Molecular Ecology (2006) **15**, 2317–2332 doi: 10.1111/j.1365-294X.2006.02935.x

Recency, range expansion, and unsorted lineages: implications for interpreting neutral genetic variation in the sharp-tailed grouse (*Tympanuchus phasianellus*)

A. W. SPAULDING,* K. E. MOCK,† M. A. SCHROEDER§ and K. I. WARHEIT‡
*Department of Biology, Utah State University, 5305 Old Main Hill, Logan, Utah 84322-5305, USA, †Department of Forest, Range, and Wildlife Sciences, Utah State University, 5230 Old Main Hill, Logan, Utah 84322-5230, USA, §Wildlife Research Division, Washington Department of Fish and Wildlife, PO Box 1077, Bridgeport, Washington 98813, USA, ‡Genetics Laboratory, Washington Department of Fish and Wildlife, 600 Capitol Way N., Olympia, Washington 98501-1091, USA

Abstract

Both current and historical patterns of variation are relevant to understanding and managing ecological diversity. Recently derived species present a challenge to the reconstruction of historical patterns because neutral molecular data for these taxa are more likely to exhibit effects of recent and ongoing demographic processes. We studied geographical patterns of neutral molecular variation in a species thought to be of relatively recent origin, *Tympanuchus phasianellus* (sharp-tailed grouse), using mitochondrial control region sequences (CR-I), amplified fragment length polymorphisms (AFLP), and microsatellites. For historical context, we also analysed CR-I in all species of *Tympanuchus*. Within *T. phasianellus*, we found evidence for restricted gene flow between eastern and western portions of the species range, generally corresponding with the range boundary of *T. p. columbianus* and *T. p. jamesi*. The mismatch distribution and molecular clock estimates from the CR-I data suggested that all *Tympanuchus* underwent a range expansion prior to sorting of mitotypes among the species, and that sorting may have been delayed as a result of mutation-drift disequilibrium. This study illustrates the challenge of using genetic data to detect historical divergence in groups that are of relatively recent origin, or that have a history dominated by nonequilibrium conditions. We suggest that in such cases, morphological, ecological, and behavioural data may be particularly important adjuncts to molecular data for the recognition of historically or adaptively divergent groups.

Keywords: phylogeography, range expansion, sharp-tailed grouse, *Tympanuchus*

Received 1 November 2006; revision accepted 6 February 2006

Introduction

Recently derived species present a challenge for interpreting neutral molecular variation. When the sorting of neutral alleles among lineages is driven by random genetic drift, the relative congruence between organismal and gene phylogenies is a function of time (Tajima 1983). Consequently, even when a restrictive gene-flow barrier separates groups with distinct but temporally shallow evolutionary trajectories, there may be a lack of diagnostic lineage sorting from sequence data (Avise *et al*. 1990; Avise 2000). Phylogenetic reconstruction in recently derived species may also be complicated by demographic changes such as range expansions, which can preserve molecular lineages and slow the lineage sorting process (Tajima 1983; Rogers & Harpending 1992).

In the case of recent origin or range expansion, emergent phylogenetic structure may be manifest and detected as allele frequency differences rather than diagnostic lineages (Moritz 1994; Wiens & Servedio 1998). Since the number and variability of loci have a strong effect on the accuracy of these methods (Nei *et al*. 1983), nuclear multilocus data such as microsatellites and amplified fragment length polymorphisms (AFLP) are particularly suited to this type of analysis. Allele frequency differences are not uniformly

Correspondence: Allen W. Spaulding, Fax: 435-797-1575; E-mail: aspauldo@biology.usu.edu

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

correlated with adaptive divergence in quantitative traits (Lynch *et al*. 1999; Morgan *et al*. 2001; McKay & Latta 2002), and they can be influenced by ongoing demographic processes (Templeton *et al*. 1995); therefore, historical context is also important to determine the relevance of allele frequency differences to conservation (Moritz 1994; Crandall *et al*. 2000; Moritz 2002) and studies of ecological variation (Felsenstein 1985; Gittleman & Kot 1990; Harvey & Pagel 1991). Separating population history from current demographic processes can be facilitated by information on the relationships and geographical locations of alleles (e.g. Templeton *et al*. 1995). Nucleotide sequence data are useful in this context; however, at shallow temporal scales, conclusions based on phylogenetic criteria (e.g. reciprocal monophyly (Moritz 1994) and degree of geographical overlap among clades [Templeton 2004]) may depend on the progress of random genetic drift through time. Many analyses that use both mitochondrial sequence data and multilocus data yield valuable insights that would be overlooked using only one of these types of data (Ellsworth *et al*. 1994; Oyler-McCance *et al*. 1999; Piertney *et al*. 2000; Johnson *et al*. 2003; Monsen & Blouin 2003; Bouzat & Johnson 2004; Mock *et al*. 2004; Bensch *et al*. 2006). This approach may be of particular importance when studying genetic variation in taxa that may still show the effects of ecological changes during the Quaternary period (Hewitt 2004).

Tympanuchus phasianellus (sharp-tailed grouse) is one of three extant species of *Tympanuchus* (A.O.U. 1998), collectively known as the North American prairie grouse. Species in this group, which also includes *Tympanuchus cupido* (greater prairie-chicken) and *Tympanuchus pallidicinctus* (lesser prairie-chicken), breed in some form of grassland where males compete for females on socially structured display grounds, or leks (Johnsgard 1983). The three species have strikingly different morphologies, courtship behaviours, and habitat affiliations, and occupy distinct but overlapping geographical ranges (Aldrich & Duvall 1955; Aldrich 1963; Johnsgard & Wood 1968; Johnsgard 1983, 2002; Sparling 1983). Past studies have estimated that speciation in *Tympanuchus* began during the Late Pleistocene (Ellsworth *et al*. 1994), or during the Middle–Late Pleistocene (Drovetski 2003). It has also been suggested that the range of all *Tympanuchus* may have expanded during the recent Holocene expansion of grasslands (Johnsgard 2002). Thus, species of *Tympanuchus* may be of recent origin, and may have undergone a rapid range expansion historically, although current populations have been dramatically reduced and fragmented.

In a previous study of molecular genetic variation among the prairie grouse, using allozymes and restriction fragment data from mitochondrial DNA (mtDNA), Ellsworth *et al*. (1994) found a lack of geographical and taxonomic differentiation among species of *Tympanuchus*. This result, likely due to incomplete lineage sorting among mitochondrial alleles and insufficient mutation rates in the allozyme markers, suggests that the striking morphological and behavioural differences among these species have arisen relatively rapidly via sexual selection (Ellsworth *et al*. 1994). These conclusions were also supported by Drovetski (2002) using nucleotide sequence data from W-chromosome, autosomal intron, and the mitochondrial control region. These studies suggest that prairie grouse may be a group of adaptively divergent species that exhibit little or no divergence among neutral molecular characters.

Of the three species of prairie grouse, *T. phasianellus* has the largest and most northern distribution (Aldrich 1963; Johnsgard 1983), extending from the Rocky Mountains and Great Plains regions, into the Northwest Territories and Alaska. Seven subspecies of *T. phasianellus* have been described (Connelly *et al*. 1998), based primarily on geographical range and subtle morphological differences: *T. p. columbianus*, *T. p. jamesi*, *T. p. campestris*, *T. p. phasianellus*, *T. p. kennicotti*, *T. p. caurus*, and *T. p. hueyi* (Fig. 1). Snyder (1935) proposed that some morphological variation in the range of *T. phasianellus* may be due to genetic differences between eastern and western parts of the range. Although he did not clearly define the differentiated regions, his mapping of character differences suggested a division at the southern Rocky Mountains, consistent with part of the boundary separating *T. p. columbianus* and *T. p. jamesi* (Fig. 1; Lincoln 1917).

Several populations of *T. phasianellus* have been the focus of conservation concerns, as have populations of other prairie grouse species (Johnsgard 2002). The historical range of *T. phasianellus* has been dramatically reduced since European settlement, particularly at the southern margin (Fig. 1; Aldrich & Duvall 1955; Johnsgard 1983). Of the seven described subspecies, the southernmost, *T. p. hueyi*, is now considered extinct (Dickerman & Hubbard 1994). The southwesternmost extant subspecies, *T. p. columbianus*, has experienced a range reduction of up to 90% (Johnsgard 2002), but was declined federally protected status in 2000 under the Endangered Species Act of 1973, due to recent increases in the sizes of local populations (USFWS 2000).

In this study, we use two multilocus nuclear markers and mtDNA sequence data to investigate historical changes in range and patterns of gene flow that may have shaped regional genetic variation in *T. phasianellus*. Such a study is essential to develop effective management strategies that maintain landscape-scale patterns of genetic variation in this species. To determine whether temporal (i.e. recency of common ancestry) or spatial (e.g. range expansion) effects influenced genetic variation in *T. phasianellus*, we combined our data with previously collected data for *T. cupido* and *T. pallidicinctus*. The objectives of our study were (i) to determine if, and when, *Tympanuchus* underwent a historical range expansion, (ii) to estimate the date of diversification within *Tympanuchus*, and (iii) to assess genetic

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

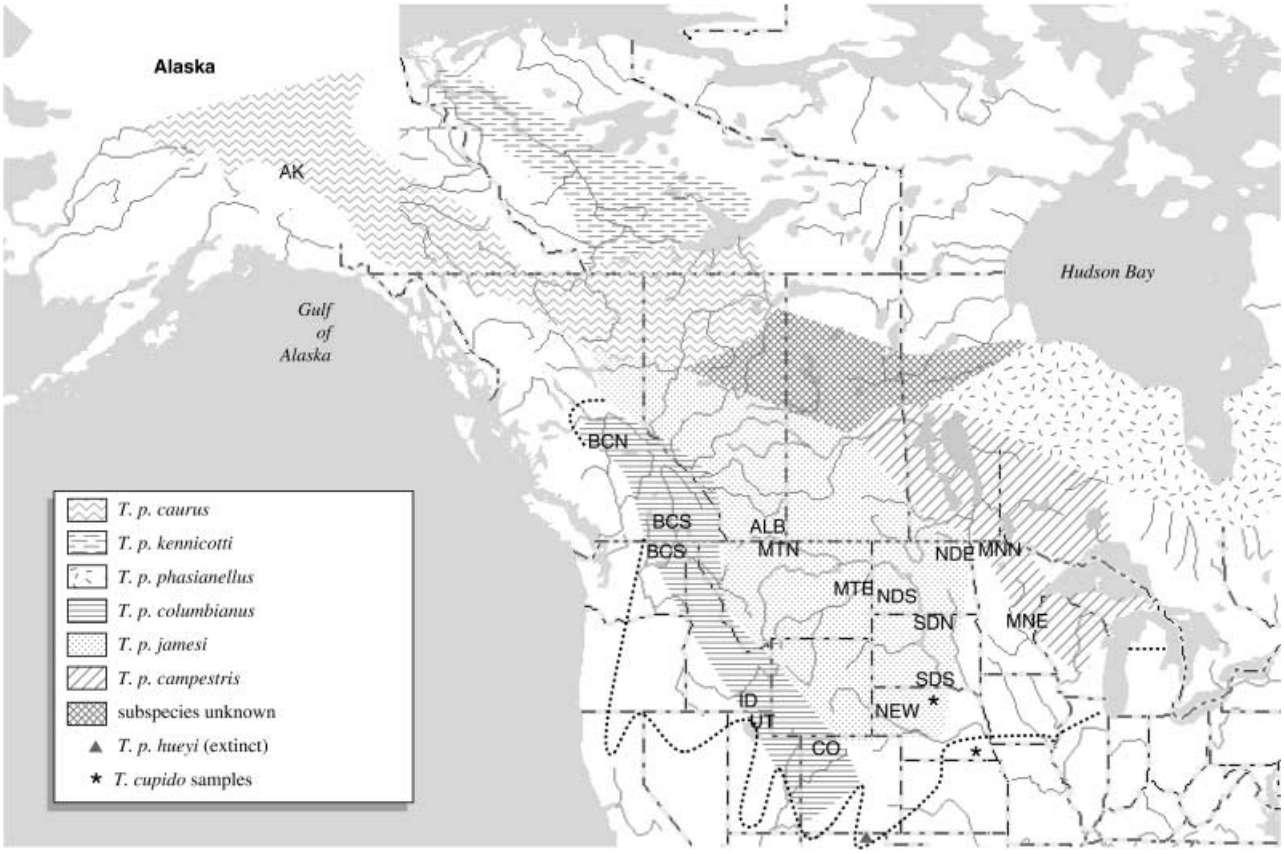

Fig. 1 Current and historical range of *Tympanuchus phasianellus* (Aldrich & Duvall 1955; Johnsgard 1983), and sampling localities. Current ranges of each extant subspecies are shown by distinctly patterned areas. The triangle at the southern tip of the historical range indicates the last known location of extinct *T. p. hueyi*. *T. phasianellus* occupying the area west of the Hudson Bay have not been assigned to any subspecies (see Snyder 1935). The dotted line shows the range of *T. phasianellus* prior to habitat losses due to European settlement and agriculture (Johnsgard 1983; Connelly *et al*. 1998). The asterisks show localities of *Tympanuchus cupido* sampled for this study.

structuring of *T. phasianellus* with reference to currently recognized subspecies boundaries in the southern and western portions of the species' range (Fig. 1). In light of the low resolution from a previous molecular study of *Tympanuchus* (Ellsworth *et al*. 1994), we chose to use highly polymorphic nuclear markers and sequence data from a rapidly evolving region of mtDNA.

Materials and Methods

Geographical sampling and DNA isolation

We acquired samples from 499 individuals, from a total of 17 localities, collected during 1998 to 2002 in the ranges of four putative subspecies of *Tympanuchus phasianellus* (Table 1 and Fig. 1). For comparison, we also sampled 43 individuals of *T. cupido* (greater prairie-chicken) from two localities in Nebraska (Table 1 and Fig. 1), in and near the range of *T. phasianellus*. Although most birds were not sampled during the breeding season, seasonal movements of *T. phasianellus* are not so extensive that we would expect them to significantly alter large-scale patterns of phylogeographical structure (Hamerstrom & Hamerstrom 1951). Each sex was well represented in the samples from each locality. To the best of our knowledge, none of the localities we sampled have been the recipients of translocated birds.

Most samples were in the form of tissues obtained with the cooperation of hunters during 1998–2001. Tissues were either shipped intact (wings or carcasses) from the field while frozen or chilled, and a sample excised in the laboratory; or excised tissues were shipped from the field at ambient temperatures in 95% EtOH or lysis buffer (Mullenbach *et al*. 1989). We extracted DNA from excised tissues using a standard salt/chloroform method (Maniatis *et al*. 1982), or a commercially available kit (QIAGEN DNeasy Tissue Kit). In southern British Columbia (BCS), Washington (WA), and Oneida County, Idaho (ID), blood samples had been collected in capillary tubes from birds trapped on leks. These samples were preserved by freezing, and we extracted DNA using a QIAGEN DNeasy kit.

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

Table 1 Sample sizes (*N*, number of individuals) and mitotypes (GenBank Accession nos DQ366600–DQ366638), for each locality: northern British Columbia (BCN), southern British Columbia (BCS), northwestern Colorado (CO), southeastern Idaho (ID), northern Utah (UT), northern Washington (WA), southern Alberta (ALB), eastern Montana (MTE), northern Montana (MTN), northeastern North Dakota (NDE), southwestern North Dakota (NDS), northern Nebraska and southern South Dakota (SDS), western Nebraska (NEW), northern South Dakota (SDN), southern Alaska (AK), northwestern Minnesota (MNN), eastern Minnesota (MNE), and southeastern Nebraska (NES). CR-I, AFLP, and microsatellite data were obtained from the same set of birds in each locality

Taxon	Locality name	Locality information	AFLP *N*	Microsatellite *N*	CR-I *N*	Mitotypes
T. p. columbianus	BCN	Leeches Lake, Cariboo; Meadow, Valenzuela and Fletcher Lakes, Lillooet, British Columbia	19	19	5	G, H, I
	BCS	near Merritt and Kamloops, Kamloops Division Yale, British Columbia	22	23	5	F, H
	CO	20-mile, Routt County, Colorado	17	17	5	A, E, F
	ID	Franklin and Oneida Counties, Idaho	16	16	2	A, F
	UT	Box Elder County, Utah	55	56	5	A, C, D, H
	WA	Douglas, Lincoln, and Okanagon Counties Washington	20	27	5	A, B, D, H
T. p. jamesi	ALB	near Milk River, Alberta	18	18	2	D, H
	MTE	Dawson and Prairie Counties, Montana	20	28	4	T, U, V, Y
	MTN	Blaine County, Montana	17	17	3	H, W
	NDE	Benson, Cavalier, Griggs, Nelson, Pembina, Rollete, and Walsh Counties, North Dakota	34	34	5	L, P, Q, R, X
	NDS	Billings, Bowman, Golden Valley, and Slope Counties, North Dakota	30	30	6	A, H, K, L, Y, Z
	SDS	Cherry and Rock Counties, Nebraska; Lyman County, South Dakota	24	22	5	H, Y, AA
	NEW	Dawes, Cheyenne, and Kimball Counties, Nebraska	11	11	5	M, N, O, S, V
	SDN	Corson and Dewey Counties, South Dakota	15	15	5	I, J, Y, AA
T. p. caurus	AK	Southeast Fairbanks County, Alaska	85	88	5	H, HH, II
T. p. campestris	MNN	Beltrami, Kittson, Marshall and Roseau Counties, Minnesota	24	23	6	X, BB, CC, DD, EE, FF
	MNE	Aitken and Pine Counties, Minnesota	11	12	5	Y, AA, GG
		Totals for ***T. phasianellus***	**438**	**456**	**78**	
	TcNES	Johnson and Pawnee Counties, Nebraska	17	17	4	V, Z, LL
T. cupido	TcSDS	Rock County, Nebraska; Lyman County, South Dakota	26	26	5	Y, JJ, KK, MM
		Totals for ***T. cupido***	**43**	**43**	**9**	

Mitochondrial data

We sequenced Domain I of the mitochondrial control region (CR-I; Baker & Marshall 1997) in 87 individuals representing four subspecies of *T. phasianellus*, and two localities of *Tympanuchus cupido*. This region had previously been shown to evolve rapidly and to be highly polymorphic in birds (Quinn 1992; Baker & Marshall 1997; Randi & Lucchini 1998). We amplified the ~500 bp fragment using primers 16775 L (Quinn 1992) and 521H (Quinn & Wilson 1993). We ran

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

amplification reactions (total volume of 50 mL, with 10 ng of isolated DNA, 1× reaction buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1 mM of each primer and 1 U *Taq* DNA polymerase) for 30 cycles: 40 s at 95 °C, 1 min at 55 °C, and 2 min at 72 °C, followed by a final 10-min extension period at 72 °C. We purified the polymerase chain reaction (PCR) products using Microcon-PCR spin columns (Millipore). Sequencing reactions for all templates were performed using the 16775 L primer and ABI BigDye Kit version 2.0. To check the accuracy of sequence data collection, we obtained complimentary strand sequences for 10% of the samples using the 521H primer. DNA sequence data were collected on an ABI 3100 Genetic Analyser. We submitted all mitochondrial sequences to GenBank, under Accession nos DQ366600–DQ366638.

We aligned the mitochondrial CR-I sequences using Megalign (Lasargene) software (DNAStar) using the CLUSTAL W algorithm and manual editing. The final alignment included 410 nucleotide sites. We constructed a mutational network of mitotypes with statistical parsimony, using the program TCS 1.15 (Clement *et al*. 2000), with no gaps, and excluding sites with missing data (337 sites remaining). Statistical support for network branches was assessed through Bayesian phylogenetic analysis, in the program MRBAYES 3.1 (Ronquist & Huelsenbeck 2003). We selected a model of nucleotide substitution for the Bayesian analysis using likelihood-ratio tests in a procedure described by Sullivan *et al*. (1997). We constructed a neighbour-joining (NJ) tree using distances under the Hasegawa–Kishino–Yano (HKY) model (Hasegawa *et al*. 1985; see model descriptions in Swofford *et al*. 1996) for calculation of all model likelihoods. Estimation of the model parameters and phylogenetic analysis were then carried out in MRBAYES 3.1 using default settings. We used the convergence diagnostic, described in the MRBAYES manual (Ronquist *et al*. 2005), to determine the appropriate length of the Markov chain Monte Carlo (MCMC) runs.

To assess regional mitotype differentiation in *T. phasianellus*, we performed an analysis of molecular variance (AMOVA) using the program ARLEQUIN 2.0 (Schneider *et al*. 2000), excluding three polymorphic nucleotide sites that had more than 5% missing data, and using 1000 random permutations to assess significance. We based a priori groupings for the AMOVA on Bayesian cluster analysis (see Results).

Since prior range expansions can have a long-term effect on molecular diversity, and since diversification in *Tympanuchus* may have taken place before or after this expansion (Ellsworth *et al*. 1994; Drovetski 2003), we tested for a range expansion in the entire genus. We aligned our sequences with CR-I sequences of 398 individuals from all three species ($N = 297$ for *T. cupido*, $N = 89$ for *T. phasianellus*, and $N = 12$ for *T. pallidicinctus*; GenBank Accession nos AF532430–AF532436 [Drovetski 2002], AJ297171, AJ297172, and AJ297174-AJ297177 [Lucchini *et al*. 2001], AY030408 [Aldridge *et al*. 2001], AY273829–AY273868 [Johnson *et al*. 2003], AY293569–AY293575 [Bouzat & Johnson 2004], AY526753–AY526860 [Palkovacs *et al*. 2004], AY569304–AY569306 [Barrowclough *et al*. 2004], AY608323–AY608334 [Johnson *et al*. 2004]. We used the aligned sequences (363 nucleotide sites), and available information on their frequencies in a mismatch distribution analysis (Rogers & Harpending 1992; Rogers 1995). Frequency data were lacking only for *T. pallidicinctus*. None of the nucleotide sites in this alignment had more than 5% missing data. For this test we used 1000 coalescent simulations to assess the significance of deviations from a simulated distribution, assuming a model of rapid population expansion.

We estimated the dates of three events using these generic-level data: (i) coalescence of the two most divergent mitotypes observed, (ii) existence of the most recent, common ancestor of *T. cupido* and *T. phasianellus*, and (iii) the rapid range expansion of all the observed mitotypes. Our molecular clock calibration was based on Drovetski's (2003) control region calibration for the entire grouse subfamily (Tetraoninae). We modified Drovetski's control region calibration because CR-I alone evolves at a higher rate. Baker & Marshall (1997) had previously compared the percent divergence among domains of the control region across several taxa representing Galliformes, Anseriformes, Charadriiformes, and Passeriformes. From their data, we estimated that the rate for CR-I alone is 1.2 times greater than the rate for the entire control region (Table 3 in Baker & Marshall 1997). Thus, we modified Drovetski's calibration of 7.23 + 1.58% sequence divergence per million years to 8.68 + 1.90% per million years.

To estimate the coalescence time of the two most distant mitotypes, we used uncorrected p-distances (proportion of differing nucleotide sites) calculated in PAUP* 4.0b10 (Swofford 1998). Although Ellsworth *et al*. (1994) used a similar estimate as an upper-bound date of species diversification in *Tympanuchus*, the availability of more sequences from this group warrants a new estimate. The date estimate, in millions of years before present (BP), is the sequence divergence divided by the divergence rate and multiplied by 10^6. We calculated a range of dates about the estimate using Drovetski's error term from the clock calibration. In order to estimate a date of divergence from the most recent common ancestor of *T. cupido* and *T. phasianellus*, we also calculated the average pairwise distances (Nei & Li 1979) between these species, using the pairwise deletion method for missing data, in the program MEGA 2.1 (Kumar *et al*. 2001). We performed the mismatch distribution using the program ARLEQUIN 2.0 (Schneider *et al*. 2000). To estimate the date of the range expansion detected using mismatch distribution analysis (see Results), we used the formula $\tau = 2ut$ (Rogers & Harpending 1992), where τ is the time since the range expansion in units of $1/2u$ generations, t is the number of generations, and u is the mutation rate. We obtained the mutation rate from $u = 2\mu k$, where and 2m is

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

the divergence rate and k is the number of nucleotides. We estimated two expansion dates, assuming one and two generations per year, respectively. We also calculated a 95% confidence interval (CI) for τ with 1000 replicates of bootstrapping (see Schneider *et al.* 2000). We calculated a range of dates for the range expansion by incorporating both the CI and the calibrated clock's error term. For example, we calculated the lower-bound date using the lower-bound of the CI for τ, and the faster sequence divergence rate of 10.58 = (8.68 + 1.90).

Nuclear data

Microsatellite locus optimization and analyses were performed at the Washington State Department of Fish and Wildlife Genetics Laboratory. We screened 28 microsatellite loci, developed in the domestic chicken (*Gallus gallus*), capercaillie (*Tetrao urogallus*), red grouse or willow ptarmigan (*Lagopus lagopus*), and black grouse (*Tetrao tetrix*), for utility in sharp-tailed grouse. Thirteen of these loci amplified successfully and 10 of these were polymorphic in our sample set (Table 2). To facilitate the multiplexing or combining of loci during the polymerase chain reaction (PCR), we modified the PCR protocols from those listed in the original publications (Table 2). Additional details for the PCR protocols are available from KW upon request. All PCR amplifications were conducted using fluorescently labelled primers. DNA fragments were separated electrophoretically using an ABI 3100 Genetic Analyser with an ABI ROX 500 size standard. Fragments were visualized and scored using ABI GENESCAN 3.7 and ABI GENOTYPER 3.7 software. Using the 10 variable microsatellite loci, we obtained genotypes for 456 individuals of *T. phasianellus* and 43 individuals of *T. cupido* (Table 1).

We generated AFLP marker profiles for 481 individuals (Table 1) using methods modified from Vos *et al.* (1995). The procedure, briefly, was as follows: (i) whole genomic DNA extractions were digested with restriction enzymes *Eco*RI and *Mse*I; and (ii) the fragments were ligated to adapters (laboratory synthesized oligomers) with T4 DNA ligase. We performed a preselective amplification on a subset of these fragments using primers complimentary to the adapters plus the *Mse*I or *Eco*RI restriction sites with an additional selective 3′ adenine nucleotide. We then performed a selective amplification, using the same primers modified by the addition of two nucleotides (arbitrarily chosen) at the 3′ end. We followed this with three different selective amplifications using the following primer pair combinations: *Eco*RI-ACG/*Mse*I-AGA, *Eco*RI-ACG/*Mse*I-ACT, and *Eco*RI-ACG/*Mse*I-ATC. In all these combinations, we used *Eco*RI primers labelled at the 5′ end with the fluorescent dye 6-FAM. Selectively amplified fragments were separated electrophoretically using an ABI 3100 Genetic Analyser with an ABI ROX 400 size standard. Combining data from these amplifications, we identified 45 polymorphic loci ranging in size from 55 to 396 bp, for 438 individuals of *T. phasianellus* and 43 individuals of *T. cupido* (Table 1). We manually scored the AFLP profiles using the program GENOGRAPHER 1.6.0 (Benham 2001). We estimated our methodological error rate by comparing results among three to four replicate reactions, carried forward from the preselective amplification stage, for 7% of the samples.

Using the microsatellite data, we tested for Hardy–Weinberg equilibrium (HWE) in each locality, and for each microsatellite locus, using the exact method implemented in GENEPOP 3.4 (Raymond & Rousset 1995). We assessed statistical significance using a MCMC method with 10 000

Table 2 *Tympanuchus phasianellus* microsatellite loci used in this study. All loci were derived from published accounts or from GenBank. Protocols were modified from those originally published to facilitate the multiplexing of loci. The PCR annealing temperatures and $MgCl_2$ concentrations are provided for each multiplex (see Warheit & Schroeder 2003)

Locus	Species	Fluorescent label	Motif	Multiplex	Annealing temp (°C)	$MgCl_2$ (mM)
ADL146*,¶	*Gallus gallus*	HEX	$(TG)_{17}$	b	56	2.0
ADL230*	*Gallus gallus*	6FAM	$(TG)_{17}$	b	56	2.0
LLSD3†	*Lagopus lagopus*	NED	$(TG)_{n}$	b	56	1.5
LLSD4†	*Lagopus lagopus*	HEX	$(TG)_{n}$	a	61**	1.5
LLSD7†	*Lagopus lagopus*	6FAM	$(AC)_{n}$	a	61**	1.5
TTD6‡	*Tetrao tetrix*	NED	$(CA)_{17}$	a	61**	1.5
TUD1§	*Tetrao urogallus*	HEX	$(CA)_{14}$	d	59	2.5
TUD2§,¶	*Tetrao urogallus*	NED	$(CA)_{13}$	f	59	2.5
TUD8§	*Tetrao urogallus*	HEX	$(GT)_{15}$	e	59	2.5
TUT1§,¶	*Tetrao urogallus*	FAM	$(CTAT)_{12}$	g	60	2.5
TUT2§	*Tetrao urogallus*	FAM	$(GATA)_{12}$	d	59	2.5
TUT3§	*Tetrao urogallus*	HEX	$(TATC)_{11}$	g	60	2.5
TUT4§	*Tetrao urogallus*	NED	$(TATC)_{8}$	g	60	2.5

*Cheng (unpublished data); †Piertney & Dallas (1997); ‡Caizergues *et al.* (2001); § Segelbacher *et al.* (2000); ¶locus not variable or did not amplify regularly, and was removed from further analyses; **touchdown (65–61 °C) included in PCR profile.

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

burn-in steps preceding the 10 000 steps sampled from the MCMC run. For each locality, we used a Bonferroni-corrected alpha value of 0.005 to account for the 10 tests (one for each microsatellite locus).

We estimated genetic relationships among the localities with NJ dendrograms, using D_A distances (Nei *et al*. 1983) for both AFLP and microsatellite data sets. The NJ method was chosen because it performs well even when branch-lengths are highly unequal (Wiens & Servedio 1998). Because AFLP is a dominant marker system, we calculated the D_A distances for these data using allele frequencies estimated under the assumption of HWE (Lynch & Milligan 1994), using the program TFPGA 1.3 (Miller 1997). We assessed support for dendrogram branches using 1000 bootstrap replicates (over loci). We performed the D_A distance calculations and NJ analysis using the program DISPAN (Ota 1993).

We made an initial assessment of genetic structure among localities using the program STRUCTURE 2.1 (Pritchard *et al*. 2000), which employs a Bayesian clustering procedure that does not require a priori assignment of individuals to geographical areas. We used this method for each of the two nuclear data sets, AFLP and microsatellites. For the microsatellite data, we used the model with admixture (Pritchard *et al*. 2000) and correlated allele frequencies among the populations (Falush *et al*. 2003). Since the data format for STRUCTURE 2.1 is for codominant markers, data for each AFLP locus in the locus-by-individual matrix were accompanied by a column of missing data. We analysed the AFLP data using a model without admixture (Pritchard *et al*. 2000), and with correlated marker frequencies. We specified that the numbers of groups (K) range from 1 to 9. We used the MCMC procedure in Structure to assess the strength of affiliation of each individual for each cluster group. For each analysis, we determined the appropriate number of burn-in steps by visual inspection of time series plots of all parameters output by the software (Pritchard & Wen 2003). For each data set, we ran the MCMC for 20 000 steps after burn-in. To check for consistency, 10 replicate runs were performed for each K. We used the largest K that met the following two conditions. (i) Groups had a minimum level of strength in clustering and correspondence with geography. Specifically, a majority of the individuals in at least one locality were required to have at least a 75% affiliation with the group. (ii) The selected K was required to be below values at which the log-likelihood began to plateau (Pritchard & Wen 2003). Long MCMC runs of 10^6 steps (after burn-in) were then run to obtain more accurate estimates of –ln P(D) at the selected value of K (Pritchard & Wen 2003).

We defined geographical regions for the AMOVA using the results of the Bayesian cluster analysis. The AMOVA F_{ST} estimator, ϕ, utilizes a matrix of Euclidian distances between haplotypes (Excoffier *et al*. 1992). For the AFLP data, we used a matrix of Euclidian distances between unique marker profiles, calculated in the program AMOVA-PREP (Miller 1998). Marker profile frequencies for each locality were also calculated in AMOVA-PREP. Individuals with missing AFLP data were excluded from the calculation of this matrix, as required by the AMOVA-PREP program. One individual was excluded from ID, NEW, and MNE; two from MTN and NDS; three from MTE and SDS; four from UT, NDE, SDN; and five from WA; resulting in $N > 10$ for each locality. The AMOVA was implemented in the program ARLEQUIN 2.0 (Schneider *et al*. 2000) using 1000 random permutations to assess significance. The localities in northwestern Colorado (CO) and southern Alberta (ALB) were excluded from these analyses since it was not clear which region they should have been assigned to (see Results).

Results

Mitochondrial data

From the 87 CR-I sequences in *Tympanuchus phasianellus* and *Tympanuchus cupido* (Table 1), we recovered 39 unique mitotypes (Table 1). Fifty of the 410 nucleotide sites in the alignment (12%) were polymorphic. Consistent with Ellsworth *et al*. (1994) and Palkovacs *et al*. (2004), we found three shared mitotypes between *T. cupido* and *T. phasianellus* (V, Y, Z; Table 1 and Fig. 2). Although not found among our *T. cupido* samples, our survey of *Tympanuchus* sequences from GenBank indicated that the most common mitotype (H) in our data was also in *T. cupido* over a broad portion of its range, in Kansas, Minnesota, Missouri, Wisconsin, Nebraska, North Dakota, South Dakota, Texas, and Massachusetts (results not shown). Nineteen branches connected this mitotype to other mitotypes in our network, producing a shallow, starlike topology (Fig. 2) that is consistent with a range expansion (Avise 2000). Also, TCS 1.15 identified the common mitotype H as ancestral based on its outgroup probability of 0. 1961 (Castelloe & Templeton 1994). The broad geographical range and presence across species boundaries is consistent with the expectation for an older mitotype (Templeton *et al*. 1995).

The nucleotide substitution model selected by the likelihood ratio tests was the HKY + G model (see Swofford *et al*. 1996 for model descriptions). MCMC runs in Bayesian phylogenetic analysis were for 3 100 000 steps. Trees sampled from the initial 775 000 steps (25%), from each of two independent MCMC runs, were discarded as burn-in; the log likelihood values had already stabilized at that point in the analysis. From the results of this phylogenetic analysis (not shown), the probability of a branch in the network was estimated as the percentage of the 46 500 retained trees having the same branch. These branch support indices appear in Fig. 2.

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

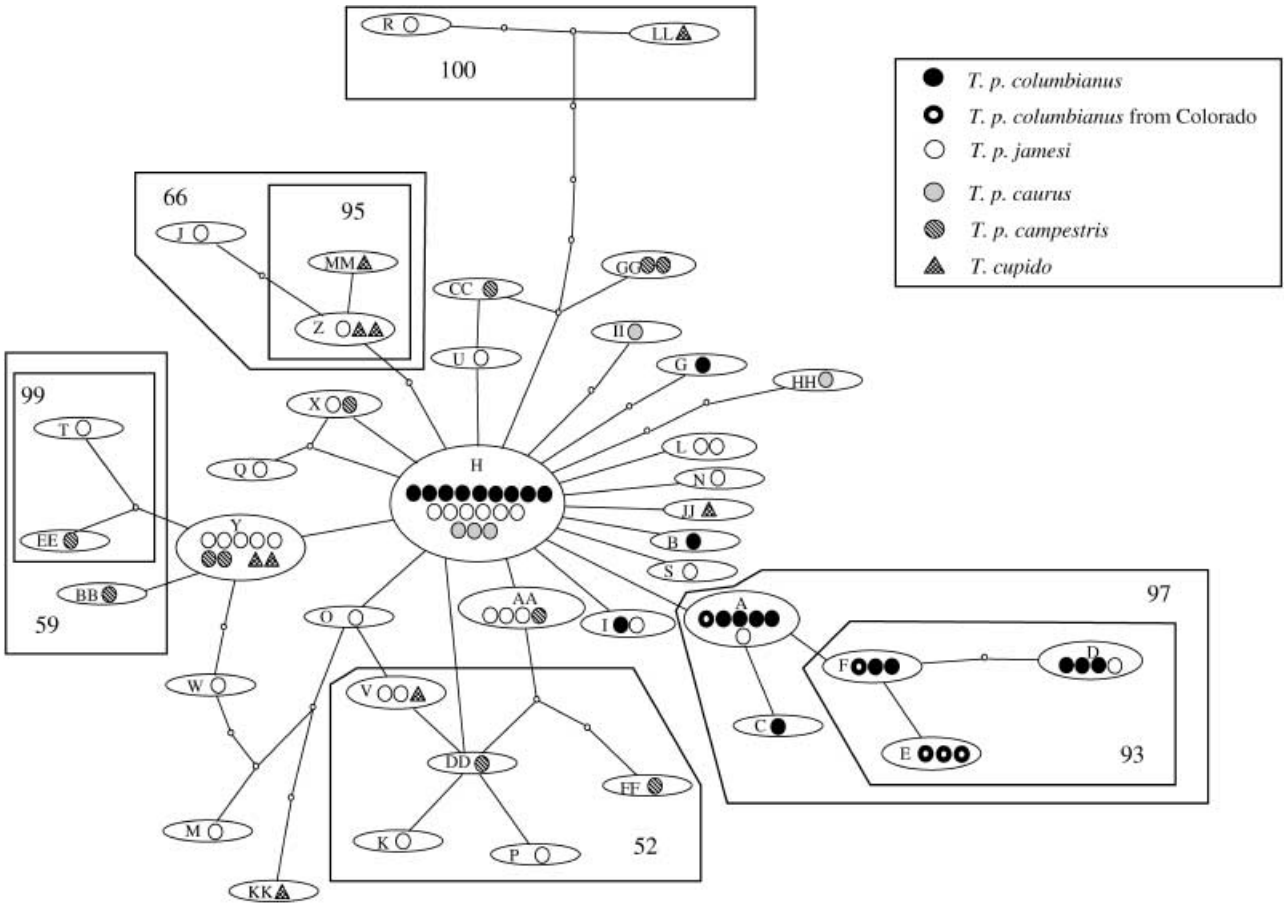

Fig. 2 Mitotype mutational network showing observed (ovals) and hypothesized (small circle) mitotypes linked by single mutational steps. Each circle (*Tympanuchus phasianellus*) or triangle (*Tympanuchus cupido*) within an oval represents one of the 87 CR-I sequences, and indicates the taxon from which it was sampled. Loops in the mutational pathways indicate unresolved relationships. Polygons encompass clades that had more than 50% posterior probability in a Bayesian phylogenetic analysis, and numbers in the polygons are the percent posterior probabilities.

Some concordance with subspecific taxonomy of *T. phasianellus* was found in that a statistically supported clade, which includes mitotype A and its descendants, was primarily associated with localities from the range of *Tympanuchus phasianellus columbianus* (Fig. 2). Two birds sampled east of the Continental Divide also had mitotypes from this clade; one (D) from near the Continental Divide in southern Alberta (ALB), and the other (A) farther from the Continental Divide in southwestern North Dakota (NDS).

In the survey of *Tympanuchus* CR-I data, the two most divergent mitotypes were found in *T. cupido* (AY526768) and *T. phasianellus* (AY030408). Based on this 6.39% sequence difference, we estimated the date of this divergence to be approximately 736 169 (604 143–942 035) years BP. Since the two divergent mitotypes are likely to have been present in the ancestral *Tympanuchus*, their coalescence may predate speciation in the group. Based on Nei's average pairwise distance, we estimated the date of the most recent common ancestor of *T. phasianellus* and *T. cupido* to be 20 975 (17 213–26 841) years BP. This type of estimate is subject to biases due to demographic changes in the range of the species (Wilson *et al*. 1985); mutation-drift disequilibrium during the range expansion may have caused the species to retain more ancestral mitotypes, thus biasing this estimate to a more recent date. Using these two estimates as conservative upper and lower bounds, we estimate the most recent common ancestor of *T. phasianellus* and *T. cupido* existed between 17 000 and 942 000 years BP.

The range expansion suggested by the starlike topology of the mitotype network was not rejected by the mismatch distribution test (SSD = 0.003, $P = 0.692$). From $\tau = 4.554$ (95% CI 2.011–14.702), the estimated date of this expansion was 64 012 (23 197–264 443) years BP under the assumption of 1 year per generation (for *T. phasianellus* (Connelly *et al*. 1998)), and 128 023 (46 395–528 887) years BP under the assumption of a 2 year generation time (for *T. cupido* (Bellinger *et al*. 2003)). Using upper and lower bounds from

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

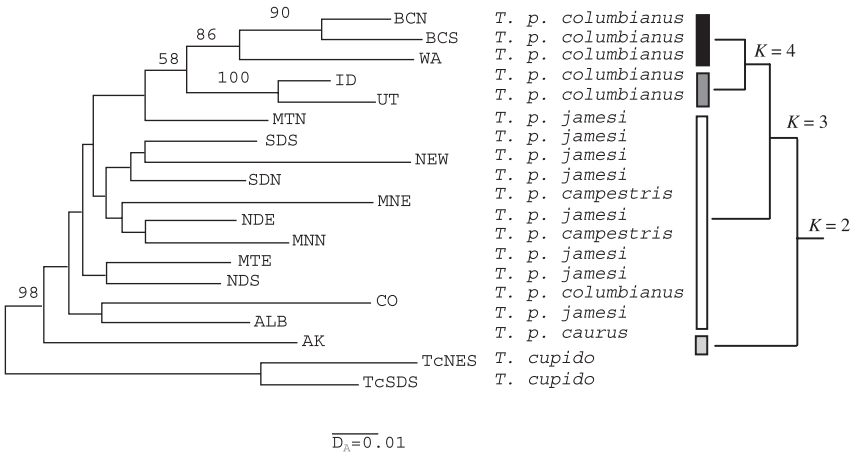

Fig. 3 Neighbour-joining dendrogram of *Tympanuchus phasianellus* localities based on microsatellite data and Nei's D_A distance (Nei *et al*. 1983). Numbers at nodes indicate the percentage of 1000 bootstrap replicates in which the node was recovered; only bootstrap values greater than 50% are shown. To the right of each locality name (described in Table 1) is the traditional subspecies range it occupies. Further to the right are the corresponding clusters from Bayesian analysis of these same data (see Fig. 5). The dichotomously branching pattern describes the order in which Bayesian clusters emerged as the specified number of groups (*K*) was increased. Both the continuity of the colour patterns and the similarity of the branching diagrams illustrate the congruence between the results of these two analyses.

these estimates, we date the range expansion to be between 23 000 and 529 000 years BP.

Nuclear data

The assumption of HWE, required for the estimation of allele frequencies from the AFLP data, did not appear to be violated. Ten of the 190 (5%) locality-by-microsatellite locus tests showed significant differences from HWE using the Bonferroni-corrected alpha levels of 0.005. These cases of disequilibrium were not concentrated at a single locus, or a single geographical locality. The average per-locus error rate for the AFLP data was 2%, a rate comparable to that of other AFLP analyses in birds (Busch *et al*. 2000; Mock *et al*. 2002).

The NJ dendrogram from the microsatellite data showed limited congruence with current taxonomic designations (Fig. 3). The branch differentiating *T. cupido* from *T. phasianellus* was well supported and is consistent with a monophyletic *T. phasianellus*, although this tree was unrooted. If we assume that the root of this tree resides along the branch differentiating *T. cupido* from *T. phasianellus*, we can make an assessment about the phylogenetic patterns within *T. phasianellus*. The NJ tree did not support the monophyly of *T. p. campestris*, *T. p. jamesi*, nor *T. p. columbianus* (Fig. 2). One branch grouped five of the six localities from the range of *T. p. columbianus*, although with only weak bootstrap support (58%). All three of the internal branches for this group were well supported, and were geographically congruent within that subspecies' range. The locality from northern Montana (MTN, *T. p. jamesi*) was grouped at the base of this *T. p. columbianus* group, but the branch was neither well supported nor geographically congruent. *T. p. jamesi* appeared highly polyphyletic as shared branches with localities of *T. p. campestris* and *T. p. columbianus* separated it into five groups. Internodes in this part of the tree were often short, and all had less than 50% bootstrap support. The single population of *T. p. caurus* in our data set was placed at the base of *T. phasianellus*, but this had less than 50% bootstrap support.

The NJ dendrogram from the AFLP data was consistent with the dendrogram from the microsatellite data in two notable respects. First, *T. cupido* was well distinguished from *T. phasianellus*, with strong bootstrap support. Second, the same five localities from the range of *T. p. columbianus* were joined by a single node. However, in this analysis, this grouping received less than 50% bootstrap support (Fig. 4), and the branching pattern exhibited less geographical congruence than the microsatellite dendrogram (compare Figs 3 and 4). The AFLP NJ dendrogram also showed moderate bootstrap support for the grouping of the Alaskan and eastern Montana localities.

Likelihood scores in the Bayesian cluster analysis of the microsatellite data stabilized after 20 000 steps of the MCMC runs. Results from the 10^6-step MCMC run with the highest −ln P(D), for $K = 4$, are presented in Fig. 5. Four geographical groups were resolved, namely: one in Alaska; another in the ranges of *T. p. campestris*, *T. p. jamesi*, and part of the range of *T. p. columbianus*; another in the southwestern part of the range of *T. p. columbianus*; and one in the Pacific Northwest portion of the range of *T. p. columbianus*. The pattern in which clusters emerged was consistent with the partitioning in the NJ dendrogram from the microsatellite data (Fig. 3).

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

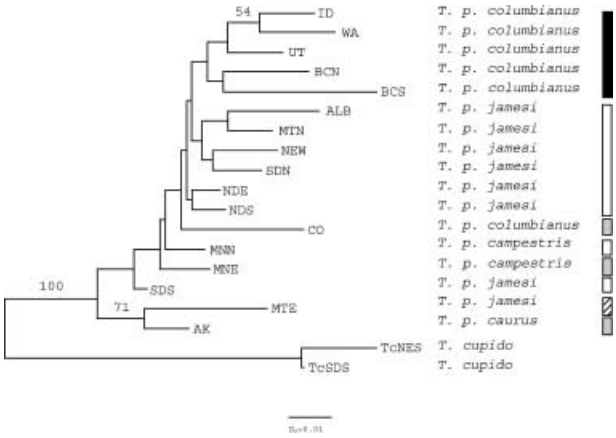

Fig. 4 Neighbour-joining dendrogram of *Tympanuchus phasianellus* localities based on AFLP data and Nei's *DA* distance. Numbers at nodes indicate the percentage of 1000 bootstrap replicates in which the node was recovered; only bootstrap values greater than 50% are shown. To the right of each locality name (described in Table 1) is the traditional subspecies range it occupies. Farther to the right are the corresponding groups from the Bayesian cluster analysis of these same data (see Fig. 5). The discontinuity of the colour patterns illustrates the incongruence between the results of these two analyses.

Bayesian cluster analysis with the AFLP data required much longer burn-ins than with the microsatellite data; likelihoods did not stabilize until after 110 000 MCMC steps. Nevertheless, results were consistent in replicate runs, for all *K*. We selected a *K* value of 4. While this analysis yielded the same number of groups as the Bayesian cluster analysis of the microsatellite data, the geographical distribution of the groups was different, and the strength of clustering was weaker and more variable (Fig. 5); this may be due to the relatively small number of AFLP loci compared to other avian studies (Wang *et al*. 2003; Mock *et al*. 2004). Clustering results obtained from the long MCMC runs of 10^6 steps were not consistent with the short runs. Interestingly, eight of the 10 replicate short MCMC runs had –ln P(D) higher than either of the long MCMC runs. The congruence of these short runs with the results for the Bayesian clustering analysis of the microsatellite data and both NJ dendrograms (Figs 3 and 4), suggests that the longer runs in STRUCTURE 2.1 were giving less accurate results for the AFLP data. The result with the highest –ln P(D) is presented in Fig. 5. The four geographical groups were the following: (i) a geographically incongruous group consisted of localities

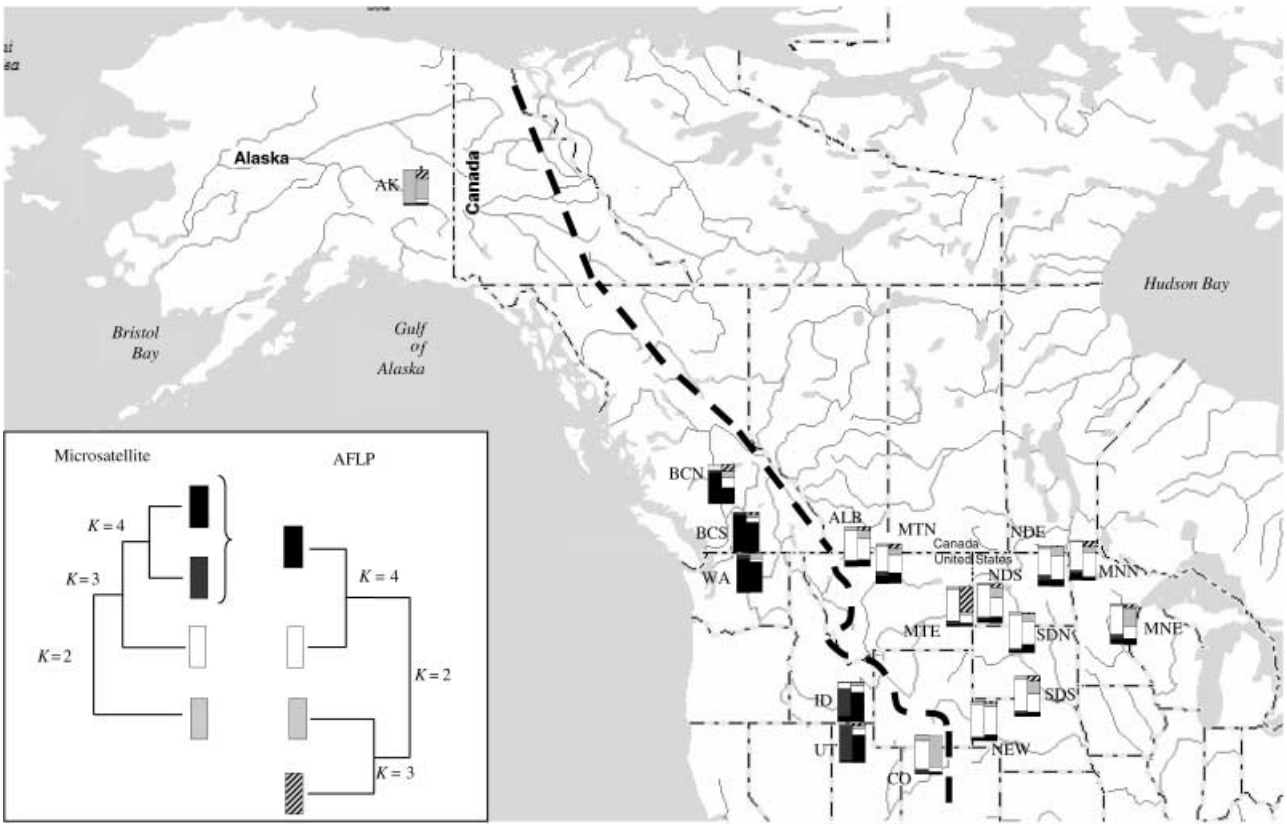

Fig. 5 Results of Bayesian cluster analysis of *Tympanuchus phasianellus* individuals. Bold, dashed line shows the Continental Divide of the Rocky Mountains. A pair of patterned bars is centred over each sample locality on the maps: microsatellite results are in the left bar, and AFLP results are in the right bar. The proportion of a bar with a specific colour or pattern depicts the average affiliation of that locality's individuals for particular clusters. The dichotomously branching diagrams (inset) describe the order in which these clusters emerged as the specified number of groups (*K*) was increased. Only results from the MCMC runs yielding the highest –ln P(D) are shown.

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

in Alaska, Colorado, and eastern Minnesota; (ii) a group including only eastern Montana; (iii) a group comprised of all localities from the range of *T. p. columbianus*, except the Colorado locality; and (iv) a group containing all remaining localities from the ranges of *T. p. jamesi* and *T. p. campestris*.

Regional variation

Both Bayesian cluster analysis and the mitotype network suggested a restriction of gene flow near the Continental Divide. Based on these results, we defined eastern and western regions for the AMOVA. The western region included five of the six localities sampled from the range of *T. p. columbianus* (BCN, BCS, WA, ID, and UT), and the eastern region included all sample localities from the range of *T. p. jamesi* and *T. p. campestris*. We were unable to assign two of the localities to a region for the AMOVA. Since CO clustered with localities east of the Continental Divide in the nuclear analyses, but was associated with a mitotype clade consisting of most of the mitotypes found west of the Continental Divide, its regional affiliation is unclear. Additionally, since one of the two individuals from ALB had a mitotype belonging to the western clade, and since ALB was grouped with CO in the microsatellite dendrogram, the regional affiliation of this locality is also unknown. Thus, we excluded both of these localities from the eastern/western AMOVA. Inclusion of these localities in the eastern region, particularly CO, increased the among-region differentiation. In a separate regional AMOVA, we tested for differentiation between the locality in Alaska (AK) and the rest of the sampled range, since the Bayesian cluster analysis also suggested a restriction of gene flow there.

In each regional AMOVA, the largest component of molecular variation was within the localities, while differentiation among localities and regions was low. Measures of geographical structure between the western and eastern regions were statistically significant for all three data sets, appearing weakest for microsatellites and strongest for CR-I (Table 3). Although AK was well distinguished in the NJ and Bayesian analyses, regional structure between Alaska and the rest of the range was marginally nonsignificant for microsatellites ($\Phi^{CT} = 0.05778$, $P = 0.06010$) and AFLP ($\Phi^{CT} = 0.06863$, $P = 0.05129$), and nonexistent for CR-I ($\Phi^{CT} = -0.03043$, $P = 0.47188$).

Discussion

Two primary patterns emerge from our study: (i) mitotype diversity in *Tympanuchus phasianellus* is dominated by the signature of a range expansion that affected all of *Tympanuchus*; and (ii) a restriction of gene flow, in the area of the Continental Divide, producing genetic structure approximately concordant with the currently recognized range boundary between *T. p. columbianus* and *T. p. jamesi*.

Lineage sorting, range expansion, and the age of Tympanuchus phasianellus

Ellsworth *et al*. (1994) and Palkovacs *et al*. (2004) have previously shown an extensive sharing of mitotypes among species of *Tympanuchus*. Ellsworth *et al*. (1994) hypothesized that this was due to a lack of lineage sorting rather than hybridization between species. In our study, both the strong nuclear differentiation between *Tympanuchus cupido* and *T. phasianellus* (Figs 3 and 4), and the wide geographical presence of the most common mitotype in ranges of both species support this hypothesis.

Ellsworth *et al*. (1994) also hypothesized that phenotypic differences evolved rapidly between *T. phasianellus* and other *Tympanuchus*. This conclusion was based on the obvious morphological and behavioural differences among the species along with the lack of mitotype sorting and the estimated upper-bound speciation date of 270 000 years BP for the two most divergent mitochondrial sequences in their study. Our data do not refute this hypothesis, but they do open the possibility that phenotypic divergence between these species was not as rapid as their estimate suggests. The estimated date for our largest mitotype divergence pushes this date back to 736 000 years BP, with an upper-bound of 942 000 years BP. However, the signal of rapid range expansion in the mismatch distribution for

Table 3 AMOVA results for eastern and western regions of *T. phasianellus*. The CO and ALB localities were excluded from these analyses (see Results)

	Among regions Φ_{CT} (*P*) (% variation)	Among localities/within regions Φ_{SC} (*P*) (% variation)	Within localities Φ_{ST} (*P*) (% variation)
CR-I	0.08843 ($P = 0.00079$) (8.84%)	0.03544 ($P = 0.09911$) (3.23%)	0.12073 ($P = 0.00198$) (87.93%)
Microsatellite	0.02727 ($P = 0.00030$) (2.73%)	0.03231 ($P < 0.00001$) (3.14%)	0.05870 ($P < 0.00001$) (94.13%)
AFLP	0.05368 ($P < 0.00001$) (5.37%)	0.08056 ($P < 0.00001$) (7.62%)	0.12992 ($P < 0.00001$) (87.01%)

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

Tympanuchus suggests that even if speciation had occurred prior to the range expansion, at the time of the expansion, mitotypes had not yet sorted and diverged appreciably among species. Thus, speciation probably did not take place much earlier than our upper bound estimate for the range expansion (529 000 years BP), and it may have been closer to our lower-bound estimate of 17 000 years BP. Based on average sequence divergences Zink (1997) and Avise & Walker (1998) have suggested that ages of avian phylogroups are commonly between 300 000 and 1 000 000 years, and as low as 75 000 years (Avise & Walker 1998). A relatively recent date in this range would seem to better explain the lack of lineage sorting in *Tympanuchus*. However, genetic drift has an important role in the sorting of mitotypes among lineages (Tajima 1983), and the mutation-drift disequilibrium from a range expansion can have a long-term effect on mitotype variation (Rogers & Harpending 1992). Thus, if *Tympanuchus* underwent a range expansion prior to sorting of mitotypes among species, sorting may have been delayed since the expansion, for 23 000–529 000 years. These issues illustrate our uncertainty about the precise age of *T. phasianellus*, and about the rate of phenotypic evolution in *Tympanuchus*.

Johnsgard (2002) hypothesized that much of the North American range of the *Tympanuchus* species was created by the rapid expansion of grasslands during the Holocene (Hoffmann & Jones 1970; Watts 1983; Webb *et al*. 1983). This is not supported by our estimation that a range expansion occurred no later than 23 000 years BP, although it is still possible that subsequent expansions occurred (Rogers & Harpending 1992). In addition, Late Pleistocene samples, identified as *T. phasianellus*, have been recovered over a broad North American range (Virginia, Pennsylvania, Tennessee, Texas, Nevada, and Oregon [Lundelius *et al*. 1983]). Given the variety of open habitats currently used by breeding *T. phasianellus* (Johnsgard 1983; Connelly *et al*. 1998), a Late-Pleistocene expansion is conceivable. Suitable habitats may have existed over a wide range of North America prior to the Holocene, including eastern prairies (Watts 1983), Great Basin steppe (Spaulding *et al*. 1983), and even in the wake of a retreating glacier (Hoffmann & Jones 1970; Heusser 1983).

Genetic Structure in Tympanuchus phasianellus

In all three data sets, there was evidence for restricted gene flow between the eastern and western regions in the range of *T. phasianellus*; the evidence was strongest in the Bayesian cluster analysis of microsatellite data (Fig. 5). That analysis suggested that admixture among individuals was low between these two regions, and this result was concordant with the averaged probabilities of regional membership calculated from the AFLP data (Fig. 5). In all three data sets, the largest component of variation in AMOVA was found within the populations (Table 3). This could be the result of lineage-preserving effects of a recent range expansion or extensive movements that were historically observed in this species (Hamerstrom & Hamerstrom 1951). Between-region structure was low but statistically significant (Table 3), and the grouping of the Western localities in both the NJ and Bayesian cluster analyses indicates that this result is not due to a pattern of isolation by distance across the species range.

Some authors have suggested that the Continental Divide has been an agent of diversification in *Tympanuchus* (Lincoln 1917; Johnsgard 2002), and another galliform species, *Meleagris gallopavo*, does show some phylogenetic concordance with this geographical feature (Mock *et al*. 2002). However, several studies of avian species indicate that the Continental Divide has not been a significant factor in producing avian phylogroups (Zink 1997), and that past glacial cycles may be a more important driver of divergence in grouse (Drovetski 2003). In any case, there is some indication in the mitotype network that the restriction of gene flow between the eastern and western regions of *T. phasianellus* may actually correspond to a recent history of fragmentation — too recent for mitochondrial lineages to have sorted. In the mitotype network, a statistically supported clade consisting of mitotype A and its descendents, C, D, E, and F, was found in localities in the range of *T. p. columbianus* (Fig. 2, Table 1). Three mitotypes of this clade (A, D, and F) are geographically dispersed in that range (Table 1) indicating their extensive movement there. This geographical association of a high-level clade, combined with the lack of geographical restriction among its subclades, is consistent with a range fragmentation (Templeton *et al*. 1995). Among the four haplotypes found to be shared between eastern and western populations, one was the common, widespread, and perhaps ancestral haplotype (H), and the remaining three mitotypes were each observed in a single location: SDN (I), NDS (A), and ALB (D). These data do not allow us to determine whether these patterns are the result of current gene flow vs. shared ancestral mitotypes. Future studies involving larger numbers of sequences and populations are likely to be informative in this regard.

NJ and Bayesian clustering analyses of the nuclear data (Figs 3, 4, and 5) indicated that the CO locality was more closely allied with *T. p. jamesi* than with *T. p. columbianus* (Fig. 1), a pattern contradicted by the mitotype network (Fig. 2). Assignment of populations to subspecies of *T. phasianellus* has long been a problem in Colorado (Lincoln 1917; Snyder 1939; Dickerman & Hubbard 1994). These conflicting patterns suggest that this problem may stem either from interregional gene flow in this area or a strong influence of genetic drift on this population (note the relatively long branches in the NJ dendrograms of Figs 3 and 4). Resolution of this issue may be possible

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

with more molecular data from additional localities, and perhaps larger sample sizes (if mtDNA is used), in the area between the eastern and western regions.

Our findings on patterns of genetic variation in *T. phasianellus* highlight the need to assess subspecies taxonomies with genetic studies of variation (Zink 2004). Our analysis found incomplete correspondence between genetic variation in *T. phasianellus* and some of the subspecies boundaries as described by Johnsgard (1983). Although the Western region we defined here is within the range of *T. p. columbianus*, a strict definition of this range, using the Continental Divide as an eastern boundary (Lincoln 1917), may not be upheld upon further genetic sampling. Distributions of other grouse species suggest that major river systems may be important barriers to gene flow in the southwest (Braun *et al*. 2003; Schroeder *et al*. 2004). There was no support for differentiation between *T. p. jamesi* and the two *T. p. campestris* localities from Minnesota, although the power of this inference was limited by low sample sizes for *T. p. campestris*. The locality in Alaska, in the range of *T. p. caurus*, was well differentiated from the other localities we sampled, although this was marginally nonsignificant, and the distinction between restricted gene flow vs. a pattern of isolation by distance was not possible due to a large geographical gap in our sampling.

Implications for management strategies and future research directions

The challenge of interpreting neutral molecular variation in recently derived species extends to species whose populations have a history of nonequilibrium conditions. For example, Palkovacs *et al*. (2004) recognized that the monophyly observed for the mtDNA of the extinct heath hen (*Tympanuchus cupido cupido*) may have been due to the extreme population bottleneck experienced by all individuals of that taxon. Loss of mutation-drift equilibrium during a sudden range expansion may have had the opposite effect for other prairie grouse taxa: a reduced effect of genetic drift might have delayed the sorting of allelic lineages among species. Whether the apparent weakness of historical signal in neutral molecular data for *Tympanuchus* is a result of rapid evolution or nonequilibrium conditions during a sudden range expansion, this group exemplifies the fact that tests for detecting historical lineages using phylogenetic criteria (e.g. reciprocal monophyly or degree of geographical overlap among clades) may be prone to type II error under some conditions. When there is reason to believe that such conditions may exist in a phylogeographical study, interpretation of the results should recognize the potential effects of population genetic processes on expected phylogeographical patterns.

There is general agreement in the conservation community that the designation of 'units' for conservation, monitoring, demographic segregation and management action should incorporate information from genetic markers as well as from studies on adaptive divergence (e.g. Crandall *et al*. 2000; Moritz 2002). Based on differences in beak shape and body size, Snyder (1935) hypothesized that there are genetic differences between populations occupying the ranges *T. p. columbianus* and *T. p. jamesi*. Our results suggest that if there has been phenotypic divergence between these ranges, it is accompanied by geographical separation. If we accept that there are some differences in quantitative genetic traits between these taxa, our results would be consistent with these two scenarios: (i) phenotypic divergence (neutral or adaptive) and neutral molecular divergence facilitated by geographical separation, and (ii) adaptive divergence without geographical separation, followed by neutral molecular divergence due to later separation (perhaps from anthropomorphic losses of suitable habitat). Distinguishing between these scenarios with inferences about history is important to the formulation of conservation management recommendations from genetic data (Crandall *et al*. 2000).

Here, we provide evidence that populations of sharp-tailed grouse in the *T. p. columbianus* range are genetically distinct from other subspecies of *T. phasianellus*, and that this distinction may be the result of processes predating modern history. We recommend that *T. p. columbianus* be managed as a distinct entity. Since the distribution and basis of adaptive and/or morphological variation among subspecies of *T. phasianellus* has not been well studied, studies of adaptive variation (e.g. morphology, behaviour, and habitat use) should be undertaken to further characterize the subspecies. Studies currently underway should clarify the range boundaries and history of *T. p. columbianus* in relation to adjoining subspecies ranges, and characterize the degree and potential impact of inbreeding in remnant populations. Until the uncertain history of the Colorado locality is investigated with additional sampling, this part of the range should not currently be considered as a source for long distance translocation programs. Our sampling range did not allow us to thoroughly assess subspecific boundaries other than between *T. p. columbianus* and *T. p. jamesi*, but additional historical boundaries may exist. As this information accumulates, a comprehensive prehistory of this species and species complex may emerge, allowing inference about the rate and pattern of adaptive diversity in these taxa.

Acknowledgements

We would like to thank the hunters who have helped in the collection of samples, and the wildlife managers who coordinated these efforts to make this study possible. Idaho and British Columbia provided permission for Washington Department of Fish and Wildlife (WDFW) field collectors to capture and handle grouse. Cherrill Bowman, Elizabeth Donelan, Alice Pichahchy, and

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

Jennifer Von Bargen, all from WDFW, provided expert laboratory assistance in Olympia. Steve Larson of USU provided technical guidance with AFLP scoring. We are grateful to Dave Hays, Janet Loxterman, John Pierce, Jim Shaklee, Mo Small, Sewall Young, Don McKinnon, Clait Braun and George Barrowclough for various discussions. We thank Mark Miller, Michael Pfrender, and the USU Herp Group for helpful comments on the manuscript. The microsatellite laboratory work, including scoring and binning of alleles was conducted by WDFW with generous funding from the States of Colorado, Utah, and Idaho, the Colville Indian Tribe, Bureau of Land Management, the National Fish and Wildlife Foundation and WDFW. Utah State University lab work was funded by the USU Community/University Research Initiative. We would also like to thank the anonymous referees for comments that were helpful in improving this manuscript.

References

A. O.U. (1998) *Check-list of North American Birds*, 7 edn. American Ornithologists' Union, Washington, D.C.

Aldrich JW (1963) Geographic orientation of American Tetraonidae. *Journal of Wildlife Management*, **27**, 529–545.

Aldrich JW, Duvall AJ (1955) Distribution of American gallinaceous game birds. *Fish and Wildlife Service Circular 34*.

Aldridge CL, Oyler-McCance SJ, Brigham RM (2001) Occurrence of greater sage-grouse × sharp-tailed grouse hybrids in Alberta. *Condor*, **103**, 657–660.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Avise JC, Ankney CD, Nelson WS (1990) Mitochondrial gene trees and the evolutionary relationships of mallard and black ducks. *Evolution*, **44**, 1109–1119.

Avise JC, Walker D (1998) Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **265**, 457–463.

Baker AJ, Marshall HD (1997) Mitochondrial control-region sequences as tools for understanding the evolution of avian taxa. In: *Avian Molecular Systematics and Evolution* (ed. Mindell DP), pp. 49–80. Academic Press, New York.

Barrowclough GF, Groth JG, Mertz LA, Gutierrez RJ (2004) Phylogeographic structure, gene flow, and species status in blue grouse (*Dendragapus obscurus*). *Molecular Ecology*, **13**, 1911–1922.

Bellinger MR, Johnson JA, Toepfer J, Dunn P (2003) Loss of genetic variation in greater prairie chickens following a population bottleneck in Wisconsin, U.S.A. *Conservation Biology*, **17**, 717–724.

Benham JJ (2001) *GENOGRAPHER*. Montana State University, Montana.

Bensch S, Irwin DE, Irwin JH, Kvist L, Akesson S (2006) Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Molecular Ecology*, **15**, 161–171.

Bouzat JL, Johnson K (2004) Genetic structure among closely spaced leks in a peripheral population of lesser prairie-chickens. *Molecular Ecology*, **13**, 499–505.

Braun CE, Delpiccolo RJ, Hoffman RW, Remington TE (2003) First specimen records for ruffed grouse in Colorado. *Western North American Naturalist*, **63**, 105–107.

Busch JD, Miller MP, Paxton EH, Sogge MK, Keim P (2000) Genetic variation in the endangered southwestern willow flycatcher. *The Auk*, **117**, 586–595.

Caizergues A, Dubois S, Loiseau A, Mondor G, Rasplus JV (2001) Isolation and characterization of microsatellite loci in black grouse (*Tetrao tetrix*). *Molecular Ecology Notes*, **1**, 36–38.

Castelloe J, Templeton AR (1994) Root probabilities for intraspecific gene trees under neutral coalescent theory. *Molecular Phylogenetics and Evolution*, **3**, 102–113.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Connelly JW, Gratson MW, Reese KP (1998) Sharp-tailed grouse: *Tympanuchus phasianellus*. In: *Birds of North America* (eds Gratson MWR, Kerry P). American Ornithologists' Union, Washington, D.C.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, **15**, 290–295.

Dickerman RW, Hubbard JP (1994) An extinct subspecies of sharp-tailed Grouse from New Mexico. *Western Birds*, **25**, 128–136.

Drovetski SV (2002) Molecular phylogeny of grouse: individual and combined performance of W-linked, autosomal, and mitochondrial loci. *Systematic Biology*, **51**, 930–945.

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

Ellsworth DL, Honeycutt RL, Silvy NJ, Rittenhouse KD, Smith MH (1994) Mitochondrial-DNA and nuclear-gene differentiation in North American prairie grouse (genus *Tympanuchus)*. *The Auk*, **111**, 661–671.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Felsenstein J (1985) Phylogenies and the comparative method. *American Naturalist*, **125**, 1–15.

Gittleman JL, Kot M (1990) Adaptation: statistics and a null model for estimating phylogenetic effects. *Systematic Zoology*, **39**, 227–241.

Hamerstrom FN, Hamerstrom F (1951) Mobility of the sharp-tailed grouse in relation to its ecology and distribution. *American Midland Naturalist*, **46**, 174–226.

Harvey PH, Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford, UK.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **21**, 160–174.

Heusser CJ (1983) Vegetational history of the northwestern United States including Alaska. In: *Late Quaternary Environment of the United States* (eds Wright HE Jr, Porter SC). University of Minnesota Press, Minneapolis.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **359**, 183–195.

Hoffmann RS, Jones JK Jr (1970) Influence of late-glacial and post-glacial events on the distribution of recent mammals on the northern Great Plains. In: *Pleistocene and Recent Environments of the Central Great Plains* (eds Wakefield D Jr, Jones JK Jr), pp. 355–394. University Press of Kansas, Lawrence, Kansas.

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

Johnsgard PA (2002) *Grassland Grouse and Their Conservation*. Smithsonian Institution Press, Washington, D.C.

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

Johnsgard PA, Wood RE (1968) Distributional changes and interaction between prairie chickens and sharp-tailed grouse in the Midwest. *Wilson Bulletin*, **80**, 173–188.

Johnson JA, Bellinger MR, Toepfer JE, Dunn P (2004) Temporal changes in allele frequencies and low effective population size in greater prairie-chickens. *Molecular Ecology*, **13**, 2617–2630.

Johnson JA, Toepfer JE, Dunn PO (2003) Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. *Molecular Ecology*, **12**, 3335–3347.

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*, **17**, 1244–1245.

Lincoln FC (1917) A review of the genus *Pedioecetes* in Colorado. *Proceedings of the Biological Society of Washington*, **30**, 83–86.

Lucchini V, Hoglund J, Klaus S, Swenson J, Randi E (2001) Historical biogeography and a mitochondrial DNA phylogeny of grouse and ptarmigan. *Molecular Phylogenetics and Evolution*, **20**, 149–162.

Lundelius EL, Graham RW, Anderson E *et al.* (1983) Terrestrial vertebrate faunas. In: *Late-Quaternary Environments of the United States* (eds Wright HE Jr, Porter SC), pp. 311–353. University of Minnesota Press, Minneapolis, Minnesota.

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91–99.

Lynch M, Pfrender ME, Spitz K *et al.* (1999) The quantitative and molecular genetic architecture of a subdivided species. *Evolution*, **53**, 100–110.

Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Publications, Cold Spring Harbor, New York.

McKay JK, Latta RG (2002) Adaptive population divergence: Markers, QTL and traits. *Trends in Ecology & Evolution*, **17**, 285–291.

Miller MP (1997) *Tools for Population Genetic Analyses (TFPGA) 1.3: A Windows Program for the Analysis of Allozyme and Molecular Population Genetic Data*. http://www.marksgeneticsoftware.net/ (accessed on: 4 April 2005).

Miller MP (1998) *AMOVA-PREP* http://www.marksgeneticsoftware.net/ (accessed on: 30 August 2005).

Mock KE, Brim-Box JC, Miller MP, Dowling ME, Hoeh WR (2004) Genetic diversity and divergence among freshwater mussel (*Anodonta*) populations in the Bonneville Basin of Utah. *Molecular Ecology*, **13**, 1085–1098.

Mock KE, Theimer TC, Rhodes OE Jr, Greenberg DL, Keim P (2002) Genetic variation across the historical range of the wild turkey (*Meleagris gallopavo*). *Molecular Ecology*, **11**, 643–657.

Monsen KJ, Blouin MS (2003) Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. *Molecular Ecology*, **12**, 3275–3286.

Morgan K, Hicks J, Spitz K *et al.* (2001) Patterns of genetic architecture for life-history traits and molecular markers in a subdivided species. *Evolution*, **55**, 1753–1761.

Moritz C (1994) Defining 'evolutionarily significant units' for conservation. *Trends in Ecology & Evolution*, **9**, 373–375.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.

Mullenbach R, Lagoda PL, Welter C (1989) An efficient salt–chloroform extraction of DNA from blood and tissues. *Trends in Genetics*, **5**, 391.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction nucleases. *Proceedings of the National Academy of Sciences, USA*, **76**, 5269–5273.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, **19**, 153–170.

Ota T (1993) *Dispan*. Institute of Molecular Evolutionary Genetics, University Park, Pennsylvania.

Oyler-McCance SJ, Kahn NW, Burnham KP, Braun CE, Quinn TW (1999) A population genetic comparison of large- and small-bodied sage grouse in Colorado using microsatellite and mitochondrial DNA markers. *Molecular Ecology*, **8**, 1457–1465.

Palkovacs EP, Oppenheimer AJ, Gladyshev E *et al.* (2004) Genetic evaluation of a proposed introduction: the case of the greater prairie chicken and the extinct heath hen. *Molecular Ecology*, **13**, 1759–1769.

Piertney SB, Dallas JF (1997) Isolation and characterization of hypervariable microsatellites in the red grouse *Lagopus lagopus scoticus*. *Molecular Ecology*, **6**, 93–95.

Piertney SB, MacColl ADC, Bacon PJ *et al.* (2000) Matrilineal genetic structure and female-mediated gene flow in red grouse (*Lagopus lagopus scoticus*): An analysis using mitochondrial DNA. *Evolution*, **54**, 279–289.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pritchard JK, Wen W (2003) *Documentation for Structure Software: Version 2*. Available from http://pritch.bsd.uchicago.edu (accessed on: 21 June 2005).

Quinn TW (1992) The genetic legacy of Mother Goose — phylogeographic patterns of lesser snow goose *Chen caerulescens caerulescens* maternal lineages. *Molecular Ecology*, **1**, 105–117.

Quinn TW, Wilson AC (1993) Sequence evolution in and around the mitochondrial control region in birds. *Journal of Molecular Evolution*, **37**, 417–425.

Randi E, Lucchini V (1998) Organization and evolution of the mitochondrial DNA control region in the avian genus *Alectoris*. *Journal of Molecular Evolution*, **47**, 449–462.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608–615.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Ronquist F, Huelsenbeck JP, van der Mark P (2005) MRBAYES 3.1 manual. http://mrbayes.csit.fsu.edu/index.php (accessed on: 24 August 2005).

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN version 2.000: a software for population genetics data analysis*. Genetics and Biometry Laboratory, Geneva, Switzerland.

Schroeder MA, Aldridge CL, Apa A, Det al. (2004) Distribution of sage-grouse in North America. *Condor*, **106**, 363–373.

Segelbacher G, Paxton RJ, Steinbrück G, Trontelj P, Storch I (2000) Characterization of microsatellites in capercaillie *Tetrao urogallus* (Aves). *Molecular Ecology*, **9**, 1919–1952.

Snyder LL (1935) A study of the sharp-tailed grouse. *Royal Ontario Museum Biological Series*, No 40.

Snyder LL (1939) Great-plains races of sharp-tailed grouse. *The Auk*, **56**, 184–185.

Sparling DW (1983) Quantitative analysis of prairie grouse vocalizations. *Condor*, **85**, 30–42.

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

Spaulding WG, Leopold EB, Van Devender TR (1983) Late Wisconsin paleoecology of the American southwest. In: *Late-Quaternary Environments of the United States* (eds Wright S, Porter SC), pp. 259–293. University of Minnesota Press, Minneapolis, Minnesota.

Sullivan JJ, Markert JA, Kilpatrick CW (1997) Phylogeography and molecular systematics of the *Peromyscus aztecus* species group (Rodentia: Muridae) inferred using parsimony and likelihood. *Systematic Biology*, **46**, 426–440.

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

Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: *Molecular Systematics*. pp. 407–514. Sinauer Associates, Sunderland, Massachusetts.

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105**, 437–460.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of the mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767–782.

USFWS (2000) Endangered and threatened wildlife and plants: 12-month finding for a petition to list the Columbian sharp-tailed grouse as threatened. In: *Federal Register*, **65**, pp. 60391–60396. Department of the Interior.

Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Wang ZAJB, Hill GE, Edwards SV (2003) Reconciling actual and inferred population histories in the house finch (*Carpodacus mexicanus*) by AFLP analysis. *Evolution*, **57**, 2852–2864.

Warheit KI, Schroeder MA (2003) *Genetic Survey of Columbian Sharp-tailed Grouse Populations in Western North America. Summary of Microsatellite Results. Report, submitted to National Fish and Wildlife Foundation*. Washington Department of Fish and Wildlife, Olympia, Washington.

Watts WA (1983) Vegetational history of the eastern United States. In: *Late-Quaternary Environments of the United States* (ed. Wright HE Jr), pp. 294–310. University of Minnesota Press, Minneapolis, Minnesota.

Webb TI, Cushing EJ, Wright HE Jr (1983) Holocene changes in the vegetation of the Midwest. In: *Late-Quaternary Environments of the United States* (ed. Wright HE Jr), pp. 142–165. University of Minnesota Press, Minneapolis, Minnesota.

Wiens JJ, Servedio MR (1998) Phylogenetic analysis and intraspecific variation. Performance of parsimony, likelihood, and distance methods. *Systematic Biology*, **47**, 228–253.

Wilson AC, Cann RL, Carr SM *et al.* (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society*, **26**, 375–400.

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

Zink RM (2004) The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 561–564.

This manuscript is part A. Spaulding's dissertation investigating sexual selection's effect on species divergence rates in grouse and wood warblers. K. Mock focuses on landscape genetics in a wide range of western North American taxa. K. Warheit directs the Molecular Genetics Laboratory at Washington Dept. of Fish and Wildlife. His genetics research focuses on the structure of fish and wildlife populations. M. Schroeder is the Washington State Upland Bird Research Scientist, conducting research on the behavior, genetics, and population dynamics of grouse, and helping efforts to augment and/or re-establish populations of sage-grouse and sharp-tailed grouse that have declined or disappeared.

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd