Conservation genomics reveals low connectivity among populations of threatened roseate terns (*Sterna dougallii*) in the Atlantic Basin

Paige A. Byerly[1,4,6] · R. Terry Chesser[2,3] · Robert C. Fleischer[4] · Nancy McInerney[4] · Natalia A. S. Przelomska[3,4,5] · Paul L. Leberg[1]

* Paige A. Byerly byerlyp@si.edu

[1] University of Louisiana at Lafayette, Lafayette, LA, USA

[2] Eastern Ecological Science Center, U.S. Geological Survey, Laurel, MD, USA

[3] National Museum of Natural History, Washington, DC, USA

[4] Smithsonian's National Zoo and Conservation Biology Institute, Washington, DC, USA

[5] Royal Botanic Gardens, Kew, Richmond, UK

[6] Present Address: Center for Conservation Genomics, Smithsonian's National Zoo and Conservation Biology Institute, Washington, DC, USA

Abstract

While the effects of barriers to dispersal such as population declines, habitat fragmentation, and geographic distance have been well-documented in terrestrial wildlife, factors impeding the dispersal of highly vagile taxa such as seabirds are less well understood. The roseate tern (*Sterna dougallii*) is a globally distributed seabird species, but populations tend to be both fragmented and small, and the species is declining across most of its range. We evaluated structuring of roseate tern populations in the Northwestern Atlantic, the Caribbean, and the Azores using both microsatellite markers and single-nucleotide polymorphisms generated through targeted sequencing of Ultra-conserved Elements. For both marker types, we found significant genetic differentiation among all 3 populations and evidence for moderate contemporary unidirectional gene flow from the Caribbean to the Azores, but not between other populations. Within the Caribbean population, we found high rates of unidirectional migration from the Virgin Islands to Florida, potentially indicative of movement from source population to sink or an artifact of dispersal among other unsampled populations in the Caribbean region. These observations have significance for species persistence in the Atlantic, as our results indicate that loss of genetic diversity within populations is unlikely to be buffered by inflow of new alleles from other breeding populations.

Keywords Conservation genomics · Genetic structuring · Laridae · Population connectivity · Seabird · *Sterna dougallii* · Roseate tern

Introduction

Connectivity, or migration of breeding individuals between groups, can be integral for the persistence of geographically isolated wildlife populations. Periodic exchanges of genetic material help maintain genetic diversity in small populations (Chesser 1991), which can be crucial for preserving both population-wide fitness (Reed and Frankham 2003) and adaptive potential (Pauls et al. 2013). Unidirectional immigration into small populations also promotes stability via the "rescue effect", with influxes of dispersing immigrants from a source population helping to stall declines and extinctions (Brown and Kodric-Brown 1977). Human activities have led to habitat fragmentation and population declines across wildlife taxa (Dirzo et al. 2014; McCauley et al. 2015), both of which can alter historical connectivity patterns (Hess 1996) and lead to the isolation of populations. Quantifying connectivity among populations has therefore become increasingly recognized as a priority action for the conservation of threatened and endangered species (Pierson et al. 2016; Funk et al. 2018; Ralls et al. 2018; Hoban et al. 2020).

Although it is tempting to conflate dispersal with an organism's movement capabilities, empirical quantification of gene flow between populations is necessary both because connectivity is not always determined by movement potential (Lombal et al. 2020) and because gene flow is not always facilitated through dispersal (Bensch et al. 1998; Peter- son et al. 2014; Barbraud & Delord 2020). This disparity between potential and actual movement is exemplified by highly mobile avian taxa such as seabirds. Most seabird species are seasonally migratory and have the potential to disperse across vast distances. Accordingly, they often exhibit an overall lack of genetic structure within the same oceanic basin due to high rates of connectivity among populations (Genovart et al. 2003; Bicknell et al. 2012; Mariano-Jelicich and Madrid 2014; Wojczulanis-Jakubas et al. 2015; Tigano et al. 2015; Yannic et al. 2016; Quillfeldt et al. 2017). How- ever, genetic structuring has also been detected among sea- bird populations on relatively small spatial scales such as between continents or even islands (Nuss et al. 2016; Sruoga and Butkauskas 2006). Restricted gene flow between seabird populations can be indicative of isolating mechanisms at work, including natal philopatry (Hailer et al. 2011; Welch et al. 2012; Danckwerts et al. 2021), assortative mating due to adaptive differences in morphology or breeding parameters (Lombal et al. 2017; Nunes and Bugoni 2018; Her- man et al. 2022), or genetic structuring resulting from range contractions and population declines (Cristofari et al. 2019; Peery et al. 2010). Increasingly, the continuous advancement of high-throughput sequencing technologies is enhancing our ability to investigate finer-scale genetic differentiation in wildlife, leading to major shifts in our understanding of movement and population structuring among seabird populations (Kersten et al. 2021).

The roseate tern (*Sterna dougallii*) is a globally distributed seabird species. Currently most populations are classified as locally threatened or endangered within their respective regions, and the species is largely declining across its range (Gochfeld and Burger 2020). In the Atlantic Basin, breeding populations are sparsely distributed across restricted range locations in Northern Europe, the Azores, the Caribbean, and northeastern North America, all of

which have enhanced protection following twentieth century population declines (Gochfeld and Burger 2020). Despite extensive long-term banding programs in some regions, roseate tern dispersal within the Atlantic Basin is still poorly under- stood (Mostello et al. 2014; Shealer and Saliva 1992; Spendelow et al. 2010). Adults appear to exhibit philopatry to their natal regions, if not specific natal colonies (Gochfeld and Burger 2020), which may inhibit movement between populations. However, roseate terns have high movement potential, and banded adults have been observed visiting other regions during the breeding season (Shealer and Saliva 1992; Nisbet and Cabot 1995; Hays et al. 1999). Populations also appear to overlap on their migratory pathways and shared wintering grounds of eastern South America (Western Atlantic populations) and Western Africa (Eastern Atlantic populations; Hays et al. 2002) which has been shown to promote interbreeding in other tern species (Szczys et al. 2017).

Despite the clear evidence for interregional travel, it has not been verified if these movements are indicative of genetic exchange among roseate tern populations. Mitochondrial and microsatellite loci have been previously used to assess roseate tern global connectivity, with populations found to be genetically differentiated between the Atlantic and Indo- Pacific oceanic basins and across the Atlantic (Lashko 2004) but not within the Northwestern Atlantic (Szczys et al. 2005; Dayton and Szczys 2021). Our own prior work has provided evidence of long-term genetic differentiation between breeding roseate terns in the Northwestern Atlantic and Caribbean (Byerly et al. 2022). However, these prior studies were all limited by number of genomic markers. As is common in terns (Faria et al. 2007; Mariano-Jelicich and Madrid 2014), roseate terns exhibit low genetic variability (Szczys et al. 2005; Dayton and Szczys 2021), which can make identifying genetic structuring challenging with highly informative yet low-quantity markers such as microsatellites.

Here, we investigated genetic connectivity among four roseate tern populations in the Atlantic Basin. To increase power and resolution of our genetic analyses, we supplemented microsatellites with single-nucleotide polymorphisms (SNPs) derived from capture and sequencing of Ultra-Conserved Elements (UCEs; Faircloth et al. 2012), which have successfully been applied to resolve relation- ships both among species and within species complexes for a variety of taxa (Everson et al. 2019; Tsai et al. 2019; Winker et al. 2018; Parker et al. 2020). A finding of minimal genetic structuring among populations of roseate terns would be consistent with other studies of large-scale seabird connectivity in the absence of biogeographical impediments to movement, and would indicate that gene flow indeed occurs both within and across the Atlantic basin. Alternatively, presence of genetic structuring between these breeding locations may indicate disruption of historical dispersal patterns resulting from extensive population declines in the twentieth century, a hypothesis that could be supported by findings of low genetic diversity and recent population bottlenecks in one or all sampled populations.

Materials and methods

Sample collection and DNA preparation

Samples were collected from four populations (the North- western Atlantic, Florida, Virgin Islands, and Azores) between 2017 and 2018. In the Caribbean (CAR), we collected samples from roseate tern chicks over 2017–18 from Leduck Island in the U.S. Virgin Islands (VI; *n* = 37) and from 3 sites in the Florida Keys (FL; *n* = 26; Fig. 1). Blood was collected by extracting < 25 µl from the tarsal veins of chicks. Samples were stored in Queen's lysis buffer. We collected blood samples only from chicks, which were captured by hand at nest sites, and sampled only chicks < 3 days old to facilitate rapid capture. Only one individual was sampled per nest. Because of high yearly site turnover and apparent mixing between colonies in the VI (Pierce 2009), genetic structuring is likely to be minimal, and the samples from Leduck Island were assumed to be representative of this population. We also opportunistically salvaged tissue from deceased chicks and eggs. Blood samples from Massachusetts (MA; *n* = 35) were collected in 2017 by the Massachusetts Division of Fisheries and Wildlife using similar methods. As the Northwestern Atlantic population has been demonstrated to be panmictic (Szczys et al. 2005; Dayton and Szczys 2021), we assumed that the MA colonies would be representative of roseate tern genetic variation within this region. Tissue samples were salvaged from chicks subsequent to depredation in the Azores (AZ; *n* = 10) in 2018 by researchers from the University of Azores.

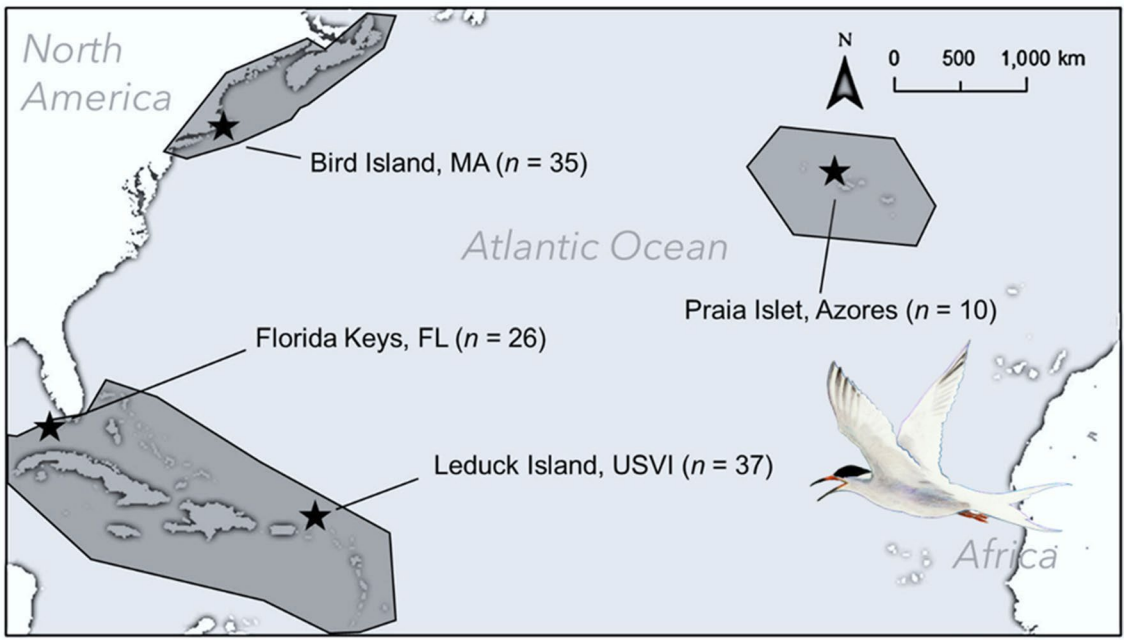

Fig. 1 Distribution of breeding roseate tern populations in the Atlantic Basin, with shading representing the approximate distribution of colonies in eastern North America (MA: Massachusetts), the Caribbean (FL: Florida, USVI: the United States Virgin Islands), and the Azores. Sampling locations within populations are represented by stars. Roseate tern illustration by C. Pavlik

We extracted DNA from all sample types using Qiagen DNeasy blood and tissue kits (Qiagen, Hilden, Germany) following recommended protocols. We quantified sample concentration using a Qubit Fluorometer (Invitrogen, Carlsbad, CA) and removed samples with total extracted DNA content of < 1.5 ng from further analysis.

Microsatellite genotyping

We amplified 14 microsatellite markers previously optimized for roseate terns in Dayton and Szczys (2021; see Table S1 for microsatellite sequences and conditions). Independent triplicates were run for low-quality (< 30 ng of input DNA) samples, samples obtained from tissue, and samples with evidence of null alleles (i.e., failure to amplify at a specific locus). PCR conditions followed Szczys et al. (2005) with an initial denaturation of 2 min at 94 °C, 35 cycles of 30 s at 72 °C, 30 s annealing at 50–58 °C, 30 s touchdown at 72 °C, and 2 min extension at 72 °C. PCR product was loaded on an ABI PRISM 3100 (Applied Biosystems Foster City, CA, USA) to separate amplicons by size and visualize alleles.

We manually detected and scored microsatellite alleles using GENEMAPPER 3.7 (Applied Biosystems). We then checked microsatellite loci for evidence of null alleles and genotyping error via 10,000 Monte Carlo simulations and 95% confidence intervals in MICRO-CHECKER (Van Oosterhout et al. 2004), which tests variant sites for evidence of excess homozygosity. We tested loci for linkage disequilibrium and deviation from Hardy–Weinberg Equilibrium (HWE) using 10,000 Monte Carlo replicates in the R package 'adegenet' 2.1.3 (Jombart 2011) and removed loci in disequilibrium. For this and all analyses involving multiple comparisons, we corrected α levels at the significance threshold α = 0.05 using a sequential Bonferroni correction (Rice 1989). Null alleles can lead to overestimation of genetic differentiation among sampling locations. We used the program FreeNA (Chapuis and Estoup 2007) to investigate the effect of potential null alleles on population differentiation in microsatellites. FreeNA estimates the null frequencies for each locus and population and applies a correction to account for the positive bias of null alleles on Weir's F_{ST} (Chapuis and Estoup 2007).

SNP genotyping

We prepared samples for genomic library construction by shearing DNA using a Biorupter with a target fragment size of < 500 bp. Fragment size was visualized via 1% agarose gel. After shearing, we prepared genomic libraries using the Blunt-end Single-Tube (BEST) method (Carøe et al. 2018), with sheared DNA fragments ligated to barcoded adapters and P7 and P5 Illumina primers, amplified for sequencing with 9–12 cycles, and purified using AMPure XP beads (Beckman Coulter, Inc.). We then quantified DNA concentration in multiplexed libraries using a Qubit Fluorometer and assessed mean fragment size using an Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA). For UCE enrichment, we combined multiplexed libraries into 750 (150 ng/library) pools. We then enriched equimolar pools for 5,060 UCE loci using the myBaits UCE Tetra- pods 5Kv1 kit (Arbor Biosciences, Ann Arbor, MI) following

the recommended protocol and with 16 cycles of post- enrichment PCR. We quantified fragment size distribution of enriched pools with a Bioanalyzer High Sensitivity dsDNA kit (Agilent Technologies, Santa Clara, CA). Enriched pools were then combined at equimolar ratios and sequenced at GENEWIZ (South Plainfield, NJ, USA) as paired-end 150 bp reads on an Illumina HiSeq system.

Sequenced genomic reads were demultiplexed using Bcl2fastq 1.8.4 (Illumina). Following demultiplexing, we processed reads using the Phyluce v.1.7.0 bioinformatics pipeline, a set of programs set up for UCE sequence analysis (Faircloth 2016). First, we cleaned reads by removing adapter contamination and low-quality bases with Illumiprocessor v2.0 (Faircloth 2016), which processes reads via the trimming tool Trimmomatic 0.32.1 (Bolger et al. 2014). For de novo contig assembly, we chose a single modern sample with an intermediary number of reads (FL1843), which we reasoned would maximize recovery of UCE contigs while minimizing the amount of duplicate contigs, and therefore processing time. To prepare our reference, we assembled the reference sample's Read 1 and 2 files de novo using the Trinity 2.0.6 assembler (Grabherr et al. 2011), mapped the output contigs to a FASTA file of the 5,060 UCE probes using Phyluce, discarding duplicate contigs and non-UCE loci in the process, and then converted the results to our final reference FASTA file. We then aligned our raw, clean reads from each sample to the reference using the 'mem' algorithm in BWA 0.7.12 (Li and Durbin 2010) and filtered the output using BWA and PicardTools 1.113 (https://sourc eforge.net/projects/picard/) within the Phyluce wrapper. We then merged all individual Binary Alignment/Map (BAM) files into a single BAM file with Picard and indexed the output merged BAM with Samtools 0.1.19 (Li et al. 2009).

To call SNPs and indels, we used HaplotypeCaller in GATK 4.1.8.1 (McKenna et al. 2010). We then used VCFtools 0.1.12b (Danecek et al. 2011) to filter SNPs by first removing indels and all loci below a Phred-scaled minimum sequence quality of 30, a minor allele count of 3, and a minimum read depth of 5, and above a maximum read depth of 100. We then filtered out individuals with > 50% missing data and restricted loci to a maximum of 10% missing data. We removed loci out of HWE while controlling for a potential Wahlund effect by grouping individuals by sampling location, testing for deviation from HWE using 1,000 Monte Carlo replicates in the package pegas v.1.1 (Paradis 2010), and removing loci out of HWE in any of the putative populations from all analyses using a sequentially Bonferroni-corrected significance threshold of $\alpha = 0.05$ (Rice 1989). To create our final dataset, we pruned loci randomly to one SNP per UCE locus to control for effects of linkage disequilibrium.

Population genetics

To measure genetic diversity, we estimated observed and expected heterozygosity (respectively, H_O and H_e) and rarefied allelic richness (A_R) for each of the four sampling locations in the R package diveRsity v.1.9.90 (Keenan et al. 2013). To evaluate differences in genetic diversity among populations, we used a randomized one-way ANOVA

with locus as a blocking factor to test for significant differences in H_O, H_e, and A_R among sampling locations and used Tukey testing to conduct multiple comparisons among population pairs if the ANOVA revealed significant interpopulation differences in a genetic diversity metric. We examined potential population differentiation with Weir and Cockerham's F_{ST} (Weir and Cockerham 1984) in diveRsity, using 10^4 boot- strap permutations to calculate bias-corrected 95% confidence intervals and test pairwise F_{ST} for significant differences among sampling locations.

Roseate tern yearly census sizes based on nest counts are not necessarily indicative of demographic changes in the population, given that individuals appear to skip breeding years; thus, effective population size (N_e) may provide a better estimate of number of breeding individuals within each population. To quantify N_e, we used the bias-corrected Linkage Disequilibrium (LD) method in the program NeEstimator v. 2.01. Estimates of N_e can be less precise with lower sample sizes and numbers of loci (Leberg 2005; England et al. 2006), and we therefore excluded the FL and AZ populations from the analysis and quantified N_e only with the SNP dataset. We estimated N_e separately for the MA and VI populations and report results from lowest allele frequency ≤ 02 following the recommendations of Waples and Do (2010) and with parametric 95% confidence intervals.

To test sampling locations for signals of recent (< 4 N_e) population bottlenecks we used BOTTLENECK 1.2.02 (Cornuet and Luikart 1996) to quantify differences between observed and expected heterozygosity under simulated mutation models. In BOTTLENECK, significant heterozygosity excess is taken as evidence of recent genetic bottle- necks under the assumption that allelic richness is lost from a population before heterozygosity with a rapid population decline. We removed AZ from the SNP analysis due to insufficient sample size (Cornuet and Luikart 1996). For both microsatellites and SNPs, we used the sign test to investigate significant differences between heterozygosity excess and expected heterozygosity excess under mutation-drift equilibrium. We examined microsatellites under both a stepwise mutation model (SMM) and an intermediate two-phase model (TPM) with 90% stepwise mutation, and examined SNPs under an infinite allele mutation (IAM) model, with 1000 iterations for all models.

Population structure

We inferred population structuring or number of genetic clusters among our sampling locations for both microsatellite and SNP data using two methods. The first, STRUCTU RE v. 2.3.4. (Pritchard 2000), uses Bayesian clustering to assign individuals to the most likely number of genetic populations by grouping genotypes to minimize deviations from HWE and linkage equilibrium. For both marker types, we ran STRUCTURE simulations for 1–5 possible clusters (K) using the admixture model with correlated allele frequencies and with sampling location both included and omitted as prior information. We ran 10 iterations each of K = 1–5 with 10^4 MCM repetitions after a burn-in of 10^4 generations.

Because primary methods of determining optimal *K* are intended to be used as ad hoc estimators and not the sole determination of *K* (Janes et al. 2017), we used multiple methods to infer the optimal value of *K*. We used the online program STRUCTURE HARVESTER v0.6.94 (Earl and von Holdt 2012) to infer the most likely value of *K* via the Evanno et al. (2005) method, which defines the best sup- ported value of *K* as that with the highest ΔK, the rate of change in log probability of the data between successive K values. As the Evanno method does not return results from $K = 1$, and has been shown to be biased towards an optimal value of $K = 2$ (Janes et al. 2017), we validated output from the Evanno method by also reporting the mean log likelihood from each iteration of *K* (LnP(*K*); Pritchard 2000). We then aligned clusters, merged STRUCTURE runs by *K* value, and visualized output using the R package Pophelper 2.3.0 (Francis 2016).

STRUCTURE models can be sensitive to deviations from HWE and linkage equilibrium, which we suspected might occur if populations were substructured or subject to inbreeding. We therefore complemented analysis of population structuring in STRUCTURE with a Discriminant Analysis of Principal Components (DAPC) for both marker types in adegenet. DAPC is a multivariate analysis method that uses principal components (PCs) to infer genetic clustering and assign of individuals to groups by minimizing within-group variation (Jombart et al. 2010). DAPC has been empirically demonstrated to be better at detecting fine- scale clustering than STRUCTURE and, because it does not require a specified population genetics model for input, does not require loci to be in HWE (Jombart et al. 2010). Prior to implementing DAPC, we used the *K*-means clustering via the 'find.clusters' function to identify the most likely number of genetic clusters in our dataset without a priori population information, retaining all PCs, and selecting the number of clusters based on the Bayesian Information Criterion (BIC) value. For DAPC analysis, retaining a suboptimal number of PCs can cause model under- or overfitting (Jombart et al. 2010). For both marker types, we used cross-validation with 1000 permutations to identify the optimal number of PCs as those that balanced the highest mean successful assignment of samples with the lowest mean squared error rate, and used this as input for the discriminant function analysis. We visualized results both without and with a priori genetic clusters to describe the relationships among putative populations. We used number of a priori clusters ($K = 4$) to assign individuals to the most likely population of origin and to identify individuals that deviated from their prior group assignment based on retained discriminant functions. We cross-validated results for both marker sets by masking 25% of individuals from each population (but using 50% of individuals from AZ for SNP loci, as we only had 4 samples) to use as a test dataset in a DAPC analysis. We then used the *predict. dapc* function with the rest of the data as a training dataset, and evaluated assignment of the training data via posterior membership probabilities. For SNPs, we further investigated population assignment testing using Monte Carlo cross-validation in the R package assignPOP (Chen et al. 2018) with resampling iterations of 50%, 70%, and 90% of individuals with 50%, 75%, and 100% of SNP loci, and with 100 replicates run for each proportion of individuals.

Estimates of genetic differentiation and clustering can be biased by uneven sample sizes among populations (Kalinow- ski 2011; Puechmaille 2016). To validate our genetic structuring results, we randomly subsampled individuals to match the smallest group sample size from both the microsatellite and SNP datasets (n = 10 and n = 4 per sampling region, respectively). For both marker types, we reran STRUCTU RE with sampling location included and omitted and reran DAPC with a priori and predicted population groups. We then compared results with those of the full datasets to investigate the effect of uneven sample sizes on our results.

Estimates of migration

We investigated contemporary gene flow between regions using BayesAss 3.0.4 (Wilson and Rannala 2003) for both markers with modifications for SNP analysis (Mussmann et al. 2019). BayesAss estimates recent (< 3 generations) unidirectional migration rate among populations as m_{ij}, with m as the proportion of population i resulting from gene flow from population j. We ran the program for 3×10^6 iterations, with a burn-in of 1×10^6 iterations and a sampling frequency of 2000. To evaluate convergence of MCMC values we per- formed the analysis 5 times using different starting seeds and compared the posterior mean parameter estimates for each run to check concordance. Delta values were adjusted with a goal of a 20–60% acceptance rate for each parameter (Wilson and Rannala 2003).

Results

Microsatellite Genotyping

Fourteen samples were excluded based on low concentration or failure to amplify (Table 1). No loci showed signs of significant LD. One microsatellite marker (IVGUD103) was monomorphic for all samples and was removed from further analysis. No loci showed evidence of scoring errors or evidence for large allele dropout, but 3 loci (AAT-27, AAC- 20, and RGB-27) exhibited homozygote excess. Although homozygote excess can indicate presence of null alleles, no loci showed consistent homozygote excess across all populations. We found no substantial differences for pairwise F_{ST} calculated with and without the null correction in FreeNA, indicating that null alleles were unlikely to have a substantial effect on our results. Additionally, no loci exhibited high frequency of null alleles across all sites, and none exceeded an estimated frequency of > 0.21 null alleles per locus per site (Table S2); therefore, we did not remove any loci based on null allele frequencies. When the 4 sampling regions were treated as separate populations, the only locus with significant HWE deviation across > 1 sampling location was Calbo-2 for both the VI and MA. We removed this locus from the analysis.

Table 1 Comparison of summary genetic diversity statistics (and 95% confidence intervals) for roseate terns from 4

sampling locations based on microsatellite and single-nucleotide polymorphic (SNP) markers

MA = Bird Island, Massachusetts; FL = Florida Keys; VI = Leduck Island, U.S. Virgin Islands; AZ = Praia Islet, Azores; AR = rarefied allelic richness; Ho = observed heterozygosity; He = expected heterozygosity

Marker type	**Site**	***n***	**A_R**	**H_o**	**H_e**
Microsatellites	MA	35	2.54	0.35 (0.21, 0.49)	0.38 (0.25, 0.52)
	FL	16	2.76	0.29 (0.19, 0.40)	0.38 (0.23, 0.53)
	VI	33	2.52	0.37 (0.23, 0.50)	0.40 (0.26, 0.54)
	AZ	10	2.75	0.36 (0.21, 0.49)	0.39 (0.25, 0.52)
SNPs	MA	24	1.35	0.19 (0.19, 0.20)	0.20 (0.19, 0.21)
	FL	13	1.38	0.21 (0.20, 0.21)	0.21 (0.20, 022)
	VI	30	1.38	0.21 (0.21, 0.22)	0.22 (0.21, 0.23)
	AZ	4	1.26	0.10 (0.09, 0.10)	0.16 (0.15, 0.17)

SNP genotyping

Thirty-seven samples were excluded based on low con- centration, failure during library preparation, or failure to sequence (Table 1). Assembly of the specimen used to create our reference (FL1843) resulted in 21,839 contigs with a mean length of 722 base pairs (bp), for a total of 4520 UCE loci representing 3,265,327 bp. Following sample alignment and filtering, we were left with a final SNP dataset of 3,385 total SNPs with an average per-site sequencing depth of 53 x (SD ± 37.91) per individual. With thinning by locus to remove loci out of linkage equilibrium and removal of 44 loci out of HWE, this resulted in 2043 SNPs analyzed, with a mean per-individual missingness of 0.06.

Population genetics

For microsatellites, we found no evidence for differences among sampling locations (Table 1) for either H_O (F_3 = 0.174, p = 0.60), H_e (F_3 = 0.24, p = 0.91), or A_R MA = Bird Island, Massachusetts; FL = Florida Keys; VI = Leduck Island, U.S. Virgin Islands; AZ = Praia Islet, Azores; A_R = rarefied allelic richness; H_O = observed heterozygosity; H_e = expected heterozygosity (F_3 = 0.18, p = 0.95). For SNPs, genetic diversity statistics SNPs (Table 2), indicating significant deviation from panmixia between all 3 populations. Differences in N_e between sampling regions largely corresponded to census population size, with a smaller effective population size for MA (N_e = 903.9; 95% CI: 935.1, 3730.3) versus VI (N_e = 1171.9; 95% CI: 1278.0, 5038.7).

Table 2 F_{ST} with 95% confidence intervals estimated from 10^4 bootstraps based on microsatellite (below diagonal) and SNP (above diagonal) markers

	MA	VI	FL	AZ
MA	0	**0.05 (0.04, 0.06)**	**0.06 (0.04, 0.07)**	**0.12 (0.03, 0.14)**
VI	**0.09 (0.05, 0.13)**	0	0.02 (0.00, 0.04)	**0.11 (0.04, 0.13)**
FL	**0.12 (0.07, 0.17)**	− 0.00 (− 0.03, 0.04)	0	**0.12 (0.04, 0.16)**
AZ	**0.16 (0.10, 0.23)**	**0.09 (0.04, 0.15)**	**0.09 (0.03, 0.17)**	0

Bolded values represent F_{ST} values with estimated 95% confidence intervals that do not overlap 0
MA = Bird Island, Massachusetts; FL = Florida Keys; VI = Leduck Island, U.S. Virgin Islands; AZ = Praia Islet, Azores

Results from BOTTLENECK did not agree between the two marker types. For microsatellites, no significant heterozygosity excess was detected in any region under either the SMM or TPM models, indicating absence of a bottleneck for all populations. Conversely, for SNPs, all regions exhibited significant heterozygosity excess at a level of $p < 0.001$ under the IAM model, indicating presence of genetic bottlenecking for all 4 populations sampled.

Population structure

We found strong evidence for population structuring differentiating MA and CAR. STRUCTURE results for both markers returned $K = 2$ as the most likely number of genetic clusters among our four sampling locations (Fig. S1, a–d) via the Evanno method, with ΔK values supported by lower values of (LnP(K)) at $K = 1$ for all 4 treatments. For micro- satellites, MA clustered separately from the other 3 sampling locations when location was included as a prior, and clustering was consistent with AZ as differentiated from MA at all levels of K (Fig. 2a). For SNPs, MA also clustered separately from FL and the VI, but AZ showed high proportions of admixture at all values of K except $K = 4$ (Fig. 2b). When location was not included as a prior, results did not differ for SNPs (Fig. S2b), but microsatellites showed greater admixture among all 4 sampling locations at all values of K (Fig. S2a), indicating that microsatellites require location included as a prior to increase statistical power.

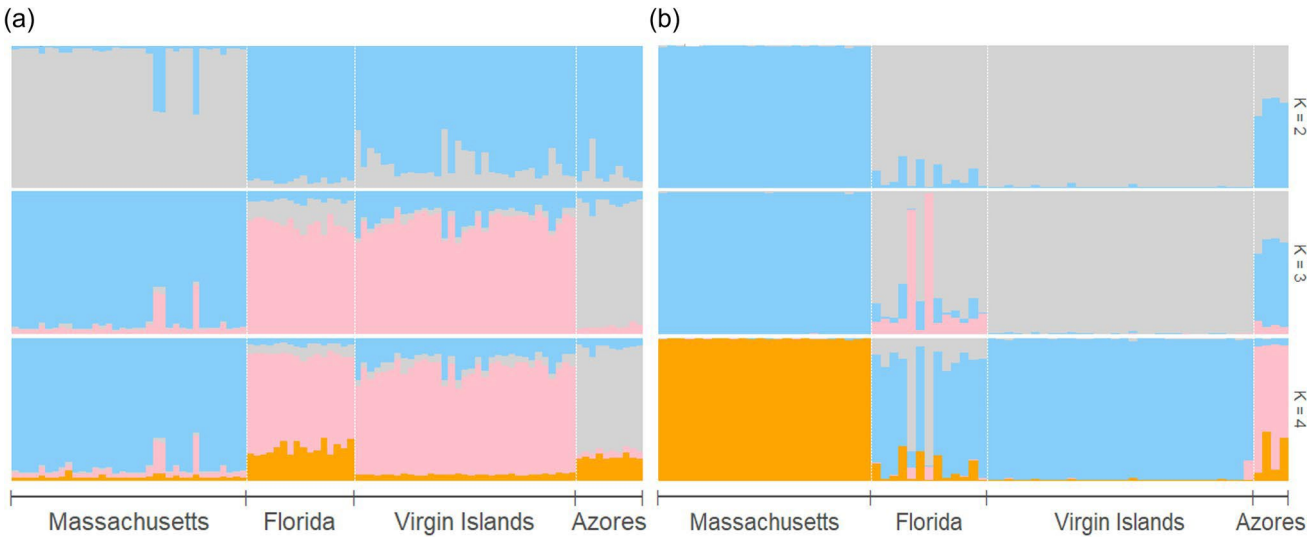

Fig. 2 Genetic clustering inferred by Bayesian structure analysis in STRUCTURE for K = 2–4 for **a** microsatellite and **b** single nucleotide polymorphism (SNP) loci for roseate terns from four sites, with the x-axis representing sampling location and the y-axis representing proportion of group membership, and with sampling location as prior information

Structure analysis with DAPC also indicated a high degree of genetic structuring between MA and CAR. Identification of optimal number of clusters with *K*-means clustering returned $K = 4$ for microsatellites, while for SNPs output from *K*-means clustering was consistent with STRUCTURE output, with 2 clusters identified as optimal. For DAPC by optimal clusters, 5 PCs were retained for microsatellites, representing 53% of conserved variance, and 15 PCs were retained for SNPs, representing 34% of conserved variance. For microsatellites, samples grouped optimally did not clearly cluster by sampling location (Fig. 3a), with mixed groupings of all 4 regions in each cluster of 1, 2, 3, and 4. For SNPs, samples were largely differentiated by region (Fig. 3b), with samples clustering as MA (cluster 1) versus FL and the VI (cluster 2), and with AZ split equally between clusters 1 and 2.

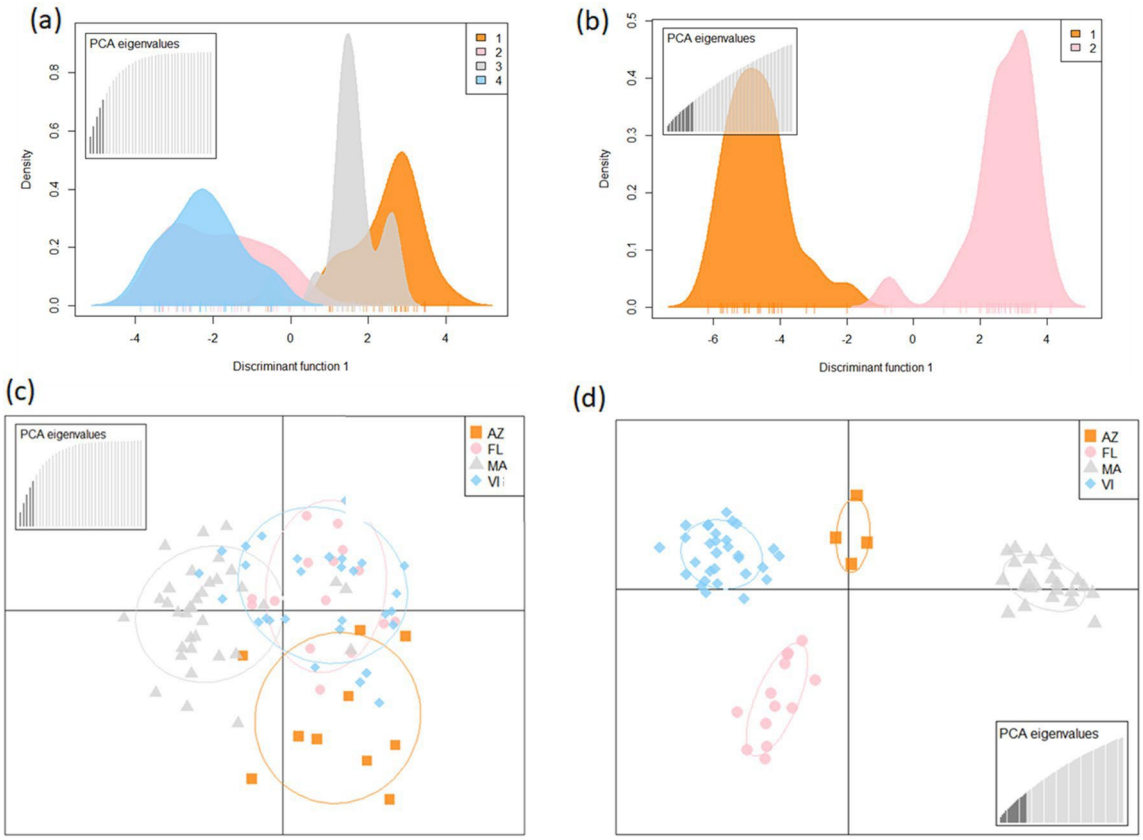

Fig. 3 Discriminant analysis of principal components (DAPC) of roseate terns from 4 sampling locations in the Atlantic Basin. Results are plotted on 1 discriminate function with grouping by optimal number of clusters for **a** microsatellite loci (with program-assigned groups 1, 2, 3, 4 consisting of mixed representatives from all 4 sampling regions) and **b** single-nucleotide polymorphism (SNP) loci (with pro- gram-assigned group 1 consisting of individuals from MA and group 2 of individuals from AZ, FL, and VI). Results are plotted on 2 discriminant functions with a priori population grouping for **c** microsatellite and **d** SNP loci, with shapes representing individuals, colors representing sample locations within ellipses designating grouping, and the x and y axes describing the first two discriminant functions. MA = Bird Island, Massachusetts; FL = Florida Keys; VI = Leduck Island, U.S. Virgin Islands; AZ = Praia Islet, Azores

When plotted as $K = 4$ with a priori population groups as input, microsatellites were differentiated slightly by region (Fig. 3c), whereas SNPs clearly clustered individuals by region (Fig. 3d). DAPC from microsatellites was able to correctly assign samples to sampling region 67% of the time, with assignment accuracy varying by population (AZ: 0.70, FL: 0.13, MA: 0.91, VI: 0.67; Fig. S3). By contrast, DAPC from SNPs correctly assigned samples to region 100% of the time. For microsatellites, cross-validation with a reduced dataset resulted in an assignment accuracy rate of 52%, with no population experiencing 100% accurate assignment (Fig. S4). Cross-validation of DAPC assignment for

SNPs accurately assigned samples to sampling region 84% of the time (Fig. S5). Assignment accuracy of the reduced dataset was lower for FL (0.85) and the VI (0.87), but 100% of samples were accurately assigned to their source population for AZ and MA. Results from Monte Carlo assignment testing based on SNPs were comparable to DAPC assignment testing for FL (0.67, with 0.32 assigning to VI), MA (0.99), and VI (0.97), but accuracy was low for AZ (0.55), with 0.30 of AZ samples assigning to VI, 0.06 to FL, and 0.06 to MA. Overall assignment accuracy increased with number of individuals included in the training dataset (Fig. S5).

Results for subsampled microsatellites were comparable to those run with the full dataset, with the reduced sample size set unable to detect genetic clustering when location was not included as a prior (Fig. S6a). With location included as a prior, microsatellite results returned a likely value of $K = 3$, with regions differentiated as MA, CAR (FL + VI), and AZ (Fig. S6b). Although STRUCTURE returned a $K = 2$ for the SNP subsampled set, which was comparable to the full set, the clustering differed from that for the full set, with AZ clustering separately from MA, FL, and VI regardless of whether location was included as a prior (Fig. S7a, b). We found a similar pattern for DAPC, with microsatellites indicating optimal number of clusters as $K = 3$ and with AZ and MA largely clustering separately from CAR (Fig. S8a–c), whereas we were only able to detect an optimal number of clusters as $K = 1$ for SNP data (Fig. S9a, b).

Estimates of migration

Consistent with the population structuring results, we did not find evidence for recent migration between MA and CAR, although results were indicative of movement within the Caribbean and potentially between VI and AZ. Bayesian estimates of gene flow were run with final delta values of
0.40 for migration rate and 0.90 for both allele frequency and inbreeding coefficient for microsatellites, and migration rate of 0.27, allele frequency 0.39, and inbreeding coefficient 0.13 for SNPs. BayesAss inferred unidirectional rates of gene flow greater than 0 from the VI into FL and AZ (Table 3). All other estimates of gene flow had confidence intervals that overlapped 0 (Table 3), indicating that gene flow was negligible between other sampling locations.

Discussion

We found evidence for substantial genetic structuring among roseate terns in the Atlantic Basin, with differentiation between all three regions supported by significant F_{ST}, low levels of admixture, and minimal evidence for migration between geographic regions. Results from both microsatellite and SNP markers showed significant differentiation between roseate terns breeding in MA and our two sampled locations in the Caribbean, concordant with the more limited results from our prior study (Byerly et al. 2022). This evidence for strong genetic differentiation between the two North American populations of breeding roseate terns provides support for their

designation as Evolutionary Significant Units (Moritz 1994). Given that both populations represent unique pools of genetic diversity, the loss of either would represent a considerable reduction in roseate tern genetic diversity on a global scale.

Results were mixed for the relationship of roseate terns breeding in the Azores to those in the rest of the Atlantic. Overall results were largely indicative of the presence of three genetic clusters in the Atlantic, with the CAR and AZ populations clustering apart from MA. The STRUCTURE analysis of SNP data did provide evidence for admixture between the MA and AZ populations and $K = 2$, but this was likely due to biases resulting from uneven sampling. When we repeated analyses with subsampled datasets to correct for differences in sample sizes, both microsatellite and SNP data showed AZ as clustering separately from the other regions at all values above $K = 1$. Although we were only able to detect $K = 1$ clusters for both STRUCTURE and DAPC with the subsampled SNP data, inferences of genetic clustering can be unreliable with sample sizes smaller than $n = 5$ (Fogelqvist et al. 2010), which likely limited the suitability of this approach. Regardless of this limitation, subsampled results more likely reflect the actual degree of structuring between AZ and the other regions, with AZ differentiated from MA.

Despite our findings of genetic differentiation between AZ and the rest of the sampling regions, estimates of migration from both microsatellites and SNPs were indicative of significant, unidirectional movement from the VI into AZ. Accuracy of Bayesian migration rate estimates increase with both sample size and number of markers (Wilson and Rannala 2003) but decrease when genetic differentiation is high ($F_{ST} \geq 0.10$; Faubet et al. 2007). The small sample size, high F_{ST}, and wide confidence intervals for migration as estimated for both SNP and microsatellite markers indicate that our BayesAss results may overestimate migration rates into AZ (Samarasin et al. 2017). However, while the actual proportion of Caribbean migrants within the AZ population is not well resolved based on our data, our findings do indicate that detectable levels of migration from the Caribbean into the Azores are likely to have occurred within the past three generations. Roseate terns breeding in the Azores are subject to high rates of nest predation from invasive species, and reproductive success can be low and highly variable (Neves 2006). Unidirectional migration rates from the VI to AZ may signal movement from a source population to a sink, with asymmetrical movement of dispersing individuals from higher productivity sites towards a site that does not provide reciprocal migrants (Pulliam 1988). Our results are consistent with prior findings of high genetic differentiation between the Northwestern and AZ populations (Lashko 2004), but indicate the possible existence of low levels of genetic exchange across the Atlantic.

Table 3 Bayesian estimates of contemporary (≤ 3 generations) gene flow among 4 roseate tern breeding locations with 95% confidence intervals

	Marker	MA	VI	FL	AZ
MA	msats	**0.94 (0.89, 0.99)**	0.03 (− 0.01, 0.08)	0.02 (− 0.02, 0.06)	0.08 (− 0.01, 0.18)
VI	msats	0.04 (0.00, 0.09)	**0.94 (0.89, 1.00)**	**0.26 (0.20, 0.33)**	**0.14 (0.02, 0.26)**
FL	msats	0.01 (− 0.01, 0.03)	0.01 (− 0.01, 0.03)	**0.69 (0.64, 0.74)**	0.03 (− 0.02, 0.08)
AZ	msats	0.01 (− 0.01, 0.02)	0.02 (− 0.01, 0.04)	0.02 (− 0.01, 0.05)	**0.75 (0.65, 0.84)**
MA	SNPs	**0.96 (0.95, 0.99)**	0.01 (0.00, 0.02)	0.01 (0.00, 0.02)	0.04 (0.00, 0.08)
VI	SNPs	0.01 (0.00, 0.02)	**0.97 (0.95, 0.99)**	**0.25 (0.22, 0.29)**	**0.08 (0.04, 0.13)**
FL	SNPs	0.01 (0.02, 0.05)	0.01 (0.00, 0.02)	**0.71 (0.68, 0.74)**	0.04 (0.00, 0.08)
AZ	SNPs	0.01 (0.00, 0.03)	0.01 (0.00, 0.02)	0.02 (0.00, 0.04)	**0.83 (0.78, 0.89)**

Bolded values represent migration rates with confidence intervals that do not overlap 0, with direction of dispersal moving from left to right in the table

MA = Bird Island, Massachusetts; FL = Florida Keys; VI = Leduck Island, U.S. Virgin Islands; AZ = Praia Islet, Azores

Within the Caribbean, we did not find evidence for significant differentiation between roseate terns in the VI and FL. Estimates of contemporary gene flow between these two populations indicate that this may be due to unidirectional dispersal from the VI into FL. However, these results were likely influenced by our inability to sample other breeding populations in the Caribbean (Beerli 2004). In the North- western Atlantic, roseate tern adults are known to move between sites within the region (Spendelow et al. 2010), and appear to follow a similar pattern of movement in the Caribbean. In the northern Caribbean, breeding colonies of roseate terns are scattered across the Bahamas and Cuba, but little is known about their abundance in these regions (Gochfeld and Burger 2020). It is probable that these more proximal unsampled "ghost" subpopulations within the Caribbean also share migrants with FL and the VI, and are the more likely sources of migrants into FL. As for AZ, this finding of movement into FL may be indicative of a source–sink relationship. Roseate terns in FL have declined considerably over the twentieth and twenty-first centuries and are nearing extirpation from the state (Gore et al. 2007; USFWS 2010). Reproductive success in FL appears to be lower than in other sites in Caribbean or the Northwestern Atlantic, as evidenced by lower hatch success and lower chick weight across growth stages (Zambrano 2007), with low reproductive success likely attributable to poor habitat quality (Zambrano et al. 2000). Based on this low productivity, it is unlikely that FL is sending migrants out to other breeding subpopulations that we did not sample, although further genetic sampling around the Caribbean could help resolve this. As roseate terns in FL continue to decline, our findings of movement into this region indicate that it may represent an ecological trap for the species in the Caribbean. Given evidence for ongoing migration into FL, preservation of existing roseate tern breeding sites and restoration for the creation of new breeding habitat would

be highly beneficial to the Caribbean population, and could help facilitate maintenance of the FL breeding population.

Genetic structuring between the Northwestern Atlantic and the other populations could be the product of philopatric behavior. Roseate tern site fidelity has not been well- documented, but regional philopatry is believed to be high (Gochfeld and Burger 2020), which could restrict dispersal between regions. However, breeding roseate terns are also somewhat nomadic in parts of their range (Tree 2005; Shealer et al. 2005; Spendelow et al. 2010) and are known to abandon colony sites of low quality (García-Quismondo et al. 2018). This relatively high intraregional movement indicates that site fidelity in roseate terns may not be as strong as it is in other seabirds that display genetic structure shaped by strong natal philopatry (Danckwerts et al. 2021). Lack of interbreeding among these populations could also be the result of assortative mating due to adaptive differences between the two breeding populations, as has been noted in other seabird species with overlap- ping winter distributions but differing breeding habitats (Lombal et al. 2017; Nunes and Bugoni 2018). Behavioral and morphological differences between the Northwestern Atlantic and Caribbean populations have been suggested, but not quantified (Nisbet and Ratcliffe 2008), and further work is needed to determine if these differences have a genetic basis. Diversity metrics recovered by both markers, as represented by allelic richness and observed heterozygosity, were low compared to super-abundant seabird species (Wojczulanis-Jakubas et al. 2015; Yannic et al. 2016; Danckwerts et al. 2021), but were largely consistent with studies of other breeding gulls and terns (Perez et al. 2020) and were comparable to past assessments of roseate tern genetic diversity (Lashko 2004; Dayton and Szczys 2021). We found evidence of population bottlenecks in all sampling locations via SNP, but not microsatellite, markers. These results are consistent with the finding of no evidence for bottlenecking in samples from Massachusetts analyzed with 16 microsatellite loci (Dayton and Szczys 2021). Evidence for bottlenecks via the SNP data suggests that differentiation between the two North American populations may have been exacerbated by genetic drift following past population declines, although our results could have been influenced by batch effects resulting from SNP filtering (Lou and Therkildsen 2021). Results from SNP markers indicated that all 4 sampling locations showed evidence of recent declines, consistent with known population declines in both the early and mid-twentieth century for the Northwestern population (USFWS 2020) and probable declines in the Florida and VI populations. Assessments of genetic diversity within our four sampling sites were mixed, but results from SNP markers, which are considered more reliable for evaluating diversity metrics (Fischer et al. 2017), indicating significantly higher observed and expected heterozygosity in the VI, the largest population sampled.

Although this finding of lower heterozygosity in the smaller populations is indicative of eroded genetic diversity resulting from population declines, the difference between MA and the two Caribbean sites, while statistically significant, is marginal, indicating comparable levels of genetic diversity between these regions. By contrast, the extremely low levels of genetic diversity for the Azores are of concern, especially given the ongoing low

productivity in the region. Given that even very rare instances of migration can influence allelic richness and heterozygosity in small and isolated populations (Gustafson et al. 2017), it is possible that low, undetectable rates of migration from the Caribbean into the Northwestern Atlantic, combined with increased recent productivity in the Northwestern Atlantic due to recovery efforts (Gochfeld and Burger 2020), have helped maintain genetic diversity in the population. Possible rare migration between regions in North America may also help explain why F_{ST} between the Northwestern Atlantic and the Caribbean was relatively low when compared to that of pairwise F_{ST} between the AZ and other regions, despite the clear evidence of genetic structuring. Another possibility is that roseate tern populations were historically connected in North America and then separated by vicariance events such as prior interglacial periods (Hewitt 2000), which could account for higher proportions of shared alleles between populations on the same sides of the Atlantic. Roseate terns are relatively long-lived (~ 20 years) and do not begin breeding until approximately 3 years of age, which would contribute to slower rates of allelic turnover between populations.

Conservation implications

Population overlaps on wintering grounds are largely unknown for roseate terns on both sides of the Atlantic, with previous studies relying on band resights of individuals roosting on beaches at night (Hays et al. 1999). Our results indicate that movement between wintering grounds may contribute to cross-oceanic dispersal in the Atlantic. A better understanding of roseate tern wintering ecology and migratory movements would contribute greatly to our understanding of dispersal among populations. The low success of tracking studies (Mostello et al. 2014; Paton et al. 2021) has previously made this challenging. Using population assignment to identify origin of individuals screened on migratory routes or wintering grounds has been investigated as a potential means of understanding overwintering dynamics in migratory birds, but fine scale assessment has largely been unsuccessful due to limited genetic structuring among populations (Connan et al. 2015; Gómez-Díaz et al. 2009) or use of uninformative genetic markers (Lovette et al. 2004). Higher-coverage genomic markers such as SNPs, which can better resolve fine-scale structure, can be used to more effectively link migratory birds to their natal regions (DeSaix et al. 2019). Given the high accuracy of population assignment with SNP markers, genetic analysis represents a promising tool for assigning individual roseate terns sampled on wintering grounds to population of origin, which can greatly expand our ability to understand population dynamics of non-breeding roseate terns in the Atlantic Basin.

The lack of detectable dispersal between the Northwest- ern Atlantic and Caribbean populations indicates that natural recolonization of either region is unlikely in the event of local extirpation. Similarly, dispersal from the Azores to the Western Atlantic appears to be rare, and past work has demonstrated high population differentiation between the United Kingdom (UK) and Azores population (Lashko 2004), indicating that

geographically segregated populations of breeding roseate terns in the Atlantic are largely isolated despite potential low levels of movement from the Caribbean to the Azores. This isolation has potential negative implications for their future persistence. If populations continue to decline, erosion of genetic diversity is unlikely to be offset by an influx of new alleles via migrants. Low diversity may compromise the adaptability of roseate terns in the face of changing climatic conditions, which is emerging as a primary threat to seabird persistence globally (Daunt and Mitchell 2013; Dias et al. 2019). Our study highlights the importance of continued conservation efforts for all breeding populations of roseate terns, with the ultimate goal of enhancing population sizes and encouraging recolonization of historically used breeding sites, thereby potentially enabling increased dispersal among breeding populations in the Atlantic Basin.

Acknowledgements Logistical support was provided by the U.S. Virgin Islands Division of Fish and Wildlife, the Florida Fish and Wild- life Conservation Commission, and the National Parks Service. We thank P. Szczys, C. Mostello, R. Zambrano, N. Warraich, C. Pavlik, D. Nellis, K. Kalasz, and V. Rodrigues Costa Neves for assisting with sampling and/or providing samples. We also thank Dr. Szczys and J. Dayton for their invaluable input for the microsatellite analysis and S. Hauser and N. Woodman for providing comments on the manuscript. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Author contributions PAB conceived and designed the study, collected samples, conducted laboratory work and data analysis, interpreted the data, and wrote the paper. RTC and RCF participated in study design, interpreting the data, and drafting the manuscript. NM participated in laboratory work. NASP participated in study design and data analysis. PLL participated in conceiving and designing the study, data analysis, interpretation of the data, and drafting the manuscript. All authors read and approved the final manuscript.

Funding This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1650114. Additional funding was provided by a Smithsonian short-term visitor award, the Ecosystems Mission Area of the U.S. Geological Survey, the David S. Lee Fund for the Conservation of Caribbean Birds, and United States Virgin Islands Division of Fish & Wildlife Cays Restoration & Research Grant VI-W-F17AF01314. Samples were collected and/or imported under appropriate United States Fish and Wildlife Service, United States Department of Agri- culture, and region-specific permits. All protocols were approved by the University of Louisiana at Lafayette Institutional Animal Care and Use Committee.

Data availability Raw sequence data were deposited in the NCBI Sequence Read Archive (PRJNA847596). Genetic data for this paper are also available through USGS ScienceBase at https://doi.org/10. 5066/P9MTXDCH (Byerly et

al. 2022). Related metadata, including individual sample identifiers with sampling region and corresponding SNP and microsatellite loci, as well as all code used to generate data, can be found at https://github.com/pabyerly/RoseateTern_Population Genetics.

Declarations

Competing interests The authors have no relevant financial or non- financial interests to disclose.

References

Barbraud C, Delord K (2020) Selection against immigrants in wild seabird populations. Ecol Lett. https://doi.org/10.1111/ele.13624

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. Mol Ecol 13:827–836

Bensch S, Hasselquist D, Bo N, Hansson B (1998) Higher fitness for philopatric than for immigrant males in a semi-isolated population of great reed warblers. Evolution 52:877–883

Bicknell AWJ, Knight ME, Bilton D, Reid JB, Burke T, Votier SC (2012) Population genetic structure and long-distance dispersal among seabird populations: implications for colony persistence. Mol Ecol 21:2863–2876

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics btu170

Brown JH, Kodric-Brown A (1977) Turnover rates in insular biogeography: effect of immigration on extinction. Ecology 58:445–449

Byerly PA, Chesser RT, Fleischer RC, McInerney N, Przelomska NAS, Leberg PL (2022) Museum genomics provide evidence for persistent genetic differentiation in a threatened seabird species in the western Atlantic. Integr Comp Biol. https://doi.org/10.1093/icb/icac107

Byerly PA, Chesser RT, Fleischer RC, McInerney N, Przelomska N, Leberg PL (2022) Conservation genomics reveals low connectivity among populations of threatened roseate terns (Sterna dougallii) in the Atlantic Basin. *USGS*. https://doi.org/10.5066/ P93VPW1V

Carøe C, Gopalakrishnan S, Vinner L, Mak SST, Sinding MHS, Samaniego JA, Wales N, Sicheritz-Pontén T, Gilbert MTP (2018) Single-tube library preparation for degraded DNA. Methods Ecol Evol 9:410–419

Chapuis M, Estoup A (2007) FreeNA manual. Mol Biol Evol 24:621–631

Chen KY, Marschall EA, Sovic MG, Fries AC, Gibbs HL, Ludsin SA (2018) assignPOP: an R package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. Methods Ecol Evol 9:439–446

Chesser RK (1991) Influence of gene flow and breeding tactics on gene diversity within populations. Genetics 129:573–583

Connan M, Teske PR, Tree AJ, Whittington PA, Mcquaid CD (2015) Subspecies assessment of Antarctic terns (*Sterna vittata*) over- wintering on the South African coast: evidence from morphology, genetics and stable isotopes. Emu 115:223–236

Cornuet J-M, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014

Cristofari R, Plaza P, Fernández CE, Trucchi E, Gouin N, Le Bohec C, Zavalaga C, Alfaro-Shigueto J, Luna-Jorquera G (2019) Unexpected population fragmentation in an endangered sea- bird: the case of the Peruvian diving-petrel. Sci Rep 9:2021

Danckwerts DK, Humeau L, Pinet P, McQuaid CD, Le Corre M (2021) Extreme philopatry and genetic diversification at unprecedented scales in a seabird. Sci Rep 11:1–12

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158

Daunt F, Mitchell I (2013). Impacts of climate change on seabirds. MCCIP Sci Rev 125–133

Dayton J, Szczys P (2021) Metapopulation connectivity retains genetic diversity following a historical bottleneck in a federally endangered seabird. Ornithol Appl 123:duab037. https:// doi.org/10.1093/ornithapp/duab037

DeSaix MG, Bulluck LP, Eckert AJ, Viverette CB, Boves TJ, Reese JA, Tonra CM, Dyer RJ (2019) Population assignment reveals low migratory connectivity in a weakly structured songbird. Mol Ecol 28:2122–2135

Dias MP, Martin R, Pearmain EJ, Burfield IJ, Small C, Phil- lips RA, Yates O, Lascelles B, Borboroglu PG, Croxall JP (2019) Threats to seabirds: a global assessment. Biol Cons 237:525–537

Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJB, Collen B (2014) Defaunation in the anthropocene. Science 345:401–406

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

England PR, Cornuet JM, Berthier P, Tallmon DA, Luikart G (2006) Estimating effective population size from linkage disequilibrium: severe bias in small samples. Conserv Genet 7:303–308

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Everson KM, McLaughlin JF, Cato IA, Evans MM, Gastaldi AR, Mills KK, Shink KG, Wilbur SM, Winker K (2019) Speciation, gene flow, and seasonal migration in *Catharus* thrushes (Aves: Turdidae). Mol Phylogenet Evol 139:106564

Faircloth BC (2016) PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32:786–788

Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst Biol 61:717–726

Faria PJ, Baus E, Morgante JS, Bruford MW (2007) Challenges and prospects of population genetic studies in terns (Charadriiformes, Aves). Genet Mol Biol 30:681–689

Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. Mol Ecol 16:1149–1166

Fischer MC, Rellstab C, Leuzinger M, Roumet M, Gugerli F, Shimizu KK, Holderegger R, Widmer A (2017) Estimating genomic diversity and population differentiation—an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. BMC Genomics 18:1–15

Fogelqvist J, Niittyvuopio A, Ågren J, Savolainen O, Lascoux M (2010) Cryptic population genetic structure: the number of inferred clusters depends on sample size. Mol Ecol Resour 10:314–323

Francis RM (2016) POPHELPER: an R package and web app to analyse and visualize population structure. Mol Ecol Resour 17:27–32

Funk WC, Forester BR, Converse SJ, Darst C, Morey S (2018) Improving conservation policy with genomics: a guide to integrating adaptive potential into U.S. Endangered Species Act: decisions for conservation practitioners and geneticists. Conserv Genet 20:115–134

García-Quismondo M, Nisbet ICT, Mostello C, Reed JM (2018) Modeling population dynamics of roseate terns (*Sterna dougallii*) in the Northwest Atlantic Ocean. Ecol Model 368:298–311

Genovart M, Oro D, Bonhomme F (2003) Genetic and morphological differentiation between the two largest breeding colonies of Audouin's gull *Larus audouinii*. Ibis 145:448–456

Gochfeld M, Burger J (2020) Roseate Tern (*Sterna dougallii*), version 1.0. In: Billerman SM (ed) Birds of the world. Cornell Lab of Ornithology, Ithaca. https://doi.org/10.2173/bow.roster.01

Gómez-Díaz E, González-Solís J, Peinado MA (2009) Population structure in a highly pelagic seabird, the Cory's shearwater *Calonectris diomedea*: an examination of genetics, morphology and ecology. Mar Ecol Prog Ser 382:197–209

Gore JA, Hovis JA, Sprandel GL, Douglass NJ (2007) Distribution and abundance of breeding seabirds along the coast of Florida, 1998–2000. Final Performance Report. Florida Fish and Wildlife Conservation Commission, Tallahassee

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, Di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full- length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652

Gustafson KD, Vickers TW, Boyce WM, Ernest HB (2017) A single migrant enhances the genetic diversity of an inbred puma population. R Soc Open Sci 4:170115

Hailer F, Schreiber EA, Miller JM, Levin II, Parker PG, Chesser RT, Fleischer RC (2011) Long-term isolation of a highly mobile sea- bird on the Galapagos. Proc Biol Sci R Soc 278:817–825

Hays H, Lima PC, Monteiro L, DiCostanzo J, Cormons GD, Nisbet ICT, Saliva JE, Spendelow JA, Burger J, Pierce J, Gochfeld M (1999) A nonbreeding concentration of roseate and common terns in Bahia, Brazil. J Field Ornithol 70:455–464

Hays H, Neves VC, Lima P (2002) Banded roseate terns from different continents trapped in the Azores. J Field Ornithol 73:180–184

Herman RW, Winger BM, Dittmann DL, Harvey MG (2022) Fine- scale population genetic structure and barriers to gene flow in a widespread seabird (*Ardenna pacifica*). Biol J Linnean Soc 1–12

Hess GR (1996) Linking extinction to connectivity and habitat destruction in metapopulation. Am Nat 148:226–236

Hewitt G (2000) The genetic legacy of the quaternary ice ages. Nature 405:907–913

Hoban S, Bruford M, D'Urban Jackson J, Lopes-Fernandes M, Heuertz M, Hohenlohe PA, Paz-Vinas I, Sjögren-Gulve P, Segelbacher G, Vernesi C, Aitken S, Bertola LD, Bloomer P, Breed M, Rodríguez-Correa H, Funk WC, Grueber CE, Hunter ME, Jaffe R, Liggins L, Mergeay J, Moharrek F, O'Brien D, Ogden R, Palma- Silva C, Pierson J, Ramakrishnan U, Simo-Droissart M, Tani N, Waits L, Laikre L (2020) Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. Biol Cons 248:108654

Janes JK, MillerDupuis JMJR, Malenfant RM, Gorrell JC, Culling- ham CI, Andrew RL (2017) The *K* = 2 conundrum. Mol Ecol 26:3594–3602

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

Kalinowski ST (2011) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity 106:625–632

Keenan K, Mcginnity P, Cross TF, Crozier WW, Prodöhl PA (2013) DiveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 4:782–788

Kersten O, Star B, Leigh DM, Anker-Nilssen T, Strøm H, Danielsen J, Descamps S, Erikstad KE, Fitzsimmons MG, Fort J, Hansen ES, Harris MP, Irestedt M, Kleven O, Mallory ML, Jakobsen KS, Boessenkool S (2021) Complex population structure of the Atlantic puffin revealed by whole genome analyses. Commun Biol 4:1–12

Lashko A (2004) Population genetic relationships in the roseate tern: globally, regionally and locally. PhD thesis, James Cook University, Australia

Leberg P (2005) Genetic approaches for estimating the effective size of populations. J Wildl Manag 69:1385–1399

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079

Lombal AJ, Wenner TJ, Lavers JL, Austin JJ, Woehler EJ, Hutton I, Burridge CP (2017) Genetic divergence between colonies of flesh-footed shearwater *Ardenna carneipes* exhibiting different foraging strategies. Conserv Genet 19:27–41

Lombal AJ, O'dwyer JE, Friesen V, Woehler EJ, Burridge CP (2020) Identifying mechanisms of genetic differentiation among populations in vagile species: historical factors dominate genetic differentiation in seabirds. Biol Rev 95:625–651

Lou RN, Therkildsen NO (2021) Batch effects in population genomic studies with low-coverage whole genome sequencing data: causes, detection and mitigation. Mol Ecol Resour 22:1678–1692

Lovette IJ, Clegg SM, Smith TB (2004) Limited utility of mtDNA markers for determining connectivity among breeding and overwintering locations in three Neotropical migrant birds. Conserv Biol 18:156–166

Mariano-Jelicich R, Madrid E (2014) Microsatellite variability among black skimmer (*Rynchops niger intercedens*) populations in Southern South America. Waterbirds 37:175–182

McCauley DJ, Pinsky ML, Palumbi SR, Estes JA, Joyce FH, Warner RR (2015) Marine defaunation: animal loss in the global ocean. Science 347:1255641

Moritz C (1994) Defining 'Evolutionarily Significant Units' for conservation. Tree 9:373–375

Mostello CS, Nisbet ICT, Oswald SA, Fox JW (2014) Non-breeding season movements of six North American roseate terns *Sterna dougallii* tracked with geolocators. Seabird 27:1–21

Mussmann SM, Douglas MR, Chafin TK, Douglas ME (2019) BA3-SNPs: contemporary migration reconfigured in Bayes- Ass for next-generation sequence data. Methods Ecol Evol 10:1808–1813

Neves V (2006) Towards a conservation strategy of the roseate tern *Sterna dougallii* in the Azores Archipelago. PhD thesis, University of Glasgow, Scotland

Nisbet ICT, Cabot D (1995) Transatlantic recovery of a ringed roseate tern *Sterna dougallii*. Ringing Migr 16:14–15

Nisbet ICT, Ratcliffe N (2008) Comparative demographics of tropical and temperate roseate terns. Waterbirds 31:346–356

Nunes GT, Bugoni L (2018) Local adaptation drives population isolation in a tropical seabird. J Biogeogr 45:332–341

Nuss A, Carlos CJ, Moreno IB, Fagundes NJR (2016) Population genetic structure of the magnificent frigatebird *Fregata magnificens* (Aves, Suliformes) breeding colonies in the Western Atlantic ocean. PLoS ONE 11:1–15

Paradis E (2010) pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 26:419–420

Parker L, Hawkins M, Camacho-Sanchez M, Campana M, West-Roberts JA, Wilbert T, Lim HC, Rockwood L, Leonard J (2020) Little genetic structure in a Bornean endemic small mammal across a steep ecological gradient. Mol Ecol 29:4074–4090

Paton PWC, Loring PH, Cormons GD, Meyer KD, Williams S, Welch LJ (2021) Fate of common (*Sterna hirundo*) and roseate terns (*S. dougallii*) with satellite transmitters attached with backpack harnesses. Waterbirds 43:342–347

Pauls SU, Nowak C, Bálint M, Pfenninger M (2013) The impact of global climate change on genetic diversity within populations and species. Mol Ecol 22:925–946

Peery MZ, Hall LA, Sellas A, Beissinger SR, Moritz C, Bérubé M, Raphael MG, Nelson SK, Golightly RT, Mcfarlane-Tranquilla L, Newman S, Palsbøll PJ (2010) Genetic analyses of historic and modern marbled murrelets suggest decoupling of migration and gene flow after habitat fragmentation. Proc R Soc B 277:697–706

Perez GS, Goodenough KS, Horn MH, Patton RT, Ruiz EA, Velarde E, Aguilar A (2020) High connectivity among breeding populations of the elegant tern (*Thalasseus elegans*) in Mexico and Southern California revealed through population genomic analysis. Water- birds 43:17–27

Peterson DA, Hilborn R, Hauser L (2014) Local adaptation limits life- time reproductive success of dispersers in a wild salmon metapopulation. Nat Commun 5:3696

Pierce J (2009) United States Virgin Islands. In: Bradley PE, Norton RL (eds) An inventory of breeding seabirds in the Caribbean. University Press of Florida, Gainsville, pp 99–111

Pierson JC, Coates DJ, Oostermeijer JGB, Beissinger SR, Bragg JG, Sunnucks P, Schumaker NH, Young AG (2016) Genetic factors in threatened species recovery plans on three continents. Front Ecol Environ 14:433–440

Pritchard JK (2000) Inference of population structure using multi-locus genotypes. Genetics 155:945–959

Puechmaille SJ (2016) The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Mol Ecol Resour 16:608–627

Pulliam HR (1988) Sources, sinks, and population regulation. Am Nat 132:652–661

Quillfeldt P, Moodley Y, Weimerskirch H, Cherel Y, Delord K, Phil- lips RA, Navarro J, Calderón L, Masello JF (2017) Does genetic structure reflect differences in non-breeding movements? A case study in small, highly mobile seabirds. BMC Evol Biol 17:160

Ralls K, Ballou JD, Dudash MR, Eldridge MDB, Fenster CB, Lacy RC, Sunnucks P, Frankham R (2018) Call for a paradigm shift in the genetic management of fragmented populations. Conserv Lett 11:1–6

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. Conserv Biol 17:230–237

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Samarasin P, Shuter BJ, Wright SI, Rodd FH (2017) The problem of estimating recent genetic connectivity in a changing world. Con- serv Biol 31:126–135

Shealer DA, Saliva JE (1992) Northeastern roseate terns seen at Puerto Rican colony during breeding season. Colon Waterbirds 15:152–154

Shealer DA, Saliva JE, Pierce J (2005) Annual survival and movement patterns of roseate terns breeding in Puerto Rico and the U.S., Virgin Islands. Waterbirds 28:79–86

Spendelow JA, Mostello CS, Nisbet ICT, Hall CS, Welch L (2010) Interregional breeding dispersal of adult roseate terns. Water- birds 33:242–245

Sruoga A, Butkauskas D (2006) Evaluation of the genetic structure of the breeding common tern (*Sterna hirundo*)

population by means of microsatellite markers. Biologija 1:47–52

Szczys P, Hughes CR, Kesseli RV (2005) Novel microsatellite markers used to determine the population genetic structure of the endangered roseate tern, *Sterna dougallii*, in Northwest Atlantic and Western Australia. Conserv Genet 6:461–466

Szczys P, Oswald SA, Arnold JM (2017) Conservation implications of long-distance migration routes: regional metapopulation structure, asymmetrical dispersal, and population declines. Biol Conserv 209:263–272

Tigano A, Damus M, Birt TP, Morris-Pocock JA, Artukhin YB, Friesen VL (2015) The Arctic: glacial refugium or area of secondary contact? Inference from the population genetic structure of the thick-billed murre (*Uria lomvia*), with implications for management. J Hered 106:238–246

Tree AJ (2005) The known history and movements of the roseate tern *Sterna dougalii* in South Africa and the Western Indian Ocean. Mar Ornithol 33:41–47

Tsai WLE, Mota-Vargas C, Rojas-Soto O, Bhowmik R, Liang EY, Maley JM, Zarza E, McCormack JE (2019) Museum genomics reveals the speciation history of dendrortyx wood-partridges in the Mesoamerican highlands. Mol Phylogenet Evol 136:29–34

U.S. Fish and Wildlife Service (2010) Caribbean roseate tern and North Atlantic roseate tern (*Sterna dougallii dougallii),* 5-year review: summary and evaluation. U.S. Fish and Wildlife Service, Boquerón, PR, and Concord, NH

U.S. Fish and Wildlife Service (2020) Roseate tern northeastern North American population (*Sterna dougallii dougallii*), 5-year review: summary and evaluation. U.S. Fish and Wildlife Service, Concord

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evol Appl 3:244–262

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

Welch AJ, Fleischer RC, James HF, Wiley AE, Ostrom PH, Adams J, Duvall F, Holmes N, Hu D, Penniman J, Swindle KA (2012) Population divergence and gene flow in an endangered and highly mobile seabird. Heredity 109:19–28

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

Winker K, Glenn TC, Faircloth BC (2018) Ultraconserved elements (UCEs) illuminate the population genomics of a recent, high- latitude avian speciation event. PeerJ 6:e5735–e5735

Wojczulanis-Jakubas K, Kilikowska A, Fort J, Gavrilo M, Jakubas D, Friesen V (2015) No evidence of divergence at neutral genetic markers between the two morphologically different subspecies of the most numerous Arctic seabird. Ibis 157:787–797

Yannic G, Yearsley JM, Sermier R, Dufresnes C, Gilg O, Aebischer A, Gavrilo MV, Strøm H, Mallory ML, Guy

Morrison RI, Gilchrist HG, Broquet T (2016) High connectivity in a long-lived high-Arctic seabird, the ivory gull *Pagophila eburnea*. Polar Biol 39:221–236

Zambrano R (2007) Reproductive success and nestling growth at a roof and ground colony of roseate terns (*Sterna dougallii*) in Florida. Masters Thesis, Florida Atlantic University, Boca Raton, Florida Zambrano R, Smith HT, Robson M (2000) Summary of breeding roseate terns in the Florida Keys: 1974–1998. Florida Field Nat 28:64–68