Introduction

Molecular investigations of evolutionary history (molecular phylogenetics) have proven useful for addressing questions ranging from the evolution of avian brood parasitism (Lanyon 1992) to evolutionary relationships within (e.g. Ball et al. 1988) and among species (e.g. Lee et al. 1996). The potential of phylogenetic techniques to reconstruct these evolutionary relationships has been tested under a variety of conditions (Hillis et al. 1992, 1994) indicating that these methods are effective at recovering historical patterns of relatedness.

Cases exist, however, where molecular phylogenetic analyses produce results discordant with assumed organismal history (e.g. Cohen et al. 1997). One reason for conflict might be that the markers used are not selectively neutral, which could confound inference of evolutionary history (Sibley & Ahlquist 1990). An alternative might be that one or several data sets conflict because of excessive homoplasy or incongruence between gene and species trees. We investigated a case where population history inferred from molecular data contradicts the history inferred from morphology in this species.

The song sparrow (Melospiza melodia) is a widely distributed North American passerine bird. Aldrich (1984) reviewed phenotypic variation and noted the presence of considerable geographical differentiation in several external phenotypic characters, but offered no information about how geographically segregated phenotypes might be related evolutionarily. Nonetheless, phenotypic variation has been partitioned into as many as

Geographic analysis of nucleotide diversity and song sparrow (Aves: Emberizidae) population history

ADAM J. FRY† and ROBERT M. ZINK
J. F. Bell Museum, and Department of Ecology, Evolution and Behaviour, 100 Ecology Building, University of Minnesota, St. Paul, MN 55108–6097, USA

Abstract

Mitochondrial DNA (mtDNA) control-region (CR) sequences were analysed to address three questions regarding the evolution of geographical variation in song sparrows. (i) Are mtDNA sequences more informative about phylogenetic relationships and population history than previously published restriction fragment (RFLP) data? (ii) Are song sparrow CR sequences evolving in a selectively neutral manner? (iii) What do the haplotype cladogram and geographical pattern of nucleotide diversity (π) suggest about the recent evolutionary history of song sparrow populations? Results from phylogenetic analyses of CR sequences corroborate RFLP results and reveal instances in which haplotypes do not group by locality. Neutrality tests (Tajima 1989a) suggest that song sparrow mtDNA is evolving in a selectively neutral manner, although exceptions are noted. A novel geographical pattern of π suggests a model of song sparrow population history involving multiple Pleistocene refugia and colonization of some formerly glaciated regions from multiple sources. Moreover, application of coalescence theory to the haplotype cladogram suggests that two different haplotypes (48NF and 151HA) may have predominated in different parts of the song sparrow’s range. This model provides insight into the current distribution of song sparrow mtDNA haplotypes and may explain the discordance between evolutionary history inferred from mtDNA and morphology in this species.

Keywords: analysis of molecular variance (AMOVA), control-region sequences, Melospiza melodia, mtDNA, nucleotide diversity, Pleistocene refugia

Received 28 October 1997; revision received 20 April 1998; accepted 20 April 1998
34 subspecies (American Ornithologists’ Union 1957), and an individual song sparrow’s phenotype is, in most cases, an excellent predictor of its geographical home. Such a well-ordered pattern of geographical variation is usually taken as evidence that natural selection has adapted individual phenotypes to local environments during periods of isolation (Mayr 1963; but see James [1983]). Given sufficient time, such isolation should also result in the geographical partitioning of genetic variation. However, phylogenetic analysis of mitochondrial DNA (mtDNA) restriction fragment length polymorphisms (RFLPs) (Zink & Dittmann 1993) produced hundreds of equally most parsimonious trees (EMPTs) that revealed little phylogeographic structure, suggesting a paradox. That is, how could morphology be highly ordered geographically and mtDNA not? Lack of geographical structure in a rapidly evolving molecular marker such as mtDNA suggests high levels of gene flow, either historical or contemporary. Such gene flow is expected to have a homogenizing effect on other characters such as morphology, preventing geographical differentiation (but see Ehrlich & Raven 1969). To explore the discrepancy between song sparrow mtDNA and external phenotype, we sequenced mtDNA from the rapidly evolving control-region (CR) (Taberlet 1996; Baker & Marshall 1997). Our goal was to determine if song sparrow mtDNA was evolving in a manner consistent with neutral expectations and to provide finer phylogeographic resolution than that afforded by RFLP data. Furthermore, sequence data allowed us to address questions concerning population expansion and the nature of postglacial colonization by song sparrows.

Materials and methods

Specimens

Ninety-five song sparrows (a subsample of 170 song sparrows used by Zink & Dittmann [1993]) from 18 localities and 16 subspecies (American Ornithologists’ Union 1957) were used in this study (Table 1). In addition, one Lincoln’s sparrow (Melospiza lincolni) and two swamp sparrows (M. georgiana) were used as outgroup taxa. Information regarding specific collecting localities or extraction and cesium chloride (CsCl) purification of mtDNA from these tissues is available in Zink & Dittmann (1993).

DNA sequencing

An ~760 bp fragment of the mtDNA CR was amplified from CsCl-purified mtDNA via the polymerase chain

Table 1. Song sparrow nucleotide diversity (π). Sample localities (acronym) (abbreviations follow Zink & Dittmann 1993), sample size (N), nucleotide diversity (π), standard error (SE), subspecies, and haplotypes found at each locality (numbers in parentheses are the number of individuals possessing that haplotype). RFLP values calculated from Zink & Dittmann (1993). Trinity County, California (YO) values not shown because individuals possessing that haplotype). RFLP values calculated from Zink & Dittmann (1993).
reaction (PCR) using standard thermocycling regimes. Four PCR primers were used: LCR4 and H1248 (Tarr 1995) amplified an ~760 bp product spanning the 5’ end of the central conserved box to the 3’ end of the ‘right domain’ (Southern et al. 1988; Saccone et al. 1991). We used two internal primers, LCON2 and HGT5 (J. Klicka, personal communication), and both light and heavy strands were resolved.

PCR product from 48 birds including outgroup taxa was cleaned and sequenced using the PCR Product Sequencing Kit (United States Biochemical) following the manufacturer’s directions. Sequencing reaction products were run on 6% acrylamide gel rígds, visualized by autoradiography (35S), and aligned manually. Sequences could not be confidently scored close to the two outermost primers. Thus, we report a 700 bp segment for all individuals.

Forty-seven additional song sparrows were sequenced using an automated DNA sequencer. PCR reactions involved only the external primers LCON2 and H1248 (Tarr 1995) as in most cases automated sequencing provided unambiguous sequences 700 bp long. PCR products were purified with a QIAquick PCR purification Kit (Qiagen). Approximately 75 ng of double-stranded PCR product was used in cycle-sequencing reactions using fluorescent dye terminators and AmpliTaq FS (Applied Biosystems). Unincorporated dyes were removed from sequencing reaction products with Centri-Sep columns (Princeton Separations). Reaction products were run on an ABI 377 (Applied Biosystems) automated DNA sequencer. Sequences for light and heavy strands were aligned and ambiguous bases resolved using Sequencher version 3.0 (Gene Codes Corp.). Sequences differing by at least 1 bp were considered unique haplotypes. Sequences have been deposited in GenBank under Accession nos AF053828–AF053882.

**Phylogenetic methods**

Maximum parsimony analysis of unique haplotypes was performed using test version 4.0d55 of PAUP* (hereafter referred to as PAUP*) written by D. L. Swofford. The large observed EMPTs was significantly shorter than the distribution of tree lengths generated by randomizing the data over the taxa. A significant value is taken to be indicative of the presence of phylogenetic signal.

To determine if a geographically structured tree was a better explanation of the data, we built a structured tree in macclade (Maddison & Maddison 1993) in which haplotypes from the same locality were forced into monophyletic clades and the locality clades were arrayed geographically. We then performed the Kishino–Hasegawa (Kishino & Hasegawa 1989) log-likelihood ratio test (K-H-test) in PAUP* to determine if the structured tree was a significantly worse explanation of the data than any of the 10 parsimony trees. We also compared the 10 parsimony trees against each other and the neighbour-joining tree using the K-H-test to determine if they differed significantly in their ability to explain the data.

Coalescence of haplotypes from different localities is explained as a past dispersal event (Slatkin & Maddison 1989). We used macclade (Maddison & Maddison 1993) to determine the number of dispersal events over the neighbour-joining tree and compared it to the average number of dispersal events over 1000 random trees. An observed number of dispersal events less than the mean over 1000 randomizations suggests that there are restrictions to gene flow.

**Sequence statistics**

Haplotype or gene diversity (h) for all song sparrow sequences was calculated as: \( h = (n/n - 1)(1 - \Sigma f_i^2) \) (Nei 1987) where \( f_i \) is the frequency of the i-th haplotype and \( n \) is the number of individuals sampled.

Nucleotide diversity (π) (Nei & Li 1979; Nei & Tajima 1981), defined as the average number of nucleotide differences per site in pairwise comparisons among DNA sequences (Nei 1987), is given by: \( \pi = \Sigma \pi_{ij}/(n(n-1)/2) \) where \( \pi_{ij} \) equals the proportion of nucleotide differences between the i-th and j-th sequences and \( n \) is the number of individuals. SE of \( \pi \) was calculated from eqn 10.9 of Nei (1987) for sample sizes 15 or fewer and by eqn 10.10 (Nei 1987) for larger sample sizes. We calculated \( \pi \) for all 18 localities used in this study as well as for 21 localities for which there are RFLP data (Zink & Dittmann 1993). To determine if \( \pi \) values conformed to a predictable geographical pattern, both weighted and unweighted least-squares linear regression were performed with \( \pi \) as the dependent variable and latitude, longitude, or latitude and longitude as predictor variables. The inverse of the SE of \( \pi \) was used as the weighting variable in the weighted analysis. In addition, individuals belonging to the same subspecies (American Ornithologists’ Union 1957) were grouped to increase sample sizes and \( \pi \) calculated from these groupings. These groupings were: 1 (AL, MA, MN, SA; \( n = 33 \)); 2 (HA, PR; \( n = 18 \)); 3 (MI, NC; \( n = 10 \)); 4 (NE,
Among subspecies (CR) \(0.2330\), among subspecies (RFLP) \(0.6098\) for Song Sparrow and among subspecies (RFLP) \(0.2479\) for Fox Sparrow. For comparison, \(n\) was calculated for 15 localities from the red-winged blackbird (Agelaius phoeniceus) RFLP data (Ball et al. 1988). In some cases, their localities (7 and 8, 9 and 10, and 14 and 15) were pooled to increase sample sizes.

**Tajima’s D statistic**

Departures from neutral molecular evolution (Kimura 1983) were tested using the method of Tajima (1989a). The \(D\) statistic was computed for all song sparrow samples pooled for both RFLP (Zink & Dittmann 1993) and sequence data. In addition, \(D\) was calculated for all song sparrow RFLP and sequence localities separately. Confidence intervals for \(D\) are taken from the beta distribution of Tajima’s (1989a) Table 2.

**Analysis of molecular variance (AMOVA)**

AMOVA was performed using WINAMOVA (Excoffier 1992), with localities grouped into recognized subspecies (American Ornithologists’ Union 1957). This produced three variance estimates: among subspecies, among populations within subspecies, and within populations. Significance of the variance estimates was obtained using a randomization procedure (1000 permutations). WINAMOVA also estimates \(\Phi\)-statistics (Excoffier et al. 1992), which are haplotype correlation measures analogous to \(F\)-statistics. The \(\Phi\)-statistics are defined as follows: \(\Phi_{CT}\) is the correlation of random haplotypes within subspecies relative to that of random haplotypes drawn from all populations; \(\Phi_{SC}\) is the correlation of random haplotypes within populations relative to that of random haplotypes drawn from the subspecies to which that population belongs; and \(\Phi_{ST}\) is the correlation of random haplotypes within populations relative to that of random haplotypes drawn from the whole (Excoffier et al. 1992).

The input matrix for the AMOVA analyses was obtained by calculating the number of differences between unique haplotypes in pairwise comparisons. This matrix is similar to the \(D1\) metric suggested by Excoffier et al. (1992) when haplotypes clearly differ. For comparative purposes, AMOVA analysis was performed on fox sparrow (Passerella iliaca) RFLP data from Zink (1994) which were divided according to his four ‘species-groups’.

**Results**

**Description of mtDNA sequences**

The 95 song sparrow sequences exhibited 43 (6.14%) variable positions and 29 (4.14%) phylogenetically informative sites. No insertions or deletions were noted and all base changes were transitions. Fifty-five (58%) of the 95 individuals examined possessed unique haplotypes. Sequence divergence between song and Lincoln’s sparrow was 2.76% and averaged 2.67% between song and the two swamp sparrows (which differed by 0.14%). Divergence between Lincoln’s and the two swamp sparrows averaged 1.78%.

The average divergence among all song sparrow haplotypes was 0.49% and ranged from 0.00% to 1.85%, the latter a comparison between the unique haplotype possessed by all three birds from Trinity County California (YO) and the unique single haplotype from Alberta, Canada (AL). Interestingly, this level of intraspecific sequence divergence exceeds the level of divergence for the same stretch of sequence between two undisputed species, the swamp and Lincoln’s sparrows. The two most common song sparrow haplotypes (151HA and 48NF) together accounted for 34% of all sequences examined and are widely distributed geographically (Fig. 1a).

**Table 2 AMOVA for song sparrow control-region sequences (CR), RFLP data (Zink & Dittmann 1993), and fox sparrow RFLP data (Zink 1994)**

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Variance</th>
<th>% Total</th>
<th>(P^*)</th>
<th>(\Phi^+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Song Sparrow</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among subspecies (CR)</td>
<td>(-0.2330 \times 10^{-5})</td>
<td>-0.09%</td>
<td>0.5385</td>
<td>(\Phi_{CT} = -0.001)</td>
</tr>
<tr>
<td>Among pop./subsp. (CR)</td>
<td>0.6957 \times 10^{-3}</td>
<td>27.19%</td>
<td>&lt;0.0010</td>
<td>(\Phi_{SC} = 0.272)</td>
</tr>
<tr>
<td>Within populations (CR)</td>
<td>0.1865 \times 10^{-2}</td>
<td>72.90%</td>
<td>NC</td>
<td>(\Phi_{ST} = 0.271)</td>
</tr>
<tr>
<td>Among subspecies (RFLP)</td>
<td>0.2479 \times 10^{-2}</td>
<td>15.96%</td>
<td>0.032</td>
<td>(\Phi_{CT} = 0.160)</td>
</tr>
<tr>
<td>Among pop./subsp. (RFLP)</td>
<td>0.2083 \times 10^{-2}</td>
<td>13.42%</td>
<td>0.008</td>
<td>(\Phi_{SC} = 0.160)</td>
</tr>
<tr>
<td>Within populations (RFLP)</td>
<td>0.1096 \times 10^{-1}</td>
<td>70.62%</td>
<td>&lt;0.001</td>
<td>(\Phi_{ST} = 0.294)</td>
</tr>
<tr>
<td>Fox Sparrow</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among subspecies (RFLP)</td>
<td>0.6098 \times 10^{-2}</td>
<td>73.08%</td>
<td>&lt;0.0010</td>
<td>(\Phi_{CT} = 0.731)</td>
</tr>
<tr>
<td>Among pop./subsp. (RFLP)</td>
<td>0.2854 \times 10^{-3}</td>
<td>3.42%</td>
<td>0.0010</td>
<td>(\Phi_{SC} = 0.127)</td>
</tr>
<tr>
<td>Within populations (RFLP)</td>
<td>0.1961 \times 10^{-2}</td>
<td>23.50%</td>
<td>&lt;0.0010</td>
<td>(\Phi_{ST} = 0.765)</td>
</tr>
</tbody>
</table>

*The probability of having a more extreme variance component and \(\Phi\) statistic than the observed value by chance alone. NC, not computed.

†See the Materials and methods for an explanation of \(\Phi\)-statistics.
Phylogenetic analysis

The PTP test (Faith & Cranston 1991) suggested the presence of phylogenetic signal in the song sparrow data as the observed EMPTs length (111 steps) is significantly shorter \((P = 0.01)\) than the distribution of tree lengths generated using randomized data.

The neighbour-joining (Fig. 2) and maximum parsimony trees (length = 111; CI = 0.576; RI = 0.743) were not identical but were generally congruent and both showed some tendency for haplotypes from the same locality to group together, particularly birds from Santa Cruz Island (SCI). The 5000 EMPTs yielded a completely unresolved strict consensus tree. Overall, the neighbour-joining tree (Fig. 2) showed more geographical structure than the RFLP tree (Zink & Dittmann 1993, their Fig. 4). Broad geographical groups of haplotypes, indicated by parentheses in Fig. 2, exist, although there are haplotypes within each group that appear out of place. In addition, the two most common haplotypes each group with a different broad geographical region (Fig. 2): 48NF groups with a primarily central/eastern clade and 151HA groups with a primarily western clade (Fig. 1b).

According to the \(K-H\)-test, a geographically structured tree is a significantly worse explanation \((P < 0.01)\) of the song sparrow data than the parsimony trees. However, none of the 10 parsimony trees is significantly worse than any other or the neighbour-joining tree (mean \(P > 0.33)\).

Twenty-seven dispersal events are implied over the neighbour-joining tree (Fig. 2). This value is less than the mean, 45.8 (range 40–59), found across 1000 random trees suggesting that, although the number of dispersal events is high, there are restrictions to gene flow (Maddison & Maddison 1993).

Sequence statistics

Haplotype diversity \((h)\) for the song sparrow sequence data was 0.77. Nucleotide diversity \((\pi)\) per locality (Table 1) for song sparrow sequence data ranged from 0.00095 to 0.00800 and from 0.00133 to 0.00800 for RFLP data. Nucleotide diversity for red-winged blackbird RFLP data (Ball et al. 1988) ranged from 0.00132 to 0.00507 and displays a statistically significant trend (slope of regression differs from zero; d.f. = 13, \(P < 0.01, F = 7.73\)) of decreasing \(\pi\) with increasing latitude (Fig. 3a). Plots of song sparrow sequence \(\pi\) values (Table 1) vs. latitude (Fig. 3b) or longitude (not shown) separately show no significant trend \((\pi\) vs. latitude: d.f. = 19, \(P > 0.24, F = 1.43; \pi\) vs. longitude: d.f. = 19, \(P > 0.14, F = 2.31\)). In a weighted least-squares linear regression with \(\pi\) as the dependent variable and 1/SE of \(\pi\) as the weighting variable, latitude and longitude together are not significant predictors of song sparrow sequence \(\pi\) values (d.f. = 18, \(P > 0.08, F = 2.77\)). However, in an unweighted least-squares linear regression, latitude and longitude together are significant predictors of song sparrow sequence \(\pi\) values (d.f. = 18, \(P < 0.04, F = 3.64\)).

The geographical distribution of \(\pi\) values (Fig. 4) shows highest values in formerly glaciated localities in north-central North America. Intermediate \(\pi\) values occur along the east coast of North America south of New England and the northwest coast, specifically the Queen Charlotte Islands. Areas with relatively low \(\pi\) are found in Newfoundland, Kenai Peninsula, and Michoacan, Mexico, for example. The \(\pi\) values calculated with individuals grouped into subspecies ranged from 0.00260 to 0.00730 and produced a pattern similar to the pattern in Fig. 4. For example, the largest value \((\pi = 0.00732, SE = 0.00198, group 1)\) is found in north-central North America, relatively large values are found in the northeast.
(\(\pi = 0.00628, SE = 0.00192, group 3, 4\)) and northwest
(\(\pi = 0.00568, SE = 0.00181, QC\)), and relatively small values
in the remaining areas (e.g. \(\pi = 0.00311, SE = 0.00140,\\n\text{group 5; } \pi = 0.00269, SE = 0.00148, \text{group 8}\)).

**Tajima's D statistic**

When all song sparrow sequences were pooled, a significant negative Tajima’s D results (\(D = -1.8432, P < 0.05, n = 95\)), where \(P\) equals the probability of obtaining the observed \(D\) value under the neutral mutation hypothesis. However, \(D\) for the RFLP data did not differ significantly from zero (\(D = -0.7223, P > 0.10, n = 170\)). When each locality was considered separately, both Newfoundland (\(D = -2.1169, P < 0.01, n = 7\)) and North Carolina (\(D = -2.4610, P < 0.01, n = 5\)) sequence localities produce significant negative D-values. However, none of the other RFLP or sequence localities produced a \(D\) value significantly different from zero (\(P > 0.05\)).

**AMOVA**

For both song sparrow sequence and RFLP data, AMOVA indicated a large variance component within populations (Table 2) and relatively less variance distributed among subspecies. The relatively low \(\Phi^\text{ST}\) values for both sequence and RFLP (0.271 and 0.294, respectively) suggest little geographical genetic subdivision in song sparrow mtDNA. For comparison, a relatively large \(\Phi^\text{ST}\) value of 0.765 suggests relatively strong genetic subdivision in fox sparrows (Table 2).

**Discussion**

The pictures of song sparrow evolutionary history implied by morphology and mtDNA differ. Song sparrow morphology is well-ordered geographically (Aldrich 1984) and suggests that natural selection has adapted phenotypes to local environments (Miller...
MtDNA RFLP data and sequences, however, suggest a history replete with continent-wide gene flow and relatively little geographical genetic subdivision (Table 2, Fig. 2). We explore the apparent discrepancy between molecules and morphology and suggest a model of population history that might explain the geographical distribution of song sparrow mtDNA haplotypes in the following text.

RFLP vs. sequences

There is general agreement between the picture of evolutionary history implied by the RFLP data of Zink & Dittmann (1993) and the CR sequences used in this study. RFLPs and sequences yield similar estimates of $\pi$ (Table 1) and AMOVA results (Table 2), suggesting that RFLP data are as informative about song sparrow phylogeography as are mtDNA CR sequences. Phylogenetic analysis of CR sequences, however, provided a more geographically structured haplotype tree. Similar agreement between RFLP data and mtDNA sequences has been reported in other studies (e.g. Dodge et al. 1995).

Neutral evolution of song sparrow mtDNA

Several authors (e.g. Tajima 1989b; Fu & Li 1993; Aris-Brosou & Excoffier 1996; Rand 1996) have noted that Tajima’s (1989a) test is sensitive to population history as well as to non-neutral molecular evolution. For example, a history of population growth or recent population bottleneck will tend to drive the value of $D$ (of otherwise neutral markers) toward more negative values, whereas population structure or mutation rate heterogeneity will drive $D$ toward more positive values. Thus, interpretation of $D$ may not be as straightforward as simply assuming the influence of selection.

We are hesitant to interpret the failure of CR sequences to pass Tajima’s (1989a) test as indicative of non-neutral
molecular evolution. Population expansions must have been a conspicuous feature of song sparrow history because a large part of the song sparrow’s current breeding range was formerly glaciated. A recent population expansion was also inferred from the distribution of pairwise differences (mismatch distribution) produced by song sparrow RFLP data (Zink 1997). Moreover, population expansion could account for the significant negative $D$-values found for the Newfoundland and North Carolina samples, which may have only recently been colonized.

The discrepancy between $D$-values obtained from sequence and RFLP data is difficult to reconcile. Although there were more variable sites in the sequence data (43) than the RFLP data (25), the average divergence among individuals in pairwise comparisons was similar for the two data sets (4.00 for RFLP and 3.49 for sequences). This gives the sequence data the appearance of being ‘underdivergent’ and pushing divergent events into a narrow time window, producing a star phylogeny. Data of this type drive the value of $D$ toward more negative values (Tajima 1989a). These results might best be interpreted as empirical evidence that Tajima's (1989a) $D$ value can provide information about population history as well as neutral molecular evolution.

**Phylogenetic relationships of song sparrow haplotypes**

Phylogenetic analysis of song sparrow CR sequences produced thousands of EMPTs. Maximum likelihood tests show that a geographically structured tree is a significantly worse explanation of the data than trees generated from parsimony analysis. However, the PTP test (Faith & Cranston 1991) indicates the presence of phylogenetic signal. For example, song sparrows from Santa Cruz Island (SCI) group together on the neighbour-joining tree. The distinctness of other vertebrate taxa on the California channel islands has been noted (Wayne et al. 1991; Mundy et al. 1997) and suggests a period of isolation for SCI song sparrows. However, the phylogenetic analysis (Fig. 2) does reveal many instances in which haplotypes from the same locality are not most closely related. We conclude that the haplotype cladogram represents an intermediate stage of phylogeographic structure between paraphyly and reciprocal monophyly (Avise et al. 1987).

The number of dispersal events inferred over the neighbour-joining tree (Fig. 2) suggests widespread gene exchange short of panmixia. Although we cannot distinguish ongoing gene flow from the historical associations of populations, the current level of geographical differentiation in morphology suggests that gene flow may be more of a historical phenomenon than a contemporary one.

**Geographic analysis of $\pi$**

The $\pi$ values observed here (Table 1) are consistent with values obtained for other avian species (Seutin et al. 1995) and populations in general (Stephan & Langley 1992). However, the SEs of these estimates are quite large and ‘provide a sobering reminder of the large stochastic variance associated with selectively neutral evolutionary processes’ (Kreitman 1991; p. 206). Despite this large stochastic variance, we believe that the $\pi$ values observed for song sparrow localities represent a real biological pattern. Support for this assertion comes from two observations. First, a similar geographical pattern of $\pi$ values is obtained with a subsample (95 individuals for CR sequences) of a larger sample (170 individuals from original RFLP analysis (Zink & Dittmann 1993)) Second, if small sample sizes had compromised our $\pi$ values, we would expect the observed geographical pattern of $\pi$ to be random with both large and small values scattered across Fig. 4. However, if we ignore the SE values, unweighted linear regression demonstrates that latitude and longitude can significantly predict $\pi$ values. We conclude that although large sample sizes are almost always desirable, in this case they may refine the geographical pattern of $\pi$ but not significantly alter it. Thus, the observed pattern of $\pi$ values may be useful if analysed geographically.

Recent studies have used DNA polymorphism to infer historical patterns of population expansion. The prediction is that a refugial population spreading from its leading edge will experience a series of bottlenecks that will reduce diversity (Nei et al. 1975; Hewitt 1996). Thus, mtDNA diversity should decrease with distance from a refugium. This prediction has been modelled by computer simulation (Hewitt 1996) and observed empirically (Bernatchez & Dodson 1990, 1991; Hayes & Harrison 1992; Hewitt 1996; Merila et al. 1997) and is observed in the red-winged blackbird RFLP data (Ball et al. 1988) (Fig. 3a). This pattern suggests that red-winged blackbirds may have spread northward from a southern refugium.

An interesting feature of $\pi$ is that because it takes into account the divergence between haplotypes, it can be inflated if dispersing haplotypes from different refugia meet in a zone of contact. Thus, unexpectedly large $\pi$ values may be found in areas that have received immigrants from more than one refugium. The large song sparrow $\pi$ values observed in north-central North America may reflect this latter point.

Highest $\pi$ values are found in formerly glaciated regions in north-central North America (e.g. AL, MA, MI, MN, SA) (Fig. 4). Obviously, these were not suitable refugial areas because as recently as 12 000-years ago (Teller 1987) this region was covered by glacial ice. However, the observed pattern of $\pi$ is consistent with a hypothesis of
multiple refugia and secondary contact of haplotypes from different refugia. We posit that formerly glaciated regions in north-central North America recently received immigrants from different refugial populations and the larger \( \pi \) values result from secondary admixture of haplotypes from these different refugia.

Formerly unglaciated localities with relatively high \( \pi \) values could therefore represent potential sites of Pleistocene refugia. We envision two or three primary refugial areas: the east coast of North America, particularly the region south of New England but not including North Carolina (NC), which shows low \( \pi \), and the coastal pacific northwest, particularly the Queen Charlotte Islands (QC). Other authors have identified one or both of these regions as potential refugial sites in both animals and plants (e.g. Bernatchez & Dodson 1991; Soltis et al. 1991; Hayes & Harrison 1992; O’Reilly et al. 1993; Orti et al. 1994; Talbot & Shields 1996). Southern California provides another potential refugial site, although more extensive sampling is needed to support this. However, long-term presence of song sparrows in southern California may be supported by the fact that two haplotypes from this area (SS) are basal to all other haplotypes when the tree is rooted with Lincoln’s and swamp sparrows. The remaining haplotypes from this area are found within the top-most central/east clade in Fig. 2. This could represent a dispersal event from eastern North America.

Those localities with the lowest \( \pi \) could have been recently colonized from single refugial populations following deglaciation. For example, the data support recent colonization of the Kenai Peninsula, Alaska (KE) probably from a western refugium, colonization of southern Mexico (MC) from a western refugium, and colonization of Newfoundland (NF) from an eastern refugium (Fig. 2). Zink & Dittmann (1993) suggested Newfoundland as a possible refugial site for song sparrows; however, \( \pi \) values from sequences and their RFLP data support a recent colonization hypothesis.

The multiple refugium hypothesis is consistent with the observation that the two most common haplotypes are found in different geographical groups (Fig. 2). Coalescence theory predicts that the probability that a haplotype is the oldest is equal to its frequency in the sample (Watterson 1976; Kelly 1977; Watterson & Guess 1977) and the expected rank of haplotypes by age is equal to their rank by frequency (Donnelly & Tavare 1986). This suggests that the two most common haplotypes (151HA and 48NF) may be the oldest. Furthermore, old haplotypes are expected to be geographically widespread, which we observed (Fig. 1a), but their derived (younger) and less-common relatives have a reduced chance of being part of dispersal events (assuming haplotypes disperse randomly and in proportion to their frequency). Thus, we expect old haplotypes to be very widespread but their ‘progeny’ to be found close to areas where they arose. This reasoning supports our suggestion that the two most-common song sparrow haplotypes were restricted to different sides of the continent, perhaps in the refugial areas discussed above. Contrary to predictions, however, neither common haplotype was identified as ‘basal’ by outgroup rooting (Fig. 2), suggesting that the 48NF and 151HA haplotypes may not be the absolute oldest but could still be relatively old compared to most observed song sparrow haplotypes.

**Molecules vs. morphology**

The discrepancy between morphology and mtDNA observed in song sparrows may be explained by assuming a reduction in gene flow after song sparrows reclaimed formerly glaciated regions in North America. If gene flow is not a contemporary phenomenon, the vestiges of colonization and historical gene flow should be retained in the mtDNA structure of song sparrows for a relatively long time, namely \( 4N_e \) generations. Although \( N_e \) for song sparrows is unknown, if it approaches that \( (57,500) \) suggested by Ball & Avise (1992), mtDNA haplotypes will not exhibit reciprocal monophyly relative to geographical locality for over 200,000 years (assuming a 1-year generation). However, the morphological traits often used in subspecies descriptions may be highly heritable (Schluter & Smith 1986), and probably polygenic. Therefore, these traits could evolve faster than mtDNA, which by virtue of its completely linked and maternal inheritance behaves as a single locus. Thus evolution of mtDNA and morphology could become decoupled.

Alternatively, it is possible that some elements (e.g. plumage colouration or pattern) of song sparrow phenotypic diversity have a significant environmental component (James 1983), or phenotypic diversity may predate the last round of glacial advance. Support for the latter may come from the fact that most phenotypic diversity in song sparrows is found in unglaciated parts of western North America (Zink & Dittmann 1993, their Fig. 1). Independent molecular markers would be useful for addressing whether morphological evolution has proceeded faster than mtDNA evolution or whether widespread gene flow has obfuscated the relationship between mtDNA and morphology. Furthermore, future work will involve more-extensive sampling to further refine the hypotheses we have presented.

**Acknowledgements**

Financial support was provided by National Science Foundation grant BSR-8906621 to R.M.Z., Louisiana Board of Regents grant LEQSF no. 86-LBR-(048)-08 to R.M.Z., and Dayton-Wilkie Fellowships to A.J.F. Many people helped collect song sparrow
specimens and are acknowledged in Zink & Dittmann (1993). D. L. Dittmann extracted and purified mtDNA used in this study. We thank G. F. Barrowclough, D. M. Rand, J. D. Weckstein, T. Burke and an anonymous reviewer for helpful comments on the manuscript. A.J.F. would like to thank D. F. McKinney and R. Holzenthal for support during this study.

References


© 1998 Blackwell Science Ltd, Molecular Ecology, 7, 1303–1313


