SPECIES LIMITS IN THE LE CONTE'S THRASHER

ROBERT M. ZINK AND RACHELLE C. BLACKWELL
J. F. Bell Museum, 100 Ecology Building, University of Minnesota, St. Paul, MN 55108, e-mail: rzink@biosci.cbs.umn.edu

OCTAVIO ROJAS-SOTO
Museo de Zoología "Alfonso L. Herrera," Departamento de Biología, Facultad de Ciencias, Universidad Nacional Autónoma de México, Apartado Postal 70-399, México D.F. 04510, México

Abstract. We studied patterns of mitochondrial DNA (mtDNA) and colorimetric variation in the Le Conte's Thrasher (Toxostoma lecontei) to determine patterns of geographic variation, and to clarify subspecific and species limits. We sequenced 619 base pairs of mtDNA (parts of cytochrome b and ND6) for 14 Le Conte's Thrashers from 9 localities representing the three putative subspecies (lecontei, macmillanorum, and arenicola). Haplotypes were shared between lecontei and macmillanorum, and we do not support recognition of the latter. However, haplotypes of arenicola differed by 3.5% sequence divergence from those of lecontei and macmillanorum, and we suggest that arenicola be recognized as a species. Colorimetric analyses of 105 specimens indicated that arenicola differed consistently from lecontei and macmillanorum, corroborating the mtDNA results.

Key words: Toxostoma lecontei, T. arenicola, mtDNA, sequence analysis, colorimetric variation, species limits, phylogeography.

INTRODUCTION

The Le Conte's Thrasher (Toxostoma lecontei) inhabits desert regions in the southwestern United States and northwestern Mexico, including Baja California (Fig. 1). Toxostoma l. lecontei is relatively wide-ranging, occurring in southern California, Arizona, Nevada, Sonora, and Baja California (AOU 1957). Phillips (1965) described the isolated population in the San Joaquin valley of southern California as T. l. macmillanorum, a taxonomic decision not accepted by Sheppard (1973, 1996). Toxostoma l. arenicola, described by Anthony (1897), is isolated on the Pacific coast of Baja California from approximately 29° to 26° N. Since first described, arenicola has been accepted by the AOU (1899, 1957) as a subspecies of T. lecontei. Contact between lecontei and arenicola might occur between Bahia Santa Rosalia and San Felipe near Laguna Chapala (29°26'N, 114°27'W) in Baja California (Sheppard, pers. comm.). A primary basis for subspecific limits was coloration, with macmillanorum in the north and arenicola in the south being relatively darker than lecontei.

As part of an ongoing study of molecular evolution in the genus Toxostoma, we acquired specimens of T. l. lecontei and T. l. arenicola, as well as the putative macmillanorum (Fig. 1). We conducted a phylogeographic analysis, consisting of superimposing a phylogeny of mitochondrial DNA haplotypes onto the geographic distribution of samples (Avise 1994). We obtained sequence data from the mitochondrial genes ND6 and cytochrome b (cyt b), and we recorded quantitative dorsal colorimetric measurements. We conclude that T. l. lecontei and T. l. arenicola warrant species status, whereas macmillanorum did not appear to be a distinct taxon.

MATERIALS AND METHODS

The 14 Le Conte's Thrashers analyzed for mtDNA sequence variation (Fig. 1) were obtained: near Maricopa, Kern Co., California (n = 3); near Vicksburg Junction, La Paz Co., Arizona (n = 3); Laguna Salada, Baja California (n = 2); 10 km north of San Felipe, Baja California (n = 1); Bahia Santa Rosalia, Baja California (n = 3), and 15 km south of Bahia Santa Rosalia, Baja California (n = 2). Voucher specimens are located in the American Museum of Natural History, Louisiana State University Museum of Natural Science, and Museo de Zoología, U.N.A.M. (Mexico). Purification of
mtDNA from the California, Arizona and Bahia Santa Rosalia tissues followed procedures in Lansman et al. (1981). MtDNA was extracted from the remaining samples following incubation in Chelex/Proteinase K, a modification of Ellegren's (1992) method. The polymerase chain reaction (Hillis et al. 1990) was used to symmetrically amplify a 484 base pair (bp) product from the cytochrome b gene and an 800 bp product spanning ND6, t-glutamine and Region I of the control region (Tarr 1995). All symmetric amplifications were performed in 25 μl reaction volumes using 0.3 units Thermus Flavus polymerase (Epicentre Technologies) and a 2 mM magnesium chloride concentration. Primers L14841 and H15299 (Operon, numbers refer to primers in Kocher et al. [1989], and Hackett [1992], respectively) were used to amplify cyt b. A primary cycle of 3 min denaturation at 94°C, 1 min at 50°C, and 1 min extension at 72°C was followed by 33 cycles of 1 min at 94°C, 1 min at 55°C, 1 min at 72°C. A final extension at 72°C was carried out for 10 min. Primers ND6E (made by Operon, identified by Edwards 1993) and HCR4 (made by Operon, identified by Tarr 1995) were used to amplify the ND6/control region. Thirty-five cycles of 1 min at 94°C, 1 min at 49°C, 1 min at 72°C were followed by a 10 min 72°C final extension. Agarose gel electrophoresis verified amplification of desired products.

Double-stranded sequencing was carried out using a PCR Product Sequencing kit (United States Biochemical) and the same primers used for PCR. Light and heavy strands of all samples were sequenced for cyt b. Only the light strand was sequenced for the ND6/control region product, and only the sequence of the ND6 gene was used in the analysis (no new relevant information was contributed by control region sequence). Standard 6% acrylamide gels were run and sequence was visualized by autoradiography. Sequences were deposited in GenBank (Accession Number U75569–75580).

Sequences were aligned visually. Each distinguishable sequence was considered a unique haplotype and only one representative of each haplotype was used for analysis. Sequence data from the Crissal Thrasher (Toxostoma crissale), the sister species of the Le Conte’s Thrasher (Zink and Blackwell, unpubl. data), was used as an outgroup to root the Le Conte’s Thrasher haplotypes. We used maximum parsimony with equal weighting to infer a phylogenetic tree. A neighbor-joining tree based on the p-value also was computed using MEGA (Kumar et al. 1993). A maximum-likelihood tree was estimated with PHYLIP version 3.3 (Felsenstein 1993).

Using a Minolta Chromo-meter (model CR-200), we measured dorsal (back) coloration on 105 specimens (Appendix 1); each specimen was measured by one person three times to assess measurement error. Some samples were pooled to create a total of nine geographic units for statistical analysis (codes in Appendix 1): SR, CH, SA, LS (including SO, SF), RI (including SE), SD (including LA), KE (including KI, SB), IM, MA (including NE, PI, WA, YU). Three colorimetric measurements were recorded for each specimen: Y (lightness), and the two chromaticity coordinates (x, y); the latter two were used to compute hue (dominant wavelength) and excitation purity (chroma). The raw Y, x, and y values were used here for statistical analysis. Nested ANOVA was used to assess measurement error for each value by evaluating the relative magnitude of variance contributed by replicate measurements of each speci-
men. Regression analysis was used to assess the effects of month and year on the colorimetric variables and residual values were computed. Principal components analysis (PCA) was used on both raw and residual values to reveal broad patterns of geographic variation. Student-Newman-Keuls (SNK) test was used to reveal maximally homogeneous subsets of geographic units for each value. The SAS package was used for statistical analysis (SAS Institute 1989).

RESULTS

We compared 619 base pairs of which 60 were variable and 22 were parsimony informative relative to the outgroup. For the 433 bp of cyt b sequenced, variation (excluding the outgroup) included 11 third-base position transitions, 2 third-base position transversions, and 4 first-base position transitions. For the 186 base pairs of ND6 sequenced, variation (excluding the outgroup) included 6 third-base position transitions, 1 first-base position transition, and 2 first-base position transversions. A total of five haplotypes was observed among the 14 Le Conte’s Thrashers. In arenicola, two haplotypes that differed by one base pair (0.16%) were observed among the five individuals sequenced. The other three haplotypes, all from lecontei and putative macmillanorum differed from each other by 3 to 5 base pairs (0.48-0.80%, respectively). The commonest haplotype occurred in 7 of 9 individuals representing lecontei. All methods of phylogenetic analysis revealed two distinct groups of haplotypes, those from the Pacific coast of Baja California corresponding to arenicola, and the remainder (Fig. 2). The two groups of haplotypes differed by an average of 20.5 base pairs (3.5%). The Crissal Thrasher differed at an average of 7.4% of bases from the five haplotypes found in lecontei and arenicola.

Although dominant wavelength and excitation purity were not analyzed, representative values were 582 and 23%, respectively, which can be compared with colorimetric analyses of other species (Johnson 1980, Atwood 1988, Cicero 1996). Regression analysis revealed that the colorimetric values were significantly influenced by the month and year of collection, a phenomenon also recognized by Sheppard (pers. comm.). The rank-order correlation coefficient (−0.23) between county and year was insignificant. Therefore, the effects of month and year were removed by regression and residual values used for subsequent statistical analysis. Nested ANOVA indicated that the effects of measurement error were relatively low, less than 5% of the total variance for each color-residual value. Therefore, the three sets of residual values for each individual were averaged and used for subsequent analysis. Nested ANOVA also showed a significant difference ($P < 0.001$) among localities (but not for replicate measurements) for raw and residual $Y$ values, but not for $x$ and $y$. Analyses in which sexes were analyzed separately (not shown) gave the same results as when sexes were combined. A bivariate ($x$, $y$) chromaticity plot (not shown) indicated extensive overlap among samples. Although principal components analysis (Fig. 3) showed considerable overlap among arenicola and lecontei (plus macmillanorum), SNK analysis (Fig. 4) showed that arenicola is significantly differentiated from adjacent geographic samples (e.g., CH, SA) for each $Y$-residual value. SNK analyses (not shown) of $x$ and $y$ suggested broadly overlapping groups. For the raw (i.e., nonresidual) $Y$ values, SR grouped with KE and SA, a similarity also apparent in Figure 3. Pooling individuals from all sample regions into the three subspecies revealed that lecontei plus macmillanorum represented a statistically homogeneous subset that stood apart from arenicola for both raw and residual $Y$-values, but not for $x$ or $y$. 

![Figure 2. Phylogenetic tree of haplotypes estimated by all methods of analysis (maximum parsimony tree length 65, consistency index 0.95, retention index 0.89). Locality information in text and Figure 1. Numbers at tick marks on branches are number of unambiguous changes and total number of changes possible (in parentheses).](image-url)
DISCUSSION

The sequence data support the existence of two groups of Le Conte’s Thrasher (Fig. 2), corresponding to the recognized subspecies *arenicola* (haplotypes 1 and 2) and *lecontei* (haplotypes 3, 4, and 5). Our mtDNA data do not support the subspecies *macmillanorum* described by Phillips (1965). As a test of phylogeographic pattern, we evaluated a tree in which one haplotype each from *arenicola* and *lecontei* were exchanged among the two clusters. This topology was significantly worse according to Kishino and Hasegawa’s (1989) maximum likelihood test (*P* < 0.05), and also was 24 steps longer than the shortest maximum parsimony tree (Fig. 2). We are confident in the deep phylogeographic division. The uniformity of sequence variation and lack of phylogeographic structure in *lecontei* over 600 km, and the comparatively discrete break between it and *arenicola*, suggest that two species are involved. The amount of sequence divergence, 3.5%, is consistent with other species-level differences (Hackett 1996). For comparison, Bendire’s Thrasher (*Toxostoma bendirei*) and Gray Thrasher (*T. cinereum*), two undisputed species, are 2.1% divergent in mtDNA sequence (Zink and Blackwell, unpubl. data). A restriction-site study (Zink et al., unpubl. data) based on the entire mtDNA genome revealed the same level of sequence divergence between the major clades of *T. lecontei*, as well as between *T. crispa* and the two *lecontei* clades. Congruence between restriction site and sequence data suggests that we did not accidentally amplify and sequence a nuclear copy of mitochondrial genes (Zhang and Hewitt 1996).

Because the mtDNA haplotype phylogeny provides in essence a “single gene tree” (Neigel and Avise 1986), other characters are needed to corroborate an hypothesis of species status for the two forms (Zink and McKitrick 1995). That is, one needs to ascertain that the pattern in mtDNA is not contradicted by patterns in other characters. Morphometric analysis (Sheppard, unpubl. data) suggests a break between *lecontei* and *arenicola*. The darker coloration of *arenicola* was a major reason for its subspecific description (Anthony 1897), although variation might be clinal or intermediates might exist at areas such as Laguna Chapala (Fig. 1; Sheppard, pers. comm.). Our analysis (Fig. 4) confirms that *arenicola* is darker. Although *arenicola* is not 100% separable from other speci-
mens in principal components analysis, the specimens from localities thought to be intermediate (Laguna Chapala) are not the ones most similar (Fig. 3). In fact, as noted by Phillips (1965), coloration of macmillanorum approaches arenicola, and in our SNK analyses (not shown) of raw, but not residual, Y-values these subspecies together formed a statistically homogeneous group. The environments inhabited by macmillanorum and arenicola are relatively humid compared to that occupied by lecontei (Sheppard, pers. comm.), raising the intriguing possibility that adaptation (or environmental induction) consistent with Gloger's Rule (Zink and Remsen 1986) might have led to darker birds at the north and south ends of the distribution. Such potential color convergence occurred against different backgrounds of mtDNA divergence (Fig. 2). One interpretation of this is that differences in coloration (e.g., macmillanorum) might evolve rapidly with respect to mtDNA sequence evolution. Whatever the cause, the distinct "step" (Fig. 4) between arenicola and samples immediately adjacent to the north corroborate the independent status of arenicola. Other morphometric and colorimetric analyses should be pursued to clarify the exact boundaries of the two taxa. Sheppard (pers. comm.) washed his specimens to prevent postmortem color change, and this source of bias should be investigated further. Newly collected material would aid the quantitative assessment of color change over time.

Although vocal differences between arenicola and lecontei are unknown (Sheppard 1996), quantitative study is required. Male birds in Bahia Santa Rosalia responded aggressively in November of 1983 and 1994 to recordings of birds from Kern County, California (Zink, unpubl. data).

We doubt that a consistent or predictive quantity of molecular or morphological divergence can ever be found for species-level taxa (Johnson and Zink 1983). However, we feel that the level and pattern of the mtDNA differentiation, together with colorimetric and biogeographic separation, are consistent with phylogenetic species status (Zink and McKitrick 1995) for the two subspecies. Whether they are biological species can only be inferred at this time, although they are almost certainly "evolutionary species" (Wiley 1978). Although more samples should be obtained within T. l. lecontei, and between it and T. l. arenicola, we think that recognizing two species better represents the evolutionary diversity in this lineage. On the other hand, the same arguments also serve to reject the taxon T. l. macmillanorum.

Many avian taxa exhibit subspecific differentiation in Baja California, especially near the southern extent of the peninsula ("Cape Region," Miller et al. 1957). Toxostoma arenicola is apparently the only avian taxon endemic to the Vizcaino desert, which is the site of a major Biosphere Reserve (Gomez-Pompa and Dirzo 1995). Ongoing studies (Zink et al., unpubl. data) of California Gnatcatchers (Polioptila californica), which span the range of the Le Conte's Thrasher, show no mtDNA breaks consistent with that found in Le Conte's Thrasher. Although there are some endemic plants (Ortega and Arriaga 1991), we were unable to find other Vizcaino desert endemics among nonavian taxa, including agaves (Scott 1978), reptiles (Grismer 1994), mammals (Riddle and Hafner, in press), or butterflies (Brown et al. 1992), although other examples might be discovered as the flora and fauna are better studied. At this time the ecological and biogeographic reasons for the evolution of arenicola, but not other taxa, are unknown.

A previously proposed common name for arenicola, "Desert Thrasher" is inappropriate (Anthony 1897), given that T. l. lecontei also is certainly a "desert" species. We suggest a descriptive name such as Coastal Le Conte's Thrasher (Sheppard, pers. comm.) or Vizcaino Thrasher, the latter in recognition of the general area inhabited by arenicola. The name Rosalia Thrasher also has been used (AOU 1899).

ACKNOWLEDGMENTS

This work was supported financially by NSF (DEB-9317945), the Louisiana Board of Regents (LEQSF #86-LBR-8906621), and the University of Minnesota. We thank D. L. Dittmann for laboratory assistance. J. M. Sheppard provided much helpful information during all stages of this work. For field assistance we thank G. F. Barrowclough and A. J. Fry. We thank the personnel of the American Museum of Natural History, Western Foundation of Vertebrate Zoology, Los Angeles County Museum (K. Garrett), and San Diego Natural History Museum (P. Unit). We thank the Minolta Corporation for donating the Chroma-meter, and E. Becker for technical advice. We thank C. Cicero, A. J. Fry, A. Navarro, A. Gomez-Pompa, J. M. Sheppard, J. T. Klicka, J. D. Weckstein, and S. M. Lanyon for comments on the manu-
script. B. Riddle provided unpublished data. A. J. Fry prepared Figure 1. Permits were provided by the U.S. Fish and Wildlife Service, U.S. Department of Agriculture, and Mexico (SEDESOL).

LITERATURE CITED


Specimens used in colorimetric analysis. Museum codes: SDNH (San Diego Museum of Natural History), LACM (Los Angeles County Museum), AMNH (American Museum of Natural History), and WFVZ (Western Foundation of Vertebrate Zoology). Subspecies: lecontei (LECO), arenicola (AREN), macmillanorum (MACM). Locality: San Augustin, Baja California (SA), San Felipe, Baja California (SF); Chapala, Baja California (CH); Laguna Salada, California (KE); San Bernardino County, California (SR); Maricopa County, Arizona (MA); Pinal County, Arizona (PI); Yuma County, Arizona (YU); Sonora, Mexico (SO).

Data are presented in the following order: museum, catalogue number, subspecies, month of collection (1 = January, etc.), year of collection, sex (M = male; F = female), and locality (see above codes). Specimens used in calorimetric analysis. Museum codes: SDNH (San Diego Museum of Natural History), LACM (Los Angeles County Museum), AMNH (American Museum of Natural History), and WFVZ (Western Foundation of Vertebrate Zoology). Subspecies: lecontei (LECO), arenicola (AREN), macmillanorum (MACM). Locality: San Augustin, Baja California (SA), San Felipe, Baja California (SF); Chapala, Baja California (CH); Laguna Salada, California (KE); San Bernardino County, California (SR); Maricopa County, Arizona (MA); Pinal County, Arizona (PI); Yuma County, Arizona (YU); Sonora, Mexico (SO).

Note added to proof:
Upton and Murphy (in press) describe a phylogeographic break in the lizard Uta stansburiana that corresponds to the break between Toxostoma lecontei and T. arenicola. These authors believe that a temporary seaway fragmented southern and northern Baja peninsular populations of this lizard ca. 1 million years ago. Although the northern and southern lizard clades differed by almost 10% sequence divergence, relative to the 3.5% estimated for the thrasher lizard clades, differences in rates of evolution could explain the apparent discrepancy.