Molecular Analysis of Genetic Variation among Large- and Small-Bodied Sage Grouse Using Mitochondrial Control-Region Sequences

Author(s): N. W. Kahn, C. E. Braun, J. R. Young, S. Wood, D. R. Mata and T. W. Quinn
Published by: University of California Press on behalf of the American Ornithologists' Union
Stable URL: http://www.jstor.org/stable/4089344
Accessed: 27/03/2013 11:36

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

University of California Press and American Ornithologists' Union are collaborating with JSTOR to digitize, preserve and extend access to The Auk.

http://www.jstor.org
Molecular Analysis of Genetic Variation Among Large- and Small-bodied Sage Grouse Using Mitochondrial Control-Region Sequences

N. W. Kahn, C. E. Braun, J. R. Young, S. Wood, D. R. Mata, and T. W. Quinn

Department of Biological Sciences, University of Denver, Denver, Colorado 80208, USA; and Colorado Division of Wildlife, Fort Collins, Colorado 80526, USA

Sage Grouse (Centrocercus urophasianus) are lek breeders whose numbers are in serious decline in some areas owing to fragmentation and loss of shrubsteppe habitat (Aldrich 1963, Johnsgard 1983, Braun 1993, 1995). Based on plumage differences (Aldrich and Duvall 1955), they have been divided taxonomically into two subspecies: C. u. urophasianus (Eastern Sage Grouse) and C. u. phaios (Western Sage Grouse), although the validity of these subspecies designations has been questioned (Johnsgard 1983). Differences in size, filoplume length, tail banding, strut display, and vocalizations (Barber 1991, Hupp and Braun 1991, Young 1994, Young et al. 1994, Braun and Young 1995) led Braun and Young (1995) to propose that the “large-bodied” and “small-bodied” forms of Sage Grouse constitute two distinct species. The largest population of the small-bodied Sage Grouse is in Gunnison County, Colorado (2,400 to 3,000 individuals; C. E. Braun unpubl. data), in a montane basin surrounded by tall mountains (3,100 to 4,200 m) to the north and east (Hupp and Braun 1994, Young et al. 1994). Sage Grouse that were extirpated from New Mexico, Oklahoma, Arizona, and at least nine other counties in southwestern Colorado probably belonged to this small-bodied taxon as well (Young et al. 1994).

The main goals of this study were to ascertain whether genetic differences exist between the large- and small-bodied forms of Sage Grouse and to estimate a crude time frame involved in the morphological and behavioral differences between the forms. We surveyed the rapidly evolving control region of the mitochondrial DNA (mtDNA) genome because it is likely to be informative for resolving differences between closely related avian populations (Brown et al. 1979, Adachi et al. 1993).

Methods.—Sage Grouse wings obtained from hunter-killed birds were collected from Blue Mountain (n = 21), Cold Springs (n = 25), Eagle (n = 26), Gunnison Basin (n = 31), Middle Park (n = 21), and North Park (n = 23) in Colorado, and from Rich County in Utah (n = 17). All wings were kept frozen at −20°C. Our choice of sampling locations was based on discrete boundaries that have been mapped by the Colorado Division of Wildlife, the agency charged with management of game species. The population names reflect the respective geographic locations around which these boundaries occur. Samples from Gunnison Basin were small-bodied birds; all other samples were large-bodied birds whose ranges were within that of the eastern subspecies.

For DNA extraction, approximately 0.1 g of tissue was cut from the muscle dorsal to the humerus of each wing with a sterile razor blade, placed in 600 μL of extraction solution (10 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 10 mM NaCl, 1% SDS, 8.0 mg/mL DTT, 0.40 mg/mL Proteinase K), and incubated at 55°C for 4 h on a rotating carousel. Each sample was then extracted twice with an equal volume of phenol: chloroform: isoamyl alcohol (25: 24: 1) and once with chloroform: isoamyl alcohol (24: 1). The resulting aqueous solution of genomic DNA was ethanol-precipitated and resuspended in TE (Tris-HCl 10 mM, EDTA 1 mM pH 8.0), to a concentration of approximately 300 μg/mL.

We used two previously described primers, 16775L (Quinn 1992) and 521H (Quinn and Wilson 1993), to amplify a highly variable section of region I, and a third primer, 418H (Quinn and Mindell 1996), to perform a “semi-nested-PCR” to increase specificity of the amplification. Double-stranded PCR (DSPCR) was performed in 25 μL total reaction volume (Quinn 1992) using 16775L and 521H primers for 30 cycles with the following thermal profile: denaturation, 92°C for 40 s; annealing, 55°C for 1 min; extension, 72°C for 2 min. Electrophoresis of each DSPCR product resolved one discrete band of approximately 500 base pairs (bp); approximately one-third of each...

<table>
<thead>
<tr>
<th>Location</th>
<th>( h ) with 2( n )</th>
<th>( h ) with 1( n )</th>
<th>( \tau )a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gunnison</td>
<td>0.06345</td>
<td>0.06452</td>
<td>0.00108</td>
</tr>
<tr>
<td>Blue Mountain</td>
<td>0.82661</td>
<td>0.84728</td>
<td>0.02777</td>
</tr>
<tr>
<td>Cold Springs</td>
<td>0.70031</td>
<td>0.74588</td>
<td>0.02989</td>
</tr>
<tr>
<td>Eagle</td>
<td>0.63039</td>
<td>0.64300</td>
<td>0.03054</td>
</tr>
<tr>
<td>Middle Park</td>
<td>0.70594</td>
<td>0.72359</td>
<td>0.02410</td>
</tr>
<tr>
<td>North Park</td>
<td>0.84282</td>
<td>0.86198</td>
<td>0.03927</td>
</tr>
<tr>
<td>Rich County, Utah</td>
<td>0.83474</td>
<td>0.86983</td>
<td>0.03927</td>
</tr>
</tbody>
</table>

a Nucleotide diversity.

b Nucleotide diversity.

resulting band was removed and melted at 65°C in 100 μL of dH₂O; 2 μL of this dilution were then used as template in a 50-μL single-stranded, nested PCR (SSnPCR) using primers 16775L (at 1.0 μM) and 418H (at 0.02 μM) and amplified for 35 cycles with the following thermal profile: denaturation, 92°C for 40 s; annealing, 50°C for 1 min; extension, 72°C for 2 min.

Cleaned SSnPCR products were sequenced using Sequenase 2.0 kits (US Biochemicals) following the manufacturer’s recommendations and visualized following the procedure of Quinn (1992). Although we sequenced a region of approximately 380 bp in the first 20 of the 31 Gunnison Sage Grouse samples, and in the first 50 of the 133 northwestern Colorado Sage Grouse samples, we found that a region of 141 bp contained approximately 92% of all variable sites; this is where subsequent sequencing efforts were concentrated.

Sequences were manually aligned using the Eye-ball Sequence Editor program (ESSE-2; Cabot and Beckenbach 1989), and each unique sequence was assigned a different haplotype designation. Percent-divergence values were corrected to account for saturation effects using the Kimura correction (Kimura 1980). Calculations of substitution-rate distance matrices and standard errors and of transition/transition ratios were performed on MEGA version 1.0 (Sudhir et al. 1993) in the noncoding mode, using the complete-deletion option for gaps. We used the neighbor-joining algorithm in MEGA 1.0 to construct distance trees and estimate branch lengths. Nucleotide diversity (π), and nucleon diversity (\( f \)) were calculated for each haplotype (Table 1; Nei 1987). Maximum-parsimony analyses were performed using the heuristic search algorithm of PAUP 3.0 (Swofford 1991). The program was run with starting trees determined by stepwise addition, zero-length branches set to collapse, and only minimal trees kept. Options were as follows: a random-addition sequence was used (100 repetitions), multiple states were interpreted as uncertainty, branch-swapping used the “tree bisection-reconstruction” (TBR) algorithm with all minimum trees saved (MUL- PARS), the steepest descent option was used, the initial maxtrees was set to equal 100 with an automatic increase by 100 if and when the limit was reached. Bootstrap values were calculated by heuristic search, using 100 repetitions, maxtrees were set to equal 100 with an automatic increase by 100 if and when the limit was reached. Bootstrap consensus retained groups with frequency greater than 50%; branch-swapping used the TBR algorithm, with all minimum trees saved (MULPARS) using the steepest descent option and saving no more than 100 trees greater than or equal to 59 steps for each repetition.

DNA sequence from Blue Grouse (Dendragapus obscurus; GenBank accession number AF132471) was used as an outgroup because in classic phylogenies this species is considered to be a sister taxon based on morphological and behavioral traits (Johnsgard 1983).

Results.—Among the 164 birds sequenced, 23 of the 25 variable sites were informative for parsimony analysis across the 141-bp region. Collectively, 21 haplotypes were identified and have been deposited in GenBank under accession numbers AF132450 to AF132470. The Gunnison Basin population, from which 31 small-bodied birds were sampled, was composed almost solely (96.8%) of haplotype A birds, with one exception (haplotype G; Fig. 1, Table 2). The A and G haplotypes differ by a single base substitution. The G haplotype was not found in any of the large-bodied populations.

In contrast, all populations of large-bodied Sage Grouse contained a wide collection of haplotypes (Fig. 1, Table 2); each included three or four of the widespread haplotypes A, B, C, and D. In addition, each large-bodied population contained several rare haplotypes that were either not found in other populations, or were shared between neighboring populations only. Because “double sequence” was absent in all samples, intraspecific variation was high, and the tissue sources were muscle (and hence relatively rich in mitochondrial genomes), we concluded that these sequences were from the mitochondrial genome and not from nuclear homologues (Quinn 1992). This was further tested by sequencing one of the Gunnison small-bodied birds from blood-extracted DNA; this procedure also resulted in sequence of the A haplotype.

A \( \chi^2 \) statistic used to test for homogeneity of haplotype frequencies among all sampling locales in a 4 × 7 contingency table was significant (\( P < 0.005 \)). Removing the Gunnison (small-bodied) sample from the analysis showed that haplotype frequencies did not deviate from expectation among large-bodied locations, although the test approached significance (0.05 < \( P < 0.10 \)).

Maximum-parsimony analysis found 144 equally parsimonious trees (length 68 steps), the consensus of which contained two deep clades of haplotypes (clades I and II; Fig. 2). Each clade contained a group...
FIG. 1. Map of Colorado showing the approximate locations of Sage Grouse populations sampled in this study. Chevrons indicate mountain ranges, solid lines indicate rivers, and pie charts show the proportion of each haplotype found within a given population.

of haplotypes that was much less divergent from each other (\( \delta = 2.0\% \)) than from haplotypes in the other clade (\( \delta = 17.5\% \)). The maximum-parsimony consensus tree was topologically congruent with the tree generated by neighbor-joining distance analysis. The major characteristic of the neighbor-joining tree was the presence of two deeply divergent clades, each containing numerous shallow branches. The amount of divergence separating the two major clades was, on average, 15% greater than the average divergence among haplotypes within either clade. Haplotypes representative of both clades occurred in all populations of large-bodied grouse (Table 2). Some of the same haplotypes, from both clades, also have been found within the range of the western subspecies (Young and Benedict unpubl. data). Only haplotypes from clade I were found in small-bodied grouse. The ratio of transitions to transversions was approximately equal in all pairwise comparisons (\( \delta = 4.23 \)).

Discussion.—The 21 haplotypes detected formed two distinct clades; within each clade, branching was relatively shallow (Fig. 2). Two levels of interpretation can be made: (1) the deep divergence among the haplotypes provides insight into distant evolutionary history; and (2) the shallow differences (through analysis of haplotype frequencies) reveal recent history of extant populations. In theory, deeply divergent haplotypes can exist within a single large population without biogeographic isolation (Slatkin and Hudson 1991); however, because of the deep split between the two clades accompanied by only shallow branching within each clade (Fig. 2), more likely they represent an ancestral isolation of the Sage Grouse into two distinct populations. The fact that haplotypes from both clades are found at all large-bodied sampling sites further implies that extensive intermixing has occurred between previously isolated populations. Perhaps a long period of allopatry was followed by renewed contact between the two groups, allowing females to disperse and mix haplotypes. Avise et al. (1987) described such a pattern as "phylogeographic category II."

If mtDNA evolves at a clock-like rate, then the sequence data could be used to estimate the time at which the two clades began diverging. Following the estimation method outlined in Quinn (1992), we calculated that the sequenced 141-bp region that we describe evolves approximately 10.1 times faster than the average mitochondrial rate. This equates to a divergence rate of approximately 20.2% per million years for this specific sequenced region, using the average mitochondrial rate of 2% for geese (Shields and Wilson 1987). A few caveats about this estimate should be made. First, the estimate assumes a molecular clock that has the same rate in grouse as it does in geese. Also, because rates along the control region differ remarkably in different segments (Quinn and Wilson 1993, Wenink 1994), this estimate should not be extrapolated to other areas within the control region (or elsewhere). Such calibrations remain approximate and should be recalculated for the specific sequenced region, as was done here. Thus, because the two major clades differ by 17.5%, and some multiple substitutions likely remain uncorrected (Wakeley 1993), they began diverging at least 850,000 years ago, within the Pleistocene, which was a period of dynamic glacial expansion and contrac-
FIG. 2. Phylogram of the strict-consensus tree derived from the 144 shortest trees of equal length (68 steps) shows that the 21 haplotypes are separated into two deep monophyletic clades. The strict-consensus tree has a consistency index of 0.718, a retention index of 0.881, and a rescaled consistency index of 0.633. Bootstrap values >50% are shown along the branches of the tree. Haplotypes from clade I and clade II are found in all large-bodied Sage Grouse populations (Fig. 1).

Haplotype frequencies and distributions among the large-bodied populations of Sage Grouse provide interesting observations about their more recent evolutionary history. The 20 haplotypes observed can be categorized into three types of geographic distribution: (1) widely distributed, (2) narrowly distributed, and (3) rare (observed in a single population only). The frequencies among narrowly distributed haplotypes are generally low, and interestingly, they are typically shared only between adjacent sampling areas (Fig. 1). Given their close relationship to the four widespread haplotypes (A, B, C, and D), it is likely that narrowly distributed haplotypes represent new local mutational events that are being gradually transported to adjacent populations with which the original source exchanges genes. Rare haplotypes likely represent very recent mutational events that are still localized to the population in which they arose. Thus, narrowly distributed and rare haplo-
types may represent genetic markers that have evolved in situ since the post-isolation mixing of refugial populations, with significantly lower levels, or lack, of gene flow. This scenario is very similar to the "two-tiered history" of Song Sparrows (Melospiza melodia) proposed by Zink and Dittmann (1993). Under their scenario, either a pulse of high gene flow or continued low-level dispersal follows glacial retreat, with subsequent mutational changes to some haplotypes that remain geographically localized. Thus, the narrowly distributed haplotypes may provide valuable information about which regional populations are exchanging genes most frequently.

Haplotypes frequencies in the small-bodied population are quite different than those found in large-bodied populations, and haplotype diversity is much lower. Only two closely related haplotypes were observed: A (at high frequency) and G (at low frequency). Because G differs from A by a single transition and has been found nowhere else, it may represent a narrowly distributed or rare small-bodied haplotype (although 133 large-bodied birds were sequenced, G might also be present at very low frequency among large-bodied populations, remaining undetected due to sampling error). Small-bodied Sage Grouse may have originated from a founder event, or from isolation at the edge of a range, that occurred after both the Pleistocene divergence and the subsequent post-refugia mixing described earlier. If the original founder group was small, then it may have carried only the A haplotype, or it may have rapidly lost all but the A haplotype after the founding event through genetic drift. A larger founding group could have carried many haplotypes, and therefore may have established haplotype-diverse populations of small-bodied birds throughout a southwestern range that later experienced bottlenecks, perhaps due to habitat fragmentation and loss from human intrusion/agricultural practices. If the bottleneck scenario is true, then different small-bodied populations potentially could be fixed for different haplotypes. Although our current data set does not allow us to distinguish between these two hypotheses, it does indicate that the appearance of the small-bodied Sage Grouse was recent with respect to the events of the Pleistocene divergence and subsequent remixing.

Because the widespread A haplotype predominates in the Gunnison small-bodied population and is shared at lower frequency among large-bodied birds, the small-bodied birds must have been isolated in recent evolutionary time; otherwise, the large- and small-bodied populations would no longer carry an identical haplotype (A). This presumes a rapid evolutionary substitution rate for the sequenced region; as discussed, we believe this rate to be approximately 20.2% per million years. Otherwise, over longer periods of time, patterns of mutational change, genetic drift, and subsequent fixation eventually would lead to monophyly of distinct haplotypes within each population. If this interpretation is correct, then Sage Grouse present another example of a taxon in which plumage and size appear to have evolved faster than mtDNA monophyly (Zink and Dittmann 1993). Given the observation that a very small percentage of males in a given lek mate successfully (Young 1994), sexual selection could have been the driving force that generated the rapid morphological and behavioral changes.

The most important observation from our study is that haplotype frequencies in the Gunnison population, when compared with large-bodied populations, indicate the absence of gene flow between small- and large-bodied birds. The large-bodied Sage Grouse had relatively high levels of haplotype diversity and a much lower frequency of the A haplotype (Fig. 1, Table 1). If gene flow was occurring between small-bodied and large-bodied Sage Grouse, an array of haplotypes in Gunnison Basin similar to those found in the large-bodied populations would be expected. Because haplotype frequencies differed significantly (P < 0.005), we conclude that gene flow is effectively absent. This finding is consistent with the proposal of Braun and Young (1995) that the small-bodied Sage Grouse of the Gunnison Basin is a separate species (AOU 1998) if the biological species concept or the evolutionary species concept is applied.

Acknowledgments.—We thank Sara Oyler, Michael Wink, two anonymous reviewers, and Bob Zink for helpful comments that improved the manuscript, and Nick Benedict for supplying unpublished sequence data. Funding for this study was provided through contracts to TWQ from the Colorado Division of Wildlife. TWQ was partially supported by an NSF grant during the preparation of the manuscript.

Literature cited


Received 9 February 1998, accepted 13 November 1998. Associate Editor: R. M. Zink