Diversity and Distributions

WILEY

Diversity and Distributions

RESEARCH ARTICLE OPEN ACCESS

Genetic Differentiation in a Wide-Ranging Tropical Seabird in the Indian Ocean Is Linked With Oceanographic Factors

Helena Teixeira[1,2] | Malcolm A. C. Nicoll[3] | Audrey Jaeger[1] | Nancy Bunbury[4,5] | Arthur Choeur[1] | Luke A'Bear[4,5] | Camille Lebarbenchon[6] | Maria Bielsa[4] | Ruth Dunn[7,8] | Robin Freeman[3] | Nirmal Jivan Shah[9] | Gerard Rocamora[10,11] | Licia Calabrese[10,11] | Alice M. Trevail[12] | Julie Tourmetz[13] | Stephen Votier[7] | Matthieu Le Corre[1] | Laurence Humeau[1]

[1]UMR ENTROPIE (Université de La Réunion, IRD, CNRS, IFREMER, Université de Nouvelle-Calédonie), Saint Denis Cedex, France | [2]ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France | [3]Institute of Zoology, Zoological Society of London, London, UK | [4]Seychelles Islands Foundation, Victoria, Mahé, Seychelles | [5]Centre for Ecology and Conservation, University of Exeter, Cornwall, UK | [6]UMR Processus Infectieux en Milieu Insulaire Tropical (PIMIT), INSERM 1187, CNRS 9192, IRD 249, Université de la Réunion, La Réunion, France | [7]The Lyell Centre, Heriot-Watt University, Edinburgh, UK | [8]Lancaster Environment Centre, Lancaster University, Lancaster, UK | [9]The Centre for Environment and Education, Nature Seychelles, Mahé, Seychelles | [10]Island Conservation Society, Mahé, Seychelles | [11]Island Biodiversity and Conservation Centre, University of Seychelles, Mahé, Seychelles | [12]Environment and Sustainability Institute, University of Exeter, Penryn, UK | [13]Société d'Etudes Ornithologiques de La Réunion (SEOR), Saint André, Réunion Island, France

Correspondence: Helena Teixeira (helena-marisa.osorio-teixeira@ird.fr)

Received: 25 March 2025 | **Revised:** 8 July 2025 | **Accepted:** 22 August 2025

Editor: Yoan Fourcade

Funding: This study was supported by the Bertarelli Foundation as part of the Bertarelli Programme in Marine Science (Project 822916), the FEDER Smac (2020–2022, N° RE0022954), the FEDER "Pathogenes associés à la Faune Sauvage Océan Indien" (Programme Opérationnel de Coopération Territoriale 2007–2013; #31189) and the CNRS – INEE/TAAF (AAP Iles Eparses "OMABIO" project). M.N. and R.F. were also supported by Research England.

Keywords: connectivity | conservation | genetic differentiation | genomics | marine biogeography | seabirds | taxonomic units | Western Indian Ocean

ABSTRACT

Aim: Knowledge of the main drivers of population differentiation is crucial for understanding evolutionary processes and preserving biodiversity. While primarily studied in terrestrial habitats, the mechanisms operating in the marine realm are less well understood. This study reconstructed the phylogeographic history of a tropical seabird to identify relevant marine barriers promoting intraspecific diversity in the Western Indian Ocean.

Location: Western Indian Ocean.

Taxon: Three subspecies of tropical shearwater*: Puffinus bailloni bailloni* , *P. b. nicolae, P. b. colstoni.*

Methods: We used restriction site-associated DNA sequencing and applied population genomics to birds from six breeding colonies to assess intraspecific diversity, population genetic structure and connectivity in the tropical shearwater. Results were complemented with data from six oceanographic variables and effective migration surfaces to evaluate the role of oceanographic factors in driving population differentiation.

Results: All analyses consistently separated the birds from the northern colonies (subsp. *nicolae* and *colstoni*) from those of the southern islands (subsp. *bailloni*), but failed to assign the *colstoni* birds as a different taxon. Results revealed remarkable levels of genetic differentiation within an ocean basin in a highly vagile species and suggested higher levels of gene flow at the northern limit of the species' distribution compared to the southern range.

Main Conclusions: Our study suggests that ocean surfaces and sea surface temperature may constitute an important barrier to gene flow for the tropical shearwater and potentially other marine species in the region. This study does not support the *colstoni*

form as a different subspecies, highlighting the need for further taxonomic reassessment. Ultimately, the results allowed us to identify Europa and Aldabra as the most threatened management units and propose conservation strategies directly applicable to these most at-risk colonies.

1 | Introduction

Since the pivotal study of the Galapagos finches by Charles Darwin, the spatial variation of biodiversity patterns has become a central research focus. Knowledge of the main drivers of population differentiation is crucial for understanding species distribution, persistence and defining effective conservation units (Humeau et al. 2020; Morris-Pocock et al. 2016). Mainly studied in terrestrial habitats, the mechanisms operating in the marine realm are less well understood due to the continuous nature of the marine environment, the absence of (obvious) physical barriers and the complex ecology of marine species (Faria et al. 2010; Morris-Pocock et al. 2010; Russo et al. 2016). Over the past few decades, advances in new technologies (e.g., biologging data, high-throughput sequencing) have enabled researchers to track the movement of individuals and assess the population connectivity of marine species. Multiple studies have inclusively reported the existence of population genetic structure in highly vagile marine species, including whales (Olavarría et al. 2007), sharks (Duncan et al. 2006; Nikolic et al. 2020), dolphins (Mendez et al. 2010; Richards et al. 2013) and sea turtles (Bourjea et al. 2007; Jensen et al. 2019), contradicting the expectations of substantial gene flow among populations in marine ecosystems. Oceanographic conditions (e.g., sea surface temperature (SST), salinity, oceanic currents) (Faria et al. 2010; Gómez-Díaz et al. 2009, 2012), historical climatic changes, particularly in high latitude species, isolation by distance (Kersten et al. 2021) and behavioural traits (e.g., dispersal capacity, assortative mating, differences in breeding phenology, natal or breeding philopatry) have emerged as the most cited mechanisms of isolation in marine species.

The Western Indian Ocean (WIO), here defined as the region extending from the coasts of East Africa (35° E) to the longitude 75° E and situated between the latitudes 0° and 25° S, hosts remarkable levels of terrestrial and marine endemism (Keesing and Irvine 2005). The region experiences a tropical to subtropical climate influenced by the Indian Ocean's seasonal monsoons, with the areas near the equator characterised by a warmer and humid climate, while regions further south experience distinct wet and dry seasons, with a peak rainfall during the northeast monsoon (December to April) (Bailey 1968; Keesing and Irvine 2005). The monsoon winds also play a crucial role in determining the direction and intensity of the surface currents in the region (Schott and Mccreary 2001). Although the WIO offers an ideal system to test biogeographic hypotheses on marine biota diversification, only a few studies have inferred the population structure of highly mobile species (Humeau et al. 2020; Jensen et al. 2019) and other marine-dependent species (Gagnaire et al. 2009; Triest et al. 2021; van der Ven et al. 2021) at a large scale. Nevertheless, these studies relied on a limited number of molecular markers, which may lack the power to fully characterise contemporary factors influencing population structure and gene flow (Choquet et al. 2023; Poelstra et al. 2022).

Among mobile marine species, seabirds stand out as an intriguing model to evaluate the mechanisms underlying population differentiation due to the dichotomy between their capacity to travel vast transoceanic distances (Egevang et al. 2010) and the tendency to return to their natal colonies to breed (Friesen 2015; Milot et al. 2008). Shearwaters (order Procellariiformes) are highly pelagic and philopatric seabirds that breed mostly on remote oceanic islands (Brooke 2004). The phylogenetic complexity of Procellariiformes, a debated group in avian systematics (Ferrer Obiol et al. 2021, 2023; Gangloff et al. 2012, 2013; Pyle et al. 2011), is notably exemplified by the tropical shearwater (*Puffinus bailloni*), which is distributed across the Indian and Pacific Oceans. Within the WIO, it is classified into four subspecies based on morphological differences, with the smallest birds found in the inner islands of Seychelles (subsp. *nicolae*; body mass 168.5 ± 18.99 g), intermediate-sized birds found exclusively at Aldabra Atoll (subsp. *colstoni*; 214.3 ± 3.78 g), and the largest birds in Réunion (subsp. *bailloni*; 217.4 ± 15.08) and Comoros (subsp. *temptator*; body weight unknown) (Bretagnolle et al. 2000; Shirihai and Bretagnolle 2015) (Figure 1). The species is additionally reported to breed in small colonies in Europa Island (Mozambique Channel), the Maldives and the Chagos Archipelago (Table 1). Subsp. *bailloni* is also distinguished from other specimens by its unique white undertail coverts (Bretagnolle et al. 2000). This broad classification is, however, controversial (Austin et al. 2004; Bretagnolle et al. 2000; Jouanin 1970; Shirihai and Christie 1996; Torres et al. 2021) and uncorroborated by other criteria such as genetic data. Outside the WIO, the species is also represented by a fifth subspecies, *P. bailloni persicus*, found only on Socotra (Yemen), and a sixth subspecies, *P. bailloni dichrous*, which breeds on several islands in the southern Pacific Ocean. The population structure of this species therefore remains unresolved at all spatial scales, along with much of its behaviour at sea and on land, hindering efforts to assess dispersal barriers and limiting our understanding of the process of genetic differentiation in the region.

Here, we use restriction site-associated DNA sequencing (DArTseq) and apply phylogenomic and population genomics methods to six tropical shearwater populations from all major breeding areas of the western Indian Ocean to (i) identify diagnosable Operational Taxonomic Units (OTUs; i.e., populations characterised by significant phylogenetic divergence) from the sampled populations; (ii) investigate levels of genetic differentiation and population structure in the WIO; and (iii) evaluate the role of oceanographic factors in driving genetic differentiation. Results will add a significant contribution to knowledge of the role of oceanographic features on the divergence of marine taxa in the southern tropics.

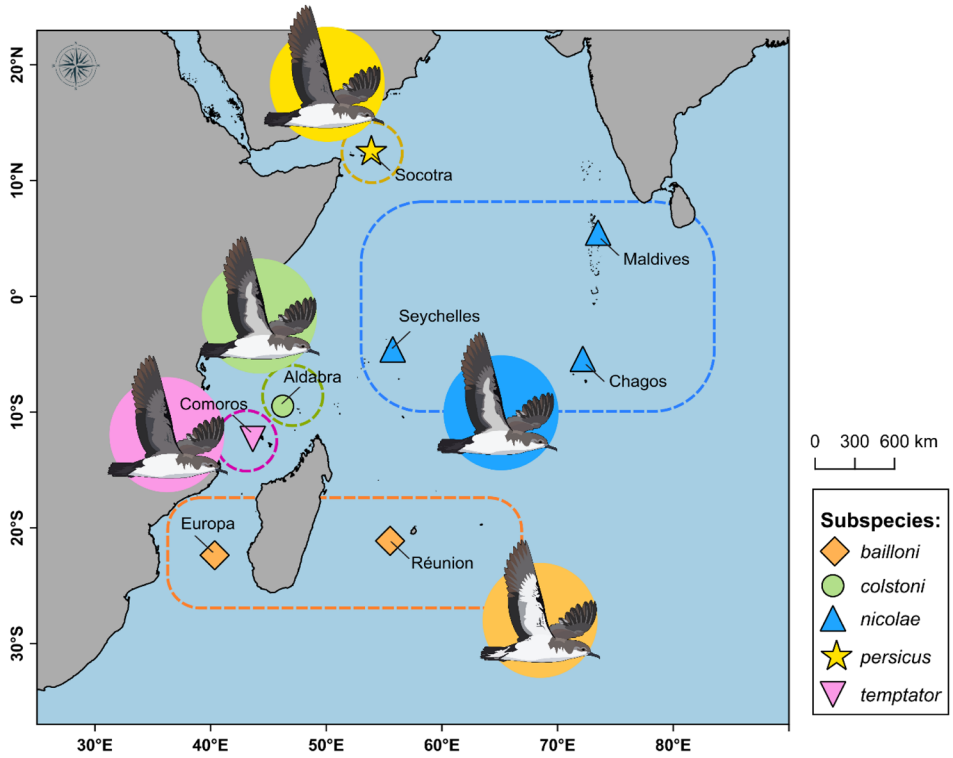

FIGURE 1 | Distribution of the tropical shearwater (*Puffinus bailloni*) across the Indian Ocean. Dashed lines indicate the range of the different forms of the tropical shearwater regarded as subspecies: *Puffinus bailloni bailloni* from Réunion and Europa, *P. b. nicolae* from the inner Seychelles, Chagos and Maldives, *P. b. colstoni* from Aldabra, *P. b. persicus* from Socotra and *P. b. temptator* from Comoros. Taxa have been described based on body size and undertail coverts colouration only (Bretagnolle et al. 2000; Shirihai and Christie 1996). The colour coding of the colonies is consistently used throughout the manuscript. Seabird illustrations modified from Trevail et al. (2023).

2 | Materials and Methods

2.1 | Study Area and Specimen Collection

The present study focuses on three described tropical shearwater subspecies in the WIO breeding in Europa and Réunion (*P. b. bailloni*), Aldabra Atoll (*P. b. colstoni*), and in Aride and Cousin (Seychelles archipelago) and South Brother (Chagos archipelago, hereafter called Chagos) Islands (*P. b. nicolae*; Figure 1). The distance between colonies ranged from 12 km (Aride and Cousin) to 3820 km (Europa and Chagos). The largest colony is located at Aride (> 20,000 breeding pairs; Calabrese 2016), while the smallest study colonies are in Europa and Aldabra (< 100 breeding pairs each) (Table 1).

Biological samples were collected between 2005 and 2023. Except for Europa and Réunion, all birds were captured by hand at their nesting burrows. Birds from Europa were caught in mist nets at their display areas—where they are known to perform aerial pursuits and actively vocalise—at dusk or at dawn. Blood samples (approx. 0.5 mL) were collected from live birds by venipuncture of the medial metatarsal or basilic veins. Additionally, muscle samples (approx. 10 g) were collected from dead birds fatally injured as a consequence of light pollution on Réunion (Chevillon et al. 2022). Samples were stored in 70% ethanol until DNA extraction. Genomic DNA was extracted using the NucleoSpin Tissue kit (Macherey-Nagel) for the extractions performed by GenoScreen (Lille, France) and the QIAmp Blood and Tissue kit (Qiagen) for the extractions performed by the authors. Individuals were sexed by PCR amplification with the primer pair 2550F and 2718R, following the method described in Fridolfsson and Ellegren (1999), which relies on the size differences between the CHD1W and CHD1Z introns located on the W and Z sex chromosomes, respectively. Results were visualised via gel electrophoresis (1.5% agarose gel).

2.2 | SNP Genotyping and Filtering

A total of 90 tropical shearwater samples (i.e., 17 from Europa Island, 18 from Réunion, 14 from Aldabra, 15 from Aride, 15 from Cousin and 11 from Chagos) were available for DArTseq, a sequencing method for reduced genomic representation (Dorey et al. 2021; Georges et al. 2018), similar to double-digest restriction-associated DNA sequencing (ddRAD) (Davey and Blaxter 2010). The method employs a combination of restriction enzymes and implicit fragment size selection to generate a large number of DNA fragments truncated to 69 bp. These fragments are then sequenced and analysed to identify single nucleotide polymorphisms (SNPs) and presence/absence variants. Genomic DNA (~20–40 ng/μL) was digested with the restriction enzymes *PstI* and *SphI* at the Diversity Arrays Technology Facility (DarT Pty Ltd., Canberra), following the routines described by Georges et al. (2018). Sequence data was processed using a proprietary DArT analytical pipeline (Jaccoud et al. 2001; Kilian et al. 2012; Sansaloni et al. 2011) and mapped to the genome assembly of the congener *Puffinus mauretanicus* (ASM2333356v1) (Cuevas-Caball et al. 2022), assuming a sequence identity ≥ 80% (*E*-value:

TABLE 1 | Characteristics of the colonies of tropical shearwaters (*Puffinus bailloni*) in the Indian Ocean.

					Morphology			*N*		
Colony	**Lat (°S)**	**Long (°E)**	**Subspecies**	**# Of breeding pairs**	**Under tail**	**Under wings**	**Body size**	**# Samples sequenced**	**# Samples post-filtering**	**# Biometrics**
Comoros (Mohéli)	−12.33	43.65	*temptator*	< 100	Brown	Brown	Larger	—	—	—
Europa	−22.36	40.36	*bailloni*	< 50	White	White	Larger	17	14	74
Réunion	−21.12	55.52	*bailloni*	> 5000	White	White	Larger	18	16	63
Seychelles (Aldabra)	−9.48	46.24	*colstoni*	< 75	Brown	White	Medium	14	13	30
Seychelles (Aride)	−4.21	55.67	*nicolae*	> 20,000	Brown	White	Smaller	15	12	—
Seychelles (Cousin)	−4.33	55.66	*nicolae*	> 5000	Brown	White	Smaller	15	9	20
Seychelles (Cousine)	−4.35	55.57	*nicolae*	> 250	Brown	White	Smaller	—	—	—
Seychelles (Ile aux Récifs)	−4.58	55.75	*nicolae*	> 500	Brown	White	Smaller	—	—	—
Chagos (South Brother)	−6.17	71.54	*nicolae*	400	Brown	White	Smaller	11	10	10
Chagos (North Brother)	−6.08	71.3	*nicolae*	1200	Brown	White	Smaller	—	—	—
Chagos (Nelson's island)	−5.4	72.18	*nicolae*	5	Brown	White	Smaller	—	—	—
Chagos (Peros Banhos)	−5.2	71.52	*nicolae*	< 20	Brown	White	Smaller	—	—	—
Maldives (Kurehdhoo)	5.5	73.5	*nicolae*	Unknown	Brown	White	Smaller	—	—	—
Yemen (Socotra)	12.5	53.9	*persicus*	Unknown	Brown	Brown	Larger	—	—	—
								90	74	197

Abbreviation: *N*, Number of individuals.

5×10^{-7}). Additional quality filtering was performed by the authors in R v4.3.3 using the "dartRverse" package (Gruber et al. 2018; Mijangos et al. 2022) at both individual and loci levels. Individuals with a call rate (i.e., proportion of scored loci per individual) below 80% and/or with a high individual heterozygosity (Figure S1) were excluded from the dataset. Further filtering was applied to single nucleotide polymorphisms (SNPs) that had a call rate (i.e., proportion of scored individuals for a locus) lower than 90%, a genotyping reproducibility (i.e., average repeatability of alleles at a locus across replicates) below 100%, and an average read depth lower than 10× (Bagley et al. 2017; Garot et al. 2019; Meier et al. 2017; Nadachowska-Brzyska et al. 2016) and higher than 80× sequences per locus. Multiple-linked SNPs per sequenced tag were also filtered out to control for short-distance linkage disequilibrium (i.e., only one SNP per fragment was retained at random). Sex-linked markers were excluded from the dataset, assuming that if an SNP is present only on the Z chromosome but not on the W chromosome, all females will be heterozygous (ZW) at that locus and all males homozygous (ZZ). OutFLANK (Whitlock and Lotterhos 2015) was additionally used to identify outlier SNPs, considering a cut-off value of 0.01. All the cut-off values were defined after plotting the data (Devloo-Delva et al. 2019; Feutry et al. 2020). To ensure that duplicated individuals (i.e., birds sampled more than once) and closely related individuals were not included in the downstream analyses, the software *ngsRelateV2* (Hanghøj et al. 2019) implemented in the *ANGSD* framework (Korneliussen et al. 2014) was used to investigate the presence of clones and first-degree relatives (i.e., parent-offspring and full siblings) in our dataset. Relatedness analyses followed the routines thoroughly described in Teixeira et al. (2024). Data for the outgroup taxa was retrieved from the DArT dataset previously generated for the Mascarene petrel (*Pseudobulweria aterrima*) (Teixeira et al. 2024), a member of the same bird family (Procellariidae). Collection information for all specimens is provided in Table S1.

2.3 | Identification of Diagnosable Operational Taxonomic Units

Due to the current taxonomic uncertainty of the tropical shearwater, two phylogenetic approaches were used to evaluate whether the described subspecies (*bailloni, nicolae* and *colstoni*) form well-supported monophyletic lineages. First, a genetic distance tree was constructed based on the UPGMA algorithm. Genetic distance between samples was estimated using the *bitwise.dist* function, which calculates the proportion of allelic differences between pairs of samples in a genlight object. Analysis was performed using the *aboot* function implemented in the "poppr" R package (Kamvar et al. 2014) considering 1000 bootstraps. The resulting tree was plotted using the *plot.phylo* function available in the "ape" (Paradis et al. 2019) R package. The second approach used SVDquartets analysis (Chifman and Kubatko 2014) implemented in PAUP* v 4.0a. SVDquartets evaluates unlinked multilocus data for subsets of four taxa (quartets) and assigns a score to each of the three possible unrooted topologies for each quartet. The topology with the lowest score is selected as the best supported, and the final set of quartets is combined to estimate the species tree (Chifman and Kubatko 2014). Analyses were performed considering 100,000 randomly selected quartets and 10,000 standard bootstraps. Phylogenies were estimated considering all individuals as a separate lineage (taxpartition = none) and assigning all individuals to their sampling location (taxpartition = pops). The SVDquartets trees were visualised using FigTree v.1.4.4 (Rambaut and Drummond 2008). The resulting trees were rooted to the Mascarene petrel.

Complementary fixed difference analysis was performed using the *gl.fixed.diff* and *gl.collapse* functions in the "dartRverse" R package (Gruber et al. 2018; Mijangos et al. 2022) to help define diagnosable operational taxonomic units (OTUs). A fixed difference between two populations at a locus occurs when the populations do not share any alleles at that locus. Thus, the accumulation of fixed differences between two populations serves as a robust indication of the absence of gene flow (Georges et al. 2018). Given that less than 10 birds were available for Cousin after stringent quality filtering, this colony was amalgamated with Aride (12 km apart) before analyses in line with the population structure and phylogenomic reconstructions (see results). Fixed differences were calculated for each pairwise colony comparison, and colonies that exhibited no fixed differences were aggregated. This process was repeated until there was no further reduction following Georges et al. (2018). To avoid conflation of true fixed differences with false positives due to finite sample sizes, the fixed differences between final OTUs were tested for statistical significance using 1000 simulations.

2.4 | Population Genetic Structure and Admixture

As an initial representation of the genetic structure of the tropical shearwater, a principal component analysis (PCA) was performed using the *gl.pcoa* and *gl.pcoa.plot* functions implemented in the "dartRverse" R package (Gruber et al. 2018; Mijangos et al. 2022). A complementary dimensionality-reduction clustering was performed with discriminant analysis of principal components (DAPC), using the *find.clusters* function in the "Adegenet" v2.1.7 R package (Jombart et al. 2010). The *K*-means clustering algorithm was employed to identify the optimal number of genetic clusters (K) from $K = 1$ to $K = 6$, and the best K value was selected based on the lowest Bayesian Information Criterion (BIC) following Jombart et al. (2010). To avoid overfitting of discriminant functions, a-score optimisation was used to evaluate the optimal number of principal components to retain in the analysis (Bagley et al. 2017). Population structure was further examined using the Bayesian assignment approach implemented in Structure v2.3.4 (Pritchard et al. 2000), with the number of genetic clusters set from 1 to 6. Analyses were performed assuming an admixture model, correlated allele frequencies and no population priors, for a run length of 100,000 iterations and 10,000 burn-in. A total of 10 independent runs were performed for each K, and ancestry proportions across runs were averaged using Clumpak (Kopelman et al. 2015). The optimal number of genetic clusters was determined based on the log likelihood given K ($L(K)$) (Pritchard et al. 2000) and the second-order rate of change of mean log likelihood (ΔK), following Evanno's method (Evanno et al. 2005), using Clumpak (Kopelman et al. 2015). Since the ΔK method is known to detect the uppermost level of genetic structure (Evanno et al. 2005), the analyses were repeated using a hierarchical approach to uncover underlying population substructure. Likewise, the PCA and DAPC analyses were also repeated for the subsp. *nicolae*

and *colstoni* dataset. Genetic differentiation among breeding colonies was additionally estimated using Wright's *F*-statistics F_{ST} (Weir and Cockerham 1984). Analyses were performed using the "StAMPP" (Pembleton et al. 2013) R package, and significance was tested using 10,000 bootstraps. To test whether sites with a low minor allele frequency (MAF; frequency of the second most common allele in a population) affected the population structure interpretations, all analyses were performed considering all sites (27,273 SNPs) and filtering out sites with an MAF lower than 0.01 (22,752 SNPs). As both datasets led to the same interpretations, we report only the results from runs assuming the most comprehensive dataset (i.e., the dataset with MAF sites).

Evolutionary relationships among colonies of tropical shearwater were further explored using TreeMix v.1.13 (Pickrell and Pritchard 2012). This approach assumes that extant populations primarily result from splits between ancestral populations and that differences in allele frequencies between groups are a result of genetic drift since the splits. First, the software estimates a bifurcating graph with the minimum amount of genetic drift between populations using a maximum likelihood approach. Subsequently, TreeMix identifies populations that poorly fit the tree model and models migration events involving such populations (Pickrell and Pritchard 2012). The SNPs dataset was initially converted to a TreeMix input file using the *gl2treemix* function in the "dartRverse" R package (Gruber et al. 2018; Mijangos et al. 2022). Analyses were then conducted considering a number of migration events (m) between 1 and 5, a total of 50 replicates for each value of m, and using the Mascarene petrel as an outgroup to root the tree. The optimal number of migration events was selected according to (i) the distribution of the mean log likelihoods across m, (ii) the *ad hoc* log likelihood statistic based on the second-order rate of change (Δm) implemented on the "OptM" R package (Fitak 2021) and (iii) the inspection of the residual plots of covariance (Pickrell and Pritchard 2012). This analysis was complemented by the three- and four-population tests ($f3$ and $f4$, respectively) to corroborate possible admixture events. The $f3$- and $f4$-statistics were calculated for each unique combination of (A; B, C) (Reich et al. 2009) and (A, B; C, Outgroup) (Keinan et al. 2007) of the six tropical shearwater colonies using TreeMix v.1.13. Significance was determined using the default Z-value.

2.5 | Identification of Potential Drivers of Genetic Structure

Two distinct approaches were used to identify the potential drivers of genetic structure in the tropical shearwater in the WIO. First, the Estimated Effective Migration Surfaces (EEMS) (Petkova et al. 2015) were used to identify and visualise potential barriers to gene flow across the landscape. EEMS identifies geographic areas where genetic dissimilarity is greater than expected under isolation-by-distance using genetic data and geo-referenced samples, without requiring environmental or topographic data. The filtered SNPs dataset was firstly converted to a VCF (Variant Call Format) file using the *gl2vcf* function implemented in the "dartRverse" R package (Gruber et al. 2018; Mijangos et al. 2022). PLINK v.2.00a4 (Purcell et al. 2007) was subsequently used to convert VCF format to binary bed files. The EEMS software (Petkova et al. 2015) was then executed based on the pairwise genetic dissimilarity matrix (i.e., average number of allelic differences between each pair of individuals across all genotyped loci) (Pedersen et al. 2018) estimated from the plink files with the *bed2diFFs* programme, for an MCMC length of 25,000,000, a burn-in of 1,000,000, considering a deme grid of 500. Resultant effective migration (m) surfaces were plotted using EEMS' *reemsplots2*.

Second, data from six oceanographic variables, including five dynamic variables and one static variable (bathymetry), were compiled to evaluate whether isolation-by-distance (IBD) or isolation-by-environment (IBE) explains the marked genetic differentiation observed for the tropical shearwater in the WIO. SST (°C), sea surface salinity, chlorophyll a concentration ($mg\,m^{-3}$), wind speed ($m\,s^{-1}$) and current speed ($m\,s^{-1}$) were extracted from the E.U. Copernicus Marine Service Information Website (https://data.marine.copernicus.eu/products), while bathymetry (m) was compiled from the ETOPO Global Relief Model (https://www.ncei.noaa.gov/products/etopo-global-relief-model) (Table S2). For each of the dynamic oceanographic variables, except wind speed, the monthly means and standard deviations were computed over a span of 20 years (2000–2020). For wind speed, a 12-year dataset (2008–2020) was used. Values were extracted within a 300 km radius circle around each breeding colony, assuming this encompasses most of the foraging range of breeding birds at each site (MLC personal comm.; but see Calabrese 2016). Mean values (±SD) were then calculated for each oceanographic variable at each site. Correlations between the pairwise genetic distances (F_{ST}) and geographical or environmental distances were tested using Mantel tests (Mantel 1967) with 10,000 random permutations of the variables, using the "ade4" R package (Dray et al. 2007). Geographical distances were calculated as great circle distance (i.e., the shortest distance between two points on the surface of a sphere; WGS84 ellipsoid) using the "sp" R package (Pebesma 2012). Environmental distances were calculated as the pairwise difference in each variable among colonies. Partial Mantel tests based on Pearson's product–moment correlation was additionally used to estimate the correlation of the genetic and environmental distances while controlling for the potential confounding effect of geographical distances. Partial Mantel tests were performed using the "vegan" R package (Oksanen 2013).

2.6 | Morphometric Analyses

Five morphometric measurements were obtained from 197 breeding adults (i.e., 74 from Europa, 63 from Réunion, 30 from Aldabra, 20 from Cousin and 10 from Chagos) from all colonies excluding Aride: wing length (i.e., flattened wing chord from the carpal joint to the tip of the longest primary), tarsus length (i.e., the portion of the leg measured from the notch at the back of the intertarsal joint to the distal edge of the large complete scale at the front of the foot, just before the toes diverge), bill length (i.e., exposed culmen), bill depth (i.e., bill depth at gonys) and hook length (i.e., distance from the base to the tip of the hooked part of the bill). The colour of undertail coverts was also recorded for each individual, whereas all the birds studied exhibited white underwing coverts. Global patterns in morphological differences among birds and colonies were investigated using PCA followed by a multivariate

analysis of variance (MANOVA) using the "ade4" R package (Dray et al. 2007). The three principal component scores from the PCA (i.e., PC1–PC3) were the variables in the MANOVA from which a Wilks' lambda (Huberty and Olejnik 2007) was derived following 10,000 permutations. Corrected pairwise permutation MANOVAs were performed to compare morphometric measurements among pairwise colonies, using the "RVAideMemoire" R package (Herve 2020). A Factor Analysis for Mixed Data (FADM) was additionally used to integrate a categorical variable (i.e., colour of undertail coverts) (Pages 2004). Morphometric measurements of all individuals are available in Table S3.

3 | Results

3.1 | Genomic Dataset

A total of 78,676 polymorphic SNP loci were called for 90 tropical shearwater individuals from six colonies across the WIO. After stringent filtering, 27,273 SNPs and 74 unrelated birds (i.e., 14 from Europa, 16 from Réunion, 13 from Aldabra, 12 from Aride, 9 from Cousin and 10 from Chagos) were retained in the dataset and considered in all downstream analyses (ingroup dataset). A subset of this dataset was also used to investigate the underlying population substructure (subsp. *nicolae* and *colstoni*: 18,065 SNPs across 44 birds; subsp. *bailloni*: 16,105 SNPs across 30 birds). Detailed information on the number of SNPs and individuals retained at each step of the quality filtering is given in Table S4. The tropical shearwater dataset was additionally co-analysed with generated Mascarene petrel sequences (Teixeira et al. 2024) to include a suitable outgroup for complementary analyses (outgroup dataset). A total of 16,152 SNP loci from 74 ingroup individuals and 20 outgroup individuals were considered for the population-based analyses (i.e., TreeMix), and 27,850 SNP loci from 74 ingroup individuals and two outgroup individuals for the phylogenomic reconstructions. See Table S5 for details about individuals used for each analysis.

3.2 | Identification of Diagnosable Operational Taxonomic Units

Two diagnosable OTUs emerged from the fixed difference analysis. Europa and Réunion, in the southern limit of the species distribution, formed one diagnosable OTU (corresponding to subsp. *bailloni*), while Aride, Cousin, Aldabra and Chagos in the northern range of the tropical shearwater formed a second diagnosable OTU (corresponding to subsp. *nicolae* and *colstoni*), with a total of 11 absolute fixed differences ($p < 0.0001$). Aldabra (subsp. *colstoni*) shared alleles at all loci with the other northern populations (Table S6).

Phylogenetic reconstructions using the UPGMA (Figure 2a) and SVDquartets (Figure 2b) consistently inferred the same tree topology, with both methods showing a deep split between the northern and southern colonies, corresponding to the two identified OTUs. In the southern WIO, both analyses supported the further split of Europa and Réunion, with each colony forming a well-supported monophyletic clade (100% bootstrap support). The phylogenies also confirmed that individuals from Aride and Cousin formed a strongly supported mixed clade (100% bootstrap support) and positioned the individuals from Chagos in a poorly supported clade sister to Aride-Cousin. Both methods consistently inferred a monophyletic clade for Aldabra (subsp. *colstoni*), but this clade was only well supported by the distance-based tree (i.e., UPGMA; 93% bootstrap support).

3.3 | Population Structure and Role of the Environmental Factors

Patterns of population structuring among tropical shearwater colonies were not consistent with the currently recognised subspecies nomenclature. The subsp. *bailloni* was consistently separated from the subsp. *nicolae* and *colstoni* in all the performed analyses. The PCA analysis performed on the entire dataset clearly discriminated the northern from the southern colonies along PC1 (21.8% of variation; Figure 3a). While Europa and Réunion were well-separated along PC2 (3.8% of variation), an overlap was observed among the northern colonies, including Aldabra (subsp. *colstoni*). However, the PCA analysis carried out only on the northern colonies separated the birds from Aldabra, Chagos and Cousin-Aride along the PC1 (4.4% of variance; Figure S2). The DAPC analyses suggested either $K = 2$ or $K = 3$ as the optimal number of genetic clusters considering the entire dataset (Figure S3a). For $K = 2$, individuals from the subsp. *nicolae* and *colstoni* were assigned to one genetic cluster, whereas birds from the subsp. *bailloni* were assigned to a second cluster. At $K = 3$, birds from Aldabra (subsp. *colstoni*) were not assigned to a third group (Figure S4). Instead, the birds from Europa and Réunion were further divided into two distinct groups, mirroring the PCA findings. Structure analyses for the entire dataset suggested $K = 2$ as the optimal number of genetic clusters (Figure S5a) and assigned birds from subsp. *nicolae* and *colstoni* and subsp. *bailloni* to different clusters with a membership coefficient of 100% for all the birds (Figure 3c). Hierarchical structure analyses confirmed the existence of a marked population substructure between Europa and Réunion, whereas the northern colonies exhibited a subtle population differentiation (Figure 3c; Figures S6–S8). While the analyses performed only for the subsp. *nicolae* and *colstoni* failed to assign the birds from Aldabra to a distinct group at $K = 2$, these birds were assigned to their own genetic cluster when considering $K = 3$ (Figure S8). Birds from Aride and Cousin clustered together across all K solutions. Pairwise F_{ST} values were highest when comparing northern vs. southern colonies ($F_{ST} = 0.23–0.28$; $p < 0.0001$), whereas comparisons involving the northern populations showed substantially lower F_{ST} values ($F_{ST} = 0.01–0.03$; $p < 0.0001$ except for Aride-Cousin), including Aldabra. A moderate genetic differentiation was observed between Réunion and Europa ($F_{ST} = 0.09$; $p < 0.0001$) (Figure 3b).

One of the most striking results was the complete absence of admixture between the subsp. *nicolae-colstoni* and subsp. *bailloni*, suggesting long-term isolation among these taxa. The patterns of population divergence inferred with TreeMix were consistent with the phylogenetic and population clustering results (Figure S9). Analyses suggested a model without migration events (Figure S10), which was corroborated by the three- and four-population tests. Results showed that none of the $f3$-statistics were significantly negative for all unique combinations of the six tropical shearwater colonies, and none of the $f4$-statistics were significantly different from zero, suggesting that the tropical shearwater colonies are not the result of historical admixture events (Tables S7 and S8).

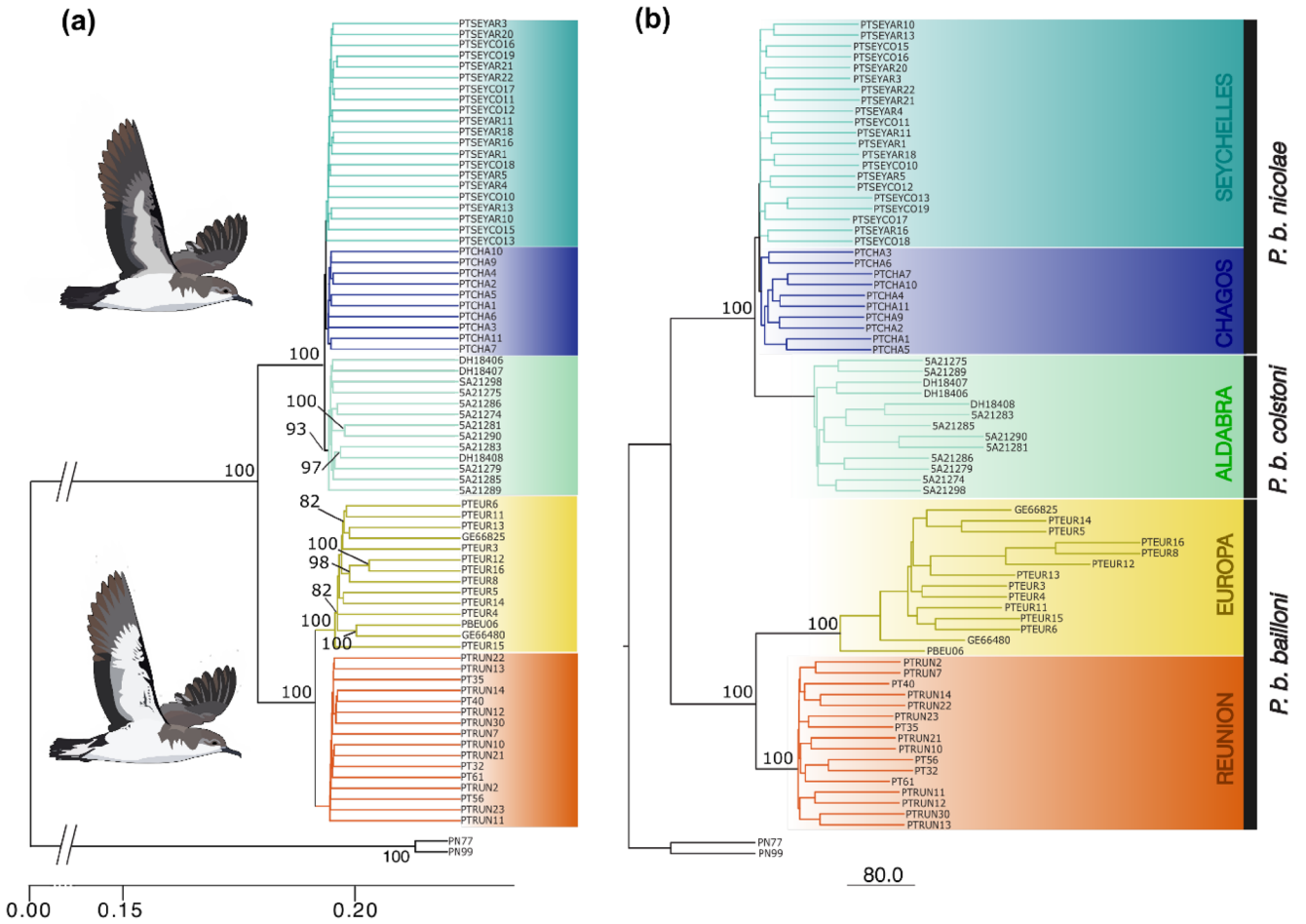

FIGURE 2 | Phylogenetic relationships of the tropical shearwater (*Puffinus bailloni*). (a) Genetic distance tree based on the UPGMA algorithm. Branch length represents the number of allelic differences between samples. Bootstrap values were estimated based on 1000 bootstraps, and (b) a phylogenetic tree generated with SVDquartets. Analyses were performed considering 100,000 randomly selected quartets and 10,000 bootstraps. For (a) and (b), node labels show bootstrap support ≥80%. Both trees included the 74 tropical shearwater (*Puffinus bailloni*) individuals that passed the quality filtering. The trees were rooted to the Mascarene petrel (*Pseudobulweria aterrima*).

Mantel tests showed a significant correlation between genetic distances and SST ($r = 0.853$; $p < 0.05$), latitude ($r = 0.937$; $p < 0.05$) and geographic distances ($r = 0.527$; $p < 0.05$). SST remained significant after controlling for the effect of geographical distances with a partial Mantel test, but not after controlling for the effect of latitude (Figure 4b). The EEMS detected an area of reduced migration (at least tenfold lower than the average migration) between the latitudes 10° and 20° S, coinciding with the Mozambique Channel and the continental land barrier created by Madagascar (Figure 4a). These findings supported a lack of gene flow between birds from subsp. *nicolae-colstoni* and subsp. *bailloni* and are consistent with the population structure and phylogenomic results.

3.4 | Morphometric Analyses

The first axis of the PCA based on five morphometric measurements (explaining 56% of the total variance) separated the individuals following a decreasing gradient in body size from the subspecies *bailloni* to the subspecies *nicolae* (Figure S11). The colonies differed significantly in morphology when tested by parametric MANOVA (Wilks' lambda = 0.20; df = 20; $p < 2.2 \times 10^{-16}$). All colonies were significantly different ($p < 0.05$), with the birds from Europa and Réunion being the largest and those from Cousin being the smallest, except for birds from Aldabra and Chagos, which share the same intermediate body size ($p = 0.135$). FADM analysis, including the qualitative character, yielded the same results.

4 | Discussion

This study integrated genomic, morphometric and oceanographic data to examine relationships among closely related seabird taxa breeding in the tropical Indian Ocean. The analyses of the tropical shearwater yield two key findings. First, our data did not support the current species nomenclature, highlighting the need for an accurate revision of the shearwater's taxonomy. Second, our study revealed high genetic differentiation within an ocean basin in a highly vagile taxon, likely driven by a combination of oceanographic, topographic and life-history factors. The taxonomic, biogeographic and conservation implications of these findings are discussed below.

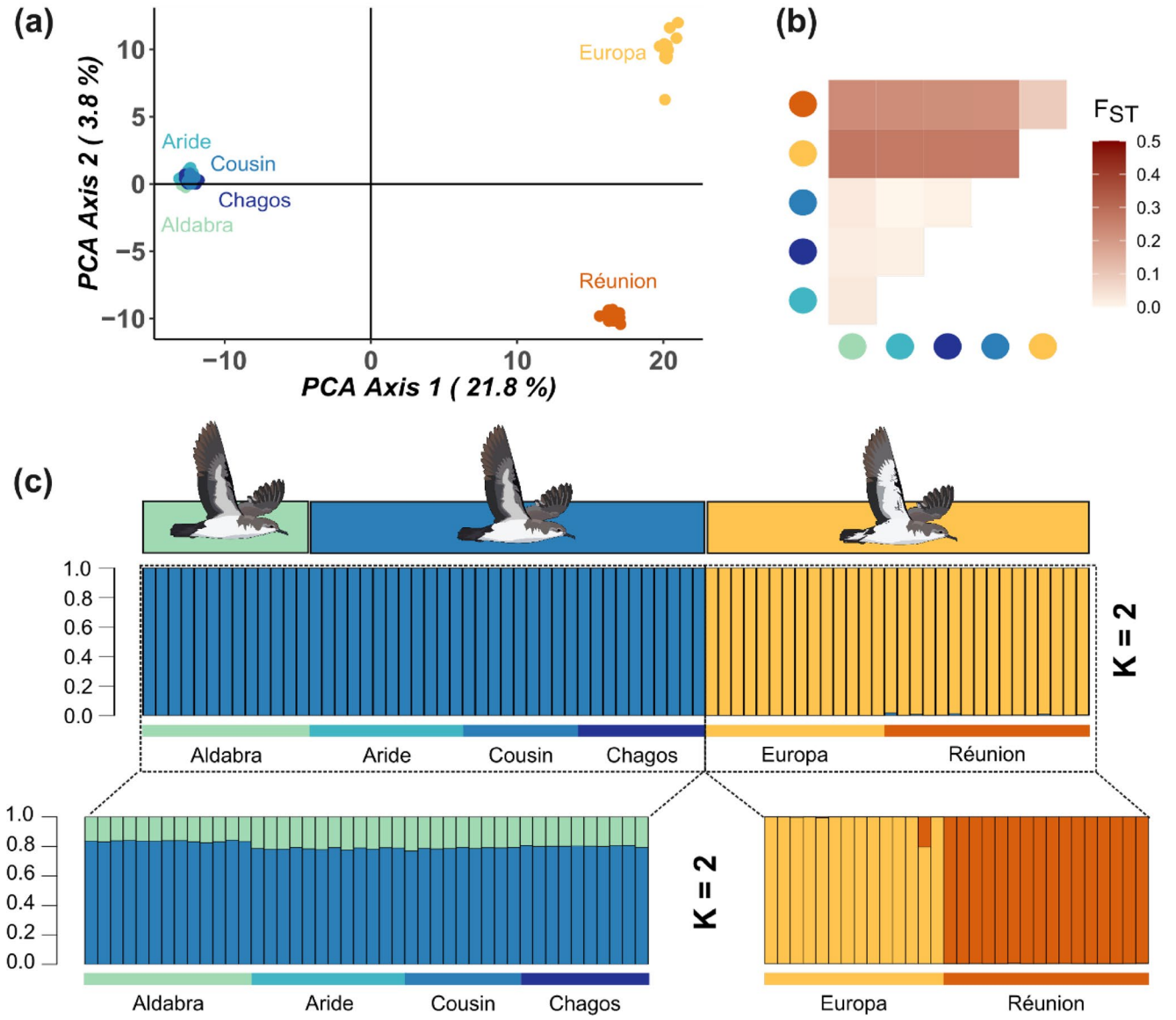

FIGURE 3 | Population genetic structure of 74 tropical shearwaters (*Puffinus bailloni*) across the Western Indian Ocean. (a) Principal coordinates analysis (PCA). Each circle represents a sample, and colours indicate the different colonies; (b) heatmap of pairwise F_{ST} estimates between pairs of colonies. Shades of red highlight moderate to higher genetic differentiation, while whitish tones indicate no genetic differentiation among colonies; (c) results from structure analyses assuming two genetic clusters ($K = 2$) for the entire dataset ($n = 74$ birds), the subsp. *bailloni* ($n = 30$ birds; yellow cluster) and the subsp. *nicolae* and *colstoni* ($n = 44$ birds; blue cluster). For the latter analyses, birds from Aldabra (subsp. *colstoni*) were not assigned to a distinct group at $K = 2$. Horizontal bars on top represent the three proposed tropical shearwater subspecies within our sampling range. Each vertical bar represents one individual, and each colour a genetic cluster. Values were obtained by averaging the posterior probabilities over 10 independent runs. Individuals were sorted following their geographic location. For plots showing different values of K, see Figures S6–S8.

4.1 | Insights Into *P. bailloni* Taxonomy

The tropical shearwater subsp. *colstoni* was proposed as a new taxon based on the premise that the birds from Aldabra were intermediate in size to those from Seychelles (*nicolae*) and Réunion (*bailloni*) (Shirihai and Christie 1996). However, the sample sizes available for the Aldabra colony were limited ($n = 5$ in Shirihai and Christie (1996) and $n = 8$ in Bretagnolle et al. (2000)), precluding robust statistical analyses and this hypothesis has never been tested using datasets other than morphometrics (but see Austin et al. 2004). Our morphometric analyses, using a comprehensive dataset, confirmed that birds from Aldabra were intermediate in size between *nicolae* and *bailloni* (Figure S11). Although the genetic analyses performed in this study revealed subtle genetic differences among all the studied populations except Aride-Cousin, the levels of genetic differentiation do not correspond to the current subspecies nomenclature. The largest genetic differences undoubtedly occurred between the northern (*nicolae* and *colstoni*) and the southern (*bailloni*) colonies. This is supported by the phylogenomic reconstructions (i.e., only two statistically well-supported OTUs), the clustering analyses for the entire dataset (which failed to assign the birds from Aldabra to a distinct genetic cluster, even when assuming

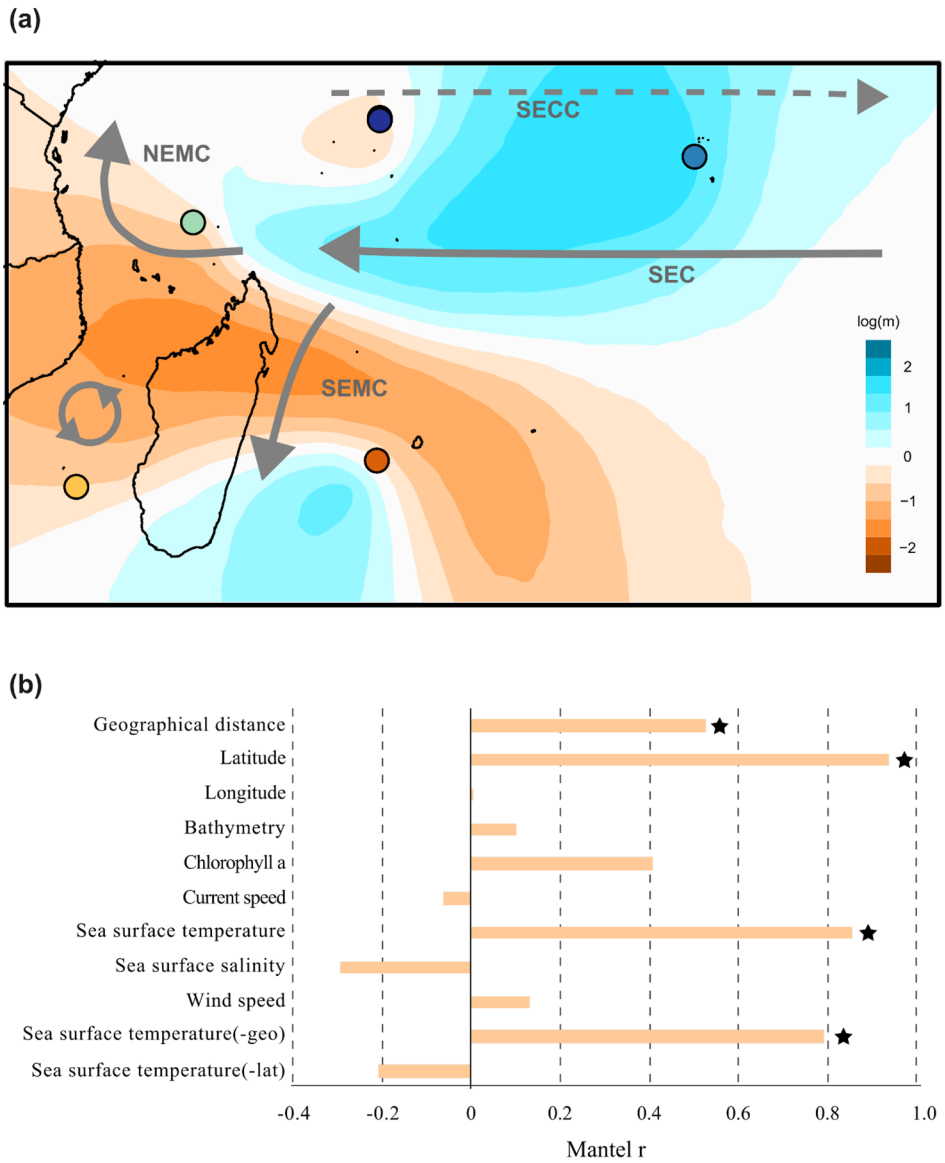

FIGURE 4 | Identification of potential drivers of genetic structure in the tropical shearwater (*Puffinus bailloni*) across the Western Indian Ocean. (a) Estimated effective migration surface (EEMS) results considering 500 demes. The effective migration rate (m) is shown on a $\log_{10}$ scale, with a colour gradient ranging from orange to blue. Shades of blue highlight higher migration rates than the global mean, while oranges indicate reduced migration. Colonies are indicated using a colour code consistent with those in Figure 1. Arrows represent the major ocean surface current patterns in the Indian Ocean during the austral summer (northeast) monsoon, while dashed lines represent currents occurring only during the austral winter monsoon. Data compiled from Schott et al. (2009). SEMC = South East Madagascar Current; NEMC = North East Madagascar Current; SEC: South Equatorial Current; SECC = South Equatorial Countercurrent; (b) Results of the Mantel tests for the correlation (r) between pairwise genetic distance (F_{ST}) and geographic distances (Isolation-by-distance), latitude, longitude, or relevant oceanographic variables (Isolation-by-environment). Mantel tests were additionally performed to assess the significance of the correlations between F_{ST} and sea surface temperature (SST), while controlling for geographical distance [SST (−geo)] and latitude [SST (−lat)]. Significant correlations are indicated with * ($p < 0.05$).

$K = 3$), and by the F_{ST} estimates (which suggest no genetic differentiation between the birds sampled in the inner Seychelles and Chagos and those from Aldabra). Following this major split, our results consistently separated the birds from Europa and Réunion (*bailloni*), suggesting they are more genetically differentiated from each other than birds from subsp. *nicolae* and *colstoni*. Finally, at a finer scale, our data revealed that the four northern populations form three genetic clusters, with birds from Aride and Cousin clustering together. The distinctness of the *bailloni* taxa from *nicolae-colstoni* aligns with recent phylogenetic studies on shearwaters based on mtDNA, nuclear markers (Torres et al. 2021) and genomic datasets (Ferrer Obiol et al. 2021), which confirmed the differentiation of birds from Réunion and the inner Seychelles. Likewise, the genetic similarity between the birds from Aldabra and the inner Seychelles is corroborated by an early phylogenetic inference of the *Puffinus* shearwaters based on a single mtDNA locus (Austin et al. 2004).

Although our study does not allow a complete review of the taxonomic relationships of the tropical shearwater, the genomic, phenotypic and spatial data presented here support

the existence of two distinct taxa within our sampled populations, corresponding to the *nicolae-colstoni* and *bailloni* forms. The complete lack of gene flow between these taxa, as demonstrated by both recent and historical admixture analyses, raises the question of whether they should be classified as distinct subspecies or separate species, though this remains to be investigated. If divergence occurred at the species level, the *bailloni* form may consist of two cryptic subspecies breeding on Europa and Réunion. In contrast, our study does not provide support for *colstoni* as a different taxon. We hypothesise that the Aldabra colony resulted from relatively recent dispersal from the other colonies of the central Seychelles and, therefore, there has not been enough time for significant genetic differentiation to accumulate, even if birds from Aldabra and the central Seychelles are no longer interbreeding. Indeed, Aldabra experienced repeated inundations during glacial periods, due to dramatic rises in sea level, with the most recent submergence occurring around 125,000 years ago (Agnarsson and Kuntner 2012; Warren et al. 2003). Consequently, the colonisation of the present-day fauna likely took place more recently (Austin et al. 2004; Warren et al. 2003). Taxonomic uncertainties in the genus *Puffinus* are not surprising, with a recent study on the North Atlantic and Mediterranean *Puffinus* (Obiol et al. 2023) suggesting the elevation of some taxa to species status, while others were reclassified as subspecies. In summary, our findings emphasise the need for a comprehensive taxonomic review of the tropical shearwater. Further research including birds from the Maldives (subsp. *nicolae*), Comoros (subsp. *temptator*), Yemen (subsp. *persicus*) and the southern Pacific (subsp. *dichrous*) is essential to reassess the taxonomic classification of this species across its complete range and subsequently delineate effective conservation units.

4.2 | Connectivity and Barriers to Gene Flow in the WIO

The present study revealed remarkable levels of genetic differentiation within an ocean basin in a highly vagile species. The highest genetic differentiation was observed between the northern (Aldabra, Aride, Cousin and Chagos) and southern (Europa and Réunion) colonies, with the F_{ST} estimates, clustering analyses, TreeMix and the $f3$- and $f4$-statistics tests suggesting reproductive isolation. These findings have major biogeographic implications, given the absence of apparent physical barriers (except for Madagascar; see below) that might promote genetic differentiation within our study area. Several non-exclusive non-physical barriers to gene flow are therefore plausible. First, Mantel tests revealed significant positive correlations between genetic distances and SST, even after controlling for the effect of geographic distance, although SST was no longer significant after controlling for latitude due to the likely confounding nature of these variables. Indeed, sea temperature (Grémillet et al. 2008) has been proposed as the main driver of genetic differentiation of the *Puffinus* shearwater complex (Torres et al. 2021), as it is known to affect not only marine productivity, prey species abundance and diversity (Grémillet et al. 2008), but also seabird distribution, breeding phenology, adult survival and growth rate (Cruz-Flores et al. 2022; Dunn et al. 2024; McDuie et al. 2018). We note that the oceanographic data used for the IBE analyses cover only the past 20 years and may not reflect the conditions during the time of the population divergence. Second, the geographical distribution of the two genetic clusters, together with the putative barriers to gene flow identified by the EEMS, suggests that the South Equatorial Current (SEC) and South Equatorial Countercurrent (SECC) (see Schott and Mccreary 2001) as well as the Intertropical Convergence Zone (ITCZ), may limit the movement of the birds across the whole at-sea range of the species. The SEC and SECC are two major oceanic currents in the Indian Ocean that flow in opposite directions: while the SEC flows year-round westward along approximately 15° S, the SECC flows eastward along the equator during the winter monsoon (Figure 4a). The ITCZ is a significant meteorological system active in the tropics, where the trade winds of the Northern and Southern Hemispheres converge and rise into the stratosphere, leading to increased humidity and, consequently, high rainfall in the region (Carvalho et al. 2022). Although the highly colonial nature of the tropical shearwater populations may have resulted in a spatial sampling bias (but see Herman et al. (2022)), as the species tends to breed in a few larger colonies, the biogeographic importance of surface currents as oceanographic barriers has been shown in other marine realms. For instance, the Eastern Pacific Barrier, the Benguela Current off southwestern Africa and the Almeria–Oran Front, at the interface between the Atlantic Ocean and Mediterranean Sea, act as significant barriers to gene flow in other seabirds by aiding or hindering their flight at sea and by shaping the distribution of prey (Gómez-Díaz et al. 2009, 2012; Morris-Pocock et al. 2010, 2016; Steeves et al. 2003, 2005a, 2005b). Surface currents were also shown to impact gene flow in other marine organisms (Baums et al. 2012; El Ayari et al. 2019; Reid et al. 2016; Romero-Torres et al. 2018). However, except for Aride (Calabrese 2016; Trevail et al. 2023), the nonbreeding and at-sea foraging distributions of our study birds are currently unknown. Tracking studies would be necessary to investigate the subspecies movement at sea and further explore this hypothesis. Third, dispersal in other petrel species was shown to be limited by species-specific life-history traits, including natal philopatry and differences in breeding phenology (i.e., allochrony) (Danckwerts et al. 2021; Teixeira et al. 2024). If philopatry alone was driving the tropical shearwater differentiation, one could expect each colony to constitute its own genetic cluster. Instead, our results revealed substantial gene flow among all the northern colonies. We also observed overlapping breeding phenology among most of the studied colonies (Figure S12, but see below), suggesting that other factors may contribute to the genetic distinctiveness of the northern and southern colonies.

At the southern limit of the species distribution, the colonies on Europa and Réunion exhibited reduced connectivity and population substructure, in contrast to the higher levels of gene flow observed in the northern range. One potential explanation is that Madagascar limits the straight-line bird movement between Europa and Réunion, as seabirds generally avoid flying over land (but see Hailer et al. (2011)). Glaciers and other continental landmasses, including the Isthmus of Panama and the African continent, have been cited as the major barrier to gene flow in seabirds (Friesen et al. 2007; Morris-Pocock et al. 2016; Silva et al. 2015; Steeves et al. 2003, 2005a, 2005b). It is also worth noting that half of Europa's breeding period occurs when only a few pairs are breeding on Réunion Island

(Figure S12), suggesting that the isolation observed between these two populations may be partly explained by temporal reproductive isolation, as has been shown in other petrels (Garrett et al. 2020; Monteiro and Furness 1998). However, the genetic isolation of Europa is not unique to the tropical shearwater. Similar genetic signatures have been observed in other marine taxa within the Mozambique Channel, including the white-tailed tropicbird (Humeau et al. 2020), the green sea turtle (Bourjea et al. 2007; Jensen et al. 2019), brooding corals (van der Ven et al. 2021) and mangrove trees (Triest et al. 2021). These findings suggest that the oceanic environment of the Mozambique Channel may constitute an additional ecological barrier for marine species, possibly influenced by the Comoros eddies generated by both anticyclonic and cyclonic eddies (reviewed in Lett et al. (2024)) and/or by differences in the SST, which is known to be coolest to the south of Madagascar (Figure S13). Multi-taxa comparative studies across the WIO are essential to assess the role of oceanographic factors in driving genetic differentiation and speciation in marine communities in the region.

4.3 | Conservation Implications

Defining effective management units is one of the most important contributions of population genetics to conservation (Humeau et al. 2020). Although the tropical shearwater is a widespread and relatively abundant seabird, colony size fluctuates substantially, with some colonies hosting more than 10,000 birds (e.g., Aride, Réunion; Calabrese 2016; Gineste et al. 2017), while others host fewer than 20 breeding pairs (e.g., various islands in the Chagos archipelago; Carr et al. 2021). Genetic evidence indicates that the tropical shearwater consists of (at least) two distinct taxa within the WIO. These findings have important conservation implications, since subspp. *bailloni* is far less abundant than subspp. *nicolae-colstoni*. Additionally, analyses uncovered a remarkable population structure within the subsp. *bailloni*, suggesting Europa and Réunion as two genetically differentiated populations. Among all the studied sites, Europa stands as the most vulnerable management unit, being one of the smallest known tropical shearwater colonies (< 50 breeding pairs) (Le Corre 2000). The population is further threatened by predation of eggs and chicks by invasive rats, which are abundant on the island (Le Corre 2000). Its small population size, coupled with spatial and genetic isolation, makes the colony particularly vulnerable to genetic bottlenecks and demographic stochasticity, all of which could erode genetic diversity and increase the risk of extinction (Frankham et al. 2002). We therefore recommend prioritising the Europa breeding colony in future conservation efforts, including the eradication of invasive predators within nesting areas (Le Corre 2000). Although our study suggests that the birds from Aldabra do not constitute an endemic seabird subspecies, this colony hosts few individuals, which are restricted to tiny islets within Aldabra's lagoon, where they are threatened by rat predation, suggesting that this colony requires further monitoring and conservation attention (NB personal comm.). It is worth noting that, although the tropical shearwater is not considered to be of conservation concern at the species level (BirdLife International; IUCN), some populations may warrant an uplisting to endangered status if the subspecies status were recognised. Altogether, our study underlines the need to firstly reassess the taxonomic and conservation status of the tropical shearwater before populations face the same fate as those from Mauritius, Rodrigues and potentially Madagascar (Mlíkovský 2005).

Author Contributions

Conceptualization: H.T., L.H., A.J., M.A.C.N., M.L.C. Data curation: H.T. Formal analysis: H.T., L.H. and A.J. Funding acquisition: M.A.C.N., M.L.C., A.J., L.H. Investigation: H.T. Methodology: H.T., L.H. Project administration: M.A.C.N., M.L.C. Resources: M.A.C.N., A.J., N.B., A.C., L.A., C.L., M.B., R.D., R.F., N.J.S., G.R., L.C., A.M.T., J.T., S.V., M.L.C. Visualization: H.T., L.H. Writing – original draft: H.T. Writing – review and editing: all authors.

Acknowledgements

This paper is an output of the Project CONNECTs, as part of the Bertarelli Programme in Marine Science (Project 822916). Our study also benefited from samples collected under the former FEDER Smac (2020–2022, N° RE0022954), FEDER "Pathogenes associés à la Faune Sauvage Océan Indien" (Programme Opérationnel de Coopération Territoriale 2007–2013; #31189) and the CNRS – INEE/TAAF (AAP Iles Eparses "OMABIO" project). Research permits to sample birds were granted by the Terres Australes et Antarctiques Françaises (TAAF) for the Europa Island colony. Collection of biological material in the Seychelles was approved by the Seychelles Bureau of Standards and the Seychelles Ministry of Agriculture, Climate Change and Environment. Fieldwork was also conducted with the approval of the Seychelles Islands Foundation (Aldabra), the Island Conservation Society (Aride) and Nature Seychelles (Cousin). Research was conducted in the Chagos Archipelago under research permit #0011SE22. Muscle samples from Réunion were collected from dead birds under the principles of the Research Centre on Biology of Bird Populations (PP 616 and banding authorisation of M.L.C.; CRPBO, National Museum of Natural History, Paris). Fieldwork was made possible thanks to the assistance of Erwan Lagadec, Maxime Amy, Christopher Jones, Merlène Saunier, Lucie Gauchet, Céline Toty, Mickael Baumann, Nicolas Ligerica and seven anonymous military at Europa. Additionally, Aride Island staff and volunteers, in particular Melinda Curran, provided logistic and fieldwork support and Pierre-André Adam (ICS Head Office) provided administrative support. We also thank everyone who participated in the SEOR rescue campaign of the tropical shearwaters, as well as Naïs Avargues for helping with the preparation of the plates for DNA extraction. We are grateful to the Diversity Arrays Technology platform for library preparation, sequencing and assistance with SNP genotyping. The authors would like to express their sincere gratitude to the Genotoul Bioinformatics Platform Toulouse Occitanie (Bioinfo Genotoul, doi: https://doi.org/10.15454/1.5572369328961167E12) for providing help, computing and storage resources for all the genomic analyses.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The raw DArTseq genotypes generated by Diversity Arrays Technology (DarT Pty Ltd., Canberra) and the associated specimen metadata are available on DRYAD. (http://datadryad.org/share/ZzUwJHnrpv73olfK6vokpd8pW-K38U7CGM41FN4begU). Additionally, the morphometric data is available in Supporting Information.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ddi.70078.

References

Agnarsson, I., and M. Kuntner. 2012. "The Generation of a Biodiversity Hotspot: Biogeography and Phylogeography of the Western Indian Ocean Islands." *Current Topics in Phylogenetics and Phylogeography of Terrestrial and Aquatic Systems* 33: 33–82.

Austin, J. J., V. Bretagnolle, and E. Pasquet. 2004. "A Global Molecular Phylogeny of the Small Puffinus Shearwaters and Implications for Systematics of the Little-Audubon's Shearwater Complex." *Auk* 121, no. 3: 847–864.

Bagley, R. K., V. C. Sousa, M. L. Niemiller, and C. R. Linnen. 2017. "History, Geography and Host Use Shape Genomewide Patterns of Genetic Variation in the Redheaded Pine Sawfly (*Neodiprion lecontei*)." *Molecular Ecology* 26, no. 4: 1022–1044. https://doi.org/10.1111/mec.13972.

Bailey, R. S. 1968. "The Pelagic Distribution of Sea-Birds in the Western Indian Ocean." *Ibis* 110, no. 4: 493–519. https://doi.org/10.1111/j.1474-919X.1968.tb00060.x.

Baums, I. B., J. N. Boulay, N. R. Polato, and M. E. Hellberg. 2012. "No Gene Flow Across the Eastern Pacific Barrier in the Reef-Building Coral *Porites lobata*." *Molecular Ecology* 21, no. 22: 5418–5433. https://doi.org/10.1111/j.1365-294X.2012.05733.x.

Bourjea, J., S. Lapègue, L. Gagnevin, et al. 2007. "Phylogeography of the Green Turtle, *Chelonia mydas*, in the Southwest Indian Ocean." *Molecular Ecology* 16, no. 1: 175–186. https://doi.org/10.1111/j.1365-294X.2006.03122.x.

Bretagnolle, V., C. Attie, and F. Mougeot. 2000. "Audubon's Shearwaters *Puffinus lherminieri* on Reunion Island, Indian Ocean: Behaviour, Census, Distribution, Biometrics and Breeding Biology." *Ibis* 142: 399–412.

Brooke, M. 2004. *Albatrosses and Petrels Across the World*, Oxford University Press.

Calabrese, L. 2016. "Foraging Ecology and Breeding Biology of Wedge-Tailed Shearwater (*Puffinus pacificus*) and Tropical Shearwater (*Puffinus bailloni*) on Aride Island Nature Reserve, Seychelles: Tools for Conservation." [PhD thesis, Université Pierre et Marie Curie].

Carr, P., S. Votier, H. Koldewey, B. Godley, H. Wood, and M. A. C. Nicoll. 2021. "Status and Phenology of Breeding Seabirds and a Review of Important Bird and Biodiversity Areas in the British Indian Ocean Territory." *Bird Conservation International* 31, no. 1: 14–34. https://doi.org/10.1017/S0959270920000295.

Carvalho, M. D. A., C. C. Lana, N. P. Sá, G. Santiago, M. C. S. Giannerini, and P. Bengtson. 2022. "Influence of the Intertropical Convergence Zone on Early Cretaceous Plant Distribution in the South Atlantic." *Scientific Reports* 12: 12600. https://doi.org/10.1038/s41598-022-16580-x.

Chevillon, L., J. Tourmetz, J. Dubos, et al. 2022. "25 Years of Light-Induced Petrel Groundings in Reunion Island: Retrospective Analysis and Predicted Trends." *Global Ecology and Conservation* 38: e02232. https://doi.org/10.1016/j.gecco.2022.e02232.

Chifman, J., and L. Kubatko. 2014. "Quartet Inference From SNP Data Under the Coalescent Model." *Bioinformatics* 30, no. 23: 3317–3324. https://doi.org/10.1093/bioinformatics/btu530.

Choquet, M., A. M. Lizano, A. Le Moan, M. Ravinet, A. K. S. Dhanasiri, and G. Hoarau. 2023. "Unmasking Microsatellite Deceptiveness and Debunking Hybridization With SNPs in Four Marine Copepod Species of *Calanus*." *Molecular Ecology* 32, no. 24: 6854–6873. https://doi.org/10.1111/MEC.17183.

Cruz-Flores, M., R. Pradel, J. Bried, et al. 2022. "Will Climate Change Affect the Survival of Tropical and Subtropical Species? Predictions Based on Bulwer's Petrel Populations in the NE Atlantic Ocean." *Science of the Total Environment* 847: 157352. https://doi.org/10.1016/j.scitotenv.2022.157352.

Cuevas-Caball, C., J. Ferrer Obiol, J. Vizueta, et al. 2022. "The First Genome of the Balearic Shearwater (*Puffinus mauretanicus*) Provides a Valuable Resource for Conservation Genomics and Sheds Light on Adaptation to a Pelagic Lifestyle." *Genome Biology and Evolution* 14, no. 5: evac067. https://doi.org/10.1093/gbe/evac067.

Danckwerts, D. K., L. Humeau, P. Pinet, C. D. McQuaid, and M. Le Corre. 2021. "Extreme Philopatry and Genetic Diversification at Unprecedented Scales in a Seabird." *Scientific Reports* 11, no. 1: 6834. https://doi.org/10.1038/s41598-021-86406-9.

Davey, J. L., and M. W. Blaxter. 2010. "RADseq: Next-Generation Population Genetics." *Briefings in Functional Genomics* 9, no. 5–6: 416–423. https://doi.org/10.1093/bfgp/elq031.

Devloo-Delva, F., G. E. Maes, S. I. Hernández, et al. 2019. "Accounting for Kin Sampling Reveals Genetic Connectivity in Tasmanian and New Zealand School Sharks, *Galeorhinus galeus*." *Ecology and Evolution* 9, no. 8: 4465–4472. https://doi.org/10.1002/ECE3.5012.

Dorey, J. B., S. V. C. Groom, A. Velasco-Castrillón, M. I. Stevens, M. S. Y. Lee, and M. P. Schwarz. 2021. "Holocene Population Expansion of a Tropical Bee Coincides With Early Human Colonization of Fiji Rather Than Climate Change." *Molecular Ecology* 30, no. 16: 4005–4022. https://doi.org/10.1111/mec.16034.

Dray, S., A.-B. Dufour, and D. Chessel. 2007. "The ade4 Package-II: Two-Table and K-Table Methods." *R News* 7: 47–52.

Duncan, K. M., A. P. Martin, B. W. Bowen, and H. G. De Couet. 2006. "Global Phylogeography of the Scalloped Hammerhead Shark (*Sphyrna lewini*)." *Molecular Ecology* 15, no. 8: 2239–2251. https://doi.org/10.1111/j.1365-294X.2006.02933.x.

Dunn, R. E., R. Freeman, M. A. Nicoll, et al. 2024. "From Route to Dive: Multi-Scale Habitat Selection in a Foraging Tropical Seabird." *Marine Biology* 171, no. 6: 124. https://doi.org/10.1007/s00227-024-04445-y.

Egevang, C., I. J. Stenhouse, R. A. Phillips, A. Petersen, J. W. Fox, and J. R. D. Silk. 2010. "Tracking of Arctic Terns *Sterna paradisaea* Reveals Longest Animal Migration." *Proceedings of the National Academy of Sciences of the United States of America* 107, no. 5: 2078–2081. https://doi.org/10.1073/pnas.0909493107.

El Ayari, T., N. El Trigui Menif, B. Hamer, A. E. Cahill, and N. Bierne. 2019. "The Hidden Side of a Major Marine Biogeographic Boundary: A Wide Mosaic Hybrid Zone at the Atlantic–Mediterranean Divide Reveals the Complex Interaction Between Natural and Genetic Barriers in Mussels." *Heredity* 122, no. 6: 770–784. https://doi.org/10.1038/s41437-018-0174-y.

Evanno, G., S. Regnaut, and J. Goudet. 2005. "Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study." *Molecular Ecology* 14, no. 8: 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Faria, P. J., F. P. Campos, J. O. Branco, C. M. Musso, J. S. Morgante, and M. W. Bruford. 2010. "Population Structure in the South American Tern *Sterna hirundinacea* in the South Atlantic: Two Populations With Distinct Breeding Phenologies." *Journal of Avian Biology* 41, no. 4: 378–387. https://doi.org/10.1111/j.1600-048X.2009.04902.x.

Ferrer Obiol, J., J. M. Herranz, J. R. Paris, et al. 2023. "Species Delimitation Using Genomic Data to Resolve Taxonomic Uncertainties in a Speciation Continuum of Pelagic Seabirds." *Molecular Phylogenetics and Evolution* 179: 107671. https://doi.org/10.1016/j.ympev.2022.107671.

Ferrer Obiol, J., H. F. James, R. T. Chesser, et al. 2021. "Integrating Sequence Capture and Restriction Site-Associated DNA Sequencing to Resolve Recent Radiations of Pelagic Seabirds." *Systematic Biology* 70, no. 5: 976–996. https://doi.org/10.1093/sysbio/syaa101.

Feutry, P., F. Devloo-Delva, A. Tran Lu Y, et al. 2020. "One Panel to Rule Them all: DArTcap Genotyping for Population Structure, Historical Demography, and Kinship Analyses, and Its Application to a Threatened Shark." *Molecular Ecology Resources* 20, no. 6: 1470–1485. https://doi.org/10.1111/1755-0998.13204.

Fitak, R. R. 2021. "OptM: Estimating the Optimal Number of Migration Edges on Population Trees Using Treemix." *Biology Methods & Protocols* 6, no. 1: bpab017. https://doi.org/10.1093/biomethods/bpab017.

Frankham, R., J. D. Ballou, D. A. Briscoe, and K. H. McInnes. 2002. *Introduction to Conservation Genetics*. Cambridge University Press.

Fridolfsson, A.-K., and H. Ellegren. 1999. "A Simple and Universal Method for Molecular Sexing of Non-Ratite Birds." *Journal of Avian Biology* 30: 116–121.

Friesen, V. L. 2015. "Speciation in Seabirds: Why Are There So Many Species... and Why Aren't There More?" *Journal of Ornithology* 156: 27–39. https://doi.org/10.1007/s10336-015-1235-0.

Friesen, V. L., T. M. Burg, and K. D. McCoy. 2007. "Mechanisms of Population Differentiation in Seabirds: Invited Review." *Molecular Ecology* 16, no. 9: 1765–1785. https://doi.org/10.1111/j.1365-294X.2006.03197.x.

Gagnaire, P. A., Y. Minegishi, J. Aoyama, et al. 2009. "Ocean Currents Drive Secondary Contact Between *Anguilla marmorata* Populations in the Indian Ocean." *Marine Ecology Progress Series* 379: 267–278. https://doi.org/10.3354/meps07895.

Gangloff, B., H. Shirihai, D. Watling, et al. 2012. "The Complete Phylogeny of *Pseudobulweria*, the Most Endangered Seabird Genus: Systematics, Species Status and Conservation Implications." *Conservation Genetics* 13, no. 1: 39–52. https://doi.org/10.1007/s10592-011-0261-6.

Gangloff, B., F. Zino, H. Shirihai, et al. 2013. "The Evolution of North-East Atlantic Gadfly Petrels Using Statistical Phylogeography." *Molecular Ecology* 22, no. 2: 495–507. https://doi.org/10.1111/mec.12119.

Garot, E., T. Joët, M. C. Combes, and P. Lashermes. 2019. "Genetic Diversity and Population Divergences of an Indigenous Tree (*Coffea mauritiana*) in Reunion Island: Role of Climatic and Geographical Factors." *Heredity* 122, no. 6: 833–847. https://doi.org/10.1038/s41437-018-0168-9.

Garrett, L. J. H., J. P. Myatt, J. P. Sadler, et al. 2020. "Spatio-Temporal Processes Drive Fine-Scale Genetic Structure in an Otherwise Panmictic Seabird Population." *Scientific Reports* 10, no. 1: 20725. https://doi.org/10.1038/s41598-020-77517-w.

Georges, A., B. Gruber, G. B. Pauly, et al. 2018. "Genomewide SNP Markers Breathe New Life Into Phylogeography and Species Delimitation for the Problematic Short-Necked Turtles (*Chelidae: Emydura*) of Eastern Australia." *Molecular Ecology* 27, no. 24: 5195–5213. https://doi.org/10.1111/MEC.14925.

Gineste, B., M. Souquet, F. X. Couzi, et al. 2017. "Tropical Shearwater Population Stability at Reunion Island, Despite Light Pollution." *Journal of Ornithology* 158, no. 2: 385–394. https://doi.org/10.1007/s10336-016-1396-5.

Gómez-Díaz, E., J. González-Solís, and M. A. Peinado. 2009. "Population Structure in a Highly Pelagic Seabird, the Cory's Shearwater *Calonectris diomedea*: An Examination of Genetics, Morphology and Ecology." *Marine Ecology Progress Series* 382: 197–209. https://doi.org/10.3354/meps07974.

Gómez-Díaz, E., J. A. Morris-Pocock, J. González-Solís, and K. D. McCoy. 2012. "Trans-Oceanic Host Dispersal Explains High Seabird Tick Diversity on Cape Verde Islands." *Biology Letters* 8, no. 4: 616–619. https://doi.org/10.1098/rsbl.2012.0179.

Grémillet, D., S. Lewis, L. Drapeau, et al. 2008. "Spatial Match-Mismatch in the Benguela Upwelling Zone: Should We Expect Chlorophyll and Sea-Surface Temperature to Predict Marine Predator Distributions?" *Journal of Applied Ecology* 45, no. 2: 610–621. https://doi.org/10.1111/j.1365-2664.2007.01447.x.

Gruber, B., P. J. Unmack, O. F. Berry, and A. Georges. 2018. "Dartr: An r Package to Facilitate Analysis of SNP Data Generated From Reduced Representation Genome Sequencing." *Molecular Ecology Resources* 18, no. 3: 691–699. https://doi.org/10.1111/1755-0998.12745.

Hailer, F., E. A. Schreiber, J. M. Miller, et al. 2011. "Long-Term Isolation of a Highly Mobile Seabird on the Galapagos." *Proceedings of the Royal Society B: Biological Sciences* 278, no. 1707: 817–825. https://doi.org/10.1098/rspb.2010.1342.

Hanghøj, K., I. Moltke, P. A. Andersen, A. Manica, and T. S. Korneliussen. 2019. "Fast and Accurate Relatedness Estimation From High-Throughput Sequencing Data in the Presence of Inbreeding." *GigaScience* 8, no. 5: 34. https://doi.org/10.1093/gigascience/giz034.

Herman, R. W., B. M. Winger, D. L. Dittmann, and M. G. Harvey. 2022. "Fine-Scale Population Genetic Structure and Barriers to Gene Flow in a Widespread Seabird (*Ardenna pacifica*)." *Biological Journal of the Linnean Society* 137, no. 1: 125–136. https://doi.org/10.1093/BIOLINNEAN/BLAC091.

Herve, M. 2020. *RVAideMemoire: Testing and Plotting Procedures for Biostatistics. R Package Version 0.9-73*. R Foundation.

Huberty, C. J., and S. Olejnik. 2007. *Applied MANOVA and Discriminant Analysis*. John Wiley & Sons.

Humeau, L., M. Le Corre, S. J. Reynolds, et al. 2020. "Genetic Structuring Among Colonies of a Pantropical Seabird: Implication for Subspecies Validation and Conservation." *Ecology and Evolution* 10, no. 21: 11886–11905. https://doi.org/10.1002/ece3.6635.

Jaccoud, D., K. Peng, D. Feinstein, and A. Kilian. 2001. "Diversity Arrays: A Solid State Technology for Sequence Information Independent Genotyping." *Nucleic Acids Research* 29, no. 4: E25.

Jensen, M. P., N. N. FitzSimmons, J. Bourjea, T. Hamabata, J. Reece, and P. H. Dutton. 2019. "The Evolutionary History and Global Phylogeography of the Green Turtle (*Chelonia mydas*)." *Journal of Biogeography* 46, no. 5: 860–870. https://doi.org/10.1111/jbi.13483.

Jombart, T., S. Devillard, and F. Balloux. 2010. "Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations." *BMC Genetics* 11: 94. https://doi.org/10.1186/1471-2156-11-94.

Jouanin, C. 1970. "Note Taxonomique sur les Petits Puffins, *Puffinus lherminieri*, de l'Océan Indien Occidental." *Oiseau et RFO* 40: 303–306.

Kamvar, Z. N., J. F. Tabima, and N. J. Gŕunwald. 2014. "Poppr: An R Package for Genetic Analysis of Populations With Clonal, Partially Clonal, and/or Sexual Reproduction." *PeerJ* 2014, no. 1: 1–14. https://doi.org/10.7717/PEERJ.281/TABLE-6.

Keesing, J., and T. Irvine. 2005. "Coastal Biodiversity in the Indian Ocean: The Known, the Unknown and the Unknowable." *Indian Journal of Marine Sciences* 34, no. 1: 11–26.

Keinan, A., J. C. Mullikin, N. Patterson, and D. Reich. 2007. "Measurement of the Human Allele Frequency Spectrum Demonstrates Greater Genetic Drift in East Asians Than in Europeans." *Nature Genetics* 39, no. 10: 1251–1255. https://doi.org/10.1038/ng2116.

Kersten, O., B. Star, D. M. Leigh, et al. 2021. "Complex Population Structure of the Atlantic Puffin Revealed by Whole Genome Analyses." *Communications Biology* 4, no. 1: 922. https://doi.org/10.1038/s42003-021-02415-4.

Kilian, A., P. Wenzl, E. Huttner, et al. 2012. "Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms." *Methods in Molecular Biology* 888: 67–89. https://doi.org/10.1007/978-1-61779-870-2_5.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. "Clumpak: A Program for Identifying Clustering Modes and Packaging Population Structure Inferences Across K." *Molecular Ecology Resources* 15, no. 5: 1179–1191. https://doi.org/10.1111/1755-0998.12387.

Korneliussen, T. S., A. Albrechtsen, and R. Nielsen. 2014. "ANGSD: Analysis of Next Generation Sequencing Data." *BMC Bioinformatics* 15, no. 1: 356. https://doi.org/10.1186/s12859-014-0356-4.

Le Corre, M. 2000. "Taxonomic Affinities of Audubon's Shearwater From Europa Island." *Condor* 102, no. 1: 187–190.

Lett, C., B. S. Malauene, T. B. Hoareau, D. M. Kaplan, and F. Porri. 2024. "Corridors and Barriers to Marine Connectivity Around Southern Africa." *Marine Ecology Progress Series* 731: 105–127. https://doi.org/10.3354/meps14312.

Mantel, N. 1967. "The Detection of Disease Clustering and a Generalized Regression Approach." *Cancer Research* 27, no. 2: 209–220.

McDuie, F., S. J. Weeks, and B. C. Congdon. 2018. "Oceanographic Drivers of Near-Colony Seabird Foraging Site Use in Tropical Marine Systems." *Marine Ecology Progress Series* 589: 209–225. https://doi.org/10.3354/meps12475.

Meier, J. I., V. C. Sousa, D. A. Marques, et al. 2017. "Demographic Modelling With Whole-Genome Data Reveals Parallel Origin of Similar *Pundamilia* Cichlid Species After Hybridization." *Molecular Ecology* 26, no. 1: 123–141. https://doi.org/10.1111/mec.13838.

Mendez, M., H. C. Rosenbaum, A. Subramaniam, C. Yackulic, and P. Bordino. 2010. "Isolation by Environmental Distance in Mobile Marine Species: Molecular Ecology of Franciscana Dolphins at Their Southern Range." *Molecular Ecology* 19, no. 11: 2212–2228. https://doi.org/10.1111/j.1365-294X.2010.04647.x.

Mijangos, J. L., B. Gruber, O. Berry, C. Pacioni, and A. Georges. 2022. "dartR v2: An Accessible Genetic Analysis Platform for Conservation, Ecology and Agriculture." *Methods in Ecology and Evolution* 13, no. 10: 2150–2158. https://doi.org/10.1111/2041-210X.13918.

Milot, E., H. Weimerskirch, and L. Bernatchez. 2008. "The Seabird Paradox: Dispersal, Genetic Structure and Population Dynamics in a Highly Mobile, but Philopatric Albatross Species." *Molecular Ecology* 17, no. 7: 1658–1673. https://doi.org/10.1111/j.1365-294X.2008.03700.x.

Mlíkovský, J. 2005. "Subfossil Birds of Andrahomana, Southeastern Madagascar." *Annalen Des Naturhistorischen Museums in Wien* 107A: 87–92.

Monteiro, L. R., and R. W. Furness. 1998. "Speciation Through Temporal Segregation of Madeiran Storm Petrel (*Oceanodroma castro*) Populations in the Azores?" *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 353: 945–953.

Morris-Pocock, J. A., D. J. Anderson, and V. L. Friesen. 2016. "Biogeographical Barriers to Dispersal and Rare Gene Flow Shape Population Genetic Structure in Red-Footed Boobies (*Sula sula*)." *Journal of Biogeography* 43, no. 11: 2125–2135. https://doi.org/10.1111/JBI.12780.

Morris-Pocock, J. A., T. E. Steeves, F. A. Estela, D. J. Anderson, and V. L. Friesen. 2010. "Comparative Phylogeography of Brown (*Sula leucogaster*) and Red-Footed Boobies (*S. sula*): The Influence of Physical Barriers and Habitat Preference on Gene Flow in Pelagic Seabirds." *Molecular Phylogenetics and Evolution* 54, no. 3: 883–896. https://doi.org/10.1016/j.ympev.2009.11.013.

Nadachowska-Brzyska, K., R. Burri, L. Smeds, and H. Ellegren. 2016. "PSMC Analysis of Effective Population Sizes in Molecular Ecology and Its Application to Black-And-White *Ficedula* Flycatchers." *Molecular Ecology* 25, no. 5: 1058–1072. https://doi.org/10.1111/mec.13540.

Nikolic, N., F. Devloo-Delva, D. Bailleul, et al. 2020. "Genome Scans Discriminate Independent Populations of the Blue Shark *Prionace glauca*." In *WPEB16 – 15th Working Party on Ecosystems and Bycatch*. https://archimer.ifremer.fr/doc/00645/75701/.

Obiol, J. F., J. M. Herranz, J. R. Paris, et al. 2023. "Species Delimitation Using Genomic Data to Resolve Taxonomic Uncertainties in a Speciation Continuum of Pelagic Seabirds." *Molecular Phylogenetics and Evolution* 179: 107671.

Oksanen, J. 2013. "Package 'vegan'. Community Ecology Package, Version 2.9. *9*, 1–295."

Olavarría, C., C. Scott Baker, C. Garrigue, et al. 2007. "Population Structure of South Pacific Humpback Whales and the Origin of the Eastern Polynesian Breeding Grounds." *Marine Ecology Progress Series* 330: 257–268.

Pages, J. 2004. "Analyse Factorielle de Donnees Mixtes." *Revue de Statistique Appliquée* 52, no. 4: 93–111.

Paradis, E., S. Blomberg, B. Bolker, et al. 2019. "Package 'ape.' Analyses of Phylogenetics and Evolution." 2.4, 47.

Pebesma, E. 2012. "Spacetime: Spatio-Temporal Data in R." *Journal of Statistical Software* 51: 1–30.

Pedersen, C.-E. T., A. Albrechtsen, P. D. Etter, et al. 2018. "A Southern African Origin and Cryptic Structure in the Highly Mobile Plains Zebra." *Nature Ecology & Evolution* 2, no. 3: 491–498. https://doi.org/10.1038/s41559-017-0453-7.

Pembleton, L. W., N. O. I. Cogan, and J. W. Forster. 2013. "StAMPP: An R Package for Calculation of Genetic Differentiation and Structure of Mixed-Ploidy Level Populations." *Molecular Ecology Resources* 13, no. 5: 946–952. https://doi.org/10.1111/1755-0998.12129.

Petkova, D., J. Novembre, and M. Stephens. 2015. "Visualizing Spatial Population Structure With Estimated Effective Migration Surfaces." *Nature Genetics* 48, no. 1: 94–100. https://doi.org/10.1038/ng.3464.

Pickrell, J. K., and J. K. Pritchard. 2012. "Inference of Population Splits and Mixtures From Genome-Wide Allele Frequency Data." *Nature Precedings* 2012: 1. https://doi.org/10.1038/npre.2012.6956.1.

Poelstra, J. W., B. K. Montero, J. Lüdemann, et al. 2022. "RADseq Data Reveal a Lack of Admixture in a Mouse Lemur Contact Zone Contrary to Previous Microsatellite Results." *Proceedings of the Royal Society B: Biological Sciences* 289, no. 1980: 20220596. https://doi.org/10.1098/rspb.2022.0596.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155, no. 2: 945–959. https://doi.org/10.1093/GENETICS/155.2.945.

Purcell, S., B. Neale, K. Todd-Brown, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81, no. 3: 559–575. https://doi.org/10.1086/519795.

Pyle, P., A. J. Welch, and R. C. Fleischer. 2011. "A New Species of Shearwater (*Puffinus*) Recorded From Midway Atoll, Northwestern Hawaiian Islands." *Condor* 113, no. 3: 518–527. https://doi.org/10.1525/cond.2011.100117.

Rambaut, A., and A. Drummond. 2008. *FigTree: Tree Figure Drawing Tool, Version 1.2. 2*. Institute of Evolutionary Biology, University of Edinburgh.

Reich, D., K. Thangaraj, N. Patterson, A. L. Price, and L. Singh. 2009. "Reconstructing Indian Population History." *Nature* 461, no. 7263: 489–494. https://doi.org/10.1038/nature08365.

Reid, K., T. B. Hoareau, J. E. Graves, et al. 2016. "Secondary Contact and Asymmetrical Gene Flow in a Cosmopolitan Marine Fish Across the Benguela Upwelling Zone." *Heredity* 117, no. 5: 307–315. https://doi.org/10.1038/hdy.2016.51.

Richards, V. P., t. W. GReiG, P. FaiR, et al. 2013. "Patterns of Population Structure for Inshore Bottlenose Dolphins Along the Eastern United States." *Journal of Heredity* 104, no. 6: 765–778. https://doi.org/10.5061/dryad.v2m91.

Romero-Torres, M., E. A. Treml, A. Acosta, and D. A. Paz-Garciá. 2018. "The Eastern Tropical Pacific Coral Population Connectivity and the Role of the Eastern Pacific Barrier." *Scientific Reports* 8, no. 1: 9354. https://doi.org/10.1038/s41598-018-27644-2.

Russo, I. R. M., C. L. Sole, M. Barbato, U. Von Bramann, and M. W. Bruford. 2016. "Landscape Determinants of Fine-Scale Genetic Structure of a Small Rodent in a Heterogeneous Landscape (Hluhluwe-iMfolozi Park, South Africa)." *Scientific Reports* 6: 29168. https://doi.org/10.1038/srep29168.

Sansaloni, C., C. Petroli, D. Jaccoud, et al. 2011. "Diversity Arrays Technology (DArT) and Next-Generation Sequencing Combined: Genome-Wide, High Throughput, Highly Informative Genotyping for Molecular Breeding of Eucalyptus." *BMC Proceedings* 5, no. 7: 1–2. https://doi.org/10.1186/1753-6561-5-S7-P54.

Schott, F. A., and J. P. Mccreary. 2001. "The Monsoon Circulation of the Indian Ocean." *Progress in Oceanography* 51: 1–123.

Schott, F. A., S. P. Xie, and J. P. McCreary Jr. 2009. "Indian Ocean Circulation and Climate Variability." *Reviews of Geophysics* 47, no. 1: RG1002. https://doi.org/10.1029/2007RG000245.

Shirihai, H., and V. Bretagnolle. 2015. "The Poorly Known Mohéli Shearwater *Puffinus (persicus) temptator* at the Comoro Islands, Western Indian Ocean." *Bulletin of the British Ornithologists Club* 135, no. 3: 216–223.

Shirihai, H., and D. A. Christie. 1996. "A New Taxon of Small Shearwater From the Indian Ocean." *Bulletin of the British Ornithologists' Club* 3, no. 3: 116.

Silva, M. C., R. Matias, R. M. Wanless, et al. 2015. "Understanding the Mechanisms of Antitropical Divergence in the Seabird White-Faced Storm-Petrel (Procellariiformes: *Pelagodroma marina*) Using a Multilocus Approach." *Molecular Ecology* 24, no. 12: 3122–3137. https://doi.org/10.1111/mec.13212.

Steeves, T. E., D. J. Anderson, and V. L. Friesen. 2005a. "A Role for Nonphysical Barriers to Gene Flow in the Diversification of a Highly Vagile Seabird, the Masked Booby (*Sula dactylatra*)." *Molecular Ecology* 14, no. 12: 3877–3887. https://doi.org/10.1111/j.1365-294X.2005.02713.x.

Steeves, T. E., D. J. Anderson, and V. L. Friesen. 2005b. "The Isthmus of Panama: A Major Physical Barrier to Gene Flow in a Highly Mobile Pantropical Seabird." *Journal of Evolutionary Biology* 18, no. 4: 1000–1008. https://doi.org/10.1111/j.1420-9101.2005.00906.x.

Steeves, T. E., D. J. Anderson, H. McNally, M. H. Kim, and V. L. Friesen. 2003. "Phylogeography of Sula: The Role of Physical Barriers to Gene Flow in the Diversification of Tropical Seabirds." *Journal of Avian Biology* 34, no. 2: 217–223. https://doi.org/10.1034/j.1600-048X.2003.03086.x.

Teixeira, H., M. Le Corre, L. Michon, et al. 2024. "Past Volcanic Activity Predisposes an Endemic Threatened Seabird to Negative Anthropogenic Impacts." *Scientific Reports* 14, no. 1: 1960. https://doi.org/10.1038/s41598-024-52556-9.

Torres, L., E. Pante, J. González-Solís, et al. 2021. "Sea Surface Temperature, Rather Than Land Mass or Geographic Distance, May Drive Genetic Differentiation in a Species Complex of Highly Dispersive Seabirds." *Ecology and Evolution* 11, no. 21: 14960–14976. https://doi.org/10.1002/ece3.8180.

Trevail, A. M., M. A. C. Nicoll, R. Freeman, et al. 2023. "Tracking Seabird Migration in the Tropical Indian Ocean Reveals Basin-Scale Conservation Need." *Current Biology* 33, no. 23: 5247–5256.e4. https://doi.org/10.1016/j.cub.2023.10.060.

Triest, L., T. Van der Stocken, D. De Ryck, et al. 2021. "Expansion of the Mangrove Species *Rhizophora mucronata* in the Western Indian Ocean Launched Contrasting Genetic Patterns." *Scientific Reports* 11, no. 1: 4987. https://doi.org/10.1038/s41598-021-84304-8.

van der Ven, R. M., J. F. Flot, C. Buitrago-López, and M. Kochzius. 2021. "Population Genetics of the Brooding Coral *Seriatopora hystrix* Reveals Patterns of Strong Genetic Differentiation in the Western Indian Ocean." *Heredity* 126, no. 2: 351–365. https://doi.org/10.1038/s41437-020-00379-5.

Warren, B. H., E. Bermingham, R. C. K. Bowie, R. P. Prys-Jones, and C. Thébaud. 2003. "Molecular Phylogeography Reveals Island Colonization History and Diversification of Western Indian Ocean Sunbirds (Nectarinia: *Nectariniidae*)." *Molecular Phylogenetics and Evolution* 29, no. 1: 67–85. https://doi.org/10.1016/S1055-7903(03)00063-0.

Weir, B. S., and C. C. Cockerham. 1984. "Estimating F-Statistics for the Analysis of Population Structure." *Evolution* 38, no. 6: 1358–1370. https://doi.org/10.2307/2408641.

Whitlock, M. C., and K. E. Lotterhos. 2015. "Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model Through Trimming the Distribution of FST." *American Naturalist* 186: S24–S36. https://doi.org/10.1086/682949.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Metadata file containing information about the 90 Tropical shearwaters (*Puffinus bailloni*; TSH) genotyped and respective sampling details. **Table S2:** Oceanographic variables used for the isolation-by-environmental tests. **Table S3:** Metadata file containing values of five morphometric measurements and colour of undertail coverts obtained from 197 breeding adults from all study colonies but Aride. **Table S4:** Mean values (± SD) of five morphometric measurements and colour of undertail coverts per colony. **Table S5:** Quality filtering steps applied to the DArTseq dataset for loci and individuals. **Table S6:** Matrix of fixed genetic differences among colonies used for the identification of the operational taxonomic units. **Table S7:** Results for the three-population ($f3$) tests. **Table S8:** Results for the four-population ($f4$) statistics for the Tropical shearwater. **Figure S1:** Individual heterozygosity estimated following (Devloo-Delva et al. 2019). Individuals with a heterozygosity higher than 0.14 were excluded from our dataset to control for potential cross-contamination. **Figure S2:** Principal component analysis (PCA) illustrating genetic similarity between the 44 individuals from subsp. *nicolae* and *colstoni* using 18,065 genome-wide SNPs. **Figure S3:** Bayesian Information Criterion (BIC) used to identify the best K value for the discriminant analysis of principal components (DAPC) analyses considering (a) the entire dataset ($n = 74$); (b) the subsp. nicolae and colstoni ($n = 44$); and (c) the subsp. *bailloni* ($n = 30$) using the "Adegenet" R package following Jombart et al. (2010). Values with the lowest BIC score indicate the best fit to the data. **Figure S4:** Discriminant analysis of principal components (*DAPC*) results considering two ($K = 2$) and three ($K = 3$) genetic clusters. **Figure S5:** Detection of the number of genetic clusters (K) using the log likelihood mean values L(K) (black circles) and ΔK statistic following the method of Evanno (Evanno et al. 2005) (black triangles) as derived from STRUCTURE for (a) the entire dataset (K ranging from 1 to 6); (b) subsp. *bailloni* (K ranging from 1 to 5); and (c) the subsp. *nicolae* and *colstoni* (K ranging from 1 to 5). Each value was obtained by averaging the posterior probabilities over 10 independent runs. **Figure S6:** Cluster assignment of 74 Tropical shearwater (*Puffinus bailloni*) individuals from two ($K = 2$) to six ($K = 6$) genetic clusters using 27,273 genome-wide SNPs with Structure. **Figure S7:** Cluster assignment of 30 individuals from subsp. *bailloni* from two ($K = 2$) to five ($K = 5$) genetic clusters using 16,105 genome-wide SNPs with Structure. **Figure S8:** Cluster assignment of 44 individuals from subsp. *nicolae* and *colstoni* from two ($K = 2$) to five ($K = 5$) genetic clusters using 18,065 genome-wide SNPs with structure. **Figure S9:** Maximum likelihood tree obtained for the 74 Tropical shearwater (*Puffinus bailloni*) individuals using TreeMix (a) without migration edges ($m = 0$); and (b) with one migration event ($m = 1$). **Figure S10:** Estimation of the best number of migration edges (m) for the TreeMix maximum likelihood tree considering the distribution of the mean likelihood values across m (black circles)

and the Δm values indicating the second-order rate of change in the log likelihood between models with different m (black triangles) following Fitak (2021). **Figure S11:** Scatterplot from a principal component analysis (PCA) of morphological variation based on five morphometric measurements based on 197 tropical shearwaters (*Puffinus bailloni*) from all study sites but Aride. **Figure S12:** The breeding phenology of the studied Tropical shearwater (*Puffinus bailloni*) colonies is displayed by different shades of darkness: black = known breeding season, grey = breeding reported only for few pairs, and white = no breeding. **Figure S13:** Monthly mean variation (± SD) of the sea surface temperature (SST, °C) among the six study sites.