DR. JANE YOUNGER (Orcid ID : 0000-0001-5974-7350)

Article type : Original Article

The challenges of detecting subtle population structure and its importance for the conservation of emperor penguins

Jane L. Younger*[1,2], Gemma V. Clucas*[1,3,4], Damian Kao[1], Alex D. Rogers[1], Karim Gharbi[5], Tom Hart†[1] and Karen J. Miller†[6].

* These authors contributed equally and share first-authorship.

† These authors contributed equally and share final authorship.

Author affiliations:

1. Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, UK.
2. Department of Biology, Loyola University Chicago, 1032 W. Sheridan Road, Chicago, IL, 60660, USA.
3. Ocean & Earth Sciences, University of Southampton Waterfront Campus, European Way, Southampton, SO14 3ZH, UK.
4. Department of Natural Resources and the Environment, University of New Hampshire, James Hall, 56 College Road, Durham, NH, 03824, USA.
5. Edinburgh Genomics, Ashworth Laboratories, University of Edinburgh, Edinburgh EH9 3JT, UK.
6. Australian Institute of Marine Science, Indian Ocean Marine Research Centre, The University of Western Australia (MO96), 35 Stirling Highway, Crawley, WA 6009, Australia.

*Corresponding authors: Jane Younger - jyounger@luc.edu and Gemma Clucas - gemma.clucas@unh.edu

Keywords: Antarctica, population genomics, RAD-seq, Southern Ocean, Ross Sea, dispersal

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14172

Abstract

Understanding the boundaries of breeding populations is of great importance for conservation efforts and estimates of extinction risk for threatened species. However, determining these boundaries can be difficult when population structure is subtle. Emperor penguins are highly reliant on sea ice, and some populations may be in jeopardy as climate change alters sea ice extent and quality. An understanding of emperor penguin population structure is therefore urgently needed. Two previous studies have differed in their conclusions, particularly whether the Ross Sea, a major stronghold for the species, is isolated or not. We assessed emperor penguin population structure using 4,596 genome-wide single nucleotide polymorphisms (SNPs), characterised in 110 individuals (10 – 16 per colony) from eight colonies around Antarctica. In contrast to a previous conclusion that emperor penguins are panmictic around the entire continent, we find that emperor penguins comprise at least four metapopulations, and that the Ross Sea is clearly a distinct metapopulation. By using larger sample sizes and a thorough assessment of the limitations of different analytical methods, we have shown that population structure within emperor penguins does exist and argue that its recognition is vital for the effective conservation of the species. We discuss the many difficulties that molecular ecologists and managers face in the detection and interpretation of subtle population structure using large SNP datasets, and argue that subtle structure should be taken into account when determining management strategies for threatened species, until accurate estimates of demographic connectivity among populations can be made.

Introduction

It is common practice to use genetics to delimit breeding populations and assess population connectivity in order to develop meaningful conservation strategies (Funk *et al.* 2012; Lowe & Allendorf 2010; Palsbøll *et al.* 2007). However, there are several key challenges to overcome when detecting and interpreting subtle population structure (Lowe & Allendorf 2010; Waples & Gaggiotti 2006). Firstly, clustering methods have been shown to perform poorly at identifying the true number of populations when gene flow is high, and performance is further impeded when markers with low mutation rates (such as biallelic SNPs with low minor allele frequencies) are used (Waples & Gaggiotti 2006). Secondly, measures of F_{ST}, one of the most commonly used methods for inferring population structure, are mathematically constrained by the allele frequencies in the dataset, and for some data types, including biallelic SNPs with rare minor alleles, the maximum value that F_{ST} can take is much less than 1 (Jakobsson *et al.* 2013). This can produce misleading results if this limitation is not understood, particularly in cases where the population structure is subtle, because small values of F_{ST} can be erroneously interpreted as evidence for panmixia. It is also crucial in such studies to recognise the distinctions between ecological and evolutionary populations (Waples & Gaggiotti 2006), and between genetic and demographic connectivity (Lowe & Allendorf 2010; Palsbøll *et al.* 2007). For populations to be demographically connected, their demographic rates, such as population growth rate, survival and birth rate, must be affected by immigration or emigration (Gilpin & Hanski 1997; Lowe & Allendorf 2010). Genetic connectivity, on the other hand, depends solely on the number of dispersal events among populations (Lowe & Allendorf 2010). While this may seem a subtle distinction, it is an important one from a conservation perspective, because even very large dispersal rates resulting in genetic homogeneity do not necessarily indicate demographic linkage of population dynamics, which will ultimately affect a population's response to future climate change. Emperor penguins (*Aptenodytes forsteri*) are an iconic, threatened species that are emblematic of these challenges, as evidenced by the vastly different

population structures that have been proposed for the species, ranging from complete demographic isolation of breeding colonies (e.g. Barbraud & Weimerskirch 2001) to species-wide panmixia (Cristofari *et al.* 2016).

Emperor penguins are likely under threat (Ainley *et al.* 2010; Jenouvrier *et al.* 2009, 2014) from changing Antarctic sea ice conditions (Collins *et al.* 2013; Vaughan *et al.* 2013). These birds form breeding colonies on sea ice at the majority of their known colony locations (Fretwell *et al.* 2012) and have been shown to be sensitive to fluctuations in sea ice extent and seasonal duration (Ainley *et al.* 2010; Fretwell *et al.* 2014; Trathan *et al.* 2011). In light of this threat, risk assessments for the species under future climate change scenarios have begun (Jenouvrier *et al.* 2009, 2014). Using a sea ice dependent demographic model paired with projected changes in local (colony-specific) sea ice conditions, Jenouvrier *et al*. (2014) predicted declines of > 50% by the year 2100 at two thirds of the colonies examined, concordant with a minimum of 19% global decline in emperor penguin numbers.

A major limitation of the Jenouvrier *et al.* (2014) forecasting study was that each colony was modelled as an isolated breeding unit, with no exchange of individuals among locations. The assumption of isolation is likely to have led to erroneous projections, because dispersal is known to decouple correlations between population trajectories and climatic variables even where local vital rates are climate dependent (Tavecchia *et al.* 2016). In addition to altering the relationship between demographic rates and local climate, dispersal can increase the resilience of a species in several key ways: 1) immigrants can promote population stability by compensating for low survival or birth rates of natives (Lowe & Allendorf 2010); 2) dispersal facilitates gene flow among breeding sites, replenishing the gene pool of a population with new, potentially adaptive alleles; and 3) dispersal enables range shifts (Walther *et al.* 2002). Range shifts may be particularly relevant for emperor penguins in the short term, as there is some

evidence that they may shift their colony sites and establish new colonies as local sea ice conditions become unfavourable (Ancel *et al.* 2014; Fretwell *et al.* 2014; LaRue *et al.* 2015). It is therefore crucial that connectivity among colony sites is incorporated into modelling studies and risk assessments for the species.

Delimiting the geographic boundaries of breeding populations is also essential for accurate monitoring of population trajectories and for implementing meaningful management plans (Funk *et al.* 2012; Palsbøll *et al.* 2007). While emperor penguins are not harvested by fisheries, many of their prey species are, particularly Antarctic krill (*Euphausia superba*) (Nicol *et al.* 2012; Trathan *et al.* 2015). Southern Ocean managing bodies, such as the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR), require an understanding of penguin distributions and patterns of connectivity in order to achieve their stated goal of ecosystem-based management.

There is a long-held assumption in the scientific literature that emperor penguins are strictly philopatric and colonies should be considered as isolated breeding units. This assumption was based on a banding study of 38 species of Antarctic seabirds, including a single colony of emperor penguins at Pointe Géologie, that concluded that "the fidelity rate to the hatching grounds is close to 100%" (Weimerskirch *et al.* 1985). This assumption has been carried forward into almost every published study of observed and projected climate change impacts on emperor penguins (Abadi *et al.* 2016; Ainley *et al.* 2010; Barbraud *et al.* 2011; Barbraud & Weimerskirch 2001; Jenouvrier *et al.* 2009, 2012, 2014). However, LaRue *et al*. (2015) gave evidence for six instances of either colony relocation or establishment of a new colony over five years after traditional breeding grounds were lost, and suggested that reductions in colony size at Pointe Géologie may be partly explained by temporary emigration. In addition, mobility of individual emperor penguins is known to be high, with juveniles shown to travel more than

7,000 km in just eight months (Thiebot *et al.* 2013), often in the vicinity of many different colonies (Kooyman *et al.* 1996; Wienecke *et al.* 2010). Given these observations, genetic exchange among emperor penguin colonies seems possible, and the assumption that colonies are demographically unlinked should be interrogated.

A study of variation in emperor penguin mitochondrial DNA (mtDNA) found no statistically significant genetic differentiation across *ca.* 8,000 km of coastline, from the Adélie Land Coast to the Weddell Sea (Younger *et al.* 2015b). This genetic homogeneity could be the result of frequent dispersal of individuals among colonies that are successfully recruited into the breeding population upon arrival. However, given the nature of the genetic marker used, it is also possible that this pattern reflects that all extant colonies are derived from a common ancestral population, and insufficient time has passed for these colonies to drift apart genetically, even if they are not currently exchanging migrants (Wright 1931). For species with large effective population sizes it can take many generations for genetic drift to alter allele frequencies, and genetic analyses will require high power to detect population differentiation (Taylor & Dizon 1996). Younger *et al*. (2015b) also found that penguins from the Ross Sea were genetically differentiated from those elsewhere on the continent, suggesting there are at least two genetically distinct populations of emperor penguins. A caveat of that study was that it was based on mitochondrial DNA and therefore represents the genetic structure of a single, uniparentally-inherited, locus.

Multiple independent loci from across the genome should allow for a more sensitive investigation of genetic differentiation among colonies. This approach was used by Cristofari *et al*. (2016) in a follow-up study of emperor penguin population structure, and the authors concluded that emperor penguins are not genetically differentiated among colonies and that the species consists of one single, panmictic population around Antarctica. Unfortunately, because

of the small sample size in the Ross Sea, Cristofari *et al*. (2016) could not adequately explore the differentiation reported in Younger *et al*. (2015b). This represents a major limitation, because the Ross Sea contains genetically distinct populations of both emperor (Younger *et al.* 2015b) and Adélie (Ritchie *et al.* 2004) penguins, acted as an ice age refugium for both Antarctic penguin species (Ritchie *et al.* 2004; Younger *et al.* 2015b), and is currently home to the world's largest breeding colonies of both emperor (Fretwell *et al.* 2012) and Adélie (Lynch & LaRue 2014) penguins. Furthermore, it is the only region with a predicted stable or increasing population of emperor penguins (Jenouvrier *et al.* 2014). Finally, Cristofari *et al*. (2016) did not include colonies in West Antarctica, the region of Antarctica experiencing the most dramatic reductions in sea ice (Stammerjohn *et al.* 2012; Vaughan *et al.* 2013). As a result, their conclusion that "as a single genetic population, emperor penguins will respond to climate change through a unified evolutionary trajectory" (Cristofari *et al.* 2016) is premature.

In this study we further investigate genetic connectivity among emperor penguin colonies around Antarctica using a robust sampling design including the Ross Sea. To achieve this, we expanded our previous mitochondrial DNA analyses (Younger *et al.* 2015b) with a dataset of 110 individuals genotyped for 4,596 high coverage genome-wide SNPs generated using restriction site associated DNA sequencing (RAD-seq) (Baird *et al.* 2008). In contrast to the other recently published SNP study, which genotyped the same total number of individuals, (Cristofari *et al.* 2016), our SNP data confirmed the existence of genetic differences between the Ross Sea and other Antarctic populations. Furthermore, we conclude that subtle yet significant genetic differences revealed among colonies are indicative of multiple metapopulations rather than a single panmictic population of emperor penguins. Recognising the likely frustration with which decision makers might view two apparently similar studies that provide contrasting conclusions about genetic structure in an iconic species, we advocate applying the

precautionary principle such that, in the absence of definitive evidence for panmixia, the breeding populations identified here should be considered as separate units for management.

Materials and methods

Study species

Emperor penguins are long-lived seabirds that exhibit delayed breeding; the average age of first breeding is five to six years and generation length is ≥ 14 years (Forcada & Trathan 2009). Pairs produce a single egg per year, with no possibility of relaying if an egg is lost, and breeding success typically ranges from 0.60 to 0.85, but may fall as low as 0.02 (Barbraud & Weimerskirch 2001; Forcada & Trathan 2009). Adult survival is typically high at 0.80 – 0.95 (Jenouvrier *et al.* 2005; Mougin & van Beveren 1979) whilst the probability of surviving the first year at sea appears to be highly variable (Jenouvrier *et al.* 2005).

Sampling, DNA extraction and RAD-seq

A total of 110 individuals (10 – 16 individuals per colony) from eight emperor penguin colonies around Antarctica were genotyped for this study. To assess genetic differentiation at both continental and regional scales, two colonies from each of three major geographic regions (the Ross Sea, Weddell Sea and Prydz Bay) were included, along with an additional two colonies in East Antarctica (Figure 1).

Muscle biopsies from the pectoral region of chick carcasses were collected at Fold Island and Pointe Géologie in 2010, and from Amanda Bay in 2012 and 2013 (see Figure 1 for colony locations). Skin biopsies from the foot or back were collected from dead adults and chicks at Halley Bay in 2012. Whole chick carcasses were collected at Auster in 1993 and 1994. Blood

samples were collected from Gould Bay in 2013 and from the Ross Sea colonies in 1992 and 1993. This sampling interval represents less than two generations for emperor penguins (Forcada & Trathan 2009) and so does not present a problem for estimating contemporary population structure, because any gene flow during this interval would not have had time to alter allele frequencies throughout the colony. For detailed methods of blood sampling see Younger *et al.* (2015b). Sampling was conducted under permits from the UK Foreign and Commonwealth Office, the US National Science Foundation and the Australian Antarctic Division. Ethical approval of each of these permits was granted by the permitting institution and additional ethical approvals were received from the University of Oxford and British Antarctic Survey.

Genomic DNA (gDNA) was extracted using QIAGEN DNEasy blood and tissue kits with the following modifications to the digestion step: 30 µL proteinase K was added to blood samples and the incubation time was extended to 3 hrs; 40 µL proteinase K was added to tissue samples (plus an additional 10 µL 1 M dithiothreitol for skin samples from the foot) and the incubation step was extended to 32 hrs. RNA contamination was reduced by treating samples with 1 µL RNase A (QIAGEN) or 1 µL Riboshredder (Epicentre). DNA concentration was measured with a Qubit (ThermoFisher Scientific). The presence of high molecular weight DNA was confirmed on a 1% gel and DNA contamination was measured using a Nanodrop (ThermoFisher Scientific).

RAD-seq, including library preparation, was performed by Edinburgh Genomics, University of Edinburgh (https://genomics.ed.ac.uk/). The method used follows Gonen *et al.* (2014) after Etter *et al.* (2011). For each individual, 250 ng of gDNA was digested with the SbfI-HF (NEB) restriction enzyme, followed by ligation to barcoded P1 adapters. The individually barcoded samples were then pooled into multiplexed libraries, and sheared into fragments of 300 – 400 bp. Gel electrophoresis was used to size-select fragments. The NEB Quick Blunting Kit was used

to blunt the libraries, followed by A-tailing and ligation to P2 adapters (IDT). Yields were increased via enrichment PCR, followed by purification with Ampure beads. The enriched libraries were checked for size and concentration using Qubit and a qPCR assay. The libraries were sequenced using 125 base paired-end reads on an Illumina HiSeq 2500 in high output mode (v4 chemistry). The individuals were sequenced across eight libraries along with samples of other species (for other studies).

Bioinformatics, SNP calling and filtering

Overall read quality and the presence of adapters were assessed using FastQC. Reads were demultiplexed, trimmed and cleaned using process_radtags from the Stacks software pipeline v1.35 (Catchen *et al.* 2011, 2013). The reads were truncated to 113 bp to exclude four terminal bases as sequence quality can decrease at the end of the read. Read pairs where either of the pair had a low quality score, uncalled bases and/or a barcode or cut-site with more than one mismatch were excluded from further analysis. The remaining paired reads were then aligned to the emperor penguin reference genome (http://gigadb.org/dataset/100005) using bwa-mem (Li 2013). Terminal alignments were prevented using a clipping penalty of 100. Reads with multiple alignments, more than 5 mismatches and/or more than 2 indels were removed using a custom python script (filter.py available from the Dryad Digital Repository http://dx.doi.org/10.5061/dryad.7c0q8). PCR duplicates were removed using Picardtools (http://broadinstitute.github.io/picard).

Our SNP calling methodology followed the general guidelines of Benestan *et al*. (2016) and our methodology for king penguins (Clucas *et al.* 2016). We used the Stacks pipeline (*pstacks* – *cstacks* – *sstacks* – *rxstacks* – *cstacks* – *sstacks* – *populations*) with the following options: in *pstacks* we specified a minimum depth of six reads mapping to the same location (-m 6) and

employed the bounded SNP model with a significance level of α = 0.05, an upper bound of 0.1 and a lower bound of 0.0041 (the highest sequencing error rate recorded by phiX spikes in the sequencing lanes, --model_type bounded --bound_low 0.0041 --bound_high 0.1 --alpha 0.05); all individuals were used to build the catalog in *cstacks*; in *rxstacks* confounded loci (loci with too many haplotypes to be biologically possible such as from repetitive regions or paralogous loci) were removed with a conservative confidence limit of 0.25 (--conf_lim 0.25), excess haplotypes were removed from individuals (-–prune_haplo), and loci with a mean log likelihood < -10 were removed from further analysis (–ln_lim -10). In the *populations* module, we removed any SNPs with a minor allele frequency (MAF) < 0.01, which are likely to be the result of sequencing errors (--min_maf 0.01); loci with a heterozygosity > 0.5, which could be the result of paralogous sections of the genome being merged into a single locus (--max_obs_het 0.5); a single SNP per RAD-tag was chosen at random to remove tightly linked SNPs from the dataset (--write_random_snp); a locus must have been present in all populations to be included in the final dataset (-p 8); and a locus must have been genotyped in at least 80% of individuals per population to be included (-r 0.8). We verified that no SNPs had a mean coverage greater than 100X using vcftools v0.1.13 (Danecek *et al.* 2011), to avoid SNPs from repetitive regions of the genome, and confirmed that none were out of Hardy Weinberg equilibrium (HWE) in more than 50% of the populations when $p < 0.01$ using the *adegenet* (Jombart & Ahmed 2011) package in R and vcftools. We used PGDSpider v2.0.8.2 (Lischer & Excoffier 2012) to prepare files for population genomic analyses.

Loci under directional or balancing selection violate the assumption of neutrality that is used in most population genetic analyses. We used BayeScan v2.1 (Foll & Gaggiotti 2008), which employs a Bayesian F_{ST} outlier test, to identify loci that might be under selection and hence discard them. It has been demonstrated (Lotterhos & Whitlock 2014) that BayeScan is powerful at detecting loci genuinely under selection in a range of demographic scenarios, but with an

accompanying high false-positive rate. As our aim was to achieve a neutral set of loci, a high false positive rate was not a concern. We conservatively set the prior odds of neutrality parameter to five (for every five loci one is expected to be under selection) to ensure all loci under selection were detected. We deemed q-values of < 0.1 to be a significant result, so we can expect one in ten loci to be a false-positive neutral locus that is wrongly discarded (Lotterhos & Whitlock 2014; Storey & Tibshirani 2003).

Clustering of individuals and populations

We visualized population structure with a Discriminant Analysis of Principal Components (DAPC) (Jombart *et al.* 2010) using the *adegenet* (Jombart & Ahmed 2011) package in R. This method can benefit from groups being defined by successive *K*-means clustering (*find.clusters* function), to avoid *a priori* assignment of individuals to groups based on sampling locations. However, the optimal number of clusters suggested by *K*-means was one, and so we ran DAPC on the individuals when they were grouped by colony and by geographic region (Ross Sea, Mawson Coast, Weddell Sea and Amanda Bay/Pointe Géologie). In each case, the optimal number of principal components to retain was found by running *optim.a.score* 20 times and taking the average. As well as clustering individuals, we clustered populations using the Analysis of Molecular Variance (AMOVA) based method of *K*-means clustering in Genodive (Meirmans 2012) using simulated annealing. We compared results with the Bayesian Inference Criterion (BIC) and pseudo-F statistic (Caliński & Harabasz 1974). The BIC is useful for detecting whether there is any structure ($K > 1$ rather than $K = 1$) because the pseudo-F statistic cannot be defined for $K = 1$, but the pseudo-F statistic is known to perform better than the BIC at determining the true value of *K* (Meirmans 2012). Finally, individuals were clustered using the Bayesian clustering algorithm in the program Structure v2.3.4 (Pritchard *et al.* 2000) which estimates the proportion of ancestry each individual shares with ancestral populations, effectively showing their assignment probability to each cluster. From the results of our

population differentiation analysis, which showed that differentiation was low, the admixture model with correlated allele frequencies was deemed the most appropriate model. We ran the program initially with K = 1, for 100,000 generations, discarding the first 50,000 as burnin, allowing lambda to vary (inferlamda = 1). From this initial run lambda was estimated to be 0.40, so we set lambda to 0.40 for all subsequent runs. We then varied K from one to eight, running each value of K ten times with a random seed, for 150,000 generations, discarding the first 50,000 generations of each run as burnin. We did this both with and without sampling locations supplied as prior information. Supplying sampling locations can be useful for detecting structure when it is subtle. Structure Harvester web v0.6.94 (Earl 2012) was used to compare runs with different values of K using the Evanno method (Evanno *et al.* 2005) and to prepare files for CLUMPP (Jakobsson & Rosenberg 2007). We used CLUMPP to check for multimodality between the replicate runs at each value of K and to average across the ten replicates. The individual assignments to ancestral populations were plotted using distruct v1.1 (Rosenberg 2004).

Population differentiation

We calculated the Weir and Cockerham unbiased estimator of F_{ST} (Weir & Cockerham 1984) between all pairs of colonies, and between the geographic regions (Ross Sea, Mawson Coast, Weddell Sea and Amanda Bay/Pointe Géologie), using Genodive v2.0b27 (Meirmans & Van Tienderen 2004). Significance was estimated using 5,000 permutations of the data, and the significance level was adjusted for multiple testing using Sequential Goodness of Fit (SGoF+) (Carvajal-Rodriguez & de Uña-Alvarez 2011). Confidence intervals for the F_{ST} estimates were also estimated by bootstrapping over loci with 1,000 bootstrap replicates, using the R package diveRsity v1.9.89 (Keenan *et al*. 2013).

Phylogeography

To estimate the evolutionary relationships among the colonies and the order in which they were founded, we used a species tree approach in the SNAPP add-on (Bryant *et al.* 2012) to BEAST v.2.4.0 (Bouckaert *et al.* 2014). SNAPP uses a coalescent approach with unlinked biallelic markers (like SNPs) to infer species trees. It is highly computationally demanding, hence we selected two random individuals (i.e. four haplotypes) per colony to include in the analysis, and we repeated the analysis twice, with different individuals, to ensure the results were replicable. Within the reduced dataset any loci that were no longer polymorphic were removed, leaving datasets of 3,221 and 3,237 SNPs. We calculated the mutation rates (u and v) from the dataset rather than estimating them as part of the MCMC. The MCMC was run for 5 million generations with the first 10% of generations discarded as burnin. We checked for convergence using Tracer v1.6 (Rambaut & Drummond 2007) and found ESSs > 4,000, which is more than sufficient. We also inspected the posterior distributions of parameters for convergence. The advantage of using a Bayesian method as implemented in BEAST is that the uncertainty in the parameter of interest, here the tree, is estimated via the posterior distribution. We used DensiTree v2.0.1 to visualize the entire posterior distribution of trees as a cloudogram.

Results

Genotyping

The average number of reads per library was 14 million (range 11 – 17 million), 97% of which were retained after filtering for reads of low quality, adapter contamination, ambiguous barcodes and ambiguous RAD cut-sites. There were no significant differences in the number of reads generated per library, although the number of reads per colony did differ (Supplementary Figure 1). The differences in read depth and the library to which individuals were assigned did not contribute to any of the population structure we observed in our data (described below).

After reads were aligned to the emperor penguin reference genome, Stacks (Catchen *et al.* 2013) identified 423,479 SNPs, of which we retained 4,596 after following standard filtering steps (Benestan *et al.* 2016) (Table 1). The coverage of these SNPs ranged from 11 – 56X, with a mean of 33X. The number of individuals successfully genotyped at each colony varied between 10 and 16 (Table 2). Population structure analyses were performed using all individuals, and to confirm that unequal sample numbers among colonies were not influencing the results we repeated a subset of analyses (pairwise F_{ST}, DAPC and Structure) using a random subset of 10 individuals per colony (results provided in Supplementary Information). Genetic diversity indices (expected and observed heterozygosity, and nucleotide diversity) were similar across all colonies (Table 2).

Clustering of populations and individuals

AMOVA-based *K*-means clustering of populations (Meirmans 2012) and Bayesian clustering of individuals performed with Structure (Pritchard *et al.* 2000) both separated colonies into genetic populations. However, both methods yielded several possibilities for the number of clusters, suggesting that differentiation among emperor penguin colonies is subtle and/or hierarchical. The two different *K*-means summary statistics for assessing the number of clusters yielded different results: $K = 3$ for the pseudo-F statistic and $K = 6$ for the BIC. The Structure analysis did not converge on a single result, with different replicate runs arriving at distinct solutions. This was a result of genuine multimodality between runs, as opposed to label switching (Jakobsson & Rosenberg 2007) (Supplementary Figure 2), which can occur when population differentiation is subtle and the MCMC therefore becomes trapped in local optima. For Structure, the highest posterior mean log-likelihood was achieved at $K = 1$ (Supplementary Figure 3) both with and without location priors. The alternative method for assessing the optimal number of clusters in Structure, the Evanno method (Evanno *et al.* 2005), selected $K = 3$ with location priors and $K = 6$ without priors, consistent with the *K*-means pseudo-F statistic

and BIC results, respectively. It should be noted that the Evanno method for analysing Structure results does not allow $K = 1$ to be tested. For the three population scenario, suggested by both the *K*-means pseudo-F statistic and the Evanno method with location priors in Structure, the following populations were defined: 1) a Ross Sea population (Cape Roget and Cape Washington colonies), 2) a Mawson Coast population (Fold Island and Auster colonies), and 3) all other colonies (Pointe Géologie, Amanda Bay, Gould Bay and Halley Bay). Under the three-cluster scenario 32.2% of the total variance was explained ($r^2 = 0.322$) in the *K*-means analysis. Upon visual inspection of the Structure plots for $K = 2$ to $K = 6$ (Figure 2), it is apparent that $K = 4$ is the finest level of genetic structure that can be discerned. At $K = 5$ and $K = 6$ there is no further resolution of population structure (Figure 2). In every scenario from $K = 2$ to $K = 6$ the Ross Sea population is clearly genetically distinct from all other colonies (Figure 2). These patterns were entirely consistent whether the full dataset was used, or whether only 10 individuals per colony were included in the analysis (Supplementary Figure 4).

Individuals (as opposed to populations) could not be separated into genetically distinct groups by a successive *K*-means procedure (find.clusters from the *adegenet* package in R (Jombart & Ahmed 2011)) as the minimum value for the BIC was found at $K = 1$. This could be because the first few principal components of the principal components analysis (PCA) used to transform the data before *K*-means clustering were unable to explain much of the variation (Figure 3, inset). However, Discriminant Analysis of Principal Components (DAPC) (Jombart *et al.* 2010) distinguished the Ross Sea colonies when individuals were grouped by their colony of origin and when individuals were grouped by geographic region (Figure 3). The power to detect this differentiation was diminished but still visible when the analysis was repeated with 10 individuals per colony (Supplementary Figure 5a), and lost when individuals were shuffled between groups (Supplementary Figure 5b). The same pattern was observed with unsupervised clustering through PCA (Supplementary Figure 6).

Overall, our clustering results suggest the existence of at least four genetic populations of emperor penguins, each consisting of at least two sampled breeding colonies. The Ross Sea is the most differentiated population overall, followed by the Mawson Coast population, then the Weddell Sea and Amanda Bay/Pointe Géologie populations, which show the least amount of differentiation from one another. The selection of $K = 1$ as the most likely number of clusters by some of the methods indicates that differentiation among these four genetic populations is subtle.

Population differentiation

Pairwise genetic differentiation, as measured by Weir & Cockerham's F_{ST} estimator (Weir & Cockerham 1984), was statistically significant for 17 out of 28 pairs of colonies (Figure 4 and see Supplementary Table 1 for F_{ST} confidence intervals). The values of significant F_{ST} were small (range 0.002 – 0.006), consistent with the clustering results that indicated subtle differentiation. It should be noted that small values of F_{ST} are expected for datasets of biallelic SNPs with low minor allele frequencies, and in such cases small F_{ST} should not be interpreted as a lack of differentiation (Jakobsson *et al.* 2013). In this study we selected SNPs with minor allele frequencies ≥ 0.01, averaging 0.078 across all SNPs, which is the likely cause of the small, yet statistically significant, F_{ST} values observed. When 10 individuals per colony were randomly sub-sampled, the patterns of pairwise differentiation were largely unchanged (Supplementary Table 2).

Of note was that within geographic regions (Weddell Sea, Mawson Coast and Ross Sea) there was no genetic differentiation (Figure 4 and see Supplementary Table 3 for F_{ST} confidence intervals). Despite the *ca*. 3,200 km between them, there was no differentiation between the two East Antarctic colonies of Amanda Bay and Point Géologie (F_{ST} = -0.001, p = 0.813, Figure 4),

consistent with the clustering results. At large spatial scales there was significant genetic differentiation corresponding with the major geographic regions e.g. among colonies in the Ross Sea, colonies in the Weddell Sea, colonies from the Mawson Coast and colonies from East Antarctica (Figure 4). Every comparison in a pairwise F_{ST} analysis that grouped colonies into these four geographic regions was statistically significant, with F_{ST} values ranging from 0.002 to 0.005 (Figure 5). The Ross Sea was the most differentiated from all other populations. Overall, our pairwise F_{ST} analyses further support the conclusion that there are four genetic populations of emperor penguins, corresponding with major geographical regions.

Phylogeography

The posterior distribution of trees resulting from our SNAPP (Bryant *et al.* 2012) species tree approach formed a diffuse cloud (Supplementary Figure 7) indicating that there was no single, well-supported topology. The lack of a well-supported topology prevents us from determining the evolutionary relationships among the colonies and further reflects the lack of strong genetic differentiation among colonies.

Discussion

Our findings of low, yet statistically significant, genetic differentiation among emperor penguin colonies spanning more than 10,000 km of the Antarctic coastline refutes claims of panmixia in this species and provides evidence that emperor penguins in Antarctica are composed of at least four distinct metapopulations. These metapopulations are variously connected through gene flow, whilst colonies within each appear to be panmictic because of high rates of dispersal.

Some clustering analyses struggled to detect more than a single genetic cluster of emperor penguins, and if these results were not scrutinised further it might be easy to conclude that there is one globally panmictic population. However, scrutiny of data is crucial, because

clustering methods are known to perform poorly at estimating the number of genetic populations when gene flow is high (Waples & Gaggiotti 2006), which we reasonably expect may be the case for emperor penguins (Kooyman & Ponganis 2016; LaRue *et al.* 2015; Younger *et al.* 2015b). By examining multiple values of *K* we were able to ascertain a clear pattern of genetic differentiation among emperor penguin colonies using Structure, and this was further supported by our pairwise F_{ST} comparisons and DAPC. The four genetic populations of emperor penguins within our sampling range are the Ross Sea, Mawson Coast, Weddell Sea and East Antarctic Coast (Figure 6c). Over short distances, i.e. less than 600 km, colonies appear to be genetically connected. The two colonies in the Weddell Sea – Gould Bay and Halley Bay, located 550 km apart – are not genetically differentiated. Likewise, the two colonies in the Ross Sea, located 300 km apart, are genetically indistinguishable, as are the Fold Island and Auster colonies of the Mawson Coast, which are separated by only 190 km, as well as Amanda Bay and Auster, located 600 km apart. Where colonies are separated by more than 600 km the patterns of genetic differentiation are less predictable; for example, Amanda Bay penguins are not genetically differentiated from those *ca*. 3,200 km away at Pointe Géologie, whereas they are significantly differentiated from the Fold Island colony just 790 km away. We therefore urge that assumptions should not be made about the connectivity of emperor penguin colonies based on geographic proximity alone.

Our finding of four metapopulations is supported by the limited evidence available from satellite tracking of fledgling and moulting penguins (Kooyman *et al.* 1996, 2000; Kooyman & Ponganis 2008, 2016; Thiebot *et al.* 2013; Wienecke *et al.* 2004). During these life stages emperor penguins are not tied to breeding colonies and hence have the opportunity to disperse. Tracking of fledglings from Pointe Géologie (140°E) showed several individuals travelling west in the direction of Amanda Bay (76°E), with one individual reaching as far as 88°E before the satellite tag failed (Thiebot *et al.* 2013). Notably, none of the Pointe Géologie fledglings travelled

east toward the Ross Sea. Thiebot *et al*. (2013) attributed the prolonged westward trajectory to the influence of the Antarctic Coastal Current. Our finding that these sites are genetically homogenous suggests a potential role for oceanography in shaping patterns of connectivity among emperor penguin colonies. Fledglings from western Ross Sea colonies have been shown to travel in a trajectory consistent with the circulation of the Ross Gyre (Kooyman *et al.* 1996; Kooyman & Ponganis 2008). This circulation pattern could encourage Ross Sea penguins to remain in the region. In further support of the Ross Sea's status as a distinct metapopulation, adults from Cape Washington and Cape Roget, along with a further three colonies in the western Ross Sea, have been tracked to a common moult area at the end of the breeding season (Kooyman *et al.* 2000). Individuals from these five colonies almost certainly encountered each other during the moult period, and none of the tracked individuals left the Ross Sea (Kooyman *et al.* 2000). It has also been shown that there is large annual variability in the number of penguins breeding at individual colony sites in the Ross Sea, which may indicate dispersal of breeding individuals (Kooyman & Ponganis 2016). Similarly, the Mawson Coast metapopulation may also share moult habitat – moulting individuals from Auster and a nearby colony at Taylor Glacier were tracked to a moulting location in close proximity to the Fold Island colony (Wienecke *et al.* 2004). While moulting Fold Island individuals have not been tracked, it is very likely that they utilise the same habitat (Wienecke *et al.* 2004). The influence of oceanography and non-breeding distribution on population structure is a promising avenue for future research.

In our first study of emperor penguin genetic connectivity, based on mtDNA sequences, we found that penguins in the Ross Sea were distinct from those in East Antarctica and the Weddell Sea (Younger *et al.* 2015b), a finding that is supported here by genome-wide SNPs. However, the SNPs revealed further population subdivision within the East Antarctic and Weddell Sea regions, with three distinct genetic populations resolved among the six colonies studied, which

were grouped as a single genetic population using mitochondrial DNA alone (Figure 6a,c). This indicates that SNPs are superior to mtDNA for detecting fine-scale genetic differentiation among colonies, as expected (Baird *et al.* 2008). Analyses of genetic connectivity within other penguin species, including Adélies (Clucas *et al.* 2014; Ritchie *et al.* 2004; Roeder *et al*. 2001; Younger *et al.* 2015a) and chinstraps (Clucas *et al.* 2014; Freer *et al.* 2015) using either mtDNA or microsatellites, resolved similar levels of genetic differentiation as found in the mtDNA analysis of emperor penguins (Younger *et al.* 2015b). It is therefore possible that other penguin species have higher levels of genetic differentiation than reported to date, which could be detected using genome-wide SNPs.

A recent study of emperor penguin genetic connectivity also used SNPs and concluded that emperor penguins are "a fully panmictic species" (Cristofari *et al.* 2016) (Figure 6b). Genetic or demographic connectivity can be described as a spectrum ranging from panmixia, in which mating is random (i.e. equally likely) amongst all pairs of individuals across the colonies studied, to complete isolation, in which there is no genetic exchange among colonies (Waples & Gaggiotti 2006). Whilst Cristofari *et al*. (2016) proposed panmixia, Jenouvrier *et al*. (2014) used complete isolation to forecast population trends for the species under future climate change scenarios. Our genetic data, both here and in Younger *et al*. (2015b), support an intermediate scenario, as would be expected for most species in the natural world (Waples & Gaggiotti 2006). Specifically, the emperor penguin species is comprised of multiple metapopulations, with some degree of connectivity among metapopulations, and very high connectivity between sub-populations within each.

The difference between our results and that of Cristofari *et al*. (2016) warrants interrogation. How might two similar datasets be so differently interpreted? We attest that one of the main arguments for panmixia of emperor penguins in Cristofari *et al*. (2016) was the "very low"

values of F_{ST} between pairs of colonies, ranging from 0.0080 between the proximate East and West Mertz colonies, to 0.0240 between Cape Washington and Pointe Géologie (Cristofari *et al.* 2016). However, as discussed above, biallelic SNP datasets with rare minor alleles are expected to yield small F_{ST} values (Jakobsson *et al.* 2013), and these are not evidence of panmixia. Cristofari *et al*. (2016) did not report the significance levels of their pairwise F_{ST} values, nor the *p*-value associated with their AMOVA test, so it is unknown whether the reported values represent differentiation or homogeneity. As a comparison, we observed F_{ST} values as low as 0.002 (between Pointe Géologie and Halley Bay, Amanda Bay and Halley Bay, and Amanda Bay and Fold Island) that represented statistically significant genetic differentiation at the $\alpha = 0.05$ level after SGoF+ correction for multiple tests. To place these findings in a broader context, a recent study (Toews *et al.* 2016) compared two species of parulid warblers, *Vermivora chrysoptera* and *Vermivora cyanoptera*, and found an F_{ST} value of 0.0045 based on 11.4 million SNPs. This F_{ST} value, a measure of differentiation between two distinct bird species (Toews *et al.* 2016), is smaller than the smallest F_{ST} value reported by Cristofari *et al*. (2016) for emperor penguins.

Neither our SNAPP species tree analysis nor the neighbour-net analysis of Cristofari *et al*. (2016) recovered evidence of distinct evolutionary lineages within emperor penguins. However, a lack of genealogically distinct populations should not be misinterpreted as panmixia. Stochastic lineage sorting is driven primarily by genetic drift, and is therefore directly proportional to the effective population sizes in question (Avise *et al.* 1983), which may be very large for emperor penguins (Fretwell *et al.* 2012). Interestingly, our mtDNA study revealed three distinct, well-supported lineages within emperor penguins (Younger *et al.* 2015). This apparent discrepancy is not surprising, because the rate of lineage sorting is expected to be four times faster for mitochondrial than nuclear DNA (Pamilo & Nei 1988).

Our original mtDNA study found a high degree of genetic differentiation between the Ross Sea colonies and those elsewhere in East Antarctica (Figure 6a) (Younger *et al.* 2015b). Based on their SNP dataset, Cristofari *et al*. (2016) reported that Ross Sea emperor penguins were panmictic with all other emperor penguins across Antarctica (Figure 6b). However, their study included only four individuals from the Ross Sea, a sample size that precludes any reasonable analysis of population structure (Cristofari *et al.* 2016). We have now conclusively validated the genetic differentiation of the Ross Sea based on robustly called genome-wide SNPs and recommend that the Ross Sea emperor penguins be conservatively considered as a distinct management unit (Funk *et al.* 2012).

Ideally, we could translate our knowledge of the genetic patterns among colonies into an estimate of demographic connectivity, defined as the extent to which a population's vital rates and growth are influenced by dispersal (Lowe & Allendorf 2010). This would enable meaningful species risk assessments (Jenouvrier *et al.* 2014; Tavecchia *et al.* 2016), population monitoring, and effective conservation strategies (Funk *et al.* 2012; Palsbøll *et al.* 2007). To interpret the observed genetic patterns, we must consider two alternative scenarios. Firstly, the lack of genetic differentiation between some colonies could be the result of contemporary dispersal of individuals among breeding sites, suggesting that demographic linkage is a possibility. However, it is also possible that the observed genetic similarities could be the result of shared ancestry among now isolated colonies, which have simply not been isolated for a sufficient time to diverge genetically. Given the growing body of evidence that emperor penguins travel widely as juveniles (Kooyman *et al.* 1996; Thiebot *et al.* 2013; Wienecke *et al.* 2010), exhibit dynamism with respect to their colony locations (Ancel *et al.* 2014; LaRue *et al.* 2015), may temporarily emigrate (Kooyman & Ponganis 2016; LaRue *et al.* 2015), and that multiple colonies share moulting locations (Kooyman *et al*. 2000; Wienecke *et al*. 2004), we consider the former

scenario, that dispersal is maintaining gene flow among populations, to be the most likely explanation for the genetic similarity found here.

An accurate estimate of the rate of dispersal among colonies would substantially improve our understanding of emperor penguin demographic connectivity. Unfortunately, the low differentiation observed among colonies precludes the estimation of dispersal rates with any degree of accuracy. BayesAss, which is typically used to estimate rates of dispersal over the past few generations (Wilson & Rannala 2003), is unreliable where F_{ST} values are less than 0.05 (Faubet *et al.* 2007) – a full order of magnitude greater than observed among emperor penguin colonies. Coalescent methods, such as Migrate-n (Beerli 2009), are used to infer rates of dispersal over evolutionary timescales, making them inappropriate for estimating contemporary connectivity. It is also unclear what effect different SNP calling pipelines may have on coalescent estimates. Low frequency SNPs are believed to be important for estimating dispersal rates (P. Beerli *pers. comm.*), but are hard to distinguish from sequencing errors, meaning SNP data may not be suitable for this type of analysis. Furthermore, the coalescent method is contingent on accurate estimation of evolutionary rates for the loci in question and it is not currently possible to accurately estimate evolutionary rates for RAD loci (Harvey & Brumfield 2015; Harvey *et al.* 2016). To compound these issues further, unsampled populations that exchange migrants with the studied populations have serious confounding effects on the estimation of dispersal rates, specifically, populations appear to be exchanging migrants when they are not (Beerli 2004; Slatkin 2005). This is almost certainly the case for emperor penguins, for which there are many unsampled colonies that lie between the colonies included in all the genetic studies to date (Figure 6).

We note that Cristofari *et al*. (2016) reported dispersal rates among emperor penguin colonies based on coalescent methods, however, these estimates are likely to be affected by many of the issues described above and, at any rate, only describe dispersal over evolutionary timescales,

rather than contemporary connectivity of colonies. Even if it were possible to generate an accurate estimate of the number of emperor penguin migrants, as yet there is no generalized framework for determining the level of dispersal necessary to maintain demographic linkage (Waples & Gaggiotti 2006) and even high rates of dispersal do not guarantee demographic interdependence (Lowe & Allendorf 2010). Therefore, it is currently impossible to determine the extent to which emperor penguin colonies are demographically linked. Future studies could address this gap by combining genetic methods with data on movement behaviour, perhaps from capture-mark-recapture methods, local demographic rates and measures of the reproductive success of both immigrants and residents (Lowe & Allendorf 2010; Tavecchia *et al.* 2016).

With these caveats acknowledged, we tentatively conclude that emperor penguin colonies are most likely not demographically isolated, nor are they likely to be linked continent-wide. Within our dataset we have identified four genetic populations of emperor penguins (Figure 6c) among eight study colonies, and future management plans, monitoring schemes, and population forecasts should take this structure into account. The subtle population structure we describe here was overlooked by Cristofari *et al*. (2016), who instead concluded that emperor penguins comprise "a single global population with a shared demography". From a management point of view these are very different conclusions. If emperor penguins are misdefined as a single demographic unit spanning the whole of Antarctica, there is a risk that localised climate change or fisheries impacts may be downplayed in management plans for the species, resulting in local extinctions (Palsbøll *et al.* 2007; Taylor *et al.* 2000). Furthermore, future species risk assessments will be confounded if it is incorrectly assumed that emperor penguins at all colonies around Antarctica are demographically linked.

The inherent challenges of detecting and interpreting subtle population structure are not unique to emperor penguins, but difficulties that many molecular ecologists must grapple with. There are several key reasons why subtle population structure may be overlooked. Firstly, sampling regimes with small sample sizes or that omit large portions of a species' range will have a limited ability to detect structure in the affected regions. Secondly, structure may not be detected if the subtleties and limitations of the analytical methods used are not respected. For example, the limited ability of clustering methods (Evanno *et al.* 2005; Pritchard *et al.* 2000) to detect the true number of genetic populations (K) in the presence of high gene flow and/or when using low mutation rate markers is well documented (Waples & Gaggiotti 2006), yet many studies report and visualise only a single value of K. In another example, the mathematical constraints on F_{ST} as a result of allele frequencies (Jakobsson et al. 2013) can result in very small F_{ST} values that may be incorrectly interpreted as panmixia if the constraints are misunderstood and statistical tests not conducted. If any of these methodological oversights result in panmixia being incorrectly inferred, this may be further misinterpreted by equating genetic connectivity to demographic linkage (Gilpin & Hanski 1997; Lowe & Allendorf 2010; Palsbøll *et al.* 2007). Such conclusions could lead to the definition of inappropriately large management units, confounding conservation efforts and, in the worse case scenario, result in localised or widespread extinctions (Palsbøll *et al.* 2007; Taylor *et al.* 2000).

Here we have assessed the genetic population structure among emperor penguin colonies and synthesised the available information on the patterns of connectivity among colonies. The difference between complete panmixia and distinct metapopulations with gene flow may seem like a subtle distinction, but it is a distinction that has important implications for our understanding of emperor penguin population dynamics, behaviour, ecology, adaptability, range-shift potential, population trajectories and, most importantly, risk assessments and conservation planning. As a research community it is crucial that we clearly delimit the extent of

our knowledge, and that the subtleties and limitations of population genetic studies are fully communicated in our scientific reporting, so that managers and conservation planners can make properly informed decisions. It is important to note that there are many emperor penguin colonies that have not been included in genetic studies to date, and including these in future work will further our understanding of the true links among emperor penguin populations. In this study, we present evidence for distinct metapopulations within emperor penguins, which are connected by some level of dispersal. Until demographic connectivity can be determined, a precautionary approach would warrant that these metapopulations are considered as separate management units.

Acknowledgements

This research was funded by the Australian Antarctic Science Program (Project 4184 – KM & JY), a Holsworth Wildlife Research Endowment (JY), the Sea World Research and Rescue Foundation (JY & KM), the Natural Environment Research Council (GC), an Endeavour Research Fellowship (JY), The Darwin Initiative (DPLUS 002 - TH), an NSF Creative Science Award (DPP-87-15864), the Charities Advisory Trust (TH) and by public donations, in particular through Quark Expeditions Ltd (TH). We gratefully acknowledge the following people and organisations for their kind donations of emperor penguin samples: the Kooyman team at the Ross Sea; Barbara Wienecke and the Australian Antarctic Division at Fold Island, Amanda Bay and Auster; Francoise Amelineau at Pointe Géologie; Phil Trathan and assistants from the British Antarctic Survey at Halley Bay; and Dick Filby and Hannah McKeand from Adventure Network International at Gould Bay. We would like to thank the Edinburgh Genomics staff for library preparation and sequencing. Edinburgh Genomics is partly supported through core grants from NERC (R8/H10/56), MRC (MR/K001744/1) and BBSRC (BB/J004243/1). We acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this

work (http://dx.doi.org/10.5281/zenodo.22558). We would like to thank Barbara Wienecke and Jerry Kooyman for helpful discussions about emperor penguin population dynamics. Finally, we are very grateful to Michelle LaRue, Laura Benestan, Bruce Deagle, and two anonymous reviewers, all of whom gave considered, insightful feedback on both our analytical methods and interpretation of the ecological patterns.

References

Abadi F, Barbraud C, Gimenez O (2017) Integrated population modeling reveals the impact of climate on the survival of juvenile emperor penguins. *Global Change Biology*, **23**, 1353–1359.

Ainley D, Russel J, Jenouvrier S *et al.* (2010) Antarctic penguin response to habitat change as Earth's troposphere reaches 2 °C above preindustrial levels. *Ecological Monographs,* **80**, 49-66.

Ancel A, Cristofari R, Fretwell PT *et al.* (2014) Emperors in hiding: when ice-breakers and satellites complement each other in Antarctic exploration. *PLoS ONE,* **9**, e100404.

Avise JC, Shapira J, Daniel SW, Aquadro CF, Lansman RA (1983) Mitochondrial DNA differentiation during the speciation process in Peromyscus. *Molecular Biology and Evolution,* **1**, 38-56.

Baird NA, Etter PD, Atwood TS *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE,* **3**, e3376.

Barbraud C, Gavrilo M, Mizin Y, Weimerskirch H (2011) Comparison of emperor penguin declines between Pointe Géologie and Haswell Island over the past 50 years. *Antarctic Science,* **23**, 461-468.

Barbraud C, Weimerskirch H (2001) Emperor penguins and climate change. *Nature,* **411**, 183-186.

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology,* **13**, 827-836.

Beerli P (2009) How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use. *Population Genetics for Animal Conservation,* **17**, 42-79.

Benestan LM, Ferchaud AL, Hohenlohe PA *et al.* (2016) Conservation genomics of natural and managed populations: building a conceptual and practical framework. *Molecular Ecology,* **25**, 2967–2977.

Bouckaert R, Heled J, Kühnert D *et al.* (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology,* **10**, e1003537.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution,* **29**, 1917-1932.

Caliński T, Harabasz J (1974) A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods,* **3**, 1-27.

Carvajal-Rodriguez A, de Uña-Alvarez J (2011) Assessing significance in high-throughput experiments by sequential goodness of fit and q-value estimation. *PLoS One,* **6**, e24700.

Catchen J, Amores A, Hohenlohe PA, Cresko WA, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics,* **1**, 171-182.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology,* **22**, 3124-3140.

Clucas GV, Dunn MJ, Dyke G *et al.* (2014) A reversal of fortunes: climate change 'winners' and 'losers' in Antarctic Peninsula penguins. *Scientific Reports,* **4**, 5024.

Clucas GV, Younger JL, Kao D *et al.* (2016) Dispersal in the sub-Antarctic: king penguins show remarkably little population genetic differentiation across their range. *BMC Evolutionary Biology,* **16**, 211.

Collins M, Knutti R, Arblaster J *et al.* (2013) Long-term Climate Change: Projections, Commitments and Irreversibility. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds. Stocker TF, Qin D, Plattner G-K*, et al.*), pp. 1029–1136. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cristofari R, Bertorelle G, Ancel A *et al.* (2016) Full circumpolar migration ensures evolutionary unity in the emperor penguin. *Nature Communications,* **7**, 11842.

D'Amen M, Zimmermann NE, Pearman PB (2013) Conservation of phylogeographic lineages under climate change. *Global Ecology and Biogeography,* **22**, 93-104.

Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics,* **27**, 2156-2158.

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources,* **4**, 359-361.

Etter PD, Bassham S, Hohenlohe PA, Johnson EA, Cresko WA (2011) SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In: *Molecular Methods for Evolutionary Genetics* (eds. Orgogozo V, Rockman MV), pp. 157-178. Humana Press.

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

Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Molecular Ecology,* **16**, 1149-1166.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics,* **180**, 977-993.

Forcada J, Trathan PN (2009) Penguin responses to climate change in the Southern Ocean. *Global Change Biology*, **15,** 1618-1630.

Freer J, Mable B, Clucas G *et al.* (2015) Limited genetic differentiation among chinstrap penguin (*Pygoscelis antarctica*) colonies in the Scotia Arc and Western Antarctic Peninsula. *Polar Biology,* **38**, 1493–1502.

Fretwell PT, LaRue MA, Morin P *et al.* (2012) An emperor penguin population estimate: the first, global, synoptic survey of a species from space. *PLoS ONE,* **7**, e33751.

Fretwell PT, Trathan PN, Wienecke B, Kooyman GL (2014) Emperor penguins breeding on iceshelves. *PLoS ONE,* **9**, e85285.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution,* **27**, 489-496.

Gilpin MEH, Hanski IA (1997) *Metapopulation Biology: Ecology, Genetics, and Evolution.* Academic Press.

Gonen S, Lowe NR, Cezard T *et al.* (2014) Linkage maps of the Atlantic salmon (*Salmo salar*) genome derived from RAD sequencing. *BMC Genomics,* **15**, 166.

Harvey MG, Brumfield RT (2015) Genomic variation in a widespread Neotropical bird (*Xenops minutus*) reveals divergence, population expansion, and gene flow. *Molecular Phylogenetics and Evolution,* **83**, 305-316.

Harvey MG, Smith BT, Glenn TC, Faircloth BC, Brumfield RT (2016) Sequence capture versus restriction site associated DNA sequencing for shallow systematics. *Systematic Biology,* **65**, 910-924.

Jakobsson M, Edge MD, Rosenberg NA (2013) The relationship between F_{ST} and the frequency of the most frequent allele. *Genetics,* **193**, 515-528.

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

Jenouvrier S, Caswell H, Barbraud C *et al.* (2009) Demographic models and IPCC climate projections predict the decline of an emperor penguin population. *Proceedings of the National Academy of Sciences,* **106**, 1844-1847.

Jenouvrier S, Holland M, Stroeve J *et al.* (2012) Effects of climate change on an emperor penguin population: analysis of coupled demographic and climate models. *Global Change Biology,* **18**, 2756-2770.

Jenouvrier S, Holland M, Stroeve J *et al.* (2014) Projected continent-wide declines of the emperor penguin under climate change. *Nature Climate Change,* **4**, 715-718.

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics,* **27**, 3070-3071.

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

Keenan K, McGinnity P, Cross TF, Crozier WW, & Prodöhl PA (2013) diveRsity: An R package for the estimation of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, **4**, 782-788.

Kooyman GL, Hunke EC, Ackley SF, van Dam RP, Robertson G (2000) Moult of the emperor penguin: travel, location, and habitat selection. *Marine Ecology Progress Series,* **204**, 269-277.

Kooyman GL, Kooyman TG, Horning M, Kooyman CA (1996) Penguin dispersal after fledging. *Nature,* **383**, 397.

Kooyman GL, Ponganis PJ (2008) The initial journey of juvenile emperor penguins. *Aquatic Conservation: Marine and Freshwater Ecosystems,* **17**, S37-S43.

Kooyman GL, Ponganis PJ (2016) Rise and fall of Ross Sea emperor penguin colony populations: 2000 to 2012. *Antarctic Science*, in press.

LaRue M, Kooyman G, Lynch H, Fretwell P (2015) Emigration in emperor penguins: implications for interpretation of long-term studies. *Ecography,* **38**, 114-120.

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Lischer HEL, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics,* **28**, 298-299.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology,* **23**, 2178-2192.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology,* **19**, 3038-3051.

Lynch H, LaRue M (2014) First global census of the Adélie penguin. *The Auk,* **131**, 457-466.

Meirmans PG (2012) AMOVA-based clustering of population genetic data. *Journal of Heredity,* **103**, 744-750.

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes,* **4**, 792-794.

Mougin JL, van Beveren M (1979) Structure et dynamique de la population de manchots empereurs *Aptenodytes forsteri* de la colonie de l'archipel de Pointe Géologie, Terre Adélie. *Comptes Rendus de l'Académie des Sciences,* **289**, 157-160**.**

Nicol S, Foster J, Kawaguchi S (2012) The fishery for Antarctic krill – recent developments. *Fish and Fisheries,* **13**, 30-40.

Palsbøll PJ, Berube M, Allendorf FW (2007) Identification of management units using population genetic data. *Trends in Ecology and Evolution,* **22**, 11-16.

Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution,* **5**, 568-583.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics,* **155**, 945-959.

Rambaut A, Drummond AJ (2007) Tracer v1. 4.

Ritchie PA, Millar CD, Gibb GC, Baroni C, Lambert DM (2004) Ancient DNA enables timing of the Pleistocene origin and Holocene expansion of two Adélie penguin lineages in Antarctica. *Molecular Biology and Evolution,* **21**, 240-248.

Roeder AD, Marshall RK, Mitchelson AJ *et al.* (2001) Gene flow on the ice: genetic differentiation among Adélie penguin colonies around Antarctica. *Molecular Ecology,* **10**, 1645-1656.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes,* **4**, 137-138.

Slatkin M (2005) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology,* **14**, 67-73.

Stammerjohn S, Massom R, Rind D, Martinson D (2012) Regions of rapid sea ice change: an inter-hemispheric seasonal comparison. *Geophysical Research Letters,* **39**, L06501.

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences,* **100**, 9440-9445.

Tavecchia G, Tenan S, Pradel R *et al.* (2016) Climate-driven vital rates do not always mean climate-driven population. *Global Change Biology,* **22**, 3960–3966.

Taylor BL, Dizon AE (1996) The need to estimate power to link genetics and demography for conservation. *Conservation Biology,* **10**, 661-664.

Taylor BL, Wade PR, De Master DP, Barlow J (2000) Incorporating uncertainty into management models for marine mammals. *Conservation Biology,* **14**, 1243-1252.

Thiebot J-B, Lescroël A, Barbraud C, Bost C-A (2013) Three-dimensional use of marine habitats by juvenile emperor penguins Aptenodytes forsteri during post-natal dispersal. *Antarctic Science,* **25**, 536-544.

Toews DP, Taylor SA, Vallender R *et al.* (2016) Plumage genes and little else distinguish the genomes of hybridizing warblers. *Current Biology,* **26**, 2313-2318.

Trathan PN, Fretwell PT, Stonehouse B (2011) First recorded loss of an emperor penguin colony in the recent period of Antarctic regional warming: implications for other colonies. *PLoS ONE,* **6**, e14738.

Trathan PN, García-Borboroglu P, Boersma D *et al*. (2015) Pollution, habitat loss, fishing, and climate change as critical threats to penguins. *Conservation Biology*, **29**, 31-41.

Vaughan DG, Comiso JC, Allison I *et al.* (2013) Observations: Cryosphere. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds. Stocker TF, Qin D, Plattner G-K*, et al.*), pp. 317–382. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Walther G-R, Post E, Convey P *et al.* (2002) Ecological responses to recent climate change. *Nature,* **416**, 389-395.

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

Weimerskirch H, Jouventin P, Mougin J, Stahl J, Van BM (1985) Banding recoveries and the dispersal of seabirds breeding in French Austral and Antarctic Territories. *Emu,* **85**, 22-33.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Wienecke B, Kirkwood R, Robertson G (2004) Pre-moult foraging trips and moult locations of emperor penguins at the Mawson Coast. *Polar Biology,* **27**, 83-91.

Wienecke B, Raymond B, Robertson G (2010) Maiden journey of fledgling emperor penguins from the Mawson Coast, East Antarctica. *Marine Ecology Progress Series,* **410**, 269-282.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics,* **163**, 1177-1191.

Wright S (1931) Evolution in Mendelian populations. *Genetics,* **16**, 97-159.

Younger J, Emmerson L, Southwell C, Lelliott P, Miller K (2015a) Proliferation of East Antarctic Adélie penguins in response to historical deglaciation. *BMC Evolutionary Biology,* **15**, 1-11.

Younger JL, Clucas GV, Kooyman G *et al.* (2015b) Too much of a good thing: sea ice extent may have forced emperor penguins into refugia during the last glacial maximum. *Global Change Biology,* **21**, 2215-2226.

Data Accessibility

The Illumina short reads are available from the NCBI sequence read archive, https://www.ncbi.nlm.nih.gov/bioproject/384210. Our custom python script (filter.py) is available from the Dryad Digital Repository http://dx.doi.org/10.5061/dryad.7c0q8 and our final SNP dataset is available from the Dryad Digital Repository http://dx.doi.org/10.5061/dryad.4s7t3.

Author Contributions

JY & GC carried out molecular laboratory work, analyzed the data, interpreted the data, drafted the manuscript and participated in conceiving and designing the study. DK participated in analyzing the data. AR & KG participated in conceiving and designing the study. TH collected samples, participated in interpreting the data and conceiving and designing the study. KM

participated in conceiving and designing the study and in interpretation of the data. All authors read and approved the final manuscript.

The authors declare no competing financial interests

Figures

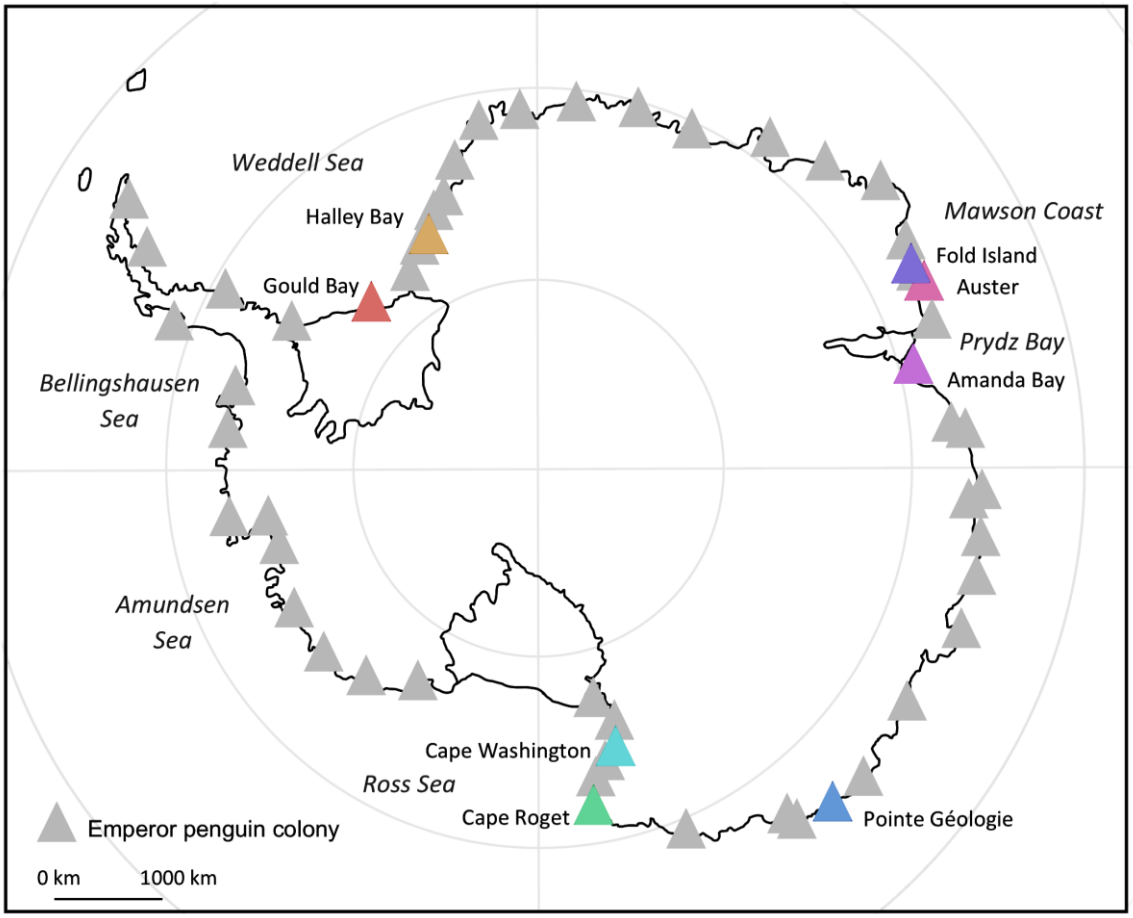

Figure 1. Map showing the locations of all known emperor penguin colonies around Antarctica (LaRue *et al.* 2015). The colonies sampled in this study are labelled.

Figure 2. Structure results showing population structure from $K = 2$ to $K = 6$. The results are averaged for the ten replicate runs that were performed at each value of K. Each individual is a vertical bar, with the colors showing the proportion of ancestry assigned to each of the clusters.

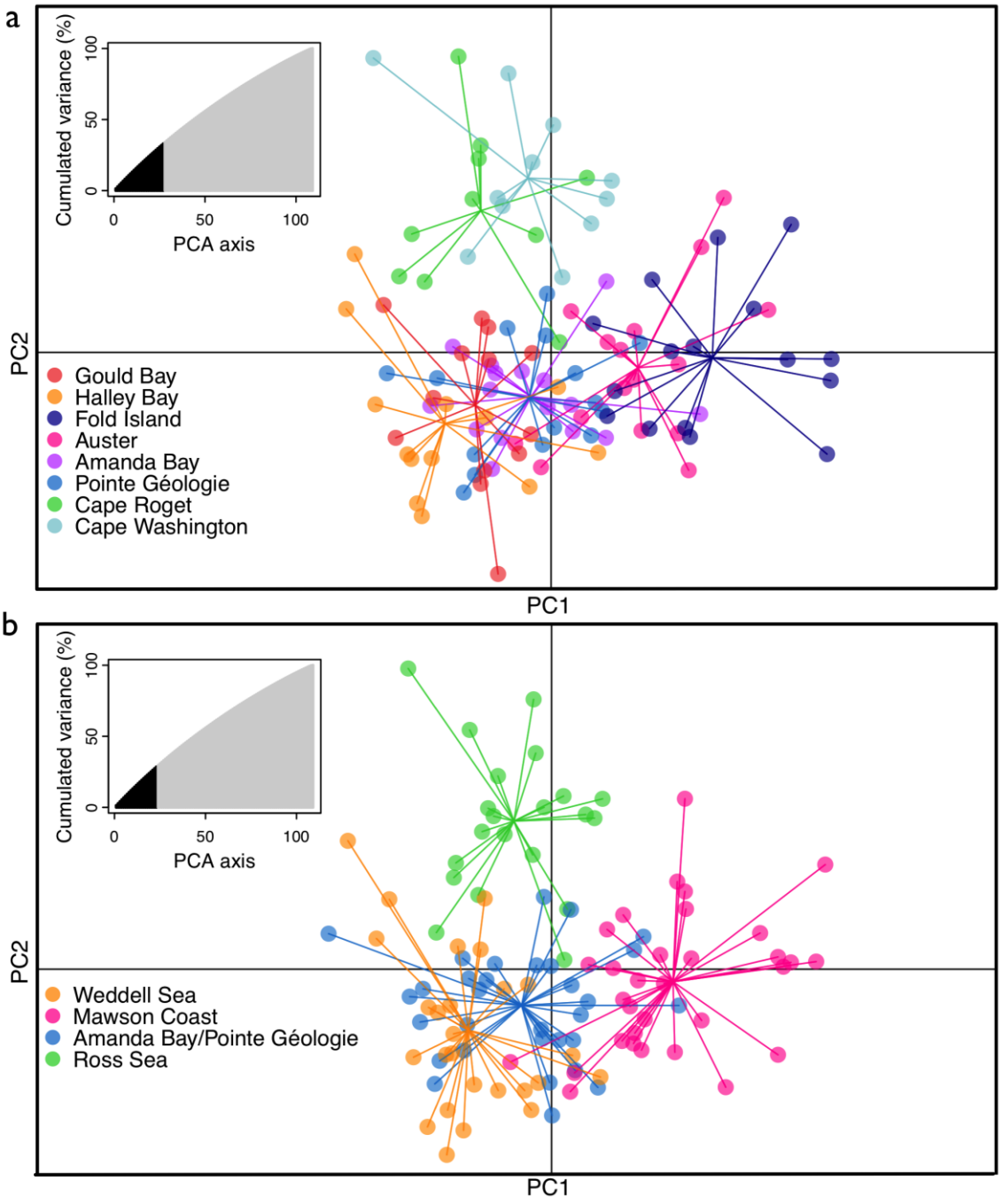

Figure 3. Results from Discriminant Analysis of Principal Components when individuals are grouped by a) colony and b) major geographic region. The numbers of retained principal components are shown in black on the inset graphs and were a) 27 and b) 23.

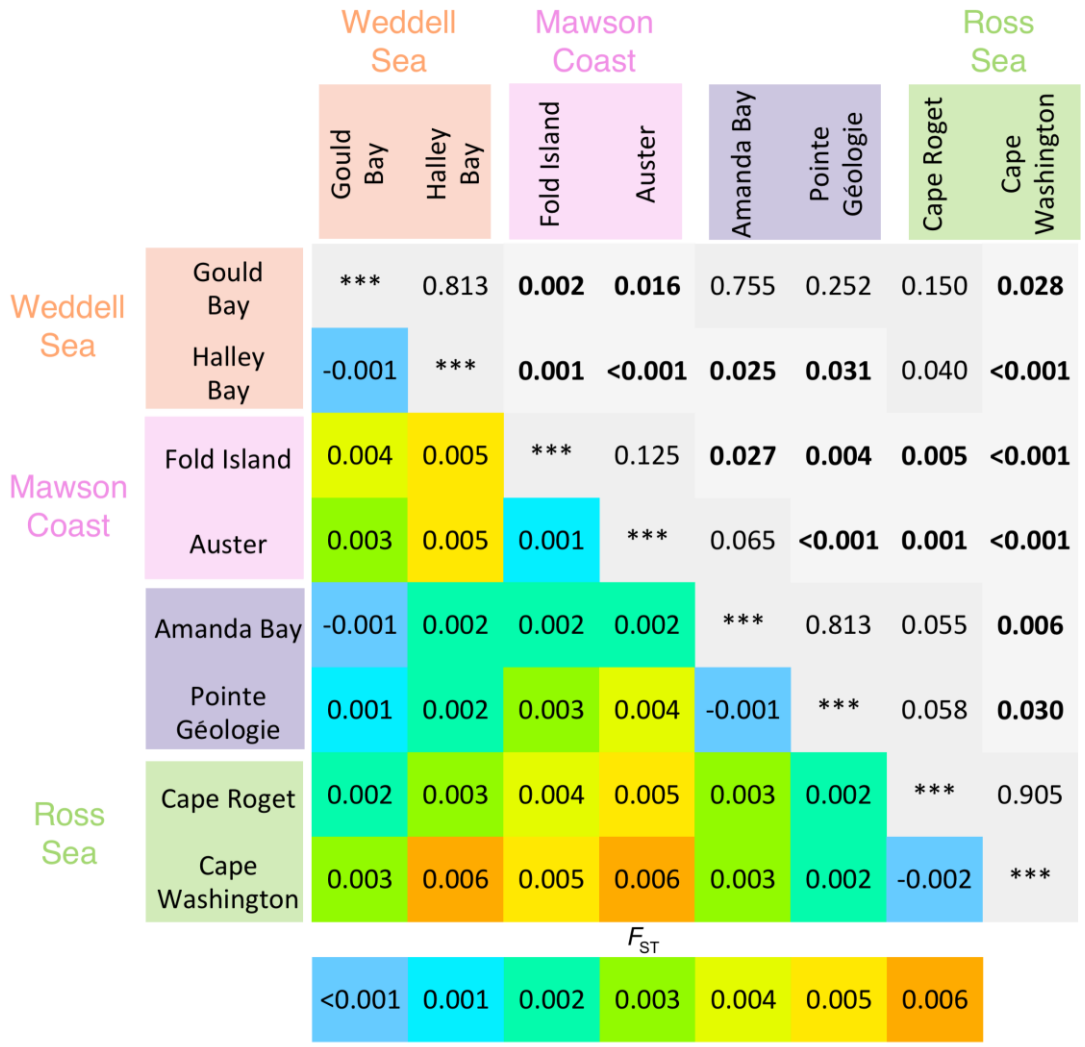

Figure 4. Genetic differentiation between all pairs of colonies. F_{ST} values are shown below the diagonal, with associated *p*-values above the diagonal. Significant *p*-values after SGoF+ correction for multiple tests are shown in bold.

	Weddell Sea	Mawson Coast	Amanda Bay & Pointe Géologie	Ross Sea
Weddell Sea	***	**<0.001**	**0.004**	**<0.001**
Mawson Coast	0.004	***	**<0.001**	**<0.001**
Amanda Bay & Pointe Géologie	0.002	0.003	***	**<0.001**
Ross Sea	0.005	0.005	0.004	***

Figure 5. Genetic differentiation between all major geographic regions. F_{ST} values are shown below the diagonal, with associated *p*-values above the diagonal.

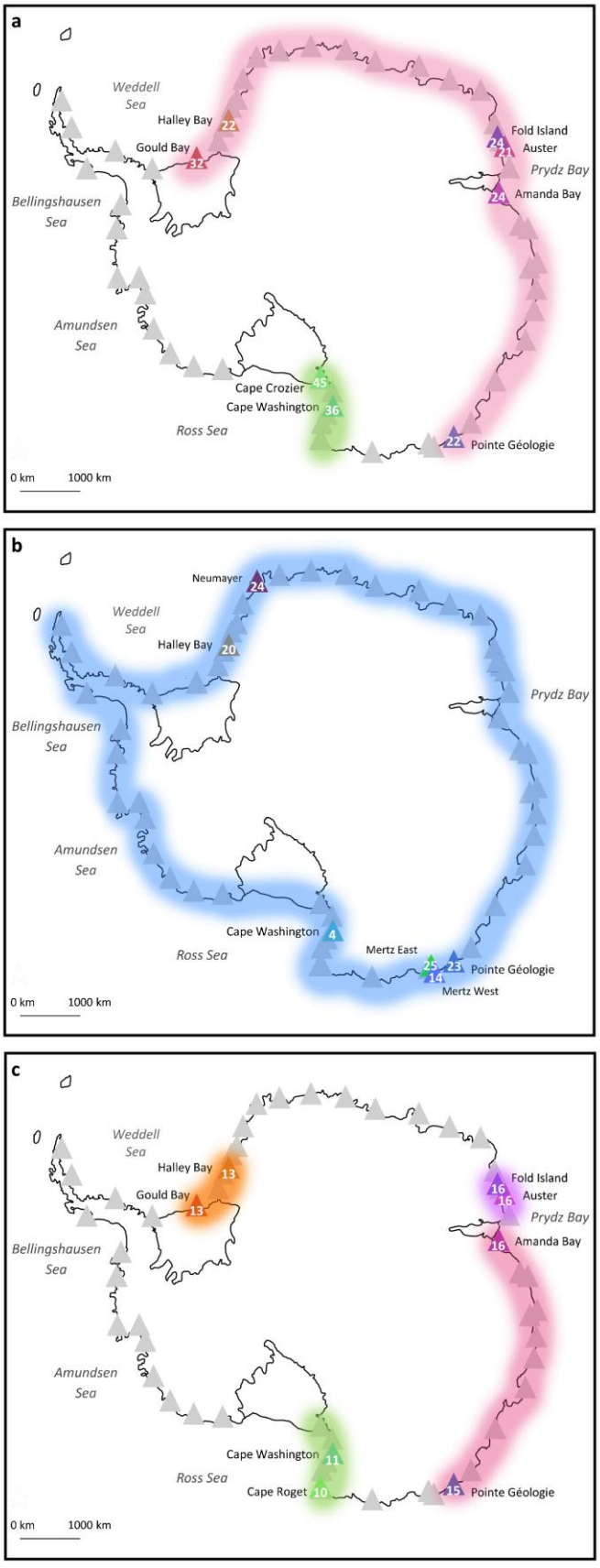

Figure 6. Our evolving understanding of emperor penguin population structure, with populations indicated with coloured shading: a) our original mitochondrial DNA study found two populations, one in the Ross Sea (green) and one spanning the Weddell Sea and East Antarctica (pink) (Younger *et al*. 2015b); b) a study based on SNPs concluded that emperor penguins are a fully panmictic species consisting of a single population spanning the Antarctic continent (blue) (Cristofari *et al*. 2016); and c) this study based on SNPs found that there are at least four metapopulations of emperor penguins, in the Ross Sea (green), Weddell Sea (orange), the Mawson Coast (purple), and Amanda Bay/Pointe Géologie (pink). Grey triangles indicate all known emperor penguin colonies (LaRue *et al*. 2015). The coloured triangles indicate colonies included in the study design, with the number of individuals sampled noted inside the triangle.

Table 1. The number of SNPs retained after applying each of the filters.

Filter	SNPs retained
Initial stacks catalog	423,479
After running rxstacks	64,320
Minor allele frequency > 0.01	30,441
Heterozygosity < 0.5	30,125
Genotyped in > 80% individuals per population	9,857
Single SNP per RAD-tag	4,600
Putatively neutral	4,596

Table 2. Genetic diversity indices for each colony, based on the dataset of 4,596 SNPs (variant sites only) retained after filtering. N – number of individuals successfully sequenced, H_E – expected heterozygosity; H_O – observed heterozygosity; π – nucleotide diversity; var – variance; stdErr – standard error of the mean.

	N	Private Alleles	H_E (mean)	H_E (var)	H_E (stdErr)	H_O (mean)	H_O (var)	H_O (stdErr)	π (mean)	π (var)	π (stdErr)
Gould Bay	13	7	0.1218	0.0178	0.002	0.1133	0.0176	0.002	0.1261	0.0191	0.002
Halley Bay	13	6	0.1223	0.0182	0.002	0.1187	0.0192	0.002	0.1265	0.0194	0.0021
Fold Island	16	11	0.1231	0.0177	0.002	0.1112	0.0164	0.0019	0.1271	0.0189	0.002
Auster	16	5	0.1223	0.0173	0.0019	0.1189	0.0177	0.002	0.1264	0.0185	0.002
Amanda Bay	16	9	0.1243	0.0173	0.0019	0.1171	0.0171	0.0019	0.1283	0.0184	0.002
Pointe Géologie	15	6	0.1227	0.018	0.002	0.1131	0.0174	0.0019	0.1269	0.0192	0.002
Cape Roget	10	1	0.1151	0.0196	0.0021	0.1071	0.0201	0.0021	0.122	0.022	0.0022
Cape Washington	11	1	0.1174	0.0186	0.002	0.1134	0.0199	0.0021	0.1232	0.0205	0.0021