Molecular Ecology (2004) **13**, 1911–1922 doi: 10.1111/j.1365-294X.2004.02115.x

Phylogeographic structure, gene flow and species status in blue grouse (*Dendragapus obscurus*)

GEORGE F. BARROWCLOUGH,* JEFF G. GROTH,* LISA A. MERTZ* and R. J. GUTIÉRREZ†
*Department of Ornithology, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024, USA, †Department of Fisheries and Wildlife, University of Minnesota, St Paul, Minnesota 55108, USA

Abstract

We investigated the genetic population structure and species status of a relatively sedentary bird that is a permanent resident of western North American forests, the blue grouse (*Dendragapus obscurus*). Phylogenetic analysis of complete mitochondrial control region DNA sequences resulted in the identification of three basal clades of haplotypes that were largely congruent with well-known biogeographical regions. These clades corresponded to the parapatric sooty (*D. o. fuliginosus*) and dusky (*D. o. obscurus*) subspecies groups of blue grouse plus a previously unrecognized division between northern and southern dusky grouse populations; the latter does not correspond closely to any currently recognized subspecies boundary. Approximately 66% of the total genetic variation was distributed among these three regions. Maximum likelihood estimates of gene flow between the regions were low or asymmetric; gene flow has been insufficient to prevent genetic divergence between dusky and sooty grouse. Estimates of gene flow among populations within sooty grouse were large except across the Columbia River valley. Among populations of dusky grouse, estimates of gene flow were heterogeneous and asymmetrical, reflecting large-scale fragmentation of the distribution due to landscape features and associated vegetation. Genetic, morphological and behavioural evidence suggest that sooty and dusky grouse are species-level taxa; the specific status of a third clade remains ambiguous.

Keywords: blue grouse, *Dendragapus*, gene flow, phylogeography, speciation

Received 23 December 2003; revision received 19 March 2004; accepted 19 March 2004

Introduction

The effects of Quaternary climate changes and glacial cycles on speciation and the distribution of the fauna and flora of western North America have long been of interest to evolutionary biologists (e.g. Mengel 1964; Remington 1968; Hubbard 1973). Unfortunately, such early studies had little power to detect sequential vicariant events and subsequent range expansions because, at that time, it was difficult to develop accurate phylogenetic hypotheses for closely related taxa. However, the recent use of phylogeographical (Avise 2000) and coalescent (Barrowclough & Groth 1999; Zink *et al*. 2000) analyses of mtDNA sequences has made possible more detailed studies of historical biogeography (Zink 1996; Hewitt 2000; Arbogast & Kenagy 2001). Although comparative reviews of the postglacial development of the European biota have appeared (Hewitt 1999), detailed hypotheses for western North America have not advanced substantially since the 1970s (but see Bermingham *et al*. 1992). In part this is because the positions of glacial refugia at lower latitudes are less geographically constrained in North America than they are in Europe; in addition, phylogeographical studies are not yet numerous. Here we use phylogeographical methods to investigate the genetic structure and population history of blue grouse, *Dendragapus obscurus*, in western North America. This species is relatively sedentary, has restrictive winter habitat requirements, and hence may act as an exemplar for faunal elements associated with coniferous forests (Zwickel 1992).

Blue grouse are well-studied game birds endemic to the mountainous regions of the western United States and Canada (Johnsgard 1983; Zwickel 1992). They are permanent residents, but many populations show relatively short distance, mainly altitudinal migratory movements between summering and wintering habitat; dispersal distances of

Correspondence: George Barrowclough. Fax: 212 7695759; E-mail: gfb@amnh.org

up to 50 km have been recorded but median dispersal is usually less than 15 km (Zwickel 1992). They require the needles of coniferous trees for winter forage, and hence are restricted to regions with high elevation forest in much of their range. That requirement, together with the lack of long-distance migratory behaviour, tends to isolate populations to relatively discrete mountain ranges and results in the elimination of large expanses of high desert, sagebrush, shrubsteppe, grassland and even isolated aspen monocultures from their distribution. Thus, in many parts of interior western North America, the range of blue grouse is fragmented extensively due to winter habitat requirements (Zwickel 1992). Blue grouse are not bred easily in captivity and are consequently not readily raised, released or translocated for management purposes; they do not fluctuate in regular cycles of population density as do red and ruffed grouse (Zwickel 1992).

One species of blue grouse is currently recognized (AOU 1998); however, the history of its taxonomic status is complicated. The various members of the 'sooty' subspecies group (*D. o. fuliginosus*, *D. o. sierrae*, *D. o. sitkensis* and *D. o. howardi*: Sierra Nevada, Cascade Mountains, and coastal ranges to southeast Alaska) were described originally as members of a single polytypic species (Baird & Ridgway 1873) that included the 'dusky' subspecies group (*D. o. obscurus*, *D. o. richardsonii*, *D. o. oreinus* and *D. o. pallidus*: Rocky Mountain and Great Basin ranges from Arizona and New Mexico north to Yukon and Northwest Territories). The sooty and dusky groups differ in several major traits. These include morphological differences such as the colour of the bare skin on the male apteria, the presence or absence of a distinct tail band, the number of rectrices and the colour of the downy young. In addition, the two groups differ in their mating vocalizations and other courtship behaviours (Bendell & Zwickel 1984; Zwickel *et al*. 1991; Zwickel 1992).

The groups were split into two species, *D. obscurus* and *D. fuliginosus* (AOU 1931), but subsequently reunited by Peters (1934) and the AOU (1944). These taxonomic actions were controversial and various authors (e.g. Grinnell & Miller 1944) expressed doubts about the classification. The literature (Munro & Cowan 1947; Jewett *et al*. 1953) cited as providing justification for lumping sooty and dusky grouse (e.g. Mayr & Short 1970) did not include any of the quantitative description and analysis that has come to be expected for avian contact zones (e.g. Rising 1983). Rather, species accounts for blue grouse refer only to the existence of a few intermediate specimens from a small number of localities in the mountains of north-central Washington and south-central British Columbia (Pitelka 1941; Munro & Cowan 1947; Aldrich 1963). Thus, the species status of blue grouse has not been resolved.

We investigated the genetic structure, species status and population history of blue grouse using control region sequences of the mitochondrial genome. This region evolves rapidly, and previous studies in birds (Barrowclough *et al*. 1999), including gallinaceous birds such as grouse (Holder *et al*. 2000), have indicated that it can provide resolution of the intraspecific structure of a gene tree with a relatively high ratio of haplotypes to individuals.

Materials and methods

Samples and sequencing

We obtained tissue samples (muscle, liver, or growing feather) from 14 localities that represent both major subspecies groups of blue grouse (Fig. 1). The approximate location and locality acronym (bold type) for the samples were: (1) *D. o. obscurus* subspecies group: New Mexico: Rio Arriba Co. and San Miguel Co. (**NM**); Colorado: Routt Co. (**CO**); Utah: Utah Co. (**UT**); Oregon: Umatilla Co., Blue Mtns. (**BLU**); Oregon: Wallowa Co., Wallowa Mtns. (**WAL**); Washington: Stevens Co., Selkirk Mtns. (**SEL**); Washington: Okanogan Co., north of Winthrop (**OKA**); British Columbia: near Clinton (**BC**); (2) *D. o. fuliginosus* subspecies group: Washington: Gray's Harbor Co. and Jefferson Co., Olympic Peninsula (**OLY**); Washington: Thurston Co., south of Puget Sound (**PUG**); Oregon: Douglas and Jackson Cos., Cascade Range (**CAS**); Oregon: Lake Co., Warner Mtns. (**WAR**); California: Humboldt Co. (**HUM**); and California: Mendocino Co. (**MEN**). Sample sizes are listed in Appendix I. Samples were frozen or stored in either a DMSO/EDTA/NaCl buffer (Seutin *et al*. 1991) or an ethanol solution. Single samples of greater prairie chicken (*Tympanuchus pinnatus*), lesser prairie chicken (*T. pallidicinctus*), sharp-tailed grouse (*T. phasianellus*) and greater sage grouse (*Centrocercus urophasianus*) were used as outgroups, based on a recent molecular-based classification of the Tetraoninae (Gutiérrez *et al*. 2000).

The entire control region was amplified using standard polymerase chain reaction (PCR) protocols for avian mtDNA (Barrowclough *et al*. 1999) with primers (Table 1) placed in the flanking tRNA-GLU (BG1) and 12sRNA (D2) genes. Internal primer pairs (BG1 to BG6 and BG7 to D4mod) were used subsequently to obtain two fragments short enough for efficient sequencing on ABI377 and ABI3100 automated

Table 1 Sequences of primers used in amplifying DNA from mitochondrial control region of blue grouse. Primers are listed in the 5′–3′ direction. Even-numbered primers anneal to the light mtDNA strand; odd-numbered primers anneal to the heavy strand

BG1	GGATTACGGCTTGAAAAGCCA
BG6	GATTCCCCATACACGCAAAC
BG7	CTCCCCTCTTTAGTCCGT
D2	GATACTTGCATGTATATGTCTAG
D4mod	GGTACCATCTTGGCATCTT

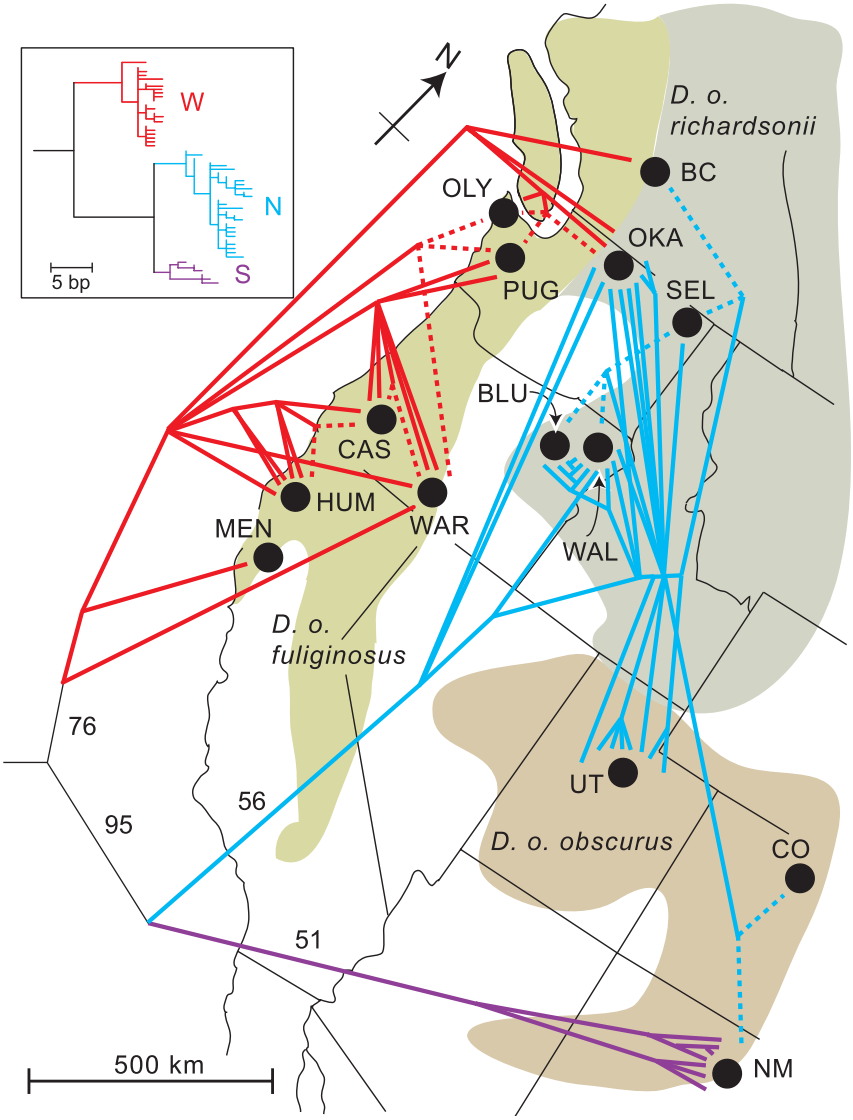

Fig. 1 Majority-rule consensus tree for blue grouse mtDNA haplotypes superimposed on the geographical range of the species. Dashed lines indicate single haplotypes found in two or more localities. Approximate distributions of three major subspecies groups are indicated. Geographical locations of samples are indicated by black dots with associated locality acronyms. Bootstrap percentages are indicated for basal nodes. Branch lengths are not drawn to a common scale for large diagram. Inset: phylogram of one of the equal length trees showing relative branch lengths.

sequencers. Each individual was sequenced on both strands and the two sequences compared using SEQUENCHER software for ambiguous sites. Sequences were aligned by eye. The sequenced portions of the glutamic acid and phenylalanine t-RNAs were removed before analysis.

Analysis

Genetic variation within population samples was estimated using the number of haplotypes found in the sample, the percentage of variable sites, and the nucleotide diversity: Π (Nei 1987); we also computed the bootstrap 95% confidence interval (1000 replicates) for the latter. Tajima's (1989) *D* statistic was computed as a measure of divergence of the sequences from expectations based on neutral theory; its significance was determined by comparison to a beta distribution.

We investigated historical demographics of populations by plotting mismatch distributions (Slatkin & Hudson 1991) and comparing them to Poisson distributions with identical means using one-sample Kolmogorov–Smirnov tests; exponentially growing populations are expected to have mismatch distributions resembling Poisson distributions and stable populations are expected to have ragged distributions (Slatkin & Hudson 1991). Ramos-Onsins & Rozas (2002) suggested that the F_S statistic (Fu 1997) has greater power for detecting population growth with moderate sample sizes than do many other estimators; consequently we computed F_S for the individual population samples.

The fraction of the total genetic variation distributed among populations was estimated using G_{ST} (Holsinger & Mason-Gamer 1996) statistics. Confidence intervals on this variance partitioning were computed by bootstrapping over individuals and maintaining the original population

sample sizes in each replicate. G_{ST} statistics were also computed using a hierarchical approach with major geographical clades treated as regions and the genetic variation partitioned into variance among regions and among populations within regions.

The computer program PAUP*, version 4.0b10 (Swofford *et al*. 1996; Swofford 1998) was used to search for shortest trees among all haplotypes and to obtain bootstrap support values for branches. Single base indels were treated as a 'fifth base' and ambiguity codes were treated as uncertainty in our analysis. When searching for minimum length trees, we added taxa in a random stepwise sequence and repeated the additions 25 times; tree-bisection–reconnection (TBR) branch swapping was used in each heuristic search. In a bootstrap analysis, a single random addition of haplotypes was used in each of 100 pseudoreplicates with a time limit of three hours per replicate on a 1 GHZ PC.

Estimates of gene flow among the populations having the largest samples were obtained both by using the reciprocal F_{ST} approach and by using the program MIGRATE, version 1.6.7 (Beerli & Felsenstein 2001). MIGRATE estimates directional (i.e. including asymmetric) gene flow using a Monte Carlo, Markov chain, maximum-likelihood procedure; the software is relatively slow and the results are sensitive to small sample sizes. Consequently, we combined some population samples, based on geographical proximity and the estimates of G_{ST}, in order to increase sample size and decrease the number of parameters being estimated; in addition, we specified that gene flow estimates be limited to geographically adjacent localities (stepping-stone model of population structure) to minimize the loss of degrees of freedom and because extremely long-distance dispersal would be unlikely for this species. In using MIGRATE, 10 short Markov chains of length 50 000 were used with sampling at every 100 trees; this was followed by a long Markov chain of length 10 000 000 with sampling every 100 trees. For both the short and long chains, 10 000 trees were discarded as initial 'burn-in'. An adaptive heating regime with four parallel chains and initial relative temperatures of 1.0, 1.5, 2.0 and 3.0 was used. A transition to transversion ratio of 10 and an error rate of 0.1% were assumed, and nucleotide ratios were estimated from the data. The runs were initiated with a UPGMA tree and initial estimates of θ and gene flow were based on estimates of F_{ST}. The program was run five times, starting with different random seeds in order to determine the stability of the results, and the results of the five runs were averaged.

Results

Blue grouse mitochondrial control region

We obtained complete control region sequences from 147 blue grouse plus four outgroups. Blue grouse control regions varied from 1147 to 1149 base pairs (bp) in length; alignment required the inference of single base indels at four positions. Alignment with the samples of *Tympanuchus* and *Centrocercus* resulted in a total of 1153 sites; this alignment was used in computing statistics such as percentage of variable sites and nucleotide diversity, etc. In one haplotype a single base was assigned an ambiguity code (*Y*) because chromatograms from both strands showed *C* and *T* peaks of approximately equal intensity. Although this suggested a low level of heteroplasmy, we did not observe the pattern of multiple heterozygous sites that might suggest amplification of a nuclear copy of the mitochondrial genome. In addition, our recovery of 60 haplotypes (see below) from 147 individuals argues strongly against nuclear copies in this study. The haplotype sequences have been deposited in GenBank (Accession nos AY570302–AY570361: blue grouse and AY569303–AY569306: outgroups).

Genetic variation within and among populations

There were 60 unique haplotypes in the 147 blue grouse sequences; 26 of the haplotypes occurred in two or more individuals and seven of those were shared among two or more population samples (Appendix I). The variation among these haplotypes occurred at 63 sites that were distributed heterogeneously along the sequence. Two-thirds of the variation occurred in the 200 bp stretch from base 150–350; Baker & Marshall (1997) have called this Domain I of the avian mitochondrial control region.

We obtained several estimates of the amount of variation within the larger population samples (Table 2). Nucleotide diversity ranged from zero to 0.0075 and was largest in the Okanogan County, Washington sample. All 17 birds from the Colorado sample had identical haplotypes; this sample had been obtained from three different groups of collectors during a 3-year period of time over many kilometres of space; consequently it is unlikely that the 17 individuals were members of one or two family groups. No pattern of variation in any of the population samples differed from neutral expectations based on Tajima's *D* statistic.

We computed G_{ST} over the largest population samples (Table 2). The estimate of 0.75 (Table 3) indicated that three-quarters of the total nucleotide diversity was distributed among, rather than within, populations; the bootstrap confidence interval was consistent with this high value. A hierarchical analysis (Table 3) indicated that most of this among-population variance was distributed among geographical regions corresponding approximately to subspecific taxa; however, even within subspecies groups, more than 20% of the variation was partitioned among population samples. These results indicated substantial genetic population structuring in blue grouse. We also computed estimates of G_{ST} between all pairs of populations; the matrix of values is not reported here. However, the point estimates

Table 2 Estimates of within-population variability of control-region sequences from blue grouse mtDNA. Number of individuals sampled, the observed number of haplotypes, percent of positions that vary, nucleotide diversity (Π) and its 95% bootstrap confidence interval, Tajima's *D* and its significance, and Fu's F_S and its significance, are all reported

Population	New Mexico	Colorado	Utah	
No. individuals	17	17	20	
No. haplotypes	8	1	8	
% Variable sites	1.47	0.00	1.21	
Π	0.0045	0.0000	0.0021	
(95% CI)[a]	(0.0028–0.0054)	NA	(0.0009–0.0029)	
D	0.081[b]	NA	–1.439[b]	
F_S	0.323[b]	NA	–1.580[b]	
	Blue Mtns	Wallowa Mtns	Selkirk Mtns	Okanogan
No. individuals	7	13	6	20
No. haplotypes	4	8	3	11
% Variable sites	0.61	1.21	0.35	2.69
Π	0.0028	0.0032	0.0015	0.0075
(95% CI)[a]	(0.0010–0.0031)	(0.0020–0.0039)	(0.0006–0.0020)	(0.0026–0.0106)
D	0.690[b]	–0.777[b]	–0.057[b]	–0.055[b]
F_S	0.812[b]	–1.599[b]	0.758[b]	0.159[b]
	Puget	Olympic	Cascade	Warner
No. individuals	11	4	6	15
No. haplotypes	4	3	5	6
% Variable sites	0.52	0.35	0.52	0.87
Π	0.0020	0.0022	0.0024	0.0028
(95% CI)[a]	(0.0011–0.0024)	(0.0000–0.0023)	(0.0008–0.0028)	(0.0012–0.0036)
D	0.539[b]	1.365[b]	0.375[b]	0.154[b]
F_S	1.127[b]	0.461[b]	–1.418[b]	0.436[b]
	Humboldt			
No. individuals	8			
No. haplotypes	6			
% Variable sites 1.47	0.69			
Π	0.0025			
(95% CI)[a]	(0.0014–0.0027)			
D	–0.358[b]			
F_S	–1.724[b]			

[a]2.5%–97.5% bootstrap confidence interval.
[b]$P > 0.05$.

between the closely spaced Blue Mountains and Wallowa Mountains samples and between the nearby Olympic Peninsula and Puget Sound samples were zero. Consequently, those population pairs were combined for analyses of gene flow and mismatch distributions.

Phylogeography

We searched for minimum length trees for the 60 haplotypes of blue grouse plus the four outgroups using a heuristic search and TBR branch swapping with PAUP*. This analysis yielded 47 817 trees of length 233. A strict consensus of these resulted in a tree consisting of monophyletic blue grouse in which there were three major (based on branch lengths) clades; these had modest to substantial support based on the bootstrap analysis (Fig. 1).

When superimposed on geography, the three major divisions delineate parapatric clades of northeastern, western and southern localities. The western, largely Pacific coast clade (labelled W in the phylogram in the Fig. 1 inset) corresponds to the traditional sooty grouse. The northeastern (labelled N) clade plus the southern clade (labelled S and restricted to New Mexico) correspond to the dusky grouse. There was some sharing of haplotypes from these clades at three localities where the ranges of the taxa are parapatric. First, in the New Mexico sample, one individual (of 17) had

Table 3 Hierarchical estimates of G_{ST} among blue grouse population samples based on mtDNA control region sequences[a]

Overall G_{ST} (among 12 populations)	0.75 (0.75–0.81)
Among three regions	0.66 (0.62–0.73)
Among pops. within *D. o. fuliginosus*	0.21 (0.17–0.46)
Among pops. within *D. obscurus*[b]	0.30 (0.28–0.43)

[a]95% bootstrap confidence intervals in parentheses.
[b]Excluding New Mexico sample, see text.

a haplotype identical to the haplotype fixed in the Colorado sample. Second, the British Columbia sample of two birds included a haplotype also found in the Selkirk Mountains population (dusky grouse) and a haplotype related closely to haplotypes found at the Puget Sound and Olympic Peninsula sites (sooty grouse). Finally, the Okanogan sample of 20 birds included 17 individuals with haplotypes belonging to the dusky grouse clade and three individuals with haplotypes within the sooty grouse clade.

Gene flow

An estimate of gene flow can be obtained from F_{ST}-like statistics (in our case, G_{ST}) given the assumption of an island model of population structure (e.g. Wright 1969; Rockwell & Barrowclough 1987; Crochet 2000). For that model with mitochondrial genes, the number of females exchanged per generation between populations is $mN_{e\,♀} = 0.5(G_{ST}^{-1} - 1)$. The estimates of pairwise gene flow among populations obtained in this manner (not shown) were of the order of one to four females per generation among geographically adjacent populations of sooty grouse. Within dusky grouse the estimates were of the order of one to two females per generation for populations from Utah to Washington state, but were less than one between New Mexico and Colorado and between Colorado and Utah. The estimates of gene flow between dusky and sooty grouse population samples all were less than one female per generation.

The F_{ST} approach to estimating gene flow has a number of limitations (Bossart & Prowell 1998; Whitlock & McCauley 1999), including the implicit assumption that gene flow is symmetrical, all populations are identical in size, and an island model of population structure applies. Maximum likelihood methods relax those three assumptions and, hence, are more generally applicable (Neigel 2002), but still retain a number of limitations as currently implemented (Wakeley 2001). We obtained maximum likelihood estimates of gene flow between populations of blue grouse using MIGRATE and a stepping-stone model of population structure (Fig. 2). Initial runs of the program yielded variable results; point estimates of gene flow only became stable to different random seeds after long chains were made very long (10 000 000 steps) and adaptive heating was used. A few runs of MIGRATE resulted in estimates of μN_e and mN_e that were infinite; those runs were not retained. The values of mN_e shown in Fig. 2 are averages of five runs each starting with different random seeds.

For population samples within the western clade (sooty grouse), the averaged estimates generally indicated gene flow greater than 1 between localities in Oregon and northern California. There was no evidence of significant gene flow between the Puget Sound sample and those in Oregon and California; these are separated by the wide Columbia River valley.

Estimates of gene flow among populations within the northeastern and southern clades (dusky grouse) were heterogeneous and asymmetrical. For example, the estimate of gene flow from Colorado to New Mexico was one, but vanishingly small in the opposite direction. In addition, the Utah population seems to generate emigrants but receive few immigrants.

With regard to gene flow between sooty and dusky grouse, there was no evidence of any gene flow between the mountain ranges of northeastern Oregon and those of south-central Oregon (Warner Mountains). Similarly, although estimates of gene flow into the Okanogan sample were greater than one from both the Puget Sound and Selkirk Mountains samples, they were much less than one from Okanogan to either of those populations. Thus, the apparent introgression between these taxa is confined to central Washington.

Population demography

We used several methods to infer the long-term demographic history of these populations. Mismatch distributions (Fig. 2) appear ragged for most populations with sample sizes greater than 10; such results usually are taken to indicate a stationary long-term population size (Barrowclough & Groth 1999). An exception is the combined Wallowa and Blue Mountain samples, which appears to be unimodal with some sampling error. The Selkirk Mountains and Humboldt samples have unusual mismatch distributions that suggest population growth, but they also were represented by small samples. The distribution shown for the Okanogan sample (Fig. 2) does not include comparisons involving the three haplotypes from the western clade; those would make the distribution more ragged, but are the signature of introgression rather than population growth (e.g. Petit *et al.* 2003). Kolmogorov–Smirnov tests (not shown) generally confirm these observations; of the samples greater than 10, only the Wallowa population had a mismatch distribution not significantly different ($P < 0.05$) from a Poisson. The results of the Kolmogorov–Smirnov tests must not be given too much significance by themselves, however, as the data used to produce such distributions are not independent (e.g. Barrowclough & Groth 1999).

The results of Fu's (1997) F_S statistic (Table 2) indicate that none of the samples departed from the expectations of

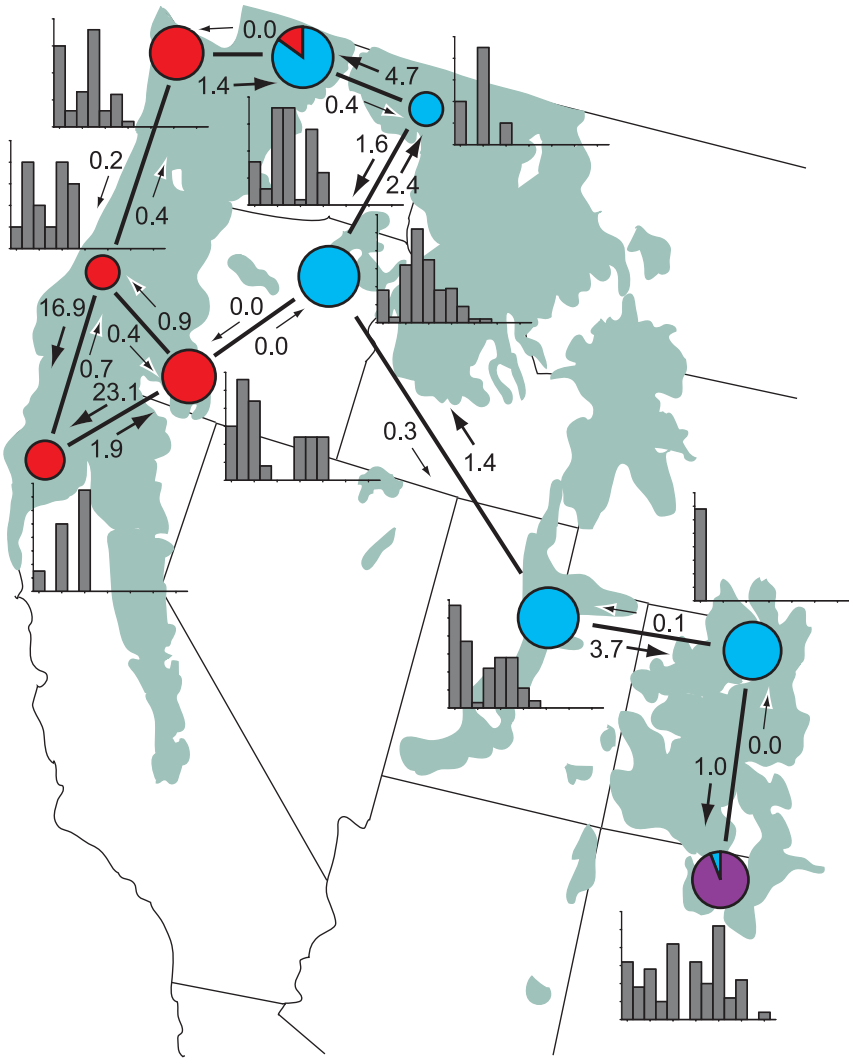

Fig. 2 Estimates of gene flow and demography for blue grouse population samples. Area shaded in green indicates appropriate habitat. Pie diagrams show proportion of locality sample represented by each of the three major haplotype clades shown in Fig. 1; area of pie diagram is proportional to sample size. Histograms show mismatch distributions for each population sample (Wallowa and Blue Mountain localities and Olympic Peninsula and Puget Sound localities were combined for this analysis). Arrows indicate directional gene flow (in units of $mN_{e\varphi}$ per generation) between adjacent localities based on coalescent likelihood estimates; size of arrow scaled to gene flow estimate (i.e. large arrows correspond to one or more females per generation, thin arrows correspond to estimates of less than one female per generation.

stable population sizes, although the value for the combined Wallowa and Blue Mountain samples came close ($0.1 > P > 0.05$).

Discussion

Genetic population structure

We found a major genetic division between populations representing what were referred to traditionally as sooty and dusky grouse. These results include an estimate of G_{ST} of 0.66 and separate clades of haplotypes, with moderate bootstrap support, that correspond to geographically parapatric ranges. Estimates of gene flow between the two regions are nearly zero. Wright (1969) showed that levels of gene flow of less than approximately one individual per generation were insufficient to prevent genetic differentiation.

There is evidence for some genetic structuring within sooty grouse ($G_{ST} = 0.21$). The estimates of gene flow among the population samples south of the Columbia River are quite large in some cases. However, the river valley may have acted as a historical barrier to gene flow; our estimates of gene flow from the southern end of Puget Sound to the Oregon samples are small. Nevertheless, the barrier cannot be complete as an identical haplotype was found on both sides of the river.

Genetic structuring also exists among the population samples of dusky grouse ($G_{ST} = 0.30$). The estimates of gene flow among most of the populations sampled are asymmetrical and in many cases small. The New Mexico sample has been isolated from the other dusky samples for a sufficient period of time to become nearly monophyletic. The Colorado sample was comprised of a single haplotype and reflects either a small effective size or has recently been through a bottleneck.

It is perhaps not surprising that a large, relatively sedentary bird with modest population density would have substantial genetic structure across a large geographical area

involving a fragmented distribution. For example, Baba *et al*. (2002) described significant genetic structuring, also based on control region sequences, in another grouse, *Bonasa bonasia*, across Eurasia. In that case, samples from Japan and extreme eastern Russia were quite divergent genetically from samples from the western Palaearctic. In a second case involving the mtDNA control region and another grouse, the rock ptarmigan (*Lagopus lagopus*), population samples on the Aleutian Islands and both sides of the Bering Strait had very substantial structuring with an estimated F_{ST} of 0.93 (Holder *et al*. 2000). The proportions of total genetic variation distributed among populations for the blue grouse and rock ptarmigan are among the largest known for birds (e.g. Zink 1997).

Historical reconstruction

We have found substantial genetic structuring of blue grouse populations; an understanding of the historical processes that led to this pattern would be useful in developing generalizations concerning genetic differentiation and speciation. Here we rely on the geographical position of robust nodes of the tree of haplotypes and estimates of demographic parameters to reconstruct the Pleistocene and Recent history of blue grouse populations.

There are three major clades of haplotypes, each of which corresponds to a geographical region; these are defined by two nodes on the tree of haplotypes. The most basal — hence earliest — of the relevant nodes divides the phylogenetic tree into eastern and western populations that correspond to the dusky and sooty forms of blue grouse. The second division — hence a later event — split the eastern clade of dusky grouse into northern and southern forms. The east–west division corresponds to similar patterns found in other birds (e.g. Johnson & Cicero 2002; Ruegg & Smith 2002) and many other organisms (Remington 1968) and has been interpreted to be the result of Pleistocene glaciations on the distributions of flora and associated faunas (Mengel 1964). Other birds, such as migratory wood warblers, are not as genetically divergent between the Rocky Mountains and the Pacific ranges (e.g. Kimura *et al*. 2002). The second division in central or southern Colorado appears to be novel and does not correspond closely to any well-known suture zone or position of geographical replacement of closely related taxa. However, southern Colorado is characterized by extensive fragmentation of alpine and high elevation vegetation (DeChaine & Martin 2004) and our results suggest that there has been long-term geographical isolation among blue grouse populations in the southern Rocky Mountains. This is somewhat surprising because the 'canonical' glacial refugia hypothesis is probably reversed in that region for blue grouse.

Blue grouse are cold-adapted birds whose high elevation coniferous habitat was probably less fragmented during glacial epochs than at present. During ice ages, Douglas fir (*Pseudotsuga menziesii*) and other similarly adapted trees, together with their faunas, are known to have had more widespread distributions at lower elevations than they have today (Grayson 1993; DeChaine & Martin 2004); consequently, blue grouse populations are also apt to have been more widespread and continuous. In some sense, southern populations of blue grouse are now restricted to refugia and unlikely to expand until the next ice age.

Drovetski (2003) recently suggested that the mitochondrial control region has been evolving at a rate of 5–10% per million years in grouse. After correcting for within population variation, our data indicate that sooty and dusky grouse are approximately 1.7% divergent and the northern and southern clades of dusky grouse are approximately 0.5% divergent; using Drovetski's (2003) average divergence rate, these correspond to divergence times of 240 000 and 70 000 years ago, respectively. The timing of avian speciation events is controversial (e.g. Klicka & Zinck 1999); nevertheless, given Drovetski's rate of divergence, our data are consistent with blue grouse divergences in the late Pleistocene.

The estimates of nucleotide diversity are similar for all the sooty grouse populations and no mismatch distributions or indications of growth can be used to distinguish between refugial and expansion portions of the range. In addition, the basal haplotypes do not congregate in any one population sample or geographical region. Thus, evidence for the location of a Pleistocene refuge for sooty grouse is lacking; either the signature of a refuge has been lost, we failed to sample it, or the entire sampled region comprised a single refuge. The latter is plausible for cold-tolerant birds such as blue grouse because suitable vegetation may have been present in much of this region during the late Pleistocene (e.g. Pielou 1991). Samples from the southern portion of the range in the Sierra Nevada and from north coastal British Columbia might help to establish this history.

Similarly, the dusky grouse populations show little evidence of growth, with the possible exception of the Wallowa population. Only the Okanogan sample had nucleotide diversity markedly greater than other populations, but this was caused by the presence of haplotypes introgressing from adjacent populations of sooty grouse; it was not due to basal haplotypes within the dusky grouse clade. Thus, there was no evidence for refugial vs. expansion populations among our dusky grouse samples. Again, these are cold-adapted birds that may well have occurred in much of their range in the United States throughout recent glacial events. Samples from the Canadian Rocky Mountains might prove informative in reconstructing the history of these birds but, at present, the available data suggest that both sooty and dusky blue grouse have been present throughout the sampled portion of their range for a period of time equal to at least the past several tens of

thousands of years or more (that is, the time it takes for coalescence to result in ragged mismatch distributions).

Species status

The geographical extent of possible hybridization between sooty and dusky grouse has not been documented quantitatively. Rather, there are anecdotes in the literature about apparent hybrid individuals from Washington (Jewett *et al*. 1953) and British Columbia (e.g. Munro & Cowan 1947). Our data from Washington indicate that introgression seems restricted to an area close to the point of morphological parapatry. Haplotypes from dusky grouse were not found in western Washington (Puget Sound), nor were haplotypes of sooty grouse found in eastern Washington (Selkirk Mountains).

We demonstrated that there are three basal clades of blue grouse mtDNA haplotypes; two of these correspond to taxa that were once recognized as species (e.g. AOU 1931) and were referred to as sooty grouse and dusky grouse in publications at that time (e.g. Bent 1932; Pitelka 1941). A number of authors, including Bent (1932), Bendell & Zwickel (1984) and Zwickel *et al*. (1991), described differences in downy young colouration, adult plumage pattern, bare skin colour and vocal and other courtship behaviours that are diagnostic for each taxon. For neutral or nearly neutral characters, such as mitochondrial DNA sequence variation, it takes approximately $2N_{e\,♀}$ generations for isolated populations to become monophyletic. Consequently, the existence of these clades indicates that allopatric dusky and sooty grouse were isolated for many tens of thousands of years at some time in the past. Their parapatric ranges probably represent an example of a northward extension of Remington's (1968) Pacific-Rocky Mountain suture zone. Our mapping of haplotypes onto geography and our estimates of gene flow suggest that current contact between the two taxa has not resulted in introgression across any substantial portion of the range. Instead, introgression is restricted locally and estimates of gene flow are insufficient to prevent further genetic differentiation. The molecular and morphological differences are sufficient for recognition of phylogenetic species (Zink & McKitrick 1995). In terms of the biological species concept (BSC), the differences in mating displays and vocalizations suggest that aspects of reproductive isolating mechanisms are in place. Earlier advocates of the BSC lumped these taxa; however, in terms of the new, 'comprehensive' BSC of Johnson *et al*. (1999), they would probably be considered species. For example, they are more divergent than are greater sage-grouse and Gunnison sage-grouse (Young *et al*. 2000) in terms of genetic differentiation (hence period of isolation), mating system and morphology.

The species status of the two haplotype clades within dusky grouse is more complicated. The haplotype trees and estimates of gene flow suggest limited genetic introgression between our population samples from New Mexico and Colorado, and probably a long historical division between them. Perplexingly, the major morphological change within the *D. o. obscurus* group of subspecies occurs between nominate *D. o. obscurus* and the remaining races to the north and west (*D. o. pallidus* and *D. o. richardsonii*). That nomenclatural division is largely based on a change in colour pattern on the tail; however, the phenotypic transition occurs 400–500 km north of the position of the haplotype transition. Additional samples will be required to elucidate the situation at the southern end of the distribution of the dusky grouse.

Acknowledgements

Paul Sweet, Chris Vogel and Robert Zink assisted with some of the fieldwork. We also are grateful to Dave Anderson, Jenny Barnett, Larry Caudill, Bob Dickerman, Mike Gregg, Dean Mitchell, Michael Pope, Greg Schirato, Michael Schroeder and the Division of Genetic Resources of the Museum of Southwestern Biology for providing us with additional tissue samples from critical localities. James Bendell and Fred Zwickel encouraged our research and provided information on blue grouse based on their long experience with these birds; they, along with Richard Banks, Patricia Brito, Parker Cane, Ned Johnson, and two anonymous reviewers, offered useful comments on earlier versions of this manuscript. This research was supported in part by the Leonard J. Sanford Trust, the Ambrose Monell Molecular Laboratory and the Lewis B and Dorothy Cullman Program for Molecular Systematic Studies, a joint initiative of the New York Botanical Garden and the American Museum of Natural History.

References

Aldrich JW (1963) Geographic orientation of AmericanTetraonidae. *Journal of Wildlife Management*, **27**, 528–545.

American Ornithologists' Union (AOU) (1931) *Check-list of North American Birds*, 4th edn. American Ornithologists' Union, Lancaster, PA.

American Ornithologists' Union (AOU) (1944) Nineteenth supplement to the *Check-list of North American Birds*. *Auk*, **61**, 441–464.

American Ornithologists' Union (AOU) (1998) *Check-list of North American Birds*, 7th edn. American Ornithologists' Union, Washington, DC.

Arbogast BS, Kenagy GJ (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, **28**, 819–825.

Avise JC (2000) *Phylogeography*. Harvard University Press, Cambridge, MA.

Baba Y, Fujimaki Y, Klaus S, Butorina O, Drovetsii S, Koike H (2002) Molecular population phylogeny of the hazel grouse *Bonasa bonasia* in East Asia inferred from mitochondrial control-region sequences. *Wildlife Biology*, **8**, 251–259.

Baird SF, Ridgway R (1873) On some new forms of birds. *Bulletin of the Essex Institute*, **5**, 197–201.

Baker AJ, Marshall HD (1997) Mitochondrial control region sequences as tools for understanding evolution. In: *Avian Molecular*

Evolution and Systematics (ed. Mindell DP), pp. 51–82. Academic Press, San Diego, CA.

Barrowclough GF, Groth JG (1999) Demographic inferences from coalescent patterns mtDNA sequences from a population of Mexican spotted owls. *Proceedings of the International Ornithological Congress*, **22**, 1914–1921.

Barrowclough GF, Gutiérrez RJ, Groth JG (1999) Phylogeography of spotted owl (*Strix occidentalis*) populations based on mitochondrial DNA sequences: gene flow, genetic structure, and a novel biogeographic pattern. *Evolution*, **53**, 919–931.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences USA*, **98**, 4563–4568.

Bendell JF, Zwickel FC (1984) A survey of the biology, ecology, abundance, and distribution of the blue grouse (genus *Dendragapus*). *Proceedings of the International Grouse Symposium*, **3**, 163–192.

Bent AC (1932) Life histories of North American gallinaceous birds. *United States National Museum Bulletin*, **162**, 1–490.

Bermingham E, Rohwer S, Freeman S, Wood C (1992) Vicariance biogeography in the Pleistocene and speciation in North American wood warblers: a test of Mengel's model. *Proceedings of the National Academy of Sciences USA*, **89**, 6624–6628.

Bossart JL, Prowell DP (1998) Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends in Ecology and Evolution*, **13**, 202–206.

Crochet PA (2000) Genetic structure of avian populations—allozymes revisited. *Molecular Ecology*, **9**, 1463–1469.

DeChaine EG, Martin AP (2004) Historical cycles of fragmentation and expansion in *Parnassius smintheus* (Papilionidae) inferred using mitochondrial DNA. *Evolution*, **58**, 113–127.

Drovetski SV (2003) Plio-Pleistocene climatic oscillations, Holarctic biogeography and speciation in an avian subfamily. *Journal of Biogeography*, **30**, 1173–1181.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Grayson DK (1993) *The Desert's Past. A Natural Prehistory of the Great Basin*. Smithsonian Institution Press, Washington, DC.

Grinnell J, Miller AH (1944) The distribution of the birds of California. *Pacific Coast Avifauna*, **27**, 1–608.

Gutiérrez RJ, Barrowclough GF, Groth JG (2000) A classification of the grouse (Aves: Tetraoninae) based on mitochondrial DNA sequences. *Wildlife Biology*, **6**, 205–211.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Holder K, Montgomerie R, Friesen VL (2000) Glacial vicariance and historical biogeography of rock ptarmigan (*Lagopus mutus*) in the Bering region. *Molecular Ecology*, **9**, 1265–1278.

Holsinger KE, Mason-Gamer RJ (1996) Hierarchical analysis of nucleotide diversity in geographically structured populations. *Genetics*, **142**, 629–639.

Hubbard JP (1973) Avian evolution in the aridlands of North America. *Living Bird*, **12**, 155–196.

Jewett SG, Taylor WP, Shaw WT, Aldrich JW (1953) *Birds of Washington State*. University of Washington Press, Seattle, WA.

Johnsgard PA (1983) *The Grouse of the World*. University of Nebraska Press, Lincoln, NB.

Johnson NK, Cicero C (2002) The role of ecologic diversification in sibling speciation of *Empidonax* flycatchers (Tyrannidae): multigene evidence from mtDNA. *Molecular Ecology*, **11**, 2065–2081.

Johnson NK, Remsen JV, Cicero C (1999) Resolution of the debate over species concepts in ornithology: a new comprehensive biologic species concept. *Proceedings of the International Ornithological Congress*, **22**, 1470–1482.

Kimura M, Clegg SM, Lovette IJ *et al.* (2002) Phylogeographical approaches to assessing demographic connectivity between breeding and overwintering regions in a Nearctic–Neotropical warbler (*Wilsonia pusilla*). *Molecular Ecology*, **11**, 1605–1616.

Klicka J, Zink RM (1999) Pleistocene effects on North American songbird evolution. *Proceedings of the Royal Society of London, Series B*, **266**, 695–700.

Mayr E, Short LL (1970) Species taxa of North American birds. *Publications of the Nuttall Ornithological Club*, **9**, 1–127.

Mengel RM (1964) The probable history of species formation in some northern wood warblers (Parulidae). *Living Bird*, **3**, 9–43.

Munro JA, Cowan IM (1947) *A Review of the Bird Fauna of British Columbia*. Special Publications of the British Columbia Provincial Museum. British Columbia Provincial Museum, Victoria, BC, Canada.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY.

Neigel JE (2002) Is F_{ST} obsolete? *Conservation Genetics*, **3**, 167–173.

Peters JL (1934) *Check-List of Birds of the World*, Vol. II. Harvard University Press, Cambridge, MA.

Petit RJ, Aguinagalde I, de Beaulieu JL *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Pielou EC (1991) *After the Ice Age*. University of Chicago Press, Chicago, IL.

Pitelka FA (1941) Distribution of birds in relation to major biotic communities. *American Midland Naturalist*, **25**, 113–137.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Remington CL (1968) Suture–zones of hybrid interaction between recently joined biotas. *Evolutionary Biology*, **2**, 321–428.

Rising JD (1983) The Great Plains hybrid zones. *Current Ornithology*, **1**, 131–157.

Rockwell RF, Barrowclough GF (1987) Gene flow and the genetic structure of populations. In. *Avian Genetics* (eds Cooke F, Buckley PA), pp. 223–255. Academic Press, London.

Ruegg KC, Smith TB (2002) Not as the crow flies: a historical explanation for circuitous migration in Swainson's thrush (*Catharus ustulatus*). *Proceedings of the Royal Society of London, Series B*, **269**, 1375–1381.

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analysis. *Canadian Journal of Zoology*, **69**, 82–90.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Swofford DL (1998) PAUP*. *Phylogenetic Analysis Using Parsimony (* and Other Methods)*, Version 4.0b4a. Sinauer Associates, Sunderland, MA.

Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: *Molecular Systematics*, 2nd edn (eds Hillis DM,

Moritz C, Mable BK), pp. 407–514. Sinauer Associates, Sunderland, MA.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Wakeley J (2001) The coalescent in an island model of population subdivision with variation among demes. *Theoretical Population Biology*, **59**, 133–144.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity*, **82**, 117–125.

Wright S (1969) *Evolution and the Genetics of Populations*, Vol. 2. *The Theory of Gene Frequencies*. University of Chicago Press, Chicago, IL.

Young JR, Braun CE, Oyler-McCance SJ, Hupp JW, Quinn TW (2000) A new species of sage-grouse (Phasianidae: *Centrocercus*) from southwestern Colorado. *Wilson Bulletin*, **112**, 445–453.

Zink RM (1996) Comparative phylogeography in North American birds. *Evolution*, **50**, 308–317.

Zink RM (1997) Phylogeographic studies of North American birds. In: *Avian Molecular Evolution and Systematics* (ed. Mindell DP), pp. 301–324. Academic Press, San Diego, CA.

Zink RM, Barrowclough GF, Atwood JL, Blackwell-Rago RC (2000) Genetics, taxonomy, and conservation of the threatened California gnatcatcher. *Conservation Biology*, **14**, 1394–1405.

Zink RM, McKitrick MC (1995) The debate over species concepts and its implications for ornithology. *Auk*, **112**, 701–719.

Zwickel FC (1992) Blue Grouse. *Birds of North America*, **15**, 1–28.

Zwickel FC, Degner MA, McKinnon DT, Boag DA (1991) Sexual and subspecific variation in the numbers of rectrices of blue grouse. *Canadian Journal of Zoology*, **69**, 134–140.

Appendix I

Geographical distribution of blue grouse haplotypes[1] (columns) among population samples (rows). For locality codes see Materials and methods and Fig. 1

Locality	N	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17		
NM	17	5	2	1	1	3	2	1	2										
CO	17							17											
UT	20									2	11	1	1	1	1	2	1		
BLU	7																		
WAL	13																		
SEL	6																		
OKA	20																		
BC	2																		
OLY	4																		
PUG	11																		
CAS	6																		
WAR	15																		
HUM	8																		
MEN	1																		

Locality	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
NM																			
CO																			
UT																			
BLU	3	2	1	1															
WAL	2				1	4	1	2	1	1	1								
SEL	3											2	1						
OKA															1	4	4	1	1
BC												1		1					
OLY															1				
PUG															5				
CAS																			
WAR																			
HUM																			
MEN																			

Locality	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
NM																			
CO																			
UT																			
BLU																			
WAL																			
SEL																			
OKA	1	3	2	1	1	1													
BC																			
OLY							2		1										
PUG							4	1	1										
CAS											1	1	1	2	1				
WAR							4					3				1	3	3	1
HUM	1																		
MEN																			

Locality	56	57	58	59	60	61
NM						
CO						
UT						
BLU						
WAL						
SEL						
OKA						
BC						
OLY						
PUG						
CAS						
WAR						
HUM	3	1	1	1	1	
MEN						1

[1]Haplotype code corresponds to last two digits of GenBank Accession no.; prefix for all haplotypes is AY5703–.