

# Evidence of positive selection and a novel phylogeny among five subspecies of song sparrow (*Melospiza melodia*) in Alaska

Caitlyn C. Oliver Brown<sup>1,2,3</sup>, Keiler A. Collier<sup>1,2,4</sup>, David J. X. Tan<sup>3</sup>, Kendall Mills<sup>1,2</sup>, Fern Spaulding<sup>1,2</sup>, Travis C. Glenn<sup>5</sup>, Christin Pruett<sup>6</sup> and Kevin Winker<sup>1,2</sup>

<sup>1</sup> University of Alaska Museum, Fairbanks, Alaska, United States

<sup>2</sup> Department of Biology and Wildlife, University of Alaska Fairbanks, Fairbanks, Alaska, United States

<sup>3</sup> Department of Biology, University of New Mexico, Albuquerque, New Mexico, United States

<sup>4</sup> RENECO International Wildlife Consultants, Abu Dhabi, United Arab Emirates

<sup>5</sup> Department of Environmental Health Science and Institute of Bioinformatics, University of Georgia, Athens, Georgia, United States

<sup>6</sup> Department of Biology, Ouachita Baptist University, Arkadelphia, Arkansas, United States

## ABSTRACT

Local adaptation occurs when populations evolve traits in response to local environmental challenges. Isolated island populations often experience different selection pressures than their mainland counterparts, which enables the study of how phenotypes and genotypes respond to differing selection regimes. We studied a group of five phenotypically differentiated subspecies of song sparrow (*Melospiza melodia*) in Alaska that demonstrate striking body size, color, and migratory behavioral differences to examine the effects of local adaptation on phenotypes and genotypes. We examined the phenotypic attributes of these populations and used whole-genome data to determine relationships and test candidate loci for evidence of selection. Phenotypic measurements of museum specimens ( $n = 227$ ) quantified the dramatic size differences among these populations, with westernmost *M. m. maxima* being ~1.6 times larger than easternmost *M. m. rufina*. Using ultraconserved elements (UCEs) and McDonald-Kreitman tests, we showed that seven candidate genes associated with bill size, circadian rhythm regulation, plumage color, and salt tolerance exhibited signs of putative positive selection. Phylogenetic analysis of UCEs identified *M. m. maxima* as sister to the other Alaska *M. melodia* subspecies. This suggests *M. m. maxima* colonized earliest, perhaps before the last glacial maximum, and that Alaska was later recolonized by ancestors of the remaining four subspecies.

Submitted 31 May 2024  
Accepted 4 August 2025  
Published 13 October 2025

Corresponding author  
Caitlyn C. Oliver Brown,  
caitlyn.brown@gmail.com

Academic editor  
Shaw Badenhorst

Additional Information and  
Declarations can be found on  
page 16

DOI 10.7717/peerj.19986

© Copyright  
2025 Oliver Brown et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Biogeography, Bioinformatics, Evolutionary Studies, Genomics, Zoology

**Keywords** *Melospiza melodia*, Alaska, Local adaptation, Island populations, Population genomics

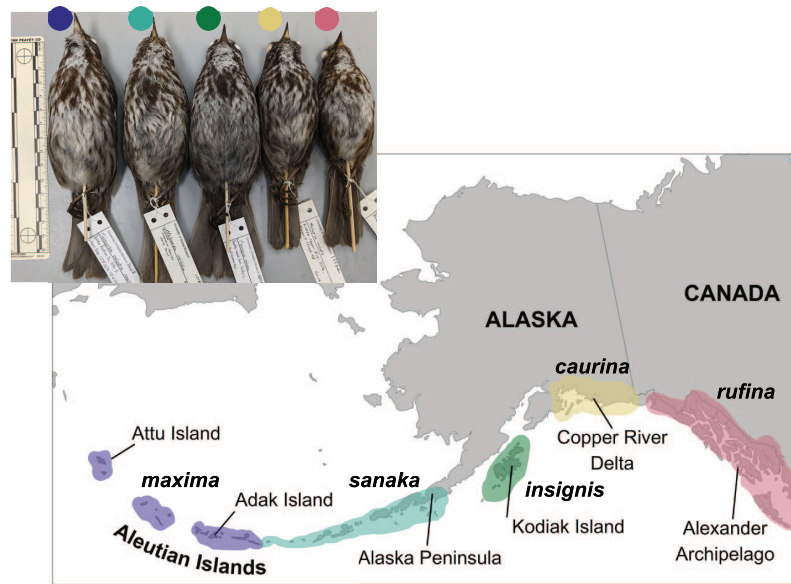
## INTRODUCTION

Our current understanding of evolution suggests that adaptation to local conditions enables populations to persist and that geographic variation among populations often reflects these processes. Local adaptation is defined as the process by which a local population evolves traits best suited for its environment (Turesson, 1922; Mayr, 1963;

Williams, 1966; Kawecki & Ebert, 2004; Savolainen, Lascoux & Merilä, 2013). Local adaptation to a specific environment can potentially lead to speciation (Savolainen, Lascoux & Merilä, 2013; Tiffin & Ross-Ibarra, 2014), but this process of adaptation can be hindered by the homogenizing effects of dispersal and gene flow (Kawecki & Ebert, 2004; Savolainen, Lascoux & Merilä, 2013). While it is still challenging to identify the genetic basis of local adaptation, the advent of high-throughput genomic data can be used to identify potential genes under selection (Tiffin & Ross-Ibarra, 2014; Bomblies & Peichel, 2022). In one method for examining local adaptation, candidate genes are chosen *a priori* due to their putative phenotypic effects of interest and adaptive potential can be tested (e.g., Walsh *et al.*, 2012). Examining local adaptation has been done in many terrestrial organisms, such as lizards (Losos, Warheit & Schoener, 1997), birds (Grant, 1981), spiders (Gillespie, 2002), and plants (Choi *et al.*, 2021).

Island systems have been influential in our understanding of the processes of evolution (MacArthur & Wilson, 1967; Losos & Ricklefs, 2009; Warren *et al.*, 2015). Islands, especially archipelagos, are natural laboratories for studying evolution, given their relatively young geological age and geographic isolation from the mainland (Losos & Ricklefs, 2009; Warren *et al.*, 2015). Because of these conditions, island populations provide a way to determine how local adaptation accrues. The majority of research on local adaptation has occurred on tropical or mid-latitude islands, however there has been a recent focus on high-latitude island systems (Whittaker & Fernandez-Palacios, 2007) such as the Aleutian Islands in the North Pacific Ocean, a volcanic archipelago that extends ~1,800 km from western Alaska towards eastern Russia (Murie, 1959). During the Pleistocene, Alaska experienced multiple cycles of glaciation and glaciers covered much of the North American continent. However, ice-free zones, or refugia, were present that could have allowed populations to persist, and the Aleutian Islands are hypothesized to contain some of these refugia (Pruett & Winker, 2005a; Winker *et al.*, 2023). Isolated avian populations on these islands often exhibit both lower genetic diversity than mainland counterparts and phenotypic differences, such as larger body size (Pruett, Li & Winker, 2018). One such species is the song sparrow (*Melospiza melodia*, Passeriformes: Passerellidae), a widely distributed songbird found only in North America with ~25 recognized subspecies (Aldrich, 1984; Patten & Pruett, 2009). These subspecies exhibit a wide range of phenotypes and life histories and reside in a variety of environments across North America (Aldrich, 1984; Patten & Pruett, 2009; Arcese *et al.*, 2020; Carbeck *et al.*, 2023). Because of this, song sparrows are an excellent candidate for the study of local adaptation.

Five song sparrow subspecies occur across southern Alaska, from the western Aleutian Islands to southeast Alaska. These subspecies, from west to east, are *M. m. maxima*, *M. m. sanaka*, *M. m. insignis*, *M. m. caurina*, and *M. m. rufina* (Fig. 1; Gibson & Withrow, 2015). Prior genetic research suggested that ancestral song sparrows colonized the Aleutian Islands sequentially, from east to west (Pruett & Winker, 2005b). Across this complex, these subspecies exhibit different life history traits, presumably due to differences in their environments. The subspecies that occur furthest west (*maxima* and *sanaka*) are non-migratory residents on the Aleutian Islands, a relatively harsh maritime environment. These two subspecies are found in grassy slopes along marine beaches, a habitat considered



**Figure 1** Distribution of five subspecies of song sparrow (*Melospiza melodia*) in Alaska. Insert at upper left is a photo of museum specimens of these subspecies. From west to east and left to right: *M. m. maxima* (dark blue), *M. m. sanaka* (teal), *M. m. insignis* (green), *M. m. caurina* (yellow), and *M. m. rufina* (pink). Photo credit: Caitlyn C. Oliver Brown. [Full-size !\[\]\(ba1b80118482ccef74a5d718ca4d7242\_img.jpg\) DOI: 10.7717/peerj.19986/fig-1](https://doi.org/10.7717/peerj.19986/fig-1)

marginal for continental subspecies, and are known to consume small marine invertebrates (Gabrielson & Lincoln, 1959; Murie, 1959; Aldrich, 1984). The subspecies that occur furthest east in Alaska (*caurina* and *rufina*) are seasonal migrants, residing largely on continental North America and migrate south, possibly to the Pacific Northwest, for the winter (Aldrich, 1984; Patten & Pruett, 2009). One subspecies (*insignis*) is primarily a year-round resident of the Kodiak Archipelago (Patten & Pruett, 2009). There are obvious phenotypic and natural history differences among these subspecies. The most notable phenotypic characteristics are differences in plumage coloration and body size: the most western subspecies in the Aleutian Islands are much larger than the eastern subspecies in southeast Alaska (Fig. 1; Aldrich, 1984; Pruett & Winker, 2005b; Patten & Pruett, 2009). These phenotypic differences suggest that these five subspecies have adapted to their local environments (Kawecki & Ebert, 2004; Patten & Pruett, 2009; Winker, 2010). We hypothesize that because *M. m. maxima* exhibits obvious phenotypic differences in plumage coloration, migration and dispersal, body size, and salt intake, candidate genes associated with these traits will be under selection.

Using museum specimens, we applied a combination of genomic and morphological analyses to study phenotypic differentiation among the subspecies of the Alaska song sparrow complex to examine the potential for adaptive evolution of high-latitude island populations. In so doing, we addressed the following questions: First, which morphological traits are significantly different among the five *M. melodia* subspecies in Alaska? Second, are candidate genes, associated with key phenotypic attributes, under positive selection in the most extreme variant, *M. m. maxima*? Third, what are the phylogenetic relationships among these subspecies?

## MATERIALS AND METHODS

### Phenotypic analyses

To examine phenotypic variation among the five *Melospiza melodia* subspecies, we obtained data from 227 voucher specimens from the University of Alaska Museum (Table S1). We restricted our analyses to adult males with complete phenotypic and locality data because our sample size of adult females was not large enough to accurately quantify sexual dimorphism. Phenotypic data included: mass, wing chord length, tail length, tarsus length, bill length, bill height, bill width, and skull length (Winker, 2000). Measurements were taken to the nearest 0.1 millimeter (mm) or to 0.1 gram (g) for mass prior to specimen preparation (Winker, 2000). Data were reviewed and analyzed in R (v. 4.1.1; R Core Team, 2021) and RStudio (v. 2022.12.0.353; Posit Team, 2025) using the devtools v2.4.5 and tidyverse v1.3.2 packages (Wickham et al., 2022, 2019). We removed outliers that were three standard deviations away from the mean to account for human measurement error. In total, 28 outlier measurement values, out of 227 individuals, were removed. Phenotypic data were then visualized with boxplots and a principal components analysis (PCA) using the packages “ggplot2” (v.3.4.1; Wickham et al., 2019), “ggpubr” (v. 0.6.0; Kassambara, 2023), and “ggord” (v. 1.1.7; Beck, 2022). For the PCA, we first z-transformed the data and then used the R function “prcomp”. When plotting the PCA, we used the first two principal components since they explained a majority of the variation. The 95% confidence intervals were calculated using the “ggord” package (Beck, 2022). ANOVA tests were used for each measurement and pairwise Tukey *post-hoc* tests were used to determine which subspecies pairs were different. R scripts and raw phenotypic data are available at <https://github.com/coliverbrown/melospiza-melodia-phenotype>. Portions of this text were previously published as part of a preprint (<https://www.biorxiv.org/content/10.1101/2024.05.21.595201v1>).

### Sampling and laboratory

For genomic analysis, we sampled high-quality voucher muscle tissue samples from wild individuals archived at the University of Alaska Museum (Table S2). Our sample size included one individual from each of the five *M. melodia* subspecies that occur in Alaska. Two taxa were included as outgroups: the swamp sparrow (*M. georgiana*) and the Lincoln’s sparrow (*M. lincolnii*). DNA extractions followed standard protocol for animal tissues using the Qiagen DNeasy Blood + Tissue Extraction Kit (Qiagen, Hilden, Germany).

Libraries were prepared using the iTru (Illumina dual-index) library protocols described in (Glenn et al., 2019). Briefly, we sheared the genomic DNA using a Bioruptor (Diagenode, Denville, NJ, USA) targeting a 500bp average fragment size. The sheared DNA was end-repaired, adenylated, and ligated to iTru adapters followed by limited-cycle polymerase chain reaction (PCR) of iTru primers to add indexes and complete the library molecules using Kapa Library Preparation Kit reagents (Kapa Biosystems, Roche, Basel, Switzerland). We sequenced the pooled libraries on an Illumina sequencer (Illumina, San Diego, CA, USA) to obtain paired-end (PE) 100 base reads. Raw sequence data have been deposited in NCBI Sequence Read Archive under BioProject PRJNA114297.

## Bioinformatics

After sequencing, we constructed whole genome assemblies for each of the sequenced individuals. Our bioinformatics pipeline centered on the package PHYLUCE (Faircloth, 2016). Raw and untrimmed FASTQ data that contained low-quality bases were removed using Illumiprocessor (Faircloth, 2013), which incorporates Trimmomatic (Bolger, Lohse & Usadel, 2014). Raw paired end reads (read1, read2, and singleton files) were then mapped to a dark-eyed junco (*Junco hyemalis*; Passeriformes: Passerellidae) reference sequence (MLZ69236; Friis et al., 2018). Resulting sequences were indexed using BWA (Li & Durbin, 2009) and SAMtools (Li et al., 2009). Next, we used PICARD (Broad Institute, 2019) to clean the alignments, add read group header information, and remove PCR and sequencing duplicates. Single nucleotide polymorphisms (SNPs) were called for each individual against the reference, and Genome Analysis Toolkit (GATK, v.4.2.1; McKenna et al., 2010) was used to call and realign around indels, call and annotate SNPs, and filter SNPs around indels. We restricted the data to high-quality SNPs by adding a quality filter (Q30) before converting the resulting VCF file to a FASTA file using GATK.

## Selection analyses

To determine whether genes associated with phenotype were under positive selection, we identified candidate genes and tested for signs of selection in *M. m. maxima*. Candidate genes were chosen through a review of literature that identified genes that showed correlations with phenotypes of interest. We chose 37 candidate genes associated with body size (Liu et al., 2015), bill size and shape (Lamichhaney et al., 2015; Chaves et al., 2016; Huang et al., 2022), migration (Vallone et al., 2007; Ruegg et al., 2014; Delmore et al., 2015), dispersal (San-Jose et al., 2023), plumage coloration, including melanin and carotenoid pathway genes (Walsh et al., 2012; San-Jose et al., 2017), and salt tolerance (Walsh et al., 2019) (Table S3). We downloaded predicted coding DNA sequence data for the candidate from the annotated *Melospiza melodia melodia* genome accessioned on GenBank (Rhie et al., 2021; annotation name: GCF\_035770615.1-RS\_2024\_02). We used these sequences (Table S3) to generate a local BLAST database and used BLASTn v2.12.0+ (Altschul et al., 1990; Camacho et al., 2009) to extract matching sequences from our *Melospiza melodia maxima* and *Melospiza georgiana* genome assemblies. Additionally, we extracted gene sequences from a *M. m. maxima* reference sequence (UAM31500; Louha et al., 2020). Using MAFFT v7 (Katoh & Standley, 2013), we manually generated multi-sequence alignments for each gene, removing indels to ensure that codon alignments were conserved.

We subsequently performed a McDonald-Kreitman test (McDonald & Kreitman, 1991) using a custom R script based on the algorithm described by (Li & Fu, 2003), which accounts for alternate mutation pathways for codons with more than one substitution difference ([https://github.com/g33k5p34k/Melospiza\\_mktest](https://github.com/g33k5p34k/Melospiza_mktest)). The McDonald-Kreitman test counts the number of synonymous ( $D_s$ ) and non-synonymous ( $D_n$ ) divergences between two species and synonymous ( $P_s$ ) and non-synonymous ( $P_n$ ) polymorphic sites within one of them (McDonald & Kreitman, 1991). In our case, we compared the polymorphic sites within *M. m. maxima* to divergences between *M. m.*



*maxima* and *M. georgiana*. These variables are used to calculate the neutrality index,  $NI = (P_n/P_s)/(D_n/D_s)$ .  $NI > 1$  indicates an excess of amino acid polymorphisms, and  $NI < 1$  is indicative of positive selection (Rand & Kann, 1996; Stoletzki & Eyre-Walker, 2011). Finally, a Fisher's exact test is calculated to determine whether the ratio of fixed substitutions between species and polymorphisms within species are significant.

## Phylogeny reconstruction

To reconstruct the phylogeny of the *M. melodia* complex and outgroups, we extracted ultraconserved elements (UCEs) from the whole-genome FASTA sequence files using PHYLUCE (Faircloth, 2016). We used UCEs because they are used to capture phylogenomic information from both shallow and deep time depths (Faircloth et al., 2012). We included the *Melospiza* individuals we sampled and the *J. hyemalis* reference sequence. Individual UCE loci were aligned using MUSCLE (Edgar, 2004) and were edge trimmed. We filtered this dataset to include loci that had all the *Melospiza* individuals and the *J. hyemalis* reference represented in a 95% matrix. The sequences were aligned by loci using MUSCLE and then concatenated into one dataset using PHYLUCE (Edgar, 2004; Faircloth, 2016). We obtained branch supports with the ultrafast bootstrap approximation (Hoang et al., 2018) implemented in IQ-TREE software (v.2.1.4; Minh et al., 2020). The phylogenetic tree was visualized and edited with FigTree (v.1.4.4; <http://tree.bio.ed.ac.uk/software/figtree/>).

## RESULTS

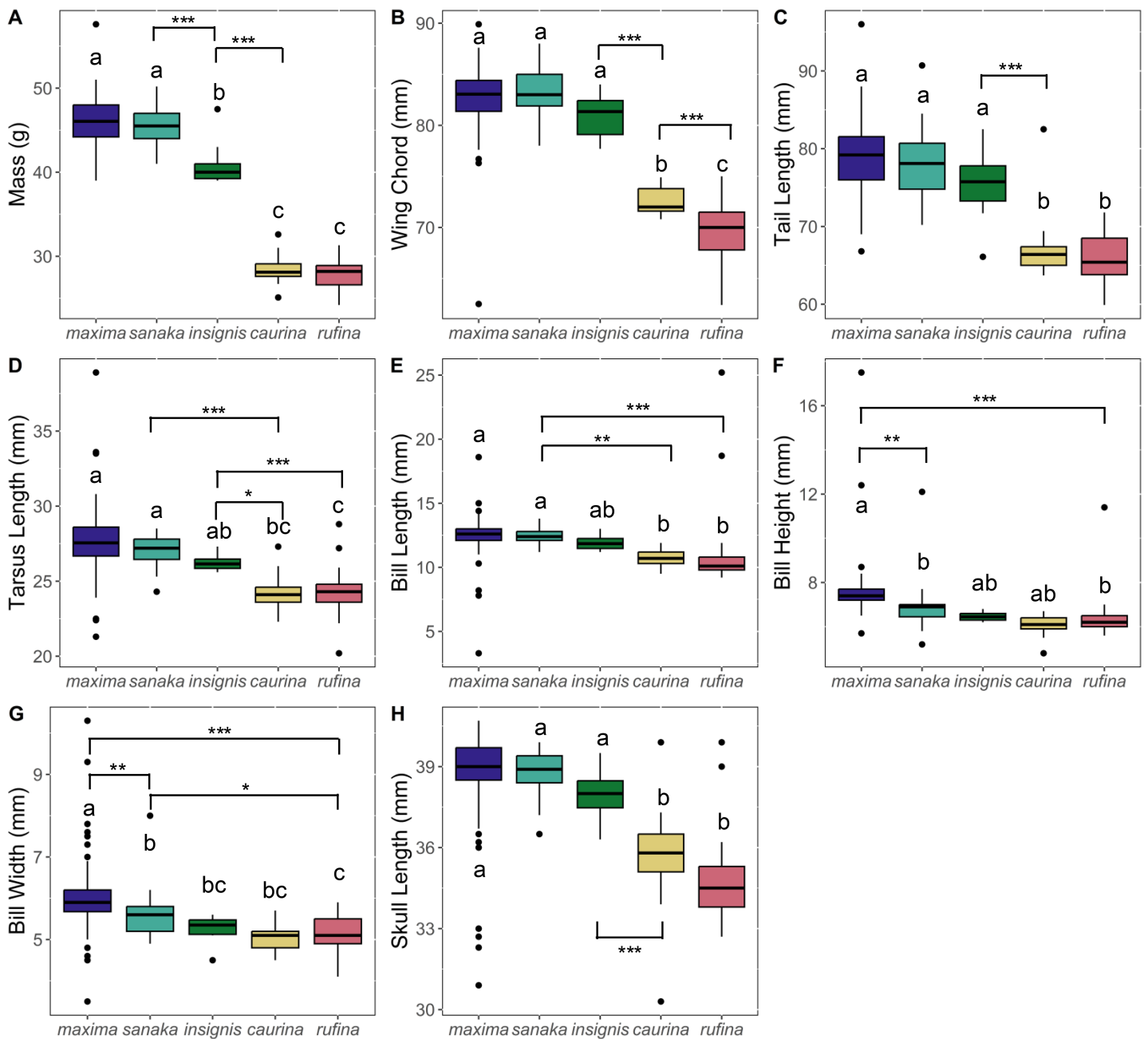
### Phenotypic variation

Phenotypic variation among the subspecies was pronounced. Overall, subspecies that occur in the western range (*maxima*, *sanaka*, and *insignis*) were larger than eastern Alaska subspecies (*rufina* and *caurina*). *Melospiza melodia maxima* was approximately 1.6 times larger in body mass than *rufina* ( $p < 0.0005$ ; Fig. 2; Table S4). For all other phenotypic measurements, differences were ~1.2 times larger (Fig. 2; Table S4). In all measurements, there was no significant difference between *rufina* and *caurina* (Fig. 2; Table S4). Additionally, *maxima* and *sanaka* showed significant differences only between bill height ( $p = 0.007$ ) and bill width ( $p = 0.006$ ) (Fig. 2; Table S4).

PCA analysis revealed that 55% of the variation among the five subspecies of *M. melodia* was explained in PC1 by five phenotypic traits: mass, wing chord, tail length, tarsus length, and skull length. Around 13% of the variation was explained in PC2 by the remaining three traits: bill length, bill width, and bill height (Fig. 3; Table S5). In the PCA graph, individuals tend to clump into two groups, mostly divided along PC1. The three western subspecies (*maxima*, *sanaka*, and *insignis*) group together, and the two remaining eastern subspecies (*caurina* and *rufina*) also group together (Fig. 3).

### Summary statistics of whole-genome sequencing

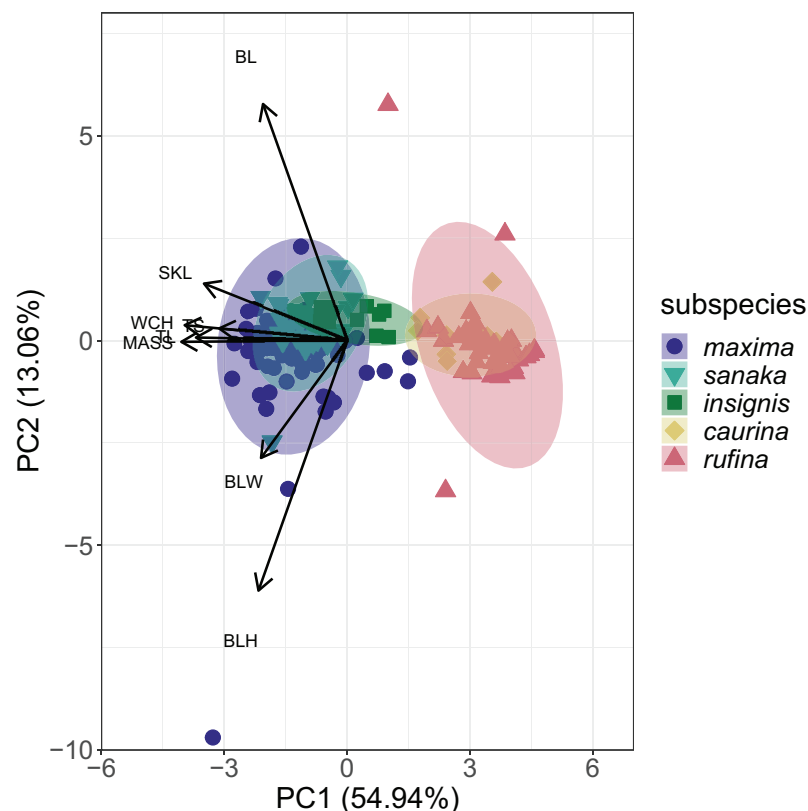
We obtained >400 million reads, ranging from 12,566,699 to 84,504,717 per individual (average = 60,119,612; Table S6). Coverage averaged 6.7×, ranging from 2.1× to 9.7× per individual (Table S6).



**Figure 2** Boxplots of phenotypic data for five *Melospiza melodia* subspecies: (A) Mass (g), (B) wing chord (mm), (C) tail length (mm), (D) tarsus length (mm), (E) bill length (mm), (F) bill height (mm), (G) bill width (mm), and (H) skull length (mm). Colors correspond to the colors in Fig. 1. Values with different letters are significantly different based on a Tukey's post-hoc test. ( $p < 0.05$ ). Asterisks denote the level of significance in pairwise differences between subspecies for each trait: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . [Full-size !\[\]\(1663bb69f307a960345edb0e712f8c02\_img.jpg\) DOI: 10.7717/peerj.19986/fig-2](https://doi.org/10.7717/peerj.19986/fig-2)

## Selection analyses

Of the 37 candidate genes tested for signals of positive selection, two had significant differences between ratios of  $P_n/P_s$  and  $D_n/D_s$ : *CD36* ( $p = 0.026$ ) and *LAP3* ( $p = 0.029$ ; Table 1). The neutrality index, which indicates direction and degree of deviation from



**Figure 3** PCA plot of phenotypic data for the five *Melospiza melodia* subspecies. PC1 explains approximately 55% of the variation with five phenotypic traits: mass (MASS), wing chord (WCH), tail length (TL), tarsus length (TS), and skull length (SKL). PC2 explains 13% of the variation with the three remaining traits: bill length (BL), bill width (BLW), and bill height (BLH). Three subspecies (*maxima*, *sanaka*, and *insignis*) group together and are larger from the second group of two subspecies (*caurina* and *rufina*). Colors correspond to Figure 1. [Full-size DOI: 10.7717/peerj.19986/fig-3](https://doi.org/10.7717/peerj.19986/fig-3)

neutrality, varied from 0 to 14.056, although 12 candidate genes had undefined neutrality indices, which occur when either  $D_n$  or  $P_s$  is zero, including *CD36* and *LAP3* (Table 1). Seven genes had a neutrality index  $< 1$ , which can be indicative of positive selection (Rand & Kann, 1996; Stoletzki & Eyre-Walker, 2011). Neutrality index values  $> 1$  (seven genes) might indicate weak purifying selection, or the maintenance of slightly deleterious mutations in the population (Rand & Kann, 1996; Stoletzki & Eyre-Walker, 2011). Finally, for six genes, there were no differences between the coding sequences (Table 1, “NA”).

## Phylogenetic reconstruction

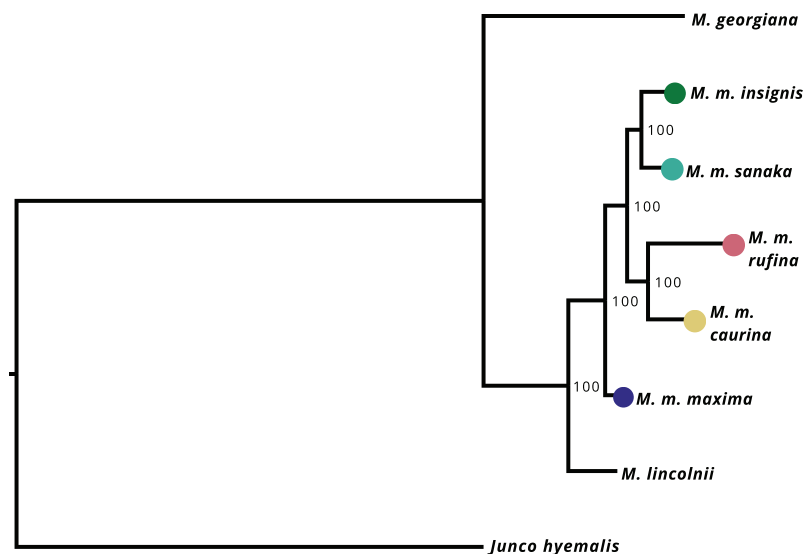
After filtering UCE loci present in the *Melospiza* and *J. hyemalis* genomes to obtain a 95% matrix, the final alignment contained 4,793 loci. All nodes within the *M. melodia* clade received 100% support from the ultrafast bootstrap (Fig. 4). The phylogenetic tree shows that two of the western subspecies (*sanaka* and *insignis*) and the two eastern subspecies (*caurina* and *rufina*) form their own clades. However, *maxima* is a sister group to all other *M. melodia* subspecies (Fig. 4).



**Table 1** McDonald-Kreitman test results.

Phenotypic trait	Candidate gene	$D_s$	$D_n$	$P_s$	$P_n$	Neutrality index (NI)	Fisher's exact test $p$ -value
Bill size	<i>ALX1</i>	4	0	1	0	Undefined	1
	<i>CCDC149</i>	4.5	3.5	1	4	5.143	0.565
	<i>DLK1</i>	3	1	9	2	0.667	1
	<i>HMGA2</i> <sup>†</sup>	NA	NA	NA	NA	NA	NA
	<i>LGI2</i> <sup>†</sup>	NA	NA	NA	NA	NA	NA
Body size	<i>AFAP1</i>	5	2	5	3	1.5	1
	<b><i>LAP3</i></b>	8	0	4	5	Undefined	<b>0.029</b>
	<i>LCORL</i> <sup>†</sup>	NA	NA	NA	NA	NA	NA
	<i>QDPR</i>	2.333	5.667	2	8	1.647	1
	<i>SLIT2</i>	10	1	12	2	1.667	1
Dispersal	<i>WAPL</i>	7.667	0.33	18	11	14.056	0.076
	<i>KCTD21</i> <sup>†</sup>	NA	NA	NA	NA	NA	NA
	<i>SLC2A1</i>	4	0	10	0	Undefined	1
Migration	<i>TGFB2</i>	2	0	3	1	Undefined	1
	<i>CLOCK</i>	2	3	3	1	0.222	0.524
	<i>CREB1</i> <sup>†</sup>	NA	NA	NA	NA	NA	NA
	<i>CRY1</i>	4	0	3	0	Undefined	1
	<i>CRY2</i>	1	0	3	1	Undefined	1
	<i>NPAS3</i>	1	0	3	0	Undefined	1
	<i>PER2</i>	15	4	9	0	0	0.273
	<i>PER3</i>	7	2	15	4	0.933	1
Plumage color – carotenoid	<i>YPEL1</i>	1	0	1	0	Undefined	1
	<i>APOD1</i>	2	0	5	0	Undefined	1
	<i>BCO1</i>	5	3	2	0	0	1
	<i>BCO2</i>	2	3	14	10	0.476	0.632
	<b><i>CD36</i></b>	3	0	3.5	14.5	Undefined	<b>0.026</b>
	<i>HPS5</i>	7	1	19.5	14.5	5.205	0.222
Plumage color – melanin	<i>STARD3</i>	4	1	4	0	0	1
	<i>AGRP</i>	0	3	2	0	0	0.1
	<i>ASIP</i> <sup>†</sup>	NA	NA	NA	NA	NA	NA
	<i>MC1R</i>	3	1	2	2	3	1
	<i>MITF</i>	1	0	10	1	Undefined	1
	<i>PCSK2</i>	2	0	3	0	Undefined	1
	<i>TYR</i>	3	2	6	3	0.750	1
	<i>MMP17</i>	1	1	9	2	0.222	0.423
Salt tolerance	<i>MYOF</i>	13	2	12	4	2.167	0.654
	<i>WNK2</i>	20.5	13.5	23	11	0.726	0.615

**Notes:**  
 The test calculates the total number of nonsilent ( $P_n$ ) and silent ( $P_s$ ) polymorphisms and nonsilent ( $D_n$ ) and silent ( $D_s$ ) divergences. The neutrality index (NI) is calculated as  $(P_n/P_s)/(D_n/D_s)$ . An undefined NI occurs when  $D_n$  or  $P_s$  is zero. Gene names and  $p$ -values in bold denote significant results ( $<0.05$ ) from Fisher's exact test.  
<sup>†</sup> No differences found between the three coding sequences tested.



**Figure 4** Maximum-likelihood phylogeny based on 4,793 UCE loci. Values on the nodes represent bootstrap values from 1,000 iterations. *M. m. maxima* is shown to be a sister group to the other *M. melodia* subspecies. [Full-size !\[\]\(1663bb69f307a960345edb0e712f8c02\_img.jpg\) DOI: 10.7717/peerj.19986/fig-4](https://doi.org/10.7717/peerj.19986/fig-4)

## DISCUSSION

Our results show that the westernmost subspecies of song sparrow (*M. m. maxima*) is approximately 1.6 times larger than eastern *rufina* (Fig. 2). The McDonald-Kreitman test showed that two candidate genes (*CD36* and *LAP3*) showed significantly different non-synonymous to synonymous substitution ratios between *M. m. maxima* and *M. georgiana*, indicating they are under selection, and the neutrality index suggested that seven more candidate genes were putatively under positive selection (Table 1). We reconstructed a novel phylogeny for the Alaska song sparrow complex, which places *M. m. maxima* outside the rest of the *M. melodia* clade (Fig. 4). This suggests that song sparrows colonized the Aleutian Islands early and underwent divergence before the eastern portion of the species' current Alaska range was recolonized.

## Phenotypic variation

The strong increase in body size among resident populations in the western Aleutians (Fig. 2) is concordant with previous research (Aldrich, 1984) and follows Bergmann's Rule, by which endothermic animals tend to be larger in cooler climates (Bergmann, 1847; Meiri & Dayan, 2003; Carbeck et al., 2023). When examining the PCA results (Fig. 3), the three western subspecies (*maxima*, *sanaka*, and *insignis*) grouped together and the remaining two eastern subspecies (*caurina* and *rufina*) grouped together. These groups are consistent with the different life-history strategies, with the larger, western subspecies being sedentary and the smaller, eastern subspecies being migratory. Sedentary birds tend to conform with Bergmann's rule more than migratory species because migratory species can seasonally escape more severe climatic conditions (Meiri & Dayan, 2003; but see Ashton, 2002). Likewise, island living could also have an impact on body size, because vertebrates

experience the “island rule”, wherein large organisms on islands evolve toward smaller sizes and small organisms evolve toward larger sizes (Lomolino, 2005; Losos & Ricklefs, 2009). Many factors are hypothesized to contribute to this general trend, such as ecological release from predators and resource limitation (Lomolino, 2005). Other passerines that inhabit the Aleutian Islands, such as the gray-crowned rosy-finch (*Leucosticte tephrocotis*) and Pacific wren (*Troglodytes pacificus*), are also much larger in the western islands (Murie, 1959).

Of the western subspecies, *maxima* and *sanaka* were similar across all measurements except for bill height and bill width (Fig. 2; Table S4). *M. m. sanaka* had significantly narrower and shallower bills than *maxima*, which is consistent with previous research (Patten & Pruett, 2009; Gibson & Withrow, 2015). However, although slender bills are often associated with insectivory and stouter bills with granivory (Aldrich, 1984), little is known about the dietary differences of the two subspecies, so the ecological or evolutionary basis of this morphological difference remains unknown (Aldrich, 1984). Additionally, foraging behavior is not the sole cause of differences in bill size. Bird bills are thermoregulatory organs, facilitating heat transfer between the organism and the environment (Tattersall, Arnaout & Symonds, 2017; Gamboa et al., 2024). In the California Channel Islands, for instance, Gamboa et al. (2024) identified temperature as the main factor affecting differences in bill size in song sparrows between islands. Potentially, differences in temperature between the western Aleutian Islands and the Alaska Peninsula may cause the observed differences in bill size. Furthermore, local adaptation is not the sole cause of population differentiation, especially on islands. Founder effects have been shown to drive phenotypic and genetic differences between populations (Kolbe et al., 2012), and it is therefore possible that the differences we found between subspecies were neutral rather than adaptive, although we are unable to distinguish between the two in this study.

### McDonald-Kreitman tests identify potential genes under selection

Our McDonald-Kreitman tests identified two genes, *CD36* and *LAP3*, that showed significantly different non-synonymous to synonymous substitution ratios between *M. m. maxima* and *M. georgiana*, indicating that they are under selection (Table 1). This pattern was driven by the absence of fixed non-synonymous substitutions between species ( $D_n = 0$ ), which suggests that these genes are under strong purifying selection ( $NI = \text{undefined}$ ). While none of the other candidate genes showed significant signatures of selection at a threshold of  $\alpha = 0.05$ , it is worth noting that our small sample size ( $N = 2$  in-group and 1 out-group samples) likely led to severe underestimates of the number of within-species synonymous substitutions ( $P_s$ ) relative to the other substitution types, thus reducing the power of the Fisher’s Exact Test.

Despite our sample size limitations, it should nonetheless still be possible to identify putatively positively selected genes based on their NI values. Because the neutrality index can alternatively be expressed by the equation  $NI = (P_n * D_s) / (P_s * D_n)$ , our systematic underestimation of the number of  $P_s$  sites should lead to higher than expected neutrality indices per gene, thereby underestimating the number of genes under positive selection.

Under our current sampling approach, loci with an  $NI < 1$  can therefore be considered strong candidates for potential positive selection.

Because *M. m. maxima* exhibit obvious phenotypic differences from *M. georgiana*, we expected to observe directional selection in candidate genes associated with plumage coloration, migration and dispersal, body size, and salt intake. Although preliminary, our results generally align with these hypotheses, except for dispersal and body size-associated genes. For bill size, our analysis infers potential positive selection in *DLK1*, which has been shown to be associated with differences in bill morphology in Darwin's finches ([Chaves et al., 2016](#)) and Hawaiian honeycreepers ([Campana et al., 2020](#)). However, while *DLK1*-associated polymorphisms have generally been found in the transcription regulation region upstream of the gene in Darwin's finches, or at amino acids 100 and 244 in the Hawaiian honeycreepers, our results instead identify a novel substitution at amino acid 260, with a methionine residue in *M. georgiana* replaced by a threonine residue in *M. m. maxima*.

As for migration-associated genes, we found signals of positive selection in the *CLOCK* and *PER3* genes ([Table 1](#)), both of which are known to regulate avian circadian rhythms ([Kwak et al., 2024](#)). The *CLOCK* gene is particularly interesting, because other studies in birds have found that the number of glutamine residue repeats in the C-terminal domain appear to be linked to migration timing ([Bazzi et al., 2016](#); [Saino et al., 2015](#)). However, we observed no glutamine repeat variation in our *M. m. maxima* or *M. georgiana* sequences and instead found two fixed amino acid substitutions at residues 6 and 221. As for the *PER3* gene, we found a single fixed substitution at amino acid residue 1,045, which occurs within the cryptochrome-binding region ([Kwak et al., 2024](#)). Our discovery of potential positive selection in these two migration-linked genes seems counterintuitive because *M. m. maxima* is sedentary whereas *M. georgiana* is migratory. However, it is possible that these polymorphisms could be linked to the suppression of migratory behavior in *maxima* individuals. Alternatively, these amino acid substitutions could be due to adaptations to the more variable photoperiods experienced by the *maxima* population at higher latitudes.

As for plumage coloration, we identified signals of potential positive selection in *BCO2*, which is associated with carotenoid coloration pathways ([Table 1](#)). *BCO2* has been a focus of carotenoid coloration studies in birds, most notably with regard to the increase in carotenoid coloration in beaks, legs, and skin ([Walsh et al., 2012](#); [Gazda et al., 2020](#); [Enbody et al., 2021](#)). [Gazda et al. \(2020\)](#) found a single nonsynonymous mutation in *BCO2* exon 9, which disabled the enzyme's carotenoid cleaving ability, resulting in an increase in carotenoid accumulation in tissues. Additionally, in a study with Darwin's finches, [Enbody et al. \(2021\)](#) found a single nucleotide polymorphism associated with yellow vs pink beak coloration of nestlings. This polymorphism, found in *BCO2* exon 4, is synonymous but nonetheless alters the expression of *BCO2*, causing carotenoid deposition in bills ([Enbody et al., 2021](#)). However, neither of these mutations were observed in our data; instead, we found amino acid substitutions at residues 31, 87, and 311. While *M. m. maxima* does not possess obvious red coloration compared to other subspecies or *M. georgiana*, the number of fixed substitutions observed in our dataset nonetheless suggests that there may be differences in carotenoid deposition between the two taxa, which may explain the darker

phenotype of *maxima*. Another possibility is that changes in carotenoid deposition might be an adaptation by *maxima* to living in harsher environments. Studies indicate that increased carotenoid concentration stimulates the immune response of birds ([Pap et al., 2009](#); [Toews, Hofmeister & Taylor, 2017](#)). In an experimental study with house sparrows, [Pap et al. \(2009\)](#) found that birds supplemented with carotenoids were able to fully compensate during parasitic infection. In humans, carotenoids also have antioxidant properties and prevent light damage to the retina ([von Lintig, 2010](#); [Toews, Hofmeister & Taylor, 2017](#)). This combination of immune response and antioxidants might allow resident subspecies *M. m. maxima* to survive winters in the Aleutian Islands.

Additionally, one gene associated with the melanogenesis pathway, *TYR*, was identified to be under potential positive selection ([Table 1](#)). Positive selection in this gene is interesting because mutations at amino acid residues 246 and 371 have been found to cause melanin deficiencies in parrot plumage ([Ghosh Roy et al., 2025](#)). While our data do not show the same substitutions as [Ghosh Roy et al. \(2025\)](#), one of our observed amino acid substitutions at position 369 is highly proximate to the H371Y substitution in the  $\alpha 8$  helix of the psittaciform *TYR* protein, which lies close to the protein's active site. We also observed another amino acid substitution at residue 508, which lies in the transmembrane region of the protein. While we do not yet know how these mutations affect *TYR* enzymatic activity, we expect the positive selection of *TYR* to cause a higher melanin accumulation in *M. m. maxima* because of its overall darker coloration compared to *M. georgiana*. *M. melodia* generally follows Gloger's rule, or the tendency for birds in more humid environments to be darker in coloration ([Patten & Pruett, 2010](#)). More research into *TYR* and melanogenesis mechanisms is needed to determine why it is potentially under positive selection in *M. m. maxima*.

Finally, we found potential positive selection in two salt-tolerance candidate genes: *MMP17* and *WNK2* ([Table 1](#)). Neither of these genes have been studied extensively in birds. *MMP17* has been linked to drinking behavior and kidney function in mice ([Srichai et al., 2011](#)). In Nelson's sparrows, a region of elevated divergence in *MMP17* was noted between upland and salt marsh populations, suggesting potential adaptation to salt marsh habitats ([Walsh et al., 2019](#)). *WNK2* has been identified as an important component in osmoregulation, regulating cell volume in response to osmotic stress ([Kahle, Rinehart & Lifton, 2010](#)). Within birds, this region had an elevated divergence between upland and salt marsh savannah sparrows ([Walsh et al., 2019](#)). We assume that the inferred positive selection of both *MMP17* and *WNK2* are adaptations to the diet and life history of *M. m. maxima*. Individuals of this subspecies routinely forage along the marine beaches and rocky tidepools of the Aleutian Islands, consuming marine invertebrates ([Gabrielson & Lincoln, 1959](#); [Murie, 1959](#); [Aldrich, 1984](#)). Compared to *M. georgiana*, which consumes terrestrial invertebrates, we expect to see adaptations to a salty environment, similar to the phenomenon described by [Walsh et al. \(2019\)](#).

Rather unexpectedly, we found no signal of positive selection in any of the seven candidate genes linked to differences in body size included in our study, despite *M. m. maxima* being on average 1.6 times larger than *M. georgiana*. This might be because other body size-linked genes could be responsible for mediating this phenotypic divergence.



Body size has been the focus of much research in the *Melospiza melodia* complex recently. [Carbeck et al. \(2023\)](#), for instance, found nine genes associated with body size that showed divergence between *maxima* and *rufina* song sparrows and then validated these candidate gene findings by predicting the genotypes of five small song sparrow subspecies (*gouldii*, *heermanni*, *maxillaris*, *pusillula*, and *samuelis*) in California. Our study did not include any candidate genes found by [Carbeck et al. \(2023\)](#), which should be further considered in future studies of song sparrows.

Additionally, we detected signals of strong negative or purifying selection ( $NI = \text{undefined}$ ) in twelve candidate genes, which was driven by the absence of fixed non-synonymous substitutions between species ([Table 1](#)). This presence of putative purifying selection in twelve candidate genes (*APOD1*, *AXL1*, *CD36*, *CRY1*, *CRY2*, *LAP3*, *MITF*, *NPAS3*, *PCSK2*, *SLC2A1*, *TGFB2*, and *YPEL1*) indicates these genes are conserved and potentially critical to the function of these organisms. It is also important to note that many genes are multi-functional and could contribute to functions outside of those we addressed here.

While these results generally align with our hypotheses, the McDonald-Kreitman (MK) test does have limitations. Modifications to the MK test exist that allow for selection to be detected in non-coding regions, but our application of the standard MK test means that we were unable to test for selection in regulatory regions surrounding protein-coding genes, which can be influenced by linked selection. Additionally, we were unable to account for the effect of slightly deleterious mutations, which can also make it difficult to identify adaptive evolution ([Andolfatto, 2008](#)). More importantly, small sample sizes and small effective population sizes could lead to elevated false positive rates ([Andolfatto, 2008](#); [Subramanian, 2016](#)). However, although our analyses suffer from both issues ([Parsch, Zhang & Baines, 2009](#)), the fact that none of our candidate genes exhibited statistically significant levels of positive selection indicates that type I error was unlikely to have been a major issue. Limitations notwithstanding, our results provide a preliminary look into how *M. m. maxima* might have locally adapted to the Aleutian Islands. Future work should increase sample sizes to include more than two sequences tested, use coalescent simulations to determine significance levels ([Andolfatto, 2008](#); [Subramanian, 2016](#)), and test more candidate genes (e.g., [Carbeck et al., 2023](#)).

## Phylogenetic reconstruction

Our phylogenetic tree shows that the westernmost subspecies *maxima* is a sister taxon to all other Alaska *M. melodia* subspecies ([Fig. 4](#)). Further, the eastern Alaska subspecies *caurina* and *rufina* are grouped together and two of the western subspecies (*sanaka* and *insignis*) form another clade ([Fig. 4](#)). This suggests that *maxima* colonized the western Aleutian Islands first, underwent differentiation from other forms, and then at some later time, the eastern part of the Alaskan range was colonized by ancestors of the remaining subspecies. One possibility for this colonization history is that following an initial colonization of Alaska, the sedentary *maxima* survived in an Aleutian glacial refugium ([Pruett & Winker, 2005a](#); [Winker et al., 2023](#)) while other Alaska populations were

extirpated, and then Alaska was re-colonized post-glacially. Seasonal migration likely aided this recolonization as glacial barriers retreated ([Pielou, 1991](#)), but some recolonizing lineages subsequently also became sedentary (e.g., *sanaka* in the Aleutians). This hypothesis is inconsistent with the colonization history postulated by [Pruett & Winker \(2005b\)](#), which suggested a probable postglacial stepwise process from east to west. Further sampling from across the species' range will be needed to confirm this hypothesis.

This is the first song sparrow phylogeny reconstructed from large sampling of the nuclear genome. Prior work using mitochondrial DNA (mtDNA) contradicts our results. [Zink & Dittmann \(1993\)](#) found no genetic structure in a mtDNA phylogeny across the species' range, and [Pruett & Winker \(2010\)](#) found no reciprocal monophyly or genetic structure within Alaska's subspecies. However, single-locus phylogenies are notoriously problematic because of issues like incomplete lineage sorting and gene flow (e.g., [Avise & Wollenberg, 1997](#); [Funk & Omland, 2003](#); [Joseph, 2021](#)), and further work is needed to understand phylogenetic relationships among all the lineages in this species.

Interestingly, *maxima* and *sanaka* are similar in size, have similar life-history patterns, and occupy similar habitats. Yet, our phylogeny suggests *M. m. sanaka* is more closely related to all of the non-*maxima* subspecies. This phenotypic similarity could be due to convergent evolution in similar environments, or, more likely, due to adaptive alleles for Aleutian life introgressing through gene flow from *maxima* to *sanaka* ([Pruett & Winker, 2005b](#); [Reding et al., 2008](#); [Graham et al., 2021](#)).

It is possible that gene flow could be obscuring our phylogenetic results, however we do not expect this to be the case. Models show that moderate to high levels of gene flow can degrade phylogenetic signal and thus produce inaccurate phylogenetic trees ([Eckert & Carstens, 2008](#); [Leaché et al., 2014](#)). Previous research by [Pruett & Winker \(2005b\)](#) found gene flow between *maxima* and *sanaka* to be at approximately 1.20–2.90 individuals per generation ([Fig. S1](#)). The exact amount of gene flow needed to obscure phylogeny is not known, but the current estimates for gene flow in the Aleutian Islands are not very high. Future research should determine whether phylogenetic relationships are being obscured by gene flow in this song sparrow complex.

## CONCLUSION

Among song sparrows in Alaska, phenotypic differences between subspecies suggest that selective pressures have acted heterogeneously on different populations. Our MK tests suggest that seven candidate genes associated with bill size, migration, plumage coloration, and salt tolerance might be under positive selection. This suggests that these traits may be under selective pressure in *maxima*, and lays the groundwork for future investigations with greater sampling depth and breadth. Additionally, we found signals of negative selection in twelve candidate genes, indicating these genes might be functionally critical and conserved rather than locally adaptive. The phylogenetic relationship that we recovered among Alaska song sparrow subspecies suggests a new hypothesis for the species' biogeographic history in this region. We suggest that ancestral populations of *M. m. maxima* colonized Alaska and the Aleutian Islands, and then persisted in a glacial refugium, likely while other

populations were extirpated. Then after a substantial period of time the rest of the state was recolonized by ancestors of the current subspecies east of *M. m. maxima*. This complex might be an ideal group for future studies of evolution of island populations and local adaptation.

## ACKNOWLEDGEMENTS

We thank Jacob Adams, Robin Andrews, Syndonia Bret-Harte, Stefanie Ickert-Bond, Rachel Richardson, Alex Sletten, Naoki Takebayashi, Laura Weingartner, Jeff Wells, and several anonymous reviewers for their invaluable input on earlier drafts. We also want to thank all the museum collectors who collected the specimens used in this project.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was supported by the Kessel Fund, the UAMN Friends of Ornithology, and the Florida Institute of Technology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:  
Kessel Fund, the UAMN Friends of Ornithology, and the Florida Institute of Technology.

### Competing Interests

K.A. Collier is employed by RENECO International Wildlife Consultants.

### Author Contributions

- Caitlyn C. Oliver Brown performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Keiler A. Collier performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- David J. X. Tan performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Kendall Mills performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Fern Spaulding performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Travis C. Glenn conceived and designed the experiments, authored or reviewed drafts of the article, generated sequence data, and approved the final draft.
- Christin Pruett conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Kevin Winker conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

## DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The raw sequence data is available at NCBI Sequence Read Archive BioProject: [PRJNA1114297](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1114297).

## Data Availability

The following information was supplied regarding data availability:

The raw data and code for phenotypic analyses are available at GitHub and Zenodo:  
 - <https://github.com/coliverbrown/melospiza-melodia-phenotype>.  
 - Caitlyn Oliver Brown. (2024). coliverbrown/melospiza-melodia-phenotype: *Melospiza melodia* morphology analysis (v1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.10994578>.

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.19986#supplemental-information>.

## REFERENCES

- Aldrich JW. 1984.** Ecogeographical variation in size and proportions of song sparrows (*Melospiza melodia*). *Ornithological Monographs* **27**(35):iii–134 DOI [10.2307/40166779](https://doi.org/10.2307/40166779).
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990.** Basic local alignment search tool. *Journal of Molecular Biology* **215**(3):403–410 DOI [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Andolfatto P. 2008.** Controlling type-I error of the McDonald–Kreitman test in genomewide scans for selection on noncoding DNA. *Genetics* **180**(3):1767–1771 DOI [10.1534/genetics.108.091850](https://doi.org/10.1534/genetics.108.091850).
- Arcese P, Sogge MK, Marr AB, Patten MA. 2020.** Song sparrow (*Melospiza melodia*), version 1.0. In: Poole AF, Gill FB, eds. *Birds of the World*. Ithaca, NY, USA: Cornell Lab of Ornithology.
- Ashton KG. 2002.** Patterns of within-species body size variation of birds: strong evidence for Bergmann’s rule. *Global Ecology and Biogeography* **11**(6):505–523 DOI [10.1046/j.1466-822X.2002.00313.x](https://doi.org/10.1046/j.1466-822X.2002.00313.x).
- Avise JC, Wollenberg K. 1997.** Phylogenetics and the origin of species. *Proceedings of the National Academy of Sciences of the United States of America* **94**(15):7748–7755 DOI [10.1073/pnas.94.15.7748](https://doi.org/10.1073/pnas.94.15.7748).
- Bazzi G, Cecere JG, Caprioli M, Gatti E, Gianfranceschi L, Podofillini S, Possenti CD, Ambrosini R, Saino N, Spina F, Rubolini D. 2016.** Clock gene polymorphism, migratory behaviour and geographic distribution: a comparative study of trans-Saharan migratory birds. *Molecular Ecology* **25**(24):6077–6091 DOI [10.1111/mec.13913](https://doi.org/10.1111/mec.13913).
- Beck MW. 2022.** ggord: ordination plots with ggplot2. R Package Version 1.1.7. Available at <https://zenodo.org/badge/latestdoi/35334615>.
- Bergmann C. 1847.** On the relationship between the heat economy of animals and their size. *Göttingen Studies* **3**:595–708 [In German].
- Bolger AM, Lohse M, Usadel B. 2014.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**(15):2114–2120 DOI [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
- Bomblies K, Peichel CL. 2022.** Genetics of adaptation. *Proceedings of the National Academy of Sciences of the United States of America* **119**(30):e2122152119 DOI [10.1073/pnas.2122152119](https://doi.org/10.1073/pnas.2122152119).
- Broad Institute. 2019.** Picard toolkit. Available at <https://broadinstitute.github.io/picard/>.

- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10(1):421 DOI 10.1186/1471-2105-10-421.
- Campana MG, Corvelo A, Shelton J, Callicrate TE, Bunting KL, Riley-Gillis B, Wos F, DeGrazia J, Jarvis ED, Fleischer RC. 2020. Adaptive radiation genomics of two ecologically divergent Hawai'ian honeycreepers: the 'akiapōlā'au and the Hawai'i 'amakihi. *Journal of Heredity* 111:21–32 DOI 10.1093/jhered/esz057.
- Carbeck K, Arcese P, Lovette I, Pruett C, Winker K, Walsh J. 2023. Candidate genes under selection in song sparrows co-vary with climate and body mass in support of Bergmann's Rule. *Nature Communications* 14(1):6974 DOI 10.1038/s41467-023-42786-2.
- Chaves JA, Cooper EA, Hendry AP, Podos J, De León LF, Raeymaekers JAM, MacMillan WO, Uy JAC. 2016. Genomic variation at the tips of the adaptive radiation of Darwin's finches. *Molecular Ecology* 25(21):5282–5295 DOI 10.1111/mec.13743.
- Choi JY, Dai X, Alam O, Peng JZ, Rughani P, Hickey S, Harrington E, Juul S, Ayroles JF, Purugganan MD, Stacy EA. 2021. Ancestral polymorphisms shape the adaptive radiation of *Metrosideros* across the Hawaiian Islands. *Proceedings of the National Academy of Sciences of the United States of America* 118(37):e2023801118 DOI 10.1073/pnas.2023801118.
- Delmore KE, Hübner S, Kane NC, Schuster R, Andrew RL, Câmara F, Guigó R, Irwin DE. 2015. Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation. *Molecular Ecology* 24(8):1873–1888 DOI 10.1111/mec.13150.
- Eckert AJ, Carstens BC. 2008. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. *Molecular Phylogenetics and Evolution* 49(3):832–842 DOI 10.1016/j.ympev.2008.09.008.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32(5):1792–1797 DOI 10.1093/nar/gkh340.
- Enbody ED, Sprehn CG, Abzhanov A, Bi H, Dobrev MP, Osborne OG, Rubin C-J, Grant PR, Grant BR, Andersson L. 2021. A multispecies BCO2 beak color polymorphism in the Darwin's finch radiation. *Current Biology* 31(24):5597–5604.e7 DOI 10.1016/j.cub.2021.09.085.
- Faircloth BC. 2013. illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. Available at <http://dx.doi.org/10.6079/J9ILL>.
- Faircloth BC. 2016. PHYLUCES is a software package for the analysis of conserved genomic loci. *Bioinformatics* 32(5):786–788 DOI 10.1093/bioinformatics/btv646.
- Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology* 61(5):717–726 DOI 10.1093/sysbio/sys004.
- Friis G, Fandos G, Zellmer AJ, McCormack JE, Faircloth BC, Milá B. 2018. Genome-wide signals of drift and local adaptation during rapid lineage divergence in a songbird. *Molecular Ecology* 27(24):5137–5153 DOI 10.1111/mec.14946.
- Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* 34(1):397–423 DOI 10.1146/annurev.ecolsys.34.011802.132421.
- Gabrielson IN, Lincoln FC. 1959. *The birds of Alaska*. Harrisburg: The Stackpole Company.
- Gamboa MP, Ghalambor CK, Sillett TS, Funk WC, Furbush RA, Husak JF, Danner RM. 2024. *Melospiza melodia* (Song Sparrow) bill size is primarily shaped by thermoregulation on the California Channel Islands. *Ornithology* 141(4):ukae028 DOI 10.1093/ornithology/ukae028.



- Gazda MA, Toomey MB, Araújo PM, Lopes RJ, Afonso S, Myers CA, Serres K, Kiser PD, Hill GE, Corbo JC, Carneiro M. 2020. Genetic basis of *de novo* appearance of carotenoid ornamentation in bare parts of canaries. *Molecular Biology and Evolution* 37(5):1317–1328 DOI 10.1093/molbev/msaa006.
- Ghosh Roy S, Brejcha J, Mojzeš P, Abdu M, Abdu U. 2025. The role of two tyrosinase-like glycoenzymes in defining the final hue of parrot plumage. *Pigment Cell & Melanoma Research* 38(3):e70010 DOI 10.1111/pcmr.70010.
- Gibson DD, Withrow JJ. 2015. Inventory of the species and subspecies of Alaska birds. *Western Birds* 46:94–185.
- Gillespie RG. 2002. Biogeography of spiders on remote oceanic islands of the Pacific: archipelagoes as stepping stones? *Journal of Biogeography* 29(5–6):655–662 DOI 10.1046/j.1365-2699.2002.00714.x.
- Glenn TC, Nilsen RA, Kieran TJ, Sanders JG, Bayona-Vásquez NJ, Finger JW, Pierson TW, Bentley KE, Hoffberg SL, Louha S, Garcia-De Leon FJ, del Rio Portilla MA, Reed KD, Anderson JL, Meece JK, Aggrey SE, Rekaya R, Alabady M, Belanger M, Winker K, Faircloth BC. 2019. Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). *PeerJ* 7(s1):e7755 DOI 10.7717/peerj.7755.
- Graham AM, Peters JL, Wilson RE, Muñoz-Fuentes V, Green AJ, Dorfsman DA, Valqui TH, Winker K, McCracken KG. 2021. Adaptive introgression of the beta-globin cluster in two Andean waterfowl. *Heredity* 127:107–123 DOI 10.1038/s41437-021-00437-6.
- Grant PR. 1981. Speciation and the adaptive radiation of Darwin’s finches: the complex diversity of Darwin’s finches may provide a key to the mystery of how intraspecific variation is transformed into interspecific variation. *American Scientist* 69:653–663 DOI 10.1515/9781400837946-013.
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* 35(2):518–522 DOI 10.1093/molbev/msx281.
- Huang J, Wang C, Ouyang J, Tang H, Zheng S, Xiong Y, Gao Y, Wu Y, Wang L, Yan X, Chen H. 2022. Identification of key candidate genes for beak length phenotype by whole-genome resequencing in geese. *Frontiers in Veterinary Science* 9:847481 DOI 10.3389/fvets.2022.847481.
- Joseph L. 2021. Species limits in birds: Australian perspectives on interrelated challenges of allopatry, introgression of mitochondrial DNA, recent speciation, and selection. *Ornithology* 138(2):ukab012 DOI 10.1093/ornithology/ukab012.
- Kahle KT, Rinehart J, Lifton RP. 2010. Phosphoregulation of the Na–K–2Cl and K–Cl cotransporters by the WNK kinases. *Biochimica et Biophysica Acta (BBA)—Molecular Basis of Disease* 1802(12):1150–1158 DOI 10.1016/j.bbadis.2010.07.009.
- Kassambara A. 2023. ggpubr: “ggplot2” based publication ready plots. Available at <https://rpkgs.datanovia.com/ggpubr/index.html>.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30(4):772–780 DOI 10.1093/molbev/mst010.
- Kawecki TJ, Ebert D. 2004. Conceptual issues in local adaptation. *Ecology Letters* 7(12):1225–1241 DOI 10.1111/j.1461-0248.2004.00684.x.
- Kolbe JJ, Leal M, Schoener TW, Spiller DA, Losos JB. 2012. Founder effects persist despite adaptive differentiation: a field experiment with lizards. *Science* 335(6072):1086–1089 DOI 10.1126/science.1209566.

- Kwak JS, León-Tapia MÁ, Diblasi C, Manousi D, Grønvold L, Sandvik GK, Saitou M. 2024. Functional and regulatory diversification of *Period* genes responsible for circadian rhythm in vertebrates. *G3: Genes, Genomes, Genetics* 14(10):jkae162 DOI 10.1093/g3journal/jkae162.
- Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerová M, Rubin C-J, Wang C, Zamani N, Grant BR, Grant PR, Webster MT, Andersson L. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518(7539):371–375 DOI 10.1038/nature14181.
- Leaché AD, Harris RB, Rannala B, Yang Z. 2014. The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology* 63(1):17–30 DOI 10.1093/sysbio/syt049.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14):1754–1760 DOI 10.1093/bioinformatics/btp324.
- Li H, Fu Y-X. 2003. NeutralityTest: novel software for performing tests of neutrality. Available at <https://www.picb.ac.cn/evolgen/softwares/download/NeutralityTest/MKtest.htm>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079 DOI 10.1093/bioinformatics/btp352.
- Liu R, Sun Y, Zhao G, Wang H, Zheng M, Li P, Liu L, Wen J. 2015. Identification of loci and genes for growth related traits from a genome-wide association study in a slow- × fast-growing broiler chicken cross. *Genes & Genomics* 37(10):829–836 DOI 10.1007/s13258-015-0314-1.
- Lomolino MV. 2005. Body size evolution in insular vertebrates: generality of the island rule. *Journal of Biogeography* 32(10):1683–1699 DOI 10.1111/j.1365-2699.2005.01314.x.
- Losos JB, Ricklefs RE. 2009. Adaptation and diversification on islands. *Nature* 457(7231):830–836 DOI 10.1038/nature07893.
- Losos JB, Warheitt KI, Schoener TW. 1997. Adaptive differentiation following experimental island colonization in *Anolis* lizards. *Nature* 387(6628):70–73 DOI 10.1038/387070a0.
- Louha S, Ray DA, Winker K, Glenn TC. 2020. A high-quality genome assembly of the North American song sparrow, *Melospiza melodia*. *G3 Genes|Genomes|Genetics* 10(4):1159–1166 DOI 10.1534/g3.119.400929.
- MacArthur RH, Wilson EO. 1967. *The theory of Island biogeography*. Princeton: Princeton University Press.
- Mayr E. 1963. *Animal species and evolution*. Cambridge: Harvard University Press.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351(6328):652–654 DOI 10.1038/351652a0.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequence data. *Genome Research* 20(9):1297–1303 DOI 10.1101/gr.107524.110.
- Meiri S, Dayan T. 2003. On the validity of Bergmann's rule. *Journal of Biogeography* 30(3):331–351 DOI 10.1046/j.1365-2699.2003.00837.x.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, 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 DOI 10.1093/molbev/msaa015.
- Murie OJ. 1959. *Fauna of the Aleutian Islands and Alaska Peninsula*. Washington, D.C.: US Fish and Wildlife Service.

- Pap PL, Vágási CI, Cziráj GÁ, Titilincu A, Pinte A, Barta Z. 2009. Carotenoids modulate the effect of coccidian infection on the condition and immune response in moulting house sparrows. *Journal of Experimental Biology* 212(20):3228–3235 DOI 10.1242/jeb.031948.
- Parsch J, Zhang Z, Baines JF. 2009. The influence of demography and weak selection on the McDonald–Kreitman test: an empirical study in *Drosophila*. *Molecular Biology and Evolution* 26(3):691–698 DOI 10.1093/molbev/msn297.
- Patten MA, Pruett CL. 2009. The song sparrow, *Melospiza melodia*, as a ring species: patterns of geographic variation, a revision of subspecies, and implications for speciation. *Systematics and Biodiversity* 7:33–62 DOI 10.1017/S1477200008002867.
- Pruett CL, Li T, Winker K. 2018. Population genetics of Alaska Common Raven show dispersal and isolation in the world’s largest songbird. *The Auk* 135(4):868–880 DOI 10.1642/AUK-17-144.1.
- Pielou EC. 1991. *After the Ice Age: the return of life to glaciated North America*. Chicago: University of Chicago Press.
- Posit Team. 2025. *RStudio: integrated development environment for R*. PBC, Boston, MA: Posit Software. Available at <http://www.posit.co/>.
- Pruett CL, Winker K. 2005a. Biological impacts of climatic change on a Beringian endemic: cryptic refugia in the establishment and differentiation of the Rock Sandpiper (*Calidris ptilocnemis*). *Climatic Change* 68(1–2):219–240 DOI 10.1007/s10584-005-1584-4.
- Pruett CL, Winker K. 2005b. Northwestern song sparrow populations show genetic effects of sequential colonization. *Molecular Ecology* 14(5):1421–1434 DOI 10.1111/j.1365-294X.2005.02493.x.
- Pruett CL, Winker K. 2010. Chapter 13: Alaska song sparrows (*Melospiza melodia*) demonstrate that genetic marker and method of analysis matter in subspecies assessments. *Ornithological Monographs* 67(1):162–171 DOI 10.1525/om.2010.67.1.162.
- Rand DM, Kann LM. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. *Molecular Biology and Evolution* 13(6):735–748 DOI 10.1093/oxfordjournals.molbev.a025634.
- R Core Team. 2021. *R: a language and environment for statistical computing*. Vienna, Austria: RFoundation for Statistical Computing. Available at <https://www.R-project.org/>.
- Reding DM, Foster JT, James HF, Pratt HD, Fleischer RC. 2008. Convergent evolution of ‘creepers’ in the Hawaiian honeycreeper radiation. *Biology Letters* 5(2):221–224 DOI 10.1098/rsbl.2008.0589.
- Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J, Lee C, Ko BJ, Chaisson M, Gedman GL, Cantin LJ, Thibaud-Nissen F, Haggerty L, Bista I, Smith M, Haase B, Mountcastle J, Winkler S, Paez S, Howard J, Vernes SC, Lama TM, Grutzner F, Warren WC, Balakrishnan CN, Burt D, George JM, Biegler MT, Iorns D, Digby A, Eason D, Robertson B, Edwards T, Wilkinson M, Turner G, Meyer A, Kautt AF, Franchini P, Detrich HW III, Svoldal H, Wagner M, Naylor GJP, Pippel M, Malinsky M, Mooney M, Simbirsky M, Hannigan BT, Pesout T, Houck M, Misuraca A, Kingan SB, Hall R, Kronenberg Z, Sović I, Dunn C, Ning Z, Hastie A, Lee J, Selvaraj S, Green RE, Putnam NH, Gut I, Ghurye J, Garrison E, Sims Y, Collins J, Pelan S, Torrance J, Tracey A, Wood J, Dagnew RE, Guan D, London SE, Clayton DF, Mello CV, Friedrich SR, Lovell PV, Osipova E, Al-Ajli FO, Secomandi S, Kim H, Theofanopoulou C, Hiller M, Zhou Y, Harris RS, Makova KD, Medvedev P, Hoffman J, Masterson P, Clark K, Martin F, Howe K, Flicek P, Walenz BP, Kwak W, Clawson H, Diekhans M, Nassar L, Paten B, Kraus RSH, Crawford AJ, Gilbert MTP, Zhang G, Venkatesh B, Murphy RW,

- Koepfli K-P, Shapiro B, Johnson WE, Di Palma F, Marques-Bonet T, Teeling EC, Warnow T, Graves JM, Ryder OA, Haussler D, O'Brien SJ, Korlach J, Lewin HA, Howe K, Myers EW, Durbin R, Phillippy AM, Jarvis ED. 2021. Towards complete and error-free genome assemblies of all vertebrate species. *Nature* 592(7856):737–746 DOI 10.1038/s41586-021-03451-0.
- Ruegg K, Anderson EC, Boone J, Pouls J, Smith TB. 2014. A role for migration-linked genes and genomic islands in divergence of a songbird. *Molecular Ecology* 23(19):4757–4769 DOI 10.1111/mec.12842.
- Saino N, Bazzi G, Gatti E, Caprioli M, Cecere JG, Possenti CD, Galimberti A, Orioli V, Bani L, Rubolini D, Gianfranceschi L, Spina F. 2015. Polymorphism at the *Clock* gene predicts phenology of long-distance migration in birds. *Molecular Ecology* 24(8):1758–1773 DOI 10.1111/mec.13159.
- San-Jose LM, Bestion E, Pellerin F, Richard M, Di Gesu L, Salmona J, Winandy L, Legrand D, Bonneaud C, Guillaume O, Calvez O, Elmer KR, Yurchenko AA, Recknagel H, Clobert J, Cote J. 2023. Investigating the genetic basis of vertebrate dispersal combining RNA-seq, RAD-seq and quantitative genetics. *Molecular Ecology* 32(12):3060–3075 DOI 10.1111/mec.16916.
- San-Jose LM, Ducrest A-L, Ducret V, Simon C, Richter H, Wakamatsu K, Roulin A. 2017. MC1R variants affect the expression of melanocortin and melanogenic genes and the association between melanocortin genes and coloration. *Molecular Ecology* 26(1):259–276 DOI 10.1111/mec.13861.
- Savolainen O, Lascoux M, Merilä J. 2013. Ecological genomics of local adaptation. *Nature Reviews Genetics* 14(11):807–820 DOI 10.1038/nrg3522.
- Srichai MB, Colleta H, Gewin L, Matrisian L, Abel TW, Koshikawa N, Seiki M, Pozzi A, Harris RC, Zent R. 2011. Membrane-type 4 matrix metalloproteinase (MT4-MMP) modulates water homeostasis in mice. *PLOS ONE* 6(2):e17099 DOI 10.1371/journal.pone.0017099.
- Stoletzki N, Eyre-Walker A. 2011. Estimation of the neutrality index. *Molecular Biology and Evolution* 28(1):63–70 DOI 10.1093/molbev/msq249.
- Subramanian S. 2016. The effects of sample size on population genomic analyses—implications for the tests of neutrality. *BMC Genomics* 17:123 DOI 10.1186/s12864-016-2441-8.
- Tattersall GJ, Arnaout B, Symonds MRE. 2017. The evolution of the avian bill as a thermoregulatory organ. *Biological Reviews* 92(3):1630–1656 DOI 10.1111/brv.12299.
- Tiffin P, Ross-Ibarra J. 2014. Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution* 29(12):673–680 DOI 10.1016/j.tree.2014.10.004.
- Toews DPL, Hofmeister NR, Taylor SA. 2017. The evolution and genetics of carotenoid processing in animals. *Trends in Genetics* 33(3):171–182 DOI 10.1016/j.tig.2017.01.002.
- Turesson G. 1922. The species and the variety as ecological units. *Hereditas* 3(1):100–113 DOI 10.1111/j.1601-5223.1922.tb02727.x.
- Vallone D, Frigato E, Vernesi C, Foà A, Foulkes NS, Bertolucci C. 2007. Hypothermia modulates circadian clock gene expression in lizard peripheral tissues. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 292(1):R160–R166 DOI 10.1152/ajpregu.00370.2006.
- von Lintig J. 2010. Colors with functions: elucidating the biochemical and molecular basis of carotenoid metabolism. *Annual Review of Nutrition* 30:35–56 DOI 10.1146/annurev-nutr-080508-141027.
- Walsh J, Benham PM, Deane-Coe PE, Arcese P, Butcher BG, Chan YL, Cheviron ZA, Elphick CS, Kovach AI, Olsen BJ, Shriver WG, Winder VL, Lovette IJ. 2019. Genomics of

- rapid ecological divergence and parallel adaptation in four tidal marsh sparrows. *Evolution Letters* 3(4):324–338 DOI 10.1002/evl3.126.
- Walsh N, Dale J, McGraw KJ, Pointer MA, Mundy NI. 2012. Candidate genes for carotenoid coloration in vertebrates and their expression profiles in the carotenoid-containing plumage and bill of a wild bird. *Proceedings of the Royal Society B: Biological Sciences* 279(1726):58–66 DOI 10.1098/rspb.2011.0765.
- Warren BH, Simberloff D, Ricklefs RE, Aguilée R, Condamine FL, Gravel D, Morlon H, Mouquet N, Rosindell J, Casquet J, Conti E, Cornuault J, Fernández-Palacios JM, Hengl T, Norder SJ, Rijsdijk KF, Sanmartín I, Strasberg D, Triantis KA, Valente LM, Whittaker RJ, Gillespie RG, Emerson BC, Thébaud C. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur–Wilson. *Ecology Letters* 18(2):200–217 DOI 10.1111/ele.12398.
- Whittaker RJ, Fernandez-Palacios JM. 2007. *Island biogeography: ecology, evolution, and conservation*. Oxford: Oxford University Press.
- Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the tidyverse. *Journal of Open Source Software* 4(43):1686 DOI 10.21105/joss.01686.
- Wickham H, Hester J, Chang W, Bryan J. 2022. devtools: tools to make developing R packages easier. Available at <https://devtools.r-lib.org/>.
- Williams GC. 1966. *Adaptation and natural selection*. Princeton, NJ, USA: Princeton University Press.
- Winker K. 2000. Obtaining, preserving, and preparing bird specimens. *Journal of Field Ornithology* 71(2):250–297 DOI 10.1648/0273-8570-71.2.250.
- Winker K. 2010. Chapter 1: subspecies represent geographically partitioned variation, a gold mine of evolutionary biology, and a challenge for conservation. *Ornithological Monographs* 67(1):6–23 DOI 10.1525/om.2010.67.1.6.
- Winker K, Withrow JJ, Gibson DD, Pruett CL. 2023. Beringia as a high-latitude engine of avian speciation. *Biological Reviews* 98(4):1081–1099 DOI 10.1111/brv.12945.
- Zink RM, Dittmann DL. 1993. Gene flow, refugia, and evolution of geographic variation in the song sparrow (*Melospiza melodia*). *Evolution Education and Outreach* 47(3):717–729 DOI 10.1111/j.1558-5646.1993.tb01228.x.