EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 47 June, 1993 No. 3

Evolution, 47(3), 1993, pp. 717–729

GENE FLOW, REFUGIA, AND EVOLUTION OF GEOGRAPHIC VARIATION IN THE SONG SPARROW (*MELOSPIZA MELODIA*)

ROBERT M. ZINK[1] AND DONNA L. DITTMANN
Museum of Natural Science, Louisiana State University, Baton Rouge, Louisiana 70803

Abstract.—We surveyed mtDNA restriction-site variation in song sparrows taken from across their continental range. Despite marked geographic variation in size and plumage color, mtDNA variation was not geographically structured. Subspecies were not identifiable by mtDNA analysis. We suggest that postglaciation dispersal scattered mtDNA haplotypes across the continent, explaining the lack of mtDNA geographic patterns. Evolution of size and plumage coloration has probably proceeded faster than mtDNA evolution, leading to the well-structured continental pattern of morphological variation. We suggest that the nonordered geographic distribution of haplotypes reflects the recency of population establishment following completion of range expansion. Dispersal distance was estimated from the mtDNA data at 6.1 km per generation, an order of magnitude greater than that (0.3 km) estimated from demographic data. Island samples were not especially different from continental ones. Rooting the haplotype cladogram with a putative primitive haplotype identified Newfoundland and the Queen Charlotte Islands as potential sites of recent refugia. We question whether study of geographic variation in song sparrows leads to insights concerning speciation.

Key words.—Geographic variation, gene flow, *Melospiza melodia*, mitochondrial DNA, population structure, refugia, speciation.

Received April 20, 1992. Accepted August 28, 1992.

The song sparrow (*Melospiza melodia*), which ranges from Newfoundland to the Aleutian Islands to the Mexican Plateau (fig. 1), is a textbook example of geographic variation in vertebrates. Variation is extreme in size and color, ranging from 46-g dark birds that breed in the Aleutian Islands, to 20-g pale birds found at the Salton Sea in southern California. The most recent taxonomic statement (American Ornithologists' Union 1957) partitions this variation into 34 subspecies, making the song sparrow one of the most polytypic bird species in North America. When examined in detail, the pattern of variation is complex (Miller 1956). There are broad regions of phenotypic uniformity (eastern North America), regions with numerous subspecies (western North America), and regions with apparent microgeographic differences (e.g., San Francisco Bay; Marshall 1948). Some islands, such as Santa Barbara Island, have a distinctive subspecies, whereas other islands, such as Newfoundland, have song sparrows of a continental subspecies. It would seem that understanding the evolution of the marked geographic variation in song sparrows would offer important insights into the processes of geographic differentiation, and ultimately speciation (Mayr 1963).

The general pattern of phenotypic variation in the song sparrow reflects the geographic distribution of environmental heterogeneity in North

[1] Present address: Bell Museum of Natural History, Ecology Building, University of Minnesota, St. Paul, MN 55108.

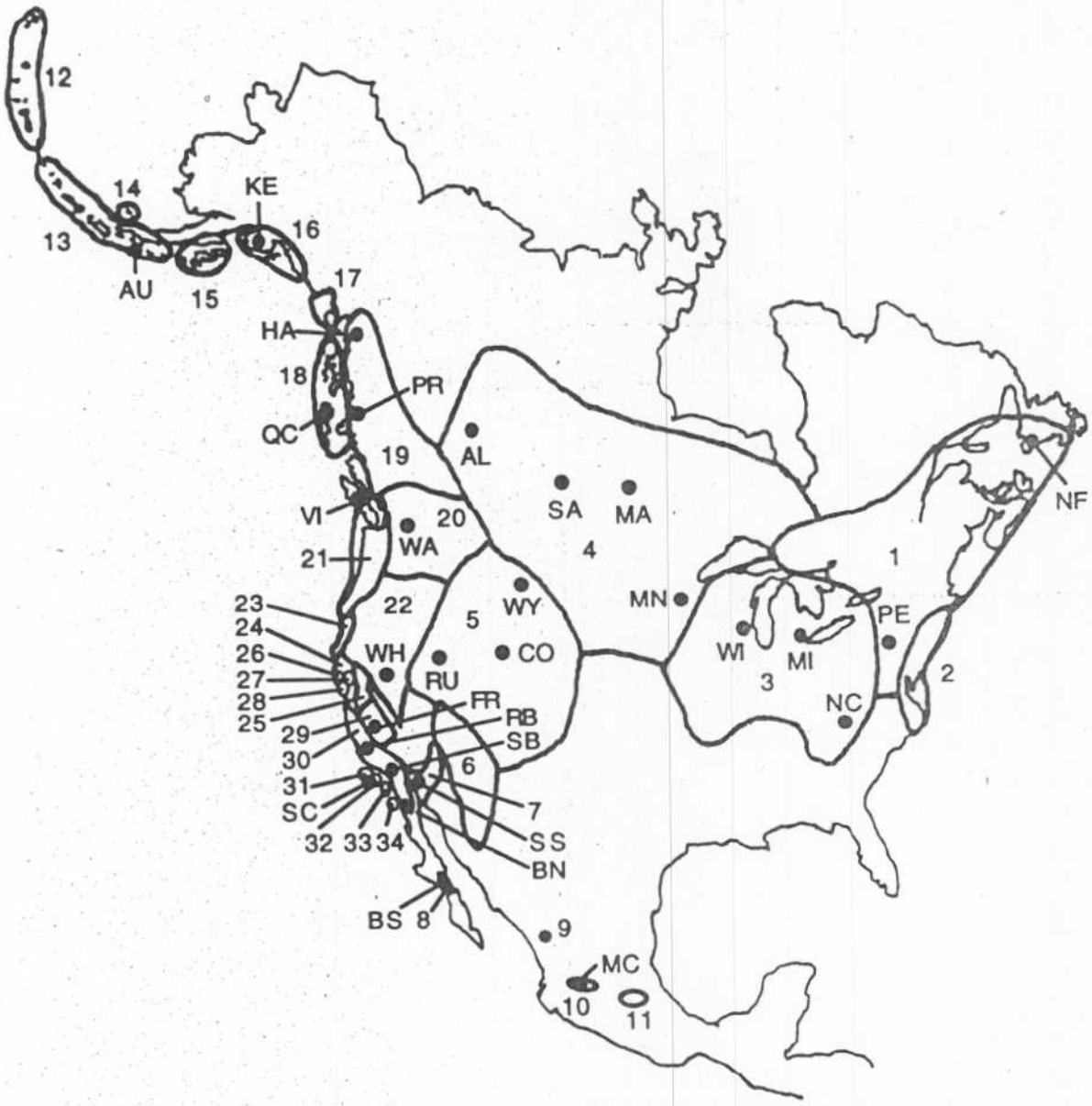

FIG. 1. Approximate breeding range of song sparrow showing geographic distribution of subspecies (1, *melodia*; 2, *atlantica*; 3, *euphonia*; 4, *juddi*; 5, *montana*; 6, *fallax*; 7, *saltonis*; 8, *rivularis*; 9, *goldmani*; 10, *adusta*; 11, *pectoralis*; 12, *maxima*; 13, *sanaka*; 14, *amaka*; 15, *insignis*; 16, *kenaiensis*; 17, *caurina*; 18, *rufina*; 19, *inexpectata*; 20, *merrilli*; 21, *morphna*; 22, *fisherella*; 23, *cleonensis*; 24, *gouldii*; 25, *mailliardi*; 26, *samuelis*; 27, *maxillaris*; 28, *pusillula*; 29, *heermanni*; 30, *cooperi*; 31, *micronyx*; 32, *clementae*; 33, *graminea*; 34, *coronatorum*). Dots indicate sample locales; two-letter codes correspond to those for sample sites in table 1.

America—more varied in the west. A typical explanation of this pattern of phenotypic variation is that it represents the effects of natural selection, fine-tuning populations to locally differing environments (Miller 1956; Mayr 1963). Indeed, at least within some populations, heritable variation in and natural selection on the traits that vary geographically have been demonstrated (Schluter and Smith 1986). Aldrich (1984) attempted to explain geographic patterns of morphological variation in song sparrows and concluded that ecological differences were more important than geographic isolation (distance) or vicariant events in causing "adaptive genetic differentiation." Aldrich's (1984) analytical design, however, did not permit him to distinguish between ecophenotypic and genetic mechanisms of generating geographic variation (Zink 1985). Alternative explanations for geographic variation might involve genetic and morphometric drift in

response to geographic isolation (Lynch 1989). Description of the geography of genetic variability ultimately forms the framework for inferring processes of geographic differentiation. If the marked morphological patterns of variation in song sparrows do not reflect a subdivided genetic population structure, then inferences derived from morphological comparisons might not reflect evolutionary processes, at least not to the degree typically assumed (Mayr 1963).

To describe the geography of genetic variation, many researchers have analyzed patterns of mitochondrial DNA (mtDNA) variation (Avise et al. 1987; Moritz et al. 1987; Dowling et al. 1990). Analyses of mtDNA have shown geographic structure in some species of birds (Shields and Wilson 1987a; Avise and Nelson 1989; Fleischer et al. 1991; Shields 1990; Zink 1991; Zink et al. 1991a; Degnan and Moritz 1993) but not all (Ball et al. 1988; Avise et al. 1990; Zink et al. 1991c; Moore et al. 1991). In contrast, nearly all allozymic surveys detected little geographic variation in north-temperate birds (Barrowclough 1983). Two previous analyses of mtDNA in song sparrows were more geographically limited than ours. Hare and Shields (1992) found mtDNA heterogeneity among song sparrows from the Aleutian Islands and the adjacent mainland, and another study (Zink 1991) found that mtDNA variation in western North America was not consistent with either geographical or subspecific boundaries.

In this study, we analyzed mtDNA variation in samples taken across the continental range, including representatives of 19 subspecies. Our aims were to determine levels and patterns of genetic differentiation, covariation of subspecies limits and mtDNA patterns, and the effects of isolation on island populations. Using outgroup comparisons, we constructed a hypothetical ancestral haplotype, which was used to root trees; the rooted trees were used to identify putative sites of refugia. We used a method proposed by Neigel et al. (1991) to estimate the magnitude of dispersal distance.

Materials and Methods

A total of 170 individuals from 29 locales representing at least 19 subspecies was studied (table 1). Specimens were collected during the breeding season, and tissue samples were preserved either in liquid nitrogen or in MSB-EDTA buffer (Lansman et al. 1981) kept at 4°C. From most specimens a study skin plus a partial skeleton, or a complete skeleton, were preserved; these are deposited at the Museum of Natural Science, Louisiana State University (other specimens are deposited at the Carnegie Museum, Pittsburgh; University of California, Santa Barbara; Burke Museum, Seattle; American Museum of Natural History, New York; North Carolina State Museum, Raleigh; University of Alaska, Fairbanks; and Museo de Zoologia "Alfonso L. Herrera," Universidad Autònoma de Mexico). Methods for isolation, purification, and restriction endonuclease digestion of mtDNA follow Lansman et al. (1981) and Dowling et al. (1990). For each individual, we recorded the presence or absence of restriction fragments produced by each of 17 endonucleases; each distinct pattern for each endonuclease was given a different letter. The sequence of letters from all restriction endonucleases constitutes an individual's composite haplotype. From the fragment profiles produced by each endonuclease, we inferred the presence or absence of restriction sites, which were analyzed for within- and among-locality variation. The matrix of presence/absence of restriction sites was used to compute the percentage of sequence divergence (p) between haplotypes following Nei and Li (1979). As an analogue to single-locus heterozygosity, we computed G (genotypic diversity of Nei 1987) as $(n/n-1)(1-\Sigma f_i^2)$, where f_i is the frequency of the ith mtDNA haplotype in a sample of n individuals. G gives the probability that random pairs of individuals surveyed have different mtDNA haplotypes. A computer program written by J. E. Neigel was used to estimate root–mean-square dispersal distance following the procedure outlined in Neigel et al. (1991). This procedure requires a phenogram based on p-values, a matrix of geographic distances among haplotypes, and generation length (we used 2 yr). Assuming 2% sequence divergence per million years (Shields and Wilson 1987b), the program estimates dispersal distance by calculating the geographic separation of haplotypes that had shared a common ancestor at various times (judged by the above calibration). To provide a quantitative measure of the distribution of genetic variability, we computed G_{ST} (Nei 1987) by considering haplotypes as alleles at a single locus (the mtDNA genome). To assess the significance of the G_{ST}, we compared it with 1000 random permutations of the geographic distribution of composite haplotypes (keeping the number of haplotypes per locality constant). To determine the geographic distribution of haplo-

TABLE 1. Sample sites, sample sizes, and subspecies of the song sparrow, *Melospiza melodia*. Sample sites shown in figure 1. Haplotype numbers refer to those in the Appendix, with numbers (in parentheses) of individuals possessing each haplotype if present in more than one individual. No haplotypes are numbered 31, 37, 38, 39, 43, or 44.

Locality (acronym)	*N*	Subspecies	Haplotype
1. Newfoundland (NF)	14	*melodia*	12(7), 13, 14, 15, 16, 17, 24, 25
2. Pennsylvania (PE)	7	*melodia*	2(4), 12, 18, 19
3. North Carolina (NC)	5	*euphonia*	2(2), 3, 11, 25
4. Wisconsin (WI)	2	*euphonia*	20, 21
5. Michigan (MI)	5	*euphonia*	2, 12, 22(2), 23
6. Minnesota (MN)	8	*euphonia/juddi*	2, 13, 21, 22, 26, 27, 28, 29
7. Manitoba (MA)	11	*juddi*	2(2), 12, 13, 18(2), 19, 26, 30(2), 32
8. Saskatchewan (SA)	5	*juddi*	13, 33, 34, 35, 36
9. Alberta (AL)	9	*juddi*	2(2), 12(2), 16, 19(2), 32, 34
10. Aleutian Is., Alaska (AU)	2	*sanaka*	9, 62
11. Kenai peninsula, Alaska (KE)	4	*kenaiensis*	4(2), 40, 41
12. southeast Alaska (HA)	8	*caurina/inexpectata*	2(2), 3, 5(3), 17, 42
13. Prince Rupert, B.C. (PR)	10	*inexpectata/rufina*	3(2), 5(5), 45, 46(2)
14. Queen Charlotte Is., B.C. (QC)	10	*rufina*	5(4), 59, 60(2), 61(2), 62
15. Vancouver Is., B.C. (VI)	5	*morphna*	3(2), 5(2), 61
16. Washington (WA)	7	*merrilli*	4(2), 5(4), 47
17. Wyoming (WY)	5	*montana*	2, 3, 4, 5, 10
18. Colorado (CO)	5	*montana*	3(2), 5(2), 8
19. Elko Co., Nev. (RU)	4	*montana*	3(2), 5, 8
20. Esmeralda Co., Nev. (WH)	4	*fisherella*	5, 6, 7, 8
21. Trinity Co., Calif. (YO)	3	*mailliardi*	3, 6, 9
22. Fresno Co., Calif. (FR)	2	*heermanni*	4, 5
23. San Bernardino Co., Calif. (SB)	4	*cooperi*	1, 2(2), 3
24. Santa Barbara Co., Calif. (RB)	4	*cooperi*	3, 9, 42, 48
25. Baja California Norte, Mexico (BN)	3	*cooperi*	6, 53, 54
26. Santa Cruz Is., Calif. (SC)	7	*clementae*	2, 5, 6, 8(2), 49, 50
27. Imperial Co., Calif. (SS)	5	*saltonis*	17, 51, 52(3)
28. Baja California Sur, Mexico (BS)	3	*rivularis*	4, 5(2)
29. Michoacan, Mexico (MC)	9	*adusta*	55, 56(3), 57(2), 58(3)
TOTAL	170		

types, we constructed several branching diagrams. We used the computer program NTSYS, written by F. J. Rohlf, to produce a UPGMA phenogram of *p*-values, and the computer program HENNIG86 (Farris 1989), to infer a network among haplotypes using the principle of maximum parsimony; the option "m*;bb" was used to provide a heuristic search for the most parsimonious topology.

For each restriction endonuclease, each fragment profile was compared against those in related species (J. C. Avise unpubl. data; Zink et al. 1991b) to construct a putative primitive haplotype (Appendix); it is generally acknowledged that multiple outgroups are better than a single species (Maddison et al. 1984). The following species were used: swamp sparrow (*Melospiza georgiana*), Lincoln's sparrow (*M. lincolnii*), dark-eyed junco (*Junco hyemalis*), white-crowned sparrow (*Zonotrichia leucophrys*), golden-crowned sparrow (*Z. atricapilla*), white-throated sparrow (*Z. albicollis*), Harris' sparrow (*Z. querula*), and rufous-collared sparrow (*Z. capensis*). For each endonuclease, whichever restriction-fragment pattern in the song sparrow was observed in another species, or was the fewest steps from one, was regarded as ancestral, and in this way a composite haplotype was constructed. For restriction endonucleases that produced a single pattern in song sparrows, that pattern was assumed to be ancestral (this would bias the distance from the ancestral haplotype to extant haplotypes, but not influence the identification of basal haplotypes in phylogenetic analysis). A restriction-site map for related species would be preferable, but limited resources prevented its construction. The ancestral haplotype was used to root the tree of extant haplotypes and to infer

where the basal haplotype(s) was(were) located, thereby suggesting sites of refugia.

RESULTS

Seven endonucleases (*Ava* I, *Bgl* I, *Bgl* II, *Eco*R I, *Hind* III, *Kpn* I, and *Xba* I) produced only a single restriction-fragment pattern. The other 10 endonucleases each yielded from two to eight patterns. On the average, 96 restriction fragments were scored for each individual, representing approximately 500 bases of endonuclease recognition sequence (3.0% of the mitochondrial genome). In total, 105 restriction sites and 56 haplotypes were detected among the 170 individuals surveyed (Appendix). The genotypic diversity (G) across the species was 0.97, which shows that the probability of randomly choosing two individuals with the same haplotype is exceedingly low. Sequence divergence (p) among the 56 haplotypes averaged 0.53% ± 0.22 (SD; n = 1540) and ranged from 0.01% to 1.23%; the average p-value among the 170 individuals was 0.40% ± 0.24 (SD; n = 14,365); these data are available from the senior author. The degree of mtDNA haplotype diversity was variable among samples: in the samples from Minnesota, Wyoming, and Esmerelda, each bird had a different haplotype, whereas the Washington and Michoacan (Mexico) samples had relatively few haplotypes (table 1). Approximately 90% of estimated mtDNA distances among individuals (fig. 2) are less than 0.7%, and the distribution does not exhibit multiple peaks, which are expected with certain types of phylogeographic structure (Avise and Ball 1991). The G_{ST} value, 0.09, was not significantly different from a random geographic array of haplotypes (P = 0.18).

Inspection of the distribution of haplotypes (table 1) nonetheless reveals some geographic differentiation. Several pairwise comparisons of samples had no haplotypes in common (e.g., samples 1 versus 11, 12, or 13). In contrast, haplotypes 3 and 5 were widely distributed in western North America and haplotype 2 was distributed widely throughout the range (table 1). Twenty haplotypes occurred in more than one site. The UPGMA phenogram (fig. 3) exhibits some geographic structure, in that three of four haplotypes from southern Mexico are most similar, several haplotypes from Newfoundland cluster together, and there is a cluster of haplotypes (numbers 16 to 29 at bottom of figure) from the northern part of the range. Overall, however, haplotypes from the same locale show little tendency to be genetically most similar. Over 100 equally parsimonious cladograms were detected (length 60, consistency index = 0.43, retention index = 0.72). A strict consensus tree (not shown) revealed no structure except for the following groups of haplotypes: 13 and 14, 22 and 27, 12 and 24, 4 and 49, 55 and 58, 59 and 61, and 40, 41, and 50. The fully resolved cladograms (not shown) resembled the phenogram in the degree to which they exhibited geographic structure. To determine if a geographic "signal" was present in the data set, the successive approximation approach to character weighting was used (Farris 1969). The 11 resultant trees had a much higher consistency index (0.93) and retention index (0.97), and an example (fig. 4) shows similar results to the phenogram (evidence of a northern cluster, clusters from Newfoundland and southern Mexico). A strict consensus tree (not shown) of these 11 trees, rooted at the hypothetical ancestral haplotype, portrayed haplotypes from Newfoundland (12, 15, 24, 25), Saskatchewan (34), and the Queen Charlotte Islands (59, 61) as basal to the remainder.

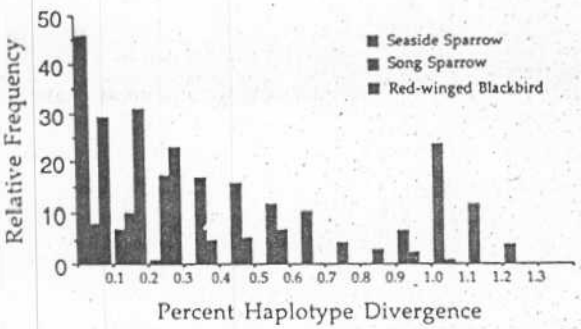

FIG. 2. Distribution of pairwise mtDNA genetic distances (p) among individuals for song sparrow, seaside sparrow (Avise and Nelson 1989), and red-winged blackbird (Ball et al. 1988).

The estimate of dispersal distance per generation was 6.1 km. This estimate was derived from the youngest (e.g., most similar) class of mtDNA lineages, which Neigel (pers. comm. 1992) considers most reliable as an indicator of gene flow, because there has been less time for geographic barriers to constrain dispersal, relative to older haplotype lineages.

DISCUSSION

Geography of mtDNA Variation across the Continent

MtDNA comparisons (fig. 2) reveal a richer picture of avian genetic population structure than that offered by allozyme analyses, which sug-

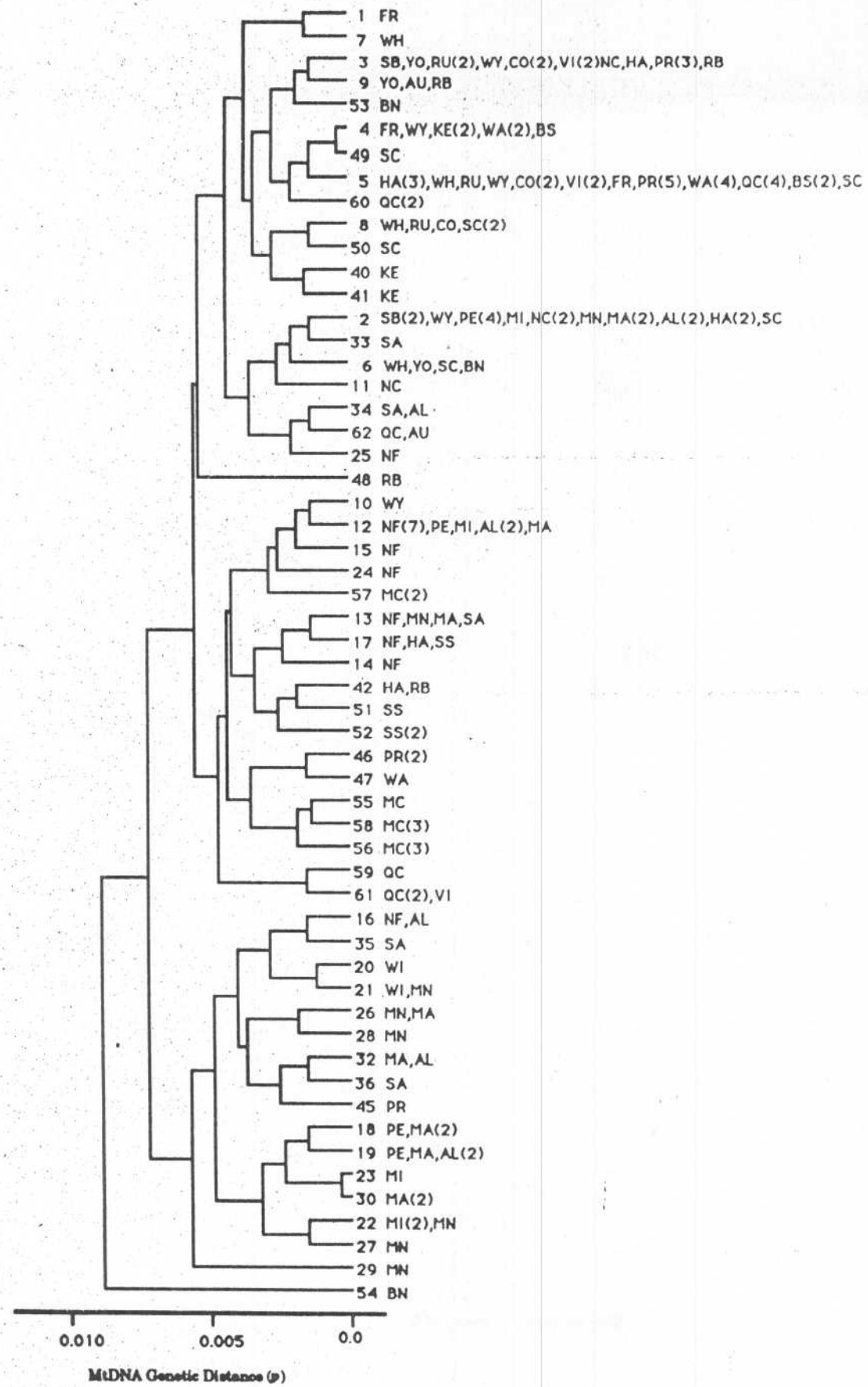

FIG. 3. UPGMA phenogram based on mtDNA genetic distances (p). Two-letter codes refer to sample sites in figure 1.

gested that most species were essentially panmictic (Barrowclough 1983). The song sparrow appears intermediate in its degree of mtDNA differentiation (fig. 2) between passerine birds such as the red-winged blackbird (*Agelaius phoeniceus*; Ball et al. 1988) and the common grackle (*Quiscalus quiscula*; Zink et al. 1991c), which exhibit little or no geographic structure, and species with marked mtDNA variation across their ranges such as the seaside sparrow (*Ammodramus maritimus*; Avise and Nelson 1989). The average p-value among individual song sparrows, 0.40, is greater than in most other avian species surveyed to date (Ball et al. 1988; Zink 1991; Zink et al. 1991c). However, the level of haplotype differentiation does not necessarily indicate the nature of geographic structure, because a species could have relatively old haplotype lineages (and hence high p-values) that are not geographically organized (Avise et al. 1987). Although the insignificant G_{ST} value for the song sparrow suggests no population structure, use of this type of statistic is possibly inappropriate (Neigel and Avise unpubl. data), because it ignores information on haplotype relationships (Slatkin 1989; Slatkin and Maddison 1989), is based on a single linkage group, assumes that mutations are infrequent, and our sample sizes per locality are small.

We think that the phylogeographic pattern (fig. 4) of haplotype relationships offers a more appropriate estimate of geographic population structure than G_{ST}. In the song sparrow, two tiers of history seem evident in the current distribution of mtDNA haplotypes. First, the geographic pattern of haplotypes does not reflect closely the geographic proximity among the samples. That is, phylogenetically most closely related haplotypes usually are not found exclusively at the same site, which is consistent with a lack of population structure or high gene flow (Slatkin and Maddison 1989). This distribution of haplotypes (figs. 3, 4) suggests either (1) recent colonization of the current range from a less widespread ancestral population, with insufficient time for geographic differentiation; (2) substantial levels of current gene flow (see below); or (3) convergent/parallel site gains or losses (which we do not consider further). Hypothesis 1 is consistent with the idea that gene flow was high as song sparrows colonized newly opened habitat following glacial retreat. Such a "pulse" of dispersal following glacial retreat was suggested by Hewitt (1988) for grasshoppers in the Pyrenees.

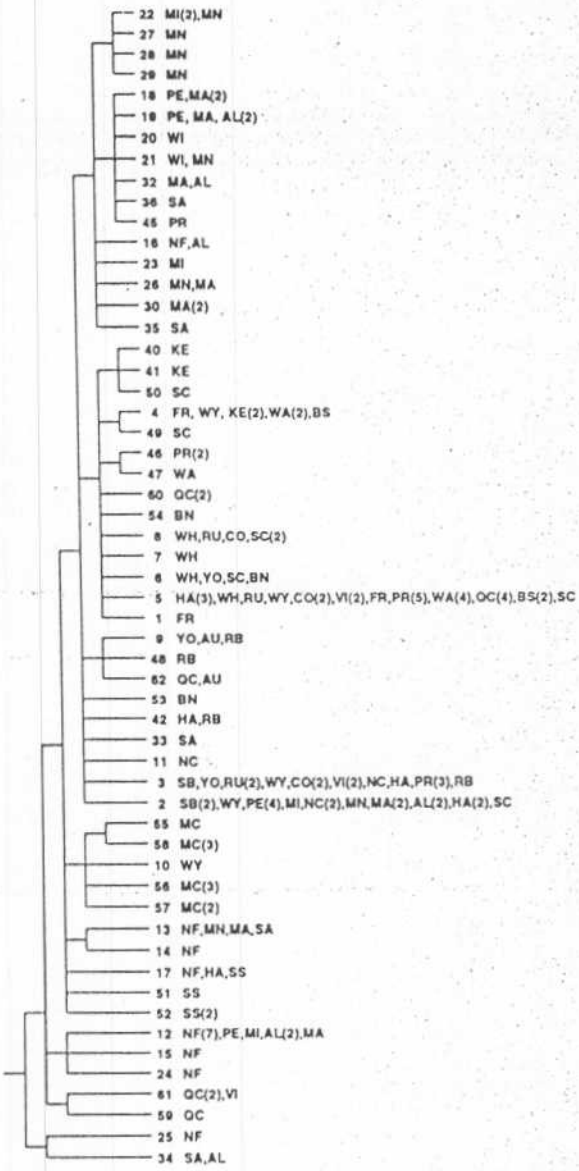

FIG. 4. Cladogram of mtDNA restriction haplotypes based on variation in restriction sites; rooted at ancestral haplotype. Two-letter codes refer to sample sites in figure 1.

The second tier of history is suggested by the geographical localization of some haplotypes, such as those from southern Mexico (Michoacan), Newfoundland, and the northern part of the breeding distribution (figs. 3, 4). These localizations of haplotypes might be due to isolation from ongoing gene flow throughout the range, or they might represent incipient geographic differentiation. We favor the latter hypothesis because although some haplotypes (2, 3, 5) are widely distributed, several samples share no haplotypes (table 1), which argues against extensive current gene flow. Thus, the two tiers suggested

by mtDNA data are (1) postglacial range expansion that scattered mtDNA haplotypes geographically, followed by (2) evolution in situ, the results of which are only beginning to be detectable as mtDNA phylogeographic patterns (figs. 3, 4).

Subspecies and mtDNA Variation

Because the song sparrow exhibits marked geographic variation in phenotypes, we expected geographically ordered mtDNA patterns. Bernatchez and Dodson (1990) found four mtDNA clusters in lake whitefish (*Coregonus clupeaformis*), and Zink (in press) found four haplotype groups of the fox sparrow (*Passerella iliaca*), across roughly the same geographic area as surveyed here. In general, other terrestrial vertebrates show mtDNA phylogeographic structure (Avise et al. 1987). The absence of mtDNA patterns that mirror the marked geographic differences in size and coloration (i.e., subspecies limits), or phylogeographic breaks found in other species, suggests at least two hypotheses. First, mtDNA and nuclear genomes might be exchanged by dispersal, and geographic variation in morphology could result from strong natural selection or environmental induction (e.g., James 1983). Much of the geographic variation in song sparrows involves overall size and intensity of coloration (Aldrich 1984), attributes that could be affected by the environment (James 1983) or by relatively minor genetic alterations.

A second hypothesis stems from our interpretation of two tiers of history, and suggests that plumage and size evolved faster, owing to their polygenic nature, than mtDNA monophyly (Neigel and Avise 1986). Geographic patterns in size and plumage coloration likely evolved since range expansion was completed (and gene flow subsided following completion of range expansion). MtDNA evolution could eventually "catch up" with morphological variation, but at present, the mtDNA gene tree exhibits paraphyly with respect to subspecies limits. Other studies have found discordance between avian mtDNA gene trees and subspecies limits (Ball et al. 1988; Zink et al. 1991c; Avise and Nelson 1989). In these studies, it appears that geographic structuring in morphology proceeds faster than in mtDNA haplotypes. As noted by Avise and Ball (1991), if phenotypic evolution proceeds faster than mtDNA evolution, subspecies might actually represent evolutionary units, although no mtDNA markers are associated with them.

Because song sparrows, like other bird species, are larger and darker in the Pacific Northwest and Aleutian Islands than elsewhere (Zink and Remsen 1986), it is tempting to suggest natural selection as the causal agent for geographic patterns of morphological variation in song sparrows. Natural selection has been demonstrated on phenotypic traits in local populations of song sparrows (Schluter and Smith 1986), and it is possible that broad-scale morphological differences are genetically based; cross-fostering experiments should be pursued (James 1983).

Gene Flow in the Song Sparrow

A basic issue in studies of geographic variation is the estimation of levels of gene flow (Slatkin 1987). A great disparity exists among estimates of avian dispersal made by different investigators. Barrowclough (1978, 1980) suggested that dispersal in most passerine birds is less than 1.0 km/generation, whereas for some species, Moore and Dolbeer (1989) suggested values of over 100 km. Estimates of gene flow based on allozyme frequencies indicated that most avian populations exchange more than one individual per generation (Zink and Remsen 1986), although the data were not used to estimate dispersal distance per se.

Hypotheses noted above for the evolution of morphology and mtDNA in song sparrows hinge on the magnitude of gene flow. The dispersal distance of 6.1 km/generation estimated from Neigel et al.'s (1991) nonequilibrium approach is an order of magnitude greater than that, 334 m, obtained by Barrowclough (1980) from demographic data on a banded population in Ohio. Neigel et al. (1991) found that their mtDNA estimate was congruent with mark-recapture estimates in *Peromyscus maniculatus*. Perhaps the discrepancy in the song sparrow is caused by different historical and spatial scales. The Ohio data come from a single population in which the habitat is fairly continuous. Our mtDNA dispersal estimate is based on the youngest mtDNA lineages from throughout the range. In the western United States and Mexico, patches of suitable song sparrow habitat are often many kilometers apart. Also, song sparrows have colonized vast areas of North America following the retreat of the last glaciers. These considerations render the estimate of gene flow, based on haplotype distributions, an average of rates of dispersal over several thousand years and across a topographically varied environment. Dispersal distances in older mtDNA lineages are lower (≤ 2.8 km), but these lineages might not be useful indicators of gene flow (Neigel and Avise unpubl. data). Also,

gene flow between some populations might be less than the continent-wide average.

Thus, the magnitude and significance of gene flow in the song sparrow is difficult to judge. The mtDNA value is of the same order of magnitude as that found in other widely distributed avian species, both with and without morphological geographic variation (Neigel and Avise unpubl. data; Zink unpubl. data). Although we doubt that the song sparrow is panmictic over its range, mark-recapture estimates or other genetic estimates are needed to establish levels of gene flow, and distinguish between the hypotheses presented above for the apparent discrepancy between mtDNA and morphological variation. We suspect the existence of a significant genetic component underlying morphological patterns.

Haplotype Evolution on Islands

Islands are often considered natural laboratories for studies of evolution; islands enforce isolation and might accelerate differentiation of populations. Allozyme evidence indicates some differentiation in the California quail (*Callipepla californica*; Zink et al. 1987) and western flycatcher (*Empidonax difficilis*; Johnson and Marten 1988) on the California channel islands. Ashley and Wills (1987) found that samples of *Peromyscus maniculatus* from the California Channel Islands contained many haplotypes not found on the adjacent mainland. The mtDNA genetic diversity in sparrow samples from four major islands [Vancouver (0.64), Queen Charlottes (0.74), Newfoundland (0.71), Santa Cruz (0.82)] averaged below the overall average of 0.97, suggesting some reduction in variability of island populations. Our data, however, indicate few effects of island isolation on mtDNA differentiation, although in most cladograms, most haplotypes from Newfoundland clustered together. However, haplotypes from Vancouver Island and the Queen Charlotte Islands were widely dispersed in the cladograms, unlike the results described for *P. maniculatus* (Ashley and Wills 1987). Thus, the island sparrow populations have not been isolated long enough to permit mtDNA divergence, there is gene exchange with mainland areas, or restriction site analysis is insufficiently sensitive.

MtDNA Gene Trees and the Search for Refugia

Pleistocene glaciations influenced many North American taxa, and current distributions must have stabilized recently (Pielou 1991). Identification of some haplotypes from Newfoundland and the Queen Charlotte Islands as "basal" suggests that these locales were refugia. If true, mtDNA genomes of modern song sparrows descended from groups of females that survived the Pleistocene somewhere in northeastern North America, perhaps on lands now inundated, and in the west on the Queen Charlotte Islands. It would be surprising if a southern refuge did not also exist. Geological evidence shows several large areas off the modern-day coast of eastern North America that could have supported bird populations at the maximum glacial extent, including Georgia Banks (off the coast of Newfoundland), Glacial Sable Island, and the Grand Banks (Pielou 1991). Pielou (1991) noted that the Queen Charlotte Islands functioned as a refugium for plants, and today several endemic subspecies of birds occur there (Zink and Remsen 1986).

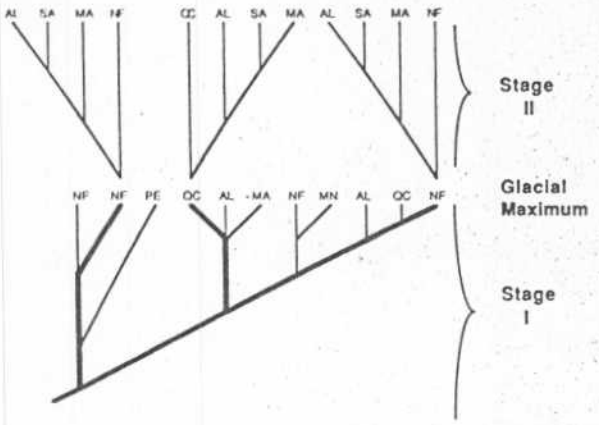

FIG. 5. Hypothetical cladogram of haplotypes (Stage I), followed by extinction of all haplotypes except two in Newfoundland (NF) and one in the Queen Charlotte Islands (QC), followed by expansion westward and eastward from these two refugia (Stage II). See text. SA = Saskatchewan, AL = Alberta, MN = Minnesota, PE = Pennsylvania, and MA = Manitoba. The thick lines show relationships of extant haplotypes prior to the glacial maximum.

To illustrate the use of rooted haplotype phylogenies to infer sites of refugia, we present a simple model (fig. 5) stylized to reflect our study of the song sparrow. We first envision a set of haplotypes that existed prior to the last glaciation (Stage I). Because of the potential effects of many glacial advances and retreats (Pielou 1991), we show each locality possessing haplotypes that are not necessarily nearest phylogenetic relatives. The (hypothetical) effect of the last glacial maximum caused the extinction of haplotypes at all locations except Newfoundland and the Queen Charlotte Islands. These haplotypes were thus the founders that recolonized North America (Stage II). If colonization proceeded from a given lo-

cality, there should be a monotonic relationship between haplotype and geographic distance—the most "derived" haplotypes are farthest from Newfoundland or the Queen Charlottes. Thorpe (1984) additionally predicted that there should be an increase in anagenesis towards the root of the tree (the refuge site). Extinction of a lineage in Newfoundland, for example, would mean that the basal haplotype of a subcluster would be some point west of Newfoundland.

Several predictions emerge from this model. The current distribution of haplotypes could be a conglomerate of lineages with different histories. Haplotypes from the same locality are not necessarily nearest relatives. In figure 5, the thick lines show how haplotypes would trace their ancestry to times prior to the glacial maximum [exactly as occurs during lineage sorting (Neigel and Avise 1986)]. For example, haplotypes from Alberta (AL) would occur terminally in three distinct parts of the overall haplotype cladogram, but would tend to be most closely related to haplotypes from other locations. Another aspect of this scenario concerns ages of haplotypes. Haplotypes that evolved from those that survived in a refugium should be no older than that refuge. The "source" haplotypes in the refugia, however, can be much older, as the thick lines in figure 5 suggest.

This model suggests an explanation for the absence of geographic patterning in song sparrow haplotypes (figs. 3, 4). However, it is difficult to reconcile the data with any simple model. The maximum difference between any pair of haplotypes, 1.2%, translates into ca. 500,000 yr (Shields and Wilson 1987b), well before the last glacial maximum, 18,000 yr ago (Pielou 1991). Although the identification of a "derived" clade of haplotypes that is northern in occurrence is consistent with the general model of dispersal from a coastal or southern refuge, these northern haplotypes trace to 0.5% divergence (fig. 3), or 250,000 years ago, greater than that predicted for Newfoundland or Queen Charlotte Islands refugia. Unknown of course is whether the calibration of 2% per million years is appropriate for these birds. We suggest that other researchers search for evidence of coastal refugia in their haplotype cladograms.

Evolutionary Significance of Geographic Variation in Song Sparrows

It is possible that the haplotypes in the song sparrow do not form a monophyletic group relative to other species. However, the most divergent haplotypes ($P = 0.0123$), 54 versus 18 and 54 versus 27, are less than 50% of the distance ($P = 0.026$; Kessler and Avise, 1985) from the song sparrow to its nearest relative [either the swamp sparrow (*Melospiza georgiana*) or Lincoln's sparrow (*M. lincolnii*)]. Therefore, the oldest detected haplotypes in extant song sparrows trace their common ancestry to a time well after the speciation event 1.2 million yr ago that produced the song sparrow.

Because of the marked geographic variation in morphology, one might consider the song sparrow as a case of speciation in action. Indeed, the mtDNA data (fig. 2) suggest that the song sparrow is at a stage of divergence intermediate between populations and species. If the song sparrow represented incipient speciation, we could infer that the rapid evolution, relative to mtDNA, of size and color differences is a first step in speciation. However, the congeners of the song sparrow are not larger, or darker or lighter than song sparrows, as might be expected if speciation is an extension of geographic differentiation (Cracraft 1989). Song sparrow subspecies differ morphometrically more among themselves than do continental song sparrows from their congeners (Zink 1982; Zink unpubl. data). Thus, it is not clear that the evolution of geographic variation in the song sparrow is representative of speciation processes that occurred in the origin of the three species in the genus. That is, species differences tend to be qualitative, and geographic differences quantitative; whether the two become connected in evolutionary time is, of course, debatable (Goldschmidt 1940; Levinton 1988).

Acknowledgments

We are grateful to the following persons for assistance in collecting specimens: G. F. Barrowclough, S. W. Cardiff, J. A. Gerwin, J. A. Marten, F. J. Pitocchelli, S. J. Hackett, K. J. Burns, D. S. Wood, S. A. Rohwer, C. S. Wood, M. P. Hare, M. A. Holmgren, and G. Voelker. We are especially grateful to P. Escalante for the samples of Mexican song sparrows. We thank J. E. Neigel for providing the computer program for estimating dispersal distance and for insight into the nature of gene trees. For assistance with collecting permits we thank G. Dick and B. McGillivray. The assistance of J. McIlhenny is acknowledged. Financial support was provided by National Science Foundation grant BSR-8906621 and Louisiana Board of Regents grant

LEQSF #86–LBR–(048)–08 to R.M.Z. For useful discussions relating to this work we thank J. C. Avise, G. F. Barrowclough, J. M. Bates, J. E. Neigel, D. P. Pashley, J. V. Remsen, and M. S. Hafner. J. M. Bates, S. Degnan, S. J. Hackett, M. S. Hafner, J. V. Remsen, and S. J. Weller provided helpful comments on the manuscript. D. R. Reynolds prepared figure 1.

Literature Cited

Aldrich, J. W. 1984. Ecogeographical variation in size and proportions of song sparrows (*Melospiza melodia*). Ornithological Monographs 35. American Ornithologists' Union, Washington, D.C.

American Ornithologists' Union. 1957. Check-list of North American birds, 5th ed. American Ornithologists' Union, Washington, D.C.

Ashley, M., and C. Wills. 1987. Analysis of mitochondrial DNA polymorphisms among Channel Island deer mice. Evolution 41:854–863.

Avise, J. C., and R. M. Ball, Jr. 1991. Mitochondrial DNA and avian microevolution. Acta Congressus Internationalis Ornithologici XX:514–524.

Avise, J. C., and W. S. Nelson. 1989. Molecular genetic relationships of the extinct dusky seaside sparrow. Science 243:646–648.

Avise, J. C., C. D. Ankney, and W. S. Nelson. 1990. Mitochondrial gene trees and the evolutionary relationship of mallard and black ducks. Evolution 44:1109–1119.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reed, and N. C. Saunders. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18:489–522.

Ball, R. M., Jr., S. Freeman, F. C. James, E. Bermingham, and J. C. Avise. 1988. Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. Proceedings of the National Academy of Sciences, USA 85:1558–1562.

Barrowclough, G. F. 1978. Sampling bias in dispersal studies based on finite area. Bird-Banding 49:333–341.

———. 1980. Gene flow, effective population sizes, and genetic variance components in birds. Evolution 34:789–798.

———. 1983. Biochemical studies of microevolutionary processes. Pp. 223–261 *in* A. H. Brush and G. A. Clark, Jr., eds. Perspectives in ornithology. Cambridge University Press, New York.

Bernatchez, L., and J. J. Dodson. 1990. Allopatric origin of sympatric populations of Lake Whitefish (*Coregonus clupeaformis*) as revealed by mitochondrial-DNA restriction analysis. Evolution 44:1263–1271.

Cracraft, J. 1989. Speciation and its ontology: The empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. Pp. 28–59 *in* D. Otte and J. A. Endler, eds. Speciation and its consequences. Sinauer, Sunderland, Mass.

Degnan, S. M., and C. Moritz. 1993. Phylogeography of mitochondrial DNA in two species of white-eye in Australia. Auk. *In press.*

Dowling, T. E., C. Moritz, and J. Palmer. 1990. Nucleic acids II. Restriction site analysis. Pp. 250–319 *in* D. M. Hillis and C. Moritz, eds. Molecular systematics. Sinauer, Sunderland, Mass.

Farris, J. S. 1969. A successive approximations approach to character weighting. Systematic Zoology 18:374–385.

———. 1989. Hennig86 reference manual. Published by the author.

Fleischer, R. C., S. I. Rothstein, and L. S. Miller. 1991. Mitochondrial DNA variation indicates gene flow across a zone of known secondary contact between two subspecies of the brown-headed cowbird. Condor 93:185–189.

Goldschmidt, R. 1940. The material basis of evolution. Yale University Press, New Haven, Conn.

Hare, M. P., and G. F. Shields. 1992. Mitochondrial DNA variation in the polytypic Alaska song sparrow. Auk. 109:126–132.

Hewitt, G. M. 1988. Hybrid zones—natural laboratories for evolutionary studies. Trends in Ecology and Evolution 3:158–166.

James, F. C. 1983. Environmental component of morphological differences in birds. Science 221:184–187.

Johnson, N. K., and J. A. Marten. 1988. Evolutionary genetics of flycatchers. II. Differentiation in the *Empidonax difficilis* complex. Auk 105:177–191.

Kessler, L. G., and J. C. Avise. 1985. A comparative description of mitochondrial DNA differentiation in selected avian and other vertebrate genera. Molecular Biology and Evolution 2:109–125.

Lansman, R. A., R. O. Shade, J. F. Shapira, and J. C. Avise. 1981. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations III. Techniques and potential applications. Journal of Molecular Evolution 17:214–226.

Levinton, J. 1988. Genetics, paleontology, and macroevolution. Cambridge University Press, Cambridge.

Lynch, M. 1989. Phylogenetic hypotheses under the assumption of neutral quantitative-genetic variation. Evolution 43:1–17.

Maddison, W. P., M. J. Donoghue, and D. R. Maddison. 1984. Outgroup analysis and parsimony. Systematic Zoology 33:83–103.

Marshall, J. T., Jr. 1948. Ecologic races of song sparrow in the San Francisco Bay region, Pt. I. Habitat and abundance. Condor 50:193–215.

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge, Mass.

Miller, A. H. 1956. Ecological factors that accelerate formation of races and species of terrestrial vertebrates. Evolution 10:262–277.

Moore, W. S., and R. A. Dolbeer. 1989. The use of banding recovery data to estimate dispersal rates and gene flow in avian species: Case studies in the red-winged blackbird and common grackle. Condor 91:242–253.

Moore, W. S., J. H. Graham, and J. T. Price. 1991. Mitochondrial DNA variation in the northern flicker (*Colaptes auratus,* Aves). Molecular Biology and Evolution 8:327–344.

Moritz, C., T. E. Dowling, and W. M. Brown. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annual Review of Ecology and Systematics 18:269–292.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nei, M., and W. H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences, USA 76:5269–5273.

Neigel, J. E., and J. C. Avise. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pp. 515–534 *in* E. Nevo and S. Karlin, eds. Evolutionary processes and theory. Academic Press, New York.

———. 1993. Application of a random-walk model to geographic distributions in animal mitochondrial DNA variation. Genetics. *In press.*

Neigel, J. E., R. M. Ball, Jr., and J. C. Avise. 1991. Estimation of single generation migration distances from geographic variation in animal mitochondrial DNA. Evolution 45:423–432.

Pielou, E. C. 1991. After the ice age: The return of life to glaciated North America. The University of Chicago Press, Chicago.

Schluter, D., and J.N.M. Smith. 1986. Natural selection on beak and body size in the song sparrow. Evolution 40:221–231.

Shields, G. F. 1990. Analysis of mitochondrial DNA of Pacific black brant (*Branta bernicla nigricans*). Auk 107:620–623.

Shields, G. F., and A. C. Wilson. 1987a. Subspecies of the Canada goose (*Branta canadensis*) have distinct mitochondrial DNAs. Evolution 41:662–666.

———. 1987b. Calibration of mitochondrial DNA evolution in geese. Journal of Molecular Evolution 24:212–217.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science 236:787–792.

———. 1989. Detecting small amounts of gene flow from phylogenies of alleles. Genetics 121:609–612.

Slatkin, M., and W. P. Maddison. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123:603–613.

Thorpe, R. S. 1984. Primary and secondary transition zones in speciation and population differentiation: a phylogenetic analysis of range expansion. Evolution 38:233–243.

Zink, R. M. 1982. Patterns of genic and morphologic variation among sparrows in the genera *Zonotrichia, Melospiza, Junco,* and *Passerella.* Auk 99:632–649.

———. 1985. Review of "Ecogeographical variation in size and proportions of song sparrows (*Melospiza melodia*)." Auk 102:913–914.

———. 1991. The geography of mitochondrial DNA variation in two sympatric sparrows. Evolution 45: 329–339.

———. 1993. The geography of mitochondrial DNA variation, population structure, and hybridization in the fox sparrow (*Passerella iliaca*). Evolution 47.

Zink, R. M., D. L. Dittmann, S. W. Cardiff, and J. D. Rising. 1991a. Mitochondrial DNA variation and the taxonomic status of the large-billed savannah sparrow. Condor 93:1016–1019.

Zink, R. M., D. L. Dittmann, and W. L. Rootes. 1991b. Mitochondrial DNA variation and the phylogeny of *Zonotrichia.* Auk 108:578–584.

Zink, R. M., D. F. Lott, and D. W. Anderson. 1987. Genetic variation, population structure, and evolution of California quail. Condor 89:395–405.

Zink, R. M., and J. V. Remsen, Jr. 1986. Evolutionary processes and patterns of geographic variation in birds. Pp. 1–69 *in* R. F. Johnston, ed. Current ornithology, vol. 4. Plenum, New York.

Zink, R. M., W. L. Rootes, and D. L. Dittmann. 1991c. Mitochondrial DNA variation, population structure, and evolution of the common grackle (*Quiscalus quiscula*). Condor 93:318–329.

Corresponding Editor: D. Schluter

Appendix

Haplotypes for song sparrows. Letters refer to restriction fragment profiles (presence/absence matrix of restriction sites is available from senior author). Sequence of restriction endonucleases: *Ava* II, *Bam*H I, *Ban* II, *Bcl* I, *Hin*c II, *Msp* I, *Nci* I, *Nde* I, *Pvu* II, *Stu* I, *Ava* I, *Bgl* I, *Bgl* II, *Eco*R I, *Hin*d III, *Kpn* I, *Xba* I. A dot signifies the "A" pattern. There are no haplotypes numbered 31, 37, 38, 39, 43, or 44.

No.	Fragment profile	No.	Fragment profile	No.	Fragment profile
1	C.D..BB.....	20	D.C..BB..B.....	45	B.G.....
2	B.D.....	21	D.C...B.....	46	.....
3	D.D.....	22	D.E..BB.....	47	C.....
4	C.D....B.....	23	D.F.BBB.....	48	D.D..CBB.....
5	C.D.....	24	D..B.B.B.....	49	..D....B.....
6	..D.....	25	D.DB...B.....	50	CBD...B.....
7	C.D..B.....	26	B.F.....	51	D.....B.....
8	C.D...B.....	27	B.E..BB.....	52	D....BB.....
9	D.D....B.....	28	B.E.....	53	D.D....C.....
10	D.....B.....	29	H.E.....	54	C.D..DC.....
11	B.B.....	30	D.F..BB.....	55	B.....B.....
12	D..B...B.....	32	B.C.....	56	B.....B.....
13	D..C.....	33	B.D.....B.....	57	I.....B.....
14	D..C.B.....	34	B.DB...B.....	58	B.....B.B.....
15	DB.B...B.....	35	D.F...B.....	59	B..B..C.....
16	D.F.....	36	B.C...B.....	60	C.D.....C.....
17	D.....	40	CBD.....	61	D..B..C.....
18	B.C..BB.....	41	EBD.....	62	B.D....B.....
19	D.C..BB.....	42	D.D...B.....		