing to add an adenine to the 3′-end, to which indexing adapters were ligated. Ligated DNA fragments were amplified and purified with beads prior to quantification using a Qubit fluorometer. Library size was measured with an Agilent TapeStation (Agilent Technologies Inc). The library was sequenced to ~40× coverage with 150 paired reads using an Illumina HiSeq X Ten platform at MedGenome (USA).

To process raw sequence reads, exact duplicates were removed using SUPER DEDUPER (https://github.com/dstreett/Super-Deduper). Reads were then filtered using CUTADAPT (Martin, 2011) and TRIMMOMATIC (Bolger et al., 2014) to trim adapter contaminations and low-quality reads. Overlapping PE reads were merged using FLASH (Magoc and Salzberg, 2011). We then aligned the resulting cleaned reads to the genome (http://gigadb.org/dataset/100005) of an emperor penguin (*Aptenodytes forsteri*) (Jarvis et al., 2014; Li et al., 2014; Zhang et al., 2011; Zhang et al., 2014), using LAST (http://last.cbrc.jp/). The resulting alignment was converted to sorted BAM format using SAMTOOLS (Li et al., 2009). We used samtools mpileup, bcftools, vcfutils.pl vcf2fq and seqtk (https://github.com/lh3/seqtk) to convert alignments into FASTA format from the reference genome. These species-specific reference fasta sequences were evaluated for completeness by comparing them against the emperor penguin draft genome.

2.3. ddRAD data processing, SNP calling and filtering

We used a custom PERL pipeline encompassing various external programs for processing the ddRAD data. The pipelines are available at https://github.com/CGRL-QB3-UCBerkeley/RAD. Raw fastq reads were first de-multiplexed based on the sequences of internal barcodes with a tolerance of one mismatch. De-multiplexed reads were removed if the expected cutting sites were not found at the beginning of the sequences,

Table 1
Macaroni and royal penguin sample sizes, locations, and genetic diversity indices. Grouping schemes: (1) all macaroni and royal populations (2) royal penguins from Macquarie and Antarctic and sub-Antarctic populations of macaroni penguins. Values in parenthesis corresponds to the standard error of each estimate. F_{IS} refers to inbreeding coefficient, and "π" to nucleotide diversity. [†]denotes royal penguins from Macquarie Island.

Grouping scheme	Population	Geographic position	Sample size	Private alleles	Observed Heterozygosity	Expected Heterozygosity	F_{IS}	π
1	Elephant	61°13′S; 55°21′W	6	0	0.269 (0.003)	0.253 (0.002)	0.015 (0.000)	0.276 (0.002)
	Bird	54°00′S; 38°02′W	11	0	0.263 (0.002)	0.262 (0.002)	0.033 (0.002)	0.275 (0.002)
	Bouvet	54°25′S; 03°22′E	7	0	0.271 (0.003)	0.256 (0.002)	0.013 (0.000)	0.276 (0.002)
	Marion	46°50′S; 37°48′E	12	0	0.266 (0.002)	0.266 (0.002)	0.035 (0.002)	0.278 (0.002)
	Crozet	45°25′S; 50°24′E	12	0	0.269 (0.002)	0.265 (0.002)	0.019 (0.002)	0.276 (0.002)
	Kerguelen	49°42′S; 69°46′E	7	0	0.262 (0.003)	0.258 (0.002)	0.039 (0.000)	0.278 (0.002)
	Macquarie[†]	54°29′S; 158°56′E	7	0	0.246 (0.002)	0.252 (0.002)	0.061 (0.000)	0.271 (0.002)
2	Macquarie[†]		7	0	0.244 (0.002)	0.247 (0.002)	0.057 (0.000)	0.266 (0.002)
	Antarctic		24	11	0.260 (0.002)	0.265 (0.001)	0.035 (0.006)	0.271 (0.002)
	Sub-Antarctic		31	25	0.261 (0.002)	0.270 (0.001)	0.048 (0.007)	0.274 (0.001)

allowing for one mismatch. The reads were then filtered using CUTADAPT (Martin, 2011) and TRIMMOMATIC (Bolger et al., 2014) to trim adapter contaminations and low-quality reads. All reads were of 91 bp length. Quality of each samples was assessed using FASTQC (Andrews, 2019) and MULTIQC (Ewels et al., 2016) software. The resulting cleaned reads for each individual and each taxon were then aligned to the macaroni penguin reference genome, using the *bwa mem* algorithms (Li and Durbin, 2009) and we sorted and indexed bam files using SAMTOOLS (Li et al., 2009).

SNP calling was made using STACKS version 2.2 (Catchen et al., 2013; Rochette et al., 2019) under the Marukilow model, with an alpha threshold for discovering SNPs and calling genotypes set at 0.05. We removed three macaroni individuals (1 each from Elephant, Bouvet and Kerguelen Islands) that exhibited low sequencing-depth. A preliminary Principal Component Analysis (PCA) was performed using PLINK 1.9 (Purcell et al., 2007) to evaluate possible outlier individuals. We detected 12 macaroni penguin outliers from Bouvet and Elephant Islands (Supplementary Fig. S1). Genetic relatedness among individuals reconstructed via an Identity-by-State analysis using SNPRELATE (Zheng et al., 2012) in R 3.4.1 (R Core Team, 2017), showed an atypical pattern for these individuals. Each of the 6 outliers from Bouvet were intimately clustered with one outlier of Elephant, and each pair conformed a very distant branch from all other samples (Supplementary Fig. S2), suggesting that this pattern corresponded to a laboratory or sequencing artifact. Thus, we removed them for subsequent analyses and retained a total of 62 individuals (Table 1).

We used the *populations* program from STACKS (Catchen et al., 2013) to filter loci. We chose a minimum percentage (90%) of individuals in a given population to process each locus, eliminated loci that were not present in all populations, and chose a minimum allele frequency of 0.05. In addition, we also filtered sites exhibiting a deviation from Hardy-Weinberg equilibrium (HWE) when present in 50% of the populations. Since several population analyses, including the algorithm underlying the program STRUCTURE, require loci to be at linkage equilibrium (Pritchard et al., 2010; Pritchard et al., 2000), we used PLINK 1.9 (Purcell et al., 2007) to identify loci at linkage disequilibrium and filtered them using the blacklist in STACKS. We only retained sites that passed all the above filters for SNP calling, using two grouping options; (1) all geographic populations sampled, and (2) three populations corresponding to royal penguins, Antarctic populations of macaroni penguins and sub-Antarctic populations of macaroni penguins. The second grouping scheme was only used for exploratory analyses of genetic diversity indices.

In addition, as several analyses required a neutral dataset, we used BAYESCAN 2.1 (Foll and Gaggiotti, 2008) to evaluate if there were loci under selection, to remove them. We group individuals into Antarctic populations versus sub-Antarctic populations, as we found that they formed separated clusters (see results). Prior odds for these analyses were set to 10 and chose a q-value of 0.05 (False discovery rate analog to a *p*-value) to assess significance.

2.4. Genetic diversity, population differentiation and clustering

Using *populations* program in Stacks, we calculated genetic diversity indices for each population of macaroni and royal penguins, and also for (1) royal penguin, (2) sub-Antarctic and (3) Antarctic populations of macaroni penguins.

Pairwise F_{ST} was calculated among colonies using ARLEQUIN V. 3.5.1.2 (Excoffier and Lischer, 2010). Statistical significance of the estimates was calculated under 10,000 permutations. The *p*-value for pairwise F_{ST} between populations was corrected using a false discovery rate correction (Benjamini et al., 2006). Results were considered significant when $p < 0.05$. We also performed a Global Mantel test to determine if there was a correlation between genetic differentiation and geographical distance between colonies. The Global Mantel test was performed in R using the vegan package (Oksanen et al., 2019) using 10,000 permutations. We used previously calculated pairwise F_{ST} values and the minimum distance between colonies obtained using Google Earth Pro v. 7.3.2.5495. Because we found a strong pattern of isolation by distance and the main clusters found corresponded to Antarctic and sub-Antarctic regions (see results), we then evaluated, separately, the effect of temperature and distance over genetic differentiation of macaroni/royal penguins populations. This approach was carried in order to control the effect of distance and temperature, respectively, to avoid mistaking a pattern of IBD with a hierarchical structuring pattern (Meirmans, 2012). Temperature data were obtained from BIO-ORACLE 2.0 (Assis et al., 2018) using the temporal maximum temperatures from monthly climatologies (2002–2009) from SeaWiFS. Distance temperature matrices were constructed using the *dist* function in Vegan (Oksanen et al., 2019). To compare genetic, geographic and temperature dissimilarity, all distance matrices values were standardized (x-mean(x)/SD(x)). Then, we performed (1) a Partial Mantel test using genetic distance (i.e. F_{ST}) and geographical distance between colonies, using temperatures as covariates and, conversely, (2) we performed a Partial Mantel test using genetic distance and maximum temperatures using geographical distances as a covariate. Partial Mantel test were carried out ECODIST (Goslee and Urban, 2007) package in R. We also evaluated independence between geographic distance and temperature dissimilarity between colonies by performing a Mantel test before to evaluate the relative contribution of temperature and geographic distance over genetic differentiation. We then performed a Multiple Matrix Regression with Randomization analysis (MMRR) (Wang, 2013) using the package PopGenReport (Adamack and Gruber, 2014; Gruber and Adamack, 2015). Finally, we conducted a Mantel Test between the genetic distance and the combined effects of geographic distances and temperatures, weighted by their relative contribution on genetic differentiation as measured by MMRR analysis.

Potential genetic clusters were determined using both non-model and model-based approaches. For all analyses, the same dataset of 4577 SNPs was used (see results). Specifically, we used PLINK 1.9 (Purcell et al., 2007) and GGPLOT2 (Wickham, 2016) in R to calculate and visualize a PCA. We used ADEGENET (Jombart, 2008; Jombart and Ahmed, 2011) in R to evaluate the number of clusters with *K*-means and a Bayesian Information Criterion (BIC), using the function find.clusters. In ADEGENET we also performed a discriminant analysis of principal components (DAPC) (Jombart et al., 2010). As the most likely number of clusters (*K*) identified by *K*-means was 1 (see results), we grouped individuals by their colonies for DAPC analysis (see Clucas et al., 2018). In addition, we performed a Bayesian clustering method implemented in STRUCTURE (Falush et al., 2003; Pritchard et al., 2000). All runs were made using PARALLELSTRUCTURE (Besnier and Glover, 2013) in R, not including the sampling location of each individual as a prior. Given that high levels of genetic similarity were previously detected between populations, suggesting that sampled individuals may have ancestors from multiple populations and that populations are closely related (Porras-Hurtado et al., 2013), we chose the admixture ancestry model and correlated allele frequencies options (Falush et al., 2003). To infer lambda, we first performed one run with 150,000 iterations and 50,000 Burn-in. This run calculated a lambda of 0.71, which we used in the subsequent structure runs. We evaluated $K = 1$–7, which corresponded to all sampled colonies. We performed 10 independent runs for each value of *K*, with 500,000 MCMC and a 50,000 burn-in period and we estimated the 90% probability intervals for admixture coefficients. The web version of STRUCTURE HARVESTER (Earl and vonHoldt, 2012) (http://taylor0.biology.ucla.edu/structureHarvester/) was used for inferring the most likely *K* using the Evanno's method (Evanno et al., 2005) and we also checked the highest posterior mean log-likelihood (mean LnP(K)). Finally, CLUMPP (Jakobsson and Rosenberg, 2007) was used for summarizing the results of all previous runs and DISTRUCT V. 1.1 (Rosenberg, 2003) was used for to visualize the results.

To determine whether geographic or taxonomic clustering better

explained the genetic variation in our data, we performed several Analyses of Molecular Variance (AMOVA) in ARLEQUIN, using 10,000 permutations. The AMOVA groups were defined as: (1) all populations of macaroni penguins against Macquarie populations of royal penguins; (2) Antarctic populations of macaroni penguins (Elephant Island, Bird Island, Bouvet Islands) against sub-Antarctic populations of macaroni and royal penguins (Marion Island, Crozet Islands, Kerguelen Islands, Macquarie Island), and (3) three groups consisting of the Antarctic populations of macaroni penguins, sub-Antarctic populations of macaroni penguins and royal penguins from Macquarie Island.

3. Results

3.1. ddRAD data processing, SNP calling and filtering

The consensus reference of macaroni penguins exhibited a mean coverage of 28×, and contained 18,969 scaffolds and N50 of 5,071,598 length. In the ddRAD data, the effective per-sample coverage mean was 72.8×, the Standard Deviation was 18.4×, the minimum coverage was 24.0×, the maximum coverage was 109.5× and the mean number of sites per locus was 90.6. We removed the eight loci that were out of HWE in at least 50% of the populations, and 695 loci that were in linkage disequilibrium. We did not detect any loci under significant selection (Supplementary Fig. S3a and b). Following filtering, we retained 4577 SNPs in the dataset comprising all sampled colonies, and 5429 SNPs in the dataset comprising royal penguins, Antarctic macaroni penguins and sub-Antarctic macaroni penguins.

3.2. Population differentiation and clustering

Observed and Expected heterozygosity, F_{IS}, and nucleotide diversity (π) were similar among colonies (Table 1). When evaluating all colonies separately, no private alleles were detected. However, when contrasting macaroni colonies from (1) the Antarctic, (2) sub-Antarctic with the third group composed by royal penguins, we found 11 and 25 private alleles, respectively, while royal penguins did not exhibit private alleles (Table 1). Pairwise F_{ST} values were very low, although significantly different from 0 for most comparisons (Fig. 2, Supplementary Table S2a, b and c). F_{ST} ranged between 0 and 0.024. The highest F_{ST} values (0.024) corresponded to comparisons of macaroni penguins sampled from Bird and Elephant Islands with royal penguins sampled from Macquarie Island. In contrast, the lowest F_{ST} values were recovered among the Antarctic islands: Bouvet, Elephant and Bird Islands (Elephant-Bouvet not significant). Among the sub-Antarctic colonies, we recovered low F_{ST} values between Marion, Crozet and Kerguelen Islands. Pairwise comparisons involving populations from Antarctic and sub-Antarctic colonies had higher values than comparisons within each region; however, they were still low (around 0.01). For all macaroni/royal populations considered together, the Global Mantel test revealed a pattern of isolation by distance (IBD), with a significant correlation between distance and F_{ST} calculations (Mantel r = 0.828, p = 0.001, Supplementary Fig. 4). However, an inverse relationship between genetic and geographic distances was observed among Antarctic localities (Supplementary Fig. 4). Nevertheless, this inverse IBD relationship was not significant (Mantel r = −0.82, p = 0.16), as only three localities were compared. Furthermore, Partial Mantel tests were highly significant for both geographical distance controlled by temperature dissimilarity (Mantel r = 0.79, p = 0.0003; Fig. 3a) and temperature controlled by geographic distance (Mantel r = 0.48, p = 0.007; Fig. 3b). The correlation coefficient between geographic distance and sea surface temperature dissimilarity was not significantly different from 0 (Mantel r = 0.40, p = 0.054, Fig. 3c), although very close to the alpha threshold. This suggests that both variables are not correlated or only weakly correlated and thus, it meets the assumptions of MMRR method (Wang, 2013). The multiple regression coefficient for geographic distance (β = 0.7, p = 0.0005) was about two times higher than the coefficient for temperature (β = 0.299, p = 0.044), suggesting that geographic distance explained more of the genetic variation, however, temperature dissimilarity also contributed significantly. The Mantel test for genetic distances using the combined effect of geographic distance and temperature distance, yielded the best correlation model (Mantel r = 0.87, p = 0.0004, Fig. 3d).

In the PCA (Fig. 4) three groups can be identified, one corresponding to Antarctic populations, a second corresponding to sub-Antarctic populations and a third composed by royal penguins and some macaroni penguins from Kerguelen. However, Kerguelen individuals were also present in the sub-Antarctic group.

When evaluating the numbers of clusters with *K*-means, the smallest number of BIC was found for one cluster, although BIC values were similar for K = 1, K = 2 and K = 3 (Supplementary Fig. S5a). When choosing two clusters, one corresponded to colonies from the Antarctic region, and the other corresponded to colonies from the sub-Antarctic region (Supplementary Fig. S5b). When three clusters were selected, one corresponded to Antarctic populations, the second to sub-Antarctic populations (not including royal penguins) and the third included royal penguins and some macaroni individuals from Kerguelen (Supplementary Fig. S6a). However, the discriminant functions from the last two groups (sub-Antarctic and royal-Kerguelen individuals) suggest that they conform a single group (Supplementary Fig. S6b).

Using prior locations accounting for all sampled colonies, DAPC (Fig. 5a) also revealed two well-differentiated groups discriminated by the first component (horizontal axis). Also, in the DAPC analyses royal penguins were close to the cluster formed by macaroni penguins from Kerguelen Island, whereas other individuals sampled from Kerguelen fall in the sub-Antarctic group constituted by penguins from Marion and Crozet Islands. These results are in accordance with PCA results.

For structure analysis, the highest posterior mean log-likelihood (mean LnP(K)) and the Evanno's method suggest that the most likely K was K = 2 (Supplementary Fig. S7a, b and Supplementary Table S3). The group composition for each k was consistent with the results from

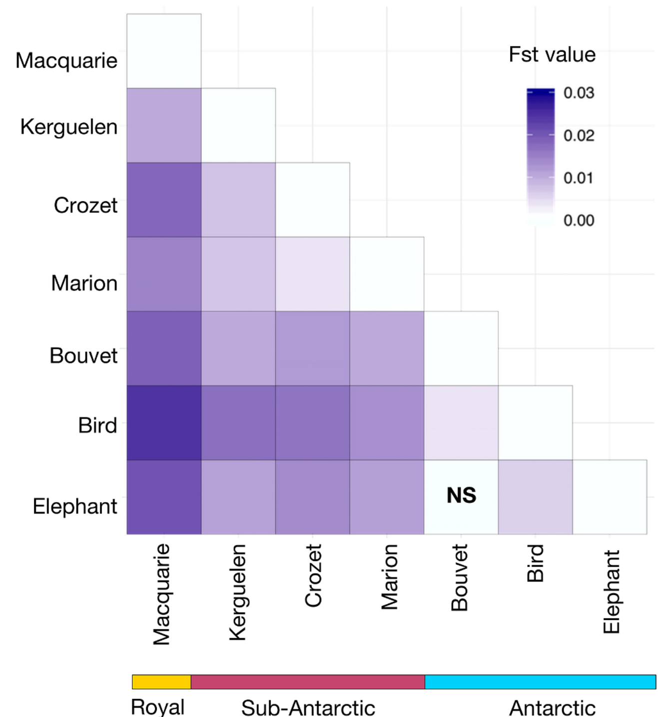


Fig. 2. Pairwise F_{ST} between populations of royal and macaroni penguins. NS denotes not significant values. The bar below the heatmap indicates which populations are from royal penguins and from sub-Antarctic and Antarctic populations of macaroni penguins.

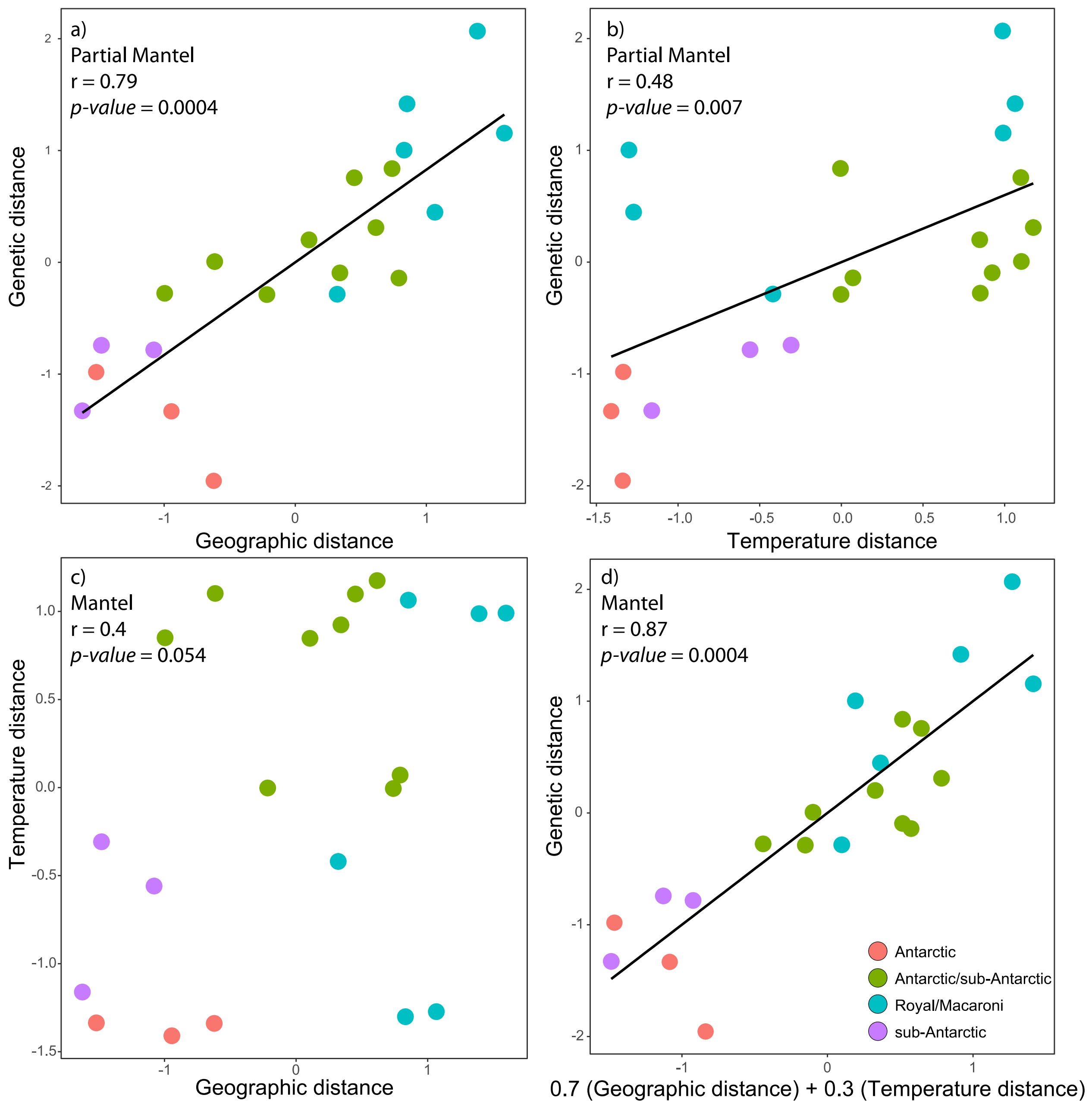


Fig. 3. Scatter plots showing the relationship of (a) isolation-by-distance, (b) isolation-by-temperature, (c) temperature and geographic distance and (d) joint effect of geographic distance and temperature based on the results of a Multiple Matrix Regression with Randomization analysis (MMRR). Regression lines drawn only for significant correlations (mantel $p < 0.05$).

previous analyses: one group included Elephant, Bird and Bouvet Islands, and the other group included Marion, Crozet, Kerguelen and Macquarie (royal penguins) Islands (Fig. 5b). These groups were clearly identified even when visualizing plots for $K > 2$ (Supplementary Fig. S8). Estimated admixture coefficients were low overall between the clusters. In the Antarctic region, five individuals (2 from Elephant island and 3 from Bouvet) and only two individuals from sub-Antarctic region (Marion and Kerguelen) showed an admixture coefficient with a 90% probability interval that did not overlap with zero or one. AMOVA (Supplementary Table S4) indicated that only 0.75% of the observed genetic variation among the individuals sampled was explained by differences between royal and macaroni penguin colonies ($F_{CT} = 0.007$, $p = 0.144$) whereas 1.0% of genetic variation was explained by differences among populations within groups ($F_{SC} = 0.010$, $p = 0.000$). When we performed AMOVA analysis between Antarctic and sub-Antarctic colonies (encompassing both macaroni and royal penguins), the genetic variation explained among groups was slightly higher (0.8%, $F_{CT} = 0.007$, $p = 0.027$), while 0.74% of the genetic variation was explained by differentiation among colonies within groups ($F_{SC} = 0.007, p = 0.001$). Highest among group component occurred in three groups (Antarctic macaroni penguins, sub-Antarctic macaroni penguins and Macquarie royal penguins), but only explained 1.01% of the whole genetic variation ($F_{CT} = 0.01$, $p = 0.015$), while 0.49% was explained by differentiation among colonies within groups ($F_{SC} = 0.004$, $p = 0.001$). Most of the genetic variation 98.5% was explained by differences among individual, indicative of weak population genetic structure.

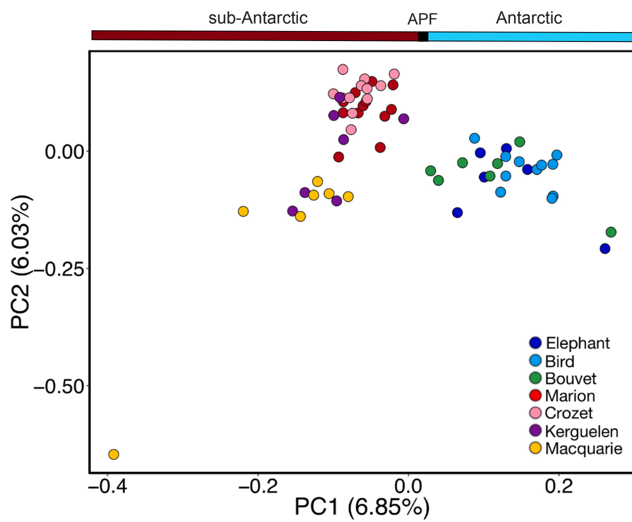


Fig. 4. PCA analysis for royal and macaroni penguin populations. Royal penguins corresponds to Macquarie population. The bar above indicate which populations are on the Antarctic and sub-Antarctic region and the position of the Antarctic Polar Front (APF).

4. Discussion

In this study, we evaluated the genetic differentiation between royal and macaroni penguins and the population structure of macaroni penguins across their distributional range. The genetic differentiation between royal and macaroni penguins was low and comparable to genetic differentiation among macaroni penguin populations. In general, we found low levels of genetic differentiation between populations of macaroni/royal penguins across their distributional range, however, geographic distance and differences in sea surface temperatures may be limiting gene flow to some degree. Moreover, two genetic groups were detected, one corresponding to Antarctic and the other to sub-Antarctic populations.

4.1. Differentiation between royal and macaroni penguins

Recent studies using short mitochondrial and nuclear sequences, and whole mitogenomes (Cole et al., 2019, 2018a; Frugone et al., 2018) have been unsuccessful in recovering reciprocal monophyly between royal and macaroni penguins. The authors of these studies suggest that the lack of genetic divergence between the two species may have been linked to recent isolation of royal penguins on the geologically young Macquarie Island where incomplete lineage sorting across the genome may underlie the sharing of alleles. Differentiation and quick fixation of morphological traits, such as facial coloration and beak size, may have

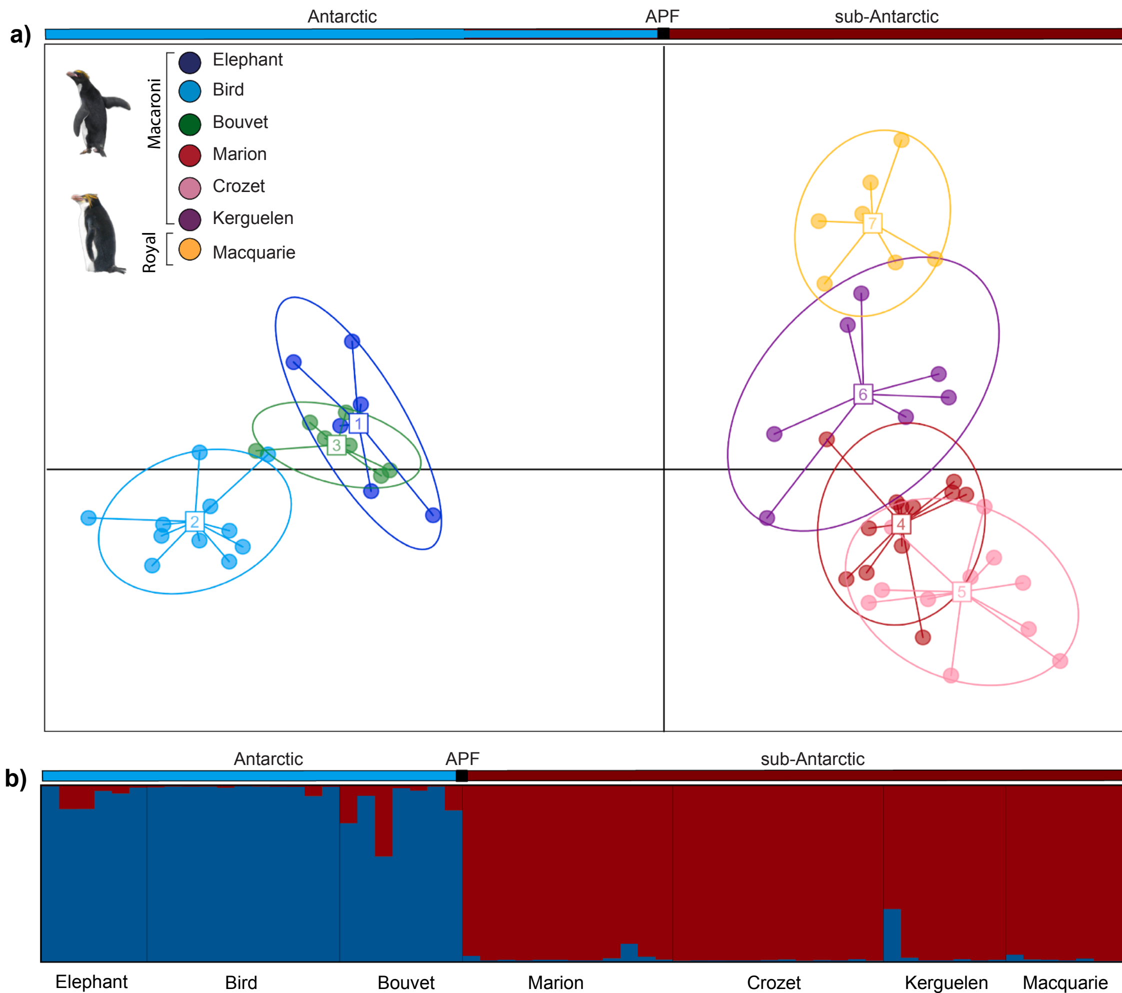


Fig. 5. Clustering analyses. (a) DAPC analysis for royal and macaroni penguin populations and (b) Structure results for analyses involving populations of macaroni and royal penguins. The analyses favored K = 2, with the two major genotypic clusters corresponding to Antarctic and Sub-Antarctic populations. The bar above indicate which populations are on the Antarctic and sub-Antarctic region and the position of the Antarctic Polar Front (APF).

occurred through a founder event or through selective pressures (e.g. sexual selection), thereby explaining the discordance between morphological traits and molecular markers. Similarly, a previous study on penguin diversification demonstrated that three *Megadyptes* subspecies, including two extinct lineages, only diverged from each other within the past half a million years, even though they are morphologically distinct (Cole et al., 2019). Other avian studies have suggested that morphological traits such as coloration could evolve much more rapidly than some genetic markers (Zink et al., 2005). Nevertheless, it is expected that reduced representation techniques, such as ddRAD methods which generate thousands of independent SNPs, should provide enhanced resolution to detect even subtle genetic differentiation (Lavinia et al., 2019; Mason and Taylor, 2015). However, in this study, using more than four thousand neutral SNP markers, we were unable to detect genetic evidence for considering royal and macaroni penguins as separate species. Instead, we found very low levels of population differentiation between royal and macaroni colonies; e.g < 1% of the total genetic variation was explained by the separation of both species (Supplementary Table S4). Secondly, our clustering analyses, including model and non-model based methods, failed to discriminate royal penguins as a separate cluster from macaroni penguins. Moreover, differentiation between Antarctic and sub-Antarctic populations of macaroni penguins was higher than differentiation between royal and macaroni penguins. For this reason, the degree of genetic similarity between royal and macaroni penguins does not support the presence of historical reproductive isolation and suggests that the two taxa should be considered the same species. Even so, royal penguins exhibited a tendency to cluster only with some macaroni penguins from Kerguelen and the variance explained by AMOVA was slightly higher when separating royal penguins as a third group. This may indicate an incipient speciation process with gene flow, as observed in other species (Árnason et al., 2018; Figueiró et al., 2017; Lavinia et al., 2019; Veale et al., 2018). Given that we evaluated genetic differentiation from a subset of the complete genome rather than using the complete genome, it is also possible that our sampling may have been insufficient to detect more differentiated loci (Campagna et al., 2017; Irwin et al., 2018; Kim et al., 2019; Toews et al., 2016). Future studies using whole genomes are needed in order to evaluate these considerations.

Overall, our results suggest that contemporary gene flow exists between macaroni and royal penguins. Reports of royal penguins (or white-faced penguins) on the Falkland/Malvinas islands (Dehnhard et al., 2012) and on Marion, Crozet and Kerguelen islands (García and Boersma, 2013) may be indicative of occasional dispersal. Furthermore, there are reports of macaroni penguins on Macquarie Island (Warham, 1971). The present study found that royal and macaroni penguins from Kerguelen Islands exhibited a tendency to form a genetic cluster, which is consistent with a previous study using mtDNA (Frugone et al. 2018) in which haplotypes (mtDNA HVR1) from royal penguins formed a separated haplogroup that included haplotypes from Kerguelen. These results suggest that gene flow between royal and macaroni penguins may occur almost exclusively between individuals from Macquarie and Kerguelen Islands, possibly in association with their geographic proximity, as suggested by the Mantel test.

Considering macaroni and royal penguins as a single evolutionary unit, hereby referred as "macaroni/royal", the maintenance of a white-faced phenotype at Macquarie Island remains unclear. Facial coloration is determined by the production and distribution of melanic pigments in feathers and is controlled by a number of recessive and dominant genes (Ng and Li, 2018; Van Grouw, 2006). It may be the case that F1 hybrids between black and white-faced penguins should retain a dark face phenotype. In this context, a marked asymmetrical migration from Macquarie Island to other colonies may explain the maintenance of the white-face phenotype on this island. If migration to Macquarie were uncommon, the black-face phenotype would not readily spread, even when taking into account the dominance of the black-faced phenotype. This hypothesis could not be tested in our study through the classical approach, such as by using the program BayesAss (Wilson and Rannala, 2003), as the low levels of genetic differentiation we observed may lead to inaccurate results when studying migration rates (Faubet et al., 2007; Meirmans, 2014). However, a comparison of the proportion of private alleles among colonies supports our hypothesis of asymmetrical gene flow, as penguins from Macquarie Island do not exhibit any private alleles in contrast with macaroni penguins from both Sub-Antarctic and Antarctic colonies (Table 1).

The use of plumage coloration to delimit species has yielded mixed results (Mason and Taylor, 2015; Ng and Li, 2018; Paxton, 2009). In agreement with our results, other researchers have found a lack of congruence between coloration (confined to a particular geographic region) and molecular data, in which phenotypes do not form a reciprocally monophyletic group (e.g., Rocha-Mendez et al., 2018). These cases highlight the importance of using molecular markers in conjunction with different traits for the purpose of defining species and evaluating the underlying evolutionary history in birds (Lavinia et al., 2019).

4.2. Genetic structure and genetic differentiation among macaroni/royal colonies

When we considered all macaroni and royal penguin colonies together, we recovered limited population structure across much of the Southern Ocean, stretching from Macquarie Island to the South Shetland Islands. High dispersal capabilities and low levels of population differentiation are common in most penguin species, except in gentoo (Clucas et al., 2018; Vianna et al., 2017) and possibly in rockhopper (*Eudyptes chrysocome*, *E. filholi* and *E. moseleyi*) penguins, which may be restricted by their dispersal capabilities (Frugone et al., 2018). In a recent comparative genomic study, Clucas et al. (2018) reported F_{ST} values for king (*Aptenodytes patagonicus*), emperor, chinstrap (*Pygoscelis antarcticus*) and Adélie penguins (*P. adeliae*) ranging from 0.002 to 0.008, despite colonies being separated by thousands of kilometers. Such low genetic structure could be related to high dispersal capabilities in these species, since few effective migrants could be responsible of genetic homogeneity (Cristofari et al., 2016). During the inter-breeding period, Adélie, chinstrap and emperor penguins travel long distances, reaching mean maximum distances of 2235 km (Dunn et al., 2011), 3900 km (Hinke et al., 2015) and > 2840 km from their respective breeding colonies (Kooyman et al., 1996; Thiebot et al., 2013; Wienecke et al., 2010). Similar distances have been reported for macaroni penguins during the inter-breeding period (mean maximum distance from the colony of 1778 ± 902 km, Thiebot et al. (2011)), but the F_{ST} values among macaroni/royal colonies exceeded the values for the species mentioned above by an order of magnitude, and almost all F_{ST} pairwise comparisons were significantly different from 0. Moreover, we found a strong pattern of isolation by distance detected across the whole macaroni/royal distribution, possibly driven by large geographic distance between colonies (Fig. 3), whereas no correlation between geographic distance and genetic differentiation has been recovered for Adélie, chinstrap and emperor penguins (Clucas et al., 2018).

Higher genetic structure in macaroni/royal penguins compared to former penguin species may be associated with variation in natal philopatry or with the establishment of new breeding sites. For example, emperor penguins sometimes change breeding sites from year to year and relocate when severe environmental disturbances have occurred (Ancel et al., 2014; Larue et al., 2015). Similar relocations have been observed in Adélie penguins (Forcada and Trathan, 2009). In both species, recolonization events may involve individuals from different colonies, promoting genetic homogenization. This proposition is supported by the subtle genetic differentiation in emperor penguins (Younger et al., 2017), and by calculations of the percentage (0.7–10% of Ne) of migrants of emperor penguins that could, theoretically, reach a certain colony each generation (Cristofari et al., 2016). Both previously mentioned studies include different views, however, they seem

to support the hypothesis that dispersal (and gene flow) between colonies may be a relatively common event in emperor penguins.

Technical constraints may make it difficult to obtain accurate estimates of philopatry in penguins. However, reported nest and mate fidelity of 69–87% and 70–73%, respectively (Williams and Rodwell, 1992), may suggest that a high proportion of macaroni penguin individuals are philopatric. Outside of the breeding period, macaroni penguins also exhibit site fidelity to their foraging areas (Thiebot et al., 2011). However, natal philopatry may not arise for all individuals within a population, and a proportion of individuals may migrate to other nearby colonies, as suggested by the pattern of isolation by distance. Indeed, the lowest genetic differentiation (F_{ST}) was between the nearest populations from Bouvet and Elephant islands, and the highest between royal penguin from Macquarie and macaroni penguins from Kerguelen islands. The pattern of isolation by distance may also indicate that despite potentially high dispersal capabilities (1778 ± 902 km; Thiebot et al. (2011)), there could be a distance-related restriction over migration for each colony across their entire range of macaroni/royal penguins.

4.3. Environmental driver of differentiation between Antarctic and sub-Antarctic provinces

In addition to the isolation by distance model evidenced among royal/macaroni colonies, we consistently detected two genetic clusters corresponding to the Antarctic and sub-Antarctic provinces, suggesting that differences in environmental characteristics of both regions may be involved in their genetic structure. The Antarctic and sub-Antarctic waters are delimitated by the Antarctic Polar Front (APF) in which an abrupt change temperature and salinity occurs. Also, the APF coincides with the boundaries of several species distributions, which are restricted to either the Antarctic or sub-Antarctic provinces (De Broyer et al., 2014) and have led to the proposition that the APF acts as a barrier to dispersal (Fraser et al., 2012; González-Wevar et al., 2017; Poulin et al., 2014). Therefore, sharp variation of environmental parameters at the AFP may be limiting gene flow between macaroni penguins from each biogeographical region and then contribute to the genetic structure of the species. As a matter of fact, we found that genetic differentiation was effectively associated to differences in sea surface temperatures between colonies, even if to a lesser extent than with the geographical distance.

The role of oceanic fronts as a barrier to dispersal in penguins has been suggested for species exhibiting both high or limited dispersal capabilities and different foraging behaviors, such as king and gentoo penguins (Clucas et al., 2018; Clucas et al., 2016; Levy et al., 2016; Vianna et al., 2017). Moreover, speciation and diversification of rockhopper penguins (*E. moseleyi*, *E. chrysocome* and *E. filholi*) has been explained by the presence of the Subtropical Front (STF) separating their colonies (Fig. 1), and by the latitudinal shift of the STF during past glacial periods (de Dinechin et al., 2009; Frugone et al., 2018). This suggests that oceanic fronts play an important role in the evolutionary histories of penguins. For testing this hypothesis on macaroni penguins populations, new sampling locations of breeding colonies from South America may allow further evaluating whether the APF is limiting dispersal more than geographic distance between the colonies.

Finally, other factors may influence dispersal between the Antarctic and sub-Antarctic groups, possibly in association with differences in environmental conditions at the breeding grounds within each region (Fraser et al., 2016; Moon et al., 2017). Other cases of genetic structuring associated with biogeographic breaks in highly dispersive taxa had shown that local adaptation to distinct environmental conditions, such as differences in resource use, could explain genetic differentiation (Pérez-Alvarez et al., 2015). This could be also the case of macaroni penguin populations inhabiting Antarctic and sub-Antarctic regions, as they experience different biotic and abiotic environmental conditions. Macaroni penguins present inter-annual fidelity to foraging areas (Thiebot et al., 2011), and that may reflect specialization in resource use and foraging strategies.

4.4. Conclusions and conservation implications

The main conclusion of our study is that genetic differentiation between macaroni and royal penguins is extremely low, and comparable to that found among macaroni colonies alone. Mainly because of its endemism to Macquarie Island and nearby Bishop and Clerk Islands, the International Union for Conservation (IUCN) has classified royal penguins as near threatened (BirdLife_International, 2018). While our results suggest that macaroni and royal penguins could be a single species, we suggest that conservation efforts should continue to consider the two taxa as separate management units. Variation in plumage coloration and morphology are often related to mate choice and may promote speciation over time (Ng and Li, 2018). Moreover, diversification with ongoing gene flow still remains as a possible explanation for the limited genetic differences found between royal and macaroni penguins, but given the genetic similarity between the two taxa, whole genome data will be required to test this possibility. Furthermore, the distribution of private alleles suggests the existence of asymmetrical gene flow from Macquarie to the other colonies. Macquarie Island royal penguins may not receive migrants from other localities, thus increasing the vulnerability of this population in terms of persistence over time.

In addition, our results suggest that the main factors associated with population differentiation in macaroni/royal penguins correspond to distance between colonies and sea surface temperatures. In general, the low levels of differentiation found among macaroni/royal penguin colonies are similar to that recovered for other penguin species, except in the cases of gentoo and possibly rockhopper penguins, which may be more constrained by their dispersal capabilities. Even so, we found that macaroni/royal penguins exhibited more genetic differentiation compared to other highly dispersive penguin species. This may be explained by higher levels of natal philopatry in macaroni penguins.

5. Data accessibility

DNA Sequences for ddRAD data: SRA accession: PRJNA523574, BioProject: PRJNA523574.

Accession number for the genome of Macaroni penguin BioSample: SAMN11566613, BioProject: PRJNA530615.

Additional information and files available on Figshare DOI: https://doi.org/10.6084/m9.figshare.c.4548920.

Author contributions

MJF wrote the paper. MJF, MEL, NS and KB performed analyses and created the figures. JAV and EP contributed to every analytical step regarding the interpretation of results and in preparing the manuscript. RCKB helped design the study and contributed to the manuscript; SNP data collection was performed in his lab and CYWC provided laboratory assistance. TLC and JMW discussed results and contributed to the manuscript. AL, PP, GPMD, MVP, FB, KP, AP and BW undertook field work, provided samples and contributed to the manuscript. All co-authors revised and commented on the manuscript and approved the final version.

Acknowledgements

This study received funding support from several institutions. We are grateful to Instituto Antártico Chileno [DT-11_17 and INACH RT_12–14]; Comisión Nacional de Ciencia y Tecnología de Chile [21160387]; FONDECYT [1150517 and 1151336]; CONICYT PIA [ACT172065 GAB and AFB170008 IEB]; French Polar Institute (IPEV) [ETHOTAAF 354]; INCT-APA from Brazil (CNPq and FAPERJ); The

Council National Research from Brazil [CNPq 574018/2008-5, CNPq 442702/2018-3]. Logistical support was provided by the Brazilian Secretary of the Inter-Ministry on Marine Resources (SECIRM) and Armada de Chile.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.106563.

References

Adamack, A.T., Gruber, B., 2014. PopGenReport: simplifying basic population genetic analyses in R. Meth. Ecol. Evol. 5, 384–387. https://doi.org/10.1111/2041-210X.12158.

Aljanabi, S.M., Martinez, I., 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucl. Acids Res. 25, 4692–4693. https://doi.org/10.1093/nar/25.22.4692.

Ancel, A., Cristofari, R., Fretwell, P.T., Trathan, P.N., Wienecke, B., Boureau, M., Morinay, J., Blanc, S., Le Maho, Y., Le Bohec, C., 2014. Emperors in hiding: When ice-breakers and satellites complement each other in antarctic exploration. PLoS One 9, e100404. https://doi.org/10.1371/journal.pone.0100404.

Andrews, S., 2019. FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Árnason, Ú., Lammers, F., Kumar, V., Nilsson, M.A., Janke, A., 2018. Whole-genome sequencing of the blue whale and other rorquals finds signatures for introgressive gene flow. Sci. Adv. 4, eaap9873. https://doi.org/10.1126/sciadv.aap9873.

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E.A., De Clerck, O., 2018. Bio-ORACLE v2.0: extending marine data layers for bioclimatic modelling. Glob. Ecol. Biogeogr. 27, 277–284. https://doi.org/10.1111/geb.12693.

Benjamini, Y., Krieger, A.M., Yekutieli, D., 2006. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93, 491–507. https://doi.org/10.1093/biomet/93.3.491.

Bertelli, S., Giannini, N.P., 2005. A phylogeny of extant penguins (Aves : Sphenisciformes) combining morphology and mitochondrial sequences. Cladistics 21, 209–239. https://doi.org/10.1111/j.1096-0031.2005.00065.x.

Besnier, F., Glover, K.A., 2013. ParallelStructure: A R package to distribute parallel runs of the population genetics program STRUCTURE on multi-core computers. PLoS One 8, e70651. https://doi.org/10.1371/journal.pone.0070651.

BirdLife_International, 2018. Eudyptes schlegeli. The IUCN Red List of Threatened Species 2018: e.T22697797A132603136. https://doi.org/10.2305/IUCN.UK.2018-2.RLTS.T22697797A132603136.en (downloaded on 27 February 2).

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Campagna, L., Repenning, M., Silveira, L.F., Fontana, C.S., Tubaro, P.L., Lovette, I.J., 2017. Repeated divergent selection on pigmentation genes in a rapid finch radiation. Sci. Adv. 3, e1602404. https://doi.org/10.1126/sciadv.1602404.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140. https://doi.org/10.1111/mec.12354.

Christidis, L., Boles, W.E., 2008. Systematics and Taxonomy of Australian Birds. CSIRO Publishing, Collingwood.

Clucas, G.V., Younger, J.L., Kao, D., Emmerson, L., Southwell, C., Wienecke, B., Rogers, A.D., Bost, C.A., Miller, G.D., Polito, M.J., Lelliott, P., Handley, J., Crofts, S., Phillips, R.A., Dunn, M.J., Miller, K.J., Hart, T., 2018. Comparative population genomics reveals key barriers to dispersal in Southern Ocean penguins. Mol. Ecol. 27, 4680–4697. https://doi.org/10.1111/mec.14896.

Clucas, G.V., Younger, J.L., Kao, D.M., Rogers, A.D., Handley, J., Miller, G.D., Jouventin, P., Nolan, P., Gharbi, K., Miller, K.J., Hart, T., 2016. Dispersal in the sub-Antarctic: king penguins show remarkably little population genetic differentiation across their range. BMC Evol. Biol. 16. https://doi.org/10.1186/s12862-016-0784-z.

Cole, T.L., Ksepka, D.T., Mitchell, K.J., Tennyson, A.J.D., Thomas, D.B., Pan, H., Zhang, G., Rawlence, N.J., Wood, J.R., Bover, P., Bouzat, J.L., Cooper, A., Fiddaman, S., Hart, T., Miller, G., Ryan, P.G., Shepherd, L.D., Wilmshurst, J.M., Waters, J.M., 2019. Mitogenomes uncover extinct penguin taxa and reveal island formation as a key driver of speciation. Mol. Biol. Evol. 36, 784-707. https://doi.org/10.1093/molbev/msz017.

Cole, T.L., Rawlence, N.J., Dussex, N., Ellenberg, U., Houston, D.M., Mattern, T., Miskelly, C.M., Morrison, K.W., Scofield, R.P., Tennyson, A.J.D., Thompson, D.R., Wood, J.R., Waters, J.M., 2018a. Ancient DNA of crested penguins: testing for temporal genetic shifts in the world's most diverse penguin clade. Mol. Phylogenet. Evol. 131, 72–79. https://doi.org/10.1016/j.ympev.2018.10.025.

Cole, T.L., Waters, J.M., Shepherd, L.D., Rawlence, N.J., Joseph, L., Wood, J.R., 2018b. Ancient DNA reveals that the 'extinct' Hunter Island penguin (Tasidyptes hunteri) is not a distinct taxon. Zool. J. Linn. Soc. 182, 459–464. https://doi.org/10.1093/zoolinnean/zlx043.

Cristofari, R., Bertorelle, G., Ancel, A., Benazzo, A., Le Maho, Y., Ponganis, P.J., Stenseth, N.C., Trathan, P.N., Whittington, J.D., Zanetti, E., Zitterbart, D.P., Le Bohec, C., Trucchi, E., 2016. Full circumpolar migration ensures evolutionary unity in the emperor penguin. Nat. Commun. 7, 11842. https://doi.org/10.1038/ncomms11842.

De Broyer, C., Koubbi, P., Griffiths, H.J., Raymond, B., Udekem d'Acoz, C.d., Van de Putte, A.P., Danis, B., Grant, S., Gutt, J., Held, C., Hosie, G., Huettmann, F., Post, A., Ropert-Coudert, Y., 2014. Biogeographic Atlas of the Southern Ocean. Scientific Committee on Antarctic Research, Cambridge.

de Dinechin, M., Ottvall, R., Quillfeldt, P., Jouventin, P., 2009. Speciation chronology of rockhopper penguins inferred from molecular, geological and palaeoceanographic data. J. Biogeogr. 36, 693–702. https://doi.org/10.1111/j.1365-2699.2008.02014.x.

Dehnhard, N., Ludynia, K., Almeida, A., 2012. A royal penguin Eudyptes schlegeli in the Falkland islands? Mar. Ornithol. 40, 95–98.

Dunn, M.J., Silk, J.R.D., Trathan, P.N., 2011. Post-breeding dispersal of Adelie penguins (*Pygoscelis adeliae*) nesting at Signy Island, South Orkney Islands. Polar Biol. 34, 205–214. https://doi.org/10.1007/s00300-010-0870-4.

Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361. https://doi.org/10.1007/s12686-011-9548-7.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

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. Mol. Ecol. Resour. 10, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x.

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

Faubet, P., Waples, R.S., Gaggiotti, O.E., 2007. Evaluating the performance of a multi-locus Bayesian method for the estimation of migration rates. Mol. Ecol. 16, 1149–1166. https://doi.org/10.1111/j.1365-294X.2007.03218.x.

Figueiró, H.V., Li, G., Trindade, F.J., Assis, J., Pais, F., Fernandes, G., Santos, S.H.D., Hughes, G.M., Komissarov, A., Antunes, A., Trinca, C.S., Rodrigues, M.R., Linderoth, T., Bi, K., Silveira, L., Azevedo, F.C.C., Kantek, D., Ramalho, E., Brassaloti, R.A., Villela, P.M.S., Nunes, A.L.V., Teixeira, R.H.F., Morato, R.G., Loska, D., Saragüeta, P., Gabaldón, T., Teeling, E.C., O'Brien, S.J., Nielsen, R., Coutinho, L.L., Oliveira, G., Murphy, W.J., Eizirik, E., 2017. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. Sci. Adv. 3, e1700299. https://doi.org/10.1126/sciadv.1700299.

Foll, M., Gaggiotti, O., 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180, 977–993. https://doi.org/10.1534/genetics.108.092221.

Forcada, J., Trathan, P.N., 2009. Penguin responses to climate change in the Southern Ocean. Glob. Chang. Biol. 15, 1618–1630. https://doi.org/10.1111/j.1365-2486.2009.01909.x.

Fraser, C.I., Kay, G.M., Plessis, M.d., Ryan, P.G., 2016. Breaking down the barrier: dispersal across the Antarctic Polar Front. Ecography 40, 235–237. https://doi.org/10.1111/ecog.02449.

Fraser, C.I., Nikula, R., Ruzzante, D.E., Waters, J.M., 2012. Poleward bound: biological impacts of Southern Hemisphere glaciation. Trends Ecol. Evol. 27, 462–471. https://doi.org/10.1016/j.tree.2012.04.011.

Friesen, V.L., Burg, T.M., Mccoy, K.D., 2007. Mechanisms of population differentiation in seabirds. Mol. Ecol. 16, 1765–1785. https://doi.org/10.1111/j.1365-294X.2006.03197.x.

Frugone, M.J., Lowther, A., Noll, D., Ramos, B., Pistorius, P., Dantas, G.P.M., Petry, M.V., Bonadonna, F., Steinfurth, A., Polanowski, A., Raya Rey, A., Lois, N.A., Pütz, K., Trathan, P., Wienecke, B., Poulin, E., Vianna, J.A., 2018. Contrasting phylogeographic pattern among *Eudyptes* penguins around the Southern Ocean. Sci. Rep. 8, 17481. https://doi.org/10.1038/s41598-018-35975-3.

García, P., Boersma, D., 2013. Penguins: Natural History and Conservation. University of Washington Press.

Garg, K.M., Tizard, R., Ng, N.S.R., Cros, E., Dejtaradol, A., Chattopadhyay, B., Pwint, N., Packert, M., Rheindt, F.E., 2016. Genome-wide data help identify an avian species-level lineage that is morphologically and vocally cryptic. Mol. Phylogenet. Evol. 102, 97–103. https://doi.org/10.1016/j.ympev.2016.05.028.

González-Wevar, C.A., Hüne, M., Segovia, N.I., Nakano, T., Spencer, H.G., Chown, S.L., Saucède, T., Johnstone, G., Mansilla, A., Poulin, E., 2017. Following the antarctic circumpolar current: patterns and processes in the biogeography of the limpet Nacella (Mollusca: Patellogastropoda) across the Southern Ocean. J. Biogeogr. 44, 861–874. https://doi.org/10.1111/jbi.12908.

Goslee, S.C., Urban, D.L., 2007. The ecodist package for dissimilarity-based analysis of ecological data. J. Stat. Softw. 22, 1–19. https://doi.org/10.18637/jss.v022.i07.

Grosser, S., Burridge, C.P., Peucker, A.J., Waters, J.M., 2015. Coalescent modelling suggests recent secondary-contact of cryptic penguin species. PLoS One 10, e0144966. https://doi.org/10.1371/journal.pone.0144966.

Gruber, B., Adamack, A.T., 2015. Landgenreport: a new r function to simplify landscape genetic analysis using resistance surface layers. Mol. Ecol. Resour. 15, 1172–1178. https://doi.org/10.1111/1755-0998.12381.

Hart, T., Fitzcharles, E., Trathan, P.N., Coulson, T., Rogers, A.D., 2009. Testing and improving the accuracy of discriminant function tests: a comparison between morphometric and molecular sexing in macaroni penguins. Waterbirds 32, 437–443. https://doi.org/10.1675/063.032.0309.

Hindell, M.A., 1988. The diet of the royal penguin *Eudyptes schlegeli* at Macquarie island. Emu 88, 219–226. https://doi.org/10.1071/MU9880219.

Hinke, J.T., Polito, M.J., Goebel, M.E., Jarvis, S., Reiss, C.S., Thorrold, S.R., Trivelpiece, W.Z., Watters, G.M., 2015. Spatial and isotopic niche partitioning during winter in chinstrap and Adelie penguins from the South Shetland Islands. Ecosphere 6. https://doi.org/10.1890/es14-00287.1.

Hull, C., 1996. Morphometric indices for sexing adult royal *Eudyptes schlegeli* and rockhopper *E. chrysocome* penguins at Macquarie island. Mar. Ornithol. 24, 23–27.
Hull, C.L., 1999. Comparison of the diets of breeding royal (*Eudyptes schlegeli*) and rockhopper (*Eudyptes chrysocome*) penguins on Macquarie Island over three years. J. Zool. 247, 507–529. https://doi.org/10.1111/j.1469-7998.1999.tb01013.x.
Irwin, D.E., Milá, B., Toews, D.P.L., Brelsford, A., Kenyon, H.L., Porter, A.N., Grossen, C., Delmore, K.E., Alcaide, M., Irwin, J.H., 2018. A comparison of genomic islands of differentiation across three young avian species pairs. Mol. Ecol. 27, 4839–4855. https://doi.org/10.1111/mec.14858.
Jakobsson, M., 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, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233.
Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y.W., Faircloth, B.C., Nabholz, B., Howard, J.T., Suh, A., Weber, C.C., da Fonseca, R.R., Li, J., Zhang, F., Li, H., Zhou, L., Narula, N., Liu, L., Ganapathy, G., Boussau, B., Bayzid, M.S., Zavidovych, V., Subramanian, S., Gabaldón, T., Capella-Gutiérrez, S., Huerta-Cepas, J., Rekepalli, B., Munch, K., Schierup, M., Lindow, B., Warren, W.C., Ray, D., Green, R.E., Bruford, M.W., Zhan, X., Dixon, A., Li, S., Li, N., Huang, Y., Derryberry, E.P., Bertelsen, M.F., Sheldon, F.H., Brumfield, R.T., Mello, C.V., Lovell, P.V., Wirthlin, M., Schneider, M.P.C., Prosdocimi, F., Samaniego, J.A., Velazquez, A.M.V., Alfaro-Núñez, A., Campos, P.F., Petersen, B., Sicheritz-Ponten, T., Pas, A., Bailey, T., Scofield, P., Bunce, M., Lambert, D.M., Zhou, Q., Perelman, P., Driskell, A.C., Shapiro, B., Xiong, Z., Zeng, Y., Liu, S., Li, Z., Liu, B., Wu, K., Xiao, J., Yinqi, X., Zheng, Q., Zhang, Y., Yang, H., Wang, J., Smeds, L., Rheindt, F.E., Braun, M., Fjeldsa, J., Orlando, L., Barker, F.K., Jønsson, K.A., Johnson, W., Koepfli, K.-P., O'Brien, S., Haussler, D., Ryder, O.A., Rahbek, C., Willerslev, E., Graves, G.R., Glenn, T.C., McCormack, J., Burt, D., Ellegren, H., Alström, P., Edwards, S.V., Stamatakis, A., Mindell, D.P., Cracraft, J., Braun, E.L., Warnow, T., Jun, W., Gilbert, M.T.P., Zhang, G., 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346, 1320. https://doi.org/10.1126/science.1253451.
Jombart, T., 2008. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129.
Jombart, T., Ahmed, I., 2011. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521.
Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 94. https://doi.org/10.1186/1471-2156-11-94.
Kim, K.-W., Jackson, B.C., Zhang, H., Toews, D.P.L., Taylor, S.A., Greig, E.I., Lovette, I.J., Liu, M.M., Davison, A., Griffith, S.C., Zeng, K., Burke, T., 2019. Genetics and evidence for balancing selection of a sex-linked colour polymorphism in a songbird. Nat. Commun. 10, 1852. https://doi.org/10.1038/s41467-019-09806-6.
Kooyman, G.L., 2002. Evolutionary and ecological aspects of some Antarctic and sub-Antarctic penguin distributions. Oecologia 130, 485–495. https://doi.org/10.1007/s00442-001-0836-x.
Kooyman, G.L., Kooyman, T.G., Horning, M., Kooyman, C.A., 1996. Penguin dispersal after fledging. Nature 383, 397-397. https://doi.org/10.1038/383397a0.
Ksepka, D.T., Bertelli, S., Giannini, N.P., 2006. The phylogeny of the living and fossil Sphenisciformes (penguins). Cladistics 22, 412–441. https://doi.org/10.1111/j.1096-0031.2006.00116.x.
Larue, M.A., Kooyman, G., Lynch, H.J., Fretwell, P., 2015. Emigration in emperor penguins: implications for interpretation of long-term studies. Ecography 38, 114–120. https://doi.org/10.1111/ecog.00990.
Lavinia, P.D., Barreira, A.S., Campagna, L., Tubaro, P.L., Lijtmaer, D.A., 2019. Contrasting evolutionary histories in Neotropical birds: divergence across an environmental barrier in South America. Mol. Ecol. 28, 1730–1747. https://doi.org/10.1111/mec.15018.
Levy, H., Clucas, G.V., Rogers, A.D., Leache, A.D., Ciborowski, K.L., Polito, M.J., Lynch, H.J., Dunn, M.J., Hart, T., 2016. Population structure and phylogeography of the Gentoo Penguin (*Pygoscelis papua*) across the Scotia Arc. Ecol. Evol. 6, 1834–1853. https://doi.org/10.1002/ece3.1929.
Li, C., Zhang, Y., Li, J., Kong, L., Hu, H., Pan, H., Xu, L., Deng, Y., Li, Q., Jin, L., Yu, H., Chen, Y., Liu, B., Yang, L., Liu, S., Zhang, Y., Lang, Y., Xia, J., He, W., Shi, Q., Subramanian, S., Millar, C.D., Meader, S., Rands, C.M., Fujita, M.K., Greenwold, M.J., Castoe, T.A., Pollock, D.D., Gu, W., Nam, K., Ellegren, H., Ho, S.Y., Burt, D.W., Ponting, C.P., Jarvis, E.D., Gilbert, M.T., Yang, H., Wang, J., Lambert, D.M., Wang, J., Zhang, G., 2014. Two Antarctic penguin genomes reveal insights into their evolutionary history and molecular changes related to the Antarctic environment. GigaScience 3, 27. https://doi.org/10.1186/2047-217x-3-27.
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 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
Magoc, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12.
Mason, N.A., Taylor, S.A., 2015. Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. Mol. Ecol. 24, 3009–3025. https://doi.org/10.1111/mec.13140.
Meirmans, P.G., 2012. The trouble with isolation by distance. Mol. Ecol. 21, 2839–2846. https://doi.org/10.1111/j.1365-294X.2012.05578.x.
Meirmans, P.G., 2014. Nonconvergence in Bayesian estimation of migration rates. Mol. Ecol. Resour. 14, 726–733. https://doi.org/10.1111/1755-0998.12216.
Moon, K.L., Chown, S.L., Fraser, C.I., 2017. Reconsidering connectivity in the sub-Antarctic. Biol. Rev. 92, 2164–2181. https://doi.org/10.1111/brv.12327.
Ng, C.S., Li, W.-H., 2018. Genetic and molecular basis of feather diversity in birds. Genome Biol. Evol. 10, 2572–2586. https://doi.org/10.1093/gbe/evy180.
Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2019. vegan: Community Ecology Package. R package version 2.5-5.
Paxton, E.H., 2009. The utility of plumage coloration for taxonomic and ecological studies. Open Ornithol. J. 2, 17–23. https://doi.org/10.2174/1874453200902010017.
Pérez-Alvarez, M.J., Olavarría, C., Moraga, R., Baker, C.S., Hamner, R.M., Poulin, E., 2015. microsatellite markers reveal strong genetic structure in the endemic chilean dolphin. PLoS One 10, e0123956. https://doi.org/10.1371/journal.pone.0123956.
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest radseq: an inexpensive method for de novo snp discovery and genotyping in model and non-model species. PLoS One 7. https://doi.org/10.1371/journal.pone.0037135.
Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., Lareu, M., 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. Front. Genet. 4. https://doi.org/10.3389/fgene.2013.00098.
Poulin, E., González-Wevar, C., Díaz, A., Gérard, K., Hüne, M., 2014. Divergence between Antarctic and South American marine invertebrates: what molecular biology tells us about Scotia Arc geodynamics and the intensification of the Antarctic Circumpolar Current. Glob. Planet. Change 123, 392–399. https://doi.org/10.1016/j.gloplacha.2014.07.017.
Pritchard, J., Wen, X., Falush, D., 2010. Documentation for structure software: Version 2.3.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. https://doi.org/10.1086/519795.
Core Team, R., 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rocha-Mendez, A., Sanchez-Gonzalez, L.A., Arbelaez-Cortes, E., Navarro-Siguenza, A.G., 2018. Phylogeography indicates incomplete genetic divergence among phenotypically differentiated montane forest populations of *Atlapetes albinucha* (Aves, Passerellidae). Zookeys 125–148. https://doi.org/10.3897/zookeys.809.28743.
Rochette, N.C., Rivera-Colón, A.G., Catchen, J.M., 2019. Stacks 2: Analytical methods for Paired-end sequencing improve RADseq-based population genomics. bioRxiv 615385. https://doi.org/10.1101/615385.
Rosenberg, N.A., 2003. Distruct: a program for the graphical display of population structure. Mol. Ecol. Notes 4, 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x.
Shafer, A.B.A., Peart, C.R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C.W., Wolf, J.B.W., 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. Meth. Ecol. Evol. 8, 907–917. https://doi.org/10.1111/2041-210X.12700.
Thiebot, J.-B., Lescroël, A., Barbraud, C., Bost, C.-A., 2013. Three-dimensional use of marine habitats by juvenile emperor penguins *Aptenodytes forsteri* during post-natal dispersal. Antarct. Sci. 25, 536–544. https://doi.org/10.1017/S0954102012001198.
Thiebot, J.B., Cherel, Y., Trathan, P.N., Bost, C.A., 2011. Inter-population segregation in the wintering areas of macaroni penguins. Mar. Ecol. Prog. Ser. 421, 279–290. https://doi.org/10.3354/meps08907.
Toews, David P.L., Taylor, Scott A., Vallender, R., Brelsford, A., Butcher, Bronwyn G., Messer, Philipp W., Lovette, Irby J., 2016. plumage genes and little else distinguish the genomes of hybridizing warblers. Curr. Biol. 26, 2313–2318. https://doi.org/10.1016/j.cub.2016.06.034.
Van Grouw, H., 2006. Not every white bird is an albino: sense and nonsense about colour aberrations in birds. Dutch Bird. 28, 79–89.
Veale, A.J., Foster, B.J., Dearden, P.K., Waters, J.M., 2018. Genotyping-by-sequencing supports a genetic basis for wing reduction in an alpine New Zealand stonefly. Sci. Rep. 8, 16275. https://doi.org/10.1038/s41598-018-34123-1.
Vianna, J.A., Noll, D., Dantas, G.P.M., Petry, M.V., Barbosa, A., Gonzalez-Acuna, D., Le Bohec, C., Bonadonna, F., Poulin, E., 2017. Marked phylogeographic structure of Gentoo penguin reveals an ongoing diversification process along the Southern Ocean. Mol. Phylogenet. Evol. 107, 486–498. https://doi.org/10.1016/j.ympev.2016.12.003.
Wang, I.J., 2013. Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. Evolution 67, 3403–3411. https://doi.org/10.1111/evo.12134.
Warham, J., 1971. Aspects of breeding behaviour in the royal penguin *Eudyptes chrysolophus* schlegeli. Notornis 18.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.
Wienecke, B., Raymond, B., Robertson, G., 2010. Maiden journey of fledgling emperor penguins from the Mawson Coast, East Antarctica. Mar. Ecol. Prog. Ser. 410, 269–282. https://doi.org/10.3354/meps08629.
Williams, T.D., Rodwell, S., 1992. Annual variation in return rate, mate and nest-site fidelity in breeding gentoo and macaroni penguins. Condor 94, 636–645. https://doi.org/10.2307/1369249.
Wilson, G.A., Rannala, B., 2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163, 1177–1191.
Wilson, R., 1997. A method for restraining penguins. Mar. Ornithol. 25, 72–73.
Younger, J.L., Clucas, G.V., Kao, D., Rogers, A.D., Gharbi, K., Hart, T., Miller, K.J., 2017.

The challenges of detecting subtle population structure and its importance for the conservation of emperor penguins. Mol. Ecol. 26, 3883–3897. https://doi.org/10.1111/mec.14172.

Zhang, G., Lambert, D.M., Wang, J., 2011. Genomic data from the emperor penguin (*Aptenodytes forsteri*). GigaScience. http://gigadb.org/dataset/100005.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D.M., Lee, C., Storz, J.F., Antunes, A., Greenwold, M.J., Meredith, R.W., Odeen, A., Cui, J., Zhou, Q., Xu, L., Pan, H., Wang, Z., Jin, L., Zhang, P., Hu, H., Yang, W., Hu, J., Xiao, J., Yang, Z., Liu, Y., Xie, Q., Yu, H., Lian, J., Wen, P., Zhang, F., Li, H., Zeng, Y., Xiong, Z., Liu, S., Zhou, L., Huang, Z., An, N., Wang, J., Zheng, Q., Xiong, Y., Wang, G., Wang, B., Wang, J., Fan, Y., da Fonseca, R.R., Alfaro-Nunez, A., Schubert, M., Orlando, L., Mourier, T., Howard, J.T., Ganapathy, G., Pfenning, A., Whitney, O., Rivas, M.V., Hara, E., Smith, J., Farre, M., Narayan, J., Slavov, G., Romanov, M.N., Borges, R., Machado, J.P., Khan, I., Springer, M.S., Gatesy, J., Hoffmann, F.G., Opazo, J.C., Hastad, O., Sawyer, R.H., Kim, H., Kim, K.W., Kim, H.J., Cho, S., Li, N., Huang, Y., Bruford, M.W., Zhan, X., Dixon, A., Bertelsen, M.F., Derryberry, E., Warren, W., Wilson, R.K., Li, S., Ray, D.A., Green, R.E., O'Brien, S.J., Griffin, D., Johnson, W.E., Haussler, D., Ryder, O.A., Willerslev, E., Graves, G.R., Alstrom, P., Fjeldsa, J., Mindell, D.P., Edwards, S.V., Braun, E.L., Rahbek, C., Burt, D.W., Houde, P., Zhang, Y., Yang, H., Wang, J., Jarvis, E.D., Gilbert, M.T., Wang, J., 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. Science 346, 1311–1320. https://doi.org/10.1126/science.1251385.

Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., Weir, B.S., 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28, 3326–3328. https://doi.org/10.1093/bioinformatics/bts606.

Zink, R.M., Rising, J.D., Mockford, S., Horn, A.G., Wright, J.M., Leonard, M., Westberg, M.C., 2005. Mitochondrial DNA variation, species limits, and rapid evolution of plumage coloration and size in the Savannah sparrow. Condor 107, 21–28. https://doi.org/10.1650/7550.