Received: 14 October 2020 | Revised: 30 April 2021 | Accepted: 13 May 2021
DOI: 10.1111/jbi.14208

Journal of Biogeography WILEY

RESEARCH ARTICLE

Genomic data reveal the biogeographical and demographic history of *Ammospiza* sparrows in northeast tidal marshes

Jennifer Walsh[1,2] | **Adrienne I. Kovach**[3] | **Phred M. Benham**[4] | **Gemma V. Clucas**[1,5] | **Virginia L. Winder**[6] | **Irby J. Lovette**[1,2]

[1]Fuller Evolutionary Biology Program, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA
[2]Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA
[3]Department of Natural Resources and the Environment, University of New Hampshire, Durham, NH, USA
[4]Museum of Vertebrate Zoology, University of California, Berkeley, CA, USA
[5]Atkinson Center for a Sustainable Future, Cornell University, Ithaca, NY, USA
[6]Department of Biology, Benedictine College, Atchison, KS, USA

Correspondence
Jennifer Walsh, Fuller Evolutionary Biology Program, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA.
Email: jennifer.walsh.emond@gmail.com

Funding information
Fuller Evolutionary Biology Program; US National Science Foundation Postdoctoral Research Fellowship in Biology; Cornell University Center for Vertebrate Genomics

Handling Editor: Camila Ribas

Abstract

Aim: Shaped by climate change and sea-level rise, tidal salt marshes represent ephemeral systems that are home to only a few species. The dynamic ecological histories of these habitats, however, render it challenging to reconstruct the complete biogeographical histories of their endemic taxa. To better understand the evolutionary history of tidal marsh endemic sparrows, we employ genomic data to evaluate the timing and direction of tidal marsh colonization events and leverage these data to test a series of models describing the biogeographical history of our study taxa.

Location: North America.

Taxon: *Ammospiza* Sparrows.

Methods: We sampled 54 individuals from the *Ammospiza* clade from freshwater and saltwater habitats. We sequenced whole genomes and generated a dataset containing 21 million SNPs. Detailed phylogenomic analyses were conducted to identify relationships between tidal marsh and freshwater species. These phylogenies were then used to date divergence times and reconstruct the biogeographical history of the clade as well as identify historic signals of introgression.

Results: Phylogenies based on several million SNPs supported well-resolved clades that correspond to all previous species designations. Demographic modelling within *A. nelsoni* provided further resolution into the subspecies topology. Phylogenetic and biogeographical reconstructions support a series of saltwater to freshwater colonization events, with some endemic taxa exhibiting associations with tidal marsh habitat over longer evolutionary time-scales and some habitat transitions occurring as recently as 5,000 years ago.

Main Conclusions: Biogeographical and phylogenomic reconstructions support an original freshwater origin of the *Ammospiza* sparrows currently inhabiting tidal marshes in North America. We also found strong evidence for a saltwater to freshwater transition in *A. n. subvirgatus* as opposed to the long hypothesized freshwater origin of this group. We conclude that the increased resolution from genomic data paired with an integrative framework of phylogenomic, demographic and biogeographical reconstructions is important for resolving evolutionary histories in complex and ephemeral environments.

KEYWORDS
effective population size, incomplete lineage sorting, New World sparrows, phylogenomics, phylogeography, tidal marsh endemics

1 | INTRODUCTION

Shaped by both climate change and sea-level rise, tidal salt marshes—defined here as coastal ecosystems regularly flooded by salt water—represent ephemeral systems; their changing structure and spatial arrangement, in turn, driving biogeographical distributions of endemic tidal salt marsh taxa over evolutionary time-scales (Malamud-Roam et al., 2006). The geological age of present-day tidal marshes is one defining feature of these systems, as they have existed in their *current* locations for no more than 10,000 years (Malamud-Roam et al., 2006; Mitsch & Gosselink, 2000). The ephemeral nature of tidal marshes is another defining feature, with changes in eustatic sea level since the Last Glacial Maximum (LGM; ~21,000 years BP; Fairbanks, 1989; Ruddiman, 2001) totalling as much as 130–140 vertical meters, characterized by oscillating pulses of rapid sea-level change separated by slow change (Malamud-Roam et al., 2006; Ruddiman, 2001). In response to these fluctuating sea levels, taxa inhabiting dynamic coastal ecosystems adapted, migrated or went extinct over time (Malamud-Roam et al., 2006) potentially resulting in high lineage turnover in coastal areas on both spatial and temporal scales (Dolby et al., 2018). Yet, tidal marshes in North America, in particular, have also provided refugia for some endemic taxa throughout the Pleistocene while then having expanded rapidly since the LGM, allowing their more recent colonization by additional taxa (Chan et al., 2006; Greenberg et al., 2006).

Because this transition between freshwater and saline habitats poses huge adaptive challenges to potential colonizers, tidal marsh endemics have become exemplars of the mechanisms underlying local adaptation (Chan et al., 2006; Benham & Cheviron, 2020; Greenberg et al., 2006; Walsh, Clucas, et al., 2019; Walsh, Benham, et al., 2019; Mikles et al., 2020). Yet, the dynamic ecological histories and spatial complexities of these habitats render it challenging to reconstruct the complete biogeographical histories of their endemic taxa. The power provided by detailed phylogenomic and demographic analyses may therefore offer the essential baseline knowledge to improve our general understanding of divergence across these saline gradients.

Adaptive variation in avian taxa that occupy tidal marshes is well documented in both physiological and morphological traits, including larger bill size (Greenberg & Olsen, 2010; Grenier & Greenberg, 2006), modified kidney structure and modified drinking behaviors in response to salt water (Benham & Cheviron, 2020; Goldstein, 2006; Poulson, 1969), and darker plumage coloration (Greenberg & Droege, 1990; Luttrell et al., 2015) compared to counterparts occupying freshwater ecosystems. Due to the extreme adaptive challenges posed by tidal salt marshes, the number of salt marsh endemic taxa is sparse, with only 25 species (or species with recognized subspecies) of mammals, reptiles and breeding birds (Greenberg, 2006). Three of these species are New World sparrows (Passerellidae) in the genus *Ammospiza*, including *A. caudacuta* (Saltmarsh Sparrow) breeding in saltwater marshes from southern Maine to Virginia (Greenlaw & Rising, 1994) and *A. maritima* (Seaside Sparrow) breeding in salt and brackish marshes along the Atlantic and Gulf Coasts from northern Massachusetts to the Copano and Nueces Bays in southern Texas (Robbins, 1983). The third *Ammospiza* sparrow restricted to tidal habitats is a subspecies of the *A. nelsoni* (Nelson's Sparrow) group, *A. n. subvirgatus*, which breeds in fresh, brackish and saltwater marshes across the Canadian Maritimes to southern Maine. The other two subspecies in the *A. nelsoni* complex are found in freshwater marshes (including prairie marshes) and meadows in the continental interior from eastern British Columbia to central Manitoba and northern South Dakota (*A. n. nelsoni*) and around the James and Hudson bays (*A. n. alterus*). The final member of this genus, *A. leconteii* (LeConte's Sparrow), breeds in the freshwater marshes and wet grasslands in the central northern Great Plains region. Despite contrasting freshwater-tidal marsh breeding habitats, *A. caudacuta* and all three subspecies of *A. nelsoni* overwinter in tidal salt marshes across overlapping non-breeding ranges (Greenlaw & Woolfenden, 2007). The demographic and evolutionary history within this tidal marsh *Ammospiza* group remains unresolved, particularly as it relates to divergence and dispersal events in response to marsh loss, formation and migration since the LGM.

Three possible Pleistocene scenarios for the evolutionary history of *Ammospiza* sparrows have previously been proposed, differing in ancestral relationships and vicariance patterns (Beecher, 1955; Greenlaw, 1993; Figure 1). An initial hypothesis by Beecher (1955) involved dispersal and isolation events in the late Pleistocene, wherein *A. caudacuta* and *A. maritima* diverged from one another on the Atlantic coast. Under this scenario, *A. caudacuta* subsequently moved northwards along the coast, where it further splits into the northern *A. nelsoni*, which then splits into two inland groups and a third that remained coastal, *A. n. subvirgatus* (Figure 1a). In a second hypothesized scenario (Figure 1b), *A. caudacuta* and *A. maritima* diverged from one another along the Atlantic coast and *A. caudacuta* subsequently dispersed inland becoming isolated from the coastal populations and establishing *A. nelsoni* (Greenlaw, 1993). Following the retreat of the Wisconsin ice mass, the interior isolate spread back eastward, establishing *A. n. subvirgatus* along the Atlantic coast and coming into secondary contact with *A. caudacuta*. The last hypothesized scenario (Figure 1c) suggests that the ancestor to *A. leconteii* and the *nelsoni-caudacuta* group inhabited freshwater marshes and low wet prairies in the Interior Plains Region (Murray, 1969). This hypothesized scenario includes two dispersal events from the continental interior, the first leading to the establishment of *A. caudacuta* on the Atlantic coast and the second resulting in the current distribution of *A. n. subvirgatus* further north. Recent multilocus (four nuclear introns and mtDNA) reconstructions place *A. maritima* as sister to the *A. nelson-caudacuta* complex (Klicka et al., 2014), offering support for a shared *maritima* ancestor; however, these relationships have yet to be evaluated with genomic data. Moreover, the timing (as well as demographic impacts) of these freshwater-to-salt marsh colonization events and lineage divergence events remain unknown. This is particularly true regarding coastal versus interior ancestry and biogeographical history within *A. nelsoni*, which has been difficult to resolve given a lack of mtDNA structure (Rising & Avise, 1993) suggestive of recent subspecies origins.

In addition to many questions surrounding divergence times and biogeographical history within *Ammospiza* sparrows, contemporary introgression between *A. n. subvirgatus* and *A. caudacuta*

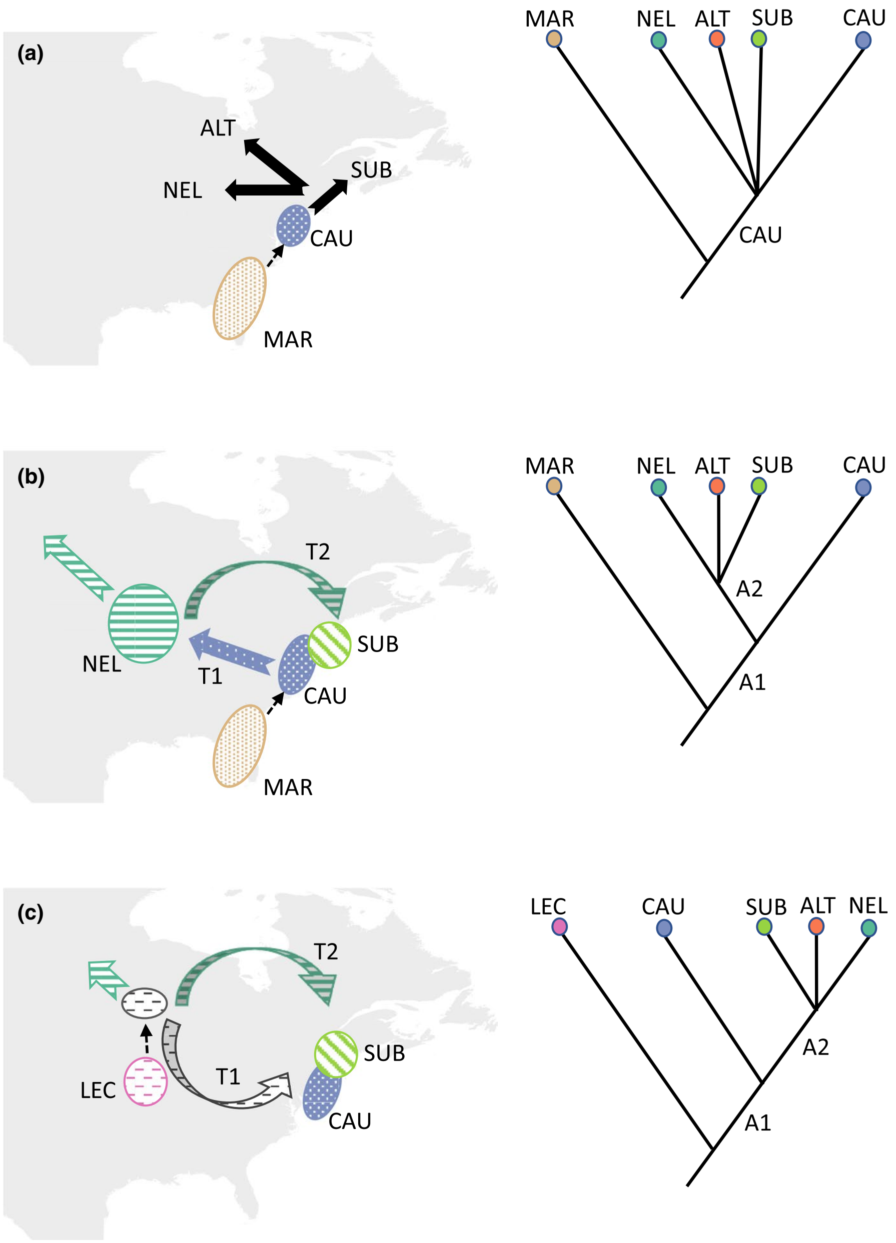

FIGURE 1 Three proposed models for the evolutionary history of *Ammospiza* sparrows in North America. (a) Beecher's (1955) late Pleistocene/Holocene model; (b) Greenlaw's (1993) Pleistocene model of vicariance involving two separate range expansions. This model assumes *A. maritima* to be sister species of *A. caudacuta*. (c) Greenlaw's (1993) alternate Pleistocene model of vicariance that assumes *A. leconteii* is sister species to *A. caudacuta*. Ellipses indicate relative geographical locations of some modern populations. Arrows represent dispersal events. Hypothesized phylogenetic relationships for each hypothesis are included. An original version of this figure was published in Greenlaw (1993) and has been adapted for this paper

may pose additional hurdles to accurate phylogenomic inference. These two taxa overlap and hybridize in tidal marshes along a 210 km stretch of the New England coast (Hodgman et al., 2002; Shriver et al., 2005; Walsh et al., 2011; Walsh, Rowe, et al., 2016), with recent work suggesting a role for adaptive divergence between them in traits related to tidal marsh adaptations (Walsh, Clucas, et al., 2019; Walsh et al., 2018). Disentangling patterns of contemporary introgression from signals of ancient hybridization within the *Ammospiza* group remains an important component of our general understanding of the evolutionary relationships within the clade. This is particularly relevant to our understanding of the frequency and chronology of saltwater–freshwater habitat transitions and the possible role of introgression in facilitating adaptation to these harsh environments.

Here, we employ whole-genome sequencing to investigate the phylogeny and demographic and biogeographical history within the group of tidal marsh *Ammospiza* sparrows. Our main objectives are to (i) reconstruct phylogenomic relationships within this group based on genome-wide SNPs; (ii) estimate divergence times between focal lineages; (iii) reconstruct the biogeographical history with a focus on freshwater–saltwater colonization events and (iv) assess the role of historical and contemporary introgression in shaping genome-wide topologies.

2 | MATERIALS AND METHODS

2.1 | Population sampling

We sampled 54 individuals from the four species of sparrows within the *Ammospiza* genus, including three subspecies of *A. nelsoni*: *Ammospiza nelsoni subvirgatus* (n = 8), *A. n. nelsoni* (n = 8), *A. n. alterus* (n = 10), *Ammospiza caudacuta* (n = 8), *Ammospiza maritima* (n = 10), *Ammospiza leconteii* (n = 4) and *Passerculus sandwichensis* (n = 6), which served as an outgroup (Supporting Information, Table S1). DNA was extracted from blood or tissue using the DNeasy blood and tissue kit (Qiagen).

2.2 | Whole-genome resequencing and variant discovery

We prepared individually indexed libraries using 1 μg of DNA following the TruSeq DNA PCR-free library preparation kit protocol, with an insert size of 350 bp. We pooled 24 libraries using concentrations of adapter-ligated DNA determined through digital PCR. The pooled libraries were sequenced on three Illumina NextSeq500 lanes at the Cornell Institute for Biotechnology core facility. Each lane contained 24 individuals, as some samples were run with samples for a different study. The quality of individual libraries was assessed using FastQC (version 0.11.5; http://www.bioinformatics.babraham.ac.uk/projects/fastqc).

Reads were aligned to the *A. caudacuta* reference genome (Walsh, Clucas, et al., 2019). We performed sequence trimming, adapter removal and quality filtering with AdapterRemoval version 2.1.1 (Lindgreen, 2012). We allowed a minimum threshold quality score of 10 (increased from the default value of 2) and merged overlapping paired-end reads. Filtered reads were aligned to the reference using bwa 0.7.4 (Li & Durbin, 2009) with default settings. Alignment statistics were obtained using qualimap version 2.1.1 (Okonechnikov et al., 2016).

BAM files were sorted and indexed using SAMtools version 1.3 (Li et al., 2009). We realigned around indels and fixed mate-pairs using GATK version 3.5 (McKenna et al., 2010). SNP variant discovery and genotyping for the 54 individuals was performed using BCFTools (-mpileup and -call). For an initial filtering step, variants that were out of Hardy–Weinberg equilibrium had a minor allele frequency less than 2%, had a mean depth <2 or greater than 50, and more than 20% missing data were filtered from the dataset using VCFtools version 0.1.14 (Danecek et al., 2011). This initial filtering step resulted in 21,622,990 SNPs. Basic descriptive statistics, including F_{ST} and nucleotide diversity, were estimated for this full dataset using VCFtools.

2.3 | Phylogenetic and demographic history estimation

Phylogenies were inferred using a filtered dataset, in which indels were treated as missing data and all invariable or ambiguous sites were removed, resulting in 4,776,571 SNPs. Maximum likelihood trees were inferred using a single-site coalescent-based method using SVDquartets (Chifman & Kubatko, 2014) as implemented in PAUP* version 4a166 (Chifman and Kubatko, 2014, 2015 Swofford, 2002). Coalescent-based methods are generally more robust when inferring accurate relationships under incomplete lineage sorting conditions (Chou et al., 2015). Node support was assessed with 100 bootstrap replicates with the remaining parameters left as default. We also conducted phylogenetic analyses using RAxML version 8.2.4 (Stamatakis, 2014) under the ASC_GTRGAMMA model with three parsimony and three random starting trees. We compared topologies from the two methods using the 'cophylo' function in the R package 'phytools' (Revell, 2012).

To further resolve the history of divergence among the three *A. nelsoni* subspecies while accounting for gene flow among populations, we compared a series of demographic models to the observed 3D site frequency spectrum (SFS) using the program *Moments* (Jouganous et al., 2017). For this analysis, we chose high-quality, biallelic SNPs that had a mean depth >6 and <20. We only retained SNPs with less than 10% missing data and we did not filter based on minor allele frequency. We further excluded all SNPs that were mapped to exons or intervals within 100 kb of exons to produce a final, putatively neutral dataset to reduce bias due to selection or linked selection (43,756 SNPs) and further filtered to exclude all missing data and retain only SNPs segregating within *A. nelsoni* populations. This left a final dataset of 18,022 SNPs. A series of 15 three-population demographic models were fit to the joint SFS to infer divergence times, effective population sizes and gene flow (Supporting Information; Figure S1). Five different models were fit to three different topologies with *A. n. nelsoni*, *A. n. alterus* or *A. n. subvirgatus* diverging from the other subspecies first. For each topology, we modelled different histories of gene flow among the three populations, including (1) no gene flow among populations, (2) gene flow among all populations throughout the history of divergence, (3) only recent gene flow among populations following both population splits, (4) gene flow after divergence, but only among geographically adjacent populations and (5) gene flow only between the two sister lineages. For all models, 12 optimizations were run from different starting parameters using the perturb function in *Moments* with a maximum number of iterations set to 20. To ensure that a global optimum for a given model had been reached, we ran one final optimization with 50 iterations using the parameter values estimated from the shorter optimization run with the highest likelihood. We calculated demographic parameter values using a generation time of 2.3 years (Walsh, Clucas, et al., 2019) and mutation rate for Passeriformes of 7.59×10^{-9} (Zhang et al., 2014). To calculate uncertainty around each parameter estimate, we generated 500 bootstrapped SFS using the bootstrap function within *Moments*. Bootstrapped spectra were then used to calculate parameter uncertainties using the Godambe information matrix (GIM) with log-transformed parameter values (Coffman et al., 2016).

Because changes in population sizes over time are indicative of past demographic events, we estimated changes in effective population sizes (N_e) among *Ammospiza* sparrows using two approaches. We first used G-PhoCS version 1.3 (Gronau et al., 2011), which uses the entire nucleotide sequence of a given locus to estimate an average N_e. For analyses in G-PhoCS, we used the filtered dataset generated above (43,756 SNPs; 2773 loci that were 100 bp in length) for demographic reconstructions. Due to high computational requirements, we randomly selected 1000 loci from this dataset and ran

G-PhoCS on 4 randomly selected individuals per species/subspecies (chosen to match the smallest sample size in the *A. leconteii* group). We ran the multi-threaded version of the program for 450,000 iterations with a 10% burn-in and the remaining parameters left as default. Once the run was completed, we visualized the traces from the parameter estimates in Tracer version 1.7.1 (Rambaut et al., 2018). We converted the median and 95% Bayesian credibility intervals for each parameter from mutation scale to individuals using a rough estimate of mutation rate of 10^{-9} per bp per generation as described in Campagna et al. (2019). We additionally used composite likelihood methods to estimate changes in historical population sizes over time using the program SMC^{++} version 1.8.0 (Terhorst et al., 2017). Because SMC^{++} requires a single input file per scaffold, we used only variants found on the N50 scaffolds as input for this analysis. We also created composite likelihoods based on three individuals per scaffold (this varies the identity of the distinguished individual to which all other samples are compared). To minimize false signals of recent bottlenecks, we removed contiguous stretches of homozygosity greater than 30 kbp. Each scaffold was converted into SMC^{++} format using the vcf2smc script and all input files were combined in a single run. Runs were conducted using a mutation rate of 7.59×10^{-9} based on the mean mutation rate of passeriforms (Zhang et al., 2014) and we used a generation time of 2.3 years (Walsh, Clucas, et al., 2019) to convert generations to years.

To estimate the timing of splitting events within the *Ammospiza* sparrows, we generated a time-calibrated tree using the multispecies coalescent approach in the program SNAPP (Bryant et al., 2012), implemented in BEAST2 v. 2.6.1 (Bouckaert et al., 2019). SNAPP integrates across all possible gene trees to infer the probability of the species tree using bi-allelic markers, such as SNPs, and does not estimate individual gene trees for each SNP. SNAPP traditionally estimates relative divergence times among lineages in coalescent units and we followed the approach of Stange et al. (2018) to estimate absolute divergence times among *Ammospiza* sparrow taxa. Briefly, this approach modifies the SNAPP settings to allow for the use of fossils to calibrate a strict molecular clock model. Second, run time of SNAPP is reduced by linking population sizes across all taxa rather than independent estimation of population sizes for each branch.

To generate an input SNP dataset for time calibration in SNAPP, we filtered the full SNP dataset in VCFtools to include a single representative from each species. For each species, we selected the individual with the least amount of missing data. We further filtered the data to exclude indels, remove sites with missing data and to ensure that SNPs were unlinked. For the latter, we used the thin function in VCFtools to only include SNPs at least 25Kb apart, as it exceeds the average decay of linkage disequilibrium estimated to be 10–17 Kb in Passerine birds (e.g. Balakrishnan & Edwards, 2009; Kawakami et al., 2014). The final dataset included 43,245 putatively unlinked SNPs. We converted the VCF file to a SNAPP xml input file using the ruby script, snapp_prep.rb (https://raw.githubusercontent.com/mmatschiner/). Given that our primary goal was to use SNAPP to estimate divergence times and not the species tree topology, we used the topology from RAxML and SVDquartets analyses, as well as the resolved topology within *A. nelsoni* from *Moments*, as a starting tree. We further constrained *P. sandwichensis* as an outgroup and *A. maritima* as sister to the *A. nelsoni-caudacuta* clade based on the above phylogenies. To calibrate the molecular clock, we used an early Pleistocene (1.6–2 million years old) fossil of *A. maritima* from Florida (Emslie, 1998). We used the lower age estimate of the fossil, 1.6 million years, to set a hard minimum divergence time between *A. maritima* and the *A. nelsoni-caudacuta* clade. We used a lognormal distribution to calibrate this node with an offset of 1.6 million years (representing the lower bound of the estimated fossil age), a mean of 1.5 and variance of 0.5. The 97.5th upper percentile of this distribution is 9.7 million years, which corresponds with the oldest known new world sparrow fossil, *Ammodramus hatcheri*, dated to 9–11 mya (Steadman, 1981) and is likely a reasonable upper bound on divergence times within *Ammospiza* sparrows. SNAPP analyses of the dataset were performed using BEAST 2.6.1 (Bouckaert et al., 2019) on the CIPRES science gateway (Miller et al., 2011). Four independent Markov Chain Monte Carlo (MCMC) chains were run for 30 million generations each and sampled every 10,000 steps. To ensure convergence and improved ESS values, we ran two additional chains for 70 million generations, sampling every 10,000 steps. We used Tracer to assess convergence and stationarity among chains independently and on logfiles combined using the LogCombiner tool within BEAST2 (Bouckaert et al., 2019). Trees from all independent runs were combined in LogCombiner with the first 10% of trees removed as burn-in. A maximum-clade credibility tree was then generated from the combined trees file using TreeAnnotator in BEAST2 and visualized using FigTree v. 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). A newick format tree was exported from FigTree for historical biogeography analyses.

2.4 | Historical biogeography inference

We estimated patterns of historical range evolution within the *Ammospiza* sparrows using the R package 'BioGeoBears'—Biogeography with Bayesian (and likelihood) evolutionary analysis in R scripts (Matzke, 2013). BioGeoBears reconstructs ancestral ranges based on the best fitting of three different biogeographical models within a likelihood framework. These models include (1) the dispersal–extinction–cladogenesis model (DEC; Ree & Smith, 2008); (2) the DIVALIKE model, a likelihood estimation of the dispersal–vicariance model (Ronquist, 1997) and (3) BAYAREALIKE, a likelihood implementation of the BayArea biogeographical model (Landis et al., 2013). The DEC and DIVALIKE models emphasize range changes at speciation events (cladogenesis), but differ in the overlap and distribution allowed between daughter and ancestral lineages. In DEC models, one daughter lineage will occupy a single region of the ancestral range (sympatrically or through vicariance) and in DIVALIKE models, a daughter lineage is allowed to inherit more than a single region (vicariant event) but does not overlap sympatrically with another daughter lineage (see Matzke, 2013 for model details). The BAYAREALIKE model instead emphasizes range changes between speciation events (anagenesis), BioGeoBears also includes a species jump parameter 'j', which accounts for founder speciation events, whereby a new lineage disperses outside of the ancestral range (Matzke, 2013). We compared the fit of each of the three models with and without the j parameter (for six models total) to the ingroup (i.e. minus *P. sandwichensis*), time-calibrated phylogeny of *Ammospiza* sparrows. Three biogeographical regions were defined: (A) northern Great Plains to Hudson Bay; (B) Gulf coast of the southern United States and (C) Atlantic coast of the eastern United States. The contemporary distribution of each terminal taxon was coded as present or absent across these three biogeographical areas. The maximum number of areas that could be occupied by ancestral species was set to three. The relative likelihoods of the fit of each model to the data were compared using Akaike's information criterion with small sample correction (AICc) to account for the small number of terminal taxa.

2.5 | Mosaic genomes and the role of introgression

To investigate the potential role of historical and contemporary introgression in shaping this species complex, we applied several methods to identify signals of contemporary gene flow. To characterize phylogeographical structure among the seven species/subspecies groups, ADMIXTURE v 1.2.3 (Alexander et al., 2009) analyses were run using a filtered dataset (49,638 SNPs) that contained no missing data and was pruned to avoid linkage using the script ldPruning.sh (https://github.com/speciationgenomics/scripts/blob/master/ldPruning.sh). We investigated one to seven possible population clusters, with 200 bootstrap resampling iterations.

We also tested for signals of introgression as indicated by discordant genealogies relative to the species tree. To explore the extent of such phylogenetic discordance across the genome, we used a topological weighting procedure implemented in the program TWISST (Martin & Van Belleghem, 2017). For this analysis, we removed invariant sites and indels and phased the data using the program BEAGLE (Browning & Browning, 2007). Local phylogenetic trees were estimated in sliding windows using PhyML with the GTR substitution model in non-overlapping windows of 100 SNPs across the genome. We used the 'complete' option in TWISST to calculate the exact weighting of each local window. For analyses in TWISST, we were interested in exploring discordant topologies among *A. leconteii*, *A. maritima*, *A. caudacuta* and *A. nelsoni*. We therefore kept *P. sandwichensis* as an outgroup but treated the three *nelsoni* subspecies as one group to reduce the number of topologies to be analysed.

Lastly, we estimated the D frequency spectrum (D_{FS}) to examine signals of introgression across different allele frequency classes following Martin and Amos (in review). D_{FS} is an extension to the ABBA BABA test D statistic, which quantifies discordant topologies based on ancestral and derived alleles. In an ABBA BABA test, equal frequencies of discordant topologies support a pattern of ILS, whereas significant deviations of D from zero in either direction indicate the direction of gene flow. D_{FS} characterizes this signal according to allele frequency bins as opposed to providing an average across all sites in the genome. Because shared derived alleles arising from introgression and incomplete lineage sorting may not be distributed evenly across allele frequencies, D_{FS} can provide insight into the age of introgression events (Martin & Amos, in review). Recent introgression events tend to result in positive D_{FS} among bins representing low-frequency derived alleles, whereas in older introgression events, positive D_{FS} is more evenly dispersed across the bins (Martin and Amos, 2021). As with the D statistic, D_{FS} is estimated based on the hypothesized topology $(((P_1,P_2,)P_3),O)$. Thus, to further examine the role of recent versus historic introgression events in shaping the *Ammospiza* complex, we estimated D_{FS} for several comparisons that were most relevant to the three biogeographical scenarios of interest: (((*A. n. subvirgatus*, *A. caudacuta*), *A. maritima*), *P. sandwichensis*); (((*A. n. subvirgatus*, *A. caudacuta*), *A. leconteii*), *P. sandwichensis*); (((*A. n. nelsoni*, *A. n. subvirgatus*), *A. caudacuta*), *P. sandwichensis*); (((*A. n. nelsoni*, *A. n. subvirgatus*), *A. maritima*), *P. sandwichensis*); and (((*A. n. nelsoni*, *A. n. subvirgatus*), *A. leconteii*), *P. sandwichensis*). For this analysis, we removed indels and only used SNPs >4 coverage.

3 | RESULTS

Whole-genome re-sequencing yielded an average of 28.8 million reads per individual, with all but one sample passing our quality filter. This individual (*A. n. alterus*) was removed from the dataset and subsequent results are reported for 53 individuals. The average alignment rate (percentage of paired reads mapped) with the reference *A. caudacuta* genome was 87.5% across all individuals (Supporting Information; Table S2). The depth of coverage averaged 3.9 X (range of coverage was 1.9–8 X). After filtering variant calls from BCFTOOLS, our final dataset contained 21,622,990 SNPs. The number of variable sites differed among groups, with the partition of variable sites as follows: 3,726,227 in *A. n. alterus*, 3,924,044 in *A. n. nelsoni*, 3,348,764 in *A. n. subvirgatus*, 6,376,757 in *A. leconteii*, 2,096,083 in *A. caudacuta*, 4,928,871 in *A. maritima* and 11,512,227 in *P. sandwichensis*. The average percent of missing data across individuals was 8.1%. Average nucleotide diversity (Supporting Information; Figure S2) was lowest in *A. caudacuta* (0.0005) and highest in *Ammospiza leconteii* (0.002) and *Passerculus sandwichensis* (0.003).

3.1 | Phylogenetic and demographic history

SVDquartets generated a phylogenetic tree with 100% bootstrap support (Figure 2). Using *P. sandwichensis* to root the phylogeny, we detected support for four, well-resolved clades corresponding to *P. sandwichensis*, *A. leconteii*, *A. caudacuta* and *A. maritima*. Relationships within *A. nelsoni* were less clear, as the tree did not recover any of the three subspecies as monophyletic (Figure 2). Based on the SVDquartets tree and consistent with previous work, *A. caudacuta* and *A. nelsoni* are sister taxa. Our whole-genome data further show *A. maritima* (not *A. leconteii*, as a prior alternate hypothesis) as the sister to the *A. caudacuta* + *A. nelsoni* clade (Figure 2). RAxML also generated a phylogenetic hypothesis that was consistent among all six runs (Supporting Information; Figure S3). The topology obtained from the RAxML analysis was similar (Supporting Information; Figures S3 and S4), confirming support for the split of *caudacuta/nelsoni* from *maritima* as opposed to *leconteii*. Similarly to the SVDquartets analysis, relationships within *A. nelsoni* were less clear. In the RAxML tree, *A. n. alterus* does not form a single clade, and is paraphyletic relative to *A. n. nelsoni*.

Results from *Moments* provided some additional resolution into the relationships within *A. nelsoni*. The best fit model of divergence and gene flow among the different *A. nelsoni* subspecies supported a divergence history where *A. n. subvirgatus* first split from inland taxa approximately 459,228 (95% CI: 450,922–457,429) years ago (Supporting Information; Table S3, Figure S5). This was followed by divergence between *A. n. nelsoni* and *A. n. alterus* approximately 5,056 years ago (95% CI: 1,225–8,897). Since divergence, gene flow was greatest from *A. n. nelsoni* populations into *A. n. alterus* populations and second from *A. n. subvirgatus* into *A. n. alterus*. Gene flow from *A. n. alterus* into other populations was relatively minimal.

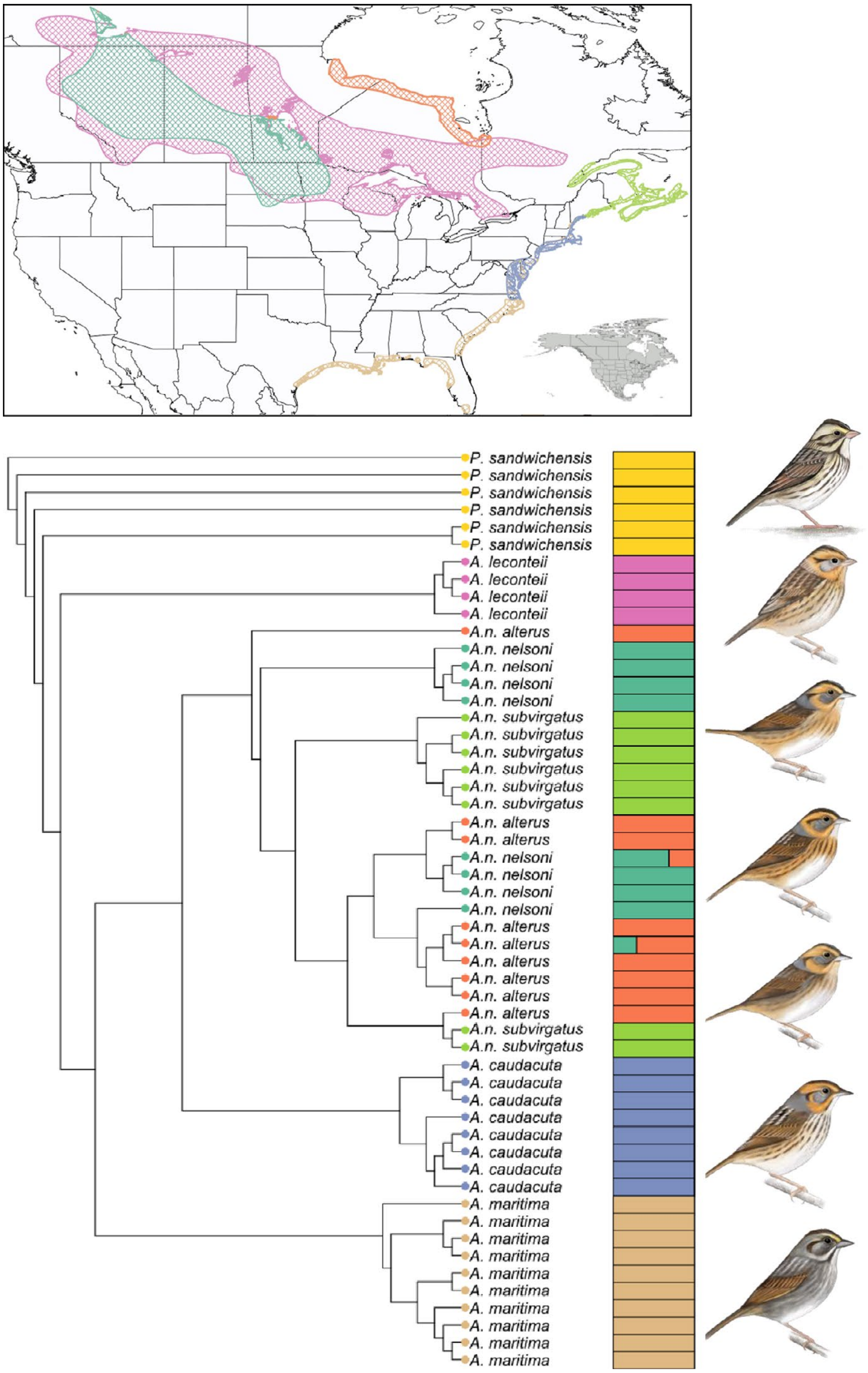

FIGURE 2 Current distribution and phylogeny of *Ammospiza* sparrows. Range map shows current distribution of *Ammospiza* sparrows across North America. The outgroup, *P. sandwichensis*, is not included on the map but is distributed broadly across North America. Colours correspond to those in the phylogeny and the admixture plot. Tree is the maximum likelihood phylogeny estimated in SVDquartets using 4.7 million SNPs. Bootstrap support was 100% for all nodes. Admixture plot shows assignment of each individual in the tree to a genetic cluster. All sparrow illustrations done by Jillian Ditner, 2020

While the divergence time estimated by *Moments* far exceeds prior estimates of divergence between interior and coastal *A. nelsoni*, several tests confirm that the most likely model supported a history of divergence between *A. n. subvirgatus* and the other subspecies first and did not support a primary division between *A. n. nelsoni* or *A. n. alterus* and the other subspecies. We discuss caveats of demographic modelling below (see Section 4), but interpret these results as offering additional support for the (*A. n. subvirgatus* (*A. n. alterus*, *A. n. nelsoni*)) topology despite discrepancies in the parameter estimates.

G-PhoCS estimated relatively small contemporary effective population sizes for all *Ammospiza* taxa (Supporting Information; Table S4): 40,000 for *A. n. alterus*, 47,000 for *A. n. nelsoni*, 90,000 for *A. n. subvirgatus*, 63,000 for *A. leconteii*, 90,000 for *A. maritima* and 57,000 for *A. caudacuta*. The corresponding estimate for the outgroup taxa *P. sandwichensis* was substantially greater (550,000). We did additional runs in G-PhoCS to estimate splitting times and migration events, but all estimates for these parameters were associated with substantial uncertainty and are thus uninformative. Effective population sizes from SMC++ suggest steady increases in N_e for all species over time until about 10,000 years ago, with larger population sizes for *P. sandwichensis*, *A. leconteii* and *A. maritima* compared to the *A. nelsoni-caudacuta* complex (Figure 3). This is congruent with estimates of small effective population sizes in tidal marsh sparrows, particularly *A. caudacuta*, with previously estimated N_e between 30,000 and 50,000 individuals (Walsh, Clucas, et al., 2019). While not unreasonable, the steep drops in N_e in the recent past should be interpreted cautiously, however, as this inference can result from long runs of homozygosity; while we tried to mask homozygosity runs in our analyses, the effectiveness of this approach is unknown. SMC++ does offer some insight into divergence within the *A. nelsoni* group, as we see similar trajectories in N_e for the three subspecies over time (Figure 3), with all three groups exhibiting almost identical estimates for N_e. Around 5,500 years ago, we see a split in N_e between the subspecies, with a decline in *nelsoni* and a continued increase in *alterus* and *subvirgatus*.

The time-calibrated tree provided additional insight into the timing of splitting events among these sparrows. When logfiles from all six MCMC chains were combined ESS (Effective Sample Size) values for all parameters exceeded 200 with the exception of the clock rate (ESS = 115). Divergence times estimates for each node were highly consistent across independent MCMC chains, across all four of the 30 million generations combined, and across all six chains combined, suggesting that our analyses had converged on a consistent estimate of divergence time for each node. All nodes of the SNAPP tree received a high posterior probability of 1.0 with the exception of the *A. n. nelsoni* and *A. n. alterus* node, which had a posterior probability of only 0.36 (Figure 4). Our analyses indicate that *P. sandwichensis* and *Ammospiza* sparrows last shared a common ancestor 6.16 million years ago (Supporting Information; Table S5; mya; 95% HPD: 4.03–9.16 mya). *A. leconteii* diverged from all other *Ammospiza* during the early Pliocene 4.75 mya (95% HPD: 3.12–7.07 mya), which was followed by later Pliocene divergence between *A. maritima* and the *A. nelsoni-caudacuta* group 2.97 mya (95% HPD: 1.93–4.39 mya). Within the sharp-tailed sparrows, *A. caudacuta* and *A. nelsoni* diverged during the Pleistocene 430,000 years ago (95% HPD: 260,000–650,000 ya). Finally, divergence among the *A. nelsoni* subspecies was estimated to have occurred very recently with mean divergence time between *A. n. subvirgatus* and the other subspecies occurring 2,100 years ago (95% HPD: 0–5,300 ya).

3.2 | Historical biogeography inference

A dispersal vicariance (DIVALIKE) biogeographical model was found to be the best fit (likelihood = −7.83; AICc = 23.65; AICc weight = 0.63; Figure 4; Table 1), although a DEC model exhibited comparable likelihood and AICc scores (likelihood = −8.48; AICc = 24.95; ΔAICc = 1.3; AICc weight = 0.33; Supporting Information; Figure S4). Adding the jump parameter 'j' to the model did not significantly improve the fit of the model (ΔAICc > 7.24; likelihood ratio test: chi-squared *D*-statistic = 2.76, *p* value = 0.096), suggesting that founder event speciation likely played only a minor role in the biogeographical history of *Ammospiza* sparrows. The results of the DIVALIKE model suggest that the ancestor to all *Ammospiza* sparrows was likely distributed in the interior great plains of North America and along the Atlantic Coast, and that it diverged between these two regions during the Miocene, ~4.75 mya, with the ancestor of *A. leconteii* occupying interior grasslands in North America and the ancestor to all other *Ammospiza* sparrows occupying coastal habitats along the Atlantic coast (Figure 4). *Ammospiza maritima* and the *A. nelsoni-caudacuta* group then most likely diverged from one another along the Atlantic Coast during the Pliocene (~2.97 mya). There was also a smaller probability (<30%) that the ancestor to *A. maritima* and the *A. nelsoni-caudacuta* group occupied the Gulf and Atlantic coasts with a vicariance event isolating *A. maritima* on the Gulf coast and the *A. nelsoni-caudacuta* group on the Atlantic coast. Next, *A. caudacuta* and *A. nelsoni* likely diverged from one another along the Atlantic coast ~0.43 mya with a smaller possibility that the ancestor of these species also occupied interior North America. Finally, following divergence from *A. caudacuta*, the ancestor to all *A. nelsoni* taxa was likely found along the Atlantic coast prior to expanding its range into interior North America and diversifying into the three recognized subspecies within the last 5,000 years. In contrast, the DEC model (Supporting Information; Figure S6) found that the ancestor of *A. caudacuta* and *A. nelsoni* was likely distributed in interior North America and the Atlantic coast. The post-divergence lineage leading to contemporary *A. caudacuta* was found on the Atlantic coast and the ancestral lineage of *A. nelsoni* likely occupied much of its current distribution prior to diverging into the current subspecies.

3.3 | Mosaic genomes and the role of introgression

Analyses suggested that signals from historic admixture events and incomplete lineage sorting (ILS) were more prevalent than those from recent introgression events, with the exception of *A.*

caudacutus and *A. n. subvirgatus*, where a present-day contact zone is independently well documented (Walsh et al., 2018; Walsh, Rowe, et al., 2016; Walsh, Shriver, et al., 2016). We saw little evidence of contemporary introgression based on Admixture results. A K of 3 had the lowest cross-validation error, corresponding to the following clusters: (1) the three subspecies of *A. nelsoni*; (2) *A. caudacutus* and (3) *P. sandwichensis*, *A. leconteii* and *A. maritima*. However, cross-validation errors fluctuated, with $K = 5$ receiving similar support and thus we use these results to confirm an overall lack of admixed individuals, as opposed to identifying clusters. To this end, when assigned to seven clusters, individuals sort out by species/subspecies group (Figure 2), with the exception of one *nelsoni* and one *alterus* individual that showed signals of admixture between the two interior *A. nelsoni* subspecies.

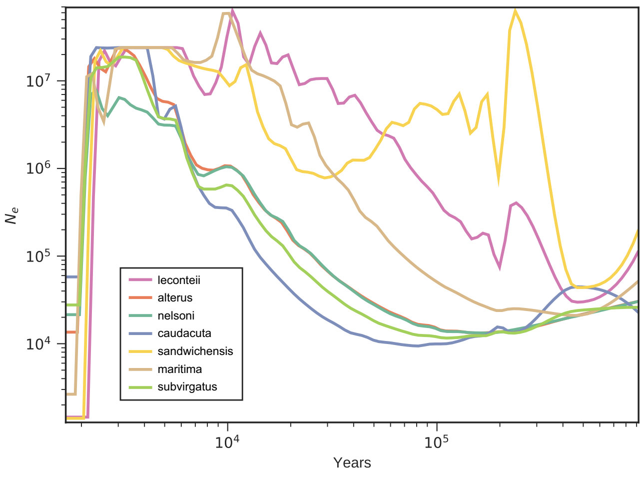

FIGURE 3 Estimates of population size histories for North American *Ammospiza* sparrows. Estimates are generated from SMC++ assuming a mutation rate of 7.59×10^{-9} (Zhang et al., 2014). A generation time of 2.3 years was used to convert generations to years

Results from Twisst support a signal of ILS among *Ammospiza* species. Twisst results supported the consensus species phylogeny from the tree-based methods above (topology 1) as the dominant genome-wide topology (Supporting Information; Figure S7). All of the other introgression-indicative topologies exhibited low weightings, genome-wide, which is consistent with a signal of incomplete lineage sorting. With regard to the timing of introgression events, estimates of D_{FS} suggest that gene flow among *Ammospiza* sparrows is constrained largely to the distant past. In comparisons assessing introgression between *A. maritima* and *A. caudacuta* (Figure 5a) as well as *A. leconteii* and *A. caudacuta* (Figure 5b), the pattern is similar to that observed with no ongoing gene flow and a possible bottleneck in *A. caudacuta*. In both of these scenarios, D_{FS} is negative in bins representing low and intermediate frequency derived alleles, but is positive in higher-frequency bins reflecting fixation of high-frequency-derived alleles via drift following ILS (Martin & Amos, 2021). The comparison assessing introgression between *A. n. subvirgatus* and *A. caudacuta* (Figure 5c) suggests a signal of more recent introgression, with D_{FS} skewed towards lower-frequency variants with low absolute values. This finding is consistent with known ongoing

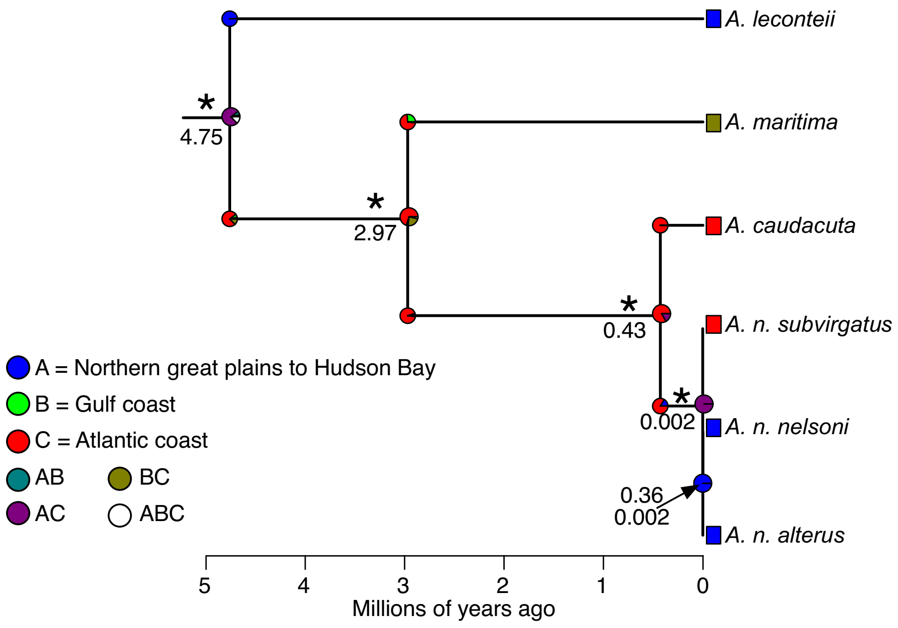

FIGURE 4 Time-calibrated phylogeny of *Ammospiza* sparrows generated in SNAPP. The outgroup (*P. sandwichensis*) was removed for BioGeoBears analysis and visualization. Coloured boxes at tips signify the contemporary distribution of each taxon. Pie charts at nodes signify the most likely ancestral distribution across North America as estimated in BioGeoBears based on a DIVALIKE model. Purple bars centred on each node signify the 95% highest posterior distribution (HPD) surrounding divergence time estimates for the node. Annotations above each branch denote posterior probability (pp) support for each node with asterisks for nodes receiving pp equal to 1. Annotations below each branch are the estimated mean divergence time for each node

TABLE 1 Comparison of different biogeographical models analysed in BioGeoBears. Models ranked by AICc score. LnL is model likelihood, *k* = number of parameters. Parameters estimated by each model: *d* = dispersal rate, *e* = extinction rate and *j* = founder event speciation

Model	LnL	*k*	AICc	ΔAICc	AICwt	*d*	*e*	*j*
DIVALIKE	−7.83	2	23.65	0.00	0.63	0.09	0.00	0.00
DEC	−8.48	2	24.95	1.30	0.33	0.07	0.00	0.00
DIVALIKE+J	−6.45	3	30.89	7.24	0.02	0.05	0.00	0.17
DEC+J	−6.66	3	31.32	7.67	0.01	0.03	0.00	0.22
BAYAREALIKE+J	−6.99	3	31.97	8.32	0.01	0.00	0.00	0.58
BAYAREALIKE	−23.13	2	54.26	30.61	0.00	2.21	1.18	0.00

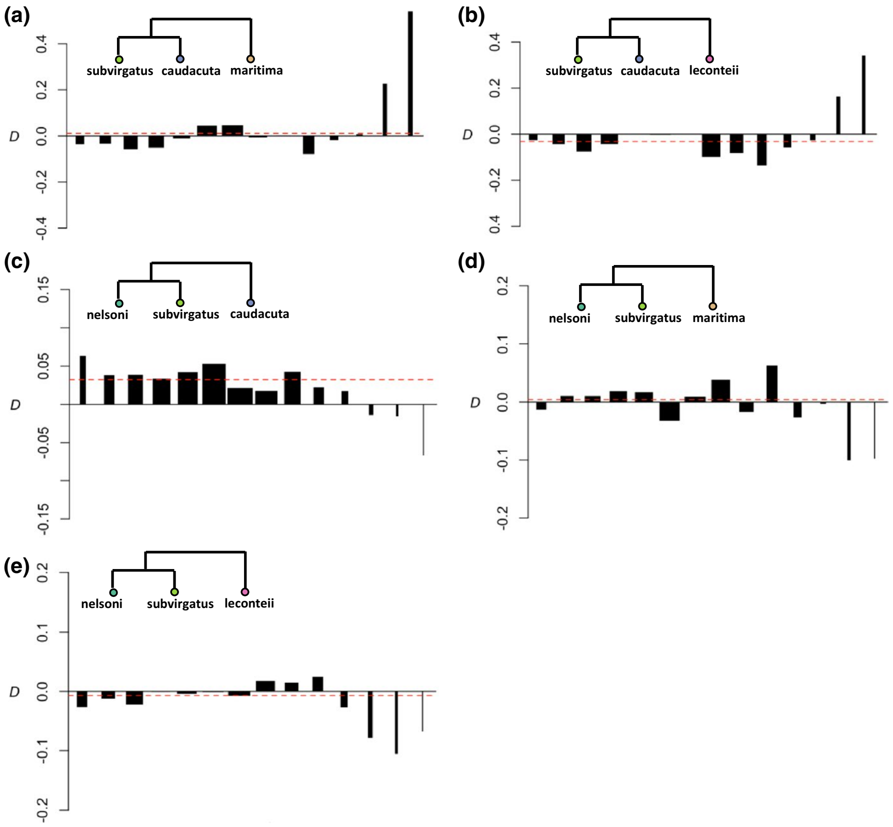

FIGURE 5 Estimates of D_{FS} for five different comparisons within North American *Ammospiza* sparrows: (a) gene flow between *A. maritima* and *A. caudacuta*; (b) gene flow between *A. leconteii* and *A. caudacuta*; (c) gene flow between *A. caudacuta* and *A. n. subvirgatus*; (d) gene flow between *A. maritima* and *A. n. subvirgatus* and (e) gene flow between *A. leconteii* and *A. n. subvirgatus*. The proportion of weighting for each bin is indicated by the width of the vertical lines. The red dashed line indicates the genome-wide *D* value

hybridization between *A. caudacuta* and *A. n. subvirgatus*. Lastly, comparisons assessing introgression events between *maritima* and *subvirgatus* (Figure 5d) and *leconteii* and *subvirgatus* (Figure 5e) again supports a signal of historic gene flow or ILS. In both of these comparisons, D_{FS} is more dispersed across the bins with the signal mainly attributed to fixed derived alleles due to drift.

4 | DISCUSSION

Demographic and biogeographical reconstructions of *Ammospiza* sparrows in North America based on whole-genome data offer new inferences about the evolutionary history of this clade. Phylogenies based on millions of SNPs supported several well-resolved clades that predominantly corresponded to their prior species designations. Within *A. nelsoni*, the relationships among the three subspecies were difficult to resolve, a finding that is unsurprising given the uncertainty arising from shallow divergence times in this clade (Giarla & Esselstyn, 2015). Analyses in *Moments* clearly supported a model where the topology of (*A. n. subvirgatus* (*A. n. alterus*, *A. n. nelsoni*)) was the best fit, however, despite discrepancies in divergence times. We also saw strong support for the placement of *A. maritima* as sister to the *A. nelsoni-caudacuta* complex, indicative of a shared salt marsh ancestor for this group and in contrast to the *A. leconteii* ancestor hypothesis posed as an alternative by Greenlaw (1993). In contrast to the more often accepted hypotheses about salt–freshwater transitions in the *A. nelsoni-caudacuta* complex, which long posited a saltwater colonization event by *A. n. subvirgatus* following a split from interior *nelsoni* populations (Greenlaw, 1993), our biogeographical reconstructions support Beecher's (1955) biogeographical scenario, which suggests the opposite—a freshwater colonization event by *A. n. nelsoni* and *alterus* following a split from *A. n. subvirgatus* (Figure 1a). Moreover, estimates of divergence times within *Ammospiza* sparrows largely predated the LGM, with the exception of splits between salt marsh and freshwater populations of *A. nelsoni*, suggesting links to saline habitats over long evolutionary time-scales.

Demographic reconstructions suggest that divergence within the *Ammospiza* clade predominantly occurred long before the formation of salt marsh habitat in its current configuration. Divergence time estimates support an initial early Pliocene split (~5mya) between *A. leconteii* and all other *Ammospiza* sparrows followed by a later Pliocene split between *A. maritima* and the *A. nelsoni-caudacuta* clade. Introgression between *A. nelsoni* and *A. caudacuta* could bias our divergence time estimates for that node towards the present while fixing population sizes across branches could also lead to over- or under-estimation of divergence times (Stange et al., 2018). However, our estimates for initial Pliocene (~3 mya) splits within *Ammospiza* sparrows are consistent with previous work based on allozymes and mtDNA, dating divergence times for major *Ammospiza* splits at 3.5–7 mya (Zink & Avise, 1990). Also consistent with previous work (Rising & Avise, 1993; Walsh, Clucas, et al., 2019), our data support a Pleistocene split (~0.43 mya) between *A. caudacuta* and *A. nelsoni*.

In contrast, the split between *A. n. subvirgatus* and the interior subspecies of *A. nelsoni* is estimated to have occurred very recently (within the last ~5,000 years and as early as 2,000 ya). While this split is shallower than previously hypothesized (~10,000 ya; Walsh, Benham, et al., 2019), effective population sizes for the three subspecies of *A. nelsoni* follow similar trajectories with an observable change approximately 5,500–8,000 ya, particularly when comparing *A. n. nelsoni* to the other two subspecies. This change in N_e may offer additional support for the split between *nelsoni* and the other two subspecies (*subvirgatus* and *alterus*) as the estimated timing is highly consistent with the recent divergence estimates from SNAPP. Moreover, the inclusion of divergence estimates from *Moments* highlights even larger discrepancies in our parameter estimates within the *A. nelsoni* complex (a split between *A. n. subvirgatus* and the interior species dated at ~459,000 ya). Several factors (e.g. ancestral size changes and undetected bottlenecks, Momigliano et al., 2021; unaccounted gene flow from 'ghost' populations, Lynch & Sethuraman, 2019; low coverage or sequencing error, Gravel et al., 2011) could be contributing to discordance between our parameter estimates. We also note here that methodological approaches that fail to account for gene flow, such as SNAPP, can result in topological uncertainty (Thom et al., 2018), which could also be contributing to the observed discordance between our parameter estimates. However, the divergence estimates obtained from SNAPP are highly consistent with previous work done in this system (Walsh, Benham, et al., 2019; Walsh, Clucas, et al., 2019) and we ultimately feel confident in the (*A. n. subvirgatus* (*A. n. alterus*, *A. n. nelsoni*)) topology despite incongruent divergence time estimates from *Moments*. We conclude that our findings warrant further investigation with larger sample sizes with greater geographical continuity across the *A. nelsoni* range, which would facilitate analyses dedicated to elucidating demographic history on a subspecies scale. Moreover, while demographic history analyses pose challenges that can impede robust inferences and reconstructions, methodological approaches are rapidly improving and may offer increased resolution, particularly at recent timescales, in the near future. Overall, this variation in divergence time estimates within the *A. nelsoni* clade is expected, as reconstructing demographic histories at shallow evolutionary time-scales can be challenging (Pease et al., 2016; Xu & Yang, 2016).

Biogeographical reconstructions suggest that the common ancestor to *Ammospiza* sparrows likely inhabited freshwater regions in the interior great plains of North America. Our top model suggests one salt marsh colonization event within *Ammospiza* sparrows; the split between the ancestor of *A. leconteii*, which remained in the grasslands, and the ancestor to all other *Ammospiza* sparrows in this study that occupied coastal habitats along the Atlantic (~4.75 mya). This is in contrast to the hypothesized biogeographical scenarios of Greenlaw (1993) that posited two or three independent salt marsh colonization events within *Ammospiza* sparrows. While our reconstructions are not perfectly aligned with the Beecher (1955) hypothesis (namely, Beecher hypothesized that *P. sandwichensis* was the closest relative to *A. maritima*), they do support a single salt marsh colonization event and the expansion of *A. nelsoni* northward along the Atlantic coast and then inland. Beecher (1955) suggests that marine connections between James Bay and the mouth of the St. Lawrence following deglaciation after the LGM may have facilitated dispersal of *A. nelsoni* inland from the coast (likely after the establishment of *A. n. subvirgatus*). The existence of this proposed marine connection is supported by the discovery of several marine plant species at inland locations (McLaughlin, 1932; Peattie, 1922; Potter, 1932), suggesting dispersal events from the North Atlantic coast to the interior in other taxonomic groups.

Our findings are in contrast to recent work favouring the hypothesis that *A. nelsoni* had freshwater origins in the interior of North America, coming into secondary contact with *A. caudacuta* in coastal habitats as a recent colonizer of saline habitats (Greenlaw, 1993; Nocera et al., 2007; Shriver et al., 2011; Walsh et al., 2018). If our reconstructions are accurate, *A. nelsoni* adapted to *freshwater* habitats instead of vice versa, which has implications for how we interpret signals and patterns of adaptive divergence within this complex. Notably, while *A. n. nelsoni* and *A. n. alterus* are seasonally freshwater adapted, they winter in salt marsh habitats (Greenlaw & Woolfenden, 2007), suggesting a retention of ancestral adaptations, or perhaps a plastic response, that allow these populations to occupy coastal marshes during the non-breeding season. While we discuss possible caveats and implications of this point in greater detail below, we highlight here that reconstructing biogeographical and evolutionary dynamics in ephemeral systems poses challenges given the propensity for high turnover (Thomaz & Knowles, 2020). Repeated cycles of isolation and re-connection can confound our interpretation of spatial and temporal patterns of diversification (Dufresnes et al., 2020; Thomaz & Knowles, 2020), reinforcing the importance of a multifaceted approach to biogeographical reconstructions (Johnson et al., 2018), particularly in historically dynamic ecosystems.

While whole-genome data offer strong support for salt marsh origins (via an *A. maritima* ancestor) for the *A. nelsoni-caudacuta* complex, biogeographical reconstructions within the group offer an incongruent picture to the currently hypothesized evolutionary scenarios as they relate specifically to dynamics within the hybrid zone. In contrast to clear signals of ILS in the other species comparisons, estimates of D_{FS} support recent introgression between *A. n.*

subvirgatus and *A. caudacuta*, consistent with previous work that documents the ongoing hybridization between them (Hodgman et al., 2002; Shriver et al., 2005; Walsh, Rowe, et al., 2016; Walsh, Shriver, et al., 2016). Thus, while contemporary introgression between *A. n. subvirgatus* and *A. caudacuta* is indisputable, characterizing the role of introgression in shaping the adaptive landscape between these species remains an elusive task. Previous research scanned the *A. caudacuta* and *A. nelsoni* genomes for differentially introgressed loci and identified several annotated genes that may play an adaptive role in hybrids inhabiting salt marsh habitats, hypothesizing adaptive introgression from *A. caudacuta* into *A. nelsoni* (Walsh et al., 2018). This hypothesis is consistent with secondary contact between *A. n. subvirgatus* and *A. caudacuta* following a recent freshwater–salt marsh colonization event by *A. n. subvirgatus*. However, our biogeographical reconstructions shed new light on this proposed scenario. Although our findings suggest saline to freshwater movements of *A. n. subvirgatus*, marshes along the Atlantic coast of North America would have undergone substantial change since the hypothesized Pleistocene split between these species (Cronin, 1988; Wares, 2002). Changing sea levels would have led to cycles of isolation and connection along the Atlantic coast (Wares, 2002), supporting a scenario where hybridization could still be a result of secondary contact following periods of isolation and allopatric divergence between the two species. While the glacial maxima resulted in much of the Atlantic coastline being uninhabitable, several species survived in unglaciated refugia both to the south and to the north of the glacial margin (Cronin, 1988; Holder et al., 1999; Wares, 2002). This suggests a strong possibility that *A. caudacuta* may have retreated to salt marsh refugia in the south, whereas *A. nelsoni* may have retreated north where extensive regions (both terrestrial and coastal) of the Canadian Maritimes remained unglaciated (Berggren & Hollister, 1974; Dyke and Prest, 1987; Holder et al., 1999). On a more recent time-scale, geomorphic and ecological processes are known to vary latitudinally (Bertness, 1999; Díaz-Ferguson et al., 2010; Salgado & Pennings, 2005) creating scenarios where select genotypes may be better suited to local environmental conditions—a pattern that has previously been observed in the *A. nelsoni-caudacuta* complex (Walsh, Rowe, et al., 2016). As such, a scenario of adaptive introgression following secondary contact could still be highly relevant here, but perhaps the spatial and temporal characteristics of these processes are more nuanced than previously thought.

It is important to note some of the caveats and limitations to our analyses, which warrant some caution as we interpret the evolutionary and biogeographical history of the *Ammospiza* clade. While our reconstructions were well supported at the species level, there were several hurdles to resolving subspecies relationships within *A. nelsoni*. In both the SVDquartets and RAxML phylogenies, *A. n. alterus* and *A. n. nelsoni* tended to cluster together, but were not reciprocally monophyletic in any phylogeny and there were clear signals of mixing between all three subspecies. Analyses from *Moments* further supported a scenario of gene flow between geographically adjacent populations. While our model selection approach offers strong support for the (*A. n. subvirgatus* (*A. n. alterus*, *A. n. nelsoni*)) topology, the lack of resolution in our parameter estimates is consistent with phylogenetic reconstructions at the subspecies and population level (Kutschera et al., 2014) and highlight the challenges with interpreting phylogenetic relationships at this scale, particularly in light of recent gene flow and ILS (Leaché et al., 2014; Tajima, 1983). Despite these caveats, our results point to the intricate and multidimensional contributions of ecology, demography, biogeography and evolution in shaping genomic landscapes. Extensive empirical and theoretical work has shown that competing processes can shape genomic signals in similar ways (Burri, 2017; Ravinet et al., 2017). To this end, incorporating information on demographic and biogeographical history is paramount to our interpretations of the genomic architecture of speciation and ecological divergence (Ravinet et al., 2017). This multifaceted approach offers insight into patterns and drivers of diversification across dynamic and ephemeral ecosystems, such as tidal marshes, and highlights the utility of high-resolution data in reconstructing demographic and biogeographical histories in complex environments. For *Ammospiza* sparrows, this approach has provided new insights about these lineages, shedding light on a less widely accepted biogeographical hypothesis—with broad implications regarding our understanding of the genomic architecture of divergence and introgression in this clade. More broadly, the approach used allowed us to tease apart the mechanisms of diversification, and the role of hybridization in facilitating adaptation to an ephemeral landscape. Specifically, our findings highlight the fact that complex biogeographical and demographic histories and processes can lead to similar genomic signals and gaining resolution into these events is paramount to our interpretation of adaptive divergence between species spanning unstable ecological gradients.

ACKNOWLEDGEMENTS

The authors thank members of the Fuller Evolutionary Biology Program, particularly L. Campagna, for feedback and comments on an earlier draft of this manuscript. We thank S. Aguillon for the help with figure modifications. We thank J. Ditner for all illustrations with input from K. McGowan. G. Shriver provided the *Ammospiza maritima* blood samples used in this study. All blood samples collected for this study were done so under state and Federal banding permits. All field sampling was done in accordance with IACUC. Tissue samples for *Ammospiza leconteii* were generously provided by the Sam Noble Oklahoma Museum of Natural History. Funding for this work was provided by the Fuller Evolutionary Biology Program, by the US National Science Foundation Postdoctoral Research Fellowship in Biology under grant No. DBI 1523719, and the Cornell University Center for Vertebrate Genomics.

DATA AVAILABILITY STATEMENT

A filtered VCF containing SNP data for the 53 individuals used in this study are available on Dryad: https://doi.org/10.5061/dryad.73n5tb2x6. Raw data will be made available upon request from the author.

ORCID

Jennifer Walsh https://orcid.org/0000-0002-8146-7029

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Balakrishnan, C. N., & Edwards, S. V. (2009). Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics*, *181*(2), 645–660.

Beecher, W. J. (1955). Late-Pleistocene isolation in salt-marsh sparrows. *Ecology*, *36*(1), 23–28. https://doi.org/10.2307/1931426

Benham, P. M., & Cheviron, Z. A. (2020). Population history and the selective landscape shape patterns of osmoregulatory trait divergence in tidal marsh Savannah sparrows (*Passerculus sandwichensis*). *Evolution*, *74*(1), 57–72.

Berggren, W. A., & Hollister, C. D. (1974). Paleogeography, paleobiogeography and the history of circulation in the Atlantic Ocean.

Bertness, M. D. (1999). *The ecology of Atlantic shorelines* (Vol. 417). Sinauer Associates.

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., & Matschiner, M. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *15*(4), e1006650.

Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics*, *81*(5), 1084–1097. https://doi.org/10.1086/521987

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

Burri, R. (2017). Linked selection, demography and the evolution of correlated genomic landscapes in birds and beyond. *Molecular Ecology*, *26*(15), 3853–3856. https://doi.org/10.1111/mec.14167

Campagna, L., McCracken, K. G., & Lovette, I. J. (2019). Gradual evolution towards flightlessness in steamer ducks. *Evolution*, *73*(9), 1916–1926. https://doi.org/10.1111/evo.13758

Chan, Y. L., Hill, C. E., Maldonado, J. E., & Fleischer, R. C. (2006). Evolution and conservation of tidal-marsh vertebrates: Molecular approaches. *Terrestrial Vertebrates of Tidal Marshes: Evolution, Ecology, and Conservation*.

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*(23), 3317–3324. https://doi.org/10.1093/bioinformatics/btu530

Chifman, J., & Kubatko, L. (2015). Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. *Journal of theoretical biology*, *374*, 35–47.

Chou, J., Gupta, A., Yaduvanshi, S., Davidson, R., Nute, M., Mirarab, S., & Warnow, T. (2015). A comparative study of SVDquartets and other coalescent-based species tree estimation methods. *BMC Genomics*, *16*(10), 1–11. https://doi.org/10.1186/1471-2164-16-S10-S2

Coffman, A. J., Hsieh, P. H., Gravel, S., & Gutenkunst, R. N. (2016). Computationally efficient composite likelihood statistics for demographic inference. *Molecular Biology and Evolution*, *33*(2), 591–593. https://doi.org/10.1093/molbev/msv255

Cronin, T. W. (1988). Evolution of marine climates of the US Atlantic coast during the past four million years. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *318*(1191), 661–678.

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

Díaz-Ferguson, E., Robinson, J. D., Silliman, B., & Wares, J. P. (2010). Comparative phylogeography of North American Atlantic salt marsh communities. *Estuaries and Coasts*, *33*(4), 828–839. https://doi.org/10.1007/s12237-009-9220-6

Dolby, G. A., Ellingson, R. A., Findley, L. T., & Jacobs, D. K. (2018). How sea level change mediates genetic divergence in coastal species across regions with varying tectonic and sediment processes. *Molecular Ecology*, *27*(4), 994–1011. https://doi.org/10.1111/mec.14487

Dufresnes, C., Nicieza, A. G., Litvinchuk, S. N., Rodrigues, N., Jeffries, D. L., Vences, M., Perrin, N., & Martínez-Solano, Í. (2020). Are glacial refugia hotspots of speciation and cytonuclear discordances? Answers from the genomic phylogeography of Spanish common frogs. *Molecular Ecology*, *29*(5), 986–1000. https://doi.org/10.1111/mec.15368

Dyke, A., & Prest, V. (1987). Late Wisconsinan and Holocene history of the Laurentide ice sheet. *Géographie Physique et Quaternaire*, *41*(2), 237–263. https://doi.org/10.7202/032681ar

Emslie, S. D. (1998). Avian community, climate, and sea-level changes in the Plio-Pleistocene of the Florida Peninsula. *Ornithological Monographs*, 1–113. https://doi.org/10.2307/40166707

Fairbanks, R. G. (1989). A 17,000-year glacio-eustatic sea level record: Influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature*, *342*(6250), 637–642. https://doi.org/10.1038/342637a0

Giarla, T. C., & Esselstyn, J. A. (2015). The challenges of resolving a rapid, recent radiation: Empirical and simulated phylogenomics of Philippine shrews. *Systematic Biology*, *64*(5), 727–740. https://doi.org/10.1093/sysbio/syv029

Goldstein, D. L. (2006). Osmoregulatory biology of saltmarsh passerines. *Studies in Avian Biology*, *32*, 110.

Gravel, S., Henn, B. M., Gutenkunst, R. N., Indap, A. R., Marth, G. T., & Clark, A. G., … 1000 Genomes Project. (2011). Demographic history and rare allele sharing among human populations. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(29), 11983–11988. https://doi.org/10.1073/pnas.1019276108

Greenberg, R. (2006). Tidal marshes: home for the few and they highly selected. In R. Greenberg, J. E. Maldonado, S. Droege, & M. V. McDonald (Eds.), *Terrestrial vertebrates of tidal marshes: Evolution, ecology, and conservation. Studies in Avian Biology* (Vol. 32, pp. 2–9).

Greenberg, R., & Droege, S. (1990). Adaptations to tidal marshes in breeding populations of the swamp sparrow. *Condor*, *92*(2), 393–404. https://doi.org/10.2307/1368236

Greenberg, R., Maldonado, J. E., Droege, S. A. M., & Victoria McDonald, M. (2006). Tidal marshes: A global perspective on the evolution and conservation of their terrestrial vertebrates. *BioScience*, *56*(8), 675–685.

Greenberg, R., & Olsen, B. (2010). Bill size and dimorphism in tidal-marsh sparrows: Island-like processes in a continental habitat. *Ecology*, *91*(8), 2428–2436. https://doi.org/10.1890/09-1136.1

Greenlaw, J. S. (1993). Behavioral and morphological diversification in Sharp-tailed Sparrows (Ammodramus caudacutus) of the Atlantic Coast. *The Auk*, *110*(2), 286–293.

Greenlaw, J. S., & Rising, J. D. (1994). Saltmarsh sparrow (*Ammodramus caudacutus*), The Birds of North America Online. In A. Poole (Ed.). Cornell Lab of Ornithology; Retrieved from the Birds of North American online http://bna.birds.cornell.edu/bna/species/112

Greenlaw, J. S., & Woolfenden, G. E. (2007). Wintering distributions and migration of Saltmarsh and Nelson's Sharp-tailed sparrows. *The Wilson Journal of Ornithology*, *119*(3), 361–377. https://doi.org/10.1676/05-152.1

Grenier, J. L., & Greenberg, R. U. S. S. E. L. L. (2006). Trophic adaptations in sparrows and other vertebrates of tidal marshes. *Studies in Avian Biology*, *32*, 130.

Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., & Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, *43*(10), 1031. https://doi.org/10.1038/ng.937

Hodgman, T. P., Shriver, W. G., & Vickery, P. D. (2002). Redefining range overlap between the sharp-tailed sparrows of coastal New England. *Wilson Bulletin.*, *114*, 38–43.

Holder, K., Montgomerie, R., & Friesen, V. L. (1999). A test of the glacial refugium hypothesis using patterns of mitochondrial and nuclear DNA sequence variation in rock ptarmigan (*Lagopus mutus*). *Evolution*, *53*(6), 1936–1950.

Johnson, J. S., Krutovsky, K. V., Rajora, O. P., Gaddis, K. D., & Cairns, D. M. (2018). Advancing biogeography through population genomics. In O. Rajora *Population Genomics* (pp. 539–585). Springer. https://doi.org/10.1007/13836_2018_39

Jouganous, J., Long, W., Ragsdale, A. P., & Gravel, S. (2017). Inferring the joint demographic history of multiple populations: Beyond the diffusion approximation. *Genetics*, *206*(3), 1549–1567. https://doi.org/10.1534/genetics.117.200493

Kawakami, T., Smeds, L., Backström, N., Husby, A., Qvarnström, A., Mugal, C. F., Olason, P., & Ellegren, H. (2014). A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. *Molecular Ecology*, *23*(16), 4035–4058. https://doi.org/10.1111/mec.12810

Klicka, J., Barker, F. K., Burns, K. J., Lanyon, S. M., Lovette, I. J., Chaves, J. A., & Bryson Jr., R. W. (2014). A comprehensive multilocus assessment of sparrow (Aves: Passerellidae) relationships. *Molecular Phylogenetics and Evolution*, *77*, 177–182. https://doi.org/10.1016/j.ympev.2014.04.025

Kutschera, V. E., Bidon, T., Hailer, F., Rodi, J. L., Fain, S. R., & Janke, A. (2014). Bears in a forest of gene trees: phylogenetic inference is complicated by incomplete lineage sorting and gene flow. *Molecular Biology and Evolution*, *31*(8), 2004–2017. https://doi.org/10.1093/molbev/msu186

Landis, M. J., Matzke, N. J., Moore, B. R., & Huelsenbeck, J. P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, *62*(6), 789–804. https://doi.org/10.1093/sysbio/syt040

Leaché, A. D., Harris, R. B., Rannala, B., & Yang, Z. (2014). The influence of gene flow on species tree estimation: A simulation study. *Systematic Biology*, *63*(1), 17–30. https://doi.org/10.1093/sysbio/syt049

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lindgreen, S. (2012). AdapterRemoval: Easy cleaning of next-generation sequencing reads. *BMC Research Notes*, *5*(1), 337. https://doi.org/10.1186/1756-0500-5-337

Luttrell, S. A., Gonzalez, S. T., Lohr, B., & Greenberg, R. (2015). Digital photography quantifies plumage variation and salt marsh melanism among Song Sparrow (*Melospiza melodia*) subspecies of the San Francisco Bay. *The Auk*, *132*(1), 277–287.

Lynch, M., & Sethuraman, A. (2019). The effects of gene flow from unsampled 'ghost' populations on the estimation of evolutionary history under the isolation with migration model. *bioRxiv*, 733600.

Malamud-Roam, K. P., Malamud-Roam, F. P., Watson, E. B., Collins, J. N., & Ingram, B. L. (2006). The quaternary geography and biogeography of tidal saltmarshes. *Studies in Avian Biology*, *32*, 11.

Martin, S. H., & Amos, W. (2021). Signatures of Introgression across the Allele Frequency Spectrum. *Molecular Biology and Evolution*, *38*(2), 716–726. https://doi.org/10.1093/molbev/msaa239

Martin, S. H., & Van Belleghem, S. M. (2017). Exploring evolutionary relationships across the genome using topology weighting. *Genetics*, *206*(1), 429–438. https://doi.org/10.1534/genetics.116.194720

Matzke, N. J. (2013). Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, *5*(4). https://doi.org/10.21425/F55419694

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

McLaughlin, W. T. (1932). Atlantic coastal plain plants in the sand barrens of northwestern Wisconsin. *Ecological Monographs*, *2*(3), 335–383. https://doi.org/10.2307/1943067

Mikles, C. S., Aguillon, S. M., Chan, Y. L., Arcese, P., Benham, P. M., Lovette, I. J., & Walsh, J. (2020). Genomic differentiation and local adaptation on a microgeographic scale in a resident songbird. *Molecular Ecology*, *29*(22), 4295–4307. https://doi.org/10.1111/mec.15647

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2011, November). The CIPRES science gateway: a community resource for phylogenetic analyses. *In Proceedings of the 2011 TeraGrid Conference: extreme digital discovery* (pp. 1–8).

Mitsch, W. J., & Gosselink, J. G. (2000). The value of wetlands: Importance of scale and landscape setting. *Ecological Economics*, *35*(1), 25–33. https://doi.org/10.1016/S0921-8009(00)00165-8

Momigliano, P., Florin, A. B., & Merilä, J. (2021). Biases in demographic modelling affect our understanding of recent divergence. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msab047

Murray, B. G. (1969). A comparative study of the Le Conte's and Sharp-tailed sparrows. *The Auk*, *86*(2), 199–231. https://doi.org/10.2307/4083496

Nocera, J., Forbes, G., & Milton, G. R. (2007). Habitat relationships of three grassland breeding bird species: Broadscale comparisons and hayfield management implications. *Avian Conservation and Ecology*, *2*(1). https://doi.org/10.5751/ACE-00127-020107

Okonechnikov, K., Conesa, A., & García-Alcalde, F. (2016). Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*, *32*(2), 292–294.

Pease, J. B., Haak, D. C., Hahn, M. W., & Moyle, L. C. (2016). Phylogenomics reveals three sources of adaptive variation during a rapid radiation. *PLoS Biology*, *14*(2), e1002379. https://doi.org/10.1371/journal.pbio.1002379

Peattie, D. C. (1922). The Atlantic coastal plain element in the flora of the Great Lakes. *Rhodora*, *24*(280), 57–70.

Potter, D. (1932). *Botanical evidence of a post-pleistocene marine connection between Hudson Bay and the St. Lawrence basin* (pp. 69–112). Contributions from the Gray Herbarium of Harvard University.

Poulson, T. L. (1969). Salt and water balance in Seaside and Sharp-tailed sparrows. *The Auk*, *86*(3), 473–489. https://doi.org/10.2307/4083409

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

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, *30*(8), 1450–1477. https://doi.org/10.1111/jeb.13047

Ree, R. H., & Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, *57*(1), 4–14. https://doi.org/10.1080/10635150701883881

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*(2), 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Rising, J. D., & Avise, J. C. (1993). Application of genealogical-concordance principles to the taxonomy and evolutionary history of the Sharp-tailed Sparrow (*Ammodramus caudacutus*). *The Auk*, *110*(4), 844–856. https://doi.org/10.2307/4088638

Robbins, C. S. (1983). Distribution and migration of Seaside Sparrows. In *The Seaside Sparrow: Its Biology and Management. Occasional Papers of the North Carolina Biological Survey, Raleigh* (pp. 31–40).

Ronquist, F. (1997). Phylogenetic approaches in coevolution and biogeography. *Zoologica Scripta*, *26*(4), 313–322. https://doi.org/10.1111/j.1463-6409.1997.tb00421.x

Ruddiman, W. F. (2001). *Earth's climate: Past and future*. Macmillan.

Salgado, C. S., & Pennings, S. C. (2005). Latitudinal variation in palatability of salt-marsh plants: Are differences constitutive? *Ecology*, *86*(6), 1571–1579. https://doi.org/10.1890/04-1257

Shriver, W. G., Gibbs, J. P., Vickery, P. D., Gibbs, H. L., Hodgman, T. P., Jones, P. T., & Jacques, C. N. (2005). Concordance between morphological and molecular markers in assessing hybridization between sharp-tailed sparrows in New England. *The Auk*, *122*, 94–107. https://doi.org/10.1093/auk/122.1.94

Shriver, W. G., Hodgman, T. P., & Hanson, A. R. (2011). Nelson's sparrow (*Ammodramus nelsoni*). In A. Poole (Ed.), *The Birds of North America Online*. Cornell Lab of Ornithology.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stange, M., Sánchez-Villagra, M. R., Salzburger, W., & Matschiner, M. (2018). Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus. *Systematic Biology*, *67*(4), 681–699. https://doi.org/10.1093/sysbio/syy006

Steadman, D. W. (1981). A reexamination of *Palaeostruthus hatcheri* (Shufeldt), a late Miocene sparrow from Kansas. *Journal of Vertebrate Paleontology*, *1*(2), 171–173.

Swofford, D. L. (2002). *Phylogenetic analysis using parsimony (*and Other Methods). Version 4.*, Sunderland, MA: Sinauer Associates.

Swofford, D. L. (2003). PAUP* ver 4.0. b10. In *Phylogenetic analysis using parsimony and other methods*. Sinauer Associates, Sunderland.

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, *105*(2), 437–460.

Terhorst, J., Kamm, J. A., & Song, Y. S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, *49*(2), 303–309. https://doi.org/10.1038/ng.3748

Thom, G., Amaral, F. R. D., Hickerson, M. J., Aleixo, A., Araujo-Silva, L. E., Ribas, C. C., Choueri, E., & Miyaki, C. Y. (2018). Phenotypic and genetic structure support gene flow generating gene tree discordances in an Amazonian floodplain endemic species. *Systematic Biology*, *67*(4), 700–718. https://doi.org/10.1093/sysbio/syy004

Thomaz, A. T., & Knowles, L. L. (2020). Common barriers, but temporal dissonance: Genomic tests suggest ecological and paleolandscape sieves structure a coastal riverine fish community. *Molecular Ecology*, *29*(4), 783–796. https://doi.org/10.1111/mec.15357

Walsh, J., Benham, P. M., Deane-Coe, P. E., Arcese, P., Butcher, B. G., Chan, Y. L., Cheviron, Z. A., Elphick, C. S., Kovach, A. I., Olsen, B. J., Shriver, W. G., Winder, V. L., & Lovette, I. J. (2019). Genomics of rapid ecological divergence and parallel adaptation in four tidal marsh sparrows. *Evolution Letters*, *3*(4), 324–338. https://doi.org/10.1002/evl3.126

Walsh, J., Clucas, G. V., MacManes, M. D., Thomas, W. K., & Kovach, A. I. (2019). Divergent selection and drift shape the genomes of two avian sister species spanning a saline-freshwater ecotone. *Ecology and Evolution*, *9*(23), 13477–13494. https://doi.org/10.1002/ece3.5804

Walsh, J., Kovach, A. I., Lane, O. P., O'Brien, K. M., & Babbitt, K. J. (2011). Genetic barcode RFLP analysis of the Nelson's and Saltmarsh sparrow hybrid zone. *Wilson Journal of Ornithology*, *123*(2), 316–322.

Walsh, J., Kovach, A. I., Olsen, B. J., Shriver, W. G., & Lovette, I. J. (2018). Bidirectional adaptive introgression between two ecologically divergent sparrow species. *Evolution*, *72*(10), 2076–2089. https://doi.org/10.1111/evo.13581

Walsh, J., Rowe, R. J., Olsen, B. J., Shriver, W. G., & Kovach, A. I. (2016). Genotype-environment associations support a mosaic hybrid zone between two tidal marsh birds. *Ecology and Evolution*, *6*(1), 279–294. https://doi.org/10.1002/ece3.1864

Walsh, J., Shriver, W. G., Olsen, B. J., & Kovach, A. I. (2016). Differential introgression and the maintenance of species boundaries in an advanced generation avian hybrid zone. *BMC Evolutionary Biology*, *16*(1), 1–18. https://doi.org/10.1186/s12862-016-0635-y

Wares, J. P. (2002). Community genetics in the Northwestern Atlantic intertidal. *Molecular Ecology*, *11*(7), 1131–1144. https://doi.org/10.1046/j.1365-294X.2002.01510.x

Xu, B., & Yang, Z. (2016). Challenges in species tree estimation under the multispecies coalescent model. *Genetics*, *204*(4), 1353–1368. https://doi.org/10.1534/genetics.116.190173

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., Storz, J. F., Antunes, A., Greenwold, M. J., Meredith, R. W., Odeen, A., Cui, J., Zhou, Q., Xu, L., Pan, H., Wang, Z., Jin, L., Zhang, P., Hu, H., … Froman, D. P. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, *346*(6215), 1311–1320. https://doi.org/10.1126/science.1251385

Zink, R. M., & Avise, J. C. (1990). Patterns of mitochondrial DNA and allozyme evolution in the avian genus Ammodramus. *Systematic zoology*, *39*(2), 148–161.

BIOSKETCH

Jennifer Walsh is broadly interested in combining genomic and ecological approaches to study speciation and local adaptation. This work is a part of a larger research programme focused on characterizing adaptive variation in tidal marsh sparrows. She collaborates with many of the other authors on avian genomics and evolution at Cornell University (see The Fuller Evolutionary Biology Program; http://www.eeb.cornell.edu/lovette/).

Authors' contributions: J.W., P.M.B., G.V.C. and A.I.K. conceived and designed the study with input from I.J.L. A.I.K., V.L.W., P.M.B. and J.W. conducted the field work and collected the blood samples. J.W. conducted all the laboratory work and J.W., G.V.C., and P.M.B. carried out all the bioinformatic analyses. Data analysis and interpretation was conducted by J.W., G.V.C. and P.M.B. with input from I.J.L. and A.I.K. J.W. wrote the manuscript with input from all co-authors.

SUPPORTING INFORMATION

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

How to cite this article: Walsh J, Kovach AI, Benham PM, Clucas GV, Winder VL, Lovette IJ. Genomic data reveal the biogeographical and demographic history of *Ammospiza* sparrows in northeast tidal marshes. *J Biogeogr*. 2021;00:1–15. https://doi.org/10.1111/jbi.14208