

Assigned Associate Editor: Willows-Munro, Sandi

MISS EMILY LOUISA CAVILL (Orcid ID : 0000-0003-2198-8102)

DR. LICIA CALABRESE (Orcid ID : 0000-0003-0953-1281)

Article type : Original Article

Running Head: Conservation genomics of an endangered bird

Conservation genomics of the endangered Seychelles Magpie-Robin (*Copsychus sechellarum*): A unique insight into the history of a precious endemic bird

Emily L. Cavill,^{1,2*} Shyam Gopalakrishnan,¹ Lara C. Puetz,¹ Ângela M. Ribeiro,³ Sarah S.T. Mak,¹ Rute R. da Fonseca,⁴ George Pacheco,¹ Bronwyn Dunlop,⁶ Wilna Accouche,⁷ Nirmal Shah,⁸ Anna Zora,⁹ Licia Calabrese,^{10,15} Martin Genner,² Gareth Jones,² Chunxue Luo,¹² Guojie Zhang^{11, 12, 13, 14} & M. Thomas P. Gilbert^{1,5}

1. Center for Evolutionary Hologenomics, The GLOBE Institute, University of Copenhagen, Denmark
2. School of Biological Sciences, University of Bristol, UK
3. Natural History Museum of Denmark, University of Copenhagen, Denmark.
4. Center for Macroecology, Evolution and Climate, The GLOBE Institute, University of Copenhagen, Denmark
5. University Museum, Norwegian University of Science and Technology, Norway
6. Cousine Island Company, Seychelles
7. Green Islands Foundation, Seychelles
8. Nature Seychelles, Seychelles
9. Frégate Island Private, Seychelles
10. Island Conservation Society, Seychelles

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/IBI.13023](https://doi.org/10.1111/IBI.13023)

This article is protected by copyright. All rights reserved

- Accepted Article
11. *Villum Center for Biodiversity Genomics, University of Copenhagen, Denmark*
 12. *China National Genebank, BGI-Shenzhen, China*
 13. *State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, China*
 14. *Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, China*
 15. *Island Biodiversity & Conservation Center, University of Seychelles, Seychelles.*

*Corresponding author.

Email: emily.cavill@live.co.uk

The Seychelles Magpie-Robin *Copsychus sechellarum* is an IUCN Red-List Endangered species endemic to the Seychelles, whose population was reduced to eight individuals on a single island in the 1960s. Translocations from the remaining population to four additional islands have been an integral factor in their recovery, but the potential genetic consequences of their translocation history have not previously been explored. We resequenced the genomes of 141 individuals sampled across the five current island populations, and analysed the data to characterise their population structure, as well as to explore suspected inbreeding. Overall very low levels of heterozygosity were observed, all coupled with long homozygous segments that suggest recent inbreeding, likely the consequence of a population bottleneck in the 1960s. Three of the four translocated populations displayed less genetic diversity than the founder population from which they were taken, a familiar pattern observed as a result of the evolutionary force of genetic drift following founder events. Furthermore, and perhaps surprising given the recent time since the new populations were established, population structure was observed within these same three populations. New awareness of inbreeding in the Seychelles Magpie-Robin populations, and continued genetic monitoring, will allow for genetically informed management decisions. This is particularly prudent in maximising the success of the future conservation translocation planned for this species.

Keywords: endangered species, conservation genetics, avian conservation, passerine, inbreeding, whole genome resequencing

Driven to near-extinction in the mid-1900s, the Seychelles Magpie-Robin *Copsychus sechellarum* is one of many species endemic to the Seychelles to have been devastatingly affected by human actions over past centuries (Government of Seychelles 2014). The decline of this now IUCN ‘Endangered’ listed bird species was due to agricultural intensification, the introduction of exotic predators, namely rats and cats (Gaymer *et al.* 1969, Wilson & Wilson 1978, Watson *et al.* 1992) and exacerbated by continued specimen collection as the population dwindled (in Oustalet 1878, as highlighted in Burt *et al.* 2016). The population, which by 1965 had been reduced to only eight sighted individuals on a single island, Frégate, had failed to recover naturally by 1990, resulting in management intervention from Birdlife International and The Royal Society for the Protection of Birds (RSPB), through the implementation of a recovery plan. As a result of this successful intervention, IUCN downlisted the species conservation status in 2005 from ‘Critically Endangered’ to ‘Endangered’, thus rendering the Seychelles Magpie-Robin one of only 32 bird species to be reclassified positively (downlisted) between 1988-2008 (Birdlife International 2008). However, this reclassification was performed without insight into the genetic consequences of the species recovery plan, and thus the recovery plan did not consider the genetic vulnerability of the species.

An integral component of the successful Seychelles Magpie-Robin recovery plan was facilitation of population expansion through translocations, which ultimately elevated the population size to nearly 400 individuals today. However, the success of the independent translocations was extremely variable. Six translocation attempts were carried out from Frégate to Aride Island (1978: six birds; 1979: four birds; 1992: two birds; 1994: one bird; 1995: six birds; 1999: one bird); and one fertilised egg was translocated from Cousin to Aride Island in 2000, all of which failed to establish successful populations (Burt *et al.* 2016).

Other translocations were more successful, principally (though not exclusively) involving birds from Frégate Island (Fig. 1). Specifically, between 1994 and 1996, nine birds were translocated from Frégate to Cousin Island; between 1995-1996, seven birds were translocated to Cousine Island, six from Frégate complemented with a seventh from Cousin Island; in 2002, the two surviving magpies on Aride were moved to Frégate Island, and 15 new birds were moved from Frégate to Aride Island; and in 2008 a total of 20 birds were translocated to Denis Island: 16 birds from Frégate and four from Cousin.

In addition to sourcing birds from different islands, different strategies were also employed when choosing birds to be moved. Specifically, the donor birds for the Cousin and Cousine translocation were targeted as unrelated (as knowingly possible) adult breeding pairs. In contrast, for Aride some social groups were moved together (Bristol *et al.* 2005). Later, and based on observations derived from monitoring the results of the previous translocations, the donor birds for Denis Island were targeted to account for age (to include high number of sub-adults and other young adults), reproductive history (including reproductive, hatching and fledging success), alongside the exclusion (where possible) of known relatives (The Seychelles Magpie-Robin Recovery Team (S.M.A.R.T. *pers. comms.*).

Today the species inhabits five islands in total. Of these five islands, Frégate Island (Hartlaub 1877, Pike 1872) and Aride Island (Hartlaub 1877, Oustalet 1878) are the only islands with documented history of Magpie-Robin populations previous to the recent translocations. There is believed to be very limited gene flow between these five islands, as despite intense monitoring only 16 cases of island to island dispersal have been recorded, of which only seven resulted in establishment (Burt *et al.* 2016, S.M.A.R.T. *pers. comms.*). If an individual survives to adulthood, the average lifespan is 4.2 years (Nature Seychelles *unpubl. data*) but Magpie-Robins have been documented to live for up to 15 years and have an estimated generation time of 3.6 years (Birdlife International 2017). A breeding pair may occupy a territory for substantially longer than this, living with up to eight non-breeding subordinates within the territory. Related subordinate Magpie-Robins have been observed to act as ‘helpers’ in the group – feeding the chicks. However, the presence of subordinates is largely recognised as causing conflict within the territory, therefore given the small islands and limited territories and resources, this can ultimately slow the process of recovery (López-Sepulcre *et al.* 2009).

For two decades following the first translocation carried out by Birdlife International and RSPB, the future of this species looked promising with continual population growth on all established islands (Fig. S1). However, in 2014, the population of Aride suffered a sudden and catastrophic crash, where the population was reduced to just six birds from a population of 30, with no apparent environmental cause.

This crash ignited recognition of the need for genetic research on this species. We sequenced the genomes of 141 individuals representing the extant populations of the five islands. Our study had

three main aims: 1) to explore how the genetic architecture of the four translocated populations differs from the founder population of Frégate Island; 2) to describe whether these translocations may have had potentially problematic consequences on the populations' genomic diversity as a result of bottlenecks or inbreeding; and 3) if so, to identify whether there may have been underlying genetic components potentially responsible for the 2014 Aride Island population crash.

METHODS

Sampling

Blood samples are routinely collected from Magpie-Robins by trained local conservation staff on all five islands, as part of the long-term monitoring program. Blood sampling is carried out by the withdrawal of blood (maximum 70 µL) through brachial venipuncture, the extracted blood is stored in one of lysis buffer, Queen's buffer or ethanol, and samples are subsequently frozen at -20 °C. A total of 141 blood samples were used for this research: 118 were provided by S.M.A.R.T. and 23 by the Netherlands Institute of Ecology (NIOO). Due to the small current population sizes of the species, this sampling covers a representative sample of the populations at the time of this research, spanning each of the five islands inhabited by the Magpie-Robins today: Frégate (31/145), Denis (32/78), Cousin (31/50), Cousine (17/50), and Aride (30 total: 7/10 of the surviving population and 24 samples from NIOO that were archived blood samples from birds recently deceased and representative of the Aride population collapse). Samples were exported with the express purposes of undertaking this research project with permission from the Ministry of Environment, Energy and Climate Change of the Seychelles (MTA-2017), and were imported to Denmark for research under permission from the Ministry of Environment and Food of Denmark (J.nr. 2017-12-711-06242).

Laboratory protocols

Two different genomic DNA extraction kits were used based on compatibility with sample storage buffers. Samples stored in lysis and Queen's buffers were extracted using the Thermo Scientific Kingfisher Blood DNA kit (ThermoFisher Scientific Inc., MA, USA) with the adjustment of using 200 µL of sample input. The Qiagen DNeasy Blood and Tissue kit (Qiagen, CA, USA) for nucleated blood was used for samples stored in ethanol. Extractions were quantified using the

Qubit Fluorometer 2.0 Broad Range DNA assay kit (Invitrogen, CA, USA) and DNA quality was assessed using the genomic DNA kit for the Agilent TapeStation 2200 (Agilent Technologies, CA, USA).

DNA was sheared following the standard protocol of the Covaris M220 Focused-Ultrasonicator with the adaptation of using a treatment time of 175 seconds to produce 250 base pair fragments. One hundred and forty-four BGISEq compatible genomic shotgun libraries were built following the ‘BEST 2.0’ protocol (Carøe *et al.* 2017, as modified by Mak *et al.* 2017) with the adjustment of using double volume reactions as DNA input was high (as high as 500ng/ μ L for some samples). The libraries were single-indexed and PCR-amplified using the recommended cycles ascertained by qPCR to avoid excessive duplicates (8–18 PCR cycles) and final PCR-amplified libraries were quantified using the Qubit Fluorometer 2.0 and the High Sensitivity DNA assay kit. All genomic shotgun libraries were pooled equimolarly in batches of eight and submitted to BGI-Shenzhen for sequencing using the BGISEq500 platform (BGI 2019) and paired-end reads of 100 base pair.

Data analysis

The PALEOMIX pipeline v1.2.12 (Schubert *et al.* 2014) was used to remove adapters (AdapterRemoval v2.2.2) and to remove PCR duplicates with java v1.8.0 (Lindgreen 2012) from the raw sequencing reads. Using BWA-mem (v0.7.15) (Li 2013), the resulting filtered reads for each individual were then mapped against the Seychelles Magpie-Robin reference genome (Feng *et al.* 2020) and indel-realigned BAM files were subsequently generated (Genome Assembly Tool Kit (GATK) v3.8.0).

Sex-linked scaffolds were removed in order to reduce bias introduced by differing sequencing coverage, mixed ploidy, local mutation rate and sex-derived diversity found in avian sex chromosomes (Ellegren & Fridolfsson 1999, Montell *et al.* 2001). Scaffolds were identified as autosomal (not sex-linked) following identification and isolation of Z- and W-linked scaffolds determined by coverage, where males have twice the coverage of females. Only autosomal scaffolds over 100 Kb in size were retained culminating in a total of 232 (of 3664) scaffolds (maximum genome size of 0.95 Gb for each individual) for ANGSD v0.921 (Korneliussen 2014) genotype likelihood computation (-rf “region file”). Additionally, locations of repetitive regions

(paralogues with high levels of heterozygotes that could skew results) were provided by Feng *et al.* (2020) and were removed from the final scaffold list with bedtools v2.26.0 -subtract (Quinlan & Hall 2010) and from downstream analyses using the -rf option in ANGSD.

Given the heterogeneous sequencing coverage across the individuals (c. 2-57 fold coverage), and in particular the low coverage of a subset of the samples (Table S1), we used genotype likelihood-based methods to propagate the uncertainty in genotype calls to downstream analyses (Korneliussen 2014). Using the alignments generated with the PALEOMIX pipeline, genotype likelihoods were computed with ANGSD (v0.921) using the samtools model (Li *et al.* 2009) for genotype likelihoods (-GL 1). Genome wide single nucleotide polymorphisms (SNPs) were identified using the genotype likelihood-based analysis. In order to ensure a high quality dataset, filtering parameters and quality thresholds were used as follows: minimum read quality and mapping qualities of 30 (-minQ 30, -minMapQ 30), filtering of excessive mismatches (-C 50), base alignment quality (-BAQ), minimum site depth of 3 times the average coverage of the population (-setMinDepth), a minimum minor allele frequency of 5% (-minMaf 0.05), and a minimum *P*-value of 1e-6 for a SNP to be called (-SNP_pval 1e-6).

The genotype likelihoods from the 141 individuals were used to identify a total of 604408 polymorphic sites, which were subsequently used to generate a pairwise covariance matrix with PCAngsd v0.982 (Meisner & Albrechsten 2018). The pairwise covariance matrix was used as input for the “factoextra” package in R/3.4.3 (R Development Core Team 2008) to compute and visualise the principal components of the genotype data, plotted with 95% confidence ellipses.

Using the same genotype likelihood data for all 141 individuals as input, NGSadmix (v32) (Skotte *et al.* 2013) was used to estimate ancestry proportions for all individuals for a given number of ancestry components. Up to five ancestral components were estimated ('k' ranging from 2-5) with up to a maximum of 250 iterations to obtain the highest likelihood estimate.

Genetic differentiation between the five populations was calculated as F_{ST} values using allele frequencies estimated in ANGSD. An unfolded site frequency spectrum (SFS) was estimated (-dosaf 1, -fold 0) for each population using realSFS as implemented in ANGSD. The estimated spectra were subsequently used to construct a bootstrapped (-bootstrap 100) pairwise site frequency spectrum between each of the populations. During this latter stage the SFS was folded in realSFS (-fold 1) to account for the unknown ancestral state, therefore the Magpie-Robin reference genome was supplied as both the reference (-ref) and the ancestral (-anc) genomes. The

means of the pairwise SFS bootstrapped replicates were calculated and used as a prior to estimate global pairwise weighted F_{ST} values with realSFS (fst stats).

Global autosomal heterozygosity was calculated at the individual level using a folded SFS (-dosaf 1, -fold 1) for the estimation of homozygous and heterozygous sites across the genome, creating 20 bootstrapped replicates that were used to determine confidence intervals of the estimated SFS. In order to control for some of the bias introduced by coverage when heterozygosity was estimated at the individual level, a site coverage minimum of 4X was used as below this threshold has been found to have an effect on diversity estimates (van der Valk *et al.* 2019).

Only individuals with an average genome-wide coverage of 5X and above were included in the runs of homozygosity (ROH) analysis (136/141) as in Hooper *et al.* (2020). The largest 30 scaffolds for each of these individuals were selected for ROH analysis defined by a minimum length of 10 Mb. The longest scaffold was 57.83 Mb, the shortest was 10.48 Mb, with a total length of 0.73 Gb. A panel of SNPs for each island was generated as .ped files using ANGSD (-doPlink 2). These data were then fed into PLINK (v1.09beta5.4) (Chang *et al.* 2015) to assess genome-wide ROH (--homozyg). ROH were called using the following thresholds: a maximum gap between consecutive SNPs set at 1000 Kb bases (--homozyg-gap 1000), sliding windows set to 300 Kb (--homozyg-window-kb 300), minimum number of SNPs for an eligible run set at 50 (--homozyg-density 50), a maximum of 30 'missing' sites allowed per window (--homozyg-window-missing 30), and five heterozygous calls allowed per window (--homozyg-window-het 5). From this output, F_{ROH} (autozygosity) was calculated as the cumulative length of an individual's ROH above 4 Mb as a proportion of the length of the used scaffolds (0.73 Gb). This method allows differentiation between runs of homozygosity that are 'identical by state' (IBS) and 'identical by descent' (IBD), as short ROHs (defined here as less than 4 Mb in the Magpie-Robins) are expected to exist with distant inbreeding and thus are most likely explained by IBS.

RESULTS

The five island populations exhibit clear population structure

The Principal Component Analysis (PCA) reveals obvious clustering of the five islands, with greatest differentiation among the longest-standing translocations (Fig. 2) (Cousin in 1994, Cousine in 1995 and Aride in 2002). The founder population of Frégate sits toward the centre of the clustering and on no axis do Frégate and Denis segregate. The first three principal components account for 20.9%, 17.5% and 10%, respectively, representing almost half (48.4%) of the total variation among populations. The axis of the first principal component shows that Cousin differs most from the other populations, as it has the highest degree of variation. PC2 (Fig. 2A) and PC3 (Fig. 2B) separate Aride and Cousine, respectively. Higher PCs do not show any biologically meaningful clustering (Fig. S2).

Mirroring this, the highest degrees of differentiation resulting from pairwise F_{ST} analyses were observed between Aride, Cousin, and Cousine (unweighted F_{ST} values between 0.109-0.110, Fig. 2C). The F_{ST} values when comparing the population on the principal founder island (Frégate) to Denis, to the most recent translocation are at 0.039 only approximately half of that obtained when compared to populations from the other three islands.

The results from the NGSadmix analyses at ‘ $k=4$ ’ support the population structure observed in the results of the PCA and F_{ST} analyses with Denis Island as the least genetically differentiated from the Frégate Island population. Structure through admixture is evident by a preponderance of one ancestral state in each population (depicted in Fig. 3 with ‘island’ colour) occurring systematically with each increased ‘ k ’ defining Cousin, Cousine, and Aride in ‘ $k=$ ’ 2, 3, and 4, respectively. Interestingly, at ‘ $k=5$ ’, the differences *within* the Aride population show more variation than the other three translocated populations.

A species with very low genome-wide autosomal heterozygosity

The Seychelles Magpie-Robin genome-wide autosomal heterozygosity (‘global heterozygosity’ hereafter) was found to be exceptionally low, independent of population parameters. The average global heterozygosity (number of singletons/total number of sites) of each population was around 0.00015 (Fig. 4A). This value is less than half the average value reported from similar analyses of 42 other IUCN red-listed avian species (Li *et al.* 2014). To illustrate a comparison of heterozygosity across threatened avian species Fig. 4B uses nine of the published results in Li *et al.* (2014) together with the addition of our Magpie-Robin results.

Evidence of high inbreeding in the Seychelles Magpie-Robins

Only stretches of 4 Mb or longer were used to compute F_{ROH} (an estimate of inbreeding based on ROHs), as this length is an expected signal of historic inbreeding (Gibson *et al.* 2006, McQuillan *et al.* 2008, Simon-Sánchez *et al.* 2007). We observed not only that all populations exhibited ROHs of up to 8 Mb long, but every individual from Cousine and Denis Islands, and at least 80% of the individuals from the other islands had at least one ROH of 8–16 Mb (Fig. 5). Furthermore, at least half of all individuals on each island (Aride 59%, Cousin 63%, Cousine 70.6%, Denis 74.2%, Frégate 53.3%) exhibited at least one extremely long ROH (>16 Mb). With ROH shorter than 4 Mb excluded from IBD interpretation, all populations were observed to have individuals with high F_{ROH} . Lastly, and consistent with the fact that these islands have a higher proportion of individuals exhibiting ROHs in the largest two categories, the Cousine and Denis populations also exhibited the highest median F_{ROH} , while Frégate Island shows the lowest F_{ROH} median.

DISCUSSION

This study is the first comprehensive genomic assessment of the IUCN Red–Listed Endangered Seychelles Magpie-Robin. Our research aimed to answer fundamental questions about the current genetic state of the species using genome-wide SNP markers obtained from a sample set of 141 individuals representing the populations of the five islands within the Seychelles archipelago on which this species is extant. It is clear that the recent history of the Seychelles Magpie-Robin has shaped the species' genome, with genomic signatures of population collapse and inbreeding mirroring to some degree the isolation of the five islands. The results are discussed here in relation to the genetic impacts of population history, past conservation translocations, and the implications and applications of this research on the future conservation management of this, and other, endangered species.

Observed genetic impact of translocations

We firstly observed that the Denis Island population has not yet been strongly affected by patterns of genetic drift, as it retains most genetic similarity to Frégate. This may reflect the larger size of the main donor population (92) at the time the individuals were taken, the number of individuals donated (16 from Frégate Island and four from Cousin Island, although only two of the latter survived the first year), and the relatively recent time since introduction (2008). Furthermore, during the second-year census only two of the birds moved from Cousin Island were found

(S.M.A.R.T. *pers. comm.*), so the genetic contribution of birds from Cousin Island to the next generation appears to be limited.

A second striking observation is the considerably different levels of differentiation observed when comparing the Cousin and Cousine populations to their Frégate population source. We find this observation striking as it was recently suggested that the three sources are likely to be genetically ‘one population’ (Burt *et al.* 2016), given that: (i) the translocation of individuals happened when Frégate was at its smallest population size (23 individuals), (ii) the translocations happened within a very tight time frame (within two years), and (iii) the close proximity of the three islands suggests they are most prone to migrations. Nevertheless, it is clear from our results that the Cousin and Cousine populations were likely affected differently by genetic drift. In this regard we note that Cousin Island, that exhibits the strongest signal of divergence, experienced a period of severe weather between 2007 and 2008, which resulted in the loss of many birds (S.M.A.R.T. *pers. comms.*), and furthermore in 2008 four individuals were donated to the Denis Island translocation. Together these reduced the population by almost half (46 to 27 birds), thus rendering it much more susceptible to drift than the population of Cousine.

We thirdly observed that the genetic differentiation (F_{ST} analyses) observed between the Aride, Cousin, and Cousine populations, is likely highest as they are the longest established introduced populations. Furthermore, Cousin and Cousine received fewer birds donated from Frégate at a smaller population size. Thus, the genetic variation between these populations appears to be a product of the population structure (PCA and admixture analyses) resulting from recent translocations. In a recent genetic study involving the Crested Ibis *Nipponia nippon*, a species with a similar history to the Magpie-Robins, an F_{ST} value of 0.134 was considered to be relatively high given the short time (of 20 years) since new populations were established (Li *et al.* 2014), which yields similar implications for the results of the Magpie-Robin (highest F_{ST} value of 0.110). The severity of genetic drift evident already in the Magpie-Robin populations despite the short time since translocations confirms the well-established idea that the evolutionary effects of drift act quicker upon small populations, although rapid selection has been documented to act upon small populations, particularly when faced with adverse conditions (Lamichhaney *et al.* 2016).

The group of Endangered and Vulnerable species used in the Li *et al.* (2014) avian genomics study found an average heterozygosity estimate of 0.0018 (calculated as the proportion of sites used that

were observed heterozygous), with the lowest levels found in the Crested Ibis, Bald Eagle *Haliaeetus leucocephalus* and White-tailed Eagle *Haliaeetus albicilla* (0.00043–0.00040). In the Magpie-Robin, the very low level of observed genome-wide heterozygosity (0.00015) appears even more extreme than that of other threatened avian species (Li *et al.* 2014). However, this is more likely a feature of past genetic bottlenecks (1960s and before) than of the more recent translocations, given that the Frégate population harbours a similar level of heterozygosity as all other populations. The genetic differentiation observed in three of the four islands is indicative of an artefact of the recent translocations. While estimates for the Magpie-Robins were made at the individual level, the purpose of this was to gauge whether global heterozygosity differs among islands.

The availability of a de novo sequenced, relatively high-quality Magpie-Robin reference genome assembly (75% of the genome comprises scaffolds of 10 Mb and above in length) and the genome-wide SNP dataset for a representative number of individuals from each island has allowed for reliable preliminary ROH analysis for the Magpie-Robins. However, as a full chromosomal level genome assembly does not yet exist for the Seychelles Magpie-Robin, any deeper interpretation should be applied with some caution. ROH analysis revealed extensive tracts of homozygous SNPs, a signature of recent inbreeding. While the overall degree of relatedness generally appears high in each population, the length of the ROH can be used to infer recentness of inbreeding events and long ROHs (up to 56 Mb in some individuals; almost the entire length of the longest scaffold used) indicates that within each Magpie-Robin population, close familial mating has happened in recent generations (Kirin *et al.* 2010, McQuillan *et al.* 2008, Moyse *et al.* 2016).

In 2014, the Aride population underwent a catastrophic crash, with undetermined cause, and the population has since been unable to recover to more than a dozen individuals. One aim of this research was to ascertain if there may have been an underlying genetic cause for this crash, and whether assisted gene flow to diversify the gene pool should be considered for population recovery. As the low level of heterozygosity, the long length of ROH, and high F_{ROH} is consistent between islands, this preliminary genomic investigation does not show an apparent genetic cause as to why this crash occurred (perhaps indicating a lack of recruitment or social conflicts as discussed in Burt *et al.* (2016), or perhaps an environmental cause). However, we note that one recent study demonstrated that the length of a ROH can be associated with the likelihood of

harbouring deleterious variants (Sams & Boyko 2019). Thus, given that very long ROHs were found in all five Magpie-Robin populations, future investigation into what genes are associated with the regions of long ROH may reveal evidence of suspected inbreeding depression (i.e. alleles in which the homozygous recessive state would be lethal) in any of the five extant populations, which may not only give insight into past population size fluctuations, but could also serve as an additional consideration in conservation development for this species.

Application of genomics to conservation management

Translocations from small populations and/or endangered species are increasing in frequency, yet remain a contentious issue in conservation (Berger-Tal *et al.* 2019). Obtaining a genetic understanding of both the donor and donated populations is required for initial planning and effective management strategies. However, these considerations are rarely undertaken or monitored, either because of time constraints, availability of resources, or limited understanding of the impact it can have (Armstrong & Seddon 2008, Jamieson 2009). Extensive research into the impact of inbreeding on small populations has determined that inbreeding is associated with an increased extinction risk through a reduction in reproductive success and survival (Frankham 2010, O'Grady *et al.* 2006), and inbreeding has been shown to have had a negative impact on the survival of island populations of passerines (Jamieson 2009, Keller 1998). Thus, quantifying individual inbreeding levels is integral for investigation of effects of inbreeding in order to inform active management of populations. Using genome wide markers for the ROH analyses carried out in this study has allowed for accurate quantification, and an essential genetic understanding, of such an inbred species.

In small, saturated populations it can prove difficult to ameliorate or ‘break up’ ROHs, as there is limited mate choice. If long ROHs are not shared between isolated populations, implementing an assisted gene flow program (i.e. genetic rescue) may help to increase the genetic diversity required for small populations to persist (Ralls *et al.* 2017), and thus could provide long-term benefits to participating populations (Nedda *et al.* 2019). A meta-analysis of ‘genetic rescue’ conservation actions (translocating individuals of the same species between already established populations, when these populations are isolated and little gene flow exists between them) found mostly positive benefits arose from diversification of the gene pool, and such actions seemed to combat the long-term effects of reintroductions through single event translocations (Frankham 2015). In

the context of our study, the very limited gene flow observed between these five Magpie-Robin populations, and thus the reproductive isolation post translocation, suggests that the Seychelles Magpie-Robins may make a good candidate for genetic rescue. However, while it may be possible to attempt such genetic diversification it is important to note potential risks associated with this approach, ranging from the transmission of unidentified diseases, to the potential death of the donated individuals from either stress or failure to settle into the new environments.

In at least some of the previous Magpie-Robin translocations, the individuals were chosen for translocation using observational monitoring data in order to avoid moving closely related individuals. As our genetic data shows that the populations today have a high level of background relatedness, it highlights an emphasis on the importance of molecular application to translocation donor selection. Thus, the new understanding of the population-level genetics using the molecular insight provided by our study is expected to improve future translocation success by means of selectively maximising genetic diversity for the Seychelles Magpie-Robins. The molecular understanding of inbreeding gained from this project can now be utilised as genetic consideration for the proposed future conservation translocation of Seychelles Magpie-Robins in an attempt to establish a new, sixth, population on to North Island, Seychelles (Millet *et al.* 2019).

Using either prospective or retrospective genetic monitoring, conservation genetics has already been partly attributed to the success of two other avian species in Seychelles (Bristol *et al.* 2013, Wright *et al.* 2014). Eleven of the 13 endemic avian species in Seychelles currently face some level of vulnerability (IUCN 2021), and using genetic resources to shape conservation action/policy should be seen as a natural addition to many recovery plans going forward, importantly, *alongside* maintaining direct actions of habitat restoration, removal of invasive species, protective legislation, and public education.

In 2005, the Seychelles Magpie-Robin was downgraded from ‘Critically Endangered’ to ‘Endangered’ as the census population exceeded a threshold of 50 mature individuals (Birdlife International 2019). However, this was done with no prior genetic understanding of the species. Environmental stochasticity and the vulnerability of oceanic islands to such events, coupled with the suspected lack of genetic ability the Magpie-Robins may have to adapt to these changes, could indicate a higher extinction risk than the species is currently categorised for. While we would usually celebrate a reclassification considering a species to be less endangered, funding and

research efforts were seen to plateau in conjunction with this change in conservation status. With a species with such small and variable population sizes, a lack of funds and resources to continue efforts to help save the Magpie-Robins could prove fatal for the species, as with six other avian species that have gone extinct from this archipelago. Based on the example of our study, we advocate that if a species is more vulnerable than currently defined, due to factors such as genetics, this needs to be duly considered in its IUCN conservation status.

A huge thank you is given to members of S.M.A.R.T. and all local island conservation staff, past and present, on Aride, Cousin, Cousine, Denis, and Frégate Islands, for collection and care of blood samples (and the birds!). These thanks are further extended to Nature Seychelles, specifically Eric Blais, and to Ashley Dias and the Seychelles Ministry of Environment, Energy and Climate for facilitating all necessary permits and for the careful transport of blood samples from Seychelles. We thank Kate Lessells and Christa Mateman at NIOO for long-term care and generous handover of Aride blood samples that were integral to this research. We also thank Fatima Sanchez, Andy Foote, and Mark de Manuel for advice on various program parameters, Stephen Garrett for invaluable grammar corrections of the manuscript, and Mikkel Skovrind for your excellent map creation skills, and Chris Tagg for your endearing photo of Magpie-robin feeding time. And an overwhelming thanks to Cheryl Sanchez for your enthusiasm for this research before it even began, and for your local support with getting this project off the ground. We also thank the Experiment fundraising platform and all those who donated at the early stages of this project. This research was principally supported through ERC Consolidator grant 681396 Extinction Genomics and Danish National Research Foundation Center for Evolutionary Hologenomics DNRF143 grants to MTPG. RF thanks the Danish National Research Foundation for its support of the Center for Macroecology, Evolution, and Climate (DNRF96).

Data accessibility

All raw sequence read data have been deposited into the Sequencing Read Archive (SRA: <https://www.ncbi.nlm.nih.gov/sra>) and are accessible through project number PRJNA722144.

Author contributions

ELC and MTPG project conceptualisation. Research principally supervised by MTPG at the University of Copenhagen, and by GJ and MG at the University of Bristol. BD, WA, NS, LC and AZ were responsible for Magpie-Robin blood sample and data collection. ELC carried out all lab work with assistance from LP, AMR, SM and GP. CL and GZ performed sequencing. ELC and SG performed all data analyses, with data interpretation by ELC, RF, SG, and MTPG. The sexing scripts were provided by RF. The manuscript was written by ELC and MTPG with input from all co-authors.

Competing interests

None

REFERENCES

- Berger-Tal, O., Blumstein, D., & Swaisgood, R. (2020). Conservation translocations: a review of common difficulties and promising directions. *Animal Conservation*, 23(2), 121-131.
<https://doi.org/10.1111/acv.12534>
- BGI (2019) *NGS Platforms*. <https://www.bgi.com/us/resources/sequencing-platforms/>.
Downloaded on [10.05.2020]
- Birdlife International (2008). Preventing Extinctions Brochure.
http://www.birdlife.org/sites/default/files/attachments/preventing_extinctions_brochure.pdf
Downloaded on [18 May 2019]
- BirdLife International (2019). Species factsheet: *Copsychus sechellarum*. <http://www.birdlife.org>.
Downloaded on [05.05.2019]
- Burt, A., Gane, J., Olivier, I., Calabrese, L., De Groene, A., & Liebrick, T. *et al.* (2016). The history, status and trends of the Endangered Seychelles Magpie-robin *Copsychus sechellarum*. *Bird Conservation International*, 26(4): 505-523. doi:
<https://doi.org/10.1017/s0959270915000404>

Carøe, C., Gopalakrishnan, S., Vinner, L., Mak, S., Sinding, M., & Samaniego, J. *et al.* (2017). Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution*, 9(2): 410–419. <https://doi.org/10.1111/2041-210x.12871>

Chang C. C., Chow C. C., Tellier L. C., Vattikuti S., Purcell S. M., *et al.* (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaSci* 4(7). doi: <https://doi.org/10.1186/s13742-015-0047-8>

Collar, N. J. & Andrew, P. (1988). Birds to Watch: The ICBP World-Checklist of Threatened Birds. International Council for Bird Preservation, Cambridge, UK

Feng, S., Stiller J., Deng, Y., Armstrong, J., Fang, Q., *et al.* (2020). Densely sampling genomes across the diversity of birds increases power of comparative genomics analyses. *Nature*, 587: 252–257. doi <https://doi.org/10.1038/s41586-020-2873-9>

Frankham, R. (1980). The founder effect and response to artificial selection in *Drosophila*. In: Robertson, A. (ed.) Selection Experiments in Laboratory and Domestic Animals, pp. 87—90. Commonwealth Agricultural Bureaux, Farnham Royal

Frankham, R., Ballou, J., & Briscoe, D. (2010). *Introduction to Conservation Genetics* (2nd ed.). Cambridge University Press

Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, 24(11): 2610-2618. doi: <https://doi.org/10.1111/mec.13139>

Fridolfsson, A., & Ellegren, H. (1999). A Simple and Universal Method for Molecular Sexing of Non-Ratite Birds. *Journal of Avian Biology*, 30(1): 116. doi: <https://doi.org/10.2307/3677252>

Gaymer, R., Blackman, R., Dawson, P., Penny, M., & Penny, C. (2008). The endemic birds of Seychelles. *Ibis*, 111(2): 157-176. doi: <https://doi.org/10.1111/j.1474-919x.1969.tb02524.x>

Gibson, J., Morton, N., & Collins, A. (2006). Extended tracts of homozygosity in outbred human populations. *Human Molecular Genetics*, 15(5): 789-795. doi: <https://doi.org/10.1093/hmg/ddi493>

Government of Seychelles (2014). *Fifth National Report to the United Nations Convention on Biological Diversity. Ministry of Environment and Energy*, P.O. Box 445, Botanical Gardens, Mont Fleuri, Victoria, Republic of Seychelles.

GoS (2014). Seychelles Biodiversity Strategy and Action Plan 2015-2020. Editors: John Nevill, Jacques Prescott, Nirmal Jivan Shah, and Marie-May Jeremie

Hartlaub, G. (1877). *Die Vögel Madagascars und der benachbarten Inselgruppen: ein Beitrag zur Zoologie der äthiopischen Region*. Halle, Germany: HW Schmidt

Hooper, R., Excoffier, L., Forney, K., Gilbert, M. T. P., Martin, M., et al. (2020). Runs of homozygosity in killer whale genomes provide a global record of demographic histories. bioRxiv preprint doi: <https://doi.org/10.1101/2020.04.08.031344>

IUCN (2021). *The IUCN Red List of Threatened Species. Version 2021-1*. <https://www.iucnredlist.org>. Downloaded on [01.04.2021]

Keller, L. (1998). Inbreeding and its fitness effects in an insular population of song sparrows (*Melospiza melodia*). *Evolution*, 52(1): 240-250. <https://doi.org/10.2307/2410939>

Kirin, M., McQuillan, R., Franklin, C., Campbell, H., McKeigue, P., & Wilson, J. (2010). Genomic Runs of Homozygosity Record Population History and Consanguinity. *Plos ONE*, 5(11): e13996. doi: <https://doi.org/10.1371/journal.pone.0013996>

Korneliussen, T., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, 15(1). <https://doi.org/10.1186/s12859-014-0356-4>

Lamichhaney, S., Han, F., Berglund, J., Wang, C., Almen, M., & Webster, M. *et al.* (2016). A beak size locus in Darwins finches facilitated character displacement during a drought. *Science*, 352(6284): 470-474. <https://doi.org/10.1126/science.aad8786>

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Nils Homer, *et al.* (2009) 1000 Genome Project Data Processing Subgroup Author Notes the Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16): 2078–2079. doi: <https://doi.org/10.1093/bioinformatics/btp352>

Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v1 [q-bio.GN]

Li, S., Li, B., Cheng, C., Xiong, Z., Liu, Q., & Lai, J. *et al.* (2014). Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. *Genome Biology*, 15(12). doi: <https://doi.org/10.1186/s13059-014-0557-1>

Lindgreen (2012): AdapterRemoval: Easy cleaning of Next Generation Sequencing reads, BMC Research Notes, 5:337. doi: <http://www.biomedcentral.com/1756-0500/5/337/>

Mak, S., Gopalakrishnan, S., Carøe, C., Geng, C., Liu, S., & Sinding, M. *et al.* (2017). Comparative performance of the BGISEQ-500 vs Illumina HiSeq2500 sequencing platforms for palaeogenomic sequencing. *Gigascience*, 6(8). doi: <https://doi.org/10.1093/gigascience/gix049>

McQuillan, R., Leutenegger, A., Abdel-Rahman, R., Franklin, C., Pericic, M., & Barac-Lauc, L. *et al.* (2008). Runs of Homozygosity in European Populations. *The American Journal of Human Genetics*, 83(5): 359-372. <https://doi.org/10.1016/j.ajhg.2008.10.009>

Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*, 210(2): 719-731. <https://doi.org/10.1534/genetics.118.301336>

Millett, J.E., Accouche, W., Havemann, K. and Retief, T. (2019). Feasibility Study for the Establishment of a Seychelles Magpie-Robin *Copsychus sechellarum* population on North Island, Seychelles. Unpublished Report. North Island, Seychelles.

Montell, H., Fridolfsson, A., & Ellegren, H. (2001). Contrasting levels of nucleotide diversity on the Avian Z and W sex chromosomes. *Molecular Biology and Evolution*, 18(11): 2010-2016. doi: <https://doi.org/10.1093/oxfordjournals.molbev.a003742>

Moyse, E., Nguyen, V. Vu, D. &; Farnir, F. (2016). Inbreeding and genomic runs of homozygosity in Vietnamese local chickens. Scientific congresses and symposiums: poster. <http://hdl.handle.net/2268/207658>. Downloaded on [15 April 2019]

Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y., & Wang, J. (2012). SNP Calling, genotype calling, and sample allele frequency estimation from New-Generation Sequencing data. *Plos ONE*, 7(7): e37558. doi: <https://doi.org/10.1371/journal.pone.0037558>

O'Grady, J., Brook, B., Reed, D., Ballou, J., Tonkyn, D., & Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation*, 133(1): 42-51. doi: <https://doi.org/10.1016/j.biocon.2006.05.016>

Oustalet, M. E. (1878). Etude sur la faune ornithologique des îles Seychelles. *Bulletin de la Société philomathique (Paris)*, 7(2): 161-206

Pike, N. (1872). A visit to the Seychelles. *Transactions of the Royal Society of Arts & Sciences of Mauritius*, 6: 83-142

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M., & Bender, D. et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3): 559-575. doi: <https://doi.org/10.1086/519795>

Quinlan, A. R. & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 26(6): 841–842. doi: [10.1093/bioinformatics/btq033](https://doi.org/10.1093/bioinformatics/btq033)

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL

<http://www.R-project.org>

Ralls, K., Ballou, J., Dudash, M., Eldridge, M., Fenster, C., & Lacy, R. *et al.* (2017). Call for a paradigm shift in the genetic management of fragmented populations. *Conservation Letters*, 11(2): e12412. doi: <https://doi.org/10.1111/conl.12412>

Robinson, N., Scheele, B., Legge, S., Southwell, D., Carter, O., & Lintermans, M. *et al.* (2018). How to ensure threatened species monitoring leads to threatened species conservation. *Ecological Management & Restoration*, 19(3): 222-229. doi: <https://doi.org/10.1111/emr.12335>

Sams, A. J., & Boyko, A. R. (2018). Fine-scale resolution of Runs of Homozygosity reveal patterns of inbreeding and substantial overlap with recessive disease genotypes in domestic dogs. *G3 (Bethesda)*. 9(1): 117–123. doi: 10.1534/g3.118.200836

Saremi, N., Supple, M., Byrne, A., Cahill, J., Coutinho, L., & Dalén, L. *et al.* (2019). Author Correction: Puma genomes from North and South America provide insights into the genomic consequences of inbreeding. *Nature Communications*, 10(1). doi: <https://doi.org/10.1038/s41467-019-13096-3>

Schubert, M., Ermini, L., Sarkissian, C., Jónsson, H., Ginolhac, A., & Schaefer, R. *et al.* (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nature Protocols*, 9(5): 1056-1082. doi: <https://doi.org/10.1038/nprot.2014.063>

Simon-Sanchez, J., Scholz, S., Fung, H., Matarin, M., Hernandez, D., & Gibbs, J. *et al.* (2006). Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. *Human Molecular Genetics*, 16(1): 1-14. doi: <https://doi.org/10.1093/hmg/ddl436>

Skotte, L., Korneliussen, T., & Albrechtsen, A. (2013). Estimating individual admixture proportions from Next Generation Sequencing data. *Genetics*, 195(3): 693-702. doi: <https://doi.org/10.1534/genetics.113.154138>

Inkscape Project (2020). *Inkscape*. Retrieved from <https://inkscape.org>

Watson, J., Warman, C., Todd, D., & Laboudallon, V. (1992). The Seychelles *et al.* *Copsychus sechellarum*: ecology and conservation of an endangered species. *Biological Conservation*, 61(2): 93-106. doi: [https://doi.org/10.1016/0006-3207\(92\)91099-e](https://doi.org/10.1016/0006-3207(92)91099-e)

Wilson, J. & Wilson, R. (1978). Observations on the Seychelles *et al.* (*Copsychus sechellarum*). *Bulletin of the British Ornithologists Club*, 98:15–21

Wright, D. J., Spurgin, L. G., Collar, N. J., Komdeur, J., Burke, T. & Richardson, D. S. (2014). The impact of translocation on neutral and functional genetic diversity within and amongst populations of the Seychelles warbler. *Molecular Ecology*, 23(9): 2165-2177. doi: 1 <https://doi.org/10.1111/mec.12740>

Tables and Figures (with captions)

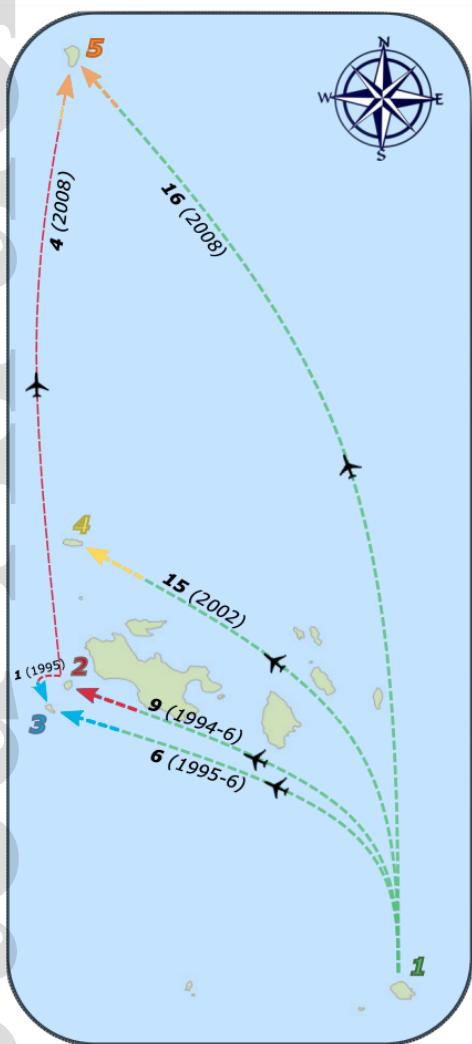


Figure 1. Seychelles Magpie-Robin conservation translocations 1994-2008: A cropped map of Seychelles created with Inkscape Version 1.0 (Inkscape Project, 2020) demonstrating the relative positions of each of the ‘Magpie-Robin’ islands. Directionality of translocations for donor-to-receiver is indicated by coloured arrows, along which the number of birds donated is defined in bold, and the year/s of translocation events are given in parentheses. Islands are labelled by number corresponding with the order of translocations, and have been assigned the colour of the ‘island identifier’ ring used when banding, which correspond as follows: 1 (green) Frégate; 2 (red)

Cousin; 3 (blue) Cousine; 4 (yellow) Aride; 5 (orange) Denis. The distance between Frégate and Denis is roughly 65 km; Aride and Denis is 45 km; Aride to Cousin is 10 km; and Cousin to Cousine is 2 km. The Seychelles archipelago has a total landmass of 459 km². The total area inhabited by the Magpie-Robins is only 4.86 km².

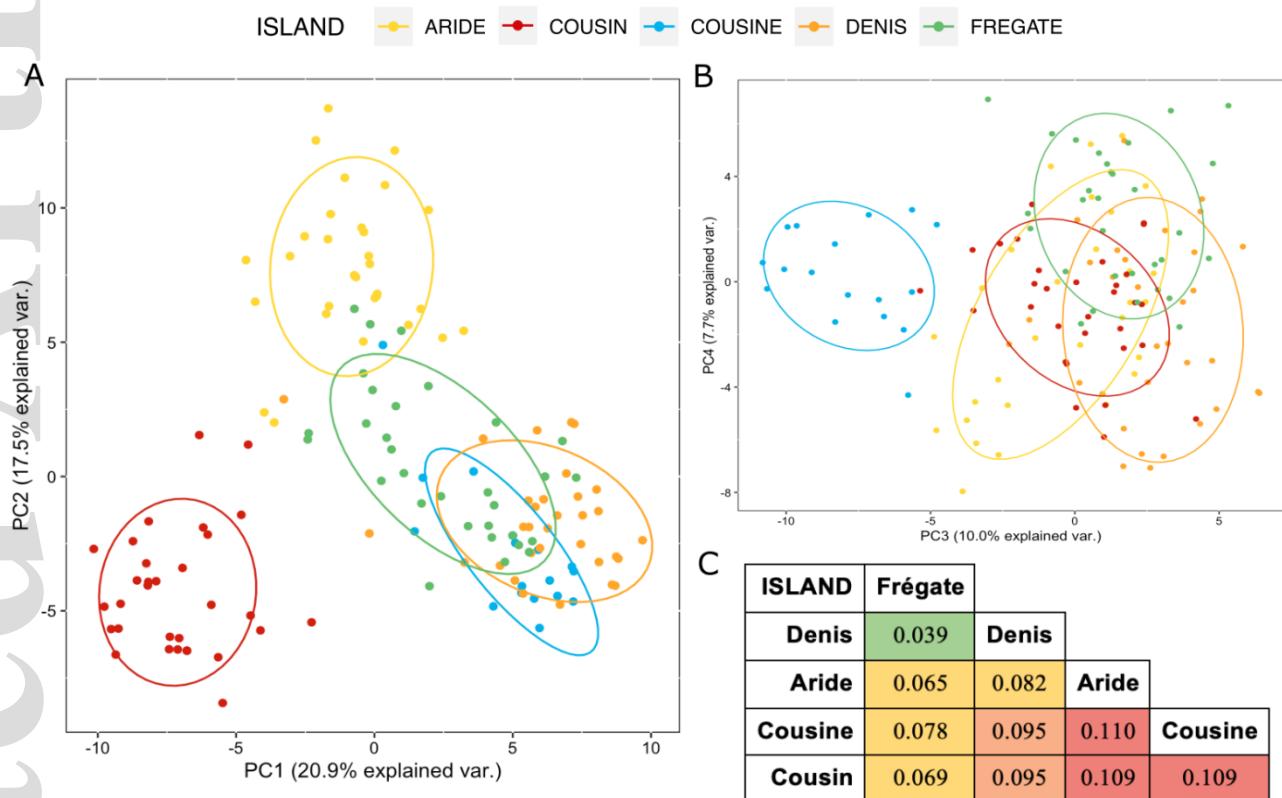


Figure 2: The five island populations exhibit clear population structure: A) PC1 and PC2 from the Principal Component Analysis, performed on the genotype likelihood from the full dataset, n=141 individuals, using PCAngsd. B) Same as A) but shows PC3 and PC4. C) Pairwise unweighted F_{ST} estimates computed with realSFS in ANGSD.

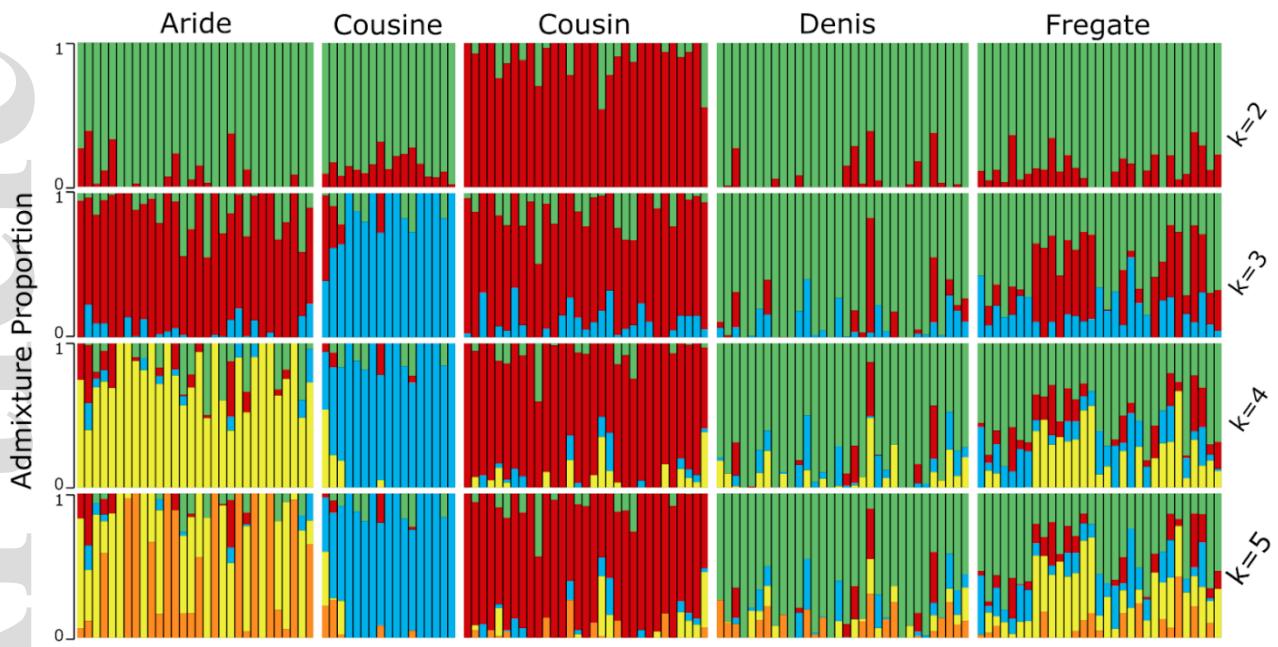


Figure 3: Population structure observed in the extant populations: NGSadmix results generated from the genotype likelihood data of all 141 individuals. Individuals are represented by columns, whereas rows depict the Admixture proportions based on the assumption of different numbers of ancestral populations. Displayed are results from $k=2$ (top) to $k=5$ (bottom).

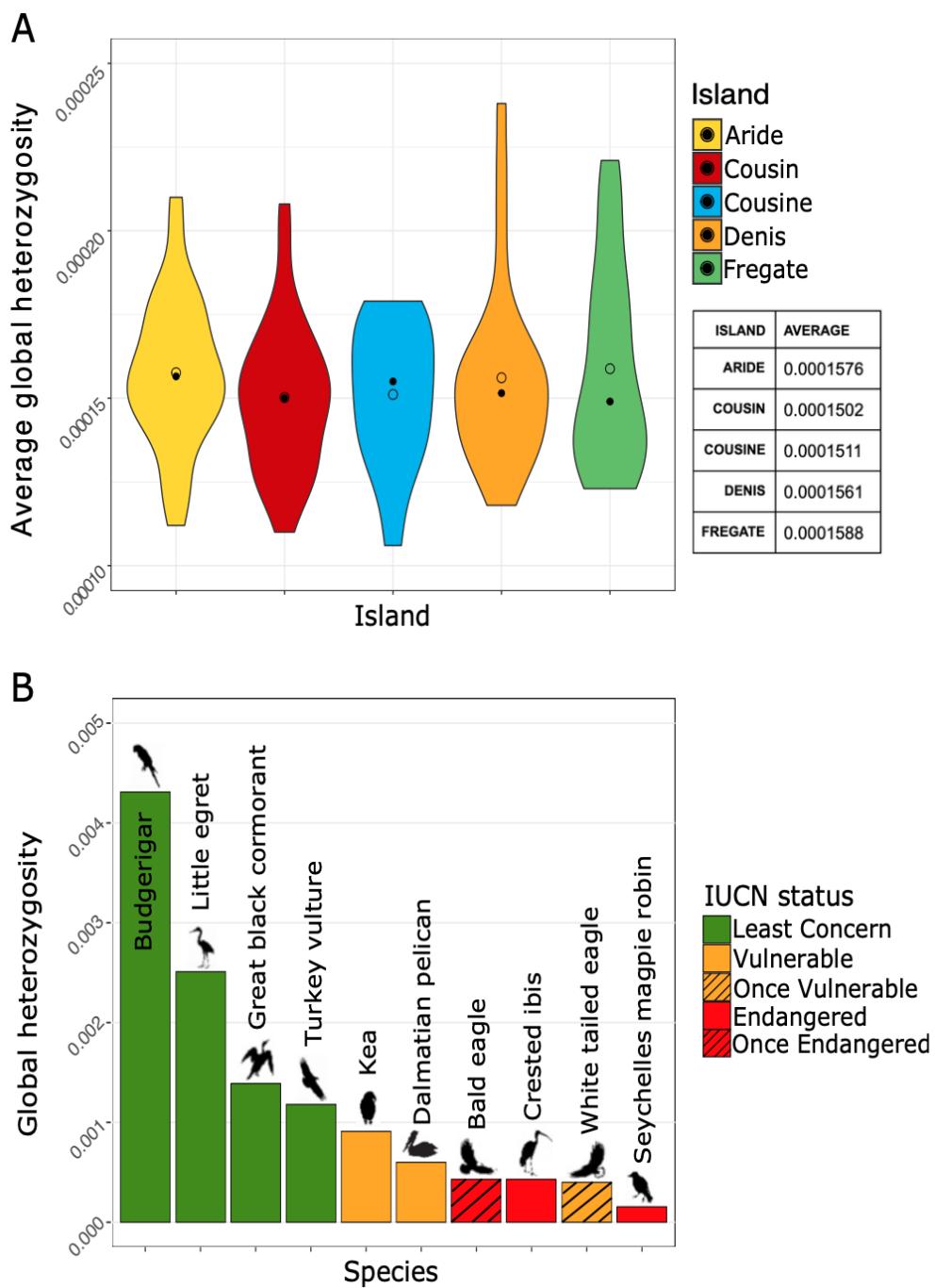


Figure 4: A species with very low genome-wide heterozygosity. A) Global autosomal average heterozygosity estimates for each Seychelles Magpie-Robin sampled, grouped by island using violin, across all five Magpie-Robin island populations. The average is calculated from bootstrapped estimates for each Seychelles Magpie-Robin sample, giving an overview of island level heterozygosity (hollow circle represents calculated mean, filled circle is calculated median). Inset shows the average heterozygosity estimate value for each island population. B) Genome-wide autosomal heterozygosity estimates for 10 different IUCN Red-Listed avian species*: nine

adapted from Li et al. (2014) results and the Seychelles Magpie-Robin estimates gained from this research are incorporated for comparison. *Scientific name of included species reading left to right: *Melopsittacus undulates*, *Egretta garzetta*, *Phalacrocorax carbo*, *Cathartes aura*, *Nestor notabilis*, *Pelecanus crispus*, *Haliaeetus leucocephalus*, *Nipponia nippon*, *Haliaeetus albicilla*, *Copsychus sechellarum*

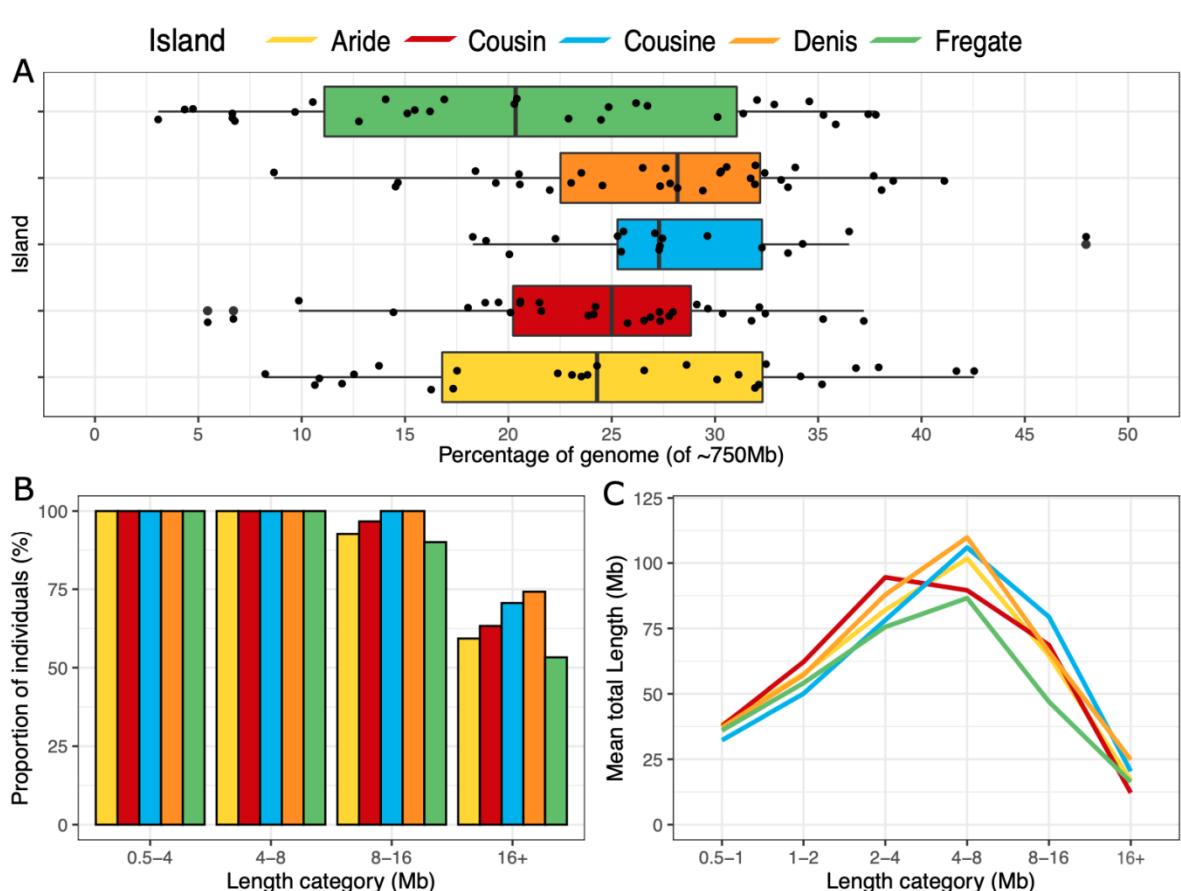


Figure 5: Evidence of high inbreeding in the Seychelles Magpie-Robins, found through ROH analysis: A) Island F_{ROH} (autozygosity) where only ROH of 4 Mb long or above were included as a factor for IBD. B) Proportions of individuals of each island who possess ROH of each size category. C) Mean total length of ROH by island for six length categories.