SUBTLE EDGE-OF-RANGE GENETIC STRUCTURING IN TRANSCONTINENTALLY DISTRIBUTED NORTH AMERICAN TREE SWALLOWS

Laura M. Stenzler1,5, Christopher A. Makarewich1, Aurélie Coulon2,3, Daniel R. Ardia4, Irby J. Lovette1,2, and David W. Winkler2

1 Fuller Evolutionary Biology Program, Laboratory of Ornithology, Cornell University, Ithaca, NY 14850
2 Museum of Vertebrates and Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853
3 UMR 7179 Muséum National d’Histoire Naturelle/Centre National de la Recherche Scientifique, 1 avenue du Petit Château, 91800 Brunoy, France
4 Department of Biology, Franklin and Marshall College, Lancaster, PA 17604

Abstract. Understanding how genetic variation in the Tree Swallow (Tachycineta bicolor) is geographically structured is informative because this broadly distributed North American bird is increasingly used as a model for studies of mating systems, life-history traits, and physiology. We explored patterns of phylogeographic differentiation across the Tree Swallow’s breeding range by using nine microsatellite loci and a mitochondrial DNA sequence marker. Contrary to this species’ high population-level variation in life-history traits and other ecologically important characteristics, we found no genetic structuring across the majority of its distribution, spanning Tennessee, New York, and Alaska, but we found that birds from California form a distinct yet subtly differentiated genetic cluster. The Tree Swallow can be characterized as a species with both continent-wide genetic panmixia and slight differentiation at one edge of its breeding distribution. This pattern of genetic variation has implications for understanding the underlying basis of geographic variation in this species’ life history and other phenotypic traits.

Key words: genetic structure, geographic variation, microsatellites, mitochondrial DNA, Tachycineta bicolor, Tree Swallow.

Resumen. Entender la estructuración geográfica de la variación genética de Tachycineta bicolor es informativo debido a que esta especie, que se distribuye ampliamente en Norteamérica, se está usando cada vez más como modelo para estudios sobre sistemas de apareamiento, caracteres de historia de vida y fisiología. Exploramos los patrones de diferenciación filogeográfica en todo el área de distribución reproductiva de T. bicolor usando nueve loci microsatelitales y un marcador de secuencia de ADN mitocondrial. De modo contrastante a la alta variación que se observa a nivel poblacional en los caracteres de historia de vida y otras características ecológicas importantes, no encontramos estructuración genética en la mayor parte del área de distribución, que incluye los estados de Tennessee, Nueva York y Alaska. Sin embargo, encontramos que las aves de California forman un agrupamiento genético distintivo aunque sutilmente diferenciado. Esta golondrina puede ser caracterizada tanto como una especie con panmixia continental, como una especie con una pequeña diferenciación geográfica en un extremo de su distribución geográfica. Este patrón de variación genética tiene implicaciones importantes para el entendimiento de la variación geográfica de las características de historia de vida y otros caracteres fenotípicos de esta especie.

INTRODUCTION

Owing to the hundreds of studies that have taken advantage of its tractability, the Tree Swallow (Tachycineta bicolor) has been termed a “model species” for research topics spanning mating systems to immunology to climate change (Jones 2003). Tree Swallows breed across nearly all of temperate North America, and past research on them has taken place at many locations across this broad distribution. Although the species is taxonomically monotypic with no described subspecies, it shows substantial geographic variation in reproductive effort (Dunn et al. 2000, Ardia 2005, Monroe et al. 2008), offspring development (McCarty 2001, Ardia 2006), and immune responses (Ardia 2005, Ardia 2007). Little is known, however, about patterns of genetic differentiation or connectivity among Tree Swallow populations or about the extent to which the observed geographic variation in phenotypic traits is related to underlying population structuring.

Manuscript received 6 October 2008; accepted 29 April 2009.
5E-mail: lms9@cornell.edu
Despite the paucity of information on genetic structuring, the ease with which cavity-nesting Tree Swallows can be banded and monitored during the breeding season has resulted in a large and well-documented collection of information on direct dispersal distances ( Hosner and Winkler 2007), which allows us to make a priori predictions about the likely genetic population structure of the species. Tree Swallows breed throughout central and northern North America and are continuously distributed from Maine to Alaska and as far south as Tennessee and southern California ( Robertson et al. 1992). Banding recoveries indicate that Tree Swallows have substantial potential for long-distance dispersal from where they fledge or breed ( Winkler et al. 2004, 2005 ). Recapture data from the U.S. Bird Banding Laboratory and the more spatially focused Swallow Dispersal Study allowed Hosner and Winkler (2007) to document dispersal distances of up to 2367 km, although the majority (85%) of known dispersals occurred among sites separated by less than 15 km. Nevertheless, even rare long-distance dispersal could be sufficient to homogenize allele frequencies across long distances ( Wight 1931, Slatkin 1987). This combination of a continuous distribution with no known major physical barriers to movement, annual migration, and documented long-distance dispersal leads to the potential for high genetic exchange and little or no spatial genetic structure in the Tree Swallow, as was found in one of the classic first phylogeographic surveys of a broadly distributed North American songbird, the Red-winged Blackbird ( Agelaius phoeniceus ), whose geographic variation in morphology is greater ( Ball et al. 1988 ). Some other North American birds also have little genetic structure over large portions of their breeding ranges ( e.g., Gibbs et al. 2000, Milot et al. 2001, Zink et al. 2006). However, geographic patterns of genetic differentiation may arise even in highly dispersive taxa as a result of evolutionary forces opposing gene flow, such as local adaptation in response to strong selection and/or genetic drift acting on small effective population sizes within locally interbreeding populations ( Slatkin 1987, Freeman-Gallant 1996), from barriers to dispersal that would have otherwise gone unrecognized ( e.g., Mackenzie et al. 2004, Zardi et al. 2007), and even from random sorting of neutral genetic markers across geography ( Irwin 2002). Overall, North American birds show a high degree of taxon-specific variation in phylogeographic patterning ( Zink 1996), rendering it necessary to examine patterns of historical differentiation on a case-by-case basis.

Numerous phylogeographic studies of North American birds have addressed the demographic influences of late Pleistocene climate changes, which must have had profound effects on the distribution of many bird species ( Hewitt 1996, Klicka and Zink 1997). Although species differ in the details of their genetic structuring, surveys employing various molecular markers have suggested that some North American birds have expanded their ranges in the Holocene from one or more past refugia ( e.g., Ruegg and Smith 2002, Milá et al. 2006, 2007). As the Tree Swallow has the broad distribution and low morphological variation often associated with this type of recent range enlargement, we explored whether its pattern of genetic variation in mitochondrial DNA is consistent with a history of demographic expansion. To examine distribution-wide levels of genetic population structuring further, we obtained samples from sites that span the range of the species north–south and east–west, and we tested for differentiation among these populations by using both microsatellite and mitochondrial sequence markers.

METHODS

SAMPLE COLLECTION

Tree Swallows frequently use man-made nest boxes for breeding, which made it possible for us to collect blood from adult birds trapped when they entered to brood or feed young. Sampled individuals were banded with a U.S. Fish and Wildlife Service band before blood was collected (100–200 µl) from the brachial vein into heparinized microhematocrit tubes and stored at room temperature in 0.5 ml of lysis buffer (0.1 M Tris [pH 8.0], 0.1 M EDTA, 10 mM NaCl, 0.5% SDS) ( White and Denismore 1992). The geographic coordinates of sampled boxes (or clusters of closely located boxes) were obtained with a handheld GPS unit. No voucher specimens were collected because all swallows were from study sites used for active monitoring of reproductive success and other ecological and behavioral variables.

We sampled birds from four sites that together span the majority of the species’ breeding range: sites in New York and Ontario were sampled in 1998 and 1999, sites in Alaska, California, and Tennessee in 2003. The spatial distribution of sampled birds differed somewhat by site: 300 birds came from 15 sites in New York and one in Ontario (11–412 km apart) ( Stenzler 2001), 37 birds were sampled from three sites in Alaska (3–8 km apart), 46 birds were from three sites in California (separated by up to 500 km), and 25 birds were sampled from one site in Tennessee.

MOLECULAR METHODS

We extracted total genomic DNA from blood samples by using DNeasy Tissue Kits (Qiagen, Valencia, CA) or Perfect gDNA (Eppendorf, Westbury, NY) blood-extraction kits according to the manufacturer’s instructions, except that proteinase-K addition was followed by an overnight incubation at 65°C. Extracted DNA was archived at −70°C until analyzed.

Individuals were genotyped via the polymerase chain reaction (PCR) at nine microsatellite loci, four of which were developed in species other than T. bicolor. The loci were HrU7 (Primmer et al. 1995), IBI 5–29 and IBI MP3-31 (Crossman 1996), TBI-81, TBI-104, and TBI-106 (Stenzler 2001), Pca3 (Dawson et al. 2000), LTR6 (McDonald and Potts 1994), and Ase29 (Richardson et al. 2000). Forward primers were
modified at the 5' end by addition of a fluorescent label (PET, 6-FAM, VIC, or NED; Applied Biosystems, Foster City, CA), and an unpaired sequence of GTTTCT was added to the 5' end of reverse primers to ensure complete adenylation of products (Brownstein et al. 1996). We used a multiplexing PCR protocol (Hailer et al. 2005), allowing individuals to be genotyped at all nine loci in three multiplexed PCR reactions. Multiplexed PCR reactions (10 μL) contained 10–100 ng of genomic DNA, 0.25 units of Jumpstart Taq polymerase (Sigma-Aldrich, St. Louis, MO), 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 3.25 mM MgCl₂, 200 μM of dNTPs (Invitrogen, Carlsbad, CA), and from 1.0 to 4.8 pmol of each forward and reverse primer to produce equal fluorescent signals. The PCR cycling profile consisted of one cycle at 95°C for 2 min, 35 cycles of 50 sec at 95°C, 1 min at 56°C or 60°C (specific to each multiplex mix), and 1 min at 72°C, followed by a final extension cycle of 30 min at 72°C. We carried out PCR by using a Dyad thermalcycler (Bio-Rad Laboratories, Hercules, CA). Fragment-size information was collected on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA), and allele sizes were estimated with GeneMapper version 3.7 (Applied Biosystems, Foster City, CA).

We also sequenced 966 base pairs (bp) of the mitochondrial gene ND2. We generated PCR products and sequenced both strands by following Lovette and Rubenstein (2007). Cycle sequencing was carried out with the BigDye 3.1 chemistry (Applied Biosystems, Foster City, CA), and sequences were collected with an Applied Biosystems model 3730 automated DNA sequencer and viewed with the software Sequencher 4.5 (Genecodes Corp., Ann Arbor, MI).

STATISTICAL ANALYSES
Where appropriate, we controlled for multiple comparisons by calculating the P values adjusted for false discovery rate (FDR) with the function compute.fdr for R 2.3.1. The library of the function is available online at http://www.studiere-search.org/depts/biostats/documents/fdr-library.R. We used the method of Benjamini and Hochberg (1995), which controls the proportion of significant results that are in fact false positives (type I errors). FDR control has the advantage of being less stringent than the widely used Bonferroni correction and hence avoids the considerable loss of power associated with this method (e.g., Moran 2003, Verhoeven et al. 2005).

MICROSATELLITES
We evaluated genotypic linkage disequilibrium by testing for nonrandom correlations between genotypes across pairs of loci by Fisher’s exact probability test implemented by the web version of GENEPOP v3.1d (Raymond and Rousset 1995). We used Markov chain methods (2000 batches, 5000 iterations per batch) to estimate the P-value of the test. Expected and observed heterozygosity and number of alleles per locus for populations defined by geographic location were also calculated with GENEPOP.

We used FSTAT v2.9.3.2 (Goudet 1995) to test for deviations from Hardy–Weinberg equilibrium at several hierarchical levels, first among birds within sampling sites, then within populations defined by geographic location after combining sampling sites within each state. This program performs an exact test based on Wright’s FST (Wright 1951), a measure of the degree of inbreeding within a group and thus a reflection of the nonrandom association of alleles. Alleles were permuted among individuals within each population. Markov chain methods (15 000 randomizations) were used to create the contingency tables, and FST was calculated and ranked for each. In addition, we evaluated Hardy–Weinberg equilibrium globally by permuting alleles among all populations and using Wright’s FST, the total inbreeding coefficient (Wright 1951), to characterize the data.

Sampling in the field was conducted without prior knowledge of geographic or genetic boundaries between groups of interbreeding individuals. Nest boxes at sampling sites were usually situated very close to each other in clusters of as few as 10 and as many as 100. Spacing between boxes within clusters was between 0 (back to back on poles) and 300 m. Thus we first considered each sampling site as an a priori population. A priori populations relatively close to each other (i.e., the 15 sites in New York/Ontario, the 3 sites in Alaska, and the 3 sites in California) were tested first for genetic differentiation. In Tennessee, 22 birds were sampled from one cluster of nest boxes (separated by ~1.7–21 km), with an additional three birds up to 30 km away; therefore, we considered all of these birds to represent one sampling site. If there was no significant population structure among sampling sites, we considered those birds as one geographically defined population for further analyses.

Using the microsatellite data, we determined values of Wright’s FST (Weir and Cockerham 1984) for pairs of populations and globally to evaluate genetic structure among a priori populations, as implemented in the software program GENETIX (Belkhir et al. 1996), which uses permutation-based statistical inference (10 000 permutations) to determine the significance of the observed FST values. In addition, we tested the null hypothesis of identical distribution of alleles across populations (i.e., no genetic differentiation) by a log-likelihood G-based exact test (Goudet et al. 1996) as implemented in the software package FSTAT v2.9.3.2 (Goudet 1995).

We also used the Bayesian clustering method implemented by the program STRUCTURE v2.0 (Pritchard et al. 2000) to infer spatial structure from the genetic data, eliminating the need to assign birds to a priori populations. The model clusters individuals into K groups, which minimize Hardy–Weinberg and linkage disequilibrium. Within a run, K is fixed to a value set by the user. Multiple runs of a range of values of
was 0.5 or higher. These K values were obtained using the program STRUCTURE. We used the admixture model with a burn-in period of 500,000 Markov-chain Monte Carlo iterations followed by 2,000,000 further iterations. We used the correlated allele frequency model, which allows migration and shared ancestry among the populations. An individual value of α (the Dirichlet parameter for the degree of admixture) was used for each population, with an alpha of 0.001 (alpha is the standard deviation of the distribution from which α is updated along the runs). Runs were conducted for K between 1 and 10 with 10 independent repetitions at each K. We first estimated the most likely ‘true’ value of K by computing the posterior probabilities of K with the method suggested by Pritchard and Wen (2004), who noted, however, the difficulties in selecting the ‘true’ value of K, especially where K > 2. For this reason, we also used the ΔK method of Evanno et al. (2005) to evaluate the value of K. We then assigned each individual to the genetic cluster for which its estimated coefficient of ancestry (Q) was 0.5 or higher.

MITOCHONDRIAL HAPLOTYPES

For mtDNA sequence data, we computed standard indices of haplotype (h) and nucleotide diversity (π) (Nei 1987) for all birds as well as per geographic population by using DnaSP (v4.50.1) (Rozas et al. 2003). To infer population-level divergence on the basis of an independent-gene genealogy we estimated genealogical relationships among the ND2 sequences in mtDNA with the program TCS v. 1.21 (Clement et al. 2000). The program creates a minimum-spanning network with 95% plausibility per link between haplotypes. We used ARLEQUIN v3.1 (Excoffier et al. 2005) to assess genetic structure among populations. This was done by calculating the population’s pairwise FST values from genetic distances computed as the average number of pairwise differences between haplotypes. In addition, an analysis of molecular variance (AMOVA) was calculated (ARLEQUIN) to partition the variation among molecular haplotypes within and among individuals and populations. Groupings consisted of all four populations or two populations (the first consisting of all California birds and the second a pool of all other birds). P-values were determined by permutation (10,000 permutations).

We tested whether the observed pattern of nucleotide polymorphism among ND2 sequences might be an indication of a past population expansion, by using the statistics FST (Fu 1997) and R (Ramos-Ortiz and Rozas 2002), which look for an excess of rare (i.e., young) alleles over expectation from a neutral model of evolution, as implemented in the program DnaSP, v4.50.1 (Rozas et al. 2003). The null hypothesis for both tests is a constant population size, versus the alternative of population growth. These two statistics are the most powerful for detecting population expansions, FST for larger and R for smaller sample sizes (Ramos-Ortiz and Rozas 2002). The significance of the tests was calculated by coalescent simulations conditional on the number of segregating sites with 10,000 replicates, also implemented by DnaSP. In addition, we used ARLEQUIN to perform Tajima’s test (Tajima 1989) D for selective neutrality. The significance of the test is calculated with a coalescent simulation under a model of selective neutrality and population equilibrium.

RESULTS

MICROSATELLITE ANALYSES

Observed heterozygosities of the nine loci ranged from 0.26 to 0.97, and the number of alleles was between 6 and 14 per locus (Table 1). There was no evidence of linkage disequilibrium among microsatellite loci, nor was there deviation from Hardy–Weinberg equilibrium, globally or within any of the four geographically defined populations (P > 0.10 for all tests

<table>
<thead>
<tr>
<th>Locus</th>
<th>All birds</th>
<th>New York</th>
<th>Alaska</th>
<th>California</th>
<th>Tennessee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k</td>
<td>H_o</td>
<td>H_e</td>
<td>k</td>
<td>H_o</td>
</tr>
<tr>
<td>HrU7</td>
<td>7</td>
<td>0.67</td>
<td>0.74</td>
<td>7</td>
<td>0.59</td>
</tr>
<tr>
<td>IBI MP 5-29</td>
<td>6</td>
<td>0.65</td>
<td>0.6</td>
<td>5</td>
<td>0.57</td>
</tr>
<tr>
<td>TBI-104</td>
<td>11</td>
<td>0.63</td>
<td>0.67</td>
<td>10</td>
<td>0.86</td>
</tr>
<tr>
<td>TBI-81</td>
<td>14</td>
<td>0.86</td>
<td>0.91</td>
<td>14</td>
<td>0.97</td>
</tr>
<tr>
<td>Ase29</td>
<td>13</td>
<td>0.88</td>
<td>0.86</td>
<td>10</td>
<td>0.78</td>
</tr>
<tr>
<td>IBI MP 3-31</td>
<td>14</td>
<td>0.90</td>
<td>0.83</td>
<td>13</td>
<td>0.81</td>
</tr>
<tr>
<td>Pca3</td>
<td>11</td>
<td>0.61</td>
<td>0.67</td>
<td>10</td>
<td>0.68</td>
</tr>
<tr>
<td>LTR6</td>
<td>11</td>
<td>0.75</td>
<td>0.75</td>
<td>10</td>
<td>0.73</td>
</tr>
<tr>
<td>TBI-106</td>
<td>9</td>
<td>0.78</td>
<td>0.70</td>
<td>6</td>
<td>0.76</td>
</tr>
<tr>
<td>All loci</td>
<td>0.75</td>
<td>0.74</td>
<td>0.77</td>
<td>0.75</td>
<td>0.72</td>
</tr>
</tbody>
</table>
with the exception on one locus in one population (Pca3 in New York; \( P = 0.007 \)).

Microsatellite-based pairwise \( F_{ST} \) values among the 15 sampling sites in New York/Ontario, among the three in California, and among the three in Alaska were universally low (–0.0132 to 0.0178) and not significant (all \( P > 0.05 \)). Exact tests for genetic differentiation among these same groups were likewise not significant. On the basis of these results, we considered the birds in New York/Ontario, in California, and in Alaska each a geographic population and grouped them accordingly for further analyses. A subset of birds from the New York/Ontario sampling area, taken from the center of New York state, was selected with a random-number table to bring the sample sizes in the four populations closer to parity for further analyses. The resulting sample consisted of 159 birds from four geographic populations: New York (\( n = 51 \)), Alaska (\( n = 37 \)), California (\( n = 46 \)), and Tennessee (\( n = 25 \)) (Fig. 1).

Pairwise \( F_{ST} \) values for these four geographic populations ranged from 0.00043 to 0.05753. The values of \( F_{ST} \) for California compared to each of the other three populations were all significant (\( P < 0.001 \)), whereas the only other significant pairwise comparison involved the New York and Alaska populations (\( P = 0.03 \); Table 2). Only the comparisons involving the California populations were significant in the exact tests for differentiation implemented in FSTAT (data not shown).

The STRUCTURE analysis further highlighted the differentiation of the California population and the homogeneity of the remaining three populations. The highest posterior probability was for subdivision into two genetic groups (\( K = 2 \)) (Fig. 2). The \( \Delta K \) method (Evanno et al. 2005) also supported this two-population subdivision, as the \( \Delta K \) value for \( K = 2 \) (7.002) was greater than the \( \Delta K \) (range: 0.67 to 1.22) for all other tested values of \( K \). We assigned each individual bird to one of the two inferred genetic groups on the basis of coefficients of ancestry calculated by the STRUCTURE run for \( K = 2 \) with the highest value of \( \ln P(D) \). The first group included a total of 32 birds, 29 from California, 2 from Alaska, and 1 from New York. The second group consisted of the remaining 127 birds. The pairwise \( F_{ST} \) between these two genetic groups was 0.05719 and highly significant (\( P < 0.001 \)). Despite the overall support for the presence of two genetic groups within the continental breeding range of the Tree Swallow, most individual birds had intermediate assignment probabilities (Fig. 3), suggesting that these two groups are not deeply differentiated. This pattern of subtle separation is particularly evident in the California population, of which the 46 genotyped individuals all fall in the middle region of the assignment distribution, with no individual strongly supported as belonging predominantly to one genetic group.

### Table 2. Pairwise \( F_{ST} \) values for four geographic populations of the Tree Swallow based on nine microsatellite loci and the sequence of ND2.

<table>
<thead>
<tr>
<th></th>
<th>New York</th>
<th>Alaska</th>
<th>California</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nine microsatellite loci</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alaska</td>
<td>(0.00436^a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>(0.04343^b)</td>
<td>(0.03857^b)</td>
<td></td>
</tr>
<tr>
<td>Tennessee</td>
<td>(0.00043^{ab})</td>
<td>(0.00368^{ab})</td>
<td>(0.05753^b)</td>
</tr>
<tr>
<td>ND2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alaska</td>
<td>(−0.00158^c)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>(0.15372^a)</td>
<td>(0.10338^a)</td>
<td></td>
</tr>
<tr>
<td>Tennessee</td>
<td>(−0.00341^c)</td>
<td>(−0.00796^c)</td>
<td>(0.12197^a)</td>
</tr>
</tbody>
</table>

\(^aP < 0.05.\)

\(^bP < 0.001.\)

\(^cNot statistically significant; 10,000 permutations.\)
GENETIC VARIABILITY OF TREE SWALLIES

MITOCHONDRIAL-HAPLOTYPE ANALYSES

We obtained 966 bp of the ND2 mitochondrial sequence from 155 of the 159 Tree Swallows; DNA was unavailable from four samples. Among the mtDNA sequences, overall haplotype diversity was 0.767, which varied little among populations. Nucleotide diversity among the four populations differed and was lowest in New York, Alaska, and Tennessee (0.0016–0.0019, mean = 0.0018) and highest in California (0.0045). Overall nucleotide diversity was 0.0027 (Table 3). Collectively there were 43 haplotypes (GenBank accession numbers EU725728 through EU725770), including one that was common in all four populations and shared by a total of 74 individuals. This common haplotype is at the center of one of the two major haplotype clusters reconstructed in the spanning network (Fig. 4). All other haplotypes in the larger cluster are at low frequencies and are separated from the common central haplotype by one to three nucleotide substitutions, with no apparent geographic structure. The second cluster of haplotypes in the spanning network was much more geographically restricted and comprised two haplotypes found in 11 Tree Swallows, 10 from California and one from Alaska. This haplotype group was separated from the larger group by a minimum of nine nucleotide substitutions.

Population pairwise $F_{ST}$ values ranged from −0.00796 to 0.15372. The values of $F_{ST}$ for California compared to those for each of the other three populations revealed significant genetic subdivision ($P < 0.05$), whereas all other pairwise comparisons were not significant (Table 2). Variation among populations was significant both for four populations (AMOVA, 9.4% among populations, 90.6% within populations, $P < 0.0001$) and when samples from New York, Tennessee, and Alaska were pooled and compared with those from California (AMOVA, 16.7% among populations, 83.3% within populations, $P < 0.0001$).

Fu’s $F$ test (Fu 1997) for New York ($F = −22.57, P < 0.0001$) and Alaska ($F = −9.25, P = 0.003$) was significant, as was Ramos-Onsins and Rozas’ $R^2$ test (2002), a more powerful test, for the smaller Tennessee ($R^2 = 0.067, P = 0.01$) population. In contrast, the null hypothesis of population stability was not rejected for New York ($R^2 = 0.001, P = 0.825$) and Alaska ($R^2 = 0.000, P = 0.673$).

Table 3. Characteristics of mtDNA ND2 sequence in Tree Swallows sampled from four geographic populations from New York, Alaska, California, and Tennessee; $h$ is haplotype diversity, $\pi$ is nucleotide diversity.

<table>
<thead>
<tr>
<th>Geographic population</th>
<th>Number of segregating sites</th>
<th>Number of haplotypes</th>
<th>$h$</th>
<th>$\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All birds</td>
<td>44</td>
<td>43</td>
<td>0.767</td>
<td>0.0027</td>
</tr>
<tr>
<td>New York</td>
<td>24</td>
<td>22</td>
<td>0.795</td>
<td>0.0016</td>
</tr>
<tr>
<td>Alaska</td>
<td>20</td>
<td>15</td>
<td>0.749</td>
<td>0.0019</td>
</tr>
<tr>
<td>California</td>
<td>18</td>
<td>11</td>
<td>0.704</td>
<td>0.0045</td>
</tr>
<tr>
<td>Tennessee</td>
<td>13</td>
<td>10</td>
<td>0.815</td>
<td>0.0018</td>
</tr>
</tbody>
</table>

FIGURE 3. Probability of membership in each of two genetic groups, based on the coefficient of ancestry ($Q$) calculated by the Bayesian clustering algorithm of STRUCTURE. Gray bars represent group 1, white bars group 2. Each bar represents an individual bird.
could not be rejected for California, where \( F_{ST} \) was not significant (\( F_{ST} = 0.269, P = 0.59 \)). The test for the entire sample of birds was highly significant for \( F_{ST} \) (\( -39.6829, P < 0.0001 \)). Tajima’s \( D \) statistic, tested against a model of selective neutrality and population equilibrium, was significant for New York (\( D = -2.31, P = 0.0007 \)), Alaska (\( D = -2.03, P = 0.005 \)), and Tennessee (\( D = -1.72, P = 0.024 \)) but not for California (\( D = 0.192, P = 0.63 \)). These results are consistent with a pattern of population expansion in New York, Alaska, and Tennessee but of greater population stability in California.

**DISCUSSION**

Our analyses of microsatellite and mitochondrial markers independently suggest that Tree Swallows have a hierarchical pattern of genetic structuring across the broad area encompassed in this study. Across sites that span most of their breeding distribution, we found little to no evidence for genetic structuring, suggesting that these geographically distant sites have been connected by recent or continuing gene flow. In the West and particularly in California, both markers supported the presence of subtle but significant population differentiation, suggesting that these western populations have a slightly more complicated history. The microsatellite-based pairwise \( F_{ST} \) between Alaska and New York was significant but very low (0.00436). From a biological point of view, this low \( F_{ST} \) plus the nonsignificance of the exact test of differentiation and of the mtDNA-based pairwise \( F_{ST} \) between these two populations lead us to believe that the birds from Alaska and New York are not genetically differentiated.

Mitochondrial variation at ND2 has a pattern consistent with the microsatellite results, with one widespread haplotype cluster common throughout the range of the species and a second divergent cluster geographically restricted to our western sites. Neutrality tests suggest a nonrandom overabundance of rare or new alleles for all areas except California. Nucleotide diversity is highest in California, possibly the result of a longer period over which mutations have accumulated there, but certainly because the California population contains individuals from both of the differentiated haplotype clusters. The microsatellite-based assignment tests and \( F_{ST} \)-based comparisons identify congruently subtle differentiation in the California population, suggesting that this pattern is not an artifact of a single-marker system. A number of historical scenarios could explain this pattern of general panmixia with modest differentiation in California; one possibility is that the ancestral California population experienced a period of isolation and differentiation before being recently reconnected to the broader, expanding population. This scenario is consistent with the low mtDNA haplotype variation in the cluster restricted to the West, with the evidence for admixture seen in California at both classes of markers and with the historical demographic signature of population expansion seen in the non-California populations. The differentiation of the California population could also be related to differences in migration behavior, as Butler (1988) reported that birds from the midwestern U.S. and Canadian prairies appear to winter along the coast of the Gulf of Mexico and in Central America, whereas east-coast breeding birds winter in Florida and the Caribbean.

It is notable that the population-genetic patterns revealed by these data are not concordant with patterns of geographic variation in life-history traits. Ardia (2005, 2006, 2007) compared life-history traits of the Tree Swallow in three of the four geographic areas we sampled (New York, Tennessee, and Alaska, but not California), finding clear geographic differences in reproductive effort, nestlings’ growth rate, body condition, and immune responses. Ardia attributed much of this geographic variation to the habitat and environmental factors that impose different selective pressures across these extremes of the Tree Swallow’s breeding range. This lack of concordance among life-history traits and genetic variation is not surprising, as microsatellite and mitochondrial sequence variation is generally selectively neutral, and these markers are unlikely by chance to be linked closely to loci under different selection at these different sites. Furthermore, genetic variation across most of the species’ range is low, and even in California, the degree of genetic differentiation is not high. The more dramatic geographic variation in life-history traits therefore likely results either from strong local selection on these traits or (perhaps most likely) from phenotypic plasticity (Charmantier et al. 2008, Gienapp et al. 2008, Crispo 2008). In an analogous study, Blondel (2007) found wide phenotypic divergence in demographic and behavioral traits in the Blue Tit (Cyanistes caeruleus) across heterogeneous landscapes in the presence of very little genetic divergence, suggesting

![Minimum-spanning network of mtDNA haplotypes from four geographic populations of the Tree Swallow.](image)
phenotypic plasticity may be one of many factors leading to divergence of adaptive traits in the absence of genetic divergence at neutral loci.

In summary, the low magnitude of genetic differentiation among Tree Swallow populations that span most of this species’ extensive breeding distribution suggests that this highly dispersive bird has experienced substantial recent gene flow across most of North America. A finer-scale evaluation of genetic variation in the western region of the species’ breeding range is necessary to explore the origin and geographic limits of the genetic structuring that caused the California population to be subtly but significantly differentiated at both microsatellite and mitochondrial loci.

ACKNOWLEDGMENTS

Funding was provided by the Cornell Laboratory of Ornithology, Evolutionary Biology Program, and by the following grants: National Science Foundation OISE-0730180, DEB-0515981, IBI-013437, DEB-0717021, DEB-0515981, and DEB-0814277, and by Cooperative Agreement EPA-CR 829374010 with the National Center for Environmental Assessment. Bird banding was done under U.S. Fish and Wildlife Service banding permit 20576, held by David W. Winkler.

LITERATURE CITED


