RESEARCH ARTICLE

Diversity and Distributions   WILEY

Taxonomy based on limited genomic markers may underestimate species diversity of rockhopper penguins and threaten their conservation

María José Frugone[1,2,3] | Theresa L. Cole[4,5] | María Eugenia López[6] |
Gemma Clucas[7,8] | Pável Matos-Maraví[9] | Nicolás A. Lois[10,11] | Pierre Pistorius[12] |
Francesco Bonadonna[13] | Phil Trathan[14] | Andrea Polanowski[15] | Barbara Wienecke[15] |
Andrea Raya-Rey[16,17,18] | Klemens Pütz[19] | Antje Steinfurth[20,21] | Ke Bi[22] |
Cynthia Y. Wang-Claypool[22] | Jonathan M. Waters[4] | Rauri C. K. Bowie[22] |
Elie Poulin[1,2] | Juliana A. Vianna[23]

[1]Laboratorio de Ecología Molecular, Departamento de Ciencias Ecológicas, Facultad de Ciencias, Universidad de Chile, Santiago, Chile

[2]Instituto de Ecología y Biodiversidad (IEB), Santiago, Chile

[3]Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile

[4]Department of Zoology, University of Otago, Dunedin, New Zealand

[5]Department of Biology, Ecology and Evolution, University of Copenhagen, Copenhagen, Denmark

[6]Department of Aquatic Resources, Swedish University of Agricultural Sciences, Drottningholm, Sweden

[7]Atkinson Center for a Sustainable Future, Cornell University, Ithaca, NY, USA

[8]Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA

[9]Biology Centre of the Czech Academy of Sciences, Institute of Entomology, České Budějovice, Czech Republic

[10]Departamento de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

[11]Instituto de Ecología Genética y Evolución de Buenos Aires, Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina

[12]DST/NRF Centre of Excellence at the Percy FitzPatrick Institute for African Ornithology, Department of Zoology, Nelson Mandela University, Port Elizabeth, South Africa

[13]CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

[14]British Antarctic Survey, Cambridge, UK

[15]Australian Antarctic Division, Kingston, Tasmania, Australia

[16]Centro Austral de Investigaciones Científicas – Consejo Nacional de Investigaciones Científicas y Técnicas (CADIC-CONICET), Ushuaia, Argentina

[17]Wildlife Conservation Society, Bronx, NY, USA

[18]Instituto de Ciencias Polares, Ambiente y Recursos Naturales, Universidad Nacional de Tierra del Fuego, Ushuaia, Argentina

[19]Antarctic Research Trust, Bremervörde, Germany

[20]FitzPatrick Institute of African Ornithology, University of Cape Town, Rondebosch, South Africa

[21]RSPB Centre for Conservation Science, Cambridge, UK

[22]Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA, USA

[23]Pontificia Universidad Católica de Chile, Center for Genome Regulation, Facultad de Agronomía e Ingeniería Forestal, Departamento de Ecosistemas y Medio Ambiente, Santiago, Chile

Correspondence

Juliana A. Vianna, Pontificia Universidad Católica de Chile, Facultad de Agronomía e Ingeniería Forestal, Departamento de Ecosistemas y Medio Ambiente, Vicuña Mackenna 4860, Macul, Santiago, Chile.
Email: jvianna@uc.cl

Abstract

Aim: Delimiting recently diverged species is challenging. During speciation, genetic differentiation may be distributed unevenly across the genome, as different genomic regions can be subject to different selective pressures and evolutionary histories. Reliance on limited numbers of genetic markers that may be underpowered can make species delimitation even more challenging, potentially resulting in taxonomic inconsistencies. Rockhopper penguins of the genus *Eudyptes* comprise three broadly recognized taxa: northern (*E. moseleyi*), southern (*E. chrysocome*) and eastern rockhopper (*E. filholi*). Their taxonomic status has been controversial for decades, with researchers disagreeing about whether *E. chrysocome* and *E. filholi* are distinct species or conspecific. Our goal is to evaluate genome-wide patterns of divergence to evaluate genetic differentiation and species delimitation in rockhopper penguins, and to assess which mechanisms may underlie previous discordance among nuclear versus mitochondrial analyses.

Location: Sub-Antarctic and temperate coastal regions of the Southern Hemisphere.

Methods: We generated reduced-representation genomic libraries using double digest restriction-site associated DNA (ddRAD) sequencing to evaluate genetic differentiation, contemporary migration rates and admixture among colonies of rockhopper penguins.

Results: The extent of genetic differentiation among the three taxa was consistently higher than population-level genetic differentiation found within these and other penguin species. There was no evidence of admixture among the three taxa, suggesting the absence of ongoing gene flow among them. Species delimitation analyses based on molecular data, along with other lines of evidence, provide strong support for the taxonomic distinction of three species of rockhopper penguins.

Main conclusions: Our results provide strong support for the existence of three distinct species of rockhopper penguins. The recognition of this taxonomic diversity is crucial for the management and conservation of this widely distributed species group. This study illustrates that widespread dispersive seabird lineages lacking obvious morphological differences may nevertheless have complex evolutionary histories and comprise cryptic species diversity.

KEYWORDS

Eudyptes, genomics, rockhopper penguins, species delimitation

1 | INTRODUCTION

Accurate delimitation of species is essential for the management and conservation of biodiversity (Agapow et al., 2004; Frankham et al., 2012). Indeed, conservation action plans and assessments are often based on biological and demographic attributes focused primarily at the species level (Agapow et al., 2004). Any failure to recognize cryptic species diversity could lead to overestimates of population and range sizes, and thereby misinform the assessment of threat levels and compromise the prioritization of conservation resources (Bickford et al., 2007). Conversely, over-splitting of lineages

through inaccurate recognition of multiple species might impair management through the exclusion of potentially valuable source stocks that could otherwise be used for relocations or genetic rescue efforts (Frankham et al., 2012). Furthermore, with limited funding for biological conservation, over-splitting dilutes and misdirects available resources (Agapow et al., 2004).

Despite its importance, species delimitation can be a challenging task (Fujita et al., 2012; Sangster, 2014). The root of this problem lies in the biological complexity and variability accompanying the speciation process, and how we define a species (De Queiroz, 2007; Schilthuizen, 2000). Consequently, there is no universal species

concept that can be applied to all living organisms to confidently delimit species. Additionally, the most basic assumption of species delimitation strategies is that "true species" will exhibit variation in traits of different biological dimensions (e.g. genetics, behaviour, morphology, ecology). This carries an additional level of complexity, as different traits along these biological dimensions could diverge at different times during the speciation process (De Queiroz, 2007). Furthermore, variation in traits may be the result of local adaptation or phenotypic plasticity and not be indicative of the speciation process (Frugone et al., 2019; Mason & Taylor, 2015). Thus, integrative taxonomy, in which traits from multiple lines of evidence (ecology, genomics, phenotype, behaviour, geography) aid the identification of divergent lineages, has become a widely accepted approach to delimiting species (Cicero et al., 2021; De Queiroz, 2007; Sangster, 2014). Using this approach, it is essential to understand the underlying processes generating trait variation (or the lack of it) that may be responsible for discordances between independent lines of evidence.

Molecular data and the availability of sophisticated analytical methods (e.g. coalescent-based species delimitation; Fujita et al., 2012) have provided a useful tool for uncovering cryptic diversity and delimiting species. However, discordances may arise among different sets of nuclear markers or between nuclear and mitochondrial DNA (mtDNA) when delimiting species, especially for recently diverged lineages (Pedraza-Marrón et al., 2019; Spinks et al., 2014). The nuclear genomes of recently diverged species may contain varying levels of genetic differentiation dispersed throughout the genome (Nosil et al., 2009; Wu, 2001). Even under complete reproductive isolation, genomic regions with reduced mutation rates, or which are subject to stabilizing selection, may remain conserved for a considerable length of time compared to other regions with higher mutation rates or those under divergent selection (Nosil et al., 2009). Furthermore, when speciation occurs in the presence of gene flow, different regions of the genome may either become homogenized, or remain differentiated if introgressed alleles carry negative fitness consequences, and if selection is stronger than gene flow (Bazykin, 1969; Key, 1968; Wu, 2001). This may lead to genomic "islands" of differentiation against a highly undifferentiated genomic background (Cutter, 2013; Feder et al., 2012; Nosil et al., 2009; Wu, 2001). Therefore, to some extent, discrepancies in species delimitation can be explained by introgressive hybridization and Incomplete Lineage Sorting (ILS) across segments of the genome (Cutter, 2013; Maddison, 1997; Pedraza-Marrón et al., 2019).

Discrepancies may also occur between mtDNA and nuclear markers due to differences in selective regimes and sex-biased gene flow, or more commonly, because of the relative degree of divergence among lineages for each type of marker (Edwards & Bensch, 2009; Toews & Brelsford, 2012; Zink & Barrowclough, 2008). The effective population size is four times smaller for mtDNA than for nuclear DNA, due to intrinsic differences in ploidy and the maternal mode of inheritance of mtDNA. Consequently, monophyly after diversification will occur for mtDNA before nuclear DNA (Zink & Barrowclough, 2008). In

this sense, mtDNA may exhibit greater differentiation and be better suited for detecting recently diverged taxa than nuclear DNA, especially if just a few nuclear genetic markers are used to assess species- and population-level divergence (Zink & Barrowclough, 2008). One way to overcome such limitations, and to estimate the underlying causes of discordance between phylogenetic hypotheses generated using mtDNA and nuclear DNA, is through the use of thousands of unlinked genetic markers. This multilocus approach encompasses regions subject to a variety of evolutionary constraints in order to infer genome-wide patterns of divergence (Esquerre et al., 2019; Guo et al., 2019).

Eudyptes is the most species-rich penguin genus (6–8 extant species), with "rockhopper penguins" comprising three recently diverged lineages: the subtropical taxon *E. moseleyi* and the sub-Antarctic taxa *E. chrysocome* and *E. filholi* (Figure 1). Using the fossil record or dates of geological events, the oldest diversification time across different studies for the split between *E. moseleyi* and the sub-Antarctic taxa is 3.06 Mya, while the divergence of the sub-Antarctic taxa from each other occurred from 0.5 to 2.26 Mya (Cole, Dutoit, et al., 2019; Cole, Ksepka, et al., 2019; de Dinechin et al., 2009; Frugone et al., 2018; Gavryushkina et al., 2017; Vianna et al., 2020). The taxonomic status of rockhopper penguins has been repeatedly revised. In the 1990s, these three lineages were considered to represent a single species (Martínez, 1992). In 2006, based on differences in their mating calls, timing of breeding (starting ~2 months earlier in *E. moseleyi*), genetics (mtDNA; HVRI) and morphological characters, rockhoppers were classified into two species (García & Boersma, 2013; Jouventin et al., 2006): the southern rockhopper (*E. chrysocome*) and the northern rockhopper (*E. moseleyi*), hereafter the two-species hypothesis. In the same year, based on genetic evidence (mtDNA; 12S, cytb and COI), Banks et al. (2006) proposed that rockhopper penguins should instead be classified as three species (hereafter the three-species hypothesis) comprising *E. moseleyi*, and *E. chrysocome* split into two species: the eastern rockhopper, *E. filholi* distributed on islands in the Indian and western Pacific oceans, and the southern rockhopper, *E. chrysocome* occupying islands in the eastern Pacific and Atlantic oceans (Figure 1).

The three-species hypothesis has been supported by recent studies including more extensive geographic sampling using mtDNA (Cole, Dutoit, et al., 2019; Cole, Ksepka, et al., 2019; Frugone et al., 2018). Conflicting with the above, Mays et al. (2019) suggested that *E. filholi* and *E. chrysocome* should be considered conspecific, in agreement with the two-species hypothesis, as their species delimitation analyses did not find sufficient differences to discriminate among the taxa, and analyses of migration rates between colonies suggested ongoing gene flow. However, such conclusions were based on the analysis of only six nuclear introns, and their species delimitation analysis could not discriminate the even more divergent and widely accepted split (around 5 Mya) between macaroni (*E. chrysolophus*) and rockhopper penguins; only one out of their four Genealogical Diversity Index (GDI) scenarios exceeded the 0.2 GDI threshold for recognizing potential species, but it did not reach the 0.7 GDI threshold for a strongly divergent species lineage. With a
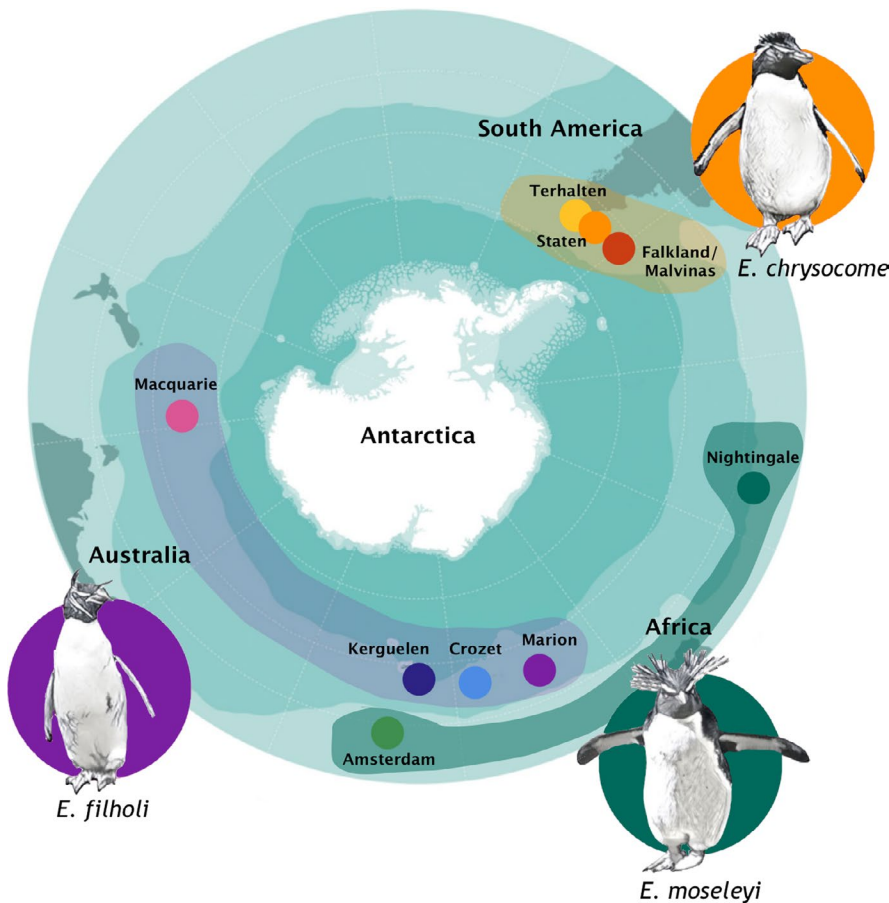
FIGURE 1 Sample locations for rockhopper penguins throughout their distribution. *Eudyptes filholi* (eastern rockhopper), *E. chrysocome* (southern rockhopper) and *E. moseleyi* (northern rockhopper)

different set of nuclear introns, Frugone et al. (2018) were also unable to discriminate macaroni and rockhopper penguins, nor could they discriminate among lineages within rockhopper penguins, despite their strong divergence and reciprocal monophyly with mtDNA markers.

If *E. chrysocome* and *E. filholi* are conspecific and are not reproductively isolated, discordances between mtDNA and nuclear markers might be explained by skewed dispersal (and gene flow) towards males (i.e. with females being more philopatric). Another possibility is that the chosen nuclear markers correspond to undifferentiated regions of the genomes of the taxa under study, a situation that may arise from different mechanisms, such as introgression or ILS. Unravelling the causes of discordance between mtDNA and nuclear DNA is important to more accurately infer species delimitations that could ultimately reinforce biodiversity conservation. Our reassessment of the evolutionary history of rockhopper penguins using a reduced genomic representation library of single nucleotide polymorphisms (SNPs), resulting in thousands of genetic markers, allow us to evaluate whether the genomic pattern of divergence is in accordance with that previously shown by mtDNA or if other processes, unrelated to genomic coverage, are responsible. In addition to evaluating genetic differentiation and species delimitation of *E. moseleyi*, *E. filholi* and *E. chrysocome*, we also evaluate contemporary migration rates and search for evidence of admixture based on prior suggestions

about ongoing gene flow among rockhopper penguins (Mays et al., 2019). Finally, we explore regional population genomics within each taxon to improve our understanding of the extent of intra- and interspecific genetic differentiation.

Currently, the IUCN recognizes two rockhopper penguin species, *E. moseleyi* and *E. chrysocome*, categorizing them as endangered and vulnerable, respectively (BirdLife International, 2020a, 2020b). The conservation actions proposed by the IUCN include the determination of the taxonomic status of *E. filholi* and *E. chrysocome*, representing a key data source for the development of conservation plans for these charismatic birds.

2 | METHODS

2.1 | Blood sampling and ddRAD-seq library preparation

We collected 96 blood samples across an extensive area covering nearly the entire distribution of the three rockhopper taxa: 24 *E. moseleyi* ($n = 12$ from Nightingale Island, Tristan da Cunha; $n = 12$ from Amsterdam Island); 30 *E. chrysocome* ($n = 13$ from Terhalten Island, Chile; $n = 13$ from Staten Island, Argentina; $n = 4$ from the Falkland/Malvinas Islands); and 42 *E. filholi* ($n = 14$ from Marion Island; $n = 10$ from Crozet Island; $n = 10$ from Kerguelen Island; $n = 8$

from Macquarie Island; Figure 1, Table S1). Access to penguin colonies, permission to collect blood samples and animal ethics approvals were granted by the responsible authorities for each sampling location (Table S2). We used blood samples to prepare double digest restriction-site associated DNA (ddRAD) libraries, following the protocol described in Peterson et al. (2012). Laboratory procedures and the library preparation protocol are described in the Appendix S1 Methods Section II.1.

2.2 | ddRAD-seq data processing, SNP calling and filtering

We used a custom PERL pipeline encompassing various external programmes for processing the ddRAD-seq data (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, allowing for one mismatch. The reads were then filtered using CUTADAPT V. 1.8.1 (Martin, 2011) and TRIMMOMATIC V. 0.36 (Bolger et al., 2014) to trim adapter contamination and low-quality reads, respectively.

To improve the efficiency and accuracy of short-read mapping, we used reference genomes for aligning ddRAD-seq data (Shafer et al., 2017). We used the genomes of *E. moseleyi* (mean coverage: 24x, 16,344 Scaffolds, N50 length: 5,071,598), *E. chrysocome* (mean coverage: 31x, 21,917 Scaffolds, N50 length: 5,071,598) and *E. filholi* (mean coverage: 26x, 19,210 Scaffolds, N50 length: 5,071,598) from Vianna et al. (2020). For our population genomic analyses (i.e. analyses performed within each of the three taxa; Table S3), we aligned the reads of each taxon to its own reference genome (Vianna et al., 2020). For analyses involving only *E. chrysocome* and *E. filholi*, we generated a second dataset, aligning the reads to *E. moseleyi* genome (see Table S3). For analyses involving the three taxa (see Table S3), reads from all individuals were aligned to an *E. chrysolophus* reference genome (mean coverage: 28x, 18,969 scaffolds, N50 length: 5,071,598; Frugone et al., 2019; Vianna et al., 2020). All alignments were performed using the *bwa mem* algorithms (Li & Durbin, 2009) sorting and indexing bam files using SAMTOOLS V. 1.9 (Li et al., 2009).

Single nucleotide polymorphisms calling was conducted using *gstacks* (Catchen et al., 2013). We used the *populations* programme in STACKS V. 2.52 (Catchen et al., 2013) and VCFTOOLS V. 0.1.13 (Danecek et al., 2011) to filter loci. For this step, we used different criteria following the assumptions and recommendation for each analysis as documented in Benestan et al. (2016) and O'Leary et al. (2018). For all datasets, we filtered out sites with observed heterozygosity >0.5 and those with a mean depth of <10x or >175x to avoid unreliable genotypes or SNPs called from repetitive regions of the genome (indicated by the hump in distribution of locus depths above 175x, Figure S1). We removed sites that exhibited >20% of missing data within each population/grouping option, aiming for a similar

distribution of missing data across them. We retained sites that were present in >70% of individuals in a population/grouping option, as established on the population map from STACKS (Table S3). Mean depth and missing data across sites and groups were calculated for each dataset using VCFTOOLS and then filtered using the blacklist in STACKS. These filtering criteria were used to generate Manhattan plots and to perform FINERADSTRUCTURE and OUTFLANK analyses (see below). We refer to the datasets obtained following these filtering criteria as the "full datasets."

To perform STRUCTURE, PCA, pairwise F_{ST} and genetic diversity analyses, we further filtered our datasets, as several population genomic analyses required loci to be in Hardy–Weinberg equilibrium (HWE) and at linkage equilibrium (e.g. Pritchard et al., 2000, 2010). We used PLINK V. 1.9 (Purcell et al., 2007) to remove loci in linkage disequilibrium calculated using a window size of 50 SNPs, two SNPs to shift the window at each step and a variance inflation factor of two. Then, using VCFTOOLS, we calculated the deviation from HWE and removed sites that exhibited significant deviations (after FDR correction) in over ~50% of the populations. We retained sites exhibiting a minimum allele frequency (MAF) of 0.05. We refer to these datasets as "unlinked datasets." For BAYESASS, SNAPP and Bayes factor delimitation (BFD*) described below, we did not filter by MAF and for SNAPP and BFD*, we retained only sites that were present across all individuals. Detailed descriptions for all filtering options are summarized in Table S3. Also, we evaluated whether our dataset included loci under selection, as several analyses of population genetics assume neutrality of the data. We ran OUTFLANK V. 0.2 (Whitlock & Lotterhos, 2015) in R 3.6.3 (R Core Team, 2020) to detect sites that could be under selection. We conducted these outlier analyses for all pairwise comparisons among the three taxa using the unlinked dataset, and also for all intraspecific datasets. In OUTFLANK, we set a trim factor (left and right) of 0.05, a minimum heterozygosity of 0.1 and a q threshold of 0.05. We used PGDSPIDER V. 2.1.1.5 to transform some VCF files into input files for multiple programs (Lischer & Excoffier, 2012). Finally, a preliminary PCA analysis revealed the presence of three outliers corresponding to *E. filholi* individuals from Kerguelen Island (Figure S2a; Appendix S1 Methods Section II.2). We further explored the genetic relatedness of these outliers to all other sampled individuals (including both *E. filholi* and *E. chrysocome*) via an identity-by-state analysis using SNPRELATE V. 1.20.1 (Zheng et al., 2012). Results showed that the distance between these individuals was higher than that observed among the rest of *E. filholi* individuals, although the analysis assigned them correctly to the *E. filholi* group (Figure S2b; Appendix S1 Methods Section II.2). Such genetic differences in these three *E. filholi* individuals may be a result of either (a) a hybrid individual from a cross with other unknown penguin species, (b) migrant individuals from unsampled colonies exhibiting higher levels of genetic differentiation (e.g. Campbell, Auckland and Antipodes Islands; Cole, Dutoit, et al., 2019; Cole, Ksepka, et al., 2019) or (c) a laboratory or sequencing artefact. Given that we do not have additional data needed to test these hypotheses, we removed them for subsequent analyses and retained a total of 93 individuals.

2.3 | Private alleles, F_{ST} and Manhattan plots

We used the *populations* program in STACKS to calculate the number of private alleles for each population and taxon using unlinked datasets of 2,302 and 1,211 SNPs, respectively (Table S3). Other genetic diversity calculations (mean expected heterozygosity and nucleotide diversity (π) and inbreeding coefficient (F_{IS}) can be found in Table S6. To estimate genomic differentiation, we calculated pairwise F_{ST} among all colonies from the three taxa using GENODIVE V. 3.04 (Meirmans & Van Tienderen, 2004). Significance was determined with 10,000 permutations of the unlinked dataset. *p*-values were corrected using the False Discovery Rate method (Benjamini et al., 2006).

To evaluate differentiation across genomic sites, we also calculated F_{ST} on a per-site basis using STACKS to produce Manhattan plots for each pair of taxa (*E. moseleyi/E. chrysocome*, *E. moseleyi/E. filholi* and *E. chrysocome/E. filholi*) and among all populations within each taxon. Calculations were made with the full datasets (Table S3). To generate one intraspecific Manhattan plot for each taxon, we calculated the mean F_{ST} for all sites that were shared across all populations. Manhattan plots were generated in R V. 3.6.3 (R Core Team, 2020) using the QQMAN V. 0.1.4 package (Turner, 2014).

2.4 | Clustering analyses

We employed several approaches to evaluate population genetic structure and genetic differentiation among the three rockhopper penguin taxa. We used PLINK V. 2.00a3 (Purcell et al., 2007) and GGPLOT2 V. 3.3.3 (Wickham, 2016) in R 3.6.3 (R Core Team, 2020) to calculate and visualize a PCA using the unlinked datasets. Missing data (*E. moseleyi-E. chrysocome-E. filholi* ~40%; *E. chrysocome-E. filholi* ~0.46%; *E. moseleyi* ~0.78%; *E. chrysocome* ~1.37%; *E. filholi* ~0.53%) was replaced using the "meanimpute" modifier in PLINK that performs a mean impute of missing genotype calls.

We also used FINERADSTRUCTURE (Malinsky et al., 2018) to infer shared ancestry among all individuals, using the full dataset that included all taxa, and RADPAINTER to generate the input file from the VCF file and to calculate the co-ancestry matrix. We re-ordered loci according to linkage disequilibrium with the script provided with the package and ran FINERADSTRUCTURE using default parameters.

In addition, we used the Bayesian clustering method in STRUCTURE (Falush et al., 2003; Pritchard et al., 2000). All runs were made using PARALLELSTRUCTURE (Besnier & Glover, 2013) on the CIPRES V. 3.3 (Miller et al., 2010) platform. We performed this analysis with (a) all nine populations across the three taxa together in a single analysis, (b) the seven populations of *E. chrysocome* and the *E. filholi* populations and (c) separately for each taxon. For these analyses, we used the unlinked datasets (Table S3). For the analyses involving more than one taxon (a and b), we chose the admixture ancestry model, uncorrelated allele frequencies (Falush et al., 2003;

Porras-Hurtado et al., 2013) and adjust the alpha parameter to 1/K due to the differences in sample size for each taxon (Wang, 2017). For analyses at the taxon level, we chose the admixture ancestry model and correlated allele frequencies (Falush et al., 2003; Porras-Hurtado et al., 2013). Given the taxonomic uncertainty of *E. chrysocome* and *E. filholi*, we also used this parameter setting for the analysis, including all populations of these taxa, to compare results yielded by the uncorrelated allele frequency model. In all cases, we did not include the geographical origin of each individual as a prior, given our interest in detecting strong population structure. We performed 10 independent runs for each value of *K*, with 500,000 MCMC permutations and a burn-in of 50,000 permutations. The number of *K* values tested for each dataset was determined based on the number of geographic populations sampled for each taxon. We used the online version of STRUCTURE HARVESTER V. 0.6.94 (Earl & von Holdt, 2012) (http://taylor0.biology.ucla.edu/structureHarvester/) to infer the most likely *K* by the highest mean ln likelihood for each value of *K* and compared them with the results from Evanno's method when testing $K > 2$ (Evanno et al., 2005). We then used CLUMPAK (Kopelman et al., 2015), an online server that automatically delivers results from CLUMPP (Jakobsson & Rosenberg, 2007) and DISTRUCT (Rosenberg, 2003), to visualize and summarize the results of all previous runs.

Finally, we conducted an analysis of molecular variance (AMOVA) to assess how much of the molecular variation could be explained by differences among populations in contrast with that which could be explained by separately grouping populations of *E. filholi* and *E. chrysocome*. The AMOVA was performed using the unlinked dataset of 922 SNPs (Table S3) using POPPR V. 2.8.5 (Kamvar et al., 2014, 2015) in R V. 3.6.3 (R Core Team, 2020). The statistical significance was assessed through 10,000 permutations.

2.5 | Species delimitation and phylogeographic reconstruction

For species delimitation analyses and phylogenetic reconstruction, we included five macaroni penguin individuals (*E. chrysolophus*) from Frugone et al. (2019) as an outgroup. This was done to test the one-species hypothesis for rockhopper penguins. We used the SNAPP package V. 1.5.1 (Bryant et al., 2012) from BEAST V. 2.6.1 (Bouckaert et al., 2014) to delimit species using SNP data. We evaluated five competing species models using the BFD* method (Leaché et al., 2014). We included models representing the one- to three-species hypotheses of rockhopper penguins (Models a, b, c in Table 1) (Banks et al., 2006; de Dinechin et al., 2009; Frugone et al., 2018; Jouventin et al., 2006; Mays et al., 2019) and two extra models that further split divergent populations of *E. moseleyi* (Model d, Table 1), and divergent populations of both *E. moseleyi* and *E. chrysocome* (Model e, Table 1). These last two models were based on previous studies showing high levels of population divergence within *E. moseleyi* (Frugone et al., 2018) and *E. chrysocome* (Frugone et al., 2018; Lois et al., 2020).

TABLE 1 Bayes factor delimitation results for all models tested

Models	Bayes Factor (subset 1)	Bayes Factor (subset 2)
a) 2 species = E. chrysolophus/E. moseleyi + E. chrysocome + E. filholi	--	--
b) 3 species = E. chrysolophus/E. moseleyi/E. chrysocome + E. filholi	−3187.32	−3407.18
c) 4 species = E. chrysolophus/E. moseleyi/E. chrysocome/E. filholi	−3844.5	−4020.74
d) 5 species = E. chrysolophus/Nightingale[a]/Amsterdam[a]/E. chrysocome/E. filholi	−3832.32	−4013.34
e) 6 species = E. chrysolophus/Nightingale[a]/Amsterdam[a]/Falklands[b]/Staten[b] + Terhalten[b]/E. filholi	−3828.72	−4000.64

Note: Bayes factor was calculated against the two-species hypothesis (a). Negative Bayes factor indicates that the alternative model (b to e) is favoured.

[a]E. moseleyi populations.

[b]E. chrysocome populations.

Because BFD* is computationally demanding, we ran two independent analyses to check for consistency by twice randomly selecting five macaroni penguin individuals as an outgroup and 18 rockhopper penguin individuals as an ingroup, which included six samples from each of the three rockhopper penguin taxa, representing 1–3 individuals from nine populations. The selected individuals for both runs and their populations of origin are summarized in Table S4a and S4b. We used unlinked loci datasets, and we retained only the sites that were present in all individuals (Table S3). The total numbers of SNPs for each dataset were 1,798 and 1,806 respectively.

We used the Python script vcf2phylip.py (Ortiz, 2019), available at https://github.com/edgardomortiz/vcf2phylip, to convert the VCF files to binary nexus format. After fixing the mutation rates (u and v) to 1 and assigning a gamma distribution G (1, 250) to the prior for the expected genetic divergence (parameter θ) and a gamma distribution G (2, 200) for the speciation rate parameter λ, we carried out stepping-stone path sampling analyses for every competing species model. Marginal likelihood estimates were obtained using alpha =0.3 and 36 MCMC steps, each consisting of 100,000 cycles with a pre-burn-in run of 10,000 cycles. All models were compared by calculating Bayes factors. We followed Kass and Raftery (1995) and considered that, for two competing models, no evidence favours any model when BF <2, and that evidence for the most-likely model is positive for 2 <BF <6, strong for 6 < BF <10 and decisive for BF >10.

To evaluate the phylogenetic relationships among rockhopper penguins, we performed the Bayesian coalescent method implemented in SNAPP (Bryant et al., 2012) and performed a maximum likelihood estimate of tree topology using IQTREE V. 2.1.2 (Nguyen et al., 2015). The dataset for SNAPP included the five macaroni penguin individuals as an outgroup and 34 rockhopper penguin individuals (12 E. moseleyi, 10 E. chrysocome and 12 E. filholi, aiming to include at least three individuals from each study population) as the ingroup. We followed filtering criterion similar to what we used for the BFD* analyses (Table S3). The MCMC was run for 600,000 generations and sampled every 500 cycles, and a 20% burn-in rate

was applied. We performed three independent runs and joined the estimated parameters and trees to evaluate consistency and convergence using TRACER V. 1.7.1 (Rambaut et al., 2018). All estimated posterior distributions of parameters, except for a few θ values of internal branches, had ESS values >200. We used DENSITREE V. 2.2.7 (Bouckaert, 2010) to plot the sampled posterior tree topologies for all combined runs. We summarized the posterior distribution into a Maximum Clade Credibility (MCC) tree using TREEANNOTATOR (part of the BEAST V. 2 2.6.1 package).

For the maximum likelihood phylogeny, we selected the same individuals analysed using SNAPP, but given that this is a concatenation-based analysis, we did not filter SNPs by linkage disequilibrium. The dataset consisted of 2,270 variable sites with no missing information. To estimate the best substitution model, we used MODELFINDER (Kalyaanamoorthy et al., 2017) implemented in IQTREE V. 2.1.2, and applied the Ascertainment Bias Correction (+ASC) as this is strongly recommended when analysing SNP data (Lewis, 2001). The best substitution model estimated using the Bayesian Information Criterion (BIC) was K3Pu+F+ASC+R2. To evaluate branch support values, we used 10,000 ultrafast bootstraps (Hoang et al., 2017) optimized by NNI, and we used 10,000 replicates for the Shimodaira-Hasegawa-like approximate likelihood test (SH-aLRT) (Guindon et al., 2010).

2.6 | Gene flow among rockhopper taxa

In the light of previous studies suggesting the existence of contemporary gene flow among rockhopper penguins (Mays et al., 2019), we used BAYESASS V. 3.0.4 (Wilson & Rannala, 2003) to test for recent migration between the three taxa using an unlinked dataset (Table S3). BAYESASS uses a Bayesian approach with MCMC to estimate migrant ancestries of each individual based on an assignment method and infers migration rates based on the frequency of individuals exhibiting migrant ancestry. We set the mixing parameters for allele frequencies and inbreeding coefficients to 0.25 and 0.05,

respectively, after achieving acceptance rates of 20%–60% on the test run. The MCMC was run for 50,000,000 iterations, discarding the first 4,000,000 as burn-in, with a sampling interval of 4,000. We performed five independent runs with different random seeds and compared the results from each run. We checked for convergence using TRACER V. 1.7.1 (Rambaut et al., 2018). We plotted the results from BAYESASS in R V. 3.6.3 (R Core Team, 2020) using TIDYR V. 1.0.2 (Wickham & Henry, 2018) and DPLYR V. 0.8.5 (Wickham et al., 2018) to prepare the data and used CIRCLIZE V. 0.4.13 (Gu et al., 2014) to plot the results.

3 | RESULTS

3.1 | SNP calling

For all ddRAD datasets, the mean effective per-sample depth of coverage ranged from 73.0× to 98.2×, with a minimum per-sample coverage of >18.2× (Table S5). We did not find any evidence for a site being under selection or out of Hardy–Weinberg equilibrium in more than ~50% of the populations or under different grouping criteria (described on Table S3). The total number of SNPs retained after each filtering step and the final number of SNPs from each dataset (full datasets ranging from 2,975 to 14,226 SNPs; unlinked datasets ranging from 898 to 4,674 SNPs) are presented in Table S3.

3.2 | Genetic diversity and differentiation

All populations of *E. moseleyi*, *E. chrysocome* and *E. filholi* exhibited private alleles, ranging from 63 (2.7%) for *E. chrysocome* on the Falkland/Malvinas Islands to 11 (0.47%) for *E. chrysocome* and *E. filholi* on Staten and Crozet islands respectively (Table S6). When we calculated genetic diversity within each taxon, we determined that around 10% of the alleles were private to each taxon (*E. moseleyi* =121, *E. chrysocome* =99 and *E. filholi* =106).

Among populations of each taxon, pairwise F_{ST} values were low, ranging from 0.002 to 0.053, and most were not significantly differentiated (Figure 2a). The highest intraspecific differentiation was found among *E. moseleyi* (0.053) populations; F_{ST} values within *E. chrysocome* (F_{ST} =0.002–0.032) and *E. filholi* (F_{ST} =0.004–0.013) were consistently low. In contrast, interspecific F_{ST} values were all significant, ranging from 0.140 to 0.497. The highest F_{ST} was recovered for pairwise comparisons between *E. moseleyi* and populations of *E. chrysocome*/*E. filholi* (Figure 2a; Table S7).

Manhattan plots revealed clear contrasts in the extent of genomic differentiation across sites within versus among taxa. Specifically, intraspecific Manhattan plots exhibited a narrower distribution of F_{ST} values than those involving different taxa. The maximum F_{ST} for the intraspecific calculations was ~0.4 between *E. moseleyi* populations. Considering 0.4 as a cut-off, Manhattan plots between *E. moseleyi* and *E. chrysocome*/*E. filholi* revealed high variation in F_{ST} values across the genome, with around 800 sites exhibiting F_{ST} values >0.4

(Figure 3a,b). *E. chrysocome* and *E. filholi* exhibited lower F_{ST} values, although 64 sites with F_{ST} >0.4 were detected (Figure 3c). In contrast, intraspecific differentiation across genomic regions was characterized by far lower F_{ST} values (Figure 3d,e,f); the only site exhibiting F_{ST} >0.4 corresponded to a comparison between two *E. moseleyi* populations.

3.3 | Clustering analyses

A PCA including all three taxa revealed three separate, tightly clustered groups according to species boundaries (Figure 2b). Specifically, the first principal component separated *E. moseleyi* from *E. chrysocome* and *E. filholi*, explaining 45.45% of the variance, and the second principal component separated *E. filholi* from the *E. chrysocome* populations, explaining 22.19% of the variance. When a PCA was conducted with just *E. filholi* and *E. chrysocome*, the first principal component separated these taxa, explaining 43.34% of the total variance, and the second principal component separated individuals from populations of *E. chrysocome* from the Falkland/Malvinas Islands from those on Terhalten and Staten islands, explaining 7.8% of the variance (Figure 2c). At the population level, PC1 (16.42%) separated *E. moseleyi* populations at Nightingale and Amsterdam islands (Figure 2d) and the *E. chrysocome* (12.01%) population on the Falklands/Malvinas Islands from the neighbouring mainland Patagonian populations (Figure 2e). In contrast, *E. filholi* did not exhibit clear geographic structure among the sampled populations (Figure 2f).

Similar results were observed when the SNPs were analysed using FINERADSTRUCTURE (Figure 4a): three groups were identified by shared ancestry, each corresponding to one of the three rockhopper taxa. These analyses revealed no evidence of hybrid individuals; that is, no individuals exhibited similar levels of co-ancestry between groups (Malinsky et al., 2018).

For STRUCTURE analyses of *E. moseleyi*, *E. chrysocome* and *E. filholi* performed in the same run (Figure 4b), the mean LnP(K) identified the best *K* was four; however, no additional groups were formed when K > 3 (Figure S3). On the other hand, the Evanno method suggested two clusters (Table S8), reflecting the relatively deep genetic differentiation between *E. moseleyi* and the remaining rockhopper taxa. Across all successive *K*, we did not find evidence of shared ancestry among the three rockhopper penguin taxa (Figure S3). We obtained similar results from testing *E. chrysocome* and *E. filholi*, without including *E. moseleyi*, despite using different allele frequency models. Both methods indicated that K = 2 was the best *K* (Table S8), and no additional structure was identified for K > 2 (Figure S4a,b).

In the intraspecific analyses, K = 2 exhibited a higher probability than K = 1 in *E. moseleyi*, corresponding to separation of the populations on Nightingale and Amsterdam islands (Table S8). We found low levels of admixture between individuals from both populations (Figure 4c). For *E. chrysocome*, the optimal *K* suggested by both methods was K = 2, where one group corresponded
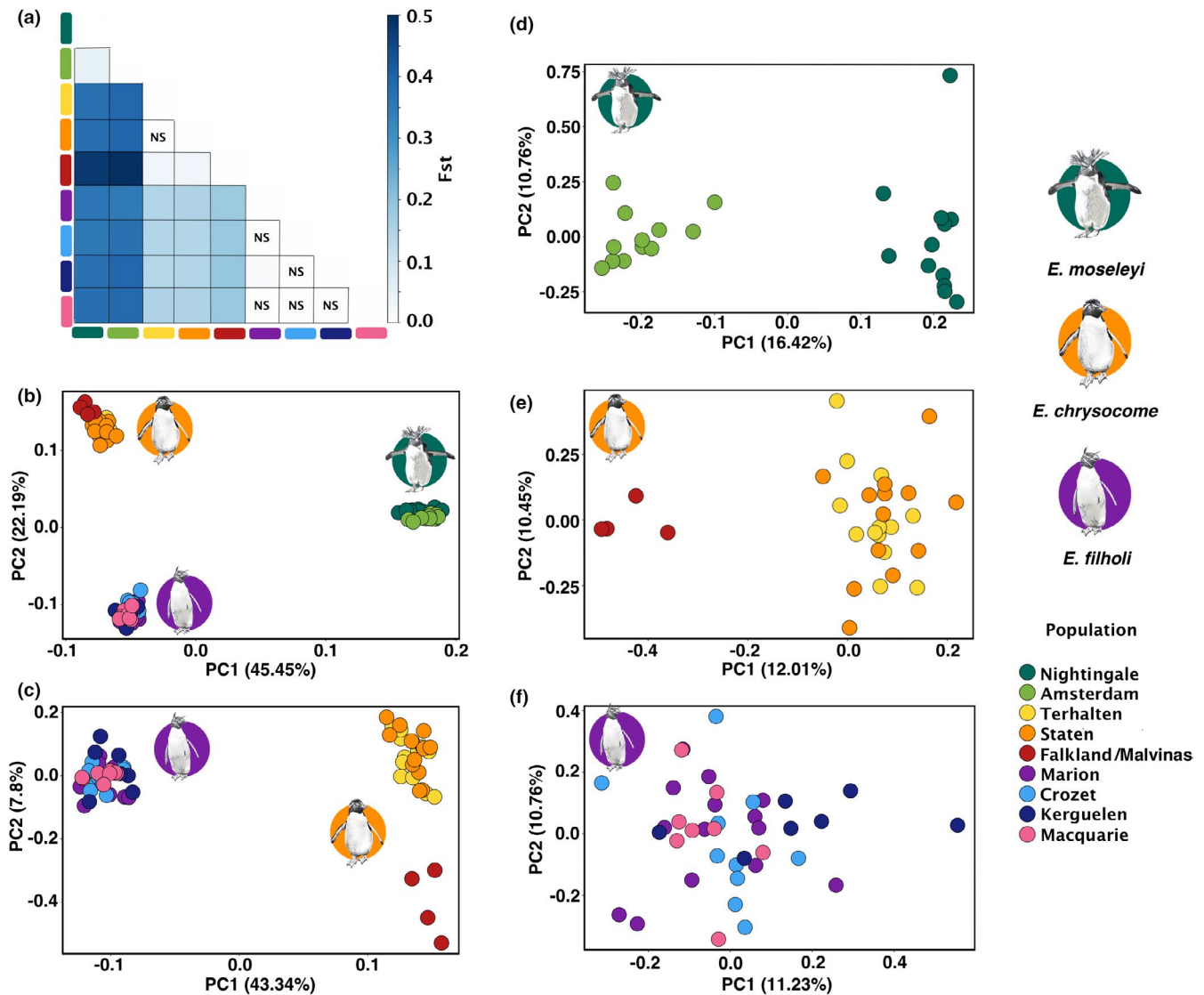
FIGURE 2 Genetic differentiation within and among rockhopper penguin taxa. (a) Pairwise F_{ST} among all sampled colonies. NS, not significant. (b) PCA results of individuals from the three taxa (*E. moseleyi*, *E. chrysocome* and *E. filholi*). (c) PCA between *E. chrysocome* and *E. filholi*. (d) Intraspecific PCA of *E. moseleyi* individuals. (e) Intraspecific PCA of *E. chrysocome* individuals. (f) Intraspecific PCA of *E. filholi* individuals. Population codes (colour) are indicated on the side of the plot. *Eudyptes moseleyi* populations: Amsterdam and Nightingale, *E. chrysocome* populations: Terhalten, Staten and Falkland/Malvinas, *E. filholi* populations: Marion, Crozet, Kerguelen and Macquarie. Replicas of these figures with an alternative colour scheme and additional visual patterns are available in Appendix S1 (section IV: Alternative figures)

to the Falkland/Malvinas population and the other to those of Staten/Terhalten islands. We found low levels of admixture between these groups (Figure 4d) but did not detect any population structure across the distributional range of *E. filholi*. The Evanno method suggested $K = 3$, but closer inspection revealed that ancestries for each cluster were evenly distributed among all individuals, indicative of a panmictic population (i.e. $K = 1$; Figure S5). The highest value of the mean LnP(K) also supported $K = 1$ (Figure 4e; Table S8). Finally, AMOVA revealed that 22.93% ($p =.026$) of the molecular variance was explained by differentiation between *E. chrysocome* and *E. filholi*, whereas only 1.63% ($p <.01$) was explained by differentiation among samples within each taxon (Table S9).

3.4 | Species delimitation and phylogeographic reconstruction

The BFD* analysis for both independent runs using a different subset of individuals yielded similar results; the most favoured model was the one with four species (i.e. the three rockhopper species hypothesis +macaroni penguins; Table 1).

The phylogenetic tree inferred using the coalescent method implemented in SNAPP recovered *E. filholi* sister to *E. chrysocome*, and this clade sister to *E. moseleyi*, all with high posterior probabilities ($PP = 1$). In contrast, posterior probabilities were much lower for the nodes within species (i.e. populations). This tree topology was highly consistent among different tree estimations (Figure 5a). The maximum
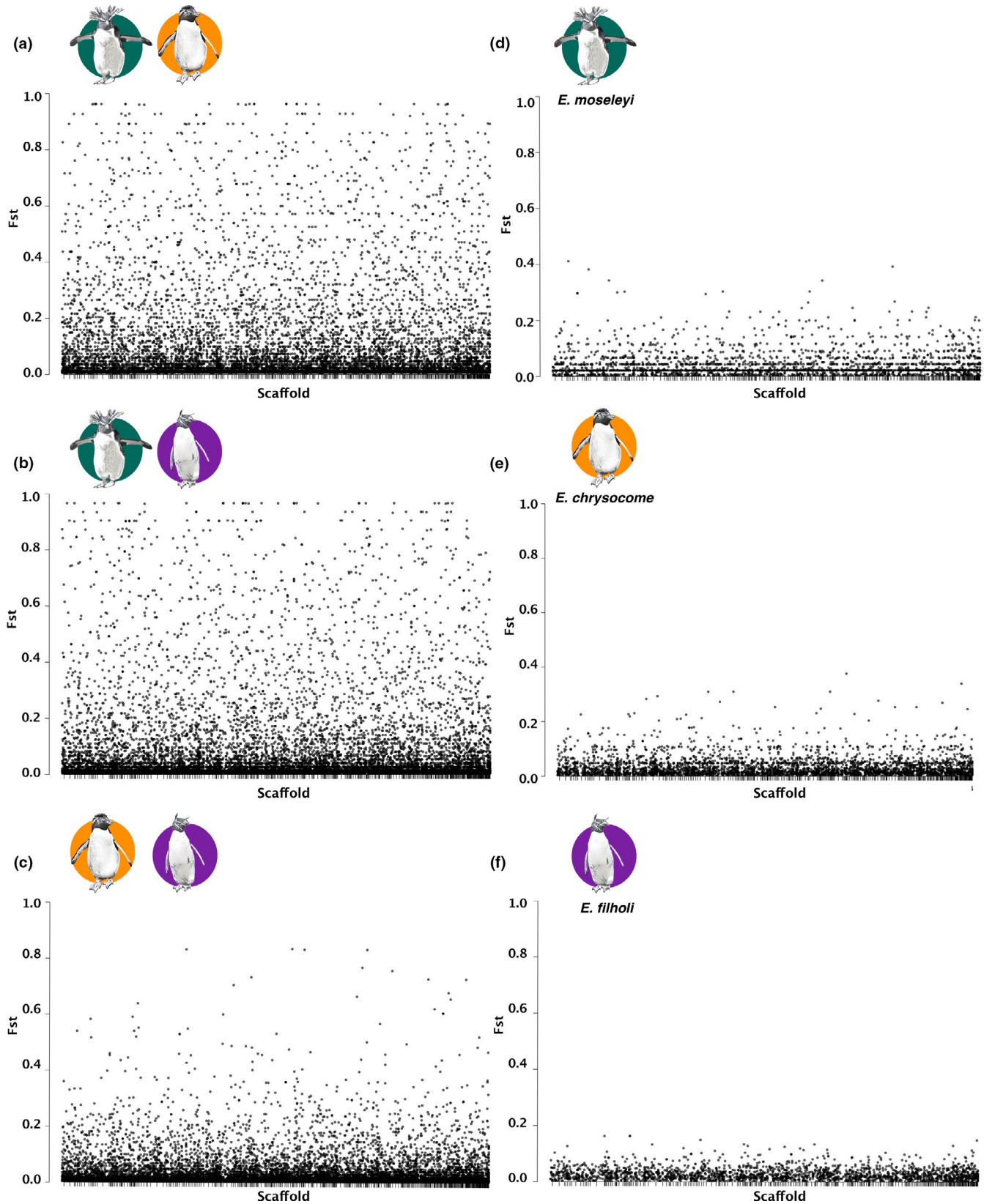
FIGURE 3 Manhattan plots comparing interspecific versus intraspecific differentiation across genomic regions. Interspecific locus-by-locus pairwise F_{ST} values are shown for (a) *E. moseleyi–E. chrysocome*, (b) *E. moseleyi–E. filholi* and (c) *E. chrysocome–E. filholi*. Intraspecific locus-by-locus F_{ST} (mean) is shown for (d) *E. moseleyi*, (e) *E. chrysocome* and (f) *E. filholi*
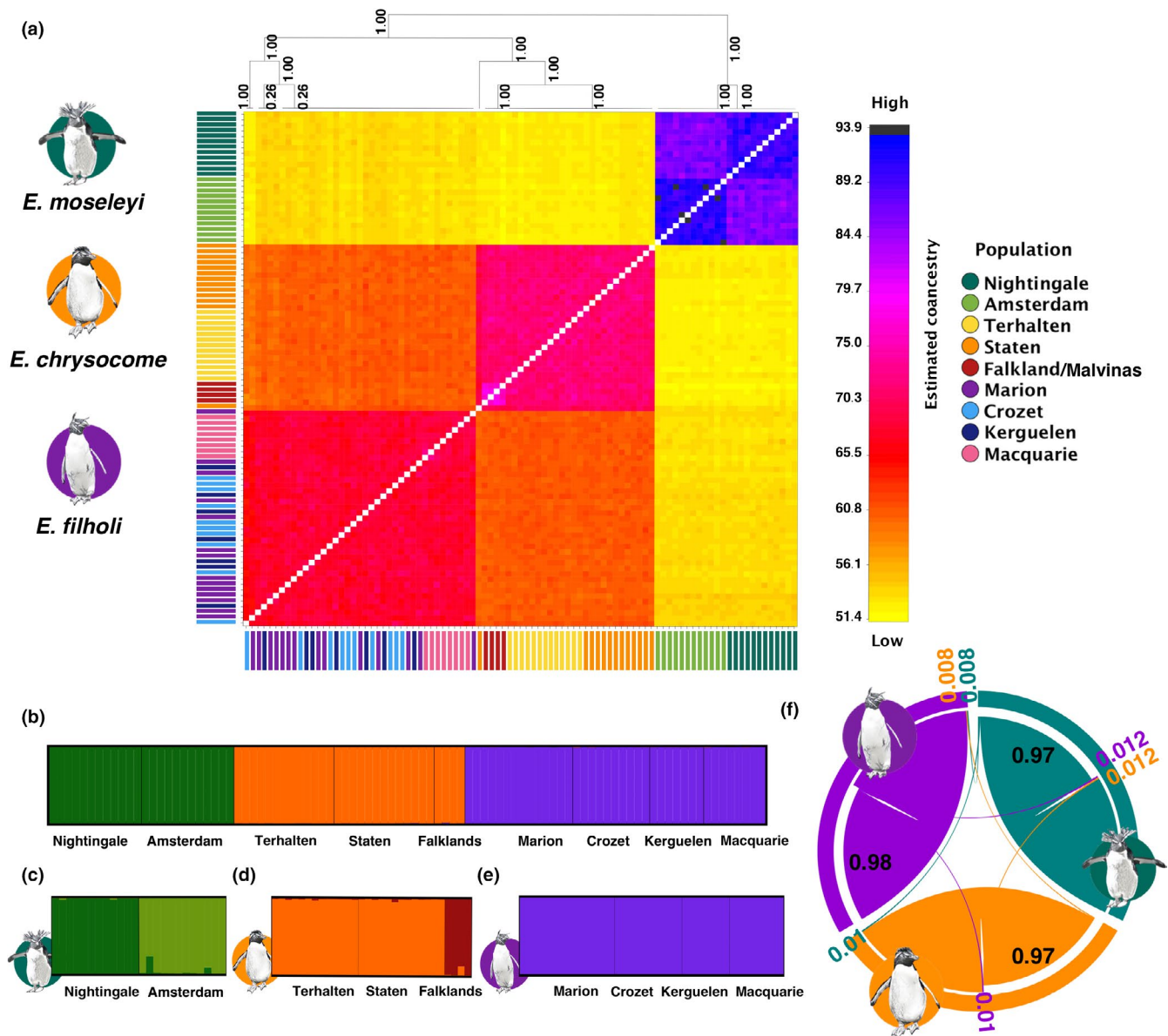
FIGURE 4 Genetic clustering analyses and estimated migration rates among rockhopper penguins. (a) Co-ancestry matrix derived by FINERADSTRUCTURE showing the estimated co-ancestry among all individuals; relationships between populations are shown in the top of the plot, with values corresponding to posterior assignment probabilities. Population codes for each individual are denoted by the colour of each bar. (b) Structure results (K = 3) involving individuals from the three rockhopper penguins taxa. Intraspecific runs of structure for (c) *E. moseleyi* (K = 2), (d) *E. chrysocome* (K = 2) and (e) *E. filholi* (K = 1). (f) Circular migration plot exhibiting, for each taxon, the fraction of migrant individuals (size of the band and numbers on each band) per generation and its origin. Replicas of this figure with an alternative colour scheme and additional visual patterns are available on Appendix S1 (section IV: Alternative figures)

likelihood reconstruction yielded similar results: a similar topology with reciprocal monophyly for each taxon with high support values (SH-aLRT and ultrafast bootstrap support values ~100%; Figure 5b).

3.5 | Gene flow among rockhopper taxa

Migration rates, estimated as the fraction of migrant individuals per generation and calculated using BAYESASS, yielded almost identical results across different runs, starting with different random seeds (Table S10). The proportion of migrant individuals between each taxon (all results expressed as mean, 95% credibility interval) are: *E. moseleyi* to *E. chrysocome*, 0.010 (−0.009–0.027); *E. chrysocome* to *E. moseleyi*, 0.012 (−0.012–0.036); *E. moseleyi* to *E. filholi*, 0.008 (−0.008–0.024); *E. filholi* to *E. moseleyi*, 0.012 (−0.012–0.036); *E. chrysocome* to *E. filholi*, 0.008 (−0.008–0.024); and *E. filholi* to *E. chrysocome*, 0.010 (−0.009–0.027). These results are summarized in Figure 4f.

4 | DISCUSSION

4.1 | Molecular evidence for species delimitation

One of the challenges in delimiting recently divergent species is the absence of a clear genomic differentiation threshold for distinguishing intraspecific versus interspecific lineages. Pragmatically, comparing patterns observed among related well-delimited species sharing similar ecological, behavioural and life-history traits with those observed in the taxa under consideration is one way to help guide taxonomic decisions, especially within an integrative taxonomic framework. We therefore investigated genomic differentiation among *E. moseleyi*, *E. chrysocome* and *E. filholi*, a species complex of penguins whose taxonomy has been the subject of considerable debate for several decades. Our results consistently provided support for the recognition of three species of rockhopper penguins. The global PCA clearly distinguished the subtropical *E. moseleyi* from sub-Antarctic *E. chrysocome* and *E. filholi* (first axis; 45% of genetic variance; Figure 2b) and also separated *E. chrysocome* from *E. filholi* (second axis 22%; Figure 2b). The genetic variance explaining the separation of *E. chrysocome* from *E. filholi* was even higher when performing the PCA not including *E. moseleyi* (First axis 43%; Figure 2c). In contrast, intraspecific comparisons explained relatively small proportions of genetic variance (11%–16%; Figure 2d,e,f). This contrast was also evident in the range of inter-taxon F_{ST} (0.140–0.497) relative to intraspecific values (0.002–0.053). Similarly, when testing F_{ST} across sites throughout the penguin genomes, several sites exhibited moderate to high genetic differentiation ($F_{ST} > 0.4$) when comparing taxa, in contrast to intraspecific comparisons where a much narrower distribution of F_{ST} values occurred (most sites exhibited $F_{ST} < 0.2$ and almost all were <0.4; Figure 3).

The range of intraspecific variation observed in our study is similar to that reported previously for rockhopper penguins (Cole, Dutoit, et al., 2019; Cole, Ksepka, et al., 2019) and within other penguin species (broader discussion on Appendix S1 section III). For example, a recent genomic study found that the intraspecific variance explained (PCA1) among individuals from different populations of king (*Aptenodytes patagonicus*), emperor (*A. forsteri*), chinstrap (*Pygoscelis antarctica*) and Adélie (*P. adeliae*) penguins ranged from 1.28% to 2.02%, and F_{ST} ranged from 0.002 to 0.003 (Clucas et al., 2018). Gentoo penguins (*P. papua*), on the other hand, exhibited greater genomic structure among colonies geographically isolated on different islands (15.81% and $F_{ST} = 0.217$). However, the analysis of gentoo penguins involved divergent lineages that were recently proposed to represent at least four distinct species, based on their genomic divergence and fine-scale morphometric differences (Clucas et al., 2018; de Dinechin et al., 2012; Pertierra et al., 2020; Tyler et al., 2020; Vianna et al., 2017). Another recent study using mitochondrial genomes from all penguin species found that the range of sequence divergence among well-defined species is 0.8%–5.2% (Cole, Dutoit, et al., 2019; Cole, Ksepka, et al., 2019). These authors found that *E. chrysocome* and *E. filholi* are 0.7% divergent,

while the range of differentiation for macaroni (*E. chrysolophus)* and royal (*E. schlegeli*) penguins, which probably comprise a single species, is 0.2% (Cole, Dutoit, et al., 2019; Cole, Ksepka, et al., 2019; Frugone et al., 2019). Thus, compared to several other penguin species, the range of variation we found among the three rockhopper taxa exceeded the expected values for intraspecific variation across genomic regions. Furthermore, our AMOVA results showed that the variance explained by grouping populations to reflect each taxon was ~23%, whereas differentiation among populations within each taxon was only around 2%.

In further support of recognizing three rockhopper penguin species, we did not find evidence of admixture among the three taxa in STRUCTURE and FINERADSTRUCTURE. Also, our BAYESASS results suggest that only ~0.01 individuals per generation are migrants from another taxon. However, because all credibility intervals overlapped with 0, we are unable to reject a scenario of zero gene flow and, hence, complete genetic isolation among the three taxa. We found high proportions of private alleles (~10%) within *E. moseleyi*, *E. chrysocome* and *E. filholi*, consistent with little or no gene flow among the three rockhopper taxa. The apparent absence of admixture, the high proportions of private alleles and the 13–17 mutational steps separating their mtDNA haplogroups (Frugone et al., 2018) together provide strong support for the historical isolation of these taxa. While it is possible that hybrid individuals have been unsampled (Crofts & Robson, 2015), hybridization among these three taxa appears to be negligible. Low frequencies of hybridization have been reported between other penguin sister lineages that are accepted as distinct species, such as the Humboldt (*Spheniscus humboldti*) and Magellanic (*S. magellanicus*) penguins (Simeone et al., 2009), and among several species of *Eudyptes* penguins (Hull & Wilthshire, 1999; Napier, 1968; White & Clausen, 2002), including hybridization between *E. filholi* and *E. chrysolophus* (Woehler & Gilbert, 1990) and *E. chrysocome* and *E. chrysolophus* (White & Clausen, 2002). Further, high levels (up to 25%) of ancestral introgression were detected among several *Eudyptes* species (Vianna et al., 2020). In this sense, the strong genomic differentiation found among the three rockhopper penguins may be the result of their geographic isolation and not necessarily implying reproductive isolation, and thus, it is not possible to discard the possibility of hybridization if they come into geographic contact.

Finally, the best model found by our species delimitation analyses using SNP data supported the existence of three species of rockhopper penguins. Despite earlier reports that the multispecies coalescent model as implemented in SNAPP and other programs such as BPP (Yang & Rannala, 2010) might delimit population structure and not species (Sukumaran & Knowles, 2017), our comprehensive population-level sampling of all three putative rockhopper penguin taxa enabled us to clearly delineate taxonomic boundaries, separating intraspecific from interspecific variation. The recognition of three rockhopper penguin species is also supported using different species delimitation methods including the tree-based ABGD (Puillandre et al., 2012), the threshold-based GMYC (Pons et al., 2006)—both of which are presented in Frugone et al. (2018)—and the BPP (Yang & Rannala, 2010) analyses presented
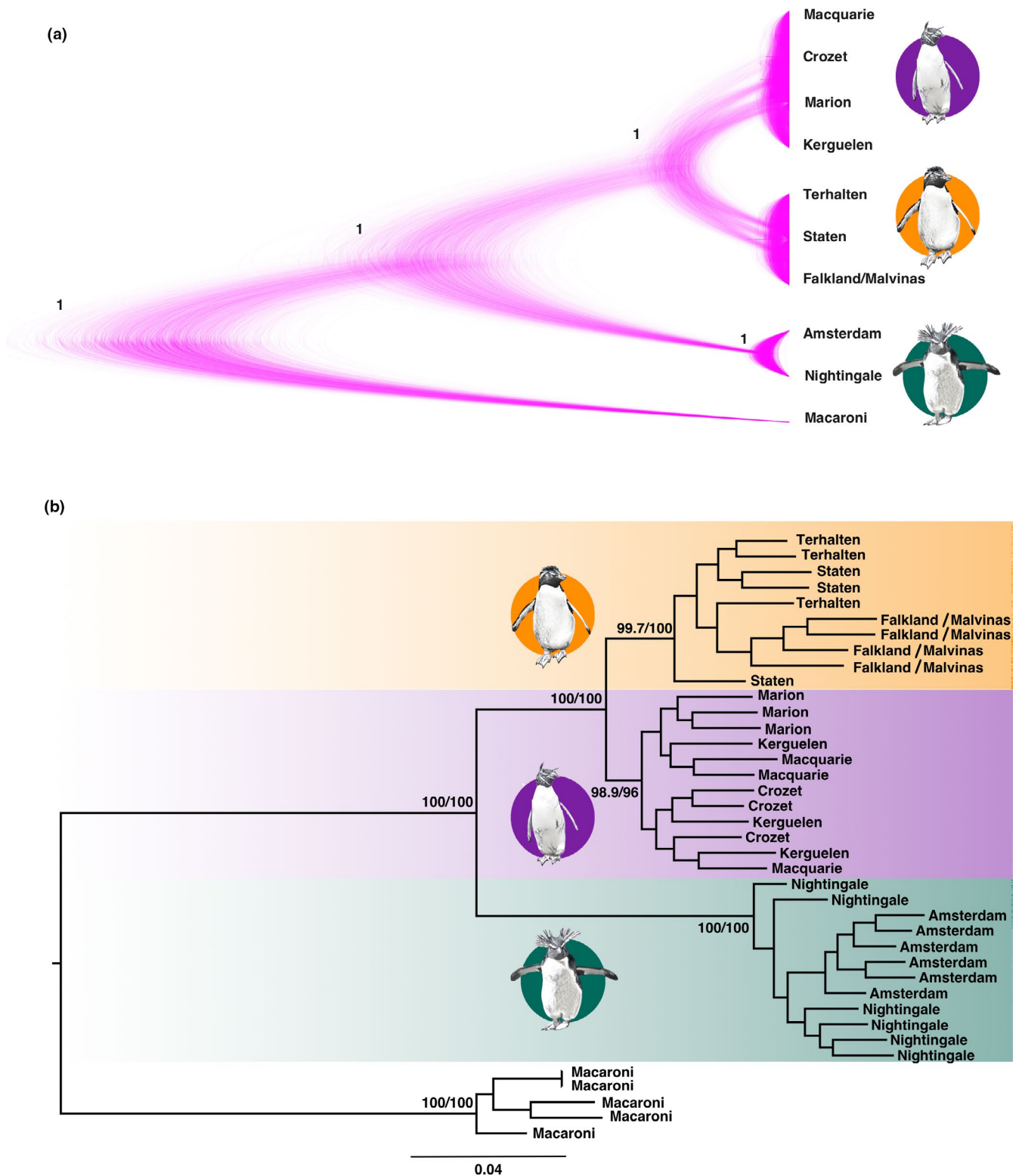
FIGURE 5 Phylogenetic relationships among rockhopper penguins. (a) Posterior distribution of different topologies yielded by SNAPP analyses. (b) Maximum likelihood reconstruction from IQ-tree. Posterior probabilities are indicated on the nodes of each phylogeny

by Mays et al. (2019). Mays et al. (2019) rejected the findings of their BPP analyses because their 6-locus dataset yielded ambiguous results and were unable to reliably distinguish any taxa based on the GDI (Jackson et al., 2017; Leache et al., 2019). This result

was surprising because it also failed to recognize even the highly divergent and widely recognized species *E. chrysolophus*, suggesting that their six-locus dataset was underpowered for evaluating species boundaries in these recently diverged penguin species.

In summary, at the molecular level, several lines of evidence support recognition of three species of rockhopper penguins: 1) genomic differentiation among *E. moseleyi*, *E. chrysocome* and *E. filholi* substantially exceeds differentiation observed within each taxon; 2) there is no genomic evidence of admixture or hybrid individuals among these taxa; 3) our analyses suggest no contemporary gene flow among the three taxa; and 4) species delimitation analyses strongly support a three-species designation to the exclusion of a model designating *E. chrysocome*/*E. filholi* as conspecific. Our results, thus, confirm the findings of previous studies revealing reciprocal monophyly with mtDNA markers, supporting this three-species designation (Banks et al., 2006; Cole, Dutoit, et al., 2019; Cole, Ksepka, et al., 2019; de Dinechin et al., 2009; Frugone et al., 2018; Mays et al., 2019).

Our analyses provide insights into the causes of discordance among nuclear/mitochondrial markers leading to different conclusions regarding species delimitation in this group. Previous studies (e.g. Frugone et al., 2018; Mays et al., 2019) included sampling from the entire range of each taxon, and the high mtDNA divergence found among taxa is therefore not a consequence of sparse sampling, which may result in isolation-by-distance being undetected (Chambers & Hillis, 2020). In this study, we used thousands of unlinked genetic markers and found that results are in accordance with the pattern exhibited across all studies using mtDNA. This finding suggests that the previous discordance among studies did not result from sex-biased dispersal of rockhopper penguins, but rather from a lack of power owing to low levels of divergence in the nuclear markers previously utilized. The value of using a few nuclear markers for species delimitation is likely to be further eroded under introgression scenarios, and/or when lineage sorting among closely related lineages is incomplete.

4.2 | Integrative taxonomy

Over-splitting of lineages potentially presents a valid concern in taxonomy given the power of modern genomic methods to detect rapidly evolved, fine-scale differentiation among populations. One could argue that should enough SNPs be used in an analysis, any two populations could be delimited through either different allele frequencies or by a few fixed mutational differences. As such, under an integrative taxonomic framework, it is important that additional lines of evidence (e.g. ecology, phenotypes, behaviour, geography, parasites) are investigated and contrasted with molecular species delimitations (Matos-Maraví et al., 2019; Solís-Lemus et al., 2015). In the case of rockhopper penguins, additional factors in support of the molecular results that reinforce the recognition of three rockhopper penguin species include: (a) the non-overlapping geographical distributions of each lineage—*E. moseleyi* (subtropical area in the Atlantic and Indian Oceans), *E. chrysocome* (sub-Antarctic area of the eastern Pacific and Atlantic Oceans) and *E. filholi* (sub-Antarctic area of the western Pacific and Indian oceans; Figure 1); (b) different dispersal capabilities in winter, with *E. filholi* traversing greater distances than

E. chrysocome (Raya Rey et al., 2007; Thiebot et al., 2012); (c) the distinct species of parasitic chewing lice found on each taxon, which is consistent with little to no gene flow among taxa (Banks et al., 2006); (d) differences in the timing of reproduction, starting about 2 months earlier for *E. moseleyi* populations, with *E. chrysocome* and *E. filholi* colonies exhibiting variation in the timing of the onset of reproduction (Clausen & Pütz, 2002; Hiscock, 2013; Hull et al., 2004; Morrison, 2016; Raya Rey et al., 2009); and (e) the presence of morphological and vocal differences among taxa. *Eudyptes moseleyi* can be distinguished from the sub-Antarctic *E. chrysocome*/*E. filholi* by its larger size and longer crests (Banks et al., 2006), whereas the main morphological differentiation between the sub-Antarctic taxa corresponds to a more pronounced pink to white gape (bare skin around the bill) on *E. filholi* which is black on *E. chrysocome* (Tennyson & Miskelly, 1989). Other morphological features differentiating *E. filholi* from *E. chrysocome* include a narrower bill and the shape of the black mark on the undersurface of the apex of the wing (Hutton, 1879). Finally, vocalizations are an important behavioural trait in penguins, since they may rely more strongly on auditory cues for mate selection and individual recognition than on morphological characters (Aubin & Jouventin, 2002; Jouventin et al., 2006; Searby & Jouventin, 2005). Differences have been found in the mating calls of *E. moseleyi* in comparison with those of *E. chrysocome* (when considered a single species to *E. filholi*), providing an additional line of evidence for delimiting those taxa (Jouventin et al., 2006). However, as far as we know, vocalizations of *E. chrysocome* and *E. filholi* have not been directly compared and should be a priority for future data collection and study.

There are several well-documented historic cases of over-lumping of species due to the absence of obvious morphological variation among taxa (Adams et al., 2014; Bickford et al., 2007; Grosser et al., 2015). The absence of apparent diagnostic morphological variation in such cryptic or sibling species may be explained by recent diversification, morphological constraints imposed by selective pressures, or by convergence (Fiser et al., 2018). Conversely, high variability in morphological characters within a species may lead to over-splitting taxa. For example, in birds, variation in meristic or plumage characters may result from local (population) adaptation across an environmental gradient (Mason & Bowie, 2020) or could be due to phenotypic plasticity (e.g. Mason & Taylor, 2015). Penguin species exhibit very low variation in their diet, and have similar ecology, behaviour and morphology when compared to other groups of birds, such as ducks (Anseriformes), shorebirds (Charadriiformes), pheasants (Galliformes) or songbirds (Passeriformes) (Winkler et al., 2015). Despite not overlapping in their distributions, both *E. chrysocome* and *E. filholi* breed in similar marine environments, although in different ocean basins, prey on similar items, and differ very slightly in their timing of reproduction (Clausen & Pütz, 2002; Hull, 1999). These aspects are also shared with other penguin species distributed in the sub-Antarctic area. For example, similarities in diet and reproductive timing are also present when comparing *E. chrysocome* and *E. filholi* with other *Eudyptes* species like the macaroni penguin (Thiebot et al., 2011; Williams & Croxall, 1991). This implies that

because *E. chrysocome*, *E. filholi* and several other penguin species overlap in ecological niche space (e.g. oceanographic conditions over their distributional range, diet and breeding period), it may constrain differentiation in some phenotypic traits (Fiser et al., 2018) if they have moved towards the same adaptive peak. The extent of niche overlap might also explain the limited morphological differentiation among *Eudyptes* penguins, not only for the rockhopper penguin, but also among other penguin species complexes (Tyler et al., 2020). In this sense, the use of genomic approaches with extensive sampling across the species complex range is critically important if inter- versus intraspecific taxonomic boundaries are to be delineated.

Most bird species were initially delimited under the Morphological Species Concept (MSC) (Cronquist, 1978) and/or Biological Species Concept (BSC) (Cracraft, 1992; Mayr, 1942). However, under the MSC, cryptic species or subtle morphological variation can lead to underestimation of species diversity, and hence, a poor understanding of the underlying evolutionary history, which can have significant impacts on the management of threatened taxa. On the other hand, in birds, hybridization might be more common than in other groups and has been documented in ~16% of their species (Ottenburghs et al., 2015). Furthermore, multispecies hybridization (i.e. hybridization between more than two species) might also be common in this group (Ottenburghs, 2019). In this sense, divergent bird species that occasionally hybridize might be considered conspecific under the BSC, given that it assumes reproductive isolation. Moreover, to test for the BSC criterion of reproductive isolation in allopatric taxa may be difficult without the use of molecular markers to estimate gene flow. The phylogenetic species concept (PSC) (Cracraft, 1983; Nelson & Platnick, 1981) clarified the taxonomy of several bird species (e.g. Cracraft, 1992) including penguins, facilitated by the advances in molecular biology and genomics. For example, macaroni and royal penguins were considered by several authors as different species mainly based on their allopatric distribution and morphological differentiation. However, genomic data revealed absence of reciprocal monophyly (Frugone et al., 2018, 2019), suggesting that they should be considered a single species. Here, we recognized the three rockhopper species under the unified species concept proposed by De Queiroz (2007) within an integrative taxonomy framework (Cicero et al., 2021; Sangster, 2014), which reconciles elements from several species concepts including the MSC, BSC and the PSC. We suggest the distinction of three rockhopper species based on the criteria described above: concordance and genealogical agreement across multiple loci (e.g. mtDNA and genomic data) with sampling encompassing most of the species distributional ranges, concordance with biogeographical patterns, morphological characters and no evidence of recent gene flow.

Population structure could be viewed as a first step in the speciation process. In general, seabirds, especially those inhabiting the Southern Ocean, exhibit very low levels of genetic structure (e.g. Burg & Croxall, 2001; Milot et al., 2008; Quillfeldt et al., 2017). Our results, however, reveal an opposite pattern and highlight that even widespread seabirds with great dispersal potential can house cryptic diversity previously undetected due to limited morphological differentiation. We also suggest that it is especially important to understand the role of introgression, as seabird colonies may have been connected multiple times in the past, and account for this by sampling across the genome to have confidence in species delimitation.

4.3 | Conservation implications

Declining numbers recently observed for populations of all three rockhopper penguin species are of major concern for contemporary conservation. Rising oceanic temperatures and associated oceanographic changes, the introduction of predators and overfishing are among the threats thought to have influenced these demographic declines (Crawford et al., 2003; Cuthbert et al., 2009; Trathan et al., 2015; Wilson et al., 2010). Currently, *E. moseleyi* is considered endangered (BirdLife International, 2020b) and *E. chrysocome*/*E. filholi*, as a single taxon, is considered vulnerable (BirdLife International, 2020a). Delimiting *E. chrysocome* and *E. filholi* as distinct species may heighten the severity of their conservation status due to more restricted ranges and smaller population sizes and acknowledge their independent evolutionary histories. *Eudyptes filholi* has a wider geographical range than *E. chrysocome*, and census data suggest smaller population sizes for *E. filholi* (422,000 breeding pairs) than *E. chrysocome* (850,000 breeding pairs). Overall, while populations of both species are decreasing, *E. filholi* has experienced particularly severe population contractions over recent decades (BirdLife International, 2020a). Whereas some authors suggest that the largest *E. chrysocome* breeding colony in the Falkland/Malvinas Islands has increased (Baylis et al., 2013), *E. filholi* populations have significantly declined across most of the species distributional range, including Marion (Crawford et al., 2003), Campbell (Morrison et al., 2015) and Crozet islands (Barbraud et al., 2020). Following the recommendations of the most recent IUCN report (BirdLife International, 2020a), we also suggest the need for continuing monitoring of population trends for these two species separately, and especially for *E. filholi*. Additionally, it is of utmost importance to prioritize research efforts aiming to understand the underlying causes of these severe population declines. In this sense, the accurate delimitation of these birds is of paramount importance to the conservation of rockhopper penguins (Bickford et al., 2007; Cuthbert et al., 2009) and will be necessary for effective management decisions.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13399.

DATA AVAILABILITY STATEMENT

DNA Sequences: BIOPROJECT: PRJNA745960. Methodological specifications and input files used in analyses: https://doi.org/10.5061/dryad.mw6m905v6.

ORCID

María José Frugone 🆔 https://orcid.org/0000-0002-8302-0386
Theresa L. Cole 🆔 https://orcid.org/0000-0002-0197-286X
María Eugenia López 🆔 https://orcid.org/0000-0001-9272-0694
Gemma Clucas 🆔 https://orcid.org/0000-0002-4305-1719
Pável Matos-Maraví 🆔 https://orcid.org/0000-0002-2885-4919
Nicolás A. Lois 🆔 https://orcid.org/0000-0001-5664-0486
Jonathan M. Waters 🆔 https://orcid.org/0000-0002-1514-7916
Rauri C. K. Bowie 🆔 https://orcid.org/0000-0001-8328-6021
Elie Poulin 🆔 https://orcid.org/0000-0001-7736-0969
Juliana A. Vianna 🆔 https://orcid.org/0000-0003-2330-7825

REFERENCES

Adams, M., Raadik, T. A., Burridge, C. P., & Georges, A. (2014). Global biodiversity assessment and hyper-cryptic species complexes: More than one species of elephant in the room? *Systematic Biology*, *63*(4), 518–533. https://doi.org/10.1093/sysbio/syu017

Agapow, P.-M., Bininda-Emonds, O. R., Crandall, K. A., Gittleman, J. L., Mace, G. M., Marshall, J. C., & Purvis, A. (2004). The impact of species concept on biodiversity studies. *The Quarterly Review of Biology*, *79*(2), 161–179. https://doi.org/10.1086/383542

Aubin, T., & Jouventin, P. (2002). How to vocally identify kin in a crowd: The penguin model. *Advances in the Study of Behavior*, *31*(31), 243–277. https://doi.org/10.1016/S0065-3454(02)80010-9

Banks, J., Van Buren, A., Cherel, Y., & Whitfield, J. B. (2006). Genetic evidence for three species of rockhopper penguins, *Eudyptes chrysocome*. *Polar Biology*, *30*(1), 61–67. https://doi.org/10.1007/s00300-006-0160-3

Barbraud, C., Delord, K., Bost, C. A., Chaigne, A., Marteau, C., & Weimerskirch, H. (2020). Population trends of penguins in the French Southern Territories. *Polar Biology*, *43*(7), 835–850. https://doi.org/10.1007/s00300-020-02691-6

Baylis, A. M. M., Wolfaardt, A. C., Crofts, S., Pistorius, P. A., & Ratcliffe, N. (2013). Increasing trend in the number of southern rockhopper penguins (*Eudyptes c. chrysocome*) breeding at the Falkland Islands. *Polar Biology*, *36*(7), 1007–1018. https://doi.org/10.1007/s00300-013-1324-6

Bazykin, A. D. (1969). Hypothetical mechanism of speciation. *Evolution*, *23*(4), 685–687. https://doi.org/10.1111/j.1558-5646.1969.tb03550.x

Benestan, L. M., Ferchaud, A.-L., Hohenlohe, P. A., Garner, B. A., Naylor, G. J. P., Baums, I. B., Schwartz, M. K., Kelley, J. L., & Luikart, G. (2016). Conservation genomics of natural and managed populations: Building a conceptual and practical framework. *Molecular Ecology*, *25*(13), 2967–2977. https://doi.org/10.1111/mec.13647

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507. https://doi.org/10.1093/biomet/93.3.491

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*(7), e70651. https://doi.org/10.1371/journal.pone.0070651

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, *22*(3), 148–155. https://doi.org/10.1016/j.tree.2006.11.004

BirdLife International (2020a). *Eudyptes chrysocome. The IUCN Red List of Threatened Species 2020. e.T22735250A182762377.*

BirdLife International (2020b). *Eudyptes moseleyi. The IUCN Red List of Threatened Species 2020. e.T22734408A184698049.*

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bouckaert, R. R. (2010). DensiTree: Making sense of sets of phylogenetic trees. *Bioinformatics*, *26*(10), 1372–1373. https://doi.org/10.1093/bioinformatics/btq110

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*(8), 1917–1932. https://doi.org/10.1093/molbev/mss086

Burg, T. M., & Croxall, J. P. (2001). Global relationships amongst black-browed and grey-headed albatrosses: Analysis of population structure using mitochondrial DNA and microsatellites. *Molecular Ecology*, *10*(11), 2647–2660. https://doi.org/10.1046/j.0962-1083.2001.01392.x

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Chambers, E. A., & Hillis, D. M. (2020). The multispecies coalescent over-splits species in the case of geographically widespread taxa. *Systematic Biology*, *69*(1), 184–193. https://doi.org/10.1093/sysbio/syz042

Cicero, C., Mason, N. A., Jiménez, R. A., Wait, D. R., Wang-Claypool, C. Y., & Bowie, R. C. K. (2021). Integrative taxonomy and geographic sampling underlie successful species delimitation. *Ornithology*, *138*(2), 1–15. https://doi.org/10.1093/ornithology/ukab009

Clausen, A. P., & Pütz, K. (2002). Recent trends in diet composition and productivity of gentoo, magellanic and rockhopper penguins in the Falkland Islands. *Aquatic Conservation-Marine and Freshwater Ecosystems*, *12*(1), 51–61. https://doi.org/10.1002/aqc.476

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. *Molecular Ecology*, 27(23), 4680–4697. https://doi.org/10.1111/mec.14896

Cole, T. L., Dutoit, L., Dussex, N., Hart, T., Alexander, A., Younger, J. L., Clucas, G. V., Frugone, M. J., Cherel, Y., Cuthbert, R., Ellenberg, U., Fiddaman, S. R., Hiscock, J., Houston, D., Jouventin, P., Mattern, T., Miller, G., Miskelly, C., Nolan, P., ... Waters, J. M. (2019). Receding ice drove parallel expansions in Southern Ocean penguins. *Proceedings of the National Academy of Sciences*, 116(52), 26690–26696. https://doi.org/10.1073/pnas.1904048116

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. *Molecular Biology and Evolution*, 36(4), 784–797.

Cracraft, J. (1983). Species concepts and speciation analysis. In R. F. Johnston (Ed.), *Current Ornithology* (pp. 159–187). Springer.

Cracraft, J. (1992). The species of the birds-of-paradise (Paradisaeidae): Applying the phylogenetic species concept to a complex pattern of diversification. *Cladistics: The International Journal of the Willi Hennig Society*, 8(1), 1–43. https://doi.org/10.1111/j.1096-0031.1992.tb00049.x

Crawford, R. J. M., Cooper, J., Dyer, B. M., Greyling, M. D., Klages, N. T. W., Nel, D. C., Nel, J. L., Petersen, S. L., & Wolfaardt, A. C. (2003). Decrease in numbers of the eastern rockhopper penguin *Eudyptes chrysocome filholi* at Marion Island, 1994/95–2002/03. *African Journal of Marine Science*, 25(1), 487–498. https://doi.org/10.2989/18142320309504037

Crofts, S., & Robson, B. (2015). First record of hybridisation between northern *Eudyptes moseleyi* and southern rockhopper penguins *E.c.chrysocome*. *Seabird*, 28, 37–42.

Cronquist, A. (1978). Once again, what is a species? In J. A. Romberger (Ed.), *Biosystematics in Agriculture* (pp. 3–20). Montclair, NJ: John Wiley & Sons.

Cuthbert, R., Cooper, J., Burle, M. H., Glass, C. J., Glass, J. P., Glass, S., Glass, T., Hilton, G. M., Sommer, E. S., Wanless, R. M., & Ryan, P. G. (2009). Population trends and conservation status of the northern rockhopper penguin *Eudyptes moseleyi* at Tristan da Cunha and Gough Island. *Bird Conservation International*, 19(1), 109–120. https://doi.org/10.1017/S0959270908007545

Cutter, A. D. (2013). Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Molecular Phylogenetics and Evolution*, 69(3), 1172–1185. https://doi.org/10.1016/j.ympev.2013.06.006

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

de Dinechin, M., Dobson, F. S., Zehtindjiev, P., Metcheva, R., Couchoux, C., Martin, A., Quillfeldt, P., & Jouventin, P. (2012). The biogeography of gentoo penguins (*Pygoscelis papua*). *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, 90(3), 352–360. https://doi.org/10.1139/z2012-016

de Dinechin, M., Ottvall, R., Quillfeldt, P., & Jouventin, P. (2009). Speciation chronology of rockhopper penguins inferred from molecular, geological and palaeoceanographic data. *Journal of Biogeography*, 36(4), 693–702. https://doi.org/10.1111/j.1365-2699.2008.02014.x

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, 56(6), 879–886. https://doi.org/10.1080/10635150701701083

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Edwards, S., & Bensch, S. (2009). Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Molecular Ecology*, 18(14), 2930–2933. https://doi.org/10.1111/j.1365-294X.2009.04270.x

Esquerre, D., Ramirez-Alvarez, D., Pavon-Vazquez, C. J., Troncoso-Palacios, J., Garin, C. F., Keogh, J. S., & Leache, A. D. (2019). Speciation across mountains: Phylogenomics, species delimitation and taxonomy of the *Liolaemus leopardinus* Glade (Squamata, Liolaemidae). *Molecular Phylogenetics and Evolution*, 139, 17. https://doi.org/10.1016/j.ympev.2019.106524

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.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(4), 1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation with gene flow. *Trends in Genetics: TIG*, 28(7), 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Fiser, C., Robinson, C. T., & Malard, F. (2018). Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology*, 27(3), 613–635. https://doi.org/10.1111/mec.14486

Frankham, R., Ballou, J. D., Dudash, M. R., Eldridge, M. D. B., Fenster, C. B., Lacy, R. C., Mendelson, J. R., Porton, I. J., Ralls, K., & Ryder, O. A. (2012). Implications of different species concepts for conserving biodiversity. *Biological Conservation*, 153, 25–31. https://doi.org/10.1016/j.biocon.2012.04.034

Frugone, M. J., López, M. E., Segovia, N. I., Cole, T. L., Lowther, A., Pistorius, P., Dantas, G. P. M., Petry, M. V., Bonadonna, F., Trathan, P., Polanowski, A., Wienecke, B., Bi, K., Wang-Claypool, C. Y., Waters, J. M., Bowie, R. C. K., Poulin, E., & Vianna, J. A. (2019). More than the eye can see: Genomic insights into the drivers of genetic differentiation in Royal/Macaroni penguins across the Southern Ocean. *Molecular Phylogenetics and Evolution*, 139, 106563. https://doi.org/10.1016/j.ympev.2019.106563

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. *Scientific Reports*, 8(1), 17481. https://doi.org/10.1038/s41598-018-35975-3

Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, 27(9), 480–488. https://doi.org/10.1016/j.tree.2012.04.012

García, P., & Boersma, D. (2013). *Penguins: Natural History and Conservation* (pp. 360). University of Washington Press.

Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., & Drummond, A. J. (2017). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Systematic Biology*, 66(1), 57–73. https://doi.org/10.1093/sysbio/syw060

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(12), e0144966. https://doi.org/10.1371/journal.pone.0144966

Gu, Z., Gu, L., Eils, R., Schlesner, M., & Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics*, 30(19), 2811–2812. https://doi.org/10.1093/bioinformatics/btu393

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. https://doi.org/10.1093/sysbio/syq010

Guo, B. C., Fang, B. H., Shikano, T., Momigliano, P., Wang, C., Kravchenko, A., & Merila, J. (2019). A phylogenomic perspective on diversity, hybridization and evolutionary affinities in the stickleback genus Pungitius. *Molecular Ecology*, *28*(17), 4046–4064.

Hiscock, J. A. (2013). *Monitoring penguins in the Antipodes Island Group: Methods and baseline data (Department of Conservation Technical Series 37)*. Department of Conservation.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2017). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Hull, C. L. (1999). The foraging zones of breeding royal (*Eudyptes schlegeli*) and rockhopper (*E. chrysocome*) penguins: An assessment of techniques and species comparison. *Wildlife Research*, *26*(6), 789–803. https://doi.org/10.1071/WR98028

Hull, C. L., Hindell, M., Le Mar, K., Scofield, P., Wilson, J., & Lea, M. A. (2004). The breeding biology and factors affecting reproductive success in rockhopper penguins *Eudyptes chrysocome* at Macquarie Island. *Polar Biology*, *27*(11), 711–720. https://doi.org/10.1007/s00300-004-0643-z

Hull, C., & Wilthshire, A. (1999). An apparent hybrid royal x rockhopper penguin at Macquarie Island. *Australian Bird Watcher*, *183*, 95–100.

Hutton, F. W. (1878). On an apparently new species of penguin, from Campbell Island. *Proceedings of the Linnean Society of New South Wales*, *3*(12), 334–335.

Jackson, N. D., Carstens, B. C., Morales, A. E., & O'Meara, B. C. (2017). Species delimitation with gene flow. *Systematic Biology*, *66*(5), 799–812. https://doi.org/10.1093/sysbio/syw117

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*(14), 1801–1806.

Jouventin, P., Cuthbert, R. J., & Ottvall, R. (2006). Genetic isolation and divergence in sexual traits: Evidence for the northern rockhopper penguin *Eudyptes moseleyi* being a sibling species. *Molecular Ecology*, *15*(11), 3413–3423. https://doi.org/10.1111/j.1365-294X.2006.03028.x

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, *6*, 208. https://doi.org/10.3389/fgene.2015.00208

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. https://doi.org/10.7717/peerj.281

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Key, K. H. L. (1968). The concept of stasipatric speciation. *Systematic Biology*, *17*(1), 14–22. https://doi.org/10.1093/sysbio/17.1.14

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Leache, A. D., Fujita, M. K., Minin, V. N., & Bouckaert, R. R. (2014). Species delimitation using genome-wide SNP data. *Systematic Biology*, *63*(4), 534–542. https://doi.org/10.1093/sysbio/syu018

Leache, A. D., Zhu, T. Q., Rannala, B., & Yang, Z. H. (2019). The spectre of too many species. *Systematic Biology*, *68*(1), 168–181. https://doi.org/10.1093/sysbio/syy051

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, *50*(6), 913–925. https://doi.org/10.1080/106351501753462876

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 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*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299. https://doi.org/10.1093/bioinformatics/btr642

Lois, N. A., Campagna, L., Balza, U., Polito, M. J., Pütz, K., Vianna, J. A., Morgenthaler, A., Frere, E., Sáenz-Samaniego, R., Raya Rey, A., & Mahler, B. (2020). Metapopulation dynamics and foraging plasticity in a highly vagile seabird, the southern rockhopper penguin. *Ecology and Evolution*, *10*, 3346–3355. https://doi.org/10.1002/ece3.6127

Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology*, *46*(3), 523–536. https://doi.org/10.1093/sysbio/46.3.523

Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference from RADseq data. *Molecular Biology and Evolution*, *35*(5), 1284–1290. https://doi.org/10.1093/molbev/msy023

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

Martínez, I. (1992). Order sphenisciformes. In J. del Hoyo, A. Elliott, & J. Sargatal (Eds.), *Handbook of the birds of the world, Ostrich to Ducks* (1, pp. 140–160). Lynxs Edicions.

Mason, N. A., & Bowie, R. C. K. (2020). Plumage patterns: Ecological functions, evolutionary origins, and advances in quantification. *The Auk*, *137*(4), 1–29. https://doi.org/10.1093/auk/ukaa060

Mason, N. A., & Taylor, S. A. (2015). Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. *Molecular Ecology*, *24*(12), 3009–3025. https://doi.org/10.1111/mec.13140

Matos-Maraví, P., Wahlberg, N., Antonelli, A., & Penz, C. M. (2019). Species limits in butterflies (Lepidoptera: Nymphalidae): Reconciling classical taxonomy with the multispecies coalescent. *Systematic Entomology*, *44*(4), 745–756. https://doi.org/10.1111/syen.12352

Mayr, E. (1942). *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. Columbia University Press.

Mays, H. L., Oehler, D. A., Morrison, K. W., Morales, A. E., Lycans, A., Perdue, J., Battley, P. F., Cherel, Y., Chilvers, B. L., Crofts, S., Demongin, L., Fry, W. R., Hiscock, J., Kusch, A., Marin, M., Poisbleau, M., Quillfeldt, P., Raya Rey, A., Steinfurth, A., … Weakley, L. A. (2019). Phylogeography, population structure, and species delimitation in rockhopper penguins (*Eudyptes chrysocome* and *Eudyptes moseleyi*). *The Journal of Heredity*, *110*(7), 801–817. https://doi.org/10.1093/jhered/esz051

Meirmans, P. G., & Van Tienderen, P. H. (2004). Genotype and genodive: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, *4*(4), 792–794. https://doi.org/10.1111/j.1471-8286.2004.00770.x

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). *Creating the CIPRES Science Gateway for inference of large phylogenetic trees* (pp. 1–8). Proceedings of the Gateway Computing Environments Workshop (GCE), *14 Nov. 2010*.

Milot, E., Weimerskirch, H., & Bernatchez, L. (2008). The seabird paradox: Dispersal, genetic structure and population dynamics in a highly mobile, but philopatric albatross species. *Molecular Ecology*, *17*(7), 1658–1673. https://doi.org/10.1111/j.1365-294X.2008.03700.x

Morrison, K. W. (2016). Individual repeatability in laying behaviour does not support the migratory carry-over effect hypothesis of egg-size dimorphism in *Eudyptes* penguins. *Journal of Avian Biology*, *47*(4), 466–475. https://doi.org/10.1111/jav.00740

Morrison, K. W., Battley, P. F., Sagar, P. M., & Thompson, D. R. (2015). Population dynamics of Eastern Rockhopper Penguins on Campbell Island in relation to sea surface temperature 1942–2012: Current warming hiatus pauses a long-term decline. *Polar Biology*, *38*(2), 163–177. https://doi.org/10.1007/s00300-014-1575-x

Napier, R. B. (1968). Erect-crested and rockhopper penguins interbreeding in the Falkland Islands. *British Antarctic Survey Bulletin*, *16*, 71–72.

Nelson, G. J., & Platnick, N. I. (1981). *Systematic and Biogeography: Cladistics and Vicariance*. Columbia University Press.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, *18*(3), 375–402. https://doi.org/10.1111/j.1365-294X.2008.03946.x

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, *27*(16), 3193–3206. https://doi.org/10.1111/mec.14792

Ortiz, E. M. (2019). *vcf2phylip v2.0: Convert a VCF matrix into several matrix formats for phylogenetic analysis.*

Ottenburghs, J. (2019). Multispecies hybridization in birds. *Avian Research*, *10*(1), 20. https://doi.org/10.1186/s40657-019-0159-4

Ottenburghs, J., Ydenberg, R. C., Van Hooft, P., Van Wieren, S. E., & Prins, H. H. T. (2015). The avian hybrids project: Gathering the scientific literature on avian hybridization. *Ibis*, *157*(4), 892–894. https://doi.org/10.1111/ibi.12285

Pedraza-Marrón, C. D. R., Silva, R., Deeds, J., Van Belleghem, S. M., Mastretta-Yanes, A., Domínguez-Domínguez, O., Rivero-Vega, R. A., Lutackas, L., Murie, D., Parkyn, D., Bullock, L. H., Foss, K., Ortiz-Zuazaga, H., Narváez-Barandica, J., Acero, A., Gomes, G., & Betancur-R, R. (2019). Genomics overrules mitochondrial DNA, siding with morphology on a controversial case of species delimitation. *Proceedings of the Royal Society - Biological Sciences*, *286*(1900), 20182924. https://doi.org/10.1098/rspb.2018.2924

Pertierra, L. R., Segovia, N. I., Noll, D., Martinez, P. A., Pliscoff, P., Barbosa, A., Aragón, P., Raya Rey, A., Pistorius, P., Trathan, P., Polanowski, A., Bonadonna, F., Le Bohec, C., Bi, K., Wang-Claypool, C. Y., González-Acuña, D., Dantas, G. P. M., Bowie, R. C. K., Poulin, E., & Vianna, J. A. (2020). Cryptic speciation in gentoo penguins is driven by geographic isolation and regional marine conditions: Unforeseen vulnerabilities to global change. *Diversity and Distributions*, *26*(8), 958–975. https://doi.org/10.1111/ddi.13072

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*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595–609. https://doi.org/10.1080/10635150600852011

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., & Lareu, M. (2013). An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Frontiers in Genetics*, *4*, 98. https://doi.org/10.3389/fgene.2013.00098

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

Pritchard, J., Wen, X., & Falush, D. (2010). *Documentation for structure software: Version 2.3*. Retrieved from https://web.stanford.edu/group/pritchardlab/structure_software/release_versions/v2.3.4/structure_doc.pdf

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology*, *21*(8), 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

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. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Quillfeldt, P., Moodley, Y., Weimerskirch, H., Cherel, Y., Delord, K., Phillips, R. A., Navarro, J., Calderón, L., & Masello, J. F. (2017). Does genetic structure reflect differences in non-breeding movements? A case study in small, highly mobile seabirds. *BMC Evolutionary Biology*, *17*(1), 160. https://doi.org/10.1186/s12862-017-1008-x

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Rey, A. R., Pütz, K., Luna-Jorquera, G., Lüthi, B., & Schiavini, A. (2009). Diving patterns of breeding female rockhopper penguins (*Eudyptes chrysocome*): Noir Island, Chile. *Polar Biology*, *32*(4), 561–568. https://doi.org/10.1007/s00300-008-0550-9

Rey, A. R., Trathan, P., Pütz, K., & Schiavini, A. (2007). Effect of oceanographic conditions on the winter movements of rockhopper penguins *Eudyptes chrysocome chrysocome* from Staten Island, Argentina. *Marine Ecology Progress Series*, *330*, 285–295. https://doi.org/10.3354/meps330285

Rosenberg, N. A. (2003). Distruct: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*(1), 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Sangster, G. (2014). The application of species criteria in avian taxonomy and its implications for the debate over species concepts. *Biological Reviews of the Cambridge Philosophical Society*, *89*(1), 199–214. https://doi.org/10.1111/brv.12051

Schilthuizen, M. (2000). Dualism and conflicts in understanding speciation. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *22*(12), 1134–1141.

Searby, A., & Jouventin, P. (2005). The double vocal signature of crested penguins: Is the identity coding system of rockhopper penguins *Eudyptes chrysocome* due to phylogeny or ecology? *Journal of Avian Biology*, *36*(5), 449–460.

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. *Methods in Ecology and Evolution*, *8*(8), 907–917. https://doi.org/10.1111/2041-210X.12700

Simeone, A., Hiriart-Bertrand, L., Reyes-Arriagada, R., Halpern, M., Dubach, J., Wallace, R., Pütz, K., & Luthi, B. (2009). Heterospecific pairing and hybridization between wild Humboldt and Magellanic penguins in Southern Chile. *The Condor*, *111*(3), 544–550. https://doi.org/10.1525/cond.2009.090083

Solís-Lemus, C., Knowles, L. L., & Ané, C. (2015). Bayesian species delimitation combining multiple genes and traits in a unified framework. *Evolution*, *69*(2), 492–507. https://doi.org/10.1111/evo.12582

Spinks, P. Q., Thomson, R. C., & Shaffer, H. B. (2014). The advantages of going large: Genome-wide SNPs clarify the complex population history and systematics of the threatened western pond turtle. *Molecular Ecology*, *23*(9), 2228–2241. https://doi.org/10.1111/mec.12736

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *114*(7), 1607. https://doi.org/10.1073/pnas.1607921114

Team, R Core (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing.* Retrieved from https://www.R-project.org/

Tennyson, A. J. D., & Miskelly, C. (1989). "Dark-faced" rockhopper penguins at the Snares Islands. *Notornis*, *36*(3), 183–189.

Thiebot, J. B., Cherel, Y., Trathan, P. N., & Bost, C. A. (2011). Inter-population segregation in the wintering areas of macaroni penguins. *Marine Ecology Progress Series*, *421*, 279–290. https://doi.org/10.3354/meps08907

Thiebot, J. B., Cherel, Y., Trathan, P. N., & Bost, C. A. (2012). Coexistence of oceanic predators on wintering areas explained by population-scale foraging segregation in space or time. *Ecology*, *93*(1), 122–130. https://doi.org/10.1890/11-0385.1

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*(16), 3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Trathan, P. N., García-Borboroglu, P., Boersma, D., Bost, C.-A., Crawford, R. J. M., Crossin, G. T., Cuthbert, R. J., Dann, P., Davis, L. S., De La Puente, S., Ellenberg, U., Lynch, H. J., Mattern, T., Pütz, K., Seddon, P. J., Trivelpiece, W., & Wienecke, B. (2015). Pollution, habitat loss, fishing, and climate change as critical threats to penguins. *Conservation Biology*, *29*(1), 31–41. https://doi.org/10.1111/cobi.12349

Turner, S. D. (2014). qqman: An R package for visualizing GWAS results using Q-Q and Manhattan plots. *bioRxiv*, 5165. https://doi.org/10.1101/005165

Tyler, J., Bonfitto, M. T., Clucas, G. V., Reddy, S., & Younger, J. L. (2020). Morphometric and genetic evidence for four species of gentoo penguin. *Ecology and Evolution*, *10*(24), 13836–13846.

Vianna, J. A., Fernandes, F. A. N., Frugone, M. J., Figueiró, H. V., Pertierra, L. R., Noll, D., Bi, K. E., Wang-Claypool, C. Y., Lowther, A., Parker, P., Le Bohec, C., Bonadonna, F., Wienecke, B., Pistorius, P., Steinfurth, A., Burridge, C. P., Dantas, G. P. M., Poulin, E., Simison, W. B., … Bowie, R. C. K. (2020). Genome-wide analyses reveal drivers of penguin diversification. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(36), 22303–22310. https://doi.org/10.1073/pnas.2006659117

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. *Molecular Phylogenetics and Evolution*, *107*, 486–498. https://doi.org/10.1016/j.ympev.2016.12.003

Wang, J. (2017). The computer program structure for assigning individuals to populations: Easy to use but easier to misuse. *Molecular Ecology Resources*, *17*(5), 981–990. https://doi.org/10.1111/1755-0998.12650

White, R., & Clausen, A. (2002). Rockhopper *Eudyptes chrysocome chrysocome* × macaroni *E. chrysolophus* penguin hybrids apparently breeding in the Falkland islands. *Marine Ornithology*, *30*, 40–42.

Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F(ST). *The American Naturalist*, *186*(Suppl 1), S24–S36. https://doi.org/10.1086/682949

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag. ISBN 978-3-319-24277-4.

Wickham, H., François, R., Henry, L., & Müller, K. (2018). *dplyr: A grammar of data manipulation. R package version 0.7 (R package version 0.7.6.* Retrieved from https://CRAN.R-project.org/package=dplyr [Computer software]

Wickham, H., & Henry, L. (2018). *tidyr: Easily tidy data with "spread" and "gather" functions. R package version 0.8 (R package version 0.8.2.* Retrieved from https://CRAN.R-project.org/package=tidyr [Computer software]

Williams, T. D., & Croxall, J. P. (1991). Annual variation in breeding biology of macaroni penguins, *Eudyptes chrysolophus*, at Bird island, South Georgia. *Journal of Zoology*, *223*, 189–202. https://doi.org/10.1111/j.1469-7998.1991.tb04759.x

Wilson, G. A., & Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, *163*(3), 1177–1191. https://doi.org/10.1093/genetics/163.3.1177

Wilson, J. W., Burle, M. H., Cuthbert, R., Stirnemann, R. L., & Ryan, P. G. (2010). Breeding success of northern rockhopper penguins (*Eudyptes moseleyi*) at Gough Island, South Atlantic Ocean. *The Emu*, *110*(2), 137–141. https://doi.org/10.1071/MU09095

Winkler, D. W., Billerman, S. M., & Lovette, I. J. (2015). *Bird families of the world: An invitation to the spectacular diversity of birds.* Lynx Edicions.

Woehler, E. J., & Gilbert, C. A. (1990). Hybrid rockhopper-macaroni penguins, interbreeding and mixed species pairs at Heard and Marion islands. *The Emu*, *90*, 198–201. https://doi.org/10.1071/MU9900198

Wu, C. I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, *14*(6), 851–865. https://doi.org/10.1046/j.1420-9101.2001.00335.x

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multi-locus sequence data. *Proceedings of the National Academy of Sciences*, *107*(20), 9264. https://doi.org/10.1073/pnas.0913022107

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*(24), 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

Zink, R. M., & Barrowclough, G. F. (2008). Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, *17*(9), 2107–2121. https://doi.org/10.1111/j.1365-294X.2008.03737.x

BIOSKETCH

María José Frugone is a postdoctoral researcher at Universidad Austral de Chile interested in the evolution and diversification of the Southern Ocean marine biota. This study was part of her PhD thesis.

Author contributions: M.J.F., R.C.K.B., J.A.V. and E.P. wrote the paper. M.J.F., M.E.L., G.C., P.M.M., N.A.L. and K.B. analysed the data. J.M.W. and T.L.C. discussed results extensively and helped prepare the manuscript. P.P., F.B., P.T., A.P., B.W., A.R.R., K.P., A.S and C.Y.W.C performed field or laboratory work and contributed to the manuscript.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.