The biogeographic and evolutionary history of an endemic clade of Middle American sparrows: *Melozone* and *Aimophila* (Aves: Passerellidae)

Luis Sandoval\(^a,\(^b,\)\(^*\), Kevin L. Epperly\(^c\), John Klicka\(^c\), Daniel J. Mennill\(^b\)

\(^*\)Escuela de Biología, Universidad de Costa Rica, CP-11501-2060, San Pedro, San José, Costa Rica
\(^b\)Department of Biological Sciences, University of Windsor, Windsor, ON, Canada
\(^c\)Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Box 353010, Seattle, WA 98195-3010, USA

**ABSTRACT**

The large number of endemic species in Middle America is frequently attributed to the interplay of geographical barriers and historical climatic changes in the region. This process promotes genetic divergence between populations, and given enough time, may yield new species. Animals that inhabit mid-elevation or highland habitats may be disproportionately affected in this way. Genetic analyses of animals in this region allow us to better understand how historical patterns of isolation have influenced the generation of new species in this biodiversity hotspot. We studied the biogeography and systematics of two closely related genera of sparrows (Passerellidae): *Melozone* and *Aimophila*. Collectively, this group is distributed from the southwestern United States and southward as far as central Costa Rica. We sampled 81 individuals of 8 *Melozone* and 2 *Aimophila* species, from 19 localities distributed throughout their ranges. We reconstructed phylogenetic relationships and time-calibrated species trees using multilocus sequence data comprised of one mitochondrial gene and five nuclear genes. We conducted an ancestral area reconstruction analysis to determine the probability of ancestral range at each divergent event. Despite analyzing six loci, we were unable to obtain a fully resolved phylogenetic tree. We recovered four main lineages: lineage 1 includes four *Melozone* species distributed north of Isthmus of Tehuantepec (*M. albicollis, M. crissalis, M. aberti, M. fusca*); lineage 2 includes three *Melozone* species distributed south of the Isthmus of Tehuantepec (*M. biarcuata, M. cabanisi, M. leucotes*); lineage 3 lineage consists of a single species endemic to the Pacific coast of Mexico (*M. kieneri*); and lineage 4 includes the more widely distributed sparrows in the genus *Aimophila*. Our analyses suggest that these genera probably originated during the late Miocene in the Madrean Highlands of southern Mexico. We identified dispersal as the prevalent cause of speciation in this clade with most lineages dispersing to their current distributions from southern Mexico either to the north following a developing and expanding Madro-Tertiary flora, or to the south across the Isthmus of Tehuantepec. A similar pattern of dispersal from this biogeographic region has been reported in other taxa including fishes, reptiles, and birds. Our results reveal that the four lineages identified represent geographically coherent and ecologically similar assemblages of taxa. Finally, when our genetic results are considered, along with apparent differences in morphology and song, the allopatric forms *M. b. cabanisi* and *M. l. occipitalis* warrant recognition as biological species.

Crown Copyright © 2017 Published by Elsevier Inc. All rights reserved.

1. Introduction

Middle America – defined here as a region that extends from northern Mexico through eastern Panama – is a biodiversity hot-spot characterized by high levels of endemism and species richness (Jiménez and Ornelas, 2015; Myers et al., 2000; Mittermeier et al., 2005). Multiple factors are presumed to have played a role in the generation of this biodiversity, including: climate changes, a varied topography, a high diversity of habitats, the immigration (and diversification) of South American taxa (Smith and Klicka, 2010), and autochthonous diversification (Haffer, 1987; Barrantes, 2009). Phylogenetic studies suggest that the autochthonous diversification of many taxa in Mesoamerica was due in part to the occurrence of geographical barriers created by regional mountain uplift, along with the expansion and contraction of forest during the cyclic climatic changes of the Pleistocene epoch (Haffer,
The combination of montane barriers and forest contractions resulted in the fragmentation of populations, promoting genetic divergence and phenotypic differentiation within these isolated populations. Mid-elevation and highland taxa were strongly affected by the expansion and contraction of montane forest habitats (Sullivan et al., 2000; Barrantes, 2009). During warming periods, montane forests presumably moved to increasingly higher elevations, effectively isolating populations of some organisms on peaks that were no longer connected by forests with nearby peaks. Alternatively, during periods of climate cooling, montane forests migrated downslope, creating potential opportunities for taxa of high elevation habitats to colonize new ranges and peaks.

Two bird genera with species distributed from mid-to-high elevations in Middle America are the *Melozone* and *Aimophila* sparrows (Stiles and Skutch, 1989; Howell and Webb, 1995; Rising, 1996, 2011). As a consequence of recent molecular phylogenetic studies, both genera have recently been revised. For example, the eight recognized species in the genus *Melozone* now includes four towhees that used to be in the genus *Pipilo* (DaCosta et al., 2009; Klicka et al., 2014; Sandoval et al., 2014). These same studies identified *Aimophila* as polyphyletic. Consequently, it now includes only three species, while eight former members of this genus were moved to *Peucaea*. As currently recognized, *Aimophila* and *Melozone* together represent a monophyletic group. However, the actual relationship between *Aimophila* and *Melozone* species is uncertain, because previous phylogenetic analyses report polytomies involving these genera (DaCosta et al., 2009; Klicka et al., 2014).

The taxonomic reconfiguration of *Aimophila* and *Melozone* has resulted in a well-circumscribed, endemic, Middle American bird assemblage. Distributions within the group (Fig. 1) range from southwestern desert in North American to Central Costa Rica. Birds of both genera are characterized as being mostly terrestrial, with year-round territorial pairs (Stiles and Skutch, 1989; Howell and Webb, 1995; Rising, 1996, 2011), and exhibiting sexually monomorphic plumage features (Rising, 2011; Sandoval and Mennill, 2013). *Melozone* species inhabit diverse habitats, including desert, xeric hillsides and canyons, dense brush and understory, semi-arid montane shrubland, dense thickets, coffee plantations, and secondary forest edges, from sea level to 2800 m (Stiles and Skutch, 1989; Howell and Webb, 1995; Rising, 1996, 2011). *Aimophila* species inhabit arid montane scrub, including sparse woodlands of oak, pine, and juniper, as well as chaparral, scrublands, and grasslands (Rising, 1996, 2011). At least some, if not all, species in both genera exhibit complex and particular vocalizations, often produced as vocal duets (e.g., Marshall, 1964; Sandoval et al., 2013, 2014).

The factors that have promoted the divergence of species within both of these genera are unknown. In this study, our first objective was to clarify relationships within this group by resolving the existing polytomies. With our phylogenetic hypothesis, our second objective was to examine the timing of divergences within these genera and investigate the correspondence between regional biogeographic events (i.e. earth history) and the generation of biodiversity within this assemblage of little-studied and taxonomically contentious songbirds.

2. Material and methods

2.1. Phylogenetic analysis

We sampled 81 individuals from 10 of the 11 recognized species in the genera *Melozone* and *Aimophila* (Chesser et al., 2010; see also Sandoval et al., 2014), from 19 localities distributed...
throughout ranges of the two genera (Fig. 1, Appendix A, Table A1). We were unable to obtain a sample for a single member of this group, *Aimophila notosticta*, a poorly known species restricted to the Mexican state of Oaxaca. Wherever possible, we used multiple samples for each taxon, so that we might identify previously unrecognized (i.e. cryptic) lineages. We used as a single outgroup taxon, Large-footed Finch (*Pezotetes capitalis*), because previous studies showed a close relationship with species in both ingroup genera, and ingroup monophyly was established previously (Klicka et al., 2014). We sequenced 1038 bp of a single mitochondrial DNA (mtDNA) locus, NADH dehydrogenase 2 (ND2), for 80 samples and the outgroup. For a subset of these samples we sequenced an additional five nuclear loci nuDNA: (1) myoglobin intron 2 (MB-I2); (2) beta fibrinogen intron 5 (FGB-I5); (3) the Z-linked aconitase 1, intron 9 (AC01-I9); (4) intron 3 and flanking exon regions of the Z-linked muscle-specific receptor tyrosine kinase gene (MUSK-I3); and (5) the Z-linked spindle 1 gene, intron 2 (SPIN1-I2). Our complete data set then was comprised of 4868 bp of sequence data obtained from one mtDNA locus and five additional nuDNA loci for 20 ingroup taxa and a single outgroup. Two *Aimophila rufescens* individuals from the same locality were required in order to assemble a full complement of nuDNA sequence for this taxon. Sequence data generated specifically for this study have been submitted to GenBank.

We extracted total genomic DNA using a DNeasy tissue extraction kit (Qiagen, Valencia, CA) following the manufacturer's protocol. Sequences were amplified via 12.5 μL polymerase chain reactions (PCR) under the following conditions: denaturation at 94 °C, followed by 35 cycles of 94 °C for 30 s; 54 °C (ND2), 60 °C (MB-I2, FGB-I5, AC01-I9), 56 °C (MUSK-I3), or 62 °C (SPIN1-I2) annealing temperature for 30 s; and 72 °C for 1 min. This was followed by a 10 min extension at 72 °C. PCR products were then sequenced at the High Throughput Genomics Center (University of Washington, Seattle, WA). We edited and manually aligned forward and reverse sequences for each individual using Sequencher software (v5.0; Gene Codes Corporation, Ann Arbor, MI). Heterozygous sites in nuclear loci were coded with the appropriate IUPAC ambiguity code.

We determined the best-fit models of evolution for each locus (ND2 - GTR+I + G; AC01-I9, FGB-I5, MUSK-I3 - HKY invariant sites; MB-I2, SPIN1-I2 - HKY) with MrModeltest software (v2.3; Nylander, 2008). We then estimated a phylogenetic tree using only the ND2 mtDNA gene sequence (N = 80 individuals) using Bayesian inference in BEAST software (v1.7.4; Drummond and Rambaut, 2007). Using this tree (Appendix A, Fig. A1), we selected a subset of samples (N = 20) which were representative of each species and any sub-grouping which was supported by a node with a posterior probability ≥ 0.95. We then estimated a maximum likelihood (ML) gene tree using only mtDNA data, a concatenated ML gene tree based on only our nuclear sequence data, and a Bayesian analysis using ‘BEAST’ software (v1.7.4; Drummond and Rambaut, 2007) using this taxon subset. After discarding the first 5 million generations (10%) of the Bayesian analysis as burn-in, the parameter values of the samples from the posterior distribution were summarized on the maximum clade credibility tree using TreeAnnotator software (v1.7.4; Drummond and Rambaut, 2007). All Bayesian trees were run for 50,000,000 generations and sampled every 5000 generations with a Yule process speciation tree prior and a relaxed uncorrelated lognormal clock. Convergence was assessed by examining effective sample size (ESS) values and likelihood plots in Tracer software (v1.5; Rambaut and Drummond, 2007). Parameter convergence was deemed sufficient when ESS reached 200 or above. Maximum likelihood trees were generated independently (mtDNA and nuDNA datasets) using RAxML software (v8.1.15; Stamatakis, 2014). Both used a GTRCAT model of evolution and for each, 1000 rapid bootstrap replicates were performed so that node support could be assessed.

All nuclear genes (bi-allelic) were phased with a probability threshold of 0.7 using PHASE software (v2.1.1), generating two haplotypes for each individual sampled (Stephens and Donnelly, 2003). For phylogeny reconstruction using the phased haplotypes we chose a coalescent approach, estimating a species tree for our 20 ingroup taxa, using ‘BEAST’ software (v1.7.4; Heled and Drummond, 2010) and all sequenced loci. The appropriate models of evolution for each locus were used, along with a Yule process tree prior and a piecewise linear and constant root model for our population size prior. A relaxed uncorrelated lognormal clock was applied to each locus using a rate of 1.25 × 10⁻² substitutions/site/Myr (2.5% change between lineages per Myr) for ND2 (Smith and Klicka, 2010), and a rate of 1.35 × 10⁻³ substitutions/site/Myr for autosomal (FGB-I5, MB-I2) or 1.45 × 10⁻³ substitutions/site/Myr for sex-linked (AC01-I9, MUSK-I3, SPIN1-I2) introns (Ellegren, 2007). The analysis was run, sampled, and analyzed as described above for prior BEAST analyses. In order to provide an alternative coalescent-based method of phylogeny reconstruction that may be less susceptible to potential error associated with recombination, we also generated a species tree using all six loci and the program SVD quartets (Chifman and Kubatko, 2014), as implemented in PAUP* v4.0a150 (preview release; Swofford, 2003). The program SVD quartets takes multi-locus, unlinked single-site data and infers the quartet trees for all subsets of four species. It then combines the set of all quartet trees into a species tree using a quartet amalgamation heuristic. We evaluated all possible quartets (N = 61,456) under a multispecies coalescent model with 500 bootstrap replicates. Trees were assembled using the quartet based phylogeny reconstruction algorithm, QFM (Reaz et al., 2014).

2.2. Biogeographical analysis

We conducted an ancestral area reconstruction analysis using our multilocus species tree and its 12 taxa/lineages. We used RASP software (v3.21; Yu et al., 2015) and chose a Bayesian binary Markov chain Monte Carlo (BMM) approach to determine the probability of each potential ancestral range at each divergence event. We chose this analysis because it allowed us to estimate the probability at a node of an ancestral area based on the average of a posterior set of trees and thereby control for phylogenetic uncertainty. Each taxon was assigned to one or more of the biogeographical regions proposed by Stotz et al. (1996), but splitting the Madrean Highlands into two regions, one to the north, and another to the south of the Isthmus of Tehuantepec (Mexican Highlands and Central American Highlands, respectively): (A) Pacific Arid Slope; (B) Mexican Highlands; (C) Gulf-Caribbean Slope; (D) Baja-Sonora; (E) Mexican Plateau; (F) Chiiriquí-Darién Highlands; and (G) Central American Highlands. The boundaries between these biogeographic regions were defined based on vegetation structure and physiographic features and each region represents an endemic area for birds and other taxa (Stotz et al., 1996). We set the number of potential ancestral areas occupied to seven (which are the number of biogeographical regions used to classify the taxon occurrence) and used the F81 + G model, the most complex model of area reconstruction available in RASP. We ran 500,000 cycles using 10 chains, sampling every 100 generations, and discarding the first 25%. We estimated the most probable cause of species divergence (dispersal or vicariance) with RASP, where dispersal assumes that the species moved from the ancestral area of distribution to a new one with a posterior divergence, and vicariance assumes that the species originated “in place” due to the “sundering” of an ancestral distribution by the appearance of a physical or biotic
barrier that interrupts gene flow. We used a multilocus species tree for this analysis, because such trees have a higher probability of capturing the genetic history of the study organism than single-locus or concatenated multilocus trees (Edwards and Beerli, 2000; Knowles and Maddison, 2002).

3. Results

3.1. Phylogenetic reconstruction

Our mtDNA ML analysis (Fig. 2), our ML analysis of five concatenated nuclear introns (Fig. 3), and our multilocus species tree analysis (Fig. 4) all suggested slightly differing phylogenetic patterns. In general, terminal nodes in the three analyses showed high support, but the basal nodes were unresolved. When poorly supported nodes were collapsed (not shown), our mtDNA tree (Fig. 2, see also Supp. Fig. A1) revealed a five-way polytomy, comprised of the following lineages: (1) Melozone kieneri, recovered here as sister to all other members of the ingroup; (2) the genus Aimophila; (3) a clade comprised of the additional four Melozone species that occur north of the Isthmus of Tehuantepec (M. aberti, M. albicollis, M. fusca, and M. crissalis; hereafter referred to as “northern Melozone”); (4) a M. biarcuata-cabanisi lineage; and (5) a lineage comprised of all M. leucotis samples. In this last lineage we uncovered significant phylogenetic diversity, suggesting that M. leucotis as currently recognized, contains two species-level taxa. Our resolution improved slightly with the addition of multiple nuclear genes. The ML tree in which only nuDNA was analyzed (Fig. 3), and our ‘BEAST species tree (Fig. 4), both yielded four-way polytomies. In each of these trees, the two lineages comprised of Melozone taxa distributed south of the Isthmus of Tehuantepec (biarcuata-cabanisi, and leucotis; hereafter “southern Melozone”) were united in a well-supported clade.

The relationships indicated for taxa within each of the four main lineages identified were, with one important exception, consistent and well-supported across analyses. Our mtDNA tree (Fig. 2) and previously published mtDNA phylogenies that included these taxa (e.g. Klicka et al., 2014), indicate that within the “northern Melozone”, M. fusca is most distantly related to the remaining three species in the assemblage (e.g. uncorrected mtDNA genetic distance between fusca and crissalis = 7.1%) and that M. aberti and M. crissalis are closely related sister taxa (uncorrected genetic distance = 2.2%). The nuclear data that we analyzed provide a very different perspective on the evolutionary history of this lineage. Our nuDNA ML tree indicates that the southern Mexican endemic form M. albicollis is instead the oldest extant taxon, and sister to the rest of the group. Melozone crissalis is reconstructed as polyphyletic, with an individual from Baja California sister (with high support) to our single M. aberti sample, while a crissalis individual from Santa Barbara California is sister (with high support) to our only M. fusca sample. A similar result was observed in our species tree.

Fig. 2. Mitochondrial gene tree of ND2 sequence data generated using maximum-likelihood. ML bootstrap values are reported above, or adjacent to each node. Bayesian posterior probabilities are shown below, or after likelihood values. Nodes that did not appear in the Bayesian tree are indicated with dashes. Country abbreviations (CR = Costa Rica; NI = Nicaragua; GT = Guatemala; MX = Mexico; US = United States; HN = Honduras) are followed by specimen collector numbers.
where *fusca*, *crissalis*, and *aberti* effectively create a three-way polytomy. Our multilocus species tree (Fig. 4) suggested a Late Miocene divergence (~6.5 mya) of the *Melozone-Aimophila* clade from their common ancestor. The four main lineages identified diverged soon after, probably within the next million years. We note that the error bars for all three of the earliest divergence events overlap at around 6 mya. Subsequent divergence events occurred periodically throughout the Pliocene and Pleistocene. The most recent of these involves the putative *aberti-crissalis-fusca* polytomy. Divergence among these three taxa appears to have occurred within the last 1–1.5 my.

Our SVD quartet analyses provided relatively little in the way of additional phylogenetic resolution. Like the other analyses that included the nuclear data, the same four main lineages were recovered, and all of the “southern *Melozone*” once again formed a well-supported clade (Appendix A, Fig. A2). The quartet analysis differed from the others in that it also suggested with moderate support (74% bootstrap) that *M. kieneri* is more closely related to the “southern *Melozone*” and *Aimophila* groups than it is to the “northern *Melozone*” assemblage. It also differs in that *aberti* and *crissalis* were recovered as sister taxa (as the mtDNA tree suggests) rather than as part of a polyphyletic *M. crissalis* (Fig. 3) o ra n *aberti-crissalis-fusca* polytomy. We note that when the same SVD quartet analysis is run without the mtDNA sequence included (five nuclear loci only, results not shown) the result is an *aberti-crissalis-fusca* polytomy.

### 3.2. Biogeographical reconstruction

Based on the multilocus species tree and our ancestral area reconstruction, the inferred ancestral area of distribution of the ancestor of the *Melozone-Aimophila* clade, was the Mexican Highlands (61.5% RASP probability; Fig. 5). The majority of the species in the four identified lineages diverged through dispersal from the inferred ancestral area of distribution to the actual occupation areas, with our analysis suggesting 15 different possible dispersal events overall, compared with only two occasions where vicariance may have played a role (Fig. 5). Of the ten nodes in our species tree (Fig. 5), an inferred dispersal event is responsible for subsequent diversification at six nodes, neither dispersal or vicariance were required for two nodes, and either dispersal or vicariance may explain diversification at the remaining two nodes.

Our analysis is ambiguous concerning the specific geographic region occupied by the ancestor of the “southern *Melozone*”–*Aimophila* clade (Mexican and Central American Highlands, 24%; South American Highlands, 22%, Mexican Highlands, 21%). In terms of the number of Isthmus of Tehauantepec crossings required to explain the observed distributions, it is most parsimonious to assume the *Aimophila* ancestor was restricted to southern Mexico, with diversification of the clade originating there. Under this scenario, *A. rufescens*, the most widely distributed of the species under study, diverged initially, near the end of the Pliocene, subsequently
dispersing widely to the north (northern Mexican Highlands, Pacific Slope, Caribbean Gulf) and south across the Isthmus (perhaps at a much later date) to the southern limits of its current range in Nicaragua. *Aimophila ruficeps* and *A. notosticta* (phylogenetic position not shown, but see DaCosta et al., 2009) diverged some time later, when *A. ruficeps* dispersed northward, eventually occupying the Baja-Sonora Plateau and portions of the Mexican and Colorado plateaus. *Aimophila notosticta* probably remained in place and today occurs as a narrowly distributed Mexican endemic, only in the arid montane scrub habitats of Oaxaca.

The RASP analysis suggests a second dispersal across the Isthmus of Tehuantepec and into the Central American Highlands by the ancestor of the “southern Melozone” clade as early as the late Miocene. A later dispersal southward from those highlands to the Chiriquí-Darién Highlands led to differentiation between the occipitalis and leucotis/nigror groups of *M. leucotis*. Speciation within the cabanisi-biarcuata lineage likely occurred in a similar fashion, although a vicariant event involving a more widespread ancestor cannot be ruled out. Biogeography of the “northern Melozone” clade is relatively less straightforward. Like the original *Aimophila-Melozone* ancestor, the ancestor of this group was also distributed in the Mexican Highlands (~57% probability). According to our species tree, an early dispersal event is indicated, when *M. kieneri* diverged early in the Pliocene by moving north and west (well before the other members of its clade) to occupy the Pacific Arid Slope. Near the end of the Pliocene, *M. albicollis*, a species now restricted to arid high elevation habitats in extreme southern Mexico, diverged from the ancestor of the aberti-crissalis-fusca assemblage. Our analysis indicates a possible vicariance event involving a widespread ancestor of all of these taxa, then distributed across the Mexican Highlands and the Baja-Sonora Plateau. This event would have disrupted gene flow between the south and north, yielding the species *M. albicollis* at the southern terminus, and the aberti-crissalis-fusca ancestor in the north. Alternatively, the aberti-crissalis-fusca ancestor dispersed to the north between one and three million years ago to occupy the Baja-Sonoran Plateau, a
region that includes what is now Baja California and much of the U. S. Desert Southwest. Additional dispersal events within the last million years then led to diversification within this group, yielding the recently derived species *M. aberti*, *M. crissalis*, and *M. fusca*.

### 4. Discussion

#### 4.1. Phylogenetic summary

Previous research on this clade (e.g. DaCosta et al., 2009; Klicka et al., 2014) has failed to resolve relationships near the base of the tree. In this study, despite nearly complete taxon sampling, and the inclusion of several additional nuclear markers, we too failed to resolve the basal nodes. Within the *Aimophila-Melozone* clade, we identified four main lineages; a “northern Melozone” clade, a southern “Melozone” clade, a clade comprised of all three members of *Aimophila*, and the *M. kieneri* lineage; however, precisely how these clades and lineages are related to one another remains unknown.

The repeated, yet unsuccessful, attempts to solve this problem suggest that the origin of these four lineages may involve a hard polytomy. That is, the ancestor of the entire clade split in the late Miocene and each daughter lineage produced by this event diverged again soon after. The short interval between the ancestor and daughter lineage divergences allowed little time for the shared
characters (i.e. synapomorphies) that would allow these basal relationships to be resolved, to accrue. The short internode branch lengths recovered at the base of our tree and the four-way polytomy observed are due, then, to a relative lack of available phylogenetic information that is a consequence of the near contemporaneous diversification of the clade’s basal lineages.

In general, gene trees derived from mtDNA data are largely consistent with species trees reconstructed using multiple nuclear loci. In our study, most nodes identified the same relationships across these different types of analyses. The addition of nuDNA did provide increased resolution at a single node, unifying the two main “southern Melozone” lineages into a well-supported clade, something our ND2 gene tree was unable to do. A more significant gene tree – species tree difference involved the placement of *M. fusca*. This (Fig. 2), and earlier mtDNA studies (Zink et al., 1998; DaCosta et al., 2009; Klicka et al., 2014), indicate that *M. fusca* diverged early in the history of the “northern Melozone” clade. Pairwise genetic distances and the ND2 “clock” rate (2.5%/my) suggest that this happened near the onset of the Pleistocene. In contrast, the nuDNA ML tree (Fig. 3) places *M. fusca* (with strong support) as sister to one of the two *M. crissalis* samples in our study from which it diverged some time in the last 0.5 my. A discrepancy between the phylogenetic signal of the mitochondrial genes and that of nuclear loci, such as we seem to have here, is referred to as a mito-nuclear discordance. Several causes of discordance have been suggested but a full accounting of these is beyond the scope of this paper (see Toews and Brelsford, 2012). We suggest, however, that the simplest explanation in our case allows both our mtDNA and nuDNA ML trees to be “correct.” The early divergence (~2.5 mya) of *M. fusca* suggested by the mtDNA tree, may in fact reflect the true branching history within this group. If so, then the nuDNA placement within the recently diverged *crissalis-fusca-aberti* cluster, may reflect relatively high levels of inter-specific gene flow currently, or in the recent past. The testing of this hypothesis will likely require high throughput sequencing and multiple samples from throughout the distributions of all three of this species.

Although our analyses do not clearly elucidate the relationship between the four identified lineages, three of them represent geographically coherent, morphologically, and ecologically similar assemblages of taxa. The *Aimophila* species are widespread, distributed from western and central United States to northern Costa Rica where they inhabit mainly patchy grassland habitats with rocks and shrubs (Rising, 1996, 2011). They differ from sparrows of the Melozone group in that they possess plumage characteristics (streaky brown backs, lighter ventral surface, rusty crown patches, dark malar stripes) typical of those seen in other more familiar North American sparrow species. In contrast, *Melozone* plumage characters are decidedly unlike those of *Aimophila* and most temperate zone forms. The “northern Melozone” birds are uniformly a dull brown or grayish-brown in color with subtle rufous markings (crissum, crown, or throat) or black markings (throat or face). The “southern Melozone” birds are even more distinctive, with bold markings, and brighter colors more similar to those of tropical sparrow taxa (e.g. *Chlorospingus, Atlapetes, and Arremon*). Species in this lineage also seem to prefer more mesic habitats than their northern counterparts, occurring in montane thickets, young secondary forests, and shade coffee plantations (Stiles and Skutch, 1989; Howell and Webb, 1995). Based on morphological and ecological traits, our last “lineage,” *M. kieneri*, is difficult to place. Like the other forms occurring north of the Isthmus of Tehuantepec it prefers drier habitats, inhabiting deciduous forest and dry forest (Howell and Webb, 1995; Sandoval and Mennill, 2014) along Mexico’s Pacific coast. In terms of plumage however, the bold and contrasting colors displayed suggest a closer relationship with suggest a closer relationship with the “southern Melozone.”

4.2. Biogeography in Middle America

All members of these two genera shared a common ancestor during the late Miocene in the Madrean (i.e. Mexican) Highlands, probably in southern Mexico. This region is a center of diversification in other taxa as well, including cichlid fish (Ričan et al., 2013), reptiles (Zaldívar-Riverón et al., 2005; Bryson et al., 2011), and other birds (Weir et al., 2008; Ornelas et al., 2014). During the late Miocene and early Pliocene, the flora that dominated the Madrean Highlands (Madro-Tertiary flora) began the expansion northward and to lowlands of the Pacific coast of Mexico and the United States (Axelrod, 1950; Davis, 1959), providing a northerly dispersal route for ancestral *Aimophila* and “northern Melozone” taxa. The earliest diversifications within these lineages occurred between 3 and 5 mya, suggesting dispersal northward was ongoing process by that time. The dry, warm climate that favored the earlier expansion and diversification of the Madro-Tertiary flora deteriorated at the start of the Pleistocene, when conditions during glacial advances became cooler and drier in highland habitats, and cooler and wetter in southwestern North America. This climate change led to a reduction of the Madro-Tertiary flora and presumably to the fragmentation of these habitats into smaller regions that were isolated from one another. As the “northern Melozone” lineage dispersed northward these isolated “islands” of suitable habitat may have been independently colonized, leading to the observed divergence between *M. aberti* and *M. crissalis* (and perhaps *M. fusca*) during the last million years.

The Isthmus of Tehuantepec in southeastern Mexico is a well-known geographical barrier for a variety of highland taxa (e.g. Parker et al., 1996; Peterson et al., 1999). It is comprised of a valley approximately 250 m above sea level and some 200 km wide at its narrowest point (Barrier et al., 1998) and is bounded on three sides by three mountain chains: the Sierra Madre Oriental and Sierra Madre del Sur to the northwest and the Chiapas-Guatemalan highlands to the southeast. Numerous montane taxa that occur on either side of the Isthmus are currently isolated from one another by this intervening valley. A comparative study of highland rodents spanning the Isthmus detected phylogeographic patterns suggesting an Isthmus-related vicariant event in multiple species (Sullivan et al., 2000). Two species of co-distributed snakes (*Atropoides, Bothriechis*; Castoe et al., 2009) and a number of birds including Common Bush-tanager (*Chlorospingus ophthalmicus*; Weir et al., 2007), Emerald Toucanet (*Aulacorhynchus prasinus*; Puebla-Olivares et al., 2008), and Chestnut-capped Brush Finch (*Arremon brunneinucha*; Cadena et al., 2007) showed similar phylogenetic breaks at the Isthmus. A recent molecular study showed that 8 of the 10 additional cloud forest bird species surveyed had diverged recently, across this zone (Barber and Klicka, 2010).

We have identified a minimum of two successful crossings of the Isthmus of Tehuantepec, one of which was followed by multiple speciation events. The first involves *A. rufescens*, the only member of its genus to occur south of Mexico. Whether the Isthmus provides a barrier for this species is unknown. *Aimophila rufescens* populations on either side of the Isthmus have not diverged morphologically and are recognized as conspecifics, suggesting that gene flow across the Isthmus is frequent enough to prevent divergence from occurring. Alternatively, if gene flow is impeded and diversification is occurring, as it is for many highland taxa in this region, we would infer that the Isthmus crossing likely occurred in the very recent history of the lineage. Unfortunately, the proper phylogeographic analysis that would allow testing of this idea has not yet been done.

The other crossing of the Isthmus of Tehuantepec documented in our study involves the “southern Melozone.” The origin of this lineage traces to a late Miocene or early Pliocene Isthmus crossing by the clade’s ancestor. Subsequent diversification occurred
around 5 mya in the Central American Highlands with each daughter lineage dispersing in a southerly direction, into northern Nicaragua (Matagalpa mountains), following the extension of this mountain range south (Ferrez Weinberg, 1992; Marshall and Lieberr, 2000). The divergence of M. biarcuata and M. cabanisi and the two reciprocally monophyletic lineages of M. leucotis occurred after the formation of the Nicaragua depression and the northern part of Costa Rica's central valley (Ferrez Weinberg, 1992). Due to the low altitude and the distance between mountains (ca. 550 km), the Nicaraguan Depression acted as a barrier to dispersal for several mid-elevation and highland species including arthropods (Duennes et al., 2012), mammals (Carleton et al., 2002; Hardy et al., 2013), and birds (Puebla-Olives et al., 2008; Weir et al., 2008; Zamudio-Beltrán and Hernández-Baños, 2015; Ortiz-Ramírez et al., 2016). The pattern reported for M. biarcuata and M. cabanisi in this paper is consistent with an inability to disperse across this barrier under recent environmental conditions. However, other species of mid-elevations and highlands apparently did move across this barrier during the early Pleistocene. The cooler and drier conditions that prevailed during glacial advances caused highland vegetation to descend to lower elevations, allowing access to previously unoccupied highland areas (Haffer, 1974, 1987; Barrantes, 2009) that were subsequently colonized. For example, sampled populations of Arremon brunninucha did not differ between Costa Rica and the mountains of Nicaragua (including Matagalpa), but these populations were differentiated from those in the north of Central America highlands (Cadena et al., 2007). This pattern is the same as the one observed for M. l. nigrior/leucotis and M. l. occipitalis in this study.

4.3. Taxonomic implications

This study revealed three distinct Melozone lineages. Despite our best efforts, we were unable to determine how they are related to one another or to the remaining group within the clade, Aimophila. Our work does show that despite uncertain relationships, none of these lineages are closely related. Each is not only well defined genetically, but also in terms of geography, ecology, and morphology. Melozone as currently recognized would be better represented by dividing it into two genera (minimally). The taxa in the “southern Melozone” retains that name, while the “northern Melozone” group is moved with M. kieneri into the resurrected genus Pyrgisoma (see DaCosta et al., 2009). This arrangement is already used in alternative taxonomies (e.g. Rising, 2011). Presently, the American Ornithologists’ Union classification committee recognizes only two species in the “southern Melozone” clade (AOU, 1998 and later supplements): M. biarcuata and M. leucotis. Our results show that each of these lineages is comprised of two geographically isolated, monophyletic groups that have long, independent evolutionary histories. Earlier work on M. biarcuata (Sandoval et al., 2014) showed that morphometric, plumage, and vocal analyses all supported the recognition of the two groups recovered here as biological species. The first includes the forms M. biarcuata biarcuata and M. b. hartwegi (unsampled in this study) of the northern Central American Highlands and the other is M. b. cabanisi, the form isolated in central Costa Rica. The M. leucotis lineage should be similarly divided. The taxon M. leucotis occipitalis occurs from southern Mexico to El Salvador. It is well differentiated genetically from the isolated forms that occur in Nicaragua and central Costa Rica, M. l. leucotis and M. l. nigrior, and also in morphology (Rising, 2011) and vocalizations (Sandoval, unpublished data). The recognition of M. leucotis and M. occipitalis as biological species on morphological grounds alone is not unprecedented (see Ridgway, 1901; Brodkorb, 1938) and our study provides the genetic and temporal perspective that was lacking previously.

Acknowledgments

LS was supported by scholarships and grants from the Ministerio de Ciencia y Tecnología (MICIT) and the Consejo Nacional para Investigaciones Científicas y Tecnológicas (CONICIT) of Costa Rica, the Government of Ontario, and the University of Windsor. Additional support was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Government of Ontario, and the University of Windsor (to DJ,M). This work was also funded in part by NSF DEB 0315469 (to J.K.). Specimens required to complete this study were provided by the Field Museum of Natural History, the Louisiana State University Museum of Natural Science, the Museum of Vertebrate Zoology at Berkeley, and the University of Washington Burke Museum. Many thanks to the Curators and Collection Managers at those institutions for their continued support and assistance. The careful and thoughtful comments provided by an anonymous reviewer of an earlier draft of this paper are much appreciated.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2017.03.008.

References


