Received: 8 March 2022 | Revised: 13 December 2022 | Accepted: 20 December 2022

DOI: 10.1111/mec.16841

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Demography and linked selection interact to shape the genomic landscape of codistributed woodpeckers during the Ice Age

Lucas R. Moreira[1,2,3] | John Klicka[4] | Brian Tilston Smith[2]

[1]Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, New York, USA

[2]Department of Ornithology, American Museum of Natural History, New York City, New York, USA

[3]Program in Bioinformatics and Integrative Biology, University of Massachusetts Chan Medical School, Worcester, Massachusetts, USA

[4]Burke Museum of Natural History and Culture and Department of Biology, University of Washington, Seattle, Washington, USA

Correspondence
Lucas R. Moreira, Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, New York, USA.
Email: lr2767@columbia.edu

Funding information
Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number: 211496/2014-6; National Science Foundation, Grant/Award Number: DEB-1655736

Handling Editor: Yanhua qu

Abstract

The influence of genetic drift on population dynamics during Pleistocene glacial cycles is well understood, but the role of selection in shaping patterns of genomic variation during these events is less explored. We resequenced whole genomes to investigate how demography and natural selection interact to generate the genomic landscapes of Downy and Hairy Woodpecker, species codistributed in previously glaciated North America. First, we explored the spatial and temporal patterns of genomic diversity produced by neutral evolution. Next, we tested (i) whether levels of nucleotide diversity along the genome are correlated with intrinsic genomic properties, such as recombination rate and gene density, and (ii) whether different demographic trajectories impacted the efficacy of selection. Our results revealed cycles of bottleneck and expansion, and genetic structure associated with glacial refugia. Nucleotide diversity varied widely along the genome, but this variation was highly correlated between the species, suggesting the presence of conserved genomic features. In both taxa, nucleotide diversity was positively correlated with recombination rate and negatively correlated with gene density, suggesting that linked selection played a role in reducing diversity. Despite strong fluctuations in effective population size, the maintenance of relatively large populations during glaciations may have facilitated selection. Under these conditions, we found evidence that the individual demographic trajectory of populations modulated linked selection, with purifying selection being more efficient in removing deleterious alleles in large populations. These results highlight that while genome-wide variation reflects the expected signature of demographic change during climatic perturbations, the interaction of multiple processes produces a predictable and highly heterogeneous genomic landscape.

KEYWORDS
bird, genetic diversity, genetic load, phylogeography, recombination, whole-genome resequencing

1 | INTRODUCTION

Pleistocene glacial cycles altered the distribution and evolution of entire communities (Hewitt, 2000, 2004). Despite the profound impact glaciations had on the evolutionary trajectory of species, the majority of research on the topic has focused on how demographic dynamics have shaped neutral genetic variation (Hewitt, 2004; Nadachowska-Brzyska et al., 2015). Population expansion (Burbrink et al., 2016; Lessa et al., 2003), genetic structuring in refugia (Anderson et al., 2006; Knowles, 2001; Shafer et al., 2010; Waltari

et al., 2007; Zink et al., 2004), and decreased diversity in expanding populations (Campbell-Staton et al., 2012; Pulgarín-R & Burg, 2012; Reid et al., 2018) are among the most common patterns recovered. However, as species rapidly expanded and colonized areas previously glaciated they would have been subject to selective pressures, such as directional (e.g., in favor of beneficial mutations) and purifying (against deleterious mutations; Davis, 2001; Gossmann et al., 2019). Understanding how natural selection, along with genetic drift, interact with features of the genome to shape the genomic landscape of diversity and differentiation will elucidate the impact of demographic changes during the Ice Age on the efficacy of natural selection.

Mutation, demography and selection play a central role shaping levels of genetic diversity, but their effects are intertwined (Jensen et al., 2019; Kern & Hahn, 2018; Li et al., 2012). Neutral genetic diversity (θ) is the product of the rate at which new alleles are generated (i.e., mutation rate μ) by the effective population size (N_e), so that diversity levels are predicted to increase as a function of the size of populations (In diploids, $\theta = 4N_e\mu$; Kimura & Crow, 1964; Kimura, 1983). On the other hand, fixation of beneficial alleles (selective sweep; Maynard & Haigh, 2007; Cutter & Choi, 2010) or removal of deleterious mutations (purifying selection; Charlesworth et al., 1993; Cutter & Choi, 2010; Cutter & Payseur, 2013; Comeron, 2014) can cause genetic diversity to decrease across the genome through the effect of linked selection (Cutter & Payseur, 2013). Demographic perturbations that cause N_e to fluctuate over time and space (e.g., glacial bottlenecks) are expected to result in a larger accumulation of mildly deleterious alleles when compared to large populations with constant N_e because of the reduced efficacy of purifying selection when genetic drift is strong (de Pedro et al., 2021; Henn et al., 2016; Rougemont et al., 2020; Wang et al., 2018; Willi et al., 2018). Hence, populations resulting from founder events, such as at the leading edge of a postglacial expansion, often show elevated genetic load (de Pedro et al., 2021; Mattila et al., 2019; Willi et al., 2018).

Levels of diversity and differentiation along the genome also vary due to the differing effects of intrinsic genomic properties (Begun & Aquadro, 1992; Dutoit, Vijay, et al., 2017; Gossmann et al., 2011; Stankowski et al., 2019; Wang et al., 2020). Genome features such as variation in mutation rate, recombination rate, distribution of functional elements, and nucleotide composition impact the rates at which genetic variants are produced, maintained, and lost (Talla et al., 2019). Regions enriched for functional elements (e.g., coding sequences), for instance, tend to exhibit significantly lower levels of genetic diversity due to the recurrent effect of natural selection (Andolfatto, 2007; Beissinger et al., 2016; Branca et al., 2011; Gossmann et al., 2011). The loss of variation is further amplified by linkage, which reduces diversity at neutrally-evolving sites in close proximity to the targets of directional or purifying selection (background selection; Charlesworth et al., 1993; Cutter & Choi, 2010; Cutter & Payseur, 2013; Comeron, 2014). The extent to which linked selection affects neighbouring sites depends on the recombination rate, which shows considerable genome-wide variation (Jensen-Seaman, 2004; Kawakami et al., 2014; Schield et al., 2020; Smukowski & Noor, 2011). Larger reductions in nucleotide diversity are expected to occur in genomic regions enriched for functional elements and with lower recombination rates. A correlation between nucleotide diversity, gene density, and recombination rate is therefore indicative that linked selection (either through background or hitchhiking selection) is at play (Talla et al., 2019). Quantifying covariance between evolutionarily independent species can help understand the interplay between these various conserved features of the genome and their impact on patterns of diversity and differentiation along the genome (Talla et al., 2019).

We aim to understand drift-selection dynamics during the Pleistocene climatic cycles by estimating the impact of demography and linked selection on the genome of Downy (*Dryobates pubescens*) and Hairy (*D. villosus*) Woodpeckers, two codistributed species that share similar ecologies and evolutionary histories. Downy and Hairy Woodpecker are year-round residents of a variety of habitats in North America, occurring in sympatry across an exceptionally broad geographic area from Alaska to Florida, although the range of the Hairy Woodpecker extends further south, reaching portions of Central America and the Bahamas (Ouellet, 1977). Despite looking very similar, the two species are not sisters and share a common ancestor more than eight million years ago, with no evidence of recent hybridization (Weibel & Moore, 2005; Dufort, 2016; Appendix S1; Figure S1). During the glacial cycles of the Pleistocene, especially when the polar ice sheets reached their maximum extent (Last Glacial Maximum; 21 kya), a large portion of the present-day distribution of Downy and Hairy Woodpeckers was covered in ice, and populations of both species were restricted to southern refugia (Graham & Burg, 2012; Klicka et al., 2011; Pulgarín-R & Burg, 2012). After the retreat of Pleistocene glaciers, Downy and Hairy Woodpeckers extended their distributions north, recolonizing higher latitudes. Phylogeographical studies in Downy and Hairy Woodpecker revealed that populations currently inhabiting previously glaciated areas show strong signatures of population expansion and population structuring consistent with multiple glacial refugia (Avise, 1992; Graham & Burg, 2012; Klicka et al., 2011; Pulgarín-R & Burg, 2012; Smith et al., 2021). This shared demographic history provides an opportunity to investigate multiple genomic factors (e.g., genetic diversity, recombination rate, and gene density) that might have impacted the distribution of diversity across populations and within the genomes of these two natural evolutionary replicates.

In this study, we generated whole-genome resequencing data for Downy and Hairy Woodpeckers to (i) investigate their population structure and demographic history; (ii) test whether the heterogeneous genomic landscape of diversity and differentiation in both taxa is correlated with intrinsic features of the genome, such as recombination rate and gene density; and (iii) test whether differences in demographic history among species and populations had an impact on the efficacy of selection. We expect that if linked selection reduces diversity at linked neutral sites along the genome, local levels of nucleotide diversity should be correlated with the rate of recombination and the density of targets of selection. In addition, we predict that if the efficiency of selection is a function of the

demographic trajectory of populations during the Ice Age, large and more stable populations (i.e., larger long-term N_e) will exhibit lower genetic load and a stronger correlation between nucleotide diversity and intrinsic genomic properties, such as recombination rate and gene density. These results have implications for our understanding of the relative importance of neutral and selective processes on the evolution of the genomic landscape of species heavily impacted by dramatic environmental fluctuations.

2 | MATERIALS AND METHODS

2.1 | Sample collection and whole genome sequencing

We collected 70 samples for both the Downy Woodpecker (*D. pubescens*) and Hairy Woodpecker (*D. villosus*) in each of seven geographically-clustered sampling locations (hereafter referred to as "sampling regions"; $n = 10$ per sampling region) across their temperate North American ranges (Figure 1): New York (Northeast), Louisiana (Southeast), Minnesota (Midwest), New Mexico and Colorado (Southern Rockies), Wyoming (Northern Rockies), Washington (Pacific Northwest), and Alaska. The samples were obtained through museum loans of vouchered specimens and augmented by field collections in Wyoming, Louisiana, and Alaska (Table S1). We extracted genomic DNA from tissue samples using the MagAttract High molecular weight DNA kit from Qiagen following the manufacturer's instructions (Qiagen). These samples were then submitted for whole genome resequencing on a paired-end Illumina HiSeq X Ten machine at RAPiD Genomics.

2.2 | Read alignment, variant calling and filtering

Raw reads were trimmed for Illumina adapters using Trimmomatic version 0.36 (Bolger et al., 2014) with the following parameters: "ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:8:true", resulting in an average of 35,689,979 paired reads per sample (average depth = 5.1× in Downy and 4.5× in Hairy Woodpecker). Read quality was assessed with FastQC version 0.11.4. (Andrews, 2010). Given the high synteny and evolutionary stasis of bird chromosomes (Damas et al., 2018; O'Connor et al., 2019; Singhal et al., 2015), we produced a chromosome-length reference genome for Downy Woodpecker by ordering and orienting the scaffolds and contigs of the Downy Woodpecker genome assembly (Jarvis et al., 2014) along the 35 chromosomes of the Zebra Finch (*Taeniopygia guttata*; version taeGut3.2.4) using Chromosemble from the Satsuma package (Grabherr et al., 2010). We verified the completeness of this new reference by searching for a set of single-copy avian orthologues using benchmarking universal single-copy orthologues (BUSCO version 2.0.1: Waterhouse et al., 2018). We finally transferred the prediction-based genome annotation of the Downy Woodpecker (Jarvis et al., 2014) by mapping the genomic coordinates of each annotated feature against the pseudochromosome reference using gmap (Wu & Watanabe, 2005).

Trimmed reads for both Downy and Hairy Woodpecker were aligned against the pseudochromosome reference genome of the Downy Woodpecker using BWA version 0.7.15 mem algorithm (Li & Durbin, 2009). On average, 97.27% of reads from Downy Woodpecker and 96.38% of reads from Hairy Woodpecker were successfully mapped, demonstrating that despite the large evolutionary distance between these two species (Dufort, 2016), sequence conservation allows efficient mapping. Resulting sequence alignment/map (SAM) files were converted to their binary format (BAM) and sequence group information was added. Next, reads were sorted, marked for duplicates, and indexed using Picard (http://www.broadinstitute.github.io/picard/). The Genome Analysis Toolkit (GATK version 3.6; DePristo et al., 2011) was then used to perform local realignment of reads near insertion and deletion (indels) polymorphisms. We first used the RealignerTargetCreator tool to identify regions where realignment was needed, then produced a new set of realigned binary sequence alignment/map (BAM) files using IndelRealigner. The final quality of mapping was assessed using QualiMap version 2.2.1 (Okonechnikov et al., 2016).

We implemented two complementary approaches for the downstream analysis of genetic polymorphism. First, we used ANGSD version 0.917 (Korneliussen et al., 2014), a method that accounts for the genotype uncertainty inherent to low depth sequencing data by inferring genotype likelihoods instead of relying on genotype calls. We estimated genotype likelihoods from BAM files using the GATK model (-GL 2; DePristo et al., 2011), retaining only sites present in at least 70% of sampled individuals (-minInd 50) and with the following filters: a minimum mapping quality of 30 (-minMapQ 30), a minimum quality score of 20 (-minQ 20), a minimum frequency of the minor allele of 5% (-minMaf 0.05), and a *p*-value threshold for the allele-frequency likelihood ratio test statistic of 0.01 (-SNP_pval 0.01). Allele frequencies were estimated directly from genotype likelihoods assuming known major and minor alleles (-doMajorMinor 1 -doMaf 1). A total of 16,736,465 and 15,463,356 SNPs were identified for Downy and Hairy Woodpecker, respectively. Because several downstream analyses lack support for genotype likelihoods, we also called genotypes using GATK version 3.8.0 (McKenna et al., 2010). First, we run HaplotypeCaller separately for each sample using the --emitRefConfidence GVCF -minPruning 1 -minDanglingBranchLength 1 options to create one gVCF per individual, then we ran GenotypeGVCFs with default settings across all samples to jointly call genotypes. In the absence of a training SNP panel for our non-model species, we applied hard filtering recommendations from the Broad Institute's Best Practices (https://www.gatk.broadinstitute.org/). We filtered SNPs with quality by depth below 2 (QD < 2.0), SNPs where reads with the alternative allele were shorter than those with the reference allele (ReadPosRankSum < −8), SNPs with evidence of strand bias (FS > 60.0 and SOR > 3.0), SNPs with root mean square of the mapping quality below 40 (MQ < 40.0), and SNPs in reads where the alternative allele had a lower mapping quality than the reference allele (MQRankSumTest < −12.5). In addition,

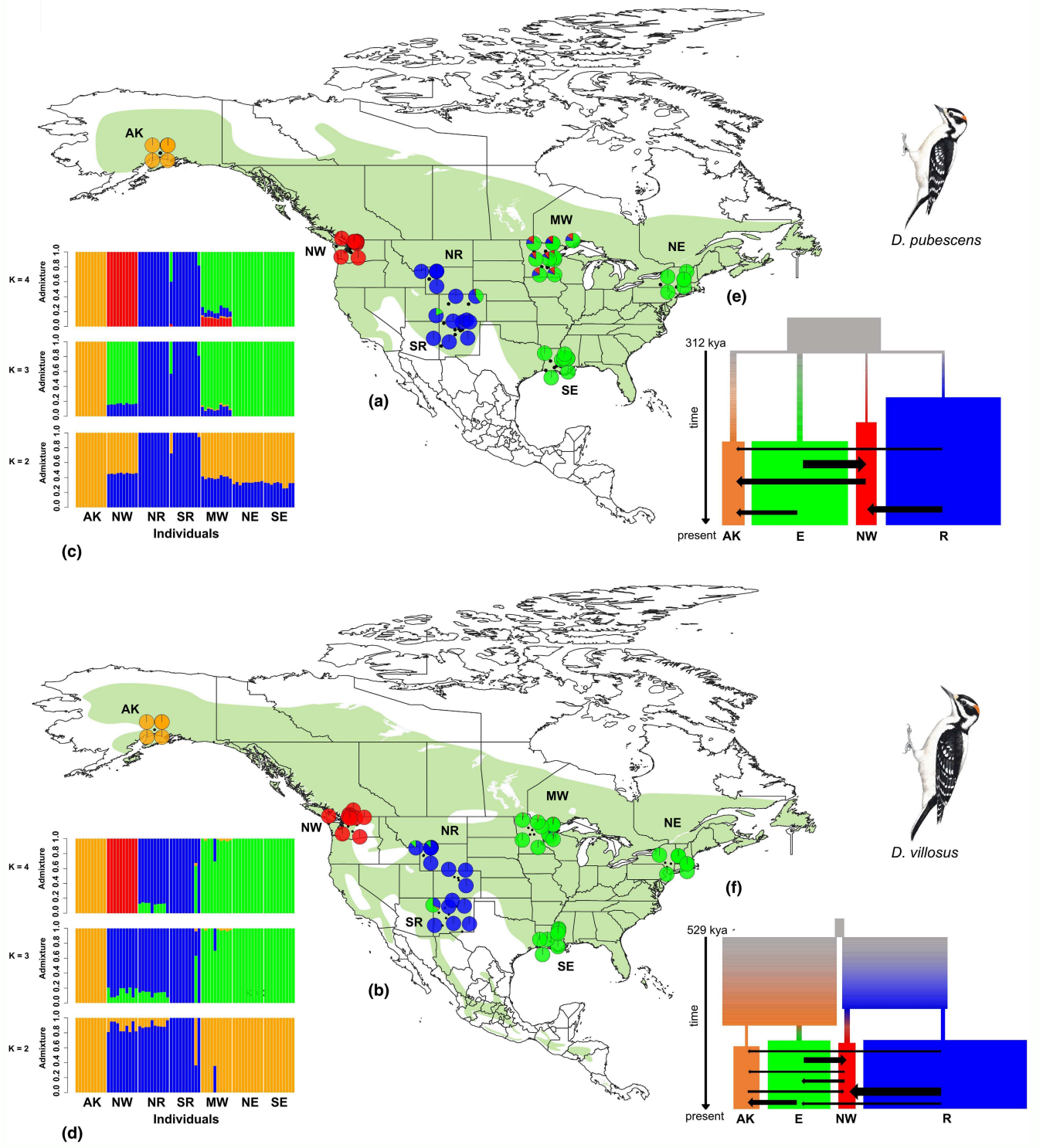

FIGURE 1 Geographic distribution of genetic variation and demographic history of the Downy (*D. pubescens*; top) and Hairy Woodpecker (*D. villosus*; bottom). (a, b) Admixture proportions from *NGSadmix* for $K = 2$–4. Each bar indicates an individual's estimated ancestry proportion for each genetic cluster, represented by different colours. (c, d) Map indicating the current range of Downy and Hairy Woodpecker (green shade), the locality of the samples, and their respective admixture proportions from *NGSadmix* (pie charts). (e, f) The best-fit demographic models from fastsimcoal2. The width of the rectangles and arrows are scaled relative to the estimated effective population sizes in haploid individuals (N_e) and the migration rate (m) in fraction of haploid individuals from donor population per generation, respectively. Only the values of migration rate $> 10^{-7} \times N_e$ migrants per generation are shown. Note that Downy Woodpecker (*D. pubescens*) is about two-thirds of the size of Hairy Woodpecker (*D. villosus*) and its bill is around one-third of the length of its head. In contrast, the bill of Hairy Woodpecker is almost as long as its head. Populations: AK, Alaska; E, East; NW, Pacific Northwest; R, Rockies; Sampling regions: AK, Alaska; MW, Midwest; NE, Northeast; NR, Northern Rockies; NW, Pacific Northwest; SE, Southeast; SR, Southern Rockies. Illustrations reproduced with permission from Lynx Edicions. [Colour figure can be viewed at wileyonlinelibrary.com]

we used VCFtools version 0.1.17 (Danecek et al., 2011) to retain only biallelic SNPs occurring in at least 75% of samples, with a minimal mean coverage of 2×, a maximum mean coverage of 100×, and a *p*-value above .01 for the exact test for Hardy–Weinberg equilibrium. We applied three different minor allele frequency (maf) thresholds – no threshold (for demographic analyses), 0.02 (for the estimation of recombination rates), and 0.05 (for all remaining analyses). After QC, we ended up with a total of 7,009,778 SNPs in Downy Woodpecker and 4,579,046 SNPs in Hairy Woodpecker.

2.3 | Population structure

We investigated how genetic diversity is distributed across populations of Downy and Hairy Woodpecker by first performing a principal components analysis (PCA) using the R package SNPRelate version 3.3 (Zheng et al., 2012). We first applied the function snpgdsLDpruning to select a subset of unlinked SNPs (LD r^2 threshold = 0.2), with <25% missing data and a maf >0.05, which resulted in a total of 71,228 SNPs for Downy Woodpecker and 71,763 SNPs for Hairy Woodpecker. We then used the function snpgdsPCA to calculate the eigenvectors and eigenvalues for the principal component analysis. We investigated population structure by looking at the first three principal components (PC1–PC3). In addition, we used NGSadmix (Skotte et al., 2013), implemented in ANGSD (Korneliussen et al., 2014), to investigate the number of genetic clusters, and associated admixture proportions for each individual. NGSadmix is a maximum likelihood approach analogous to STRUCTURE (Pritchard et al., 2000), but bases its inferences on genotype likelihoods instead of SNP calls, therefore accounting for the uncertainty of genotypes. We ran NGSadmix for $K = 2$–7 (Figure S2), but only results for 2–4 revealed interpretable population structure.

We also described the relationships among sampling regions by building a maximum likelihood tree based on the polymorphism-aware phylogenetic model (PoMo; Schrempf et al., 2016) implemented in IQ-Tree 2 (Minh et al., 2020). PoMo is a phylogenetic method that accounts for incomplete lineage sorting inherent to population-level data by incorporating polymorphic states into DNA substitution models. We used a python script (https://www.github.com/pomo-dev/cflib) to convert our vcf files containing only intergenic SNPs into the input format of PoMo (counts file). IQ-Tree was run using the HKY+P model of sequence evolution with 100 non-parametric bootstraps to assess support. We used three samples from Hairy Woodpecker as an outgroup to root the tree for Downy Woodpecker, and vice versa.

We estimated pairwise genetic differentiation (F_{ST} values) among sampling regions in each species using ANGSD version 0.917 (Korneliussen et al., 2014). We first produced site-allele-frequency likelihoods using the command -doSaf with no site filters, followed by the realSFS -fold 1 command to generate a folded site frequency spectrum (SFS). We then estimated weighted F_{ST} values using the realSFS F_{ST} command both globally and across nonoverlapping 100 kb windows.

We investigated patterns of gene flow across the landscape using the estimated effective migration surface (EEMS; Petkova et al., 2016), which is a method to visualize variation in patterns of gene flow across a habitat. EEMS compares pairwise genetic dissimilarity among localities to identify geographic areas that deviate from the null expectation of isolation by distance (IBD). Low values of relative effective migration rate (m) indicate a rapid decay in genetic similarity in relation to geographic distances, which suggests the presence of barriers to gene flow. In contrast, high values of m indicate larger genetic similarity than expected given the geographic distance, suggesting genetic connectivity. We generated pairwise identity-by-state (IBS) matrices using the -doIBS function in ANGSD (Korneliussen et al., 2014) and used these matrices to represent dissimilarity between individuals. We ran EEMS using 200 demes and performed five MCMC chain runs with 1×10^7 iterations following a burnin of 5×10^6, and a thinning of 9999, checking the posterior probabilities to ensure convergence.

2.4 | Demographic inference

We inferred past changes in effective population size (N_e) using stairway plot 2 (Liu & Fu, 2020), a method that leverages information contained in the site frequency spectrum (SFS) to estimate recent population history. Unlike methods based on the sequentially Markov coalescent (e.g., PSMC, SMC++), stairway plot 2 is both applicable to a large sample of unphased whole genome sequences and insensitive to read depth limitations. We estimated the folded site frequency spectrum for each population (i.e., genetic cluster) using the realSFS function in ANGSD (Korneliussen et al., 2014). For each population, we used the default 67% sites for training, and calculated median estimates and 95% pseudo-CI based on 200 replicates. We assumed a mutation rate of 4.007×10^{-9} mutations per site per generation, as estimated from synonymous sites of the Northern Flicker's genome (Hruska & Manthey, 2021) and a generation time of 1 year for both species. We then utilized the estimates of N_e from stairway plot 2 across the past 500 kya to calculate the harmonic mean on linear-stepped time points, representing each population's long-term N_e.

We further investigated the demographic history of the two species using fastsimcoal2 version 2.6.0.3, a composite likelihood method that uses the joint site frequency spectrum (jSFS) to perform model selection and estimate demographic parameters (Excoffier & Foll, 2011). We tested the support for two competing demographic models: (i) a model where all populations diverge synchronously from a single large refugium and expand independently with asymmetric gene flow, and (ii) a bifurcating model where populations diverge at different times from multiple refugia and expand independently with asymmetric gene flow. Since we only need a reasonably large subset of the genome to get an accurate estimate of the site frequency spectrum (Beichman et al., 2018), we generated the four-population folded jSFS from a set of high quality SNPs with no maf filtering present in

chromosome 1 (total number of sites in chr. 1 = 107,800,922; number of variable sites = Downy: 6,030,759; Hairy: 7,967,215) using easySFS.py (https://www.github.com/isaacovercast/easySFS). We projected the jSFS down to 20 chromosomes (i.e., 10 diploid samples) per population to avoid issues associated with differences in sample size and missing data. To minimize the impact of selection, we only included sites in noncoding regions of the genome. All models followed the topology of the population tree obtained from IQ-Tree 2 and assumed a mutation rate of 4.007×10^{-9} mutations per site per generation. For each model, we conducted 75 iterations of the optimization procedure, each with 40 expectation conditional maximization cycles and 100,000 genealogical simulations per cycle. We performed model selection using the run with the highest likelihood for each model. For each species, we chose the model with the largest relative Akaike information criterion (AIC_w) as the best-fit model. We obtained 95% pseudo-CI for parameter estimates by performing 100 parametric bootstrap estimates simulating jSFSs under the best model and re-estimating parameters using these simulated datasets.

2.5 | Genetic diversity, recombination rates, and linkage disequilibrium

We compared genetic diversity among populations of the two species by estimating the genome-wide pairwise nucleotide diversity θ_π using ANGSD (Korneliussen et al., 2014). We first ran the command -doSaf 1 -minMapQ 30 -minQ 20 in ANGSD to generate site-allele-frequency likelihoods based on the GATK model (McKenna et al., 2010), then we used realSFS with the option -fold 1 to estimate the folded SFS. ANGSD was also used to estimate genome-wide Tajima's *D*. We estimated recombination rates (r =recombination rate per base pair [bp] per generation) along the genome of the two species using ReLERNN, a deep learning algorithm (Adrion et al., 2020). ReLERNN takes as input a vcf file and simulates training, validation, and test datasets matching the empirical distribution of θ_W. ReLERNN then uses the raw genotype matrix and a vector of genomic coordinates to train a model that directly predicts per-base recombination rates (as opposed to a population-scaled recombination rate) across sliding windows (Adrion et al., 2020). For each population, we used the SNP data set with maf > 0.02 and ran the analysis with default settings. Because ReLERNN is robust to demographic model misspecification (Adrion et al., 2020), we simulated an equilibrium model considering a mutation rate of 4.007×10^{-9} mutations per generation (Hruska & Manthey, 2021) and assuming a generation time of 1 year. Finally, we explored the recombination history of the individuals in each sampling region by analysing their patterns of linkage disequilibrium (LD) decay using PopLDdecay (Zhang et al., 2018). We calculated pairwise linkage (measured by D'/r^2) using the default maximum distance between SNPs of 300 kb and plotted it as a function of genomic distance (in kb).

2.6 | Genomic predictors of regional variation in nucleotide diversity

To investigate the factors shaping the genomic landscape of diversity in the two woodpecker species, we tested the effect of (i) recombination rate, (ii) gene density, and (iii) GC content on regional patterns of nucleotide diversity. For each population, we computed pairwise nucleotide diversity (θ_π) across 100 kb nonoverlapping windows using ANGSD (Korneliussen et al., 2014). We first used the -doThetas function to estimate the site-specific nucleotide diversity from the posterior probability of allele frequency (SAF) using the estimated site frequency spectrum (SFS) as a prior. Then, we ran the thetaStat do_stat command to perform the sliding windows analysis. To quantify variation in recombination rates, we calculated weighted averages of recombination rates estimated in ReLERNN across 100 kb nonoverlapping windows (weighted by window length contribution). To test the effect of window size on downstream correlational analyses, we also estimated nucleotide diversity and recombination rates over a larger (1 Mb) window. We assessed gene density (i.e., density of targets of selection) as the proportion of coding sequence (in number of base pairs) for any given 100 kb nonoverlapping window and estimated GC content in each 100 kb nonoverlapping window using the function *GC* of the R package seqinr version 3.6-1 (Charif & Lobry, 2007). We fit a general linear regression in R to assess the relationship between nucleotide diversity (θ_π) and the three predictor variables – recombination rate, gene density, and base composition. We also fit a LOESS model to account for the potential nonlinearity of these relationships using the R package caret (Kuhn, 2008). Models were trained using cross-validation of 80% of the total data. To control for the collinearity among these variables, we also ran a principal component regression (PCR). PCR is a technique that summarizes the predictor variables into orthogonal components (PCs) before performing regression, therefore removing the correlation among variables. PCR was conducted using the R package pls (Wehrens & Mevik, 2007). All variables were Z-transformed before these analyses.

We also explored the association between patterns of intraspecific population differentiation (F_{ST}) and intrinsic properties of the genome (i.e., nucleotide diversity and recombination rates). To summarize the genomic landscape of differentiation into a single response variable, we employed two approaches: for each 100 kb window, we (i) calculated the average F_{ST} across all pairwise sampling region comparisons; (ii) we performed a principal component analysis and extracted that first principal component (PC1) that explained the greatest covariance among all pairwise sampling region comparisons (Downy: variance explained = 37.51%; Hairy: variance explained = 47.5%). Summaries of F_{ST} produced by these two approaches were highly correlated (Downy: Pearson's $r = 0.97$; $p < .001$; Hairy: Pearson's $r = 0.98$; $p < .001$), so we only considered the average F_{ST} for simplicity.

2.7 | Natural selection and genetic load

To understand the impact of natural selection across populations of Downy and Hairy Woodpecker, we first estimated the deleterious load of each species and populations. We used the software snpEff version 4.1 (Cingolani et al., 2012) to classify SNPs into one of four categories of functional impact, according to the predicted effect of the gene annotation – (i) modifiers: variants in noncoding regions of the genome (e.g., introns, intergenic) whose effects are hard to predict; (ii) low: variants in coding sequences that cause no change in amino acid (i.e., synonymous); (iii) moderate: variants in coding sequences that cause a change in amino acid (i.e., nonsynonymous); and (iv) high: variants in coding sequences that cause gain or loss of start and stop codon. We then selected a subset of 12 individuals with the lowest percentage of missing data (therefore, maximizing the total number of sites) in each species to polarize our SNPs. To do so, we looked for biallelic SNPs in Downy Woodpecker for which one of the alleles were fixed in Hairy Woodpecker and vice versa. The allele fixed in the outgroup was assumed to be the ancestral state. This is a sensitive step in the estimation of genetic load, so we only kept SNPs for which the ancestral state could be determined unambiguously (Simons & Sella, 2016). We ended up with a total set of 363,903 polarized SNPs across the genome.

We then characterized the site-frequency spectrum (SFS) for each type of variant (according to the impact inferred from snpEff) by estimating the total frequency of each derived allele and calculating the proportion of each allele-frequency bin. As a proxy for genetic load for each individual we estimated the ratio of the number of homozygous derived alleles of high impact (i.e., loss of function) over the number of homozygous derived alleles of low impact (i.e., synonymous). This metric is a proxy for the genetic load under a recessive model while controlling for the underlying population differences in the neutral SFS (Simons et al., 2014; Simons & Sella, 2016). It assumes that derived alleles are only deleterious when in a homozygous state. We therefore also considered an additive model (i.e., semi-dominant) that assumes that derived alleles have deleterious effects in both homozygosity and heterozygosity. For this metric, we counted the total number of derived alleles, instead of only the ones in homozygosity (Simons & Sella, 2016).

The effect of linked selection is expected to be weaker in populations that underwent more severe bottlenecks due to their smaller long-term N_e when compared to populations that maintained large N_e (Charlesworth & Willis, 2009; Kirkpatrick & Jarne, 2000). We tested this prediction by quantifying the strength of correlation between nucleotide diversity (θ_π), recombination rate, and gene density in all four populations of Downy and Hairy Woodpecker, which showed varied demographic responses to the Pleistocene glaciations. For each population, we estimated 95% CI for Pearson's correlation coefficient using the R package metan (Olivoto & Lúcio, 2020).

To assess the impact of selection over a deeper evolutionary scale, we estimated d*N*/d*S*, the ratio of nonsynonymous over synonymous substitution, using a set of 397 genes that were orthologous across Downy Woodpecker, Hairy Woodpecker and two avian outgroups – Chicken (*Gallus gallus*) and Zebra Finch (*Taeniopygia guttata*). We identified orthologous genes across all four species using the software JustOrthlogs (Miller et al., 2019) and only kept well-aligned loci. We first downloaded Ensembl genome assemblies and gene annotations for version GRCg6a and bTaeGut1_v1.p of the Chicken and Zebra Finch genome, respectively (Ensembl version 103). We then extracted coding sequences (CDS) for all identified orthologues from their respective reference genomes using a GFF3 parser included in JustOrthologs and aligned them with the frameshift-aware MACSE software (Ranwez et al., 2011). We used the parameter setting --min_percent_NT_at_ends 0.3 and -codonForInternalStop *NNN* for aligning and exporting sequences. The resulting amino-acid alignments were inspected with HMMcleaner to mask sites that were likely misaligned. We finally used codeml to estimate the overall d*N*/d*S* ratio along each branch of the tree assuming a one-ratio branch model in PAML (Yang, 2007).

3 | RESULTS

3.1 | Congruent population structure and genetic diversity

Whole genome resequencing yielded an average sequencing depth of 5.1× (1.4–12.5×) for Downy Woodpecker and 4.5× (1.1–11.7×) for Hairy Woodpecker. A total of 91.1% of BUSCO genes were present and complete in our pseudochromosome reference, indicating sufficient completeness. In addition, 99.98% of all the 14,443 annotated genes in Downy Woodpecker were successfully mapped to the pseudochromosome reference. A total of 16,736,465 and 15,463,356 single nucleotide polymorphisms (SNPs; $n = 70$ samples per species) were identified in the Downy and Hairy Woodpecker genomes, respectively, using the genotype likelihood approach implemented in ANGSD (Korneliussen et al., 2014).

Our principal component analysis (PCA) recovered congruent genetic structure across both species' ranges (Figure 2a,c; Figures S3–S5). The three first principal components (PCs) explained together 7.5% (Downy) and 8.3% (Hairy) of the total genetic variance. Geographic structure was generally characterized by a genetic distinction between boreal-eastern and western sampling regions. In Downy Woodpecker, however, the Pacific Northwest sampling region fell more closely related to the Eastern group than the Western group (Figure 2a). Consistent with these findings, *NGSadmix* (Skotte et al., 2013) supported four geographically congruent genetic clusters (hereafter referred to as "populations"; $K = 4$) in the Downy and Hairy Woodpecker: East (NE, SE, and MW), Pacific Northwest (NW), Rocky Mountains (SR and NR), and Alaska (AK; Figure 1c,d). The average genome-wide estimate of F_{ST} was slightly larger in Hairy Woodpecker (average F_{ST} = 0.1; 0.03–0.19) than Downy Woodpecker (average F_{ST} = 0.08; 0.03–0.16), indicating larger (but overlapping) levels of population differentiation (Table S2). In both species, the largest values of F_{ST} involved comparisons between Alaska and other sampling regions (Downy: F_{ST} [AK vs. NR] = 0.16; Hairy: F_{ST} [AK vs. SR] = 0.19),

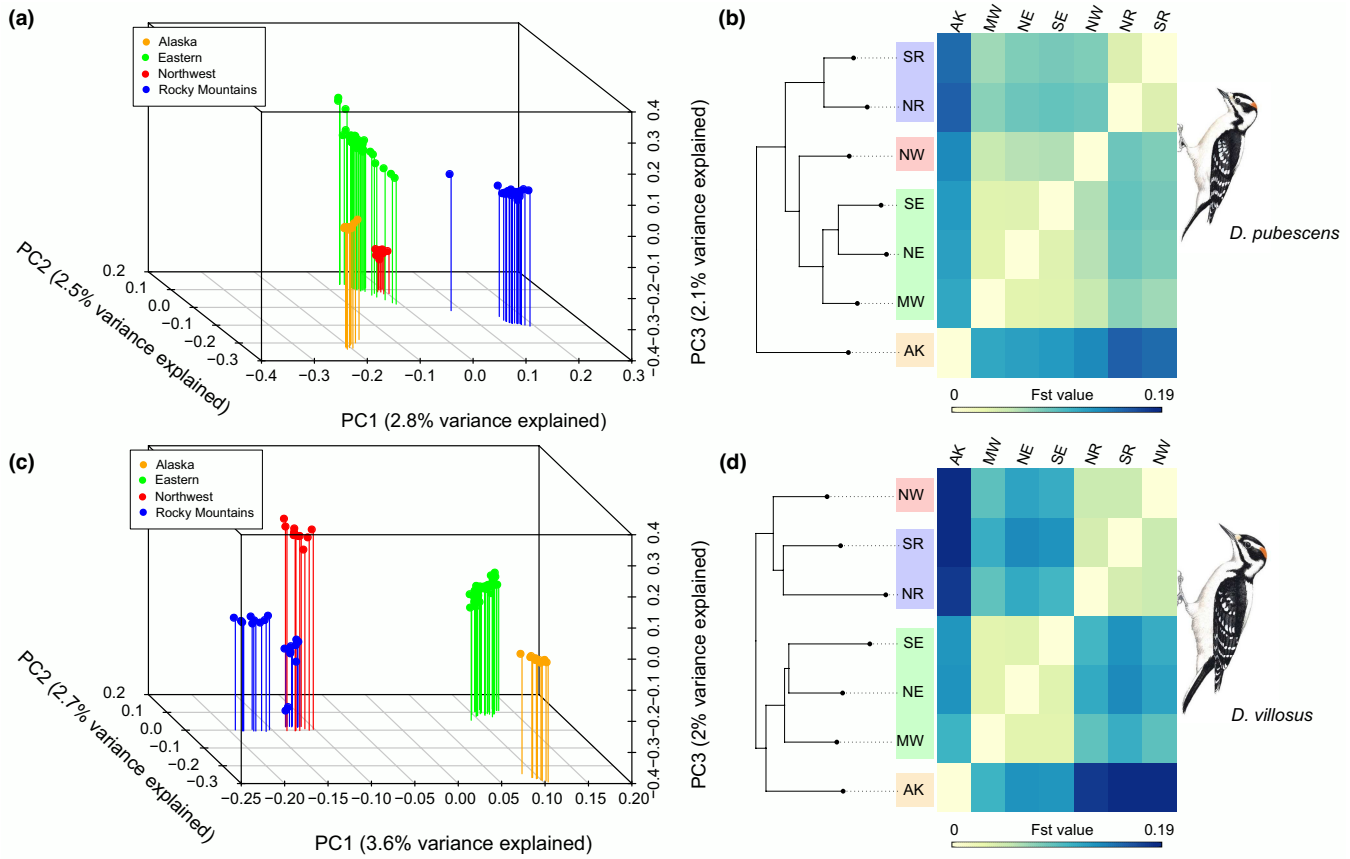

FIGURE 2 Population genetic structure in the Downy (*D. pubescens*; top) and Hairy (*D. villosus*; bottom) Woodpecker. (a, c) Principal component analysis (PCA) of Downy and Hairy Woodpecker based on 71,228 and 71,763 unlinked genome-wide SNPs, respectively, with <25% missing data and a minor allele frequency (maf) > 0.05. (b, d) Heatmap showing genome-wide pairwise F_{ST} values and associated maximum likelihood tree based on the polymorphism-aware phylogenetic model (PoMo) in IQ-Tree 2. All nodes show 100% bootstrap support. Darker colours on the heatmap correspond to larger values of F_{ST}. Sampling regions: AK, Alaska; MW, Midwest; NE, Northeast; NR, Northern Rockies; NW, Pacific Northwest; SE, Southeast; SR, Southern Rockies. Illustrations reproduced with permission from Lynx Edicions. [Colour figure can be viewed at wileyonlinelibrary.com]

and the lowest were within the East and the Rocky Mountains clusters (Downy: $F_{ST} = 0.03–0.06$; Hairy: $F_{ST} = 0.03–0.04$).

In both species, the EEMS analysis detected a pronounced reduction in effective migration near the Great Plains and along the Rocky Mountains, especially in its Northern portion. In contrast, eastern North America showed a higher degree of connectivity when compared to the west (Figure 3). This finding indicates that major topographic features and variation in habitat availability might have contributed to the maintenance of population differentiation, despite the presence of gene flow. These spatial discontinuities might also reflect the recent contact of populations expanding out of glacial refugia.

3.2 | Demographic history

Temporal variation in effective population size was generally consistent between species, being characterized by recurrent episodes of bottleneck followed by population expansion (Figure 4a,b; Appendix S1). We found that within each population, nucleotide diversity was highly correlated with the harmonic mean of the N_e estimated from stairway plot 2 over the past 500 kya (long-term N_e; linear regression: $t = 4.876$; $R^2 = 0.76$; $p < .002$; Figure S6), indicating these independent analyses were consistent.

The maximum likelihood tree from genome-wide intergenic SNPs for Hairy Woodpecker showed two distinct clades – an East + Alaska and a West clade (Figure 2b,d). The topology for Downy Woodpecker, however, revealed different relationships. First, the Pacific Northwest sampling region (NW) was more closely related to the eastern clade than to the western clade, supporting our PCA analysis. In addition, the Alaska (AK) sampling region was sister to all other sampling regions. We tested whether (i) all four populations diverged synchronously from a single ancestral refugium, or whether (ii) populations diverged at different times from multiple ancestral refugia. Our fastsimcoal2 results show differing support for these alternative demographic hypotheses among the two focal species. The best-supported model for Hairy Woodpecker was model ii (Table S3; Figure 1f; Figure S7), in which two ancestral populations (representing a split between eastern and western North America) diverged from each other around 529 kya (95% CI = 513–561 kya;

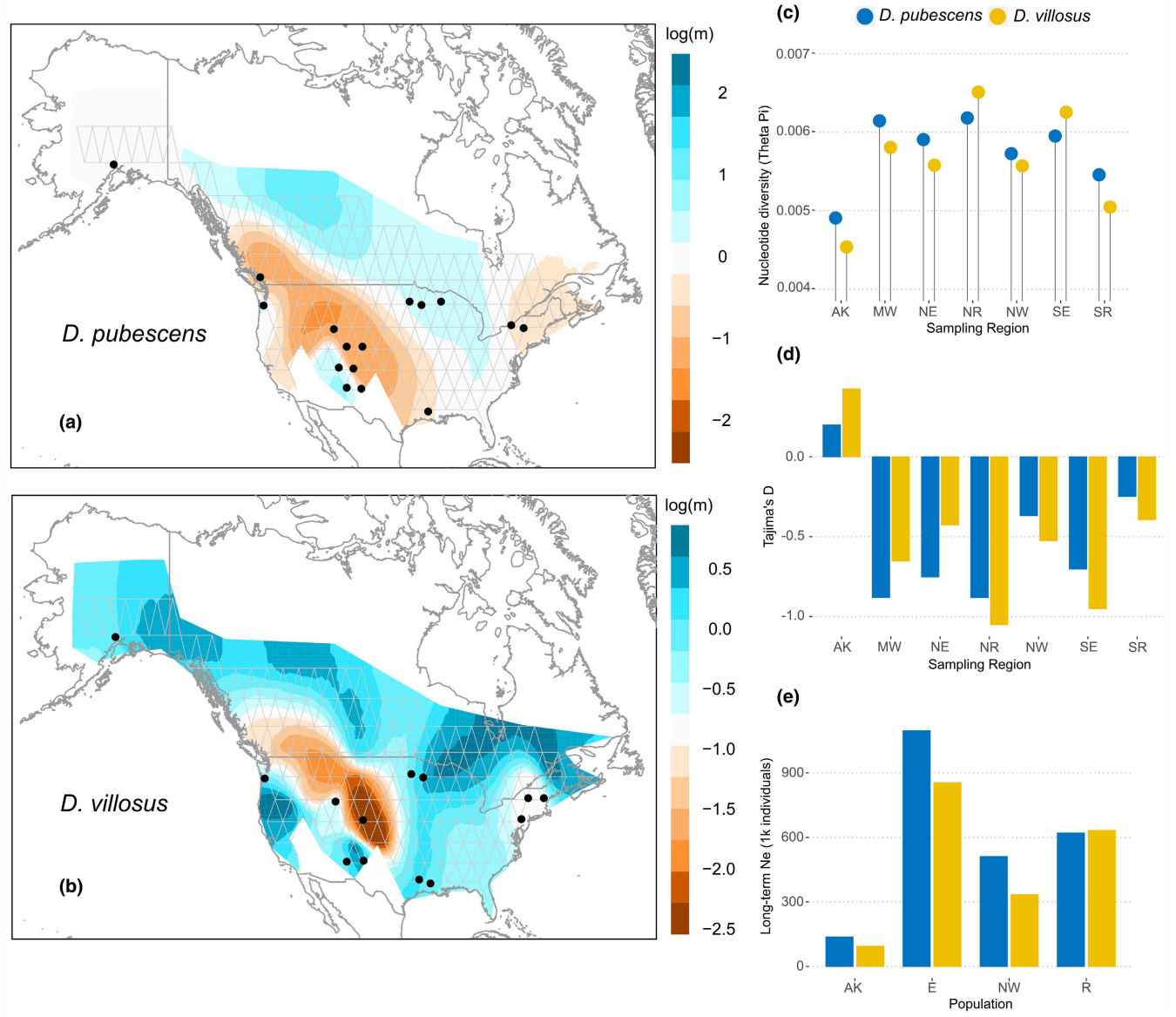

FIGURE 3 Spatial patterns of gene flow and genome-wide genetic variation in Downy (*D. pubescens*) and Hairy Woodpecker (*D. villosus*). (a) Effective migration surface inferred by EEMS in Downy Woodpecker and (b) Hairy Woodpecker. Warmer colours indicate lower and colder colours indicate higher effective migration (on a log scale) relative to the overall migration rate over the species range. Triangles represent the grid chosen to assign sampling locations to discrete demes. (c) Genome-wide pairwise nucleotide diversity (θ_π) per sampling region. (d) Genome-wide Tajima's *D* per sampling region. (e) The harmonic mean of effective population size (N_e) estimated over the past one million years with stairway plot 2 for all four genetic clusters. Populations: AK, Alaska; E, East; NW, Pacific Northwest; R, Rockies; Sampling regions: AK, Alaska; MW, Midwest; NE, Northeast; NR, Northern Rockies; NW, Pacific Northwest; SE, Southeast; SR, Southern Rockies. [Colour figure can be viewed at wileyonlinelibrary.com]

Table S4) and gave rise to the four populations, which in turn underwent strong bottlenecks. A final explosive expansion then occurred between 193–212 kya when populations grew up to 12-fold. In contrast, Downy Woodpecker showed support for model *i*, in which all populations diverge from a single major refugium (Table S3; Figure 1e; Figure S7). This divergence occurred around 312 kya (95% CI = 146–551 kya; Table S4) and was accompanied by a large bottleneck, reducing N_e to less than 10% of its original size in most populations. A final population expansion then occurred at the end of the Mid-Pleistocene (152–232 kya). Overall, estimates of N_e from fastsimcoal2 confirmed the trends observed in stairway plot 2, albeit with less resolution. We found large and variable levels of post-expansion gene flow across populations in both species (Downy: 0–4.8 migrants per generation; Hairy: 0–6.66 migrants per generation) that confirmed our EEMS migration surfaces.

3.3 | Genomic correlates of nucleotide diversity and differentiation

We found that nucleotide diversity varied widely along the genome ($\theta_{\pi\ \text{Downy}} = 6\times10^{-4}$–$1.84\times10^{-2}$; $\theta_{\pi\ \text{Hairy}} = 4\times10^{-4}$–$1.94\times10^{-2}$), but

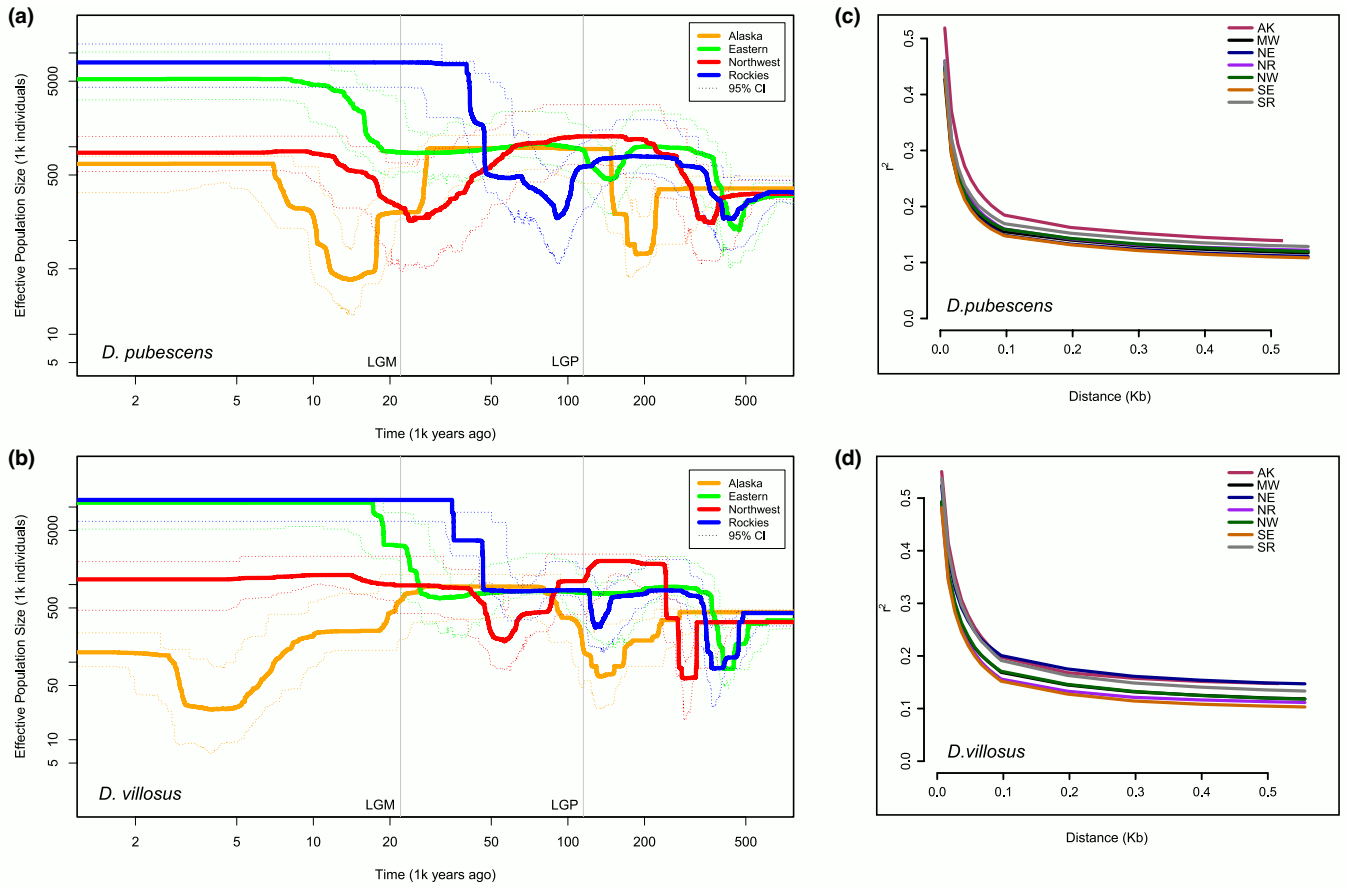

FIGURE 4 Changes in effective population size (N_e) over time and linkage disequilibrium (LD) in Downy (*D. pubescens*; top) and Hairy (*D. villosus*; bottom) Woodpecker. (a, b) Inferred history of effective population size using all four genetic clusters in Downy (a) and Hairy Woodpecker (b) obtained with stairway plot 2 using the folded site frequency spectrum (SFS). For this analysis, we specified a mutation rate of 4.007×10^{-9} mutations per site per generation. Both axes are represented in a log scale. Dotted lines represent 95% confidence intervals, and vertical lines represent the Last Glacial Period (LGP; 115 kya) and the Last Glacial Maximum (LGM; 21 kya). (c, d) Decay of linkage disequilibrium (LD) in all seven sampling locations of Downy (c) and Hairy (d) Woodpecker. Populations: AK, Alaska; E, East; NW, Pacific Northwest; R, Rockies; Sampling regions: AK, Alaska; MW, Midwest; NE, Northeast; NR, Northern Rockies; NW, Pacific Northwest; SE, Southeast; SR, Southern Rockies. [Colour figure can be viewed at wileyonlinelibrary.com]

this variation was highly correlated between Downy and Hairy Woodpecker (Pearson's $r = 0.9$; $p < .001$; Appendix S1; Figure S8). We found recombination rates to be highly correlated between the two species as well (Pearson's $r = 0.66$; $p < .001$; Figure S8). Across the genome, we estimated a mean per-base recombination rate (r) = 2.64×10^{-9} c/bp (4.2×10^{-12}–3.97×10^{-9}) in Downy Woodpecker and $r = 2.62 \times 10^{-9}$ c/bp (3.51×10^{-10}–4.03×10^{-9}) in Hairy Woodpecker. Considering the average long-term N_e of Downy and Hairy Woodpecker as approximately 1×10^6 in the East population, these recombination rates correspond to a population-scaled rate $\rho = 4N_e r = 0.01$. Mean recombination rates were 2–3-fold higher in autosomal chromosomes compared to the sex-linked Z chromosome (Figures S9, S10), consistent with differences in N_e between sex chromosomes (Sundström et al., 2004; Xu et al., 2019; Zhou et al., 2014). As a result of both high recombination rates and large N_e, we also observed that linkage disequilibrium (LD) in Downy and Hairy Woodpecker decays very rapidly. LD drops to half of its initial levels in less than 100 bp (Figure 4c,d). Consistently, the average LD was greater for sampling regions with smaller N_e or populations that have likely experienced a more recent founder event, such as Alaska and the Southern Rockies (Figure 4c,d). We found a significant positive association between nucleotide diversity (θ_π) and recombination rates in all populations of both species with a 100 kb window size (Figure 5a,b; linear regression for East – Downy: $t = 66.09$, $R^2 = 0.27$, $p < .001$; Hairy: $t = 101.7$, $R^2 = 0.47$, $p < .001$; LOESS regression for East – Downy: $R^2 = 0.31$; Hairy: $R^2 = 0.47$), and an even stronger relationship with the 1 Mb windows (Downy: $R^2 = 0.34$; Hairy: $R^2 = 0.6$; Figure S11). This association, however, was expected (to a certain extent) even if diversity was not correlated with recombination rates because recombination rates were estimated directly from θ_w.

To investigate the impact of linked selection on the genomic landscape of diversity, we also tested the prediction that regions of the genome with a higher density of targets of selection (i.e., coding sequence) exhibit lower nucleotide diversity. Our results revealed a weak but significant negative association between nucleotide diversity (θ_π) and gene density in all populations (Figure 5c,d; linear regression for East – Downy: $t = -11.67$, $R^2 = 0.011$, $p < .001$; Hairy:

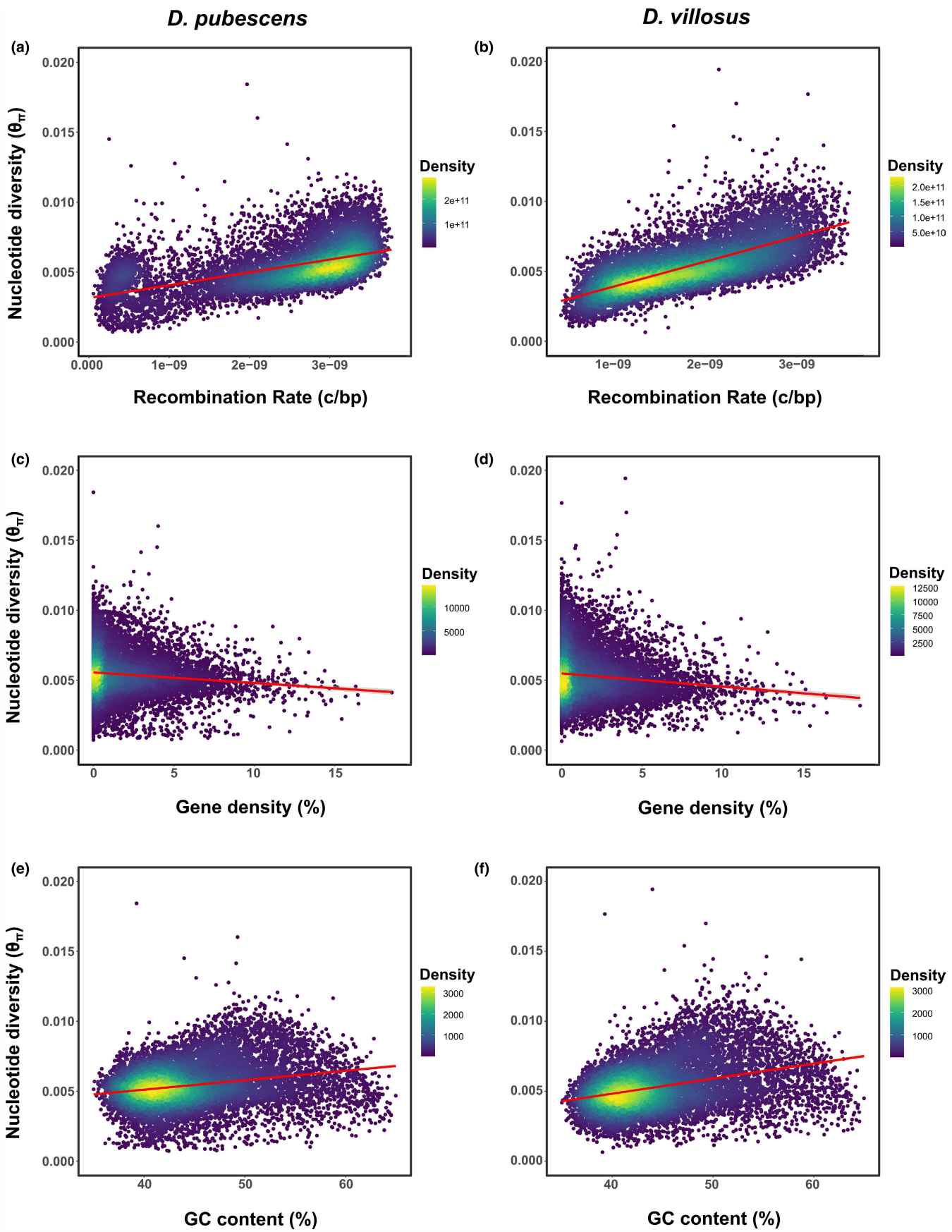

FIGURE 5 Genomic predictors of nucleotide diversity in Downy (*D. pubescens*; left) and Hairy (*D. villosus*; right) Woodpecker. Association between nucleotide diversity (θ_π) and three features of the genome: (a, b) recombination rates (Downy: $t = 66.09$, $p < .001$; Hairy: $t = 101.7$, $p < .001$), (c, d) gene density (Downy: $t = -11.67$, $p < .001$; Hairy: $t = -13.08$, $p < .001$), and (e, f) GC content (Downy: $t = 26.47$, $p < .001$; Hairy: $t = 38.89$, $p < .001$). Each point in the scatter plot represents a 100 kb window of the genome. Colours indicate the density of points. [Colour figure can be viewed at wileyonlinelibrary.com]

$t = -13.08$, $R^2 = 0.014$, $p < .001$; LOESS regression for East – Downy: $R^2 = 0.012$; Hairy: $R^2 = 0.016$). This association was not driven by the collinearity between gene density and recombination, which was positive and negligible (Downy: Pearson's $r = 0.034$, $p < .001$; Hairy: Pearson's $r = 0.058$, $p < .001$). We also found that regions with high GC content tended to show higher nucleotide diversity (Figure 5e,f; linear regression for East – Downy: $t = 26.47$, $R^2 = 0.057$, $p < .001$; Hairy: $t = 38.89$, $R^2 = 0.116$, $p < .001$; Appendix S1). GC content, however, was positively correlated with gene density in both species (Downy: Pearson's $r = 0.25$; $p < .001$; Hairy: Pearson's $r = 0.25$; $p < .001$; Appendix S1; Figures S12, S13) and correlated with recombination rates only in Hairy Woodpecker (Pearson's $r = 0.35$; $p < .001$; Appendix S1; Figures S11, S12). We then performed a principal component regression (PCR) to separate the effect of individual explanatory variables (recombination rate, GC content, and gene density) and control for the multicollinearity among predictor variables. Principal component regression summarizes variables into orthogonal components (PCs) and uses these components as predictors in a linear regression. PC2, which represented almost exclusively recombination rate in Downy Woodpecker and both recombination rate and gene density in Hairy Woodpecker (Table 1), uniquely explained 24% and 27.8% of variation in nucleotide diversity in Downy and Hairy Woodpecker, respectively (PC2 linear regression – Downy: $t = 60.11$, $R^2 = 0.24$, $p < .001$; Hairy: $t = 66.45$, $R^2 = 0.27$, $p < .001$). In Downy Woodpecker, both PC1 and PC3 represented the correlation between gene density and GC content, but PC3 had a much stronger effect (Table 1), accounting for 11.61% of the variation in nucleotide diversity. In Hairy Woodpecker, on the other hand, PC1 (represented mostly by GC content and recombination rate) explained more variation in nucleotide diversity than PC3 (Table 1). Our analyses confirm the central role that these genomic properties played in shaping patterns of nucleotide diversity along the genome.

We found that the coefficient of correlation between genetic diversity and the density of targets of selection varied across populations, albeit only slightly (Figure 6; Table S5). Alaska showed the weakest correlation (Downy: Pearson's $r = -0.1$, $t = -10.8$, $p < .001$; Hairy: Pearson's $r = -0.1$, $t = -11.6$, $p < .001$), whereas Rocky Mountains showed the strongest (Downy: Pearson's $r = -0.11$, $t = -11.9$, $p < .001$; Hairy: Pearson's $r = -0.13$, $t = -14.5$, $p < .001$). We found a stronger correlation between genetic diversity and recombination rate for populations with a large N_e (e.g., Rockies and East) as opposed to smaller populations (e.g., Alaska and Northwest; Figure 6). The exception was the East population in Downy Woodpecker, which despite its large N_e had the lowest correlation coefficient (Figure 6; Table S5).

There was considerable variation in average intraspecific population differentiation (F_{ST}) along the genome (Downy: $F_{ST} = 0.01$–0.25; Hairy: $F_{ST} = 0.01$–0.32), indicating high variability in patterns of population differentiation. As expected, we recovered a significant negative association between average F_{ST} and nucleotide diversity, suggesting that areas of genome that show elevated differentiation tend to be characterized by reduced diversity (linear regression – Downy: $t = -19.12$, $R^2 = 0.03$; $p < .001$; Hairy: $t = -53.49$, $R^2 = 0.2$; $p < .001$; Figure 7). Finally, we found a negative association between average F_{ST} and recombination rates, indicating higher differentiation in regions of low recombination (linear regression – Downy: $t = -32.18$, $R^2 = 0.08$; $p < .001$; Hairy: $t = -41.55$, $R^2 = 0.13$; $p < .001$).

3.4 | Genetic load and the efficacy of selection

To further explore the magnitude of linked selection in the genomes of Downy and Hairy Woodpecker, we classified each variant according to their functional impact as predicted by the gene annotation. We found that the majority of identified SNPs in Downy and Hairy Woodpecker were classified as modifiers (Downy: 99.35%; Hairy: 99.13%), which are variants in intergenic or intronic regions whose impacts are hard to determine but tend to be neutral to nearly neutral. Low impact variants (i.e., synonymous mutations) characterized 0.46% and 0.64% of SNPs in Downy and Hairy Woodpecker, respectively. Moderate impact variants, mutations that cause a change in amino acid sequence (i.e., nonsynonymous mutations) represented 0.17% and 0.22% of the SNPs in Downy and Hairy Woodpecker, respectively. Finally, only 0.006% (Downy) and 0.007% (Hairy) of the SNPs were classified as high impact. These variants correspond to mutations that cause loss of function, such as loss or gain of a start or stop codon and are therefore expected to occur at very low frequencies.

We investigated differences in the burden of deleterious alleles carried by populations of Downy and Hairy Woodpecker that could reflect differences in the efficacy of purifying selection. Our results revealed that the frequency distribution of mutations with moderate

TABLE 1 Principal component regression

Species	Explanatory variables	% of variance explained (R^2)		
		PC1	PC2	PC3
Downy Woodpecker	Recombination rate	0.01	23.61	0.09
	Gene density	0.78	0.00	5.80
	GC content	0.77	0.41	5.70
	Total	1.58	24.03	11.61
Hairy Woodpecker	Recombination rate	7.52	9.36	0.59
	Gene density	4.52	18.45	0.24
	GC content	10.59	0.00	0.95
	Total	22.65	27.88	1.80

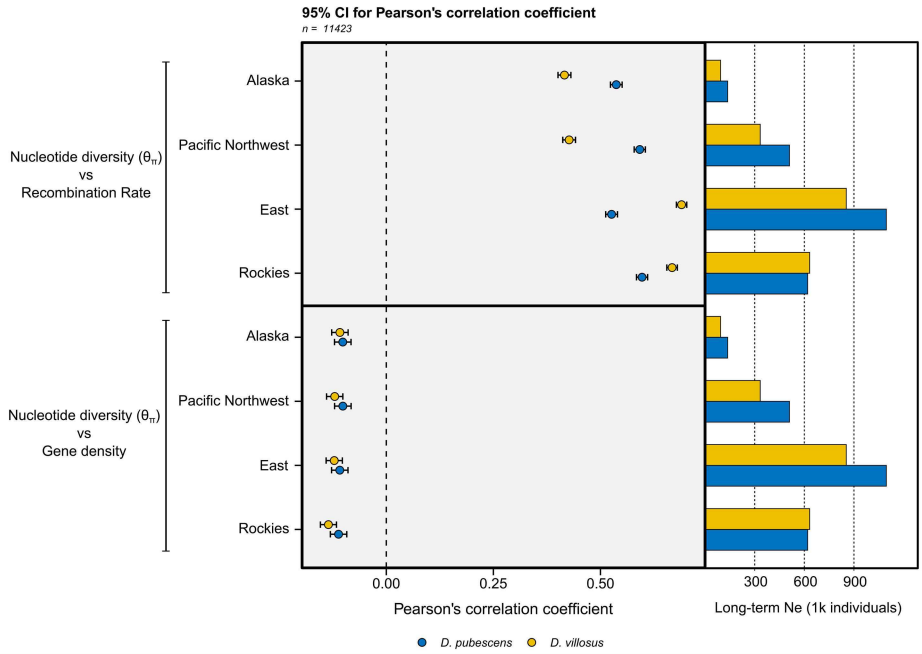

FIGURE 6 Correlation coefficients between genomic variables across populations of Downy (*D. pubescens*; pink) and Hairy (*D. villosus*; blue) Woodpecker and their respective long-term effective population sizes (N_e). Top plot shows the Pearson's correlation coefficient (mean and 95% confidence interval [CI]) for the association between nucleotide diversity and recombination rate. Bottom plot shows the Pearson's correlation coefficient (mean and 95% CI) for the association between nucleotide diversity and gene density. Right barplot shows the long-term effective population size for each population, as estimated from stairway plot 2. n = 11,423 genomic windows. [Colour figure can be viewed at wileyonlinelibrary.com]

and high impact shifted downwards compared to the mutations with low impact (Figure 8a,b). This indicates that purifying selection was successful in purging mutations that were highly deleterious. Hairy Woodpecker, however, showed a larger excess of low-frequency mutations of high impact when compared to Downy Woodpecker (Figure 7a,b), suggesting that purifying selection might have been more efficient in Hairy Woodpecker. In addition, we found that the recessive deleterious load was overall larger in Downy than Hairy Woodpecker, but this difference was not statistically significant (Kruskal-Wallis $\chi^2 = 1.33$, $df = 1$, $p = .24$; Figure 8c,d; Table S6). The recessive deleterious load was much larger in the Rocky Mountains when compared to other populations. Alaska also showed elevated recessive deleterious load in both species, generally larger than the East and Pacific Northwest (Figure 8c,d; Table S6).

Lastly, we investigated the overall impact of natural selection on protein-coding sequences of Downy and Hairy Woodpecker. dN/dS ratio was higher in Downy Woodpecker (dN/dS = 0.065) than in Hairy Woodpecker (dN/dS = 0.053), suggesting that purifying selection might have been weaker in the Downy Woodpecker lineage over deeper evolutionary times (i.e., >4 N_e generations ago; Elyashiv et al., 2010; Figuet et al., 2016; Herrera-Álvarez et al., 2020).

4 | DISCUSSION

Our genomic analyses reveal that both Ice Age demographic fluctuations and linked selection played a significant role shaping patterns of diversity and differentiation across populations and along the genomes of Downy and Hairy Woodpecker. We found that genome-wide nucleotide diversity, as well as the landscape of recombination, are highly correlated between these two species, which diverged more than 8 Ma. This degree of coupling in genome evolution suggests that intrinsic properties of the genome, such as recombination rate, might be conserved across deep evolutionary time. We posit that linked selection might underlie the genomic heterogeneity observed, as demonstrated by a significant association between nucleotide diversity, recombination rate, and gene density. Despite strong fluctuations in N_e over the Pleistocene, Downy and Hairy Woodpecker maintained large population sizes, which might have facilitated the action of natural selection. Nevertheless, given the large differences in long-term N_e observed among populations, our results indicate variation in the efficacy of selection.

4.1 | Conserved properties of the genome underlie the correlated genomic landscape of Hairy and Downy Woodpecker

We recovered large heterogeneity in patterns of nucleotide diversity (θ_π) and F_{ST} along the genomes of Downy and Hairy Woodpecker. Despite this variation, our results revealed a highly correlated genomic landscape between the two species. Such covariation in levels of genome-wide measures of diversity and differentiation across distantly related species is common (Burri et al., 2015; Delmore et al., 2018; Dutoit, Vijay, et al., 2017; Renaut et al., 2013; Stankowski et al., 2019; Van Doren et al., 2017) and

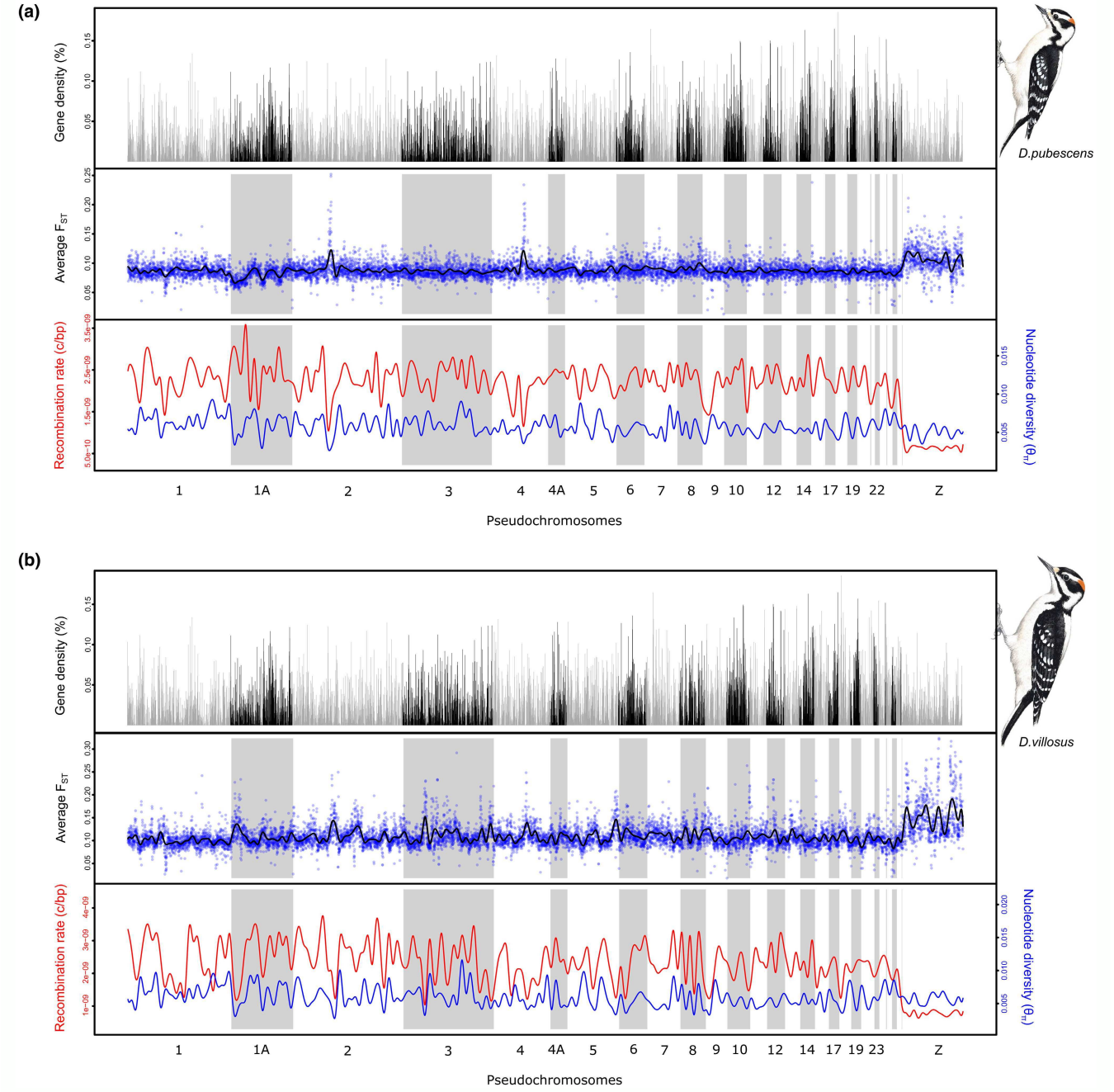

FIGURE 7 Landscape of diversity and differentiation along the genomes of Downy (*D. pubescens*; a) and Hairy (*D. villosus*; b) Woodpecker. Top plot shows the percentage of coding sequence in each nonoverlapping 100 kb window. Middle plot indicates the average pairwise F_{ST} calculated across nonoverlapping 100 kb windows. Bottom plot represents the weighted average recombination rate in c/bp (red) and the nucleotide diversity (θ_π; blue) for each nonoverlapping 100 kb window. Illustrations reproduced with permission from Lynx Edicions. [Colour figure can be viewed at wileyonlinelibrary.com]

suggests that properties of the genome, such as mutation rate, recombination rate, and density of targets of selection are conserved across deep evolutionary time (Dutoit, Burri, et al., 2017; Dutoit, Vijay, et al., 2017). For example, bird genomes are known to show large karyotypic stability, with very few chromosomal rearrangements and high synteny across highly divergent species (Damas et al., 2018; Ellegren, 2010, 2013; O'Connor et al., 2019; Singhal et al., 2015). Features of the genome, such as recombination rates and GC content, might also be conserved across species. We found that estimates of recombination rate are highly correlated between Downy and Hairy Woodpecker, although slightly higher in the latter. Linkage disequilibrium (LD), which is a function of both recombination rate and N_e, was extremely short in Downy and Hairy Woodpecker. Whereas linkage disequilibrium extends for over thousands of base pairs in humans (Ardlie et al., 2002; Reich et al., 2001), for instance, it breaks after only 100 bp in Downy and Hairy Woodpecker. Such properties have been observed in other bird species with very large N_e (Balakrishnan & Edwards, 2009;

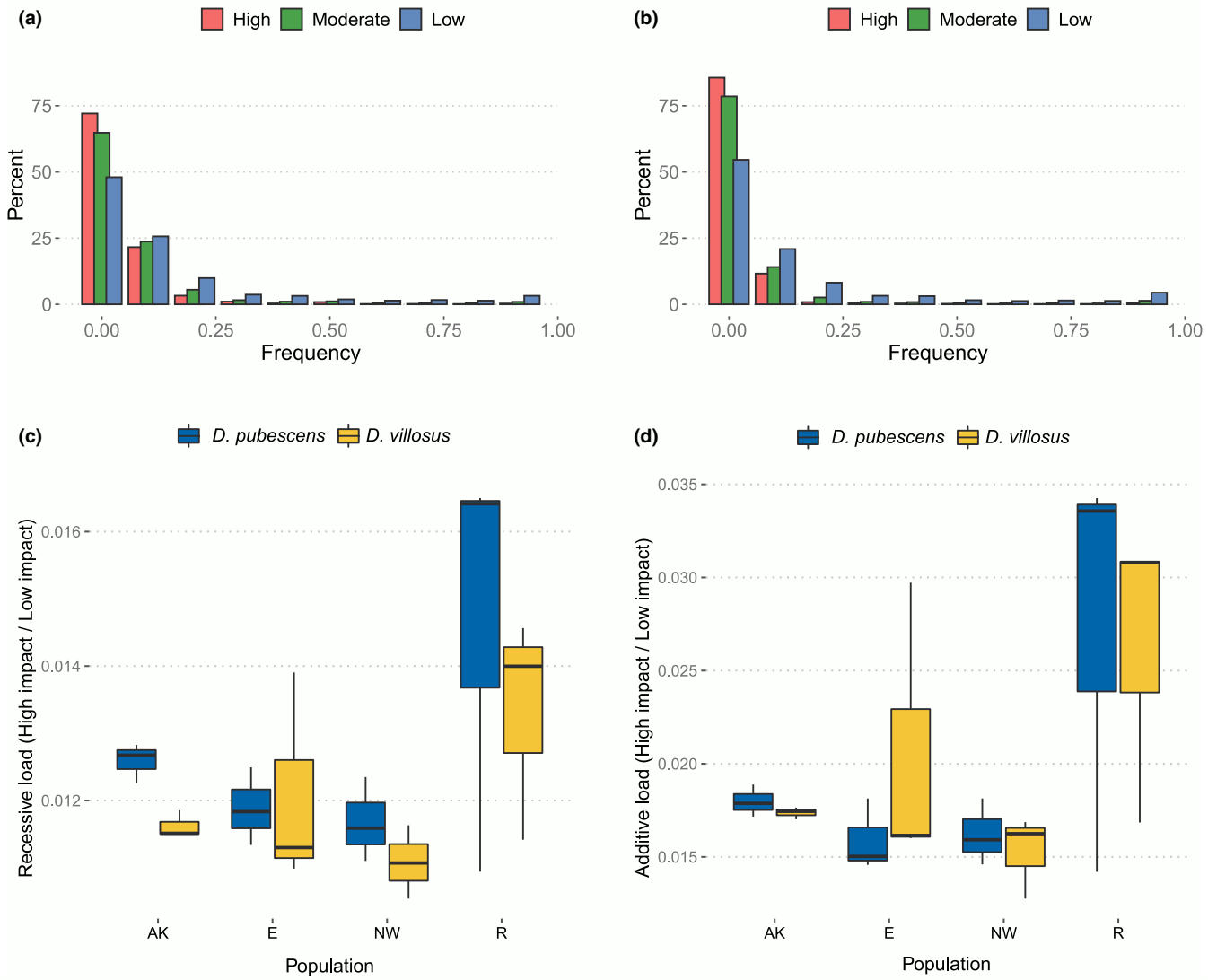

FIGURE 8 Deleterious load in Downy (*D. pubescens*) and Hairy Woodpecker (*D. villosus*). (a) Site frequency spectrum (SFS) for variants with low (neutral), moderate (mild), and high (deleterious) impact in Downy Woodpecker and (b) Hairy Woodpecker. (c) Ratio of homozygous derived variants of high impact (deleterious) over homozygous derived variants of low impact (neutral) in each genetic cluster and species (recessive model). (d) Ratio of the total number of derived variants of high impact (deleterious) over total number of derived variants of low impact (neutral) in each genetic cluster and species (additive model). Horizontal bars denote population medians. Populations: AK, Alaska; E, East; NW, Pacific Northwest; R, Rockies. [Colour figure can be viewed at wileyonlinelibrary.com]

Kardos et al., 2016). We also found large variation in recombination rates both within and among chromosomes, with the Z chromosome showing the lowest rates. Considering the lack of recombination across much of the Z chromosome in female birds (heterogametic sex; ZW), at the population level, crossing-over occurs at a much lower rate in sex chromosomes than in their autosome counterparts (Irwin, 2018; Sundström et al., 2004; Wilson Sayres, 2018). Similar to Downy and Hairy Woodpecker, recombination in the chicken (*Gallus gallus*) was approximately 2.5 times lower in the Z chromosome than in the autosomes (Levin et al., 1993; Schmid et al., 2000). As a consequence, many bird species show reduced diversity and faster divergence in the Z chromosome (Balakrishnan & Edwards, 2009; Borge et al., 2005; Mank et al., 2007; Sundström et al., 2004).

4.2 | The interplay between natural selection and recombination produces a heterogeneous genomic landscape

One of the main mechanisms proposed to explain the substantial heterogeneity in levels of polymorphism along the genome is the effect of linked selection (Charlesworth et al., 1993; Cutter & Payseur, 2013; Maynard & Haigh, 2007). Both directional selection (i.e., in favour of a beneficial allele) and purifying selection (i.e., against a deleterious allele) are expected to reduce diversity around functional elements (Charlesworth et al., 1993; Maynard & Haigh, 2007). Such a reduction is extended to all neighbouring sites that happen to be linked to the target of selection (background selection; Charlesworth et al., 1993; Comeron, 2014).

The extent to which adjacent sites are affected by linked selection is dependent on the recombination landscape, such that regions where recombination rate is lower tend to show lower genetic diversity and vice versa (Begun & Aquadro, 1992; Mugal et al., 2013; Wang et al., 2016). Similarly, the higher the density of functional elements (i.e., targets of selection), the more severe is the reduction in genetic diversity due to the effect of recurrent selection (Andolfatto, 2007; Beissinger et al., 2016; Branca et al., 2011; Gossmann et al., 2011). A correlation between nucleotide diversity and recombination may arise in the absence of linked selection. For example, biases in population-based recombination estimators, mutagenic effects of recombination itself, or the nonindependence of recombination with mutation rate and GC content could produce spurious associations. However, we do not expect these biases to impact the correlations with gene density. The lower diversity in functional regions are therefore interpreted as evidence of the effect of selection on linked neutral sites and can be used to assess the magnitude of linked selection (Corbett-Detig et al., 2015; Cutter & Payseur, 2013). In light of these results, we identified strong evidence that linked selection has contributed to patterns of genetic diversity along the genomes of Downy and Hairy Woodpecker. First, nucleotide diversity (θ_π) was positively associated with recombination rates in both species. Second, there was a weak but highly significant negative association between nucleotide diversity (θ_π) and gene density. Third, as predicted by theory, the strength of association between nucleotide diversity (θ_π), recombination rate, and gene density varied across populations of Downy and Hairy Woodpecker. Overall, more pronounced correlations were observed in populations with larger long-term N_e, supporting the impact of demographic history on linked selection.

Natural selection is also expected to impact levels of genetic differentiation along the genome (Cruickshank & Hahn, 2014; Matthey-Doret & Whitlock, 2019; Stankowski et al., 2019). We estimated a weak but significant negative association between nucleotide diversity (θ_π) and the average pairwise F_{ST}, indicating that regions of the genome that are highly differentiated between populations tend to show reduced diversity. These correlations are consistent with the effect of linked selection continuously eroding diversity near targets of selection (especially in regions of low recombination), which leads to the inflation of local levels of population differentiation (Cruickshank & Hahn, 2014). Whether the correlation between F_{ST}, nucleotide diversity, and recombination rate are due to recurrent divergent selection (e.g., local adaptation) or purifying selection (e.g., against deleterious mutations) would require the comparison between the landscape of F_{ST} and d_{xy} (Matthey-Doret & Whitlock, 2019; Vijay et al., 2017) and rigorous model-based analyses. Regardless, this and other comparative studies across both distantly and closely related bird species demonstrate that linked selection can reduce genetic diversity prior to population splits and consequently produce parallel patterns of genetic differentiation in regions of low recombination (Burri et al., 2015; Delmore et al., 2018; Irwin et al., 2016; Vijay et al., 2017).

4.3 | Dynamic population demography characterizes the evolution of Hairy and Downy Woodpecker in the Pleistocene

We found that population structure was spatially congruent between Downy and Hairy Woodpecker, but that their demographic histories and extent of genetic structuring varied. Both species were characterized by phylogeographic clusters that are consistent with previous studies (Graham & Burg, 2012; Klicka et al., 2011; Pulgarín-R & Burg, 2012) and were concordant with structuring in glacial refugia, albeit forming during different time scales (Appendix S1). The observation of common geographic patterns formed across different time periods highlights the predictability of the interaction of the physical landscape, drift, and gene flow on genetic diversity. Further evidence of the dramatic effects of Pleistocene climatic fluctuations on genetic diversity were the repeated cycles of population contraction and expansion. Yet, despite strong variation in N_e over the past 500 ky, our data indicate that Downy and Hairy Woodpecker have been resilient enough to maintain relatively large populations, which favoured the maintenance of very high genetic diversity, even in the face of repeated bottlenecks. Our estimates of current effective population size were comparable to estimates of census population size in the United States and Canada, which is approximately 13 million individuals of Downy Woodpecker and 8.5 million individuals of Hairy Woodpecker (North American Breeding Bird Survey; Sauer et al., 2017), although our estimates for Hairy Woodpecker are much larger (Table S4). While nonequilibrium population dynamics are a hallmark of species that occur in previously glaciated areas (Hewitt, 2004), the relationship between the magnitude of Pleistocene population size reductions and the efficacy of directional and purifying selection under these conditions remains a missing cornerstone for understanding how species respond to extreme climatic changes.

Consistent with theoretical predictions, nucleotide diversity within populations was strongly correlated with the long-term N_e. Alaska showed the lowest genome-wide genetic diversity, probably as a consequence of being one of the latest areas to be deglaciated and most recently founded (Figure 3e). On the other hand, the Northern Rockies sampling region exhibited the largest nucleotide diversity in both focal species (Figure 3e). Although such large estimates of genetic diversity could arise due to unaccounted population structure, data from multiple sources support the existence of a temporally fluctuating ice-free corridor along the Canadian Rocky Mountains that might have functioned as a glacial refugium (Jackson Jr, 1979; Pedersen et al., 2016; Rutter, 1984; Shafer et al., 2010). Thus, it is possible that suitable habitat might have allowed rapid growth and persistence of large populations in the Northern Rockies during the glacial periods of the Pleistocene.

4.4 | The efficacy of linked selection was affected by different evolutionary trajectories of Downy and Hairy Woodpecker

We investigated whether differences in the demographic trajectories of populations of Downy and Hairy Woodpecker in response to the Pleistocene glaciation had an impact on the efficacy of natural selection across the genome. Given that purifying selection is more efficient in larger populations (Ohta, 1973), we predicted that populations that underwent a stronger bottleneck or maintained lower levels of N_e were more likely to have accumulated highly deleterious mutations (i.e., genetic load; Henn et al., 2016; Wang et al., 2018; Willi et al., 2018; Rougemont et al., 2020; de Pedro et al., 2021). We failed to find support for this prediction. In contrast to our expectations, we found that the Rocky Mountains, the genetic cluster with the largest long-term N_e, exhibited the largest genetic load in both species. One possible explanation for this finding is that highly deleterious alleles might have been more efficiently purged from populations that experienced more severe population changes (Kirkpatrick & Jarne, 2000). For example, species whose populations underwent extreme bottlenecks show fewer mutations of high impact because extensive inbreeding makes highly deleterious alleles more likely to be exposed in homozygosity (Grossen et al., 2020; Robinson et al., 2018; Xue et al., 2015). This is not the case for Downy and Hairy Woodpecker, which, despite repeated episodes of bottleneck, still managed to maintain considerably large population sizes, making inbreeding very unlikely to have occurred. Besides, we found that Alaska, the population with the lowest long-term N_e, does not carry the fewest highly deleterious alleles, as predicted by the "purging under inbreeding" scenario. Instead, it carries a larger load than the East and the Pacific Northwest, which are populations with a higher long-term N_e. Both of these populations, however, were characterized by rapid changes in population size (Figures 1 and 4), which might explain their large load. It is possible that the large deleterious load in Alaska is due to substantial range shifts occurring during the glaciations, which lead to surfing of deleterious alleles when Alaska was recurrently recolonized (Hallatschek & Nelson, 2008).

At the species level, however, we found that genetic load was generally larger in Downy Woodpecker than Hairy Woodpecker, which is consistent with more efficient purifying selection in Hairy Woodpecker. This finding makes sense considering that Hairy Woodpecker exhibits slightly larger N_e than Downy Woodpecker. Supporting this observation, we also found a larger excess of highly deleterious mutations at low frequencies in Hairy Woodpecker, indicating that deleterious alleles were less likely to rise to high frequencies in Hairy Woodpecker than Downy Woodpecker likely due to more efficient selection. Lastly, we observed that the genome-wide ratio of nonsynonymous over synonymous substitutions (dN/dS) was higher in Downy Woodpecker than Hairy Woodpecker. Elevated genome-wide, as opposed to gene-specific, dN/dS ratio is suggestive of a reduction in the efficacy of purifying selection (Elyashiv et al., 2010; Figuet et al., 2016). This result indicates that a smaller N_e in the lineage leading to Downy Woodpecker might have allowed more fixation of slightly deleterious alleles.

In conclusion, we investigated the impact of demography and natural selection on the genomic landscape of two codistributed woodpecker species whose population histories have been profoundly impacted by the Ice Age. We found that despite a dynamic demographic history, Downy and Hairy Woodpecker were able to maintain very large N_e even during glacial periods, which might have facilitated the action of natural selection. Supporting this conclusion, our results reveal a correlation between nucleotide diversity, recombination rate, and gene density, which suggests the effect of linked selection shaping the genomic landscape. In addition, we found that the magnitude of linked selection was associated with population-specific N_e trajectories, indicating that demography and natural selection operated in concert to shape patterns of polymorphism along the genome. This study adds to the growing body of literature supporting the role of natural selection in driving patterns of genome-wide variation but highlights the difficulty of interpreting the outcome of the interplay between genetic drift and natural selection in organisms with nonequilibrium demographic dynamics and large effective population sizes.

AUTHOR CONTRIBUTIONS

This study was conceived and designed by Lucas R. Moreira and Brian Tilston Smith. A subset of samples was collected and made available by John Klicka. Lucas R. Moreira conducted all wet laboratory, bioinformatic and statistical analyses, and drafted the manuscript with input from all authors.

ACKNOWLEDGEMENTS

We thank the following individuals and institutions for providing tissue samples and specimen loans for this study: S. Birk/R. Faucett (University of Washington Burke Museum), C. M. Milensky (Smithsonian Institution), C. Dardia (Cornell University Museum of Vertebrates), G. Spellman/A. Doll (Denver Museum of Nature & Science), B. Marks/S. Hackett/J. Bates (Field Museum of Natural History), F. Sheldon/D. Dittmann (LSU Museum of Natural Science), K. Barker/T. Imfeld (Midwest Museum of Natural History), C. Witt/A. Johnson/M. Anderson (Museum of Southwestern Biology). This study would not be possible without the assistance of Lucas DeCicco, Matt Brady, and Paul Sweet, who were very generous to help collect samples in the field. All required US Federal and State permits to collect specimens specifically used in this study were obtained from the appropriate agency. We thank Kaiya Provost, Jon Merwin, Vivien Chua, Glenn Seehozer, Gregory Thom, Elkin Tenorio, William Mauck, Lukas Musher, Laís Coelho, Amanda Rocha, Bruno Almeida, Isaac Overcast, Joel Cracraft, Frank Burbrink, Molly Przeworski, Deren Eaton, Tom Trombone, and Don Melnick for their invaluable input during the development and writing of this manuscript. We are grateful for the suggestions of three anonymous reviewers that greatly improved the quality of this work. This study was funded by the Department of Ecology, Evolution, and Environmental Biology (E3B) at Columbia University, Conselho

Nacional de Desenvolvimento Científico e Tecnológico (CNPq; grant no. 211496/2014-6), the Frank M. Chapman Memorial Fund and Linda J. Gormezano Memorial Fund from the American Museum of Natural History (AMNH), the American Ornithological Society Hesse Award, the Society of Systematic Biologists Graduate Student Research Award, and a US National Science Award to BTS (DEB-1655736).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Whole-genome resequencing data are available at the Sequence Read Archive (https://ncbi.nlm.nih.gov/sra) under the BioProject PRJNA917637. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Pseudochromosome reference genome and annotation, vcf files, nucleotide diversity, GC content, gene density, and recombination rate estimates, outputs from IQ-Tree, NGSadmix, and stairway plot 2 are available in the Dryad Digital Repository (https://doi.org/10.5061/dryad.dfn2z355g). Code used in this study is available at https://github.com/lucasrocmoreira/Moreira-et-al_MolEco_2023.git. Additional data related to this paper may be requested from the authors.

ORCID

Lucas R. Moreira https://orcid.org/0000-0002-9084-3954

REFERENCES

Adrion, J. R., Galloway, J. G., & Kern, A. D. (2020). Predicting the landscape of recombination using deep learning. *Molecular Biology and Evolution*, *37*, 1–27.

Anderson, L. L., Hu, F. S., Nelson, D. M., Petit, R. J., & Paige, K. N. (2006). Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences*, *103*(33), 12447–12450.

Andolfatto, P. (2007). Hitchhiking effects of recurrent beneficial amino acid substitutions in the *Drosophila melanogaster* genome. *Genome Research*, *17*(12), 1755–1762.

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute.

Ardlie, K. G., Kruglyak, L., & Seielstad, M. (2002). Patterns of linkage disequilibrium in the human genome. *Nature Reviews. Genetics*, *3*(4), 299–309.

Avise, J. (1992). Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *The Auk*, *109*(3), 626–636.

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.

Begun, D. J., & Aquadro, C. F. (1992). Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature*, *356*(6369), 519–520.

Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using genomic data to infer historic population dynamics of nonmodel organisms. *Annual Review of Ecology, Evolution, and Systematics*, *49*, 433–456. https://doi.org/10.1146/annurev-ecolsys-110617-062431

Beissinger, T. M., Wang, L., Crosby, K., Durvasula, A., Hufford, M. B., & Ross-Ibarra, J. (2016). Recent demography drives changes in linked selection across the maize genome. *Nature Plants*, *2*, 16084.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.

Borge, T., Webster, M. T., Andersson, G., & Saetre, G.-P. (2005). Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two Ficedula flycatcher species. *Genetics*, *171*(4), 1861–1873.

Branca, A., Paape, T. D., Zhou, P., Briskine, R., Farmer, A. D., Mudge, J., Bharti, A. K., Woodward, J. E., May, G. D., Gentzbittel, L., Ben, C., Denny, R., Sadowsky, M. J., Ronfort, J., Bataillon, T., Young, N. D., & Tiffin, P. (2011). Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*. *Proceedings of the National Academy of Sciences*, *108*(42), E864–E870.

Burbrink, F., Chan, Y. L., Myers, E. A., Ruane, S., Smith, B. T., & Hickerson, M. J. (2016). Asynchronous demographic responses to Pleistocene climate change in Eastern Nearctic vertebrates. *Ecology Letters*, *19*(12), 1457–1467.

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A. A., Sætre, G.-P. P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Research*, *25*(11), 1656–1665.

Campbell-Staton, S. C., Goodman, R. M., Backström, N., Edwards, S. V., Losos, J. B., & Kolbe, J. J. (2012). Out of Florida: mtDNA reveals patterns of migration and Pleistocene range expansion of the green anole lizard (*Anolis carolinensis*). *Ecology and Evolution*, *2*(9), 2274–2284.

Charif, D., & Lobry, J. R. (2007). SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In U. Bastolla, M. Porto, H. E. Roman, & M. Vendruscolo (Eds.), *Structural approaches to sequence evolution: Molecules, networks, populations* (pp. 207–232). Springer.

Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, *134*(4), 1289–1303.

Charlesworth, D., & Willis, J. H. (2009). The genetics of inbreeding depression. *Nature Reviews. Genetics*, *10*(11), 783–796.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, *6*(2), 80–92.

Comeron, J. M. (2014). Background selection as baseline for nucleotide variation across the drosophila genome. *PLoS Genetics*, *10*(6), e1004434.

Corbett-Detig, R. B., Hartl, D. L., & Sackton, T. B. (2015). Natural selection constrains neutral diversity across a wide range of species. *PLoS Biology*, *13*(4), e1002112.

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*(13), 3133–3157.

Cutter, A. D., & Choi, J. Y. (2010). Natural selection shapes nucleotide polymorphism across the genome of the nematode *Caenorhabditis briggsae*. *Genome Research*, *20*(8), 1103–1111.

Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites: Unifying the disparity among species. *Nature Reviews. Genetics*, *14*(4), 262–274.

Damas, J., Kim, J., Farré, M., Griffin, D. K., & Larkin, D. M. (2018). Reconstruction of avian ancestral karyotypes reveals differences

in the evolutionary history of macro- and microchromosomes. *Genome Biology*, *19*(1), 155.

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.

Davis, M. B. (2001). Range shifts and adaptive responses to quaternary climate change. *Science*, *292*(5517), 673–679.

de Pedro, M., Riba, M., González-Martínez, S. C., Seoane, P., Bautista, R., Claros, M. G., & Mayol, M. (2021). Demography, genetic diversity and expansion load in the colonizing species *Leontodon longirostris* (Asteraceae) throughout its native range. *Molecular Ecology*, *30*, 1190–1205.

Delmore, K. E., Lugo Ramos, J. S., Van Doren, B. M., Lundberg, M., Bensch, S., Irwin, D. E., & Liedvogel, M. (2018). Comparative analysis examining patterns of genomic differentiation across multiple episodes of population divergence in birds. *Evolution Letters*, *2*(2), 76–87.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498.

Dufort, M. J. (2016). An augmented supermatrix phylogeny of the avian family Picidae reveals uncertainty deep in the family tree. *Molecular Phylogenetics and Evolution*, *94*, 313–326.

Dutoit, L., Burri, R., Nater, A., Mugal, C. F., & Ellegren, H. (2017). Genomic distribution and estimation of nucleotide diversity in natural populations: Perspectives from the collared flycatcher (*Ficedula albicollis*) genome. *Molecular Ecology Resources*, *17*(4), 586–597.

Dutoit, L., Vijay, N., Mugal, C. F., Bossu, C. M., Burri, R., Wolf, J., & Ellegren, H. (2017). Covariation in levels of nucleotide diversity in homologous regions of the avian genome long after completion of lineage sorting. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1849), 20162756. https://doi.org/10.1098/rspb.2016.2756

Ellegren, H. (2010). Evolutionary stasis: The stable chromosomes of birds. *Trends in Ecology & Evolution*, *25*(5), 283–291.

Ellegren, H. (2013). The evolutionary genomics of birds. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 239–259.

Elyashiv, E., Bullaughey, K., Sattath, S., Rinott, Y., Przeworski, M., & Sella, G. (2010). Shifts in the intensity of purifying selection: An analysis of genome-wide polymorphism data from two closely related yeast species. *Genome Research*, *20*(11), 1558–1573.

Excoffier, L., & Foll, M. (2011). Fastsimcoal: A continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, *27*(9), 1332–1334.

Figuet, E., Nabholz, B., Bonneau, M., Mas Carrio, E., Nadachowska-Brzyska, K., Ellegren, H., & Galtier, N. (2016). Life history traits, protein evolution, and the nearly neutral theory in amniotes. *Molecular Biology and Evolution*, *33*(6), 1517–1527.

Gossmann, T. I., Shanmugasundram, A., Börno, S., Duvaux, L., Lemaire, C., Kuhl, H., Klages, S., Roberts, L. D., Schade, S., Gostner, J. M., Hildebrand, F., Vowinckel, J., Bichet, C., Mülleder, M., Calvani, E., Zelezniak, A., Griffin, J. L., Bork, P., Allaine, D., … Ralser, M. (2019). Ice-age climate adaptations trap the alpine marmot in a state of low genetic diversity. *Current Biology: CB*, *29*(10), 1712–1720.e7.

Gossmann, T. I., Woolfit, M., & Eyre-Walker, A. (2011). Quantifying the variation in the effective population size within a genome. *Genetics*, *189*(4), 1389–1402.

Grabherr, M. G., Russell, P., Meyer, M., Mauceli, E., Alföldi, J., Di Palma, F., & Lindblad-Toh, K. (2010). Genome-wide synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics*, *26*(9), 1145–1151.

Graham, B. A., & Burg, T. M. (2012). Molecular markers provide insights into contemporary and historic gene flow for a non-migratory species. *Journal of Avian Biology*, *43*(3), 198–214.

Grossen, C., Guillaume, F., Keller, L. F., & Croll, D. (2020). Purging of highly deleterious mutations through severe bottlenecks in alpine ibex. *Nature Communications*, *11*(1), 1001.

Hallatschek, O., & Nelson, D. R. (2008). Gene surfing in expanding populations. *Theoretical Population Biology*, *73*(1), 158–170.

Henn, B. M., Botigué, L. R., Peischl, S., Dupanloup, I., Lipatov, M., Maples, B. K., Martin, A. R., Musharoff, S., Cann, H., Snyder, M. P., Excoffier, L., Kidd, J. M., & Bustamante, C. D. (2016). Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(4), E440–E449.

Herrera-Álvarez, S., Karlsson, E., Ryder, O. A., Lindblad-Toh, K., & Crawford, A. J. (2020). How to make a rodent Giant: Genomic basis and tradeoffs of gigantism in the capybara, the World's largest rodent. *Molecular Biology and Evolution*, *38*, 1715–1730. https://doi.org/10.1093/molbev/msaa285

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), 907–913.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *359*(1442), 183–195.

Hruska, J. P., & Manthey, J. D. (2021). De novo assembly of a chromosome-scale reference genome for the northern flicker Colaptes auratus. *G3*, *11*(1). https://doi.org/10.1093/g3journal/jkaa026

Irwin, D. E. (2018). Sex chromosomes and speciation in birds and other ZW systems. *Molecular Ecology*, *27*(19), 3831–3851.

Irwin, D. E., Alcaide, M., Delmore, K. E., Irwin, J. H., & Owens, G. L. (2016). Recurrent selection explains parallel evolution of genomic regions of high relative but low absolute differentiation in a ring species. *Molecular Ecology*, *25*(18), 4488–4507.

Jackson, L. E., Jr. (1979). New evidence for the existence of an icefree corridor in the Rocky Mountain foothills near Calgary, Alberta, during late Wisconsinan time. *Anatomy & Physiology: Current Research*, *79*, 107–111.

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B. B., Houde, P., Li, C., Al, E., Ho, S. Y. W., Faircloth, B. C., Nabholz, B., Howard, J. T., Suh, A., Weber, C. C., da Fonseca, R. R., Li, J., Zhang, F., Li, H., Zhou, L., Narula, N., … Al, E. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, *346*(6215), 1320–1331.

Jensen, J. D., Payseur, B. A., Stephan, W., Aquadro, C. F., Lynch, M., Charlesworth, D., & Charlesworth, B. (2019). The importance of the neutral theory in 1968 and 50 years on: A response to Kern and Hahn 2018. *Evolution; International Journal of Organic Evolution*, *73*(1), 111–114.

Jensen-Seaman, M. I. (2004). Comparative recombination rates in the rat, mouse, and human genomes. *Genome Research*, *14*(4), 528–538.

Kardos, M., Husby, A., McFarlane, S. E., Qvarnström, A., & Ellegren, H. (2016). Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in natural populations. *Molecular Ecology Resources*, *16*(3), 727–741.

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.

Kern, A. D., & Hahn, M. W. (2018). The neutral theory in light of natural selection. *Molecular Biology and Evolution*, *35*(6), 1366–1371.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.

Kimura, M., & Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics*, *49*(4), 725–738.

Kirkpatrick, M., & Jarne, P. (2000). The effects of a bottleneck on inbreeding depression and the genetic load. *The American Naturalist*, *155*(2), 154–167.

Klicka, J., Spellman, G. M., Winker, K., Chua, V., & Smith, B. T. (2011). A phylogeographic and population genetic analysis of a widespread, sedentary North American bird: The Hairy Woodpecker (*Picoides villosus*). *The Auk*, *128*(2), 346–362.

Knowles, L. L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshopprers. *Molecular Ecology*, *10*(3), 691–701.

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 356.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1–26.

Lessa, E. P., Cook, J. A., & Patton, J. L. (2003). Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proceedings of the National Academy of Sciences*, *100*(18), 10331–10334.

Levin, I., Crittenden, L. B., & Dodgson, J. B. (1993). Genetic map of the chicken Z chromosome using random amplified polymorphic DNA (RAPD) markers. *Genomics*, *16*(1), 224–230.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Li, J., Li, H., Jakobsson, M., Li, S., Sjödin, P., & Lascoux, M. (2012). Joint analysis of demography and selection in population genetics: Where do we stand and where could we go? *Molecular Ecology*, *21*(1), 28–44.

Liu, X., & Fu, Y.-X. (2020). Stairway plot 2: Demographic history inference with folded SNP frequency spectra. *Genome Biology*, *21*(1), 280.

Mank, J. E., Axelsson, E., & Ellegren, H. (2007). Fast-X on the Z: Rapid evolution of sex-linked genes in birds. *Genome Research*, *17*(5), 618–624.

Matthey-Doret, R., & Whitlock, M. C. (2019). Background selection and FST: Consequences for detecting local adaptation. *Molecular Ecology*, *28*(17), 3902–3914.

Mattila, T. M., Laenen, B., Horvath, R., Hämälä, T., Savolainen, O., & Slotte, T. (2019). Impact of demography on linked selection in two outcrossing Brassicaceae species. *Ecology and Evolution*, *9*(17), 9532–9545.

Maynard, J., & Haigh, J. (2007). The hitch-hiking effect of a favourable gene. *Genetics Research*, *89*(5–6), 391–403.

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.

Miller, J. B., Pickett, B. D., & Ridge, P. G. (2019). JustOrthologs: A fast, accurate and user-friendly ortholog identification algorithm. *Bioinformatics*, *35*(4), 546–552.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*(5), 1530–1534.

Mugal, C. F., Nabholz, B., & Ellegren, H. (2013). Genome-wide analysis in chicken reveals that local levels of genetic diversity are mainly governed by the rate of recombination. *BMC Genomics*, *14*(1), 86.

Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G., & Ellegren, H. (2015). Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. *Current Biology: CB*, *25*(10), 1375–1380.

O'Connor, R. E., Kiazim, L., Skinner, B., Fonseka, G., Joseph, S., Jennings, R., Larkin, D. M., & Griffin, D. K. (2019). Patterns of microchromosome organization remain highly conserved throughout avian evolution. *Chromosoma*, *128*(1), 21–29.

Ohta, T. (1973). Slightly deleterious mutant substitutions in evolution. *Nature*, *246*(5428), 96–98.

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.

Olivoto, T., & Lúcio, A. D. (2020). metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution/British Ecological Society*, *11*(6), 783–789.

Ouellet, H. R. (1977). *Biosystematics and ecology of Picoides villosus (L.) and P. pubescens (L.) (Aves Picidae)*. McGill University.

Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., Mendoza, M. L. Z., Beaudoin, A. B., Zutter, C., Larsen, N. K., Potter, B. A., Nielsen, R., Rainville, R. A., Orlando, L., Meltzer, D. J., Kjær, K. H., & Willerslev, E. (2016). Postglacial viability and colonization in North America's ice-free corridor. *Nature*, *537*, 1–15.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94–100.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Pulgarín-R, P. C., & Burg, T. M. (2012). Genetic signals of demographic expansion in Downy Woodpecker (*Picoides pubescens*) after the last North American Glacial Maximum. *PLoS One*, *7*(7), e40412.

Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. P. (2011). MACSE: Multiple alignment of coding SEquences accounting for frameshifts and stop codons. *PLoS ONE*, *6*(9), e22594. https://doi.org/10.1371/journal.pone.0022594

Reich, D. E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P. C., Richter, D. J., Lavery, T., Kouyoumjian, R., Farhadian, S. F., Ward, R., & Lander, E. S. (2001). Linkage disequilibrium in the human genome. *Nature*, *411*(6834), 199–204.

Reid, B. N., Kass, J. M., Wollney, S., Jensen, E. L., Russello, M. A., Viola, E. M., Pantophlet, J., Iverson, J. B., Peery, M. Z., Raxworthy, C. J., & Naro-Maciel, E. (2018). Disentangling the genetic effects of refugial isolation and range expansion in a trans-continentally distributed species. *Heredity*, *122*, 441–457. https://doi.org/10.1038/s41437-018-0135-5

Renaut, S., Grassa, C. J., Yeaman, S., Moyers, B. T., Lai, Z., Kane, N. C., Bowers, J. E., Burke, J. M., & Rieseberg, L. H. (2013). Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Communications*, *4*(1), 1827.

Robinson, J. A., Brown, C., Kim, B. Y., Lohmueller, K. E., & Wayne, R. K. (2018). Purging of strongly deleterious mutations explains long-term persistence and absence of inbreeding depression in Island foxes. *Current Biology: CB*, *28*(21), 3487–3494.e4.

Rougemont, Q., Moore, J.-S., Leroy, T., Normandeau, E., Rondeau, E. B., Withler, R. E., Van Doornik, D. M., Crane, P. A., Naish, K. A., Garza, J. C., Beacham, T. D., Koop, B. F., & Bernatchez, L. (2020). Demographic history shaped geographical patterns of deleterious mutation load in a broadly distributed Pacific Salmon. *PLoS Genetics*, *16*(8), e1008348.

Rutter, N. W. (1984). Pleistocene history of the western Canadian ice-free corridor. *Quaternary Stratigraphy of Canada: A Canadian Contribution to the IGCP Project*, *24*, 49–56.

Sauer, J. R., Niven, D. K., Hines, J. E., Ziolkowski, D. J., Jr., Pardieck, K. L., Fallon, J. E., & Link, W. A. (2017). *The north American breeding bird survey, results and analysis 1966–2015. Version 2.07*. US Geological Survey Patuxent Wildlife Research Center.

Schield, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton, R. W., Meik, J. M., Mackessy, S. P., & Castoe, T. A. (2020). Snake recombination landscapes are concentrated in functional regions despite PRDM9. *Molecular Biology and Evolution*, *37*(5), 1272–1294.

Schmid, M., Nanda, I., Guttenbach, M., Steinlein, C., Hoehn, M., Schartl, M., Haaf, T., Weigend, S., Fries, R., Buerstedde, J.-M., Wimmers, K., Burt, D. W., Smith, J., A'Hara, S., Law, A., Griffin, D. K., Bumstead, N., Kaufman, J., Thomson, P. A., … Mizuno, S. (2000). First report

on chicken genes and chromosomes 2000. *Cytogenetic and Genome Research*, *90*(3-4), 169–218.
Schrempf, D., Minh, B. Q., De Maio, N., von Haeseler, A., & Kosiol, C. (2016). Reversible polymorphism-aware phylogenetic models and their application to tree inference. *Journal of Theoretical Biology*, *407*, 362–370.
Shafer, A. B. A., Cullingham, C. I., Côté, S. D., & Coltman, D. W. (2010). Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, *19*(21), 4589–4621.
Simons, Y. B., & Sella, G. (2016). The impact of recent population history on the deleterious mutation load in humans and close evolutionary relatives. *Current Opinion in Genetics & Development*, *41*, 150–158.
Simons, Y. B., Turchin, M. C., Pritchard, J. K., & Sella, G. (2014). The deleterious mutation load is insensitive to recent population history. *Nature Genetics*, *46*(3), 220–224.
Singhal, S., Leffler, E. M., Sannareddy, K., Turner, I., Venn, O., Hooper, D. M., Strand, A. I., Li, Q., Raney, B., Balakrishnan, C. N., Griffith, S. C., McVean, G., & Przeworski, M. (2015). Stable recombination hotspots in birds. *Science*, *350*(6263), 928–932.
Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*(3), 693–702.
Smith, B. T., Gehara, M., & Harvey, M. G. (2021). The demography of extinction in eastern north American birds. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1944), 20201945.
Smukowski, C. S., & Noor, M. A. F. (2011). Recombination rate variation in closely related species. *Heredity*, *107*(6), 496–508.
Stankowski, S., Chase, M. A., Fuiten, A. M., Rodrigues, M. F., Ralph, P. L., & Streisfeld, M. A. (2019). Widespread selection and gene flow shape the genomic landscape during a radiation of monkeyflowers. *PLoS Biology*, *17*(7), e3000391.
Sundström, H., Webster, M. T., & Ellegren, H. (2004). Reduced variation on the chicken Z chromosome. *Genetics*, *167*(1), 377–385.
Talla, V., Soler, L., Kawakami, T., Dincă, V., Vila, R., Friberg, M., Wiklund, C., & Backström, N. (2019). Dissecting the effects of selection and mutation on genetic diversity in three wood white (Leptidea) butterfly species. *Genome Biology and Evolution*, *11*(10), 2875–2886.
Van Doren, B. M., Campagna, L., Helm, B., Illera, J. C., Lovette, I. J., & Liedvogel, M. (2017). Correlated patterns of genetic diversity and differentiation across an avian family. *Molecular Ecology*, *26*(15), 3982–3997.
Vijay, N., Weissensteiner, M., Burri, R., Kawakami, T., Ellegren, H., & Wolf, J. B. W. (2017). Genomewide patterns of variation in genetic diversity are shared among populations, species and higher-order taxa. *Molecular Ecology*, *26*(16), 4284–4295.
Waltari, E., Hijmans, R. J., Peterson, A. T., Nyári, Á. S., Perkins, S. L., & Guralnick, R. P. (2007). Locating Pleistocene refugia: Comparing phylogeographic and ecological niche model predictions. *PLoS One*, *2*(7), e563.
Wang, J., Street, N. R., Park, E., Liu, J., & Ingvarsson, P. K. (2020). Evidence for widespread selection in shaping the genomic landscape during speciation of Populus. *Molecular Ecology*, *29*(6), 1120–1136.
Wang, J., Street, N. R., Scofield, D. G., & Ingvarsson, P. K. (2016). Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related populus species. *Genetics*, *202*(3), 1185–1200.
Wang, X. J., Hu, Q. J., Guo, X. Y., Wang, K., Ru, D. F., German, D. A., Weretilnyk, E. A., Abbott, R. J., Lascoux, M., & Liu, J. Q. (2018). Demographic expansion and genetic load of the halophyte model plant Eutrema salsugineum. *Molecular Ecology*, *27*(14), 2943–2955.
Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E. V., & Zdobnov, E. M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, *35*(3), 543–548.
Wehrens, R., & Mevik, B.-H. (2007). *The pls package: principal component and partial least squares regression in R*. https://repository.ubn.ru.nl/bitstream/handle/2066/36604/36604.pdf
Weibel, A. C., & Moore, W. S. (2005). Plumage convergence in Picoides woodpeckers based on a molecular phylogeny, with emphasis on convergence in downy and hairy woodpeckers. *The Condor*, *107*(4), 797–809.
Willi, Y., Fracassetti, M., Zoller, S., & Van Buskirk, J. (2018). Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution*, *35*(4), 781–791.
Wilson Sayres, M. A. (2018). Genetic diversity on the sex chromosomes. *Genome Biology and Evolution*, *10*(4), 1064–1078.
Wu, T. D., & Watanabe, C. K. (2005). GMAP: A genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*, *21*(9), 1859–1875.
Xu, L., Wa Sin, S. Y., Grayson, P., Edwards, S. V., & Sackton, T. B. (2019). Evolutionary dynamics of sex chromosomes of Paleognathous birds. *Genome Biology and Evolution*, *11*(8), 2376–2390.
Xue, Y., Prado-Martinez, J., Sudmant, P. H., Narasimhan, V., Ayub, Q., Szpak, M., Frandsen, P., Chen, Y., Yngvadottir, B., Cooper, D. N., de Manuel, M., Hernandez-Rodriguez, J., Lobon, I., Siegismund, H. R., Pagani, L., Quail, M. A., Hvilsom, C., Mudakikwa, A., Eichler, E. E., … Scally, A. (2015). Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science*, *348*(6231), 242–245.
Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591.
Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M., & Yang, T.-L. (2018). *PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files* (pp. 1–3). Bioinformatics.
Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326–3328.
Zhou, Q., Zhang, J., Bachtrog, D., An, N., Huang, Q., Jarvis, E. D., Gilbert, M. T. P., & Zhang, G. (2014). Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science*, *346*(6215), 1246338.
Zink, R. M., Klicka, J., & Barber, B. R. (2004). The tempo of avian diversification during the quaternary. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *359*(1442), 215–220.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Moreira, L. R., Klicka, J., & Smith, B. T. (2023). Demography and linked selection interact to shape the genomic landscape of codistributed woodpeckers during the Ice Age. *Molecular Ecology*, *32*, 1739–1759. https://doi.org/10.1111/mec.16841