



## Cytonuclear discordance in the crowned-sparrows, *Zonotrichia atricapilla* and *Zonotrichia leucophrys*

Rebecca S. Taylor<sup>1</sup>, Ashley C. Bramwell<sup>2</sup>, Rute Clemente-Carvalho<sup>3</sup>, Nicholas A. Cairns, Frances Bonier, Katherine Dares, Stephen C. Loughheed<sup>\*</sup>

Department of Biology, Queen's University, Kingston, ON K7L 3N6, Canada

### ARTICLE INFO

#### Keywords:

Crowned-sparrows  
ddRADSeq  
Mitochondrial introgression  
Incomplete lineage sorting  
Speciation

### ABSTRACT

The golden-crowned (*Zonotrichia atricapilla*) and white-crowned (*Z. leucophrys*) sparrows have been presented as a compelling case for rapid speciation. They display divergence in song and plumage with overlap in their breeding ranges implying reproductive isolation, but have almost identical mitochondrial genomes. Previous research proposed hybridization and subsequent mitochondrial introgression as an alternate explanation, but lacked robust nuclear gene trees to distinguish between introgression and incomplete lineage sorting. We test for signatures of these processes between *Z. atricapilla* and *Z. leucophrys*, and investigate the relationships among *Z. leucophrys* subspecies, using mitochondrial sequencing and a reduced representation nuclear genomic dataset. Contrary to the paraphyly evident in mitochondrial gene trees, we confirmed the reciprocal monophyly of *Z. atricapilla* and *Z. leucophrys* using large panels of single nucleotide polymorphisms (SNPs). The pattern of cytonuclear discordance is consistent with limited, historical hybridization and mitochondrial introgression, rather than a recent origin and incomplete lineage sorting between recent sister species. We found evidence of nuclear phylogeographic structure within *Z. leucophrys* with two distinct clades. Altogether, our results indicate deeper divergences between *Z. atricapilla* and *Z. leucophrys* than inferred using mitochondrial markers. Our results demonstrate the limitations of relying solely on mitochondrial DNA for taxonomy, and raise questions about the possibility of selection on the mitochondrial genome during temperature oscillations (e.g. during the Pleistocene). Historical mitochondrial introgression facilitated by past environmental changes could cause erroneous dating of lineage splitting in other taxa when based on mitochondrial DNA alone.

### 1. Introduction

Mitochondrial DNA (mtDNA) sequence data have been a key source for phylogenetic analyses in animals for several decades (Avisé et al. 1987; Avisé 2000). While mtDNA can provide a useful measure of diversity and critical isolating mechanisms (Sloan et al., 2017), phylogenies inferred primarily from mtDNA may not reflect the species' evolutionary history (Chan and Levin 2005). Both nuclear (nDNA) and mtDNA genomes are known to introgress independently including the wholesale replacement (capture) of genomes and regions (Toews and Brelsford, 2012; Harrison and Larson 2014). Introgression between non-sister lineages can render the species trees meaningless or misleading (Mallet et al. 2016; Edwards et al. 2016). Even between sister taxa where

introgression is more common, such introgression/capture events will make the divergence times appear more recent (Leache et al. 2013). As such, relying solely on mtDNA may lead to a mis-estimation of taxonomic relationships and divergence estimates (Willis et al. 2014).

Many cases of rapid speciation have been inferred from dating of mitochondrial (mtDNA) phylogenies (e.g. Arnaiz-Villena et al., 1998; Shen et al., 2010). Among vertebrates, examples of rapid speciation are particularly common from the avian literature. These include: goldfinches (*Carduelis*; Arnaiz-Villena et al., 1998), Darwin's finches (*Geospizinae*; Grant, 1999), pheasants (Shen et al., 2010), and the southern capuchino seedeaters (*Sporophila*; Campagna et al., 2012). Rapid speciation in birds is typically accompanied by well-developed prezygotic isolating mechanisms (e.g. based on plumage or songs), despite

<sup>\*</sup> Corresponding author at: Department of Biology, Queen's University, Kingston, Ontario, Canada.

E-mail address: [steve.loughheed@queensu.ca](mailto:steve.loughheed@queensu.ca) (S.C. Loughheed).

<sup>1</sup> Present address: Biology Department, Trent University, Peterborough, ON K9J 7B8, Canada

<sup>2</sup> Present address: Department of Biological Sciences, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON M1C 0B1, Canada

<sup>3</sup> Present address: Hakai Institute, 1713 Hyacinthe Bay Rd, Harlot Bay, BC V0P 1N0, Canada

<https://doi.org/10.1016/j.ympev.2021.107216>

Received 16 July 2020; Received in revised form 27 April 2021; Accepted 27 May 2021

Available online 1 June 2021

1055-7903/© 2021 The Authors.

Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

sister species being genetically similar across most of their genome (Edwards et al., 2005). Here we investigate the divergence between golden-crowned (*Zonotrichia atricapilla*) and white-crowned (*Z. leucophrys*) sparrows, which have been touted as an example of exceedingly rapid speciation, perhaps splitting as recently as 50,000 years ago (Johnson and Cicero, 2004); however empirical evidence does not uniformly support this recent speciation hypothesis (see Weckstein et al., 2001).

The white-crowned sparrow has a geographic range spanning most of North America, while the golden-crowned is generally confined to regions west of the Rocky Mountains (Cortopassi and Mewaldt, 1965; Fig. 1). Both species are migratory, breeding in the northern parts of their ranges and overwintering in the south, although *Z. leucophrys* have some year-round resident populations along the coast of Washington, Oregon and California. Adults are distinct in their plumages, breeding behaviours, and vocalizations (del Hoyo et al., 2011). Playback experiments have shown that both sexes of *Z. leucophrys* react more strongly to conspecific than heterospecific songs (Pleszczyńska, 1979), and to natal versus non-natal vocal dialects (Baker et al., 1982). Female *Z. l. nuttalli* mate assortatively with males of their natal dialect population (Tomback and Baker, 1984). Finally, while some reports of *Z. atricapilla*-*Z. leucophrys* hybrids exist (Miller, 1940; Morton and Mewaldt, 1960; McCarthy, 2006), these represent a comparatively small number given the extensive overlap in their breeding distributions across British Columbia, the Yukon, and Alaska (Cortopassi and Mewaldt, 1965). All existing evidence thus suggests that these are good biological species.

Early genealogies based on allozyme data (Zink, 1982) and mitochondrial DNA (mtDNA) (Zink et al., 1991) implied that *Z. atricapilla* and *Z. leucophrys* were independently-evolving sister species. However, subsequent analyses of mtDNA sequences suggested that they are not reciprocally monophyletic, even sharing some haplotypes (Weckstein et al., 2001; McKereghan-Dares, 2008), and that their low global mitochondrial nucleotide diversity is comparable to a single, homogenous avian population rather than two well-diverged species. Weckstein et al. (2001) presented alternate hypotheses to explain this pattern: 1) *Z. atricapilla* and *Z. leucophrys* originated in allopatry, followed by contact, historical hybridization and subsequent introgression, with ultimate replacement of one ancestral mitochondrial genome in one species

with that of the other (i.e. mtDNA capture; see Good et al., 2008); or 2) these two species originated very recently, with insufficient time for complete lineage sorting to have occurred.

Within *Z. leucophrys* there are five named subspecies (*Z. l. nuttalli*, *Z. l. pugetensis*, *Z. l. gambelii*, *Z. l. oriantha* and *Z. l. leucophrys*). A recent genomic study of nuclear DNA investigated population structure across a contact zone of *Z. l. nuttalli* and *Z. l. pugetensis*. Differentiation was found between the two subspecies, driven by differences in male advertising song, indicating that they are separate evolutionary significant units (Lipshutz et al., 2017). However, secondary contact has led to ongoing hybridization between the two subspecies (Lipshutz et al., 2017; Toews, 2017). Genetic relationships based on nuclear DNA among the other subspecies are largely unknown.

The five crowned-sparrow species (golden-crowned and white-crowned sparrows along with Harris' sparrow, *Z. querula*, white-throated sparrow, *Z. albicollis*, and the rufous-collared sparrow, *Z. capensis*) have contributed significantly to our current understanding of speciation with research on reproductive isolation (e.g. Danner et al., 2011), mating behaviour (e.g. Tomback and Baker, 1984; Houtman and Falls, 1994; Lipshutz et al., 2017), phenology (e.g. Handford, 1980; Moore et al., 2005), migration (e.g. Landys et al., 2004; McFarlan et al., 2009; Price et al., 2010), vocal dialects, song learning, and function (e.g. Handford and Nottebohm, 1976; Lougheed and Handford, 1992; MacDougall-Shackleton and MacDougall-Shackleton, 2001; Shizuka, 2014; Toews, 2017), behavioural genetics (Merritt et al., 2020) and phylogeography, geographic range expansions, and secondary contact zones (Lougheed et al., 2013; Campagna et al., 2014). Despite their prominence in evolutionary and ecological studies, there is phylogenetic uncertainty about the relationships among the species (Carson and Spicer, 2003).

Here, we use a genome-wide panel of nuclear Single Nucleotide Polymorphisms (SNPs) combined with mitochondrial DNA sequencing to investigate the relationships between *Z. atricapilla* and *Z. leucophrys* to test the alternate propositions presented by Weckstein et al. (2001). If the species originated, as proposed in the historical hybridization hypothesis, with allopatric divergence, subsequent sympatry and hybridization, and introgression of mitochondrial DNA from one species into another, we expect signatures of genetic differentiation and monophyly in the nuclear DNA, but not the mitochondrial DNA. Conversely, under the recent speciation hypothesis we predict low genetic differentiation and a lack of reciprocal monophyly in the nuclear and mitochondrial DNA. We include samples from all subspecies within *Z. leucophrys*, and from all five crowned sparrow species so that we can assess the phylogenetic affinities of these taxa.

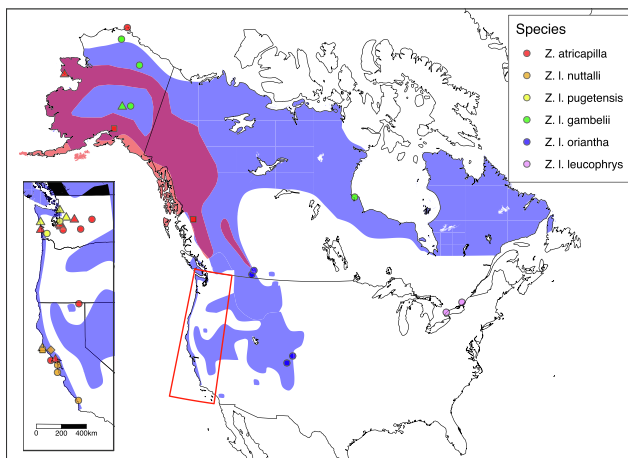
## 2. Materials and methods

### 2.1. Sampling

From museum collections and our own sample archive, we assembled 73 tissue and blood samples of *Z. atricapilla* and *Z. leucophrys* from across their respective geographic ranges, including individuals from all five subspecies of *Z. leucophrys* (Fig. 1; Table S1) for ddRAD sequencing. For mtDNA sequencing, 74 samples were used (Table S1). We added eight specimens from the remaining three *Zonotrichia* species (two from *Z. albicollis*, three from *Z. capensis*, two from *Z. querula*; Table S1; Fig. S1). Tissue samples were preserved in lysis buffer (4.0 M urea, 0.2 M NaCl, 10 mM EDTA, 0.5% n-Lauroyl Sarcosine, 100 mM Tris-CHL; pH 8.0) or 99% ethanol until extraction.

### 2.2. DNA extraction

Total genomic DNA was extracted either using the DNeasy® Blood and Tissue Kit (QIAGEN, Mississauga, Ontario, Canada) following the manufacturer's protocols, or a standard phenol-chloroform method (Sambrook & Russell, 2001). For ethanol preserved samples, excess



**Fig. 1.** The locations of sampling sites of *Z. atricapilla* and *Z. leucophrys* individuals. The background colors represent the breeding ranges, light pink for *Z. atricapilla* and purple for *Z. leucophrys* with deep pink representing range overlap. Circles indicate samples used for both ddRADseq and mitochondrial DNA sequencing, triangles represent locations where at least one sample has mitochondrial DNA only, squares represent locations where at least one sample has ddRAD data only. Many samples were collected from birds outside of the breeding season (i.e., nonbreeding or migration). More sampling details are provided in Table S1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ethanol was allowed to evaporate before extraction. Samples fixed in lysis buffer were washed in a 1x phosphate-buffered saline (PBS; 10 mM, pH 7.4) solution. All newly extracted samples were treated with RNase A (10 mg/mL) and incubated at 37 °C for 30 min. We checked the purity of our DNA using a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies Inc., Wilmington, DE, USA). If needed, due to low concentration or purity, samples were re-precipitated using a standard ethanol precipitation protocol with sodium acetate (3 M, pH 5.2) as the carrier molecule, and then re-suspended in microfiltered autoclaved water and stored at -20 °C until subsequent analyses. We quantified the concentration of double-stranded DNA in each sample using a Qubit® 2.0 Fluorometer (Life Technologies, Grand Island, NY, USA).

### 2.3. Mitochondrial DNA sequencing

Mitochondrial DNA sequence variation was assayed from control region (approximately 413 bp) and cytochrome oxidase I (COI, 666 bp) using the primers, PCR cocktail recipes, and cycling parameters given in Loughheed et al. (2013; primers: LCA1271-R and ZNGluF3) and Kerr et al. (2009; primers BirdF1 and COIbirdR2), respectively. The amplification products were visualized on 1.0% agarose gels using RedSafe™ Nucleic Acid Staining Solution, and purified using AMPure XP 2X Solid Phase Reversible Immobilization (SPRI) following the manufacturer's instructions. Purified products were sent to The Centre for Applied Genomics (SickKids Hospital, Toronto, ON, Canada) for sequencing on an ABI 3730XL using the BigDye™ terminator kit. Mitochondrial sequences are available on Genbank, the control region under accession numbers MW911496 – MW911580 and COI under accession numbers MW915858 – MW915942.

### 2.4. Double-digest restriction associated DNA (ddRAD) sequencing

Following Poland et al. (2012) double-digest RAD-Seq libraries were constructed at the Institute for Integrative Systems Biology (IBIS; at Laval University) using the restriction enzyme pairing of SbfI and MspI. Uniquely barcoded samples were pooled into a library and then size selected for roughly 100–300 bp using a BluePippin® DNA size selection system (Sage Science Inc., Beverly, MA, USA). The library was forwarded to Génome Québec (Génome Québec Innovation Centre) and assessed using a LabChip GX (Perkin-Elmer, MA, USA) instrument to determine library size quality. The library was then sequenced as single-end, 100 base-pair reads on a single Illumina HiSeq2000 lane (Illumina, San Diego, CA, USA). Raw sequence ddRAD data are available at the National Centre for Biotechnology Information (NCBI), under the BioProject accession number PRJNA720382.

### 2.5. ddRADseq read filtering

Raw reads were de-multiplexed sequentially by barcode length using Stacks 2.3 (Catchen et al., 2013). Low quality reads (with a Phred score below 10), and those with uncalled bases, were filtered out and reads were truncated to 92 bp. De-multiplexed reads were aligned to the *Zonotrichia albicollis* genome from NCBI (GCF000385455, Bioproject PRJNA197293) using the 'very\_sensitive' setting in Bowtie2 (Langmead and Salzberg, 2012). Resulting SAM files were converted to BAM files and sorted using Samtools 1.1 (Li et al., 2009). We then used the Refmap and Populations modules in Stacks randomly selecting one SNP per locus, filtering out loci with a minor allele frequency of less than 0.05. After an initial run, seven individuals were found to have high levels of missing data and were removed, leaving 66 individuals (Table S1). We re-ran Refmap and Populations without the low-quality samples using two filtering settings, one allowing for 10% missing data (all individuals as one population and setting  $r$  to 0.9) and the other without any missing data (setting  $r$  to 1.0). We replicated this process without the outgroup taxa to create input files with only *Z. atricapilla* and *Z. leucophrys*, leaving 58 individuals (18 *Z. atricapilla*, 18 *Z. l. nuttalli*, 9 *Z. l. pugetensis*, 7 *Z. l.*

*gambelli*, 4 *Z. l. oriantha*, and 2 *Z. l. leucophrys*). All analyses were conducted with both datasets (with no missing data and with 10% missing data).

### 2.6. mtDNA network and phylogenetics

We used the concatenated COI and control region sequences (total 1211 bp) to make a statistical parsimony haplotype network in PopArt (Clement et al., 2000; Leigh and Bryant, 2015; PopArt: <http://popart.otago.ac.nz>) using individuals from *Z. leucophrys* and *Z. atricapilla*. We performed Bayesian phylogenetic analyses on a dataset comprising all individuals and a dataset with only individual haplotypes identified in our network analysis. The best partitioning scheme and models of evolution for each dataset were selected using PartitionFinder2 (Lanfear et al., 2016). For the full dataset the best partition scheme and associated models of evolution were: HKY + I for control region, HKY + I for COI codon position 1, F81 for COI codon position 2, and HKY + I for COI codon position 3. For the haplotype dataset the best partition scheme and associated models of evolution were: HKY + I for control region, HKY + I for COI codon position 1, F81 for COI codon position 2, and HKY for COI codon position 3.

For Bayesian phylogenetic relationships we used MrBayes v3.2.6 x64 (Huelsenbeck and Ronquist, 2001, Ronquist et al., 2012). For all models, partitions were unlinked for substitution rates and proportion of invariant sites (I), allowing parameter values to vary independently. Each Bayesian analysis included two simultaneous runs each with random starting trees, one cold and three incrementally heated Markov chains, and default priors for all parameters. Each analysis was allowed to run until the standard deviation of split frequencies < 0.01, sampling every 100 generations (full dataset: 2 million generations; haplotype dataset: 1 million generations). We visually assessed that our runs had achieved stationarity by evaluating the likelihoods in the program Tracer version 1.7.1 (Rambaut et al., 2018). For each analysis, we constructed a 50% majority rule consensus tree (discarding the first 25% of trees as burnin).

### 2.7. Phylogenomic reconstruction and tests for introgression using SNPs

Using the PHYLIP files produced in Stacks, we ran RAxML version 8 (Stamatakis, 2014) to generate a maximum likelihood tree. We used a General Time Reversible model with a gamma distribution of substitution rates and a proportion of invariant sites (GTRGAMMAI), and used the rapid bootstrap algorithm using 1000 replicates. We ran the analysis with all species, but because of the long branch lengths (see Results), we also ran the analysis using *Z. querula* as an outgroup with only *Z. atricapilla* and *Z. leucophrys* individuals to better resolve relationships among lineages.

We ran Treemix 1.13 (Pickrell and Pritchard, 2012) to infer a maximum likelihood tree and to investigate any introgression events between *Z. atricapilla* and/or *Z. leucophrys* sub-species. We ran Treemix from 0 to 9 migration events, running it without grouping SNPs for three iterations, and grouping SNPs into windows of 20 for three iterations to account for possible linkage and because the OptM package (below) needs different likelihood scored between iterations. We plotted the results in RStudio 1.0.136 (RStudio Team, 2015) and used the R package OptM (available here: <https://cran.r-project.org/web/packages/OptM/index.html>) to calculate the ad hoc statistic delta M, which is the second order rate of change in the log-likelihood of the different migration events, to help infer how many migration events to visualize.

We also performed an ABBA/BABA test to calculate the D statistic from the VCF file using Dsuite (Malinsky et al., 2020). Given results from Treemix and ADMIXTURE (see Results), we tested for introgression from *Z. atricapilla* into the two *Z. leucophrys* clades using (((P1,P2)P3)O), where one tests whether there has been introgression from P3 into the sister clades P1 or P2. In our case, the two *Z. leucophrys* clades (see Results) comprise P1 and P2, *Z. atricapilla* is P3, and *Z. querula* is the



outgroup (O). We could not test for introgression from *Z. leucophrys* into *Z. atricapilla* as *Z. atricapilla* is a single clade.

## 2.8. Population structure

All population structure analyses were done using a dataset comprising only *Z. atricapilla* and *Z. leucophrys*. We ran an ADMIXTURE analysis (version 1.3; Alexander et al., 2009) for population assignment. To do this, we converted the VCF file into a BED file using Plink 1.9 (Purcell et al., 2007). We pruned the dataset to remove any SNPs with a correlation coefficient ( $R^2$ ) of 0.1 or greater. We then ran ADMIXTURE from K 1 to 8, and plotted the results in RStudio. We also did a principal component analysis (PCA) in RStudio using the packages ADEGENET (Jombart, 2008) and vcfR (Knaus and Grüwald, 2017). We did an additional PCA on a subset of data with *Z. l. nuttalli* and *Z. l. pugetensis* only, and another PCA on *Z. l. gambelli*, *Z. l. oriantha*, and *Z. l. leucophrys* only to investigate possible substructure (see Results). We calculated pairwise  $F_{ST}$  between *Z. atricapilla* and all *Z. leucophrys* subspecies, as well as between the three major lineages uncovered in the analyses (see Results) in RStudio using the HIERFSTAT package in R (Goudet, 2005) and the Weir and Cockerham calculation (WC84; Weir and Cockerham, 1984).

## 3. Results

### 3.1. ddRADseq data quality

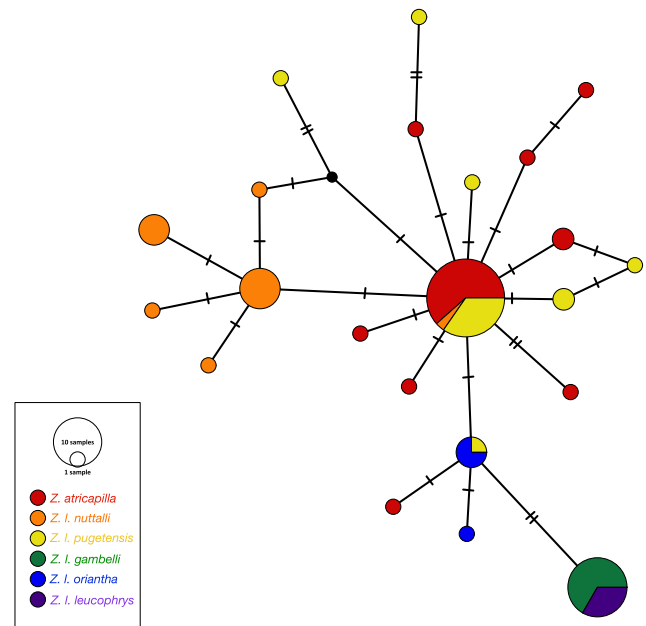
We recovered 18.4 million single-end Illumina reads. After removing individuals with low quality reads, we had 3,469 SNPs in the VCF file with no missing data and 7,096 SNPs in the VCF file allowing 10% missing data for the datasets with all five crowned sparrow species, and 4,139 SNPs in the VCF file with no missing data and 6,626 SNPs in the VCF file allowing 10% missing data in the files with *Z. atricapilla* and *Z. leucophrys* only. The depth was high for all retained individuals, ranging from 41.6 to 183.1X (aside from one *Z. capensis* individual with 481.4 X) for the file with outgroups and no missing data, and between 39.0 and 183.3X (aside from one *Z. capensis* individual which has 464.1 X) in the file allowing 10% missing data. These datasets were used for the RAXML analyses. Similarly, for the dataset comprising *Z. atricapilla* and *Z. leucophrys* only the depth per individual ranged from 42.7 to 183.0X for the file with no missing data and between 39.7 and 169.3 X for the file allowing 10% missing data. These datasets were used for all subsequent nuclear SNP analyses; however, all results that we present are from the analyses with no missing data as results from analysis of different datasets were the same.

### 3.2. Evolutionary relationships apparent from analysis of mitochondrial DNA

The most obvious feature of both the mtDNA network (Fig. 2) and our Bayesian tree from analysis of distinct mtDNA haplotypes (Fig. 3) or of all individuals (Fig. S2) is that *Z. atricapilla* and *Z. leucophrys* share mtDNA haplotypes and do not form distinct, reciprocally monophyletic clusters/clades. We also find no obvious phylogeographic structure among *Z. leucophrys* subspecies, although some haplotypes are unique to particular subspecies. *Zonotrichia albicollis* and *Z. querula* comprise sister lineages, that are in turn, sister to the *Z. atricapilla*/*Z. leucophrys* clade. The North American crowned sparrow clade is reciprocally monophyletic with the rufous-collared sparrow, *Z. capensis*.

### 3.3. Interspecific relationships and population structure in *Zonotrichia* with nuclear DNA

In contrast to our results from analysis of mtDNA, *Z. atricapilla* and *Z. leucophrys* are well-supported, reciprocally monophyletic sister taxa in the RAXML phylogenomic analysis (Fig. 4A and 4B). The *Z. leucophrys*

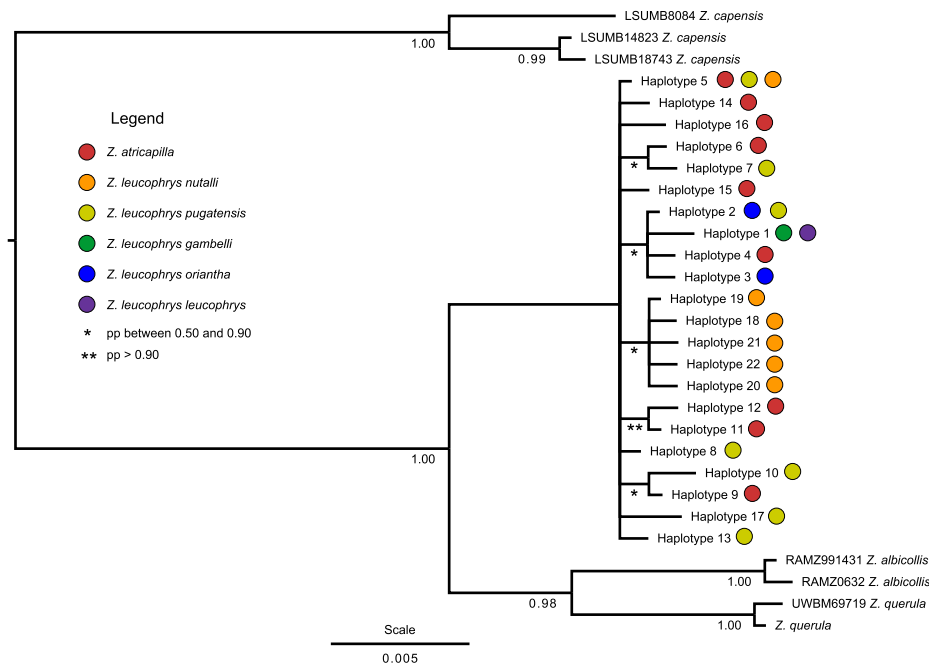


**Fig. 2.** Statistical parsimony network of 1211 bp of the mitochondrial control region and COI concatenated for *Z. atricapilla* and all *Z. leucophrys* subspecies. Black circles represent unsampled or now extinct haplotypes and the hatch marks indicate mutations. Circle size indicates the number of samples of each haplotype.

lineage was further divided into two clades: *Z. l. nuttalli* and *Z. l. pugetensis* (NP), and *Z. l. gambelli*, *Z. l. oriantha*, and *Z. l. leucophrys* (GOL). *Zonotrichia querula* was sister to *Z. atricapilla* and *Z. leucophrys*, followed by *Z. albicollis*, with *Z. capensis* recovered as sister to all North American congeners (Fig. 4A and 4B).

The Treemix maximum likelihood tree was similar, with *Z. atricapilla* well separated from all *Z. leucophrys* subspecies, and showing a large drift parameter (Fig. 5A). The NP clade remains distinct from the GOL subspecies, and shows evidence of some drift (Fig. 5A). Treemix stopped adding migration events after 5, and OptM indicated one migration event as maximizing delta M, which shows migration between the ancestor of the NP clade into *Z. l. leucophrys* (Fig. 5B; Fig S3). When adding three migration events, the analysis did infer migration from *Z. atricapilla* into *Z. l. leucophrys* (Fig. 5D); however adding three migration events only slightly increased the variance explained compared with adding one or two migration events (Fig S3). The ABBA/BABA test also detected a signature of introgression from *Z. atricapilla* into the GOL clade of *Z. leucophrys* ( $D = 0.034$ ,  $p = 0.023$ ).

The ADMIXTURE analysis for  $K = 2$  (which had the lowest cross validation error) shows *Z. atricapilla* as a different population from *Z. leucophrys*, with some *Z. atricapilla* ancestry apparent in the GOL clade (Fig. 6A). The results for  $K = 3$  split the three clades seen in the phylogenetic analyses (*Z. atricapilla*, NP, and GOL) with a small amount of GOL ancestry showing in the NP group, especially in *Z. l. pugetensis* individuals (Fig. 6B). For  $K = 4$ , we see some individuals within the GOL clade splitting off, and for  $K = 5$  we see *Z. l. pugetensis* splitting off with admixed individuals in *Z. l. nuttalli* (Fig. 6C and 6D). The PCA reflected the same pattern, with the same three major groups, *Z. atricapilla*, NP, and GOL, well separated clusters (Fig. 7A). Results from the PCA on only individuals from the NP clade, shows *Z. l. nuttalli* and *Z. l. pugetensis* to be separate clusters (Fig. 7B). Similarly, when analyzing individuals from the GOL clade alone, we see *Z. l. gambelli*, *Z. l. oriantha*, and *Z. l. leucophrys* comprising unique clusters (although sampling is limited for *Z. l. leucophrys*; Fig. 7C). Pairwise  $F_{ST}$  values between *Z. atricapilla* and all *Z. leucophrys* subspecies were high ranging from 0.27 to 0.31 (Table 1). The values were high between *Z. l. nuttalli* and all other subspecies except *Z. l. pugetensis*, and were low between the subspecies within the GOL



**Fig. 3.** Bayesian mitochondrial DNA phylogenetic analyses using individual haplotypes from all five *Zonotrichia* subspecies. The legend shows which haplotypes are present in *Z. atricapilla* and the *Z. leucophrys* subspecies.

clade (Table 1). Pairwise  $F_{ST}$  values between the three major clades were all high, 0.30 and 0.27 between *Z. atricapilla* and NP and GOL, respectively, and at 0.12 between NP and GOL.

## 4. Discussion

### 4.1. Mitonuclear discordance in crowned sparrows

Our study is the first to investigate the phylogenetic relationships among the five *Zonotrichia* species using genome-wide SNP data. Contrary to inferences using mtDNA exclusively, which indicated paraphyly between white-crowned and golden-crowned sparrows (Figs. 2 and 3; Weckstein et al., 2001; McKereghan-Dares, 2008), we find *Z. atricapilla* and *Z. leucophrys* comprise well-supported, reciprocally-monophyletic sister species, with all five *Zonotrichia* species monophyletic (Fig. 4A and 4B). *Zonotrichia atricapilla* and *Z. leucophrys* form distinct clusters in ADMIXTURE and PCA (Figs. 6 and 7) with high  $F_{ST}$  values between *Z. atricapilla* and all *Z. leucophrys* subspecies (Tables 1). Our results suggest that *Z. atricapilla* and *Z. leucophrys* are more deeply diverged in their nDNA than in their mtDNA, and that their origins probably markedly predate the previously reported 50,000 years given the reciprocal monophyly of the species using genome wide SNP data. Consistent with Weckstein et al. (2001), our analyses imply that mitochondrial introgression resulting from hybridization is the primary cause of discordance between the two genomes for the *Z. atricapilla*-*Z. leucophrys* clade. This strong discordance between nuclear and mitochondrial DNA, alongside our ABBA/BABA, Treemix (Fig. 5D) and ADMIXTURE (Fig. 6A) results, generally rules out incomplete lineage sorting (Toews and Brelsford, 2012). We also found genetic structuring within *Z. leucophrys*, uncovering two reciprocally-monophyletic groups (Fig. 4A and 4B), which form distinct clusters in ADMIXTURE and PCAs (Figs. 6 and 7).

### 4.2. Causes of mitonuclear discordance

Because the vertebrate mitochondrial genome has an effective population size that is approximately one-quarter that of nDNA (Funk and Omland, 2003), we expect the rate of stochastic lineage sorting to be

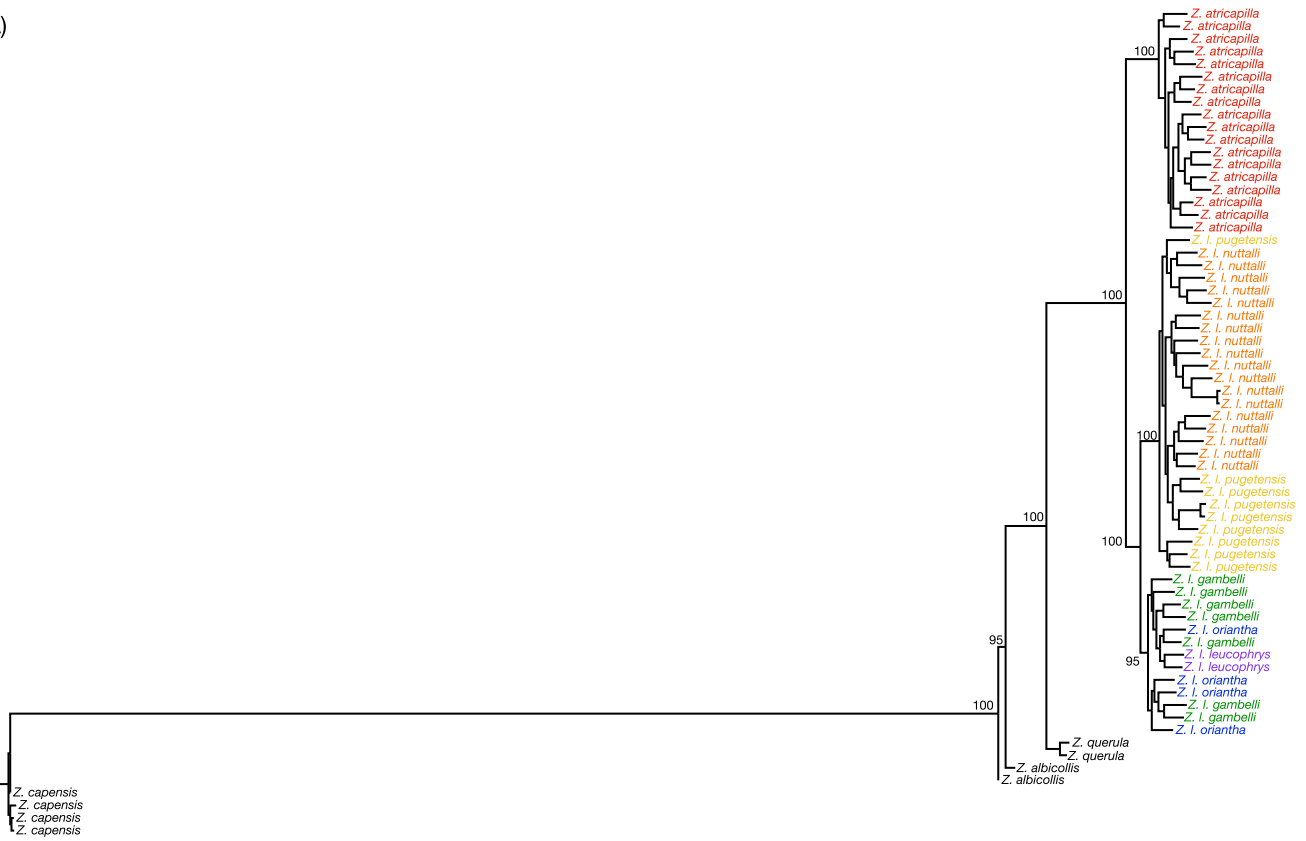
commensurably faster in mtDNA. We thus expect mtDNA to exhibit differentiation more quickly in good, reproductively-isolated species (Funk and Omland, 2003), although genome-wide nuclear datasets can reveal greater population structure than can be detected with a single locus such as mitochondrial DNA (e.g. Lavretsky et al., 2019). When hybridization and subsequent mitochondrial introgression has occurred, the opposite pattern can be evident. For example, if female hybrids backcross with males of their paternal lineage over multiple generations this could result in populations where individuals bear the paternal species' nuclear genome, and the maternal species' mtDNA (Funk and Omland, 2003; see Weckstein et al., 2001).

For taxa where the fitness of heterogametic hybrids is significantly reduced, the Z (or X) chromosome is largely thought to be responsible (Haldane's rule; Haldane, 1922; see Schilthuizen et al., 2011). This can promote speciation by preventing hybridization and fusion of incipient species ("large-Z effect"; Turelli and Orr, 1995), and thus limit opportunities for mtDNA to cross species boundaries via heterogametic hybrids (Rheindt and Edwards, 2011; Sloan et al., 2017). However, absolute inviability in hybrids of incipient bird species is relatively rare, and it generally takes a minimum of five million years for complete post-zygotic isolation in birds to evolve (Lijtmaer et al., 2003). As natural hybrids between *Z. atricapilla* and *Z. leucophrys* are rare, and as yet no hybridization experiments have been done, we lack information on their fitness.

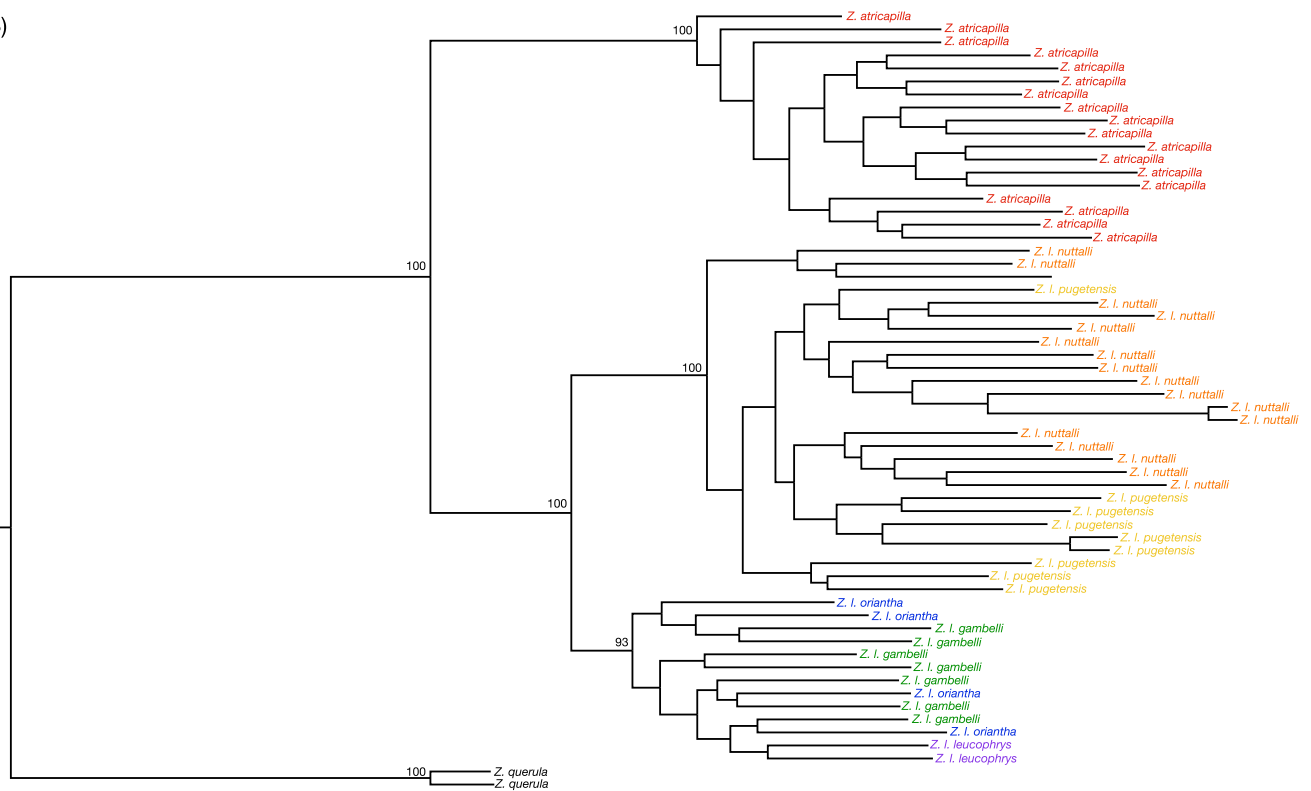
Weckstein et al. (2001) hypothesized that, at one point in their histories, hybridization took place between *Z. leucophrys* and *Z. atricapilla* with subsequent mitochondrial introgression from one into the other. Our Treemix analysis did indicate an introgression event from *Z. atricapilla* into *Z. l. leucophrys* when adding three migration events (Fig. 5D). Under this scenario, hybridization followed by continued backcrossing of hybrids would explain the discordance between the *Zonotrichia* mitochondrial gene and species trees (Zink et al., 1991; Weckstein et al., 2001). The originally reported divergence time estimate of 50,000 years thus reflects a period of historical hybridization that essentially "reset" the mitochondrial molecular clock for the two species (see Rheindt and Edwards, 2011).

Recent evidence suggests that mitochondrial introgression may be more common than previously appreciated, occurring across a wide

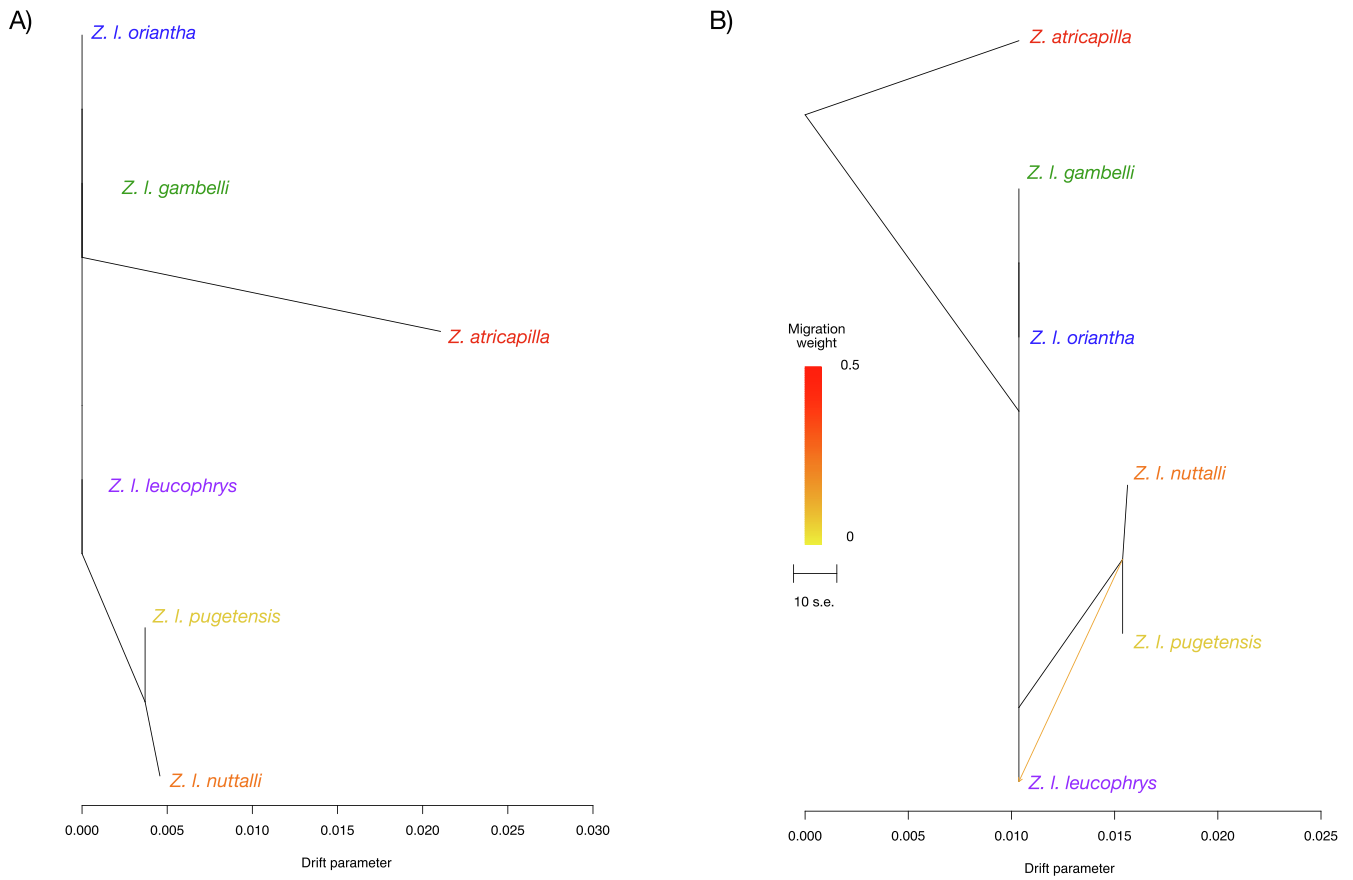
A)



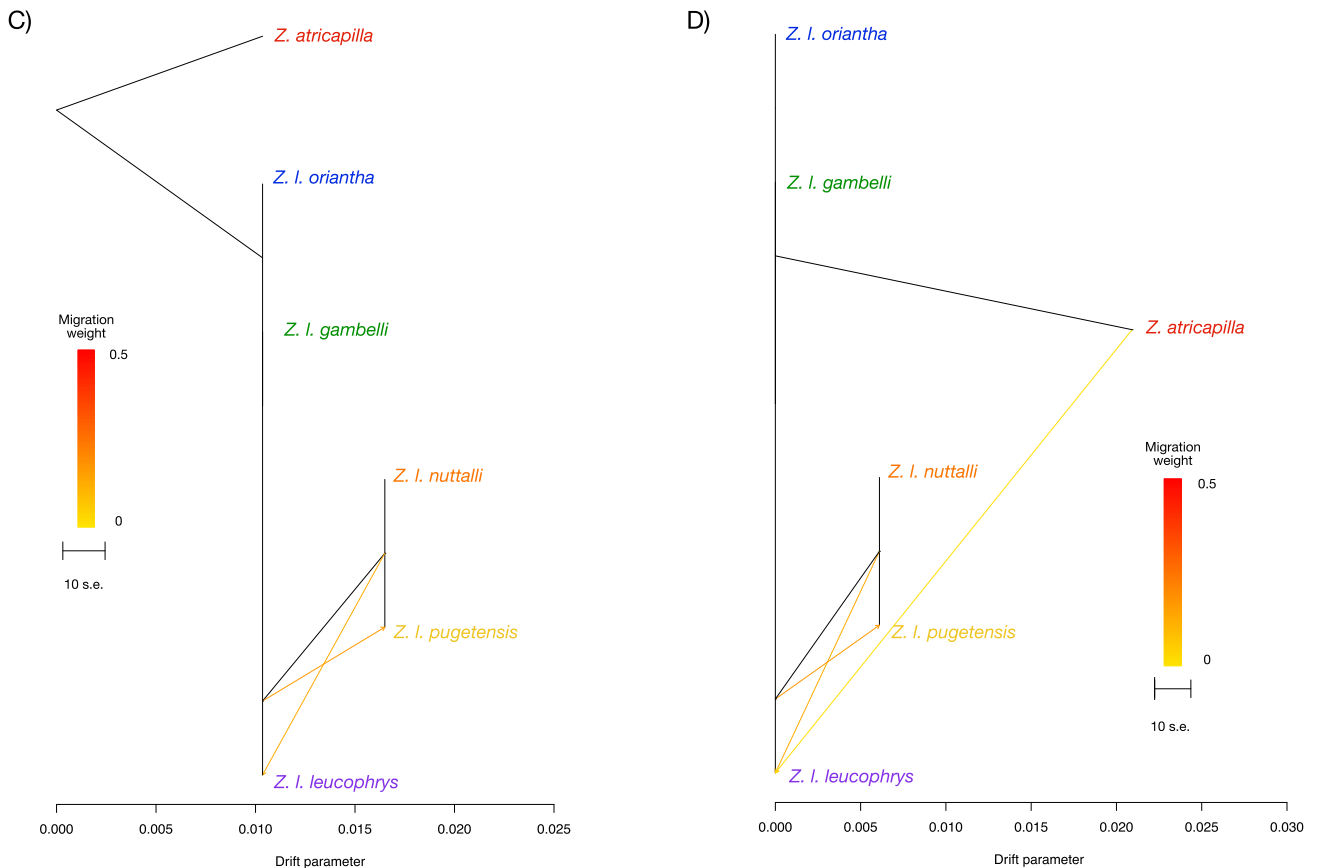
B)

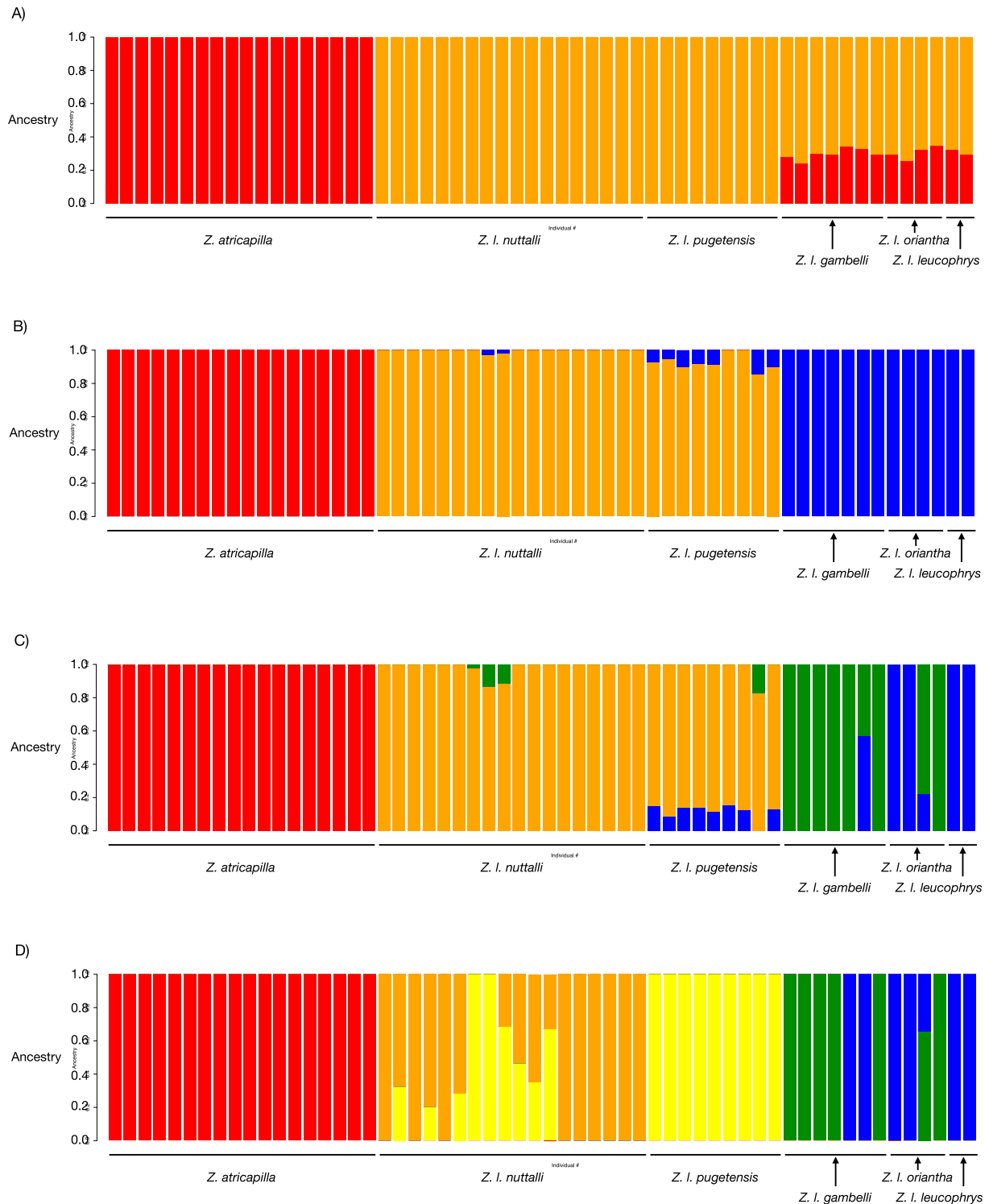


**Fig. 4.** Maximum-likelihood trees produced in RAxML with all outgroups (A) and with only one outgroup taxa to show relationships within *Z. leucophrys* more clearly (B). Node labels indicate bootstrap support values from 1000 bootstrap replicates; bootstrap values above 90% are shown. Name labels are coloured according to species and subspecies (red = *Z. atricapilla*; orange = *Z. l. nuttalli*; yellow = *Z. l. pugetensis*; green = *Z. l. gambelli*; purple = *Z. l. leucophrys*; blue = *Z. l. oriantha*; black = outgroup taxa). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** Maximum likelihood tree produced in Treemix (A), when inferring one migration event as suggested by OptM (B), when inferring two migration events (C) and three migration events (D). The bottom axis shows the drift parameter. Name labels are coloured according to species and subspecies (red = *Z. atricapilla*; orange = *Z. l. nuttalli*; yellow = *Z. l. pugetensis*; green = *Z. l. gambelli*; purple = *Z. l. leucophrys*; blue = *Z. l. oriantha*). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)





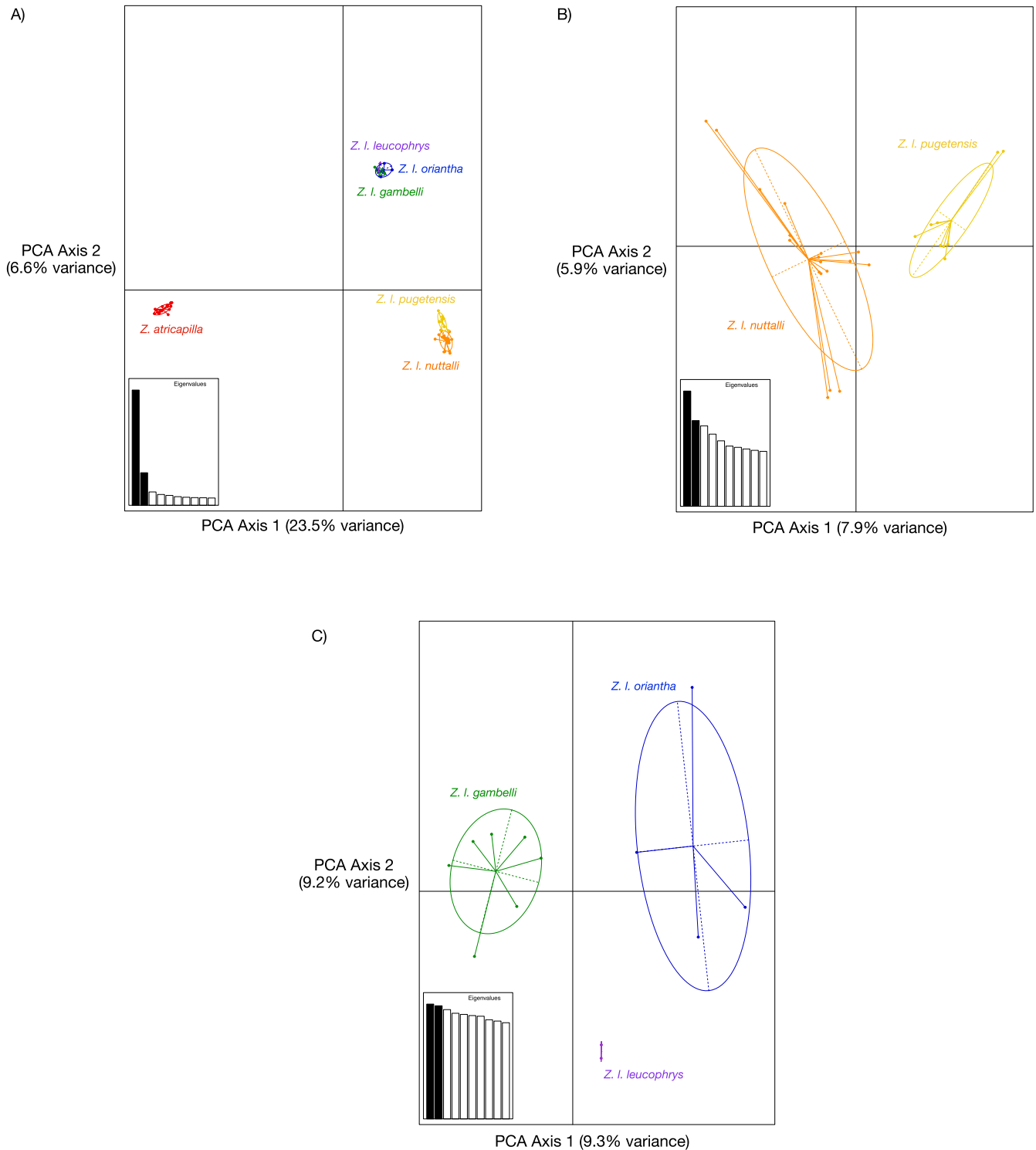
**Fig. 6.** Molecular assignments of *Z. atricapilla* and *Z. leucophrys* individuals using ADMIXTURE. (A) The plot for  $K = 2$  separating *Z. atricapilla* from *Z. leucophrys*, (B) The plot for  $K = 3$  separating *Z. atricapilla* as one cluster, *Z. l. nuttalli* and *Z. l. pugetensis* (NP) as another, and *Z. l. gambelli*, *Z. l. oriantha*, and *Z. l. leucophrys* (GOL) as the third, and (C) The plot for  $K = 4$  and (D)  $K = 5$  showing some structuring within the NP and GOL clades.

range of taxa (Toews and Brelsford, 2012). Several processes can facilitate mtDNA introgression once hybridization has occurred. Neutral processes, such as small population sizes and genetic drift (for example during bottlenecks or founder events), or large disparities in population sizes upon secondary contact, can facilitate mitochondrial introgression (Ballard and Whitlock, 2004; Toews and Brelsford, 2012). For example, mitochondrial introgression between chipmunk species in the genus

*Tamias* is thought to have occurred due to low population densities during climate induced range shifts (Good et al., 2008), and differential population sizes may have promoted mitochondrial introgression in *Neodiprion* sawflies (Linnen and Farrell, 2007).

Mitochondrial introgression can also be caused by selection (i.e. adaptive introgression; Toews and Brelsford, 2012). Increasing evidence highlights the importance of selection on the mitochondrial genome,





**Fig. 7.** Results of a principal component analysis of *Z. atricapilla* and the five *Z. leucophrys* subspecies (A), of the NP clade (B) and the GOL clade (C). Ellipses and colours are encoded as follows: red = *Z. atricapilla*; orange = *Z. l. nuttalli*; yellow = *Z. l. pugetensis*; green = *Z. l. gambelli*; purple = *Z. l. leucophrys*; blue = *Z. l. oriantha*. Eigenvalues of the different PCA axes are shown inset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

contradicting the long-held assumption that much of the variation in mtDNA is neutral (Ballard and Whitlock, 2004; Toews and Brelsford, 2012). For example, selection on the mitochondrial genotype (or perhaps other cytoplasmic factors) is likely to have caused the introgression of mtDNA between *Drosophila teissieri* and the *D. yakuba*/*D. santomea* species pair (Bachtrog et al., 2006). Similarly, a mitochondrial selective sweep is thought to have caused mitonuclear discordance in

yellowhammers (*Emberiza citrinella*) and pine buntings (*E. leucocephalos*; Irwin et al., 2009). Examples of adaptive introgression often invoke thermal adaptation as an explanation, due to the key role of mtDNA in efficient energy production and metabolism (Ballard and Whitlock, 2004; Lamb et al., 2018). For example, mitochondrial introgression between *Lepus* hare species (Melo-Ferreira et al., 2005), and between arctic char and lake trout (*Salvelinus alpinus* and *S. namaycush*; Wilson

**Table 1**Pairwise  $F_{ST}$  values from analysis of nuclear SNP data for *Z. atricapilla* and all five *Z. leucophrys* subspecies.

	<i>Z. atricapilla</i>	<i>Z. l. nuttalli</i>	<i>Z. l. pugetensis</i>	<i>Z. l. gambelii</i>	<i>Z. l. oriantha</i>
<i>Z. l. nuttalli</i>	0.31				
<i>Z. l. pugetensis</i>	0.31	0.04			
<i>Z. l. gambelii</i>	0.27	0.13	0.13		
<i>Z. l. oriantha</i>	0.28	0.13	0.13	0.01	
<i>Z. l. leucophrys</i>	0.28	0.13	0.14	0.02	0.02

and Bernatchez, 1998), are potentially due to temperature adaptation. Mitochondrial variation in multiple Australian bird species is associated with historical climatic variation (Lamb et al., 2018).

It is unclear whether the mitochondrial introgression between *Z. atricapilla* and *Z. leucophrys* was caused by neutral or selective processes, or a combination thereof. The Pleistocene epoch was characterized by repeated cycles of global cooling and warming, with the most rapid oscillations occurring from 60,000 to 25,000 years before present (Adams, 1997). In the northern hemisphere, colder periods resulted in a series of extensive glacial advances. At their maximum extent Arctic ice sheets covered most of present-day Canada (Adams, 1997). While it is only informed speculation, populations of *Z. atricapilla* and *Z. leucophrys* could have been forced into common refugia by ice sheets and shifting vegetation during the Pleistocene, leading some individuals to seek out heterospecific mates. Several rounds of bottlenecks due to climate change might have occurred during this period and similarly caused the fixation of one haplotype in both species during periods when population sizes were small. Selection may have acted in combination with periods of small population sizes and augmented this process, if one mitochondrial haplotype conferred a fitness advantage.

If effective population sizes were small, the re-assortment of nuclear loci could have occurred quickly to erase most traces of hybridization events, although we do detect some signature of hybridization in our ABBA/BABA test, as well as in the Treemix (Fig. 5D) and ADMIXTURE (Fig. 6A) results. Premating isolation mechanisms may have also been well-developed to prevent the fusion of the species. Behavioural reproductive isolation through conspecific song recognition is well-documented in *Zonotrichia* species, with juveniles predisposed to learn their own species' songs (e.g. Shizuka, 2014).

The time required for biological speciation is taxon-dependent, with some species able to interbreed even after more than 20 million years of divergence (Lijtmaer et al., 2003; Coyne and Orr, 2004). Given that little gene flow is needed for mtDNA to move between species successfully (Takahata and Slatkin, 1984), and complete reproductive isolation can take millions of years to develop, mitochondrial introgression could be far more common in birds than previously realized. Literature surveys suggest that 14–16% of bird species exhibit mitochondrial paraphyly, and it is estimated that at least 5.7% of these cases are due to introgression (Funk and Omland, 2003; McKay and Zink, 2010).

#### 4.3. Implications of mitochondrial introgression

An increasing number of studies are revealing the importance of selection on the mitochondrial genome (Toews and Brelsford, 2012). Despite this, the causes of mitochondrial introgression are not well studied in natural systems. Future work should investigate the functional differences caused by different mitochondrial haplotypes (e.g. as done in yellow-rumped warblers: *Setophaga* spp.; Toews et al., 2014), to further our understanding of selection on the mitochondrial genome and its consequences for cytonuclear discordance. In addition, the co-adaptation of mitochondrial and nuclear genes is of key importance for fitness (Hill, 2017; Sloan et al., 2017). Cases such as the *Zonotrichia* sparrows, where one mitochondrial haplotype was completely replaced by another, could indicate that in some instances the cytonuclear interactions are weaker than the strength of selection on the mtDNA (Bachtrog et al., 2006), if selection was a driver of the mitochondrial

introgression event rather than purely neutral processes.

Regardless of whether mitochondrial introgression is caused by neutral or selective processes, there are implications for the use of mtDNA alone for inferring species histories. For example, the use of both nuclear and mitochondrial DNA in Asian box turtles (*Cuora* spp.) with complicated patterns of mitochondrial introgression resulted in the recognition of a single monophyletic taxon, which had been masked by complicated patterns of mitochondrial introgression, with strong implications for the choice of individuals to use in captive breeding programs (Spinks and Shaffer, 2007). In addition, there are obvious implications for taxonomy, with introgression confounding phylogenetic inference of species (e.g. crotophytid lizards; McGuire et al., 2007), or masking hidden genetic structuring as we have found both between *Z. atricapilla* and *Z. leucophrys*, and among *Z. leucophrys* subspecies.

#### 4.4. Relationships among the five *Zonotrichia* species and the *Z. leucophrys* subspecies

In both the mitochondrial and nuclear phylogenetic reconstructions (Figs. 3 and 4), *Z. capensis* is the most distinct and sister to all other *Zonotrichia* species. However, we see different patterns between the nuclear and mitochondrial phylogenies with *Z. querula* and *Z. albicollis* sister species in the mitochondrial DNA tree (Fig. 3), with *Z. albicollis* basal to *Z. querula* in the nuclear DNA phylogeny (Fig. 4). To resolve the relationship, it may be necessary to increase the sampling of these two species to increase the power of the analyses.

We found well-supported monophyletic clades within *Z. leucophrys* that divide the subspecies into two groups (GOL and NP; Figs. 4, 6, and 7). Further, we found support for genetic differences between *Z. l. nuttalli* and *Z. l. pugetensis* (Fig. 4A, 4B, and 7B), consistent with the differentiation found by Lipshutz et al. (2017), although we do not find them to be completely reciprocally monophyletic with one *Z. l. pugetensis* individual recovered within the *Z. l. nuttalli* clade (Fig. 4A and 4B). We likely find stronger structuring and less evidence of hybridization between *Z. l. nuttalli* and *Z. l. pugetensis* in our study because we did not use individuals from within or near to the contact zone (Fig. 1). Within the GOL clade, however, the subspecies are not recovered as distinct (Fig. 4A and 4B) with low  $F_{ST}$  values among three subspecies (Table 1). We do find that PCA separated them (Fig. 7C) suggesting some weak differentiation.

## 5. Conclusions

Our findings support previous assertions of historical mitochondrial introgression between the two sister species, and further emphasize the need to re-evaluate estimates of age of divergence. For example, future studies with more data could re-estimate divergence times using the relative branch lengths within the *Zonotrichia* phylogeny, as the age of other *Zonotrichia* species (e.g. *Z. capensis*) are better established (Lougheed et al., 2013). More work should also focus on better delineating the genealogical patterns of the *Z. leucophrys* subgroups. These provide an important opportunity to study different stages of the speciation process *in situ*. Finally, we show a clear need to incorporate both nuclear and mitochondrial genomes when investigating avian evolutionary histories. For example, McKay and Zink (2010) found that in 21.3% of cases of paraphyly in birds, introgression could not be

reliably distinguished from incomplete lineage sorting (McKay and Zink, 2010). Mitochondrial DNA is effectively a single locus and lacks the resolution needed to discern close relationships and cannot reveal signatures of hybridization on its own. Genome-wide nDNA sequencing is more powerful, provides better resolution, and allows us to make stronger inferences about the histories of many taxa. Ultimately cases of mitonuclear discordance tell us something important about both the history of species (such as *Zonotrichia* sparrows) and about mitochondrial evolution. We must consider mtDNA as simply a ‘tool’ for phylogeography (Ballard and Whitlock, 2004).

## 6. Author statement

RST conducted analyses and contributed to the writing of the manuscript; ACB conducted lab work and contributed to the analysis and writing of the manuscript; NAC contributed to the writing of the manuscript; FB conducted some lab work and contributed to the writing of the manuscript and assisted with procuring samples; RC-C conducted lab work and assisted with writing of the manuscript. KD conducted lab work and contributed to the analysis and writing of the manuscript; SCL conceived the study, secured funding, procured some samples, did phylogenetic analysis, and contributed to the writing of the manuscript.

## CRediT authorship contribution statement

**Rebecca S. Taylor:** Formal analysis, Writing - review & editing, Visualization. **Ashley C. Bramwell:** Investigation, Writing - review & editing, Visualization. **Rute Clemente-Carvalho:** Investigation, Writing - review & editing. **Nicholas A. Cairns:** Writing - review & editing. **Frances Bonier:** Resources, Writing - review & editing. **Katherine Dares:** Investigation, Writing - review & editing. **Stephen C. Lougheed:** Conceptualization, Formal analysis, Resources, Writing - review & editing, Project administration, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We thank John Wingfield (University of California, Davis); Dai Shizuka (University of Nebraska); the Burke Museum of Natural History and Culture; the Royal Alberta Museum; the Royal Ontario Museum; the University of Alaska Museum; Donna Maney (Emory University); Beth MacDougall-Shackleton (Western University); Paul Martin (Queen’s University); and the Louisiana State University Museum of Natural History for contributing tissue samples. Brian Boyle (Genomic Analysis Platform, Laval University) and Génome Québec Innovation Centre performed the library preparation and Illumina sequencing, respectively. We also thank Brody Crosby for help with designing the graphical abstract. Suggestions from two anonymous reviewers improved the manuscript substantially. This research was part of two undergraduate theses (ACB and KD) at Queen’s University. This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant to S.C.L., the Baillie Family Chair in Conservation Biology (SCL), a NSERC Undergraduate Student Research Award to ACB, a NSERC Alexander Graham Bell Canada Graduate Scholarship to NAC, and an Ontario Trillium Scholarship to RST.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ympev.2021.107216>.

## References

- Adams, J.M., 1997. Global land environments since the last interglacial. Oak Ridge National Laboratory, TN, USA <http://www.esd.ornl.gov/err/qen/nerc.html>.
- Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664.
- Arnaiz-Villena, A., Álvarez-Tejado, M., Ruíz-del-Valle, V., García-de-la-Torre, C., Varela, P., Recio, M.J., Ferre, S., Martínez-Laso, J., 1998. Phylogeny and rapid northern and southern hemisphere speciation of goldfinches during the Miocene and Pliocene epochs. *Cell Mol. Life Sci.* 54, 1031–1041.
- Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* 18, 489–522.
- Avise, J.C., 2000. *Phylogeography: the history and formation of species*. Harvard University Press.
- Bachtrog, D., Thornton, K., Clark, A., Andolfatto, P., 2006. Extensive introgression of mitochondrial DNA relative to nuclear genes in the *Drosophila yakuba* species group. *Evolution* 60, 292–302.
- Baker, M.C., Spitzer-Nabors, K.J., Bradley, D.C., 1982. The response of female mountain white-crowned sparrows to songs from their natal dialect and an alien dialect. *Behav. Ecol. Sociobiol.* 10, 175–179.
- Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13, 729–744.
- Campagna, L., Benites, P., Lougheed, S.C., Lijtmaer, D.A., Di Giacomo, A.S., Eaton, M.D., Tubaro, P.L., 2012. Rapid phenotypic evolution during incipient speciation in a continental avian radiation. *Proc. R. Soc. B: Biol. Sci.* 279, 1847–1856.
- Campagna, L., Kopuchian, C., Tubaro, P.L., Lougheed, S.C., 2014. Secondary contact followed by gene flow between divergent mitochondrial lineages of a widespread Neotropical songbird (*Zonotrichia capensis*). *Biol. J. Linn. Soc.* 111, 863–868.
- Carson, R.J., Spicer, G.S., 2003. A phylogenetic analysis of the emberizid sparrows based on three mitochondrial genes. *Mol. Phylogenet. Evol.* 29, 43–57.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140.
- Chan, K.M., Levin, S.A., 2005. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* 59, 720–729.
- Clement, M., Posada, D., Crandall, K., 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9, 1657–1659.
- Cortopassi, A.J., Mewaldt, L.R., 1965. The circumannual distribution of white-crowned sparrows. *Bird-Banding* 36, 141–169.
- Coyne, J.A., Orr, H.A., 2004. *Speciation*. Sinauer Publishers, Sunderland, MA, USA.
- del Hoyo, J., Elliot, A., Christie, D.A., 2011. *Handbook of Birds of the World. Volume 16: Tanagers to New World Sparrows*. Lynx Edicions, Catalonia, Spain.
- Danner, J.E., Danner, R.M., Bonier, F., Martin, P., Small, T.W., Moore, I.T., 2011. Female, but not male, tropical sparrows respond more strongly to the local song dialect: implications for population divergence. *Am. Nat.* 178, 53–63.
- Edwards, S.V., Kingan, S.B., Calkins, J.D., Balakrishnan, C.N., Jennings, W.B., Swanson, W.J., Sorenson, M.D., 2005. Speciation in birds: genes, geography, and sexual selection. *PNAS* 102, 6550–6557.
- Edwards, S.V., Potter, S., Schmitt, C.J., Bragg, J.G., Moritz, C., 2016. Reticulation, divergence, and the phylogeography–phylogenetics continuum. *PNAS* 113, 8025–8032.
- Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and plophyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu. Rev. Ecol. Syst.* 34, 397–423.
- Good, J.M., Hird, S., Reid, N., Demboski, J.R., Stepan, S.J., Martin-Nims, T.R., Sullivan, J., 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks. *Mol. Ecol.* 17, 1313–1327.
- Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* 5, 184–186.
- Grant, P.R., 1999. *Ecology and Evolution of Darwin’s Finches*. Princeton University Press, Princeton, New Jersey.
- Haldane, J.B.S., 1922. Sex ratio and unisexual sterility in animal hybrids. *J. Genet.* 12, 101–109.
- Handford, P., 1980. Aspects of the annual cycle in highland populations of the rufous-collared sparrow, *Zonotrichia capensis*. *Auk* 97, 400–403.
- Handford, P., Nottebohm, F., 1976. Allozymic and morphological variation in population samples of rufous-collared sparrow, *Zonotrichia capensis*, in relation to vocal dialects. *Evolution* 30, 802–817.
- Harrison, R.G., Larson, E.L., 2014. Hybridization, introgression, and the nature of species boundaries. *J. Hered.* 105, 795–809.
- Hill, G.E., 2017. The mitonuclear compatibility species concept. *Auk* 134, 393–409.
- Houtman, A.M., Falls, J.B., 1994. Negative assortative mating in the white-throated sparrow, *Zonotrichia albicollis*: the role of mate choice and intra-sexual competition. *Anim. Behav.* 48, 377–383.
- Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.
- Irwin, D.E., Rubtsov, A.S., Panov, E.N., 2009. Mitochondrial introgression and replacement between yellowhammers (*Emberiza citrinella*) and pine buntings (*Emberiza leucocephalos*) (Aves: Passeriformes). *Biol. J. Linn. Soc.* 98, 422–438.
- Johnson, N.K., Cicero, C., 2004. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. *Evolution* 58, 1122–1130.
- Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405.
- Knaus, B.J., Grüwald, N.J., 2017. vcfr: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17, 44–53.

- Lamb, A.M., Gan, H.M., Greening, C., Joseph, L., Lee, Y.P., Moran-Ordóñez, Sunnucks, P., Pavlova, A., 2018. Climate-driven mitochondrial selection: A test in Australian songbirds. *Mol. Ecol.* 27, 898–918.
- Landys, M.M., Wingfield, J.C., Ramenofsky, M., 2004. Plasma corticosterone increases during migratory restlessness in captive white-crowned sparrow *Zonotrichia leucophrys gambelii*. *Horm. Behav.* 46, 574–581.
- Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, 72–773.
- Lavretsky, P., DaCosta, J.M., Sorenson, M.D., McCracken, K.G., Peters, J.L., 2019. ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. *Mol. Ecol.* 28, 2594–2609.
- Leache, A.D., Harris, R.B., Rannala, B., Yang, Z., 2013. The influence of gene flow on species tree estimation: a simulation study. *Syst. Biol.* 63, 17–30.
- Leigh, J.W., Bryant, D., 2015. PopART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6, 1110–1116.
- Linnen, C.R., Farrell, B.D., 2007. Mitonuclear discordance is caused by rampant mitochondrial introgression in *neodiprion* (hymenoptera: diprionidae) sawflies. *Evolution* 61, 1417–1438.
- Lipshutz, S.E., Overcast, I.A., Hickerson, M.J., Brumfield, R.T., Derryberry, E.P., 2017. Behavioral response to song and genetic divergence in two subspecies of white-crowned sparrows (*Zonotrichia leucophrys*). *Mol. Ecol.* 26, 3011–3027.
- 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, 2078–2079.
- Lijtmaer, D.A., Mahler, B., Tubaro, P.L., 2003. Hybridization and postzygotic isolation patterns in pigeons and doves. *Evolution* 57, 1411–1418.
- Lougheed, S.C., Campagna, L., Dávila, J.A., Tubaro, P.L., Lijtmaer, D.A., Handford, P., 2013. Continental phylogenography of an ecologically and morphologically diverse Neotropical songbird, *Zonotrichia capensis*. *BMC Evol. Biol.* 13, 58.
- Lougheed, S.C., Handford, P., 1992. Vocal dialects and the structure of geographic variation in morphological and allozymic characters in the Rufous-collared Sparrow, *Zonotrichia capensis*. *Evolution* 46, 1443–1456.
- MacDougall-Shackleton, E.A., MacDougall-Shackleton, S.A., 2001. Cultural and genetic evolution in mountain white-crowned sparrows: song dialects are associated with population structure. *Evolution* 55, 2568–2575.
- Malinsky, M., Matschiner, M., Svardal, H., 2020. Dsuite – Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* 21, 584–595.
- Mallet, J., Besansky, N., Hahn, M.W., 2016. How reticulated are species? *BioEssays* 38, 140–149.
- McCarthy, E.M., 2006. *Handbook of Avian Hybrids of the World*. Oxford University Press, New York, NY, USA.
- McFarlan, J.T., Bonen, A., Guglielmo, C.G., 2009. Seasonal upregulation of fatty acid transporters in flight muscles of migratory white-throated sparrows (*Zonotrichia albicollis*). *J. Exp. Biol.* 212, 2934–2940.
- McGuire, J.A., Linkem, C.W., Koo, M.S., Hutchison, D.W., Lappin, A.K., Orange, D.I., Lemos-Espinal, J., Riddle, B.R., Jaeger, J.R., 2007. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. *Evolution* 61, 2879–2897.
- McKay, B.D., Zink, R.M., 2010. The causes of mitochondrial DNA gene tree paraphyly in birds. *Mol. Phylogenet. Evol.* 54, 647–650.
- McKerghan-Dares, K., 2008. *Diversification and speciation of crowned-sparrows, genus Zonotrichia, in the New World (Undergraduate thesis, Queen's University, Kingston, Ontario)*. Retrieved from <http://hdl.handle.net/1974/1164>.
- Melo-Ferreira, J., Boursot, P., Suchentrunk, F., Ferrand, N., Alves, P.C., 2005. Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Mol. Ecol.* 14, 2459–2464.
- Merritt, J.R., Grogan, K.E., Zinzow-Kramer, W.M., Sun, D., Ortlund, E.A., Yi, S.V., Maney, D.L., 2020. A supergene-linked estrogen receptor drives alternative phenotypes in a polymorphic songbird. *PNAS* 117, 21673–21680.
- Miller, A.H., 1940. A hybrid between *Zonotrichia coronata* and *Zonotrichia leucophrys*. *Condor* 42, 45–48.
- Moore, I.T., Bonier, F., Winfield, J.C., 2005. Reproductive asynchrony and population divergence between two tropical bird populations. *Behav. Ecol.* 16, 755–762.
- Morton, M.L., Mewaldt, L.R., 1960. Further evidence of hybridization between *Zonotrichia atricapilla* and *Zonotrichia leucophrys*. *Condor* 62, 485–486.
- Pickrell, J.K., Pritchard, J.K., 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLOS Genet.* 8, e1002967.
- Pleszczyńska, W., 1979. Interspecific and intraspecific song discrimination by *Zonotrichia albicollis* and *Z. leucophrys*. *Can. J. Zool.* 58, 1433–1438.
- Poland, J.A., Brown, P.J., Sorrells, M.E., Jannink, J.-L., 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLOS ONE* 7, e32253.
- Price, E.R., McFarlan, J.T., Guglielmo, C.G., 2010. Preparing for migration? The effects of photoperiod and exercise on muscle oxidative enzymes, lipid transporters, and phospholipids in white-crowned sparrows. *Physiol. Biochem. Zool.* 83, 252–262.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904.
- Rheindt, F.E., Edwards, S.V., 2011. Genetic introgression: an integral but neglected component of speciation in birds. *Auk* 128, 620–632.
- Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542.
- RStudio Team., 2015. *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA. Available at <http://www.rstudio.com/>.
- Sambrook, J., Russell, D.W., 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, New York, USA.
- Schilthuizen, M.C., Giesbers, W.G., Beukeboom, L.W., 2011. Haldane's rule in the 21st century. *Hereditas* 107, 95–102.
- Shen, Y.-Y., Liang, L., Sun, Y.-B., Yue, B.-S., Yang, X.-J., Murphy, R.W., Zhang, Y.-P., 2010. A mitogenomic perspective on the ancient, rapid, radiation in the Galliformes with an emphasis on the Phasianidae. *BMC Evol. Biol.* 10, 132.
- Shizuka, D., 2014. Early song discrimination by nestling sparrows in the wild. *Anim. Behav.* 92, 19–24.
- Sloan, D.B., Havird, J.C., Sharbrough, J., 2017. The on-again, off-again relationship between mitochondrial genomes and species boundaries. *Mol. Ecol.* 26, 2212–2236.
- Spinks, P.Q., Shaffer, H.B., 2007. Conservation phylogenetics of the Asian box turtles (Geomydidae, Cuora): mitochondrial introgression, numts, and inferences from multiple nuclear loci. *Conserv. Genet.* 8, 641–657.
- Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313.
- Takahata, N., Slatkin, M., 1984. Mitochondrial gene flow. *PNAS* 81, 1764–1767.
- Toews, D.P.L., 2017. From song dialects to speciation in white-crowned sparrows. *Mol. Ecol.* 26, 2842–2844.
- Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* 21, 3907–3930.
- Toews, D.P.L., Mandic, M., Richards, J.G., Irwin, D.E., 2014. Migration, mitochondria, and the yellow-rumped warbler. *Evolution* 68, 241–1155.
- Tomback, D.F., Baker, M.C., 1984. Assortative mating by white-crowned sparrows at song dialect boundaries. *Anim. Behav.* 32, 465–469.
- Turelli, M., Orr, H.A., 1995. The dominance theory of Haldane's rule. *Genetics* 140, 389–402.
- Weckstein, J.D., Zink, R.M., Blackwell-Rago, R.C., Nelson, D.A., 2001. Anomalous variation in mitochondrial genomes of white-crowned (*Zonotrichia leucophrys*) and golden-crowned (*Z. atricapilla*) sparrows: pseudogenes, hybridization, or incomplete lineage sorting? *Auk* 118, 231–236.
- Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- Willis, S.C., Farias, I.P., Ort, G., 2014. Testing mitochondrial capture and deep coalescence in Amazonian cichlid fishes (Cichlidae: *Cichla*). *Evolution* 68, 256–268.
- Wilson, C.C., Bernatchez, L., 1998. The ghost of hybrids past: fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Mol. Ecol.* 7, 127–132.
- Zink, R.M., 1982. Patterns of genic and morphologic variation among sparrows in the genera *Zonotrichia*, *Melospiza*, *Junco*, and *Passerella*. *Auk* 99, 632–649.
- Zink, R.M., Dittmann, D.L., Rootes, W.L., 1991. Mitochondrial DNA variation and the phylogeny of *Zonotrichia*. *Auk* 108, 578–584.