Journal of Heredity 2008:99(5):453–463
doi:10.1093/jhered/esn025
Advance Access publication May 8, 2008

Phylogeography of a Widespread North American Migratory Songbird (*Setophaga ruticilla*)

GABRIEL J. COLBECK, H. LISLE GIBBS, PETER P. MARRA, KEITH HOBSON, AND MICHAEL S. WEBSTER

From the School of Biological Sciences, Washington State University, Pullman, WA 99164 (Colbeck and Webster); the Department of Evolution, Ecology and Organismal Biology, The Ohio State University, 370 Aronoff Laboratory, Columbus, OH 43210 (Gibbs); the Smithsonian Migratory Bird Center, National Zoological Park, Washington, DC 2008 (Marra); Canadian Wildlife Service, Saskatoon, SK, Canada, S7N 0X4 (Hobson).

Address correspondence to Gabriel J. Colbeck at the address above, or e-mail: gcolbeck@mail.wsu.edu.

Abstract

Genetic analyses for many widespread North American species have revealed significant east–west differentiation, indicating that many survived through the Pleistocene in 2 glacial refugia—1 in the eastern and 1 in the western part of the continent. It remains unclear, however, whether other areas may have served as important glacial refugia. Moreover, many such species exhibit widespread genetic similarity within eastern and western regions because of recent expansion from small refugial populations, making it difficult to evaluate current-day levels of gene flow. In this study, we used mitochondrial DNA (mtDNA) control region sequence and amplified fragment length polymorphism markers to survey genetic variation in a widespread migratory bird, the American redstart (*Setophaga ruticilla*). mtDNA analyses revealed a pattern that contrasts with that found for most other widespread species studied to date: most redstart populations across North America appear to have spread out from a single glacial refugium, possibly located in the southeastern United States, whereas populations in far-eastern Canada may have survived in a second glacial refugium located on the now-submerged Atlantic coastal shelf off the coast of Newfoundland. A pattern of isolation by distance in mtDNA suggested some constraints on current-day gene flow among extant redstart populations. This study thus reveals a recent evolutionary history for this species that differs from that of most other widespread North American passerines and provides evidence for limited gene flow in a species with potentially large dispersal distances.

Pleistocene glacial cycles and the climate changes that they precipitated have had profound impacts on patterns of genetic diversity within extant species (Avise and Walker 1998; Hofreiter et al. 2004). Changes in climate have likely had diverse effects on different populations, resulting in range shifts as different geographic areas become more or less suitable and isolation between populations when barriers to dispersal, such as glaciers, arise. The extent to which different species were affected in similar ways by these past climatic events remains unclear. At one end of the spectrum, coexisting extant species may have been affected in similar ways and therefore share common recent histories. Alternatively, different species may have been affected in different ways by Pleistocene events, such that their recent histories are divergent despite overlapping distributions in the present.

To understand the effects of Pleistocene climate changes, it is necessary to address 2 interrelated issues. First, it is necessary to determine the number of refugia that may have harbored populations of extant species. One possibility is that populations survived through Pleistocene glacial cycles without being subdivided, subsequently expanding from that location to colonize the continent as glaciers receded. This "single-refugium" hypothesis has been supported for some bird species with relatively narrow current ranges (e.g., Mila et al. 2000; Veit et al. 2005; Davis et al. 2006) and also for some widespread species showing limited migration (Zink 1996). In contrast, most widespread highly migratory species (i.e., species that currently cover much of North America) exhibit patterns of significant genetic differentiation between populations on either side of a geographic barrier (i.e., mountains, deserts, plains, bodies of water). These patterns indicate the existence of 2 or more major glacial refugia during Pleistocene glacial cycles (the multiple-refugia hypothesis; see References).

Second, in addition to knowing the number of glacial refugia involved, it is necessary to infer the locations of those refugia. Most studies of widespread North American

species have yielded a pattern of significant genetic differentiation between populations on either side of the Rocky Mountains (or Great Plains) with little genetic differentiation within these phylogroups (Table 1). This pattern suggests the existence of 2 major glacial refugia during the Pleistocene—1 in the east and 1 in the west. East–west differentiation is particularly common among species of widespread, migratory birds (Milot et al. 2000; Kimura et al. 2002; Ruegg and Smith 2002; Lovette et al. 2004; Peters et al. 2005), including all warblers with continental distributions studied to date (reviewed in Kelly and Hutto 2005), and has also been found in other taxa (McGowan et al. 1999; Rueness et al. 2003; Ayoub and Riechert 2004; Runck and Cook 2005).

Although the exact locations of these eastern and western refugia are often unclear, some studies have suggested more specific locations. For example, studies of plants (Boys et al. 2005), insects (Berlocher and Dixon 2004), fish (Bernatchez 1997), mammals (Paetkau and Strobeck 1996; Kyle and Strobeck 2003), and frogs (Lee-Yaw et al. 2008) have suggested the existence of a glacial refugium on the now-submerged Atlantic coastal shelf near Newfoundland (Figure 1). Some studies of birds have similarly suggested a glacial refugium in the maritime region of far-eastern Canada (Gill et al. 1993; Holder et al. 1999; Boulet et al. 2005), but the possibility that this was an important refugium for migratory birds is relatively unexplored (but see Zink and Dittmann 1993; Boulet and Gibbs 2006). Other important glacial refugia have also been suggested, for example, the Queen Charlotte Islands (Byun et al. 1997; Burg et al. 2005, 2006) and areas of Alaska (Anderson et al. 2006; Burg et al. 2006), as well as southern regions below the glacial maximum (Mila et al. 2000).

Genetic tools can be used to determine historical biogeography, as described above, but this can be a difficult task as signatures of historical processes and events can be obscured by current-day processes. For example, current genetic differentiation can be due to historical barriers to gene flow and/or to current-day processes such as isolation by distance (Wilke and Pfenninger 2002; Smith and Farrell 2005). Alternatively, extensive current gene flow would randomize genotypes across the species' range, effectively erasing the geographic signature of earlier events. Recent postglacial population expansion can also lead to little or no genetic differentiation among locations, even in cases where current gene flow is limited (Timmermans et al. 2005; Spaulding et al. 2006). Indeed, in North American birds the lack of genetic differentiation within phylogroups on either side of the Rockies (above) might be due to postglacial expansion from glacial refugia, high levels of current-day gene flow, or both. To separate these possibilities, it is necessary to combine extensive geographic sampling, sensitive molecular markers that can evolve over relatively short time periods, and appropriate analytical methodologies. Thus, a major issue surrounding structuring of genetic variation is assessing the relative importance of recent historical events, such as glacial cycles, and current-day processes, such as gene flow.

The American redstart (*Setophaga ruticilla*) is a widespread North American songbird occurring on both sides of the Rocky Mountains (Figure 1). If this species has a recent history similar to that of other widespread migratory birds (above), we would expect American redstarts to show evidence of 2 Pleistocene glacial refugia, that is, genetic differentiation between eastern and western North America. However, all of the widely distributed migratory songbirds listed above with east–west genetic divides consist of previously described subspecies. American redstarts lack recognized subspecific designations and thus may have a different recent evolutionary history. For example, populations may have survived in only a single glacial refugium, subsequently colonizing the rest of the continent (i.e., Zink 1996), or in multiple refugia defined by barriers other than the Rocky Mountains/Great Plains.

We had 2 primary goals in this study. First, we utilized widespread sampling and 2 types of molecular markers—mitochondrial DNA (mtDNA) control region and amplified fragment length polymorphism (AFLP)—to distinguish between the alternative historical scenarios outlined above for the American redstart. Second, we used a number of analytical approaches to separate the effects of historical events from current-day processes in determining the current pattern of genetic variation. To achieve these goals, we used standard measures of genetic differentiation (F_{ST}, analysis of molecular variance [AMOVA]) to identify genetic subdivisions across the range of the redstart and coalescent analyses (Nielsen and Wakeley 2001) to determine if those subdivisions (for mtDNA) are likely due to historical and/or current-day processes. We also tested for the signature-limited current gene flow by testing for patterns of isolation by distance.

Table 1. Phylogeographic breaks found in species with broad distributions similar to that of the American redstart

Species	Study	Location of divide	Timing of divergence (years before present/sequence divergence)
Yellow warbler	Milot et al. (2000)	Great Plains/Rockies	100 000/2.15%
Swainson's thrush	Ruegg and Smith (2002)	Rockies	10 000/0.69%
Wilson's warbler	Kimura et al. (2002)	Great Plains/Rockies	33 654–62 500/0.7–1.3%
Yellow-breasted chat	Lovette et al. (2004)	Great Plains/Rockies	Not given/1.8%
Common yellowthroat	Lovette et al. (2004)	Great Plains/Rockies	Not given/2.0%
Nashville warbler	Lovette et al. (2004)	Great Plains/Rockies	Not given/1.7–2.3%
Wood duck	Peters et al. (2005)	Great Plains/Rockies	10 000–124 000/not given
American redstart	This study	Newfoundland	40 300–2 171 000/1.0%

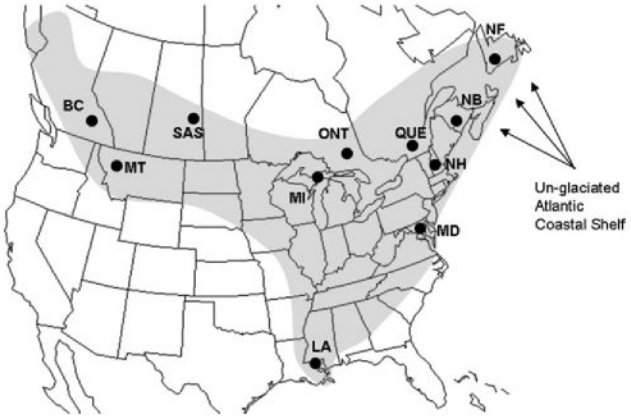

Figure 1. The breeding distribution of the American redstart is shaded in gray. Sampling sites (black dots) are labeled with the following abbreviations: LA, Louisiana; MD, Maryland; NH, New Hampshire; MI, Michigan; MT, Montana; BC, British Columbia; SAS, Saskatchewan; ONT, Ontario; QUE, Quebec; NB, New Brunswick, NF, Newfoundland. Arrows indicate the approximate location of the proposed Atlantic coastal shelf refugium.

Materials and Methods

Sample Collection

We selected collection sites to provide even coverage across the breeding range of the American redstart (Figure 1, Table 2). All individuals (adults) were captured in mist nets and released after sampling. We captured breeding males with decoys and song playback, whereas females were captured directly from their nests while they were incubating or provisioning nestlings. US samples were collected in the summer of 2002, whereas Canadian samples were collected from May to June of 1994 and 1995. Approximately 50 μl of blood was collected from the brachial vein of each individual and stored in 200 μl lysis buffer for US samples (Densmore and White 1991) and Queen's lysis buffer (Seutin et al. 1991) for Canadian samples. We isolated whole genomic DNA from US blood samples using a standard phenol–chloroform procedure and from Canadian samples using either phenol–chloroform or DNAZOL genomic DNA isolation reagent (Invitrogen, Carlsbad, CA). DNA was quantified using a fluorometer and diluted to a final working concentration of 20 μg/ml.

mtDNA Sequencing

We amplified a 290-bp sequence of the control region I gene using redstart-specific primer sequences F: TTAAGGG-TATGTATAGTATG, and R: TTCTTGAAGGCTGTT-GGTCG, which were designed from preliminary sequence information generated for redstarts using control region primers DPdl-L5 and DPdl-H4 designed for yellow warblers (*Dendrocia petechia*; Milot et al. 2000). Polymerase chain reactions (PCRs) were conducted using 1 μl of undiluted whole genomic DNA, 3 μl of $MgCl_2$, 3 μl 10× PCR buffer, 1 μl deoxyribonucleotide triphosphates, 0.5 μl forward and reverse primers (10 μM), 2 μl *Taq* polymerase (Promega, Madison, WI), and 10 μl distilled, de-ionized H_2O. Reactions consisted of 35 cycles of denaturation at 94 °C for 1 min, a 55 °C annealing temperature held for 90 s, and a 72 °C extension temperature held for 2 min. We are confident that this procedure generated mtDNA sequences as opposed to nuclear homologs because 1) sequences were clear with no double peaks (Sorenson and Quinn 1998) and 2) sequences generated from blood-extracted DNA and purified mtDNA isolated from a single bird were identical (Gibbs HL, unpublished data).

We cleaned PCR products using shrimp alkaline phosphatase procedure (Amersham, Piscataway, NJ). We then obtained DNA sequences from purified PCR products via cycle sequencing with ABI Big dye version 3.1 (Applied Biosystems, Foster City, CA) reaction mixture using protocols suggested by the manufacturer and running products on an ABI 3730 capillary sequencer. Sequence data were visualized using Sequencher software versions 3.0 (Gene Codes, Ann Arbor, MI). All mtDNA sequences generated have been deposited in GenBank (accession numbers EF999143–EF999326).

Phylogenetic Analysis of Mitochondrial Sequences

We performed all phylogenetic analyses on the Washington State University Phylogenetics, Population and Evolutionary Ecology Computer Cluster. We then used PAUP 4.0b10

Table 2. Sampling locations and number of individuals analyzed in AFLP and mtDNA analyses

Location	Latitude	Longitude	AFLP	mtDNA
Revelstoke National Park, British Columbia	50.59	118.12		34
Prince Albert National Park, Saskatchewan	53.57	106.22		19
Queen's University Field Station, Ontario	44.35	76.19		16
La Mauricie National Park, Quebec	45.35	73.41		18
Fundy National Park, New Brunswick	45.37	65.02		13
Gros Morne National Park, Newfoundland	49.36	57.31		20
Bogue Chita National Wildlife Refuge, Louisiana	30.19	89.55	27	11
Jugg Bay, Maryland	38.45	76.42	28	10
Hubbard Brook, New Hampshire	43.57	71.43	19	9
Raco, Michigan	46.24	84.33	33	14
Swan Lake, Montana	48.02	114.01	33	11

(Swofford 2003) to determine phylogenetic relationships among haplotypes using maximum likelihood. For likelihood analyses, we used DT-MODSEL (Minin et al. 2003) to select the substitution model that best fits the data using an information criterion and executed searches using a heuristic search strategy with a single random addition sequence starting from a random tree. The level of support for each node of the generated tree was evaluated with 100 bootstrap replicates. We used an mtDNA control region sequence from the chestnut-sided warbler (*Dendroica pennsylvanica*) as an out-group (GenBank accession number AF206016; Milot et al. 2000).

We used TCS 1.13 (Clement et al. 2000) to generate a haplotype network for all of the sequences. The program was set to estimate the upper limit of the number of mutational steps between haplotypes.

Gene Flow and Coalescent Modeling

We assessed the genetic structure of redstart populations with ARLEQUIN 2.0 (Excoffier et al. 1992). First, to test whether redstarts exhibit population structure similar to that of other widespread migratory warblers (Kelly and Hutto 2005), we compared 2 groups: eastern and western. Second, to test for the signature of a separate Atlantic coastal shelf refugium (Figure 1), we grouped all mainland populations together separate from Newfoundland. Patterns of gene flow were evaluated using pairwise and global F_{ST} values along with Mantel tests for isolation by distance. In order to make comparisons between mtDNA and AFLPs, it is necessary to take into account their differences in effective population size. We therefore used the equation F_{ST} (nuc) $= F_{ST}$ (mt)/4 $- 3(F_{ST}$ (mt)) (Brito 2007) to calibrate our mtDNA F_{ST} values. We used ARLEQUIN to estimate gene diversity in eastern versus western sample sites. We also used ARLEQUIN to investigate population expansion by calculating mismatch distributions and evaluating their significance by comparing them with a distribution calculated under a model of sudden population growth (Rogers 1995). MEGA 3.1 (Kumar et al. 2001) was used to infer mean net corrected sequence divergence between populations east versus west of Rockies and between Newfoundland versus continental populations.

Phylogeographic subdivisions that are associated with a potential barrier to gene flow could have arisen because that barrier has weakly limited gene flow for a long time period or because it has severely limited gene flow for a short time period. We used a coalescent modeling approach (Nielsen and Wakeley 2001) to determine which of these scenarios was more consistent with our data. Divergence times and migration rates between population groups (both eastern vs. western and Newfoundland vs. continental) were estimated using coalescent modeling. We used the MDIV software package, which implements the coalescent model described in Nielsen and Wakeley (2001). Aligned sequence data from sample sites were used to estimate the parameters Θ $(=2N_e\mu)$, M $(=N_e m$ = number of migrants between populations per generation), and T (the divergence time between populations, where 1 time unit $= N_e$ generations). These analyses each used a finite-sites model and a 3 000 000-generation Markov chain Monte Carlo run with a 500 000-generation burn-in to explore the solution space. M_{max} was set to 3, and T_{max} was set to 10. The coalescent-scaled parameter T was converted to T_{div} (time in years since 2 populations diverged) using the formula $T_{div} = T\Theta/(2\mu)$ and assuming 2 years per generation (Sillett TS, personal communication) and both a low (0.076 mutations per site per million years) and a high (0.3 mutations per site per million years) estimate of mutation rates (Davis et al. 2006).

AFLP Analysis of Genetic Structure

Our general AFLP procedures followed Vos et al. (1995) with the following modifications. We first digested 1-µl DNA samples (containing 50–250 ng DNA) by incubating with restriction enzymes *Eco*RI and *Mse*I (New England Biolabs, Ipswich, MA) at 37 °C for 3 h, followed by denaturation at 60 °C for 5 min. Adapters (Invitrogen) designed to ligate to the *Eco*RI and *Mse*I restriction sites were then added to the reaction mix along with T4 ligation enzyme (New England Biolabs) and incubated at 16 °C for 2 h. This mixture was then rerestricted and ligated simultaneously as described by Vos et al. (1995). This "two-step" restriction ligation was done in order to ensure complete and uniform DNA restriction and ligation across reactions, as when we simply employed the simultaneous restriction/ligation described by Vos et al., we observed variable results between reactions. Our preselective PCR using the primers *Eco*RI + *A* and *Mse*I + *C* had an annealing temperature of 59 °C and an $MgCl_2$ concentration of 3.0 mM. We used this annealing temperature, which is slightly above temperatures commonly reported in AFLP studies, to maximize the specificity of the primers. For final selective amplification, 1 µl of the preselective PCR product mix was added to a mixture containing fluorescently labeled *Eco*RI + 3 primers (Applied Biosystems) and unlabeled *Mse*I + 3 or +2 primers. We used a touchdown PCR for the selective amplification (Brunelli J, personal communication), consisting of 7 initial cycles, starting with an annealing temperature of 65 °C and decreasing 1 °C per cycle to finish with an annealing temperature of 59 °C, followed by 25 cycles with an annealing temperature of 59 °C. This touchdown methodology was used to maximize the specificity of the primers.

We ran products from this selective amplification on an ABI 3730 capillary sequencer with LIZ size standard (Applied Biosystems) and collected the digital gel data using ABI Prism Gene Mapper analysis software (version 3.75). Each lane file was analyzed for the presence and absence of AFLP products by eye. To maximize repeatability of fragment scoring, we scored only those fragments between 120 and 400 bp, and we considered a locus to be scorable if at least one individual had a peak above 3000 reflectance units (rfu). At each scorable locus, individuals were scored as having a band at the locus if a distinctive peak was discernable above the background noise. The smallest

scorable peaks were usually 200 rfu, although most peaks were at least 1000 rfu. These scoring criteria were established by conducting 2 different restriction/ligation reactions on 16 individuals and carrying them all the way through to selective amplification with the selective primer pair *Eco* AGG/*Mse* CCG. Bands meeting the above criteria were 100% repeatable between separate reactions, but repeatability fell off sharply when considering loci smaller than 120 bp, larger than 400 bp, or with maximum reflectance peaks below 3000 rfu in the brightest individual profiles.

We screened 24 different primer combinations but used only the 8 combinations that gave clear, consistent banding patterns (ECO-ACA, -AAC, -ATG, -AGG paired with MSE-CCG, -CG). In total, 180 bands met our repeatability standards, and 113 of those were polymorphic (polymorphic bands determined by AFLPOP; Duchesne and Bernatchez 2002). Population genetic analyses of all 180 AFLP bands (pairwise and global F_{ST}, AMOVA, and Mantel tests for isolation by distance) were conducted using GENALEX (Peakall and Smouse 2006). We conducted a Bayesian cluster analysis using the program STRUCTURE (Pritchard et al. 2000), treating all 113 polymorphic loci as haplotypic data. Although AFLPs are not haplotypic data, STRUCTURE has been shown to perform well treating AFLP data in this manner (Evanno et al. 2005). For the STRUCTURE runs, we varied the assumed number of populations (K) from 1 to 5 (the number of sample sites in the analysis), and each run consisted of a 100 000-iteration burn-in followed by 1 000 000 iterations. For a given value of K, these run parameters produced nearly identical results in repeated runs.

Results

mtDNA and Phylogenetic Analysis

We sequenced 290 nucleotides of the control region I of the mitochondrial genome from 184 individuals and found 106 distinct haplotypes based on 38 variable sites. A maximum-likelihood analysis with these 106 haplotypes yielded a tree with little structure (not shown), and no nodes received any bootstrap support above 50. Sequence divergence between most sampled populations was small (<0.10%), including sequence divergence between populations east and west of the Rockies (0.10%). In contrast, mean net uncorrected sequence divergence between continental sample sites and Newfoundland was 1.00%.

Haplotype Analyses

The constructed parsimony network (Figure 2) showed high levels of homoplasy, with many (n = 106) unique haplotypes connected in multiple ways to other haplotypes and no true "central" haplotypes. The high number of connections between haplotypes made it impossible to group them, precluding the use of nested-clade analysis. Nonetheless, one cluster of haplotypes did separate out from the others: the bulk of the samples from Newfoundland (n = 17) clustered together in an unambiguous part of the network. Clustered with these Newfoundland haplotypes were 4 haplotypes collected from eastern populations (3 individuals from New Brunswick and 1 individual from Maryland) and 1 haplotype collected from a western population (Montana, the lone "W" in Figure 2). This "Newfoundland" cluster is separated from all other haplotypes by 2 mutational steps.

Coalescent Analyses

We conducted coalescent analyses to test divisions between putative refugia: east versus west and mainland versus Newfoundland. Estimates of time since divergence between Newfoundland and the mainland, calculated using high and low mutation rates (see Materials and Methods) and the 95% confidence intervals of the estimates of T, Θ, and M, as determined by MDIV, yielded a lower estimate of 40.3 Kya and an upper estimate of 2171.0 Kya (Figure 3a). Migration rates between Newfoundland and the mainland were estimated to be between 0.46 and 2.20 females per generation (Figure 3b). In contrast, divergence estimates between eastern and western populations yielded an irregular probability distribution near zero, indicating that it is likely that these populations have never separated from each other (Nielsen and Wakeley 2001; Davis et al. 2006). The program was also unable to calculate an upper limit for migration rates between these populations, indicating high levels of contemporary gene flow.

Population Genetic Differentiation and Gene Flow

An AMOVA conducted on mtDNA haplotypes with populations partitioned according to an east–west split showed that 1.19% of the variation was explained between groups (Phi_{ST} = 0.012, P = 0.17). When the split is defined as occurring between Newfoundland and the mainland, 4.29% of the variation was explained between groups (Phi_{ST} = 0.043, P = 0.08). The mismatch distribution for mainland sampling sites was unimodal and did not significantly differ from the distribution expected under population expansion (Figure 4a; Rogers and Harpending 1992). The mismatch distribution for the Newfoundland sample site appeared to be different from the distribution expected under population expansion (Figure 4b; P = 0.06), potentially suggesting a lack of population growth (Slatkin and Hudson 1991). The mismatch distribution for the western sites also conformed to the population expansion distribution (data not shown).

Pairwise F_{ST} values calculated from mtDNA haplotype frequencies were much larger (up to an order of magnitude) for comparisons between Newfoundland and all other sites (Table 3). After Bonferroni correction for multiple comparisons, only comparisons involving Newfoundland and sites further to the west remained significant (Newfoundland compared with Louisiana, Montana, British Columbia, Saskatchewan, Ontario; all $P < 0.005$). Mantel tests revealed significant isolation by distance (Figure 5a) when considering all sample sites (r = 0.47, P = 0.004) and all sites excluding Newfoundland (r = 0.47, P = 0.013). The high correlation coefficients from the Mantel tests indicate that a large proportion (roughly 28% based on R^2 values

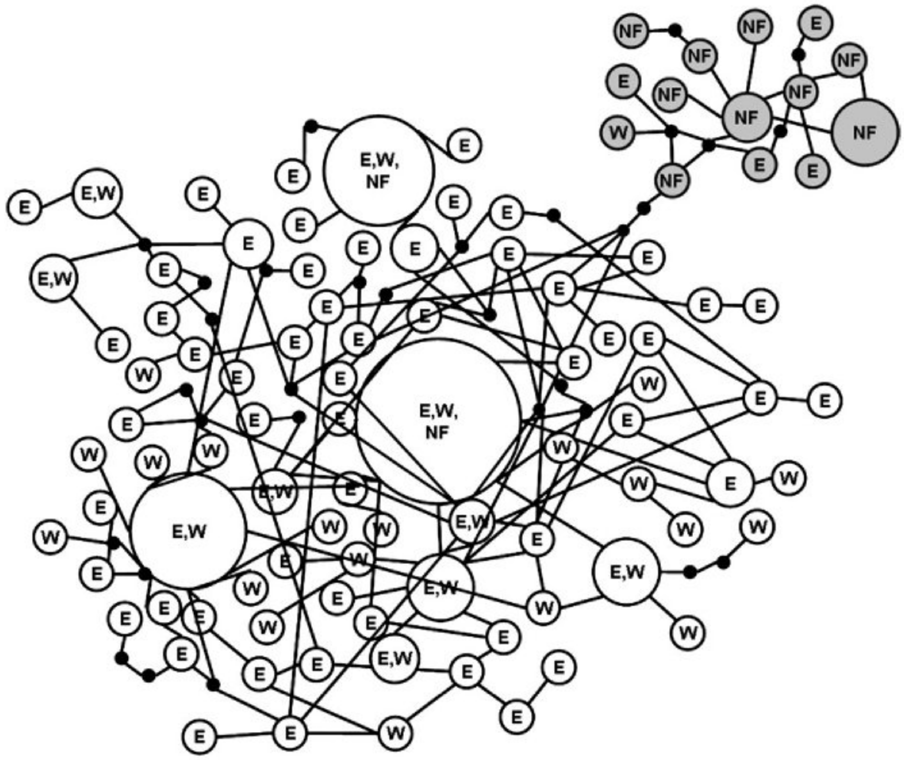

Figure 2. Haplotype network produced by TCS. Each circle represents a single haplotype (small circles indicate a single individual, and larger circles indicate that the haplotype was sampled from multiple individuals), and circles connected by a line are 1 mutational step apart. Letters indicate the locations where the haplotype was found (W, western North America; E, eastern North America; and NF, Newfoundland). Haplotypes in the "Newfoundland Clade" are shaded gray. Filled circles represent missing haplotypes.

from a regression analyses of F_{ST} and distance) of the variation in F_{ST} is explained by distance, supporting the hypothesis of limited gene flow among populations. For the sake of comparison with the AFLPs, isolation by distance was examined for the US sample sites only and showed a nonsignificant trend toward isolation by distance (Figure 5b; $r = 0.50$, $P = 0.12$).

AFLP Population Structure

For the 5 US sample sites, an AMOVA revealed that the majority of variation (94%) occurred among individuals within populations. However, a moderate but significant global $Phi_{ST} = 0.06$ ($P = 0.01$) indicated that some structuring exists. Post hoc groupings of all populations revealed that a grouping of Montana and Michigan versus Louisiana, Maryland, and New Hampshire explained a small (2%) but significant ($P = 0.01$) amount of variation. No pairwise Phi_{ST} values remained significant after Bonferroni correction for multiple comparisons. mtDNA F_{ST} values corrected for effective population size (Table 4) show that AFLP differentiation is on average 1 order of magnitude larger than mtDNA differentiation. We did not find a pattern of isolation by distance among the US populations (Figure 5b; $r = 0.62$, $P = 0.06$), though the number of populations included was small. In the STRUCTURE cluster analysis, the range of priors for K yielded a flat probability distribution with all potential K values having similar probabilities. $K = 4$ had the highest value, but each of the 5 sample sites had similar frequencies of each individual K, and most individuals were not cleanly assigned to one K. We were therefore unable to make any inference about the number of genetically distinct populations. Thus, although a clear picture of isolation by distance did not emerge, our AFLP analyses show some population structure, with eastern and southern populations being more similar to each other than to midwestern and western populations.

Discussion

Phylogeographic History

Phylogeographic studies of North American migratory birds and other species have shown that most widespread species survived Pleistocene glacial cycles in 2 or more glacial refugia: 1 in the east and 1 in the west separated by the Great Plains and/or Rocky Mountains (e.g., Milot et al. 2000; Ruegg and Smith 2002; Clegg et al. 2003; Lovette et al. 2004; Peters et al. 2005). In contrast, our results for the

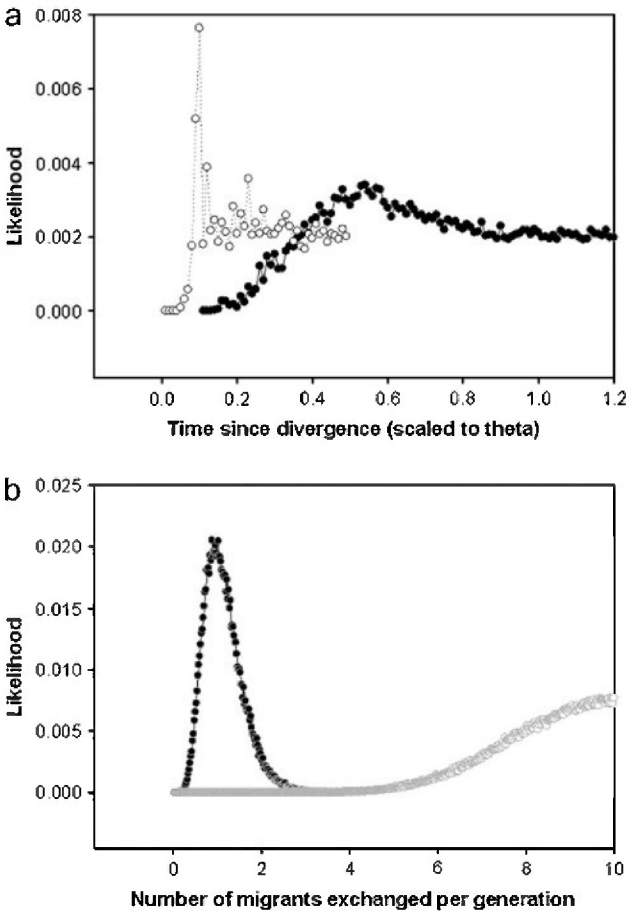

Figure 3. Likelihood graphs produced by the MDIV analysis showing time since divergence between those populations (a) and migration rate between 2 putative historical populations (b). Filled circles are the results of the analysis when run according to a Newfoundland–mainland split, whereas the open circles are the results of an east–west split.

American redstart show little or no genetic differentiation among populations throughout continental North America, a pattern similar to that seen in some widespread species with limited migration (Zink 1996). In contrast, our results for the American redstart show little or no genetic differentiation among populations throughout continental North America. This pattern agrees somewhat with morphological patterns as American redstarts are one of only 2 species of wood warbler with continental distributions (the other being the black-and-white warbler, *Mniotilta varia*) that lack formal subspecies designations (Sherry and Holmes 1997; Kelly and Hutto 2005). Thus, this species appears to have had just a single continental refugium during the Pleistocene, with expansion from this refugium to populate the rest of the continent. Because the majority of the northern portion of the redstart's range was glaciated during the Pleistocene (Figure 1), including most or all of the western portion of the range, it is likely that this continental refugium would have been in the southeast. However, although our data indicate that most or all continental redstart populations have spread out from a single glacial refugium, we cannot make firm conclusions about its exact location.

In contrast to the lack of differentiation among continental populations, our results also show that populations in far-eastern Canada are somewhat genetically differentiated from those on the continental mainland (Figures 2 and 3a, Table 3), indicating the possibility of a second refugium in the far northeast. The amount of genetic differentiation between continental and Newfoundland populations seen in this study, dating from 40 000 years ago to 2 000 000 years ago and falling within the Pleistocene, is comparable to the differentiation between eastern and western phylogroups found in other studies (Table 1). Moreover, patterns of genetic differentiation and mismatch distributions (Figure 4b) indicate that populations in this region have been relatively stable with relatively modest spread into the remainder of the continent.

Although the Atlantic coastal shelf of Canada has been proposed to have been an important refugium for other organisms (Paetkau and Strobeck 1996; Bernatchez 1997; Kyle and Strobeck 2003; Berlocher and Dixon 2004; Boys et al. 2005; Lee-Yaw et al. 2008), its importance for migratory songbirds remains unclear. This is largely because most studies have failed to incorporate samples from these areas in their studies. To the best of our knowledge, only 2 studies have sampled populations in Newfoundland: song sparrows, *Melospiza melodia* (Zink and Dittmann 1993), and yellow warblers, *D. petechia* (Boulet and Gibbs 2006). Both studies found indications that individuals could have survived glacial maxima in a northeastern refugium. Thus, studies of additional species, with extensive sampling in eastern Canada as well as other areas, are necessary before firm conclusions can be drawn about the relative importance of this putative glacial refugium for migratory songbirds.

In sum, our results not only support the "2-refugia" hypothesis but also show that most populations in North America arose from a single continental refugium. Thus, 2 or more Pleistocene refugia may be the norm for highly mobile species that currently have continental distributions. However, the exact locations of these refugia may differ, even among species with largely overlapping current distributions. The maritime regions of far-eastern Canada may have been an important glacial refugium for other North American birds, but more studies are needed before firm conclusions can be drawn.

Current-Day Gene Flow and Range Expansion

Patterns of genetic divergence found in many species affected by Pleistocene glacial cycles are often accompanied by low levels of differentiation within regions. This pattern can be explained by population contraction during glacial maxima (and thus loss of genetic diversity) followed by rapid range expansion leading to genetic homogeneity across a wide geographic area or alternatively by high levels of current-day gene flow. Historical range expansion may

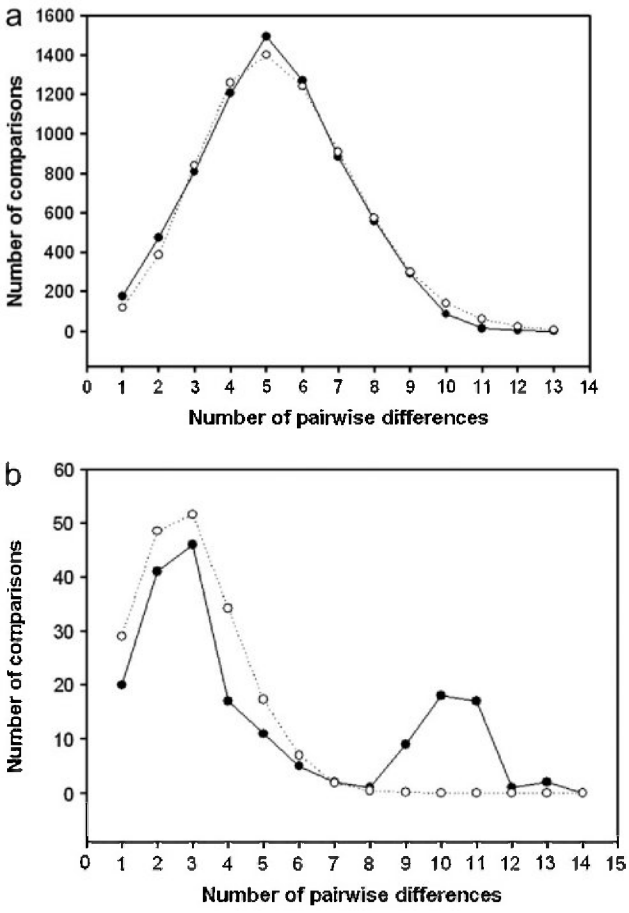

Figure 4. Distribution of pairwise nucleotide differences between individuals sampled in (a) the mainland and (b) Newfoundland. Filled circles are the observed distribution; open circles are the distribution expected under range expansion.

obscure any signal of current gene flow patterns (Mila et al. 2000), and perhaps for this reason only one study of a widespread North American bird that has identified range expansion has additionally found a pattern of isolation by distance (Gibbs et al. 2000). In the current study, mismatch distributions (Figure 4a) indicated range expansion across continental North America, and coalescent analyses indicated high levels of gene flow between populations east and west of the Rockies. Nevertheless, several results suggest that current-day gene flow is somewhat limited in this species. First, our mtDNA analyses detected significant isolation by distance across North America (Figure 5), a pattern supported by significant F_{ST} values only between Newfoundland and populations further to the west (Table 3), and AMOVA analyses of AFLP data suggest a weak pattern of limited gene flow within the continental United States. Taken together, these analyses suggest not only that continental redstarts expanded from a single glacial refugium (see Ruegg and Smith 2002) but also that gene flow across the current range is somewhat limited.

Additionally, we found much larger levels of genetic differentiation (corrected for effective population size) for nuclear markers than for mitochondrial markers (Table 4). Male-biased dispersal would lead to the opposite pattern, and so our results suggest female-biased dispersal, as has long been thought to be the norm for birds (Greenwood 1980; Clark et al. 1997). It is somewhat puzzling that nuclear markers would show greater differentiation than mitochondrial markers, as female dispersal should affect nuclear and mitochondrial gene flow equally. We currently do not have an explanation for this difference.

Comparisons of Marker Utility

Studies of widespread species that have revealed low levels of differentiation within refugial populations using mtDNA could be hindered by issues of marker sensitivity—mtDNA may simply lack useful variation at the regional level (Lovette et al. 2004). More recently, AFLP has been purported to generate a large number of highly variable markers with little development time (Mueller and Wolfenbarger 1999), yet very few researchers have used AFLPs to estimate population structure in animals (Bensch and Akesson 2005). In our comparison of the different

Table 3. Results of pairwise F_{ST} estimates obtained from mtDNA analyses

	LA	MD	NH	MI	MT	BC	SAS	ONT	QUE	NB
MD	−0.015									
NH	0.023	0.036								
MI	0.002	0.013	−0.013							
MT	0.021	0.023	0.051	0.035						
BC	0.030	0.040	0.034	0.027	0.002					
SAS	−0.001	0.020	0.027	0.011	0.036	0.043				
ONT	−0.020	0.001	0.024	0.005	0.013	0.020	−0.003			
QUE	0.002	0.014	−0.008	−0.007	0.018	0.011	0.009	−0.008		
NB	0.008	0.020	−0.008	−0.020	0.042	0.034	0.013	0.001	−0.010	
NF	**0.075***	**0.083**	**0.060**	**0.053**	**0.096***	**0.077***	**0.075***	**0.071***	**0.058**	**0.053**

Bold values are comparisons between Newfoundland and other sites; values significantly different from zero at $P < 0.005$ are denoted with an asterisk. LA, Louisiana; MD, Maryland; NH, New Hampshire; MI, Michigan; MT, Montana; BC, British Columbia; SAS, Saskatchewan; ONT, Ontario; QUE, Quebec; NB, New Brunswick; NF, Newfoundland.

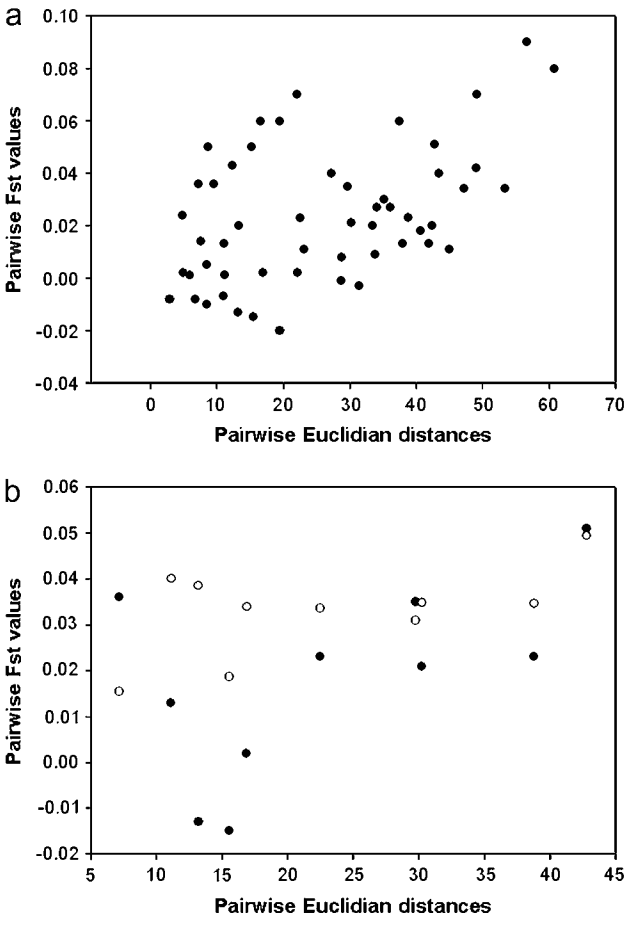

Figure 5. Scatter plot of pairwise distances (x axis) and pairwise F_{ST} values (mtDNA) for all sample sites (a). Isolation by distance for US sample sites only, filled circles are based on mtDNA and open circles are based on AFLPs (b).

markers, we discovered that, in contrast to mtDNA, AFLPs suffered from one important setback: this method seems highly sensitive to the quality of DNA and hence the sample storage and preparation methods. Laboratory analyses were conducted on samples from the United States within months of collection in the field, whereas extracted Canadian DNA samples were frozen for 7–8 years before analysis. The differences in storage time and storage buffer represent the only obvious differences between the 2 sets of samples as Canadian DNA originally extracted with a kit was reextracted with phenol–chloroform and problems persisted. We thus stress that one must be aware of this issue when considering DNA storage and preparation methodologies for AFLP analysis.

Table 4. Results of pairwise Phi_{ST} estimates obtained from AFLP analyses (above diagonal) and F_{ST} values obtained from mtDNA (calibrated for comparison with nuclear markers—see Brito 2007) below diagonal

	LA	MD	NH	MI	MT
LA		0.019	0.034	0.034	0.035
MD	−0.004		0.015	0.040	0.035
NH	0.006	0.009		0.039	0.049
MI	0.001	0.003	−0.003		0.031
MT	0.005	0.006	0.013	0.009	

LA, Louisiana; MD, Maryland; NH, New Hampshire; MI, Michigan; MT, Montana.

Conclusions

Although dispersal distances for migratory songbirds like the American redstart are potentially very large, our data show that weak but significant population structuring exists and that this genetic structure is due to both historical and current-day processes. Historically, our analyses revealed that continental populations have recently expanded from a smaller mainland refugial population, whereas redstart populations in the maritime regions of Canada are potentially descended from a northeastern refugial population. These results contrast with patterns seen in most other widespread North American migratory birds, which appear to have survived through the Pleistocene glaciations in eastern and western refugia (see also Zink 1996). Our analyses also revealed a weak pattern of limited contemporary gene flow, seen in both the mtDNA and AFLP analyses, layered on top of the pattern generated by Pleistocene events.

Boulet and Gibbs (2006) used genetic analyses to infer that at least 2 eastern refugia existed for another widespread migratory North American songbird, the yellow warbler. Furthermore, 2 of their sample sites, 1 in Newfoundland and 1 in New Brunswick, contained common haplotypes along with predominantly maritime haplotypes, raising the possibility that 1 of the eastern refugia was in the far northeast. Taken together with the genetic differentiation of the Newfoundland sample site presented in this study, these studies support the existence of an Atlantic coastal shelf refugium for at least 2 species of migratory songbird. Future phylogeographic studies should sample in this region to determine the importance of this refugial area for other extant species.

Funding

National Science Foundation (USA) to P.P.M. and M.S.W.; Max Bell Foundation and Environment Canada to H.L.G. and K.A.H.

Acknowledgments

The authors would like to thank numerous field assistants and laboratory technicians: Anne-Marie Barber, Lorie Collins, Robert Dawson, Lilie DeSousa, Jose Diaz, James Goetz, Emmanuel Milot, David Okines, and John Woods. Valuable input and advice regarding laboratory work and data analysis were provided by David Althoff, Joe Brunelli, Matt Horning,

Martin Morgan, Tobin Peever, Kari Segraves, Andrew Storfer, Jack Sullivan, Lisette Waits, and Suann Yang.

References

Anderson LL, Hu FS, Nelson DM, Pete RJ, Paige KN. 2006. Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. Proc Natl Acad Sci USA. 103:12447–12450.

Avise J, Walker CD. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. Proc R Soc Lond B Biol Sci. 265:457–463.

Ayoub NA, Riechert SE. 2004. Molecular evidence for Pleistocene glacial cycles driving diversification of a North American desert spider, *Agelenopsis aperta*. Mol Ecol. 13:3453–3465.

Bensch S, Akesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? Mol Ecol. 14:2899–2914.

Berlocher SH, Dixon PL. 2004. Occurrence of Rhagoletis species in Newfoundland. Entomol Exp Appl. 113:45–52.

Bernatchez L. 1997. Mitochondrial DNA analysis confirms the existence of two glacial races of rainbow smelt *Osmerus mordax* and their reproductive isolation in the St Lawrence River estuary (Quebec, Canada). Mol Ecol. 6:73–83.

Boulet M, Gibbs HL. 2006. Lineage origin and expansion of a Neotropical migrant songbird after recent glaciation events. Mol Ecol. 15:2505–2525.

Boulet M, Potvin C, Shaffer F, Breault A, Bernatchez L. 2005. Conservation genetics of the threatened horned grebe (*Podiceps auritus L.*) population of the Magdalen Islands, Quebec. Conserv Genet. 6:539–550.

Boys J, Cherry M, Dayanandan S. 2005. Microsatellite analysis reveals genetically distinct populations of red pine (*Pinus resinosa Pinaceae*). Am J Bot. 92:833–841.

Brito PH. 2007. Contrasting patterns of mitochondrial and microsatellite genetic structure among western European populations of tawny owls (*Strix aluco*). Mol Ecol. 16:3423–3437.

Burg TM, Gaston AJ, Winker K, Friesen VL. 2005. Rapid divergence and postglacial colonization in western North American Steller's jays (*Cyanocitta stelleri*). Mol Ecol. 14:3745–3755.

Burg TM, Gaston AJ, Winker K, Friesen VL. 2006. Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. Mol Ecol. 15:2409–2419.

Byun SA, Koop BF, Reimchen TE. 1997. North American black bear mtDNA phylogeography: implications for morphology and the Haida Gwaii glacial refugium controversy. Evolution. 51:1647–1653.

Clark AL, Saether BE, Roskaft E. 1997. Sex biases in avian dispersal: a reappraisal. Oikos. 79:429–438.

Clegg SM, Kelly JF, Kimura M, Smith TB. 2003. Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant, Wilson's warbler (*Wilsonia pusilla*). Mol Ecol. 12:819–830.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. Mol Ecol. 9:1657–1659.

Davis LA, Roalson EH, Cornell KL, McClanahan KD, Webster MS. 2006. Genetic divergence and migration patterns in a North American passerine bird: implications for evolution and conservation. Mol Ecol. 15:2141–2152.

Densmore LD, White PS. 1991. The systematics and evolution of the Crocodilia as suggested by restriction endonuclease analysis of mitochondrial and nuclear ribosomal DNA. Copeia. 3:602–615.

Duchesne P, Bernatchez L. 2002. AFLPOP: a computer program for simulated and real population allocation, based on AFLP data. Mol Ecol Notes. 2:380–383.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes—application to human mitochondrial-DNA restriction data. Genetics. 131:479–491.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology. 14:2611–2620.

Gibbs HL, Dawson RJG, Hobson KA. 2000. Limited differentiation in microsatellite DNA variation among northern populations of the yellow warbler: evidence for male-biased gene flow? Mol Ecol. 9:2137–2147.

Gill FB, Mostrom AM, Mack AL. 1993. Speciation in North-American chickadees. 1. Patterns of mtDNA genetic-divergence. Evolution. 47:195–212.

Greenwood PJ. 1980. Mating systems, philopatry and dispersal in birds and mammals. Anim Behav. 28:1140–1162.

Hofreiter M, Serre D, Rohland N, Rabeder G, Nagel D, Conard N, Munzel S, Paabo S. 2004. Lack of phylogeography in European mammals before the last glaciation. Proc Natl Acad Sci USA. 101:12963–12968.

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

Kelly JF, Hutto RL. 2005. An east-west comparison of migration in North American wood warblers. Condor. 107:197–211.

Kimura M, Clegg SM, Lovette IJ, Holder KR, Girman DJ, Mila B, Wade P, Smith TB. 2002. Phylogeographical approaches to assessing demographic connectivity between breeding and overwintering regions in a Nearctic-Neotropical warbler (*Wilsonia pusilla*). Molecular Ecology. 11:1605–1616.

Kumar S, Tamura K, Jakobsen IB, Nei M. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics. 17:1244–1245.

Kyle CJ, Strobeck C. 2003. Genetic homogeneity of Canadian mainland marten populations underscores the distinctiveness of Newfoundland pine martens (*Martes americana atrata*). Can J Zool. 81:57–66.

Lee-Yaw JA, Irwin JT, Green DM. 2008. Postglacial range expansion from northern refugia by the wood frog (*Rana sylvatica*). Mol Ecol. 17:867–884.

Lovette IJ, Clegg SM, Smith TB. 2004. Limited utility of mtDNA markers for determining connectivity among breeding and overwintering locations in three neotropical migrant birds. Conserv Biol. 18:156–166.

McGowan C, Howes LA, Davidson WS. 1999. Genetic analysis of an endangered pine marten (*Martes americana*) population from Newfoundland using randomly amplified polymorphic DNA markers. Can J Zool. 77:661–666.

Mila B, Girman DJ, Kimura M, Smith TB. 2000. Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. Proc R Soc Lond B Biol Sci. 267:1033–1040.

Milot E, Gibbs HL, Hobson KA. 2000. Phylogeography and genetic structure of northern populations of the yellow warbler (*Dendroica petechia*). Mol Ecol. 9:667–681.

Minin V, Abdo Z, Joyce P, Sullivan J. 2003. Performance-based selection of likelihood models for phylogeny estimation. Syst Biol. 52:674–683.

Mueller UG, Wolfenbarger LL. 1999. AFLP genotyping and fingerprinting. Trends Ecol Evol. 14:389–394.

Nielsen R, Wakeley J. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. Genetics. 158:885–896.

Paetkau D, Strobeck C. 1996. Mitochondrial DNA and the phylogeography of Newfoundland black bears. Can J Zool. 74:192–196.

Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes. 6:288–295.

Peters JL, Gretes W, Omland KE. 2005. Late Pleistocene divergence between eastern and western populations of wood ducks (*Aix sponsa*) inferred by the 'isolation with migration' coalescent method. Mol Ecol. 14:3407–3418.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics. 155:945–959.

Rogers AR. 1995. Genetic-evidence for a Pleistocene population explosion. Evolution. 49:608–615.

Rogers AR, Harpending H. 1992. Population-growth makes waves in the distribution of pairwise genetic-differences. Mol Biol Evol. 9:552–569.

Ruegg KC, Smith TB. 2002. Not as the crow flies: a historical explanation for circuitous migration in Swainson's thrush (*Catharus ustulatus*). Proc R Soc Lond B Biol Sci. 269:1375–1381.

Rueness EK, Stenseth NC, O'Donoghue M, Boutin S, Ellegren H, Jakobsen KS. 2003. Ecological and genetic spatial structuring in the Canadian lynx. Nature. 425:69–72.

Runck AM, Cook JA. 2005. Postglacial expansion of the southern red-backed vole (*Clethrionomys gapperi*) in North America. Mol Ecol. 14:1445–1456.

Seutin G, White BN, Boag PT. 1991. Preservation of avian blood and tissue samples for DNA analyses. Can J Zool. 69:82–90.

Sherry TW, Holmes RT. 1997. American redstart (*Setophaga ruticilla*). In: Poole A, Gill F, editors. The birds of North America. Philadelphia (PA): The Academy of Natural Sciences No.277.

Slatkin M, Hudson RR. 1991. Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. Genetics. 129:555–562.

Smith CI, Farrell BD. 2005. Phylogeography of the longhorn cactus beetle Moneilema appressum LeConte (*Coleoptera: cerambycidae*): was the differentiation of the Madrean sky islands driven by Pleistocene climate changes? Mol Ecol. 14:3049–3065.

Sorenson MD, Quinn TW. 1998. Numts: a challenge for avian systematics and population biology. Auk. 115:214–221.

Spaulding AW, Mock KE, Schroeder MA, Warheit KI. 2006. Recency, range expansion, and unsorted lineages: implications for interpreting neutral genetic variation in the sharp-tailed grouse (*Tympanuchus phasianellus*). Mol Ecol. 15:2317–2332.

Swofford DL. 2003. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sunderland (MA): Sinauer Associates.

Timmermans M, Ellers J, Marien J, Verhoef SC, Ferwerda EB, Van Straalen NM. 2005. Genetic structure in Orchesella cincta (Collembola): strong subdivision of European populations inferred from mtDNA and AFLP markers. Mol Ecol. 14:2017–2024.

Veit ML, Robertson RJ, Hamel PB, Friesen VL. 2005. Population genetic structure and dispersal across a fragmented landscape in cerulean warblers (*Dendroica cerulea*). Conserv Genet. 6:159–174.

Vos P, Hogers R, Bleeker M, Reijans M, Vandelee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. 1995. Aflp—a new technique for DNA-fingerprinting. Nucleic Acids Res. 23:4407–4414.

Wilke T, Pfenninger M. 2002. Separating historic events from recurrent processes in cryptic species: phylogeography of mud snails (*Hydrobia spp.* Mol Ecol. 11:1439–1451.

Zink RM. 1996. Comparative phylogeography in North American birds. Evolution. 50:308–317.

Zink RM, Dittmann DL. 1993. Gene flow, refugia, and evolution of geographic-variation in the song sparrow (*Melospiza melodia*). Evolution. 47:717–729.

Received July 4, 2007
Accepted April 8, 2008

Corresponding Editor: Rob Fleischer