Conserv Genet (2007) 8:659–670
DOI 10.1007/s10592-006-9212-z

ORIGINAL PAPER

Genetic differentiation of an endangered capercaillie (*Tetrao urogallus*) population at the Southern edge of the species range

R. Rodríguez-Muñoz · P. M. Mirol · G. Segelbacher · A. Fernández · T. Tregenza

Received: 27 February 2006 / Accepted: 31 August 2006 / Published online: 9 November 2006

Abstract The low-latitude limits of species ranges are thought to be particularly important as long-term stores of genetic diversity and hot spots for speciation. The Iberian Peninsula, one of the main glacial refugia in Europe, houses the southern distribution limits of a number of boreal species. The capercaillie is one such species with a range extending northwards to cover most of Europe from Iberia to Scandinavia and East to Siberia. The Cantabrian Range, in North Spain, constitutes the contemporary south-western distribution limit of the species. In contrast to all other populations, which live in pure or mixed coniferous forests, the Cantabrian population is unique in inhabiting pure deciduous forests. We have assessed the existence of genetic differentiation between this and other European populations using microsatellite and mitochondrial DNA (mtDNA) extracted from capercaillie feathers. Samples were collected between 2001 and 2004 across most of the current distribution of the Cantabrian population. Mitochondrial DNA analysis showed that the Cantabrian birds form a distinct clade with respect to all the other European populations analysed, including the Alps, Black Forest, Scandinavia and Russia, which are all members of a discrete clade. Microsatellite DNA from Cantabrian birds reveals the lowest genetic variation within the species in Europe. The existence of birds from both mtDNA clades in the Pyrenees and evidence from microsatellite frequencies for two different groups, points to the existence of a Pyrenean contact zone between European and Cantabrian type birds. The ecological and genetic differences of the Cantabrian capercaillies qualify them as an Evolutionarily Significant Unit and support the idea of the importance of the rear edge for speciation. Implications for capercaillie taxonomy and conservation are discussed.

Keywords ESU · Glaciation · Refugia · Taxonomy · Hybrid zone · Contact zone · Grouse

R. Rodríguez-Muñoz · T. Tregenza (✉)
Centre for Ecology and Conservation, School of Biosciences, University of Exeter, Cornwall Campus, Penryn TR10 9EZ, UK
e-mail: T.Tregenza@exeter.ac.uk

R. Rodríguez-Muñoz
e-mail: R.Rodriguez-Munoz@exeter.ac.uk

P. M. Mirol
CIGEBA, Facultad de Ciencias Veterinarias, Universidad Nacional de La Plata, B1900AVW La Plata, Argentina
e-mail: pmirol@macn.gov.ar

G. Segelbacher
Max Planck Institute of Ornithology, Vogelwarte Radolfzell, Schlossallee 2, D-78315 Radolfzell, Germany
e-mail: segelbac@orn.mpg.de

G. Segelbacher
Department of Wildlife Ecology and Management, University Freiburg, Tennenbacher Str. 4, D-79106 Freiburg, Germany

A. Fernández
Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas. , Pabellón del Perú, Avda. Maria Luisa s/n, 41 013 Sevilla, Spain
e-mail: albertofg.uo@uniovi.es

Present Address:
P. M. Mirol
Museo Argentino de Ciencias Naturales, Angel Gallardo 470, C1405DJR Buenos Aires, Argentina

Introduction

Besides its intrinsic interest in relation to the temporal processes driving speciation and species distributions (Barraclough and Nee 2001), the identification of species subunits from genetic analyses has recently become a matter of interest from a conservation perspective. Historically, nature conservation has focused on the protection of ecosystems and species as a whole. New insights support the importance of preserving genetic variation within species, defining Evolutionarily Significant Units (ESUs) as distinct population segments (Moritz 2002). Differences among populations of the same species can arise either from processes of divergent selection or from neutral processes particularly where there are colonisation events or long-term historical isolation (Moritz 2002; Tregenza 2002). The importance of quaternary glaciations in moulding the current phylogeographical patterns is well established (Avise and Walker 1998; Hewitt 2000). Temperature cycles linked to glacial periods have produced alternate expansions and contractions together with North and Southward shifts of species ranges (Hewitt 2000). During peaks of ice extension, most species have been relegated to a few milder climate areas located at low latitude (Taberlet et al. 1998). The importance of the low latitude range edges as long-term stores of genetic diversity and hot spots for speciation has been recently stressed (Martin and McKay 2004; Hampe and Petit 2005).

The Iberian Peninsula was one of the main glacial refugia in Europe (Taberlet et al. 1998), with a number of well-studied contemporary contact zones between historically Iberian populations and the descendants of more Easterly refugia (Butlin et al. 1992; Guillaume et al. 2000). Close to the Pyrenees, to the West, the Cantabrian Range runs parallel to the North Iberian coast, acting as a boundary between two very distinct biogeographical regions, the Atlantic to the North and the Mediterranean to the South (Gómez and Lunt 2006). This particular position has allowed this range to maintain stable populations of many Boreal and Central European species during warm periods as well as acting as a refugia during the glaciations (Gómez and Lunt 2006).

The capercaillie (*Tetrao urogallus*) is a bird species inhabiting coniferous and mixed forests from Western Europe to Eastern Siberia (Storch 2000). Phenotypic characters have led to the description of twelve different subspecies (de Juana 1994), but there is still a shortage of genetic analyses supporting current classification. Moreover, a recent study based on mtDNA found no clear evidence for the presence of the subspecies previously described in Finland from morphology, display song and allozyme loci (Liukkonen-Anttila et al. 2004). Although no accurate data are available on historical distribution, recent isolation of the Cantabrian population from its nearest neighbour in the Pyrenees is thought to date from at least the 18th century (Castroviejo et al. 1974). In contrast with the coniferous forests inhabited by all other populations, Cantabrian forests are deciduous, with a dominance of oak (*Quercus petraea* and *Quercus pyrenaica*) and beech (*Fagus sylvatica*). Evidence from palynological studies, indicates that conifers were replaced by deciduous forests during the Holocene, with only a small relict Scots pine (*Pinus sylvestris*) forest having survived (García Antón et al. 1997). The absence of conifers is a major habitat difference between capercaillies living in the Cantabrian range and those living elsewhere (Quevedo et al. 2006). Pine needles are one of the main food sources for capercaillies elsewhere in the world, whereas those inhabiting the Cantabrian Range have a distinct diet, with beech buds and holly tree leaves substituting the pine needles as the main source of food during the winter (Castroviejo 1975; Rodríguez and Obeso 2000). Based on this ecological divergence and some significant differences in morphology between the Cantabrian and the Pyrenean capercaillies, the former population was described as belonging to a different subspecies (Castroviejo 1967), the Cantabrian capercaillie (*T. u. cantabricus*). Because of its geographic location, at the edge of the species distribution and in one of the main glacial refugia across Europe (Hewitt 2000), the study of this population is particularly interesting. The Cantabrian capercaillie presumably survived through the Ice Ages somewhere in Iberia, and its contemporary distribution together with its relationships with other capercaillie populations are present day evidence of its evolutionary history.

It has been persuasively argued that the capercaillie is a good indicator or 'umbrella' species, since it is confined to undisturbed environments with high species diversity (Suter et al. 2002; Pakkala et al. 2003). Additionally, within the Cantabrian Range the capercaillie is, together with the brown bear, the major flagship species in relation to the conservation of protected areas. Thus, in practical terms, preservation of this population is closely linked to the preservation of the whole forest ecosystem. Over the last two decades both the population size and the distribution area of the Cantabrian capercaillie have decreased dramatically (Obeso 2003). Understanding the phylogeography and the genetic status of this population are two key issues for the rationalization of any recovery action to be taken.

By combining microsatellite and mtDNA analysis, we aim to assess (1) the degree of genetic differentiation of the Cantabrian capercaillie population, (2) its present genetic diversity, and (3) the position of the Pyrenees in the current phylogeographic scenario of the capercaillie in Europe.

Materials and methods

Sampling procedure

Between 2001 and 2004, we collected feathers found around lekking grounds, across most of the current range of the species in the Cantabrian Range (Fig. 1). Most samples were collected during early spring and late summer, in connection with the mating and moulting seasons, respectively. We selected sampling sites to cover the entire contemporary range of the population, but the very low density of the eastern and central parts of the range reduced the number of samples found in those areas. To assess phylogeographic status, we included samples from the Pyrenees as well as from other European populations (Segelbacher et al. 2003b).

Laboratory protocol

Total DNA was extracted from feathers using DNeasy Tissue Kit (QUIAGEN) following the manufacturer's instructions with some modifications already described for this species (Segelbacher and Storch 2002). Feathers were processed by cutting the tips into small pieces to be used together with the remains of skin attached around them.

Mitochondrial sequencing

The amplification of a fragment of the control region (CR) of the mtDNA was carried out with primers PHDL (5′-AGG ACT ACG GCT TGA AAA GC-3′, (Fumihito et al. 1995) and PH-H521 (5′-TTA TGT GCT TGA CCG AGG AAC CAG-3′, (Randi and Lucchini 1998). These primers amplify domains IA and IB of the CR, where variability is higher than in other domains in this region (Lucchini et al. 2001). The control region of the Pyrenean samples was PCR-amplified with the heavy strand primer (5′-GTG AGG TGG ACG ATC AAT AAA T-3′) annealing in the control region (3′ position—nucleotide 401 in *Gallus gallus*) and the light strand primer (5′-TTG TTC TCA ACT ACG GGA AC-3′) annealing in the adjacent tRNAglu region (3′ position—nucleotide 16741). PCR were performed in 50 µl volumes including 0.5 µM of each primer, 200 µM dNTPs, 1.5 mM $MgCl_2$ and 1U *Taq* polymerase (Bioline) in 1 × reaction buffer (as provided in the Bioline kit), with BSA at a final concentration of 0.1 µg/µl. The amplification consisted of an initial hot start of 15 min at 95°C, after which the enzyme was added, followed by 30 cycles of 1 min at 94°C, 1 min at 57°C and 1 min at 72°C, and a final incubation of 3 min at 72°C. The products were

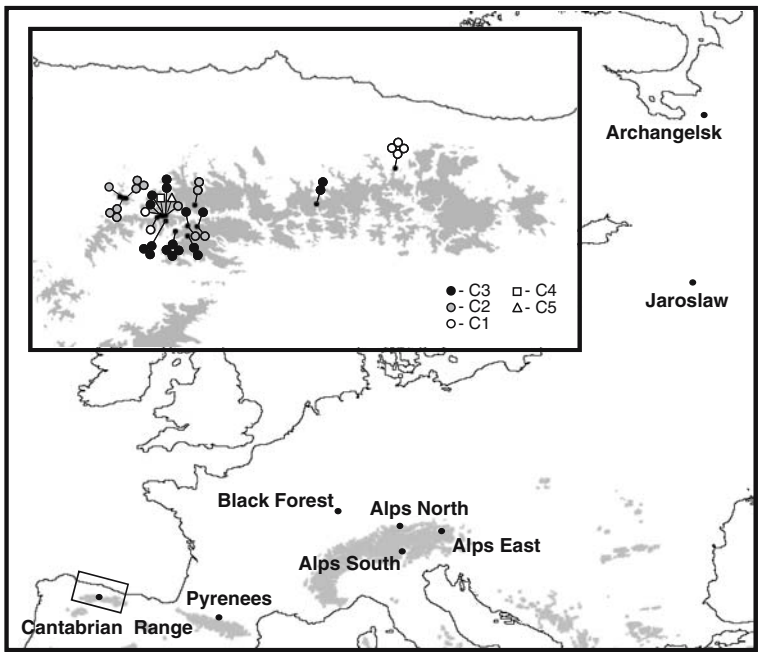

Fig. 1 Sample areas included in the analyses. Inset shows detail of the Cantabrian Range including the location of sample sites and the number of birds of each haplotype sampled at each site (each symbol represents a bird)

purified using QIAquick columns (QIAGEN) and sequenced with an automated sequencer. Sequences (GenBank accession numbers DQ398960–DQ396971) were aligned with CLUSTAL-X (Thompson et al. 1997).

Microsatellite typing

Individual samples were genotyped at 8 tetranucleotide microsatellite loci (Tut1, Tut2, Tut3, Tut4, BG 4, BG5, BG15, BG18; Segelbacher 2002). PCR amplifications and genotyping were conducted as described elsewhere (Segelbacher et al. 2000; Piertney and Höglund 2001; Segelbacher and Storch 2002). PCR fragments were resolved by electrophoresis by labelling the primers with fluorescent dye and running the fragments on an ABI 377 genetic analyser. We used internal standards (reference individuals) and external standards (ABI ROX350) for each single run to exclude any scoring errors. To detect whether contamination with exogenous DNA or PCR products had occurred, negative controls were included in each set of extractions and PCR amplifications. Amplification of the cloned locus aided in size determination and also served as a positive control. To avoid contamination, DNA extractions, pre-PCR pipetting and post-PCR pipetting were carried out in different rooms and aerosol-resistant filter pipette tips were used throughout. Each feather sample was typed at least three times to check for genotyping errors (e.g. false homozygotes.

Data analyses

Mitochondrial DNA

We analysed population genetic structure using analysis of molecular variance (AMOVA) in ARLEQUIN (Schneider et al. 1997), which was also used to calculate the frequency distributions of the nucleotide differences in all possible pairwise comparisons (mismatch distributions). Slatkin and Hudson (1991) demonstrated that the mismatch distributions of stable populations have a ragged profile due to stochastic lineage loss. In contrast, an exponentially growing population has a smooth unimodal distribution approaching a Poisson distribution. This reflects a star like genealogy in which all of the coalescent events occurred in a short period of time. We used Harpending's raggedness index (t) to test the significance of the distribution, which produces larger values for multimodal distributions. We also used ARLEQUIN to calculate Fu's test of selective neutrality (Fu 1997). This test evaluates the probability of observing a random neutral sample with a number of alleles similar or smaller than the observed value, given the observed number of pairwise differences, taken as an estimator of θ, and detects deviations from the pattern of polymorphism expected from a neutral model of evolution in a demographically stable population. The F statistic is very sensitive to population demographic expansion, which generally leads to large negative values.

We used the program FLUCTUATE (Kuhner et al. 1998) to make simultaneous estimates of present day θ and the population growth rate g, assuming an exponential model of growth and using a maximum likelihood approach. The parameters used for the simulations were obtained by running a hierarchy of likelihood-ratio tests in MODELTEST 3.0 (Posada and Crandall 1998) to choose the model of evolution which best fitted the data. MODELTEST calculates the likelihood ratio test statistic δ and its associated P-value using a χ^2 distribution in order to reject or fail different null hypothesis about the process of DNA substitution.

We inferred the phylogenetic relationships between species using the maximum likelihood approaches in PAUP 4.0b10 (Swofford 1998). We selected the model of nucleotide substitution using MODELTEST 3.0, as above. Heuristic searches were conducted with 1,000 random sequence addition replicates. Because classic phylogenetic methods are not directed toward analysis of intraspecific data, we constructed networks based on statistical parsimony using the program TCS 1.06 (Clement et al. 2000). Phylogenetic methods assume that ancestral haplotypes are no longer present; yet coalescent theory predicts that ancestral haplotypes will be the most frequent sequences sampled in a population level study. Statistical parsimony is particularly useful to estimate robust networks when few nucleotide differences exist among haplotypes, and to assign outgroup weights to haplotypes, allowing hypothesis testing about geographical origin (Emerson et al. 2001; Posada and Crandall 2001).

Microsatellites

To obtain standard estimates of genetic diversity within and between sample sites we used genotype and allele frequencies of the microsatellite loci. We assessed relative genetic variation in each population using allele frequency data to calculate the mean number of alleles and observed heterozygosity H_O, and gene diversity H_E (Nei 1972), and F_{IS} using the GENETIX software (Belkhir et al. 2004). We assessed the relative amount of genetic variation in each population using allele frequency data from which allelic richness

(Petit et al. 1998) was determined using FSTAT Version 2.93 (Goudet 2001). Allelic richness is a measure of the number of alleles independent of sample size, and hence allows comparison of this quantity between samples of different sizes. We used the clustering method described by Pritchard et al. (2000) to infer population structure and to assign individuals of different haplotypes to populations using multilocus genotype data as implemented in the program STRUCTURE 2.1. Assuming Hardy–Weinberg equilibrium and complete linkage equilibrium between loci within populations, allele frequencies and assignment of individuals to populations were inferred using a Bayesian approach. All runs were based on 100,000 iterations after a burn-in period of 20,000 iterations. A minimum of ten independent runs were conducted in order to assess the consistency of results across runs, using admixture and non-admixture models without incorporation of population information. As results of independent runs and between models did not differ much, we assume convergence was reached. Patterns of differentiation were visualized by Factorial Correspondence Analysis (FCA) of individual multilocus scores computed using GENETIX. This analysis is an ordination method that projects individuals into a multidimensional space according to their allelic composition. In this analysis we have omitted the Russian samples (Archangelsk and Jaroslaw, Table 1) because for the analysis of the Pyrenean samples we only used the possible neighbouring populations (Cantabrian Range and Alps). To perform the FCA, we used a subset of 90 birds from the Alps together with all the available samples from the Cantabrian Range and the Pyrenees.

We used Bottleneck 1.2.02 (Cornuet and Luikart 1996) to detect recent population bottlenecks using our allele frequency data. As suggested by Luikart et al (1998), in the absence of any information on the pattern of mutations in the microsatellites in use, we ran a TPM model (95% step-wise mutations with a variance of 5%); a Wilcoxon-test and observed mode shift in allele sizes provides evidence of past population size bottlenecks.

Table 1 Sampled localities and number of individuals used for mitochondrial DNA and microsatellite analyses

Region	mtDNA	Microsatellites	Total
Cantabrian Range	37	20	37
Pyrenees	22	16	22
Europe	13	275	2,756
Black Forest	3	0	3
Alps N	3	130	130
Alps S	5	36	36
Alps E	2	47	47
Archangelsk (Russia)	0	44	44
Jaroslawl (Russia)	0	18	18
Total	72	311	18

Results

Mitochondrial DNA

We examined a total of 72 samples, including 37 from the Cantabrian Range, 22 from a common locality in the Pyrenees and 13 from four other populations in Europe (three in the Alps and one in the Black Forest, Fig. 1). Sequences from the Cantabrian Range and Alps/Black Forest samples were 402 bp long, whilst those from the Pyrenees were 250 bp long, due to differences in the primers used in our two laboratories. We found five different haplotypes in the Cantabrian Range (C1 to C5). Haplotype C3 is the most frequent (17 individuals) and is found in 5 out of 9 different collection sites, followed by C2 (12 individuals) and C1 (6 individuals). Haplotypes C4 and C5 are unique. In the Alps/Black Forest we found seven different haplotypes (E1 to E7). Haplotype E1 is found in five individuals, E4 and E6 in two individuals each, and the rest are unique. In the Pyrenees we found four haplotypes, C4p (11 individuals), C2p (two individuals), E1p (seven individuals) and E2/E4/E6p (two individuals). There was total correspondence among those four haplotypes and C4, C2, E1 and E2/E4/E6, although because sequences from the Pyrenees are 250 bp long, we lose the differences between these three last haplotypes. Our results indicate that there are 13 individuals with Cantabrian-like haplotypes and nine with Alps/Black Forest like haplotypes in the Pyrenees. The most common Cantabrian-like haplotype in the Pyrenees (C4) is found only in one bird in the Cantabrian Range, whilst the most common haplotype in the Cantabrian Range (C3) has not been found in the Pyrenees. It is noteworthy that all the samples from the Pyrenees were collected in the same locality, which means that both haplogroups are living in sympatry.

There are five substitutions fixed between haplotypes C and E, plus one deletion. Nucleotide diversity is very similar in the Cantabrian Range and the Alps/Black Forest, but much higher in the Pyrenees (Table 2). Although the sequences obtained for the Pyrenean sample are shorter, it is very unlikely that in the 152 bp difference between these sequences and the rest, they could have accumulated enough shared differences between them that would have changed their inclusion within the E or C haplogroups. We detected

Table 2 Mitochondrial variation summary statistics and results from the FLUCTUATE analysis

Region	n	H	π	θ_S	θ_π	θ	g	Tajima's D	Fu's F
Cantabrian Range	37	5	0.004089 (0.002742)	0.00298 (0.00256)	0.00408 (0.00439)	0.0026	18.3	0.95597 $P = 0.177$	0.99118 $P = 0.723$
Alps/Black Forest	13	7	0.004529 (0.003128)	0.00481 (0.00256)	0.00453 (0.00313)	0.0448	1489.4	–0.21576 $P = 0.436$	–2.59417 $P = 0.023$
Pyrenees	22	4	0.018286 (0.010419)	0.01207 (0.00524)	0.01829 (0.01042)	–	–	1.78189 $P = 0.043$	5.46258 $P = 0.979$

n, Number of sequences surveyed; H, number of different haplotypes found in each region; π, nucleotide diversity; θ_S, from Waterson (1975), θ_π, from Tajima (1989); θ and g, estimated with FLUCTUATE; D and F, estimated with ARLEQUIN. FLUCTUATE analysis was not run on the Pyrenees due to the hybrid nature of the population. Standard errors are shown between brackets

only three mutations in the European and Cantabrian birds in those 152 bases, and it is highly improbable that the Pyrenean birds could have that portion of the CR evolving at a much higher rate than the rest of the sample. Genetic differentiation among populations from the Cantabrian Range and the Alps/Black Forest was tested using AMOVA. The Pyrenees location was excluded from this analysis because it shows haplotypes from both other geographic locations. There were significant differences between both geographical locations, Cantabrian Range and Alps/Black Forest ($\Phi_{ST} = 0.56$, $P < 0.0001$), 19.6% of the variation was due to differences among groups while differences among populations within geographic regions accounted for 36.6% of the variation.

The phylogenetic relationships between the populations were reconstructed using maximum likelihood and parsimony approaches, with *T. parvirostris* (AJ297179), *T. mlokosiewiczi* (AJ297173) and *T. tetrix* (AJ297153) as outgroups (Lucchini et al. 2001). The analysis was based on the 402 bp long sequences C1 to C5 and E1 to E7. The parameters for the maximum likelihood analysis were obtained using MODELTEST. The model favoured was HKY+G, with a transition/transversion ratio of 3.57 and a shape parameter (alpha) of 0.0136. The maximum likelihood tree (Fig. 2a) shows one monophyletic clade for *T. urogallus*. Within that clade, Cantabrian and European haplotypes appear paraphyletic, with the Cantabrian haplotypes closer to the root. This result might be taken as a preliminary indication of a possible ancestrality of the Cantabrian birds with respect to the European birds, although it needs a more thorough examination. The parsimony consensus tree (Fig. 2b) and the neighbour-joining tree (not shown) display a clear dichotomy of both lineages from a common ancestor, with high bootstrap values. Maximum likelihood methods of phylogenetic reconstruction take into account information about branch lengths as well as the model of evolutionary change, so they are consistent under many situations in which parsimony and distance are inconsistent (Hillis et al. 1994; Kuhner and Felsenstein 1994; Huelsenbeck 1995). Under this framework, the ancestrality of the Cantabrian in relation to the European birds should be given careful consideration. The TCS network (Fig. 3) shows that the Cantabrian and European haplotypes form two well-differentiated lineages separated by five fixed substitutions plus one deletion. The highest outgroup weight (0.023) was obtained for a Cantabrian haplotype, C4, and although this result should be taken with caution due to the differences in sample size between Cantabrian and European birds, it is in clear agreement with the maximum likelihood tree.

The hybrid nature of the Pyrenees can also be seen in the mismatch distributions (Fig. 4). The distribution for the Alps/Black Forest is smooth and unimodal (Fig. 4c) indicating a population at demographic expansion; Harpending's raggedness index is low and not significant ($r = 0.071$, $P = 0.56$). This result is also suggested by the FLUCTUATE analysis, where estimates of the growth parameter are positive (Table 2) and compatible with those of an expanding population. In the Cantabrian Range, the mismatch distribution is multimodal (Fig. 3a) with a significant raggedness index of 0.37 ($P = 0.02$), indicating a stationary population. However, the FLUCTUATE analysis shows a positive, although very low, growth rate and a very low theta (Table 2). The mismatch distribution in the Pyrenees shows two distinct modes of number of nucleotide differences, one with small numbers of differences (1–3) corresponding to pairwise comparisons within the C and E groups of haplotypes, and the other (9–13 differences) corresponding to comparisons between C and E. The raggedness value is intermediate between the other two ($r = 0.26$, $P = 0.06$), which might also be a consequence of the presence of both clades in the population.

Because an excess of low frequency mutations accompanies range expansion, another possible way to detect demographic expansion is through neutrality tests. Table 2 shows the results obtained for the three

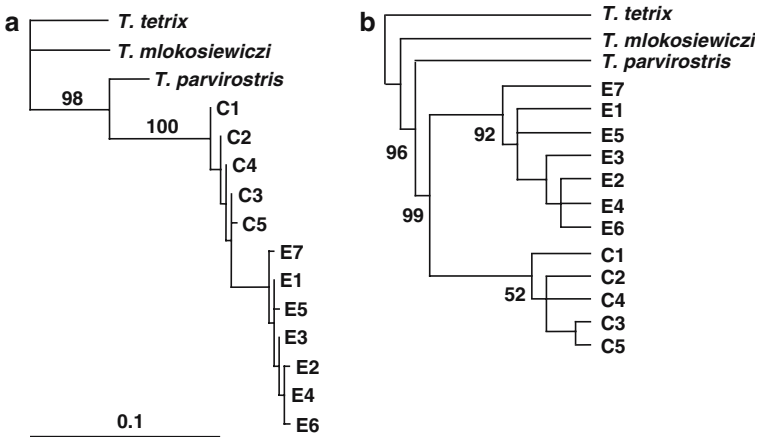

Fig. 2 Mitochondrial DNA phylogenies for European capercaillies (C, haplotypes found in the Cantabrian Range; E, haplotypes found in the Alps and the Black Forest). (**a**) Maximum likelihood topology; (**b**) Parsimony tree. Bootstrap values higher than 50% are shown at internodes. Sequences from the Pyrenees were excluded (see text)

regions. Only the statistics corresponding to Alps/Black Forest are negative, with a significant Fu's *F*. Thus the standard neutral model can be rejected for the Cantabrian Range and the Pyrenees but not for the Alps/Black Forest, and the observed patterns are concordant with the hypothesis of demographic expansion in this group and stability in the two others.

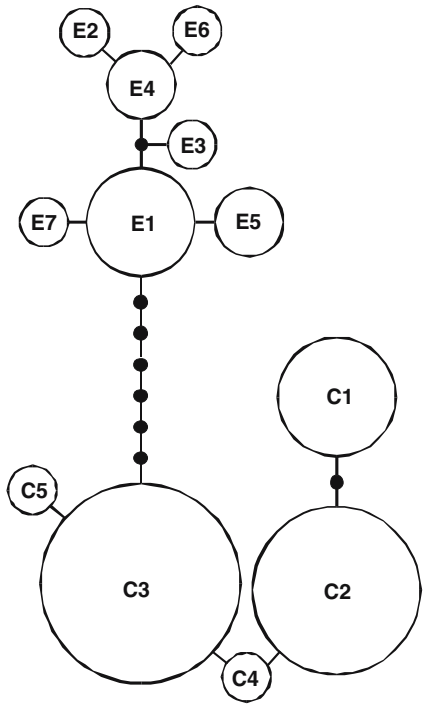

Fig. 3 Network obtained for the *T. urogallus* haplotypes using statistical parsimony. Circles represent haplotypes, with the area of the circle proportional to the frequency of the haplotype. Small black circles on the lines connecting haplotype circles indicate substitutions. We checked all other sequences published to date in GeneBank corresponding to Northern Europe and Russia, and all of them clusterize with clade E. We have excluded Pyrenean sequences because they were sequenced for only 250 bp

Microsatellites

We investigated genetic variation diversity of all populations by analysing departures from Hardy–Weinberg distribution and linkage equilibrium. We did not find any evidence for linkage disequilibrium at any loci, but the Pyrenean ($P = 0.0004$), and the Northern Alpine ($P = 0.004$) population showed a significant deviation from Hardy–Weinberg expecta-

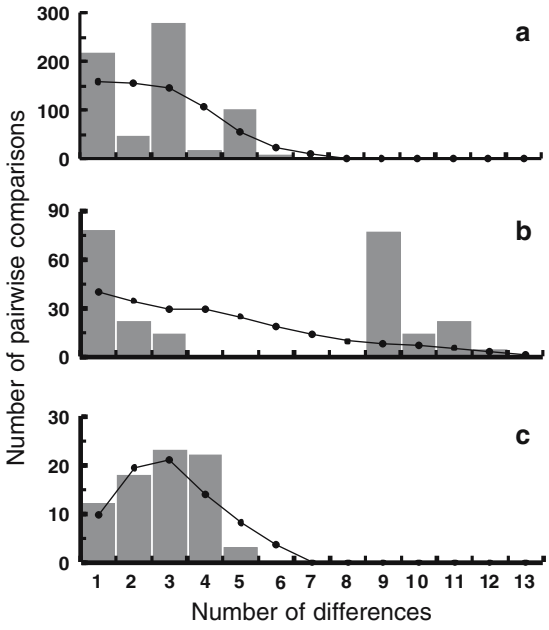

Fig. 4 Mismatch distributions among haplotypes of *Tetrao urogallus* and their fit to the stepwise growth model according to ARLEQUIN (**a**) Cantabrian Range (402 bp sequences); (**b**) Pyrenees (250 bp sequences); (**c**) North and Central Europe (402 bp sequences)

tions ($P = 0.007$). The lowest degree of heterozygosity, allelic richness and mean number of alleles were detected in the Cantabrian Range (Table 3). Pyrenean birds also displayed very low genetic diversity compared to Alpine and Boreal populations (Table 3). Pairwise population F_{ST} values ranged from 0.023 to 0.256 (Table 4) and were significant for all pairings even after Bonferroni adjustment. The highest F_{ST} values were found between the Cantabrian and the other populations indicating that birds from the Cantabrian Range are genetically the most distinct. The lowest values were found between populations from the Alpine metapopulation system.

A FCA was conducted using 126 specimens available from the Cantabrian Range ($N = 20$, the Pyrenees ($N = 16$) and the Alps (North, $N = 32$; East, $N = 36$; South, $N = 22$), since these are the neighbour populations for which we had reasonably high sample sizes (Fig. 5). The first axis of the FCA explains 7.0% of the variation and the second explains 5.2%. There is a clear cluster of points corresponding to the Cantabrian birds, and a second cluster corresponding to the European specimens. Individuals from the Pyrenees appear dispersed between both clades. This finding was also supported by our results of the assignment using STRUCTURE. For $k = 2$, all but one of the Pyrenean birds could be clearly assigned to one of two main clusters. Alpine birds could be attributed to one main cluster, whereas Cantabrian birds form a second cluster. When we classify Pyrenean individuals according to their mitochondrial haplotypes using STRUCTURE, all mitochondrial clade E individuals but one are found within the left side of the E group. However, half of the mitochondrial clade C individuals appear within the C cluster, while the other half appear mixed with the E cluster.

Our bottleneck analysis revealed evidence for recent contraction of population size only in the Cantabrian population (mode of allele sizes shifted from the expected L-shaped distribution). All other populations did not show any sign of recent bottleneck events.

Discussion

Genetic differentiation

Our mtDNA analysis reveals the existence of two clearly distinct capercaillie clades in Western Europe. Drovetski (2003) estimated mutation rates for the grouse control region to be 7.23% per million years based on a molecular clock calibrated with the fossil record. According to these times, the 1.24% sequence divergence found between the Cantabrian and the European clades would correspond to an isolation time of 171,000 years. This genetic difference matches those observed in morphology and ecology between the Cantabrian and all other capercaillie populations. Cantabrian capercaillies have a lighter colour and smaller beak and they are the only subspecies inhabiting pure deciduous forests (Castroviejo 1975). The absence of Cantabrian haplotypes from the Central, East and North European populations, and the existence of birds from both clades in the Pyrenees suggest that there is a contact zone in that range presumably dating from shortly after the end of the last glaciation around 9,000 years ago (Hewitt 2001). Moreover, it suggests the existence of at least two quaternary capercaillie refugia, one of them in the Iberian Peninsula, and the other presumably in the Italian Peninsula, the Balkans or further East. Several other taxa (*Chorthippus parallelus, Erinaceus europaeus, Ursus arctos*) have been shown to follow a similar biogeographical pattern (Hewitt 2000). A hybrid zone formed in the Pyrenees has been extensively described in the grasshopper *Chorthippus parallelus* (Butlin and Hewitt 1985; Cooper et al. 1995), as the result of secondary contact between expanding populations from Spain and the Balkans. In this case, mainland Europe was colonized by animals migrating from the Balkans, while the animals that survived in Spain did not expand Eastwards from the Pyrenees. Other examples of species that survived in refugia between the Pyrenees and the Cantabrian Range are the urodele *Salamandra*

Table 3 Genetic diversity of capercaillie populations at 8 microsatellite loci

Population	n	R	A	H_o	H_e	F_{IS}
Cantabrian Range	20	2.45	2.62	0.36	0.36	–0.022
Pyrenees	16	3.25	3.37	0.48	0.53	0.091
Alps N	130	4.84	7.13	0.72	0.70	–0.030
Alps S	36	4.59	5.88	0.71	0.71	0.000
Alps E	47	4.75	6.37	0.64	0.65	0.011
Archangelsk	44	4.68	6.00	0.72	0.72	–0.004
Jaroslawl	18	4.73	5.13	0.68	0.73	0.071

n, Number of individuals analysed; R, allelic richness; A, mean number of alleles per locus; H_o, mean observed heterozygosity; H_e, expected heterozygosity; F_{IS}, inbreeding coefficient

Table 4 Pairwise FST values for all population comparisons. All values are significant after Bonferroni correction ($P < 0.05$). Indicative adjusted nominal level (0.05) for multiple comparisons is: 0.0024

	Pyrenees	Alps N	Alps S	Alps SE	Archangelsk	Jaroslawl
Cantabrian Range	0.159	0.187	0.213	0.166	0.256	0.229
Pyrenees		0.121	0.121	0.103	0.156	0.122
Alps N			0.029	0.037	0.099	0.102
Alps S				0.022	0.108	0.098
Alps E					0.127	0.134
Archangelsk						0.027

salamandra (Alcobendas et al. 1996) and the lizard *Lacerta vivipara* (Guillaume et al. 2000). The capercaillie however, is the largest and most mobile animal that has been found with a likely Pyrenean contact zone suggesting that the post-glacial barrier represented by the lack of suitable habitat in the Pyrenean mountain chain and its coastal margins might have been sufficient to prevent Iberian populations of even large vertebrates from recolonizing Europe at the end of the last glaciation. This points to the possibility that Iberian populations of many animals may be genetically distinct from their northerly conspecifics, supporting the view that low latitude populations are important as hot spots for genetic differentiation (Hampe and Petit 2005) as shown by Martin and McKay (2004). An alternative explanation for the existence of both haplotypes in the Pyrenees might be that past reintroduction events using specimens from allochthonous populations. However, although unsuccessful attempts to reinforce the Pyrenean population were carried out in the Spanish side during the 1970's these used local birds and to our knowledge, no birds have been ever been introduced from other populations.

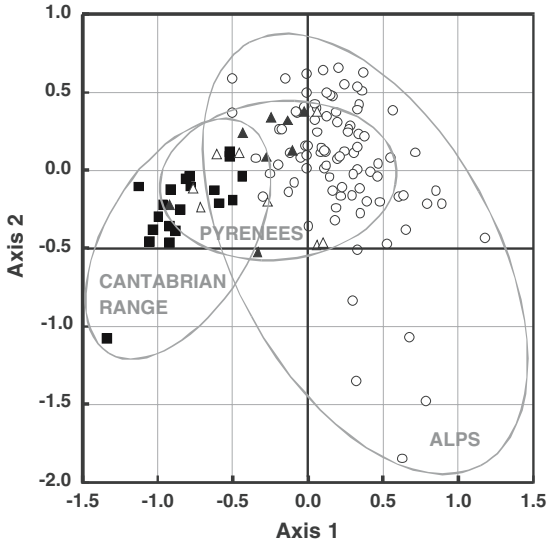

Fig. 5 Distribution of microsatellite allele frequencies shown as Factorial Correspondence Analysis scores. Alps, circles; Cantabrian Range, squares; Pyrenees, triangles. mtDNA clade C, black; mtDNA clade E, white. Ellipses include all individuals from each of the 3 population designations

An interesting result is the disagreement between the most common C haplotype found in the Cantabrian Range and the one found in the Pyrenees. It is generally believed that populations at the edge of the refugial area will lead the expansion when the climate ameliorates, and long distance dispersers will rapidly fill new territories (Hewitt et al. 1989; Hewitt 1993). These dispersers are expected to contribute disproportionately to the genetic composition of the founded populations (Ibrahim et al. 1996), and so not all refugial diversity will be represented. This is even more important if refugia are structured into a number of smaller units (Gómez and Lunt 2006). Within this framework, it is possible to understand the finding of C4 as the most common haplotype in the Pyrenees if it arrived there as part of a leading edge colonization, whilst being lost or driven to low frequency in the Cantabrian Range as a consequence of a recent bottleneck. Furthermore, if the indications of ancestrality of the Cantabrian haplotypes are correct, the pre- and post-glacial scenario could be slightly different than the one proposed before. It is possible that Cantabrian-like haplotypes were broadly distributed in pre-glacial Europe, and that the distribution of these types was fragmented with the advance of the ice sheet. Cantabrian-like haplotypes could have survived in different Mediterranean refugia, with new European-like haplotypes originating in the Italian or Balkan refugia. When the surviving populations expanded during the warm periods, individuals carrying the derived European and the Cantabrian haplotypes met in the Pyrenees originating the contact zone. This scenario would explain the topology found with the maximum likelihood approach, and would predict the existence of Cantabrian-like haplotypes in the Balkans or Italy.

The microsatellite markers do not suggest the existence of any barrier to gene flow between haplotype classes in the Pyrenees although the population is not in Hardy–Weinberg equilibrium. Genetic variability

levels are intermediate between those found in the Cantabrian Range and in the rest of Europe. Although the microsatellite analysis seems to indicate clinal variation in frequency across the range, it would be necessary to extend the mitochondrial characterization to populations to both sides of the Pyrenees in order to asses the clinal or sharp nature of the contact zone. FCA analysis shows most of the individuals with European mitochondrial type close to the European cluster, but this is not reciprocal for individuals with Cantabrian haplotypes. STRUCTURE analysis shows in many cases an equivocal assignment of C individuals to the E cluster and vice-versa. More exhaustive research is needed to elucidate the actual reproductive condition of the Pyrenean birds, which must include a broader geographic representation of the Pyrenean distribution of *T. urogallus.*

Genetic diversity

Based on microsatellite allelic variation and heterozygosity we found extremely low genetic diversity of capercaillies within the Cantabrian Range. Diversity was lower than in previously studied populations (Segelbacher et al. 2003a) including Pyrenean and isolated central European populations. As well as being the most genetically depauperate, the Cantabrian population was also the most genetically distinct population in Europe. The low genetic diversity may be the result of long-term isolation of the Cantabrian population suggesting that the population has been very small for enough generations to mean that diversity has been lost through genetic drift, as has been suggested for small isolated populations of black grouse (*T. tetrix*) (Höglund et al. submitted).

Differences between Cantabrian and other European populations reside in the frequency of different alleles, but we have not found any exclusive allele for any of the two groups. In this context, the population in the Pyrenees shows intermediate frequency values between populations to the east and to the west, suggesting that it has mixed ancestry, either historically, or because it is currently a contact zone between divergent clades from the Iberian peninsula and from elsewhere in Europe.

Mitochondrial DNA shows very similar nucleotide diversity in the Cantabrian and the Alps/Black Forest populations, although haplotype frequencies are very different. This result is not all surprising, even when microsatellite frequencies indicate very low diversity of Cantabrian populations compared with European populations, the similar levels of mitochondrial diversity could be indicating the persistence of ancestral polymorphisms, which have not yet been affected by population decline.

In the Cantabrian Range we found only five haplotypes in a sample of 37 birds, two of which are unique. Six out of nine locations sampled revealed only one haplotype, two had two different haplotypes and the last location had four haplotypes. Populations in Europe included 13 specimens from three locations in the Alps, which had two, three and four different haplotypes, and one location with one haplotype in the Black Forest. The shape of the individual networks (Fig. 3) is clearly different, the Alps/Black forest network indicates a population in expansion, as it is also shown by the mismatch and the FLUCTUATE analysis. Fu's *F*, which is very sensitive to population demographic expansion, is also significant (Table 2). On the other hand, the Cantabrian network suggests a stationary or declining population, with few and frequent haplotypes. The multimodal mismatch distribution (Fig. 4a) also indicates that this population is in a demographic equilibrium although this analysis cannot distinguish between equilibrium and decline, and it seems more likely that the Cantabrian population is actually decreasing its effective population size or has gone to a recent bottleneck, as indicated by microsatellites.

Implications for conservation

Our results clearly demonstrate that the Cantabrian capercaillie qualifies to be considered as an Evolutionarily Significant Unit (Moritz 2002). The combination of genetic data with the available information on recent population trends and distribution changes (Obeso 2003) suggests that its present status should be defined as critical. In addition to its interest as an 'umbrella' or 'indicator' species (Suter et al. 2002; Pakkala et al. 2003), there are strong social and political factors acting at a regional scale that confer to the capercaillie a key role in the overall conservation of the Cantabrian Range. Despite the existence of abundant legislation for the protection of a number of other endangered species, only capercaillie and brown bear carry any weight in assessing the impact of human activities in natural environments in this area. The extinction of capercaillie in the Eastern parts of the Cantabrian Range has already started to be used by developers to argue that there is no longer any need to conserve their former habitats. Because the whole area inhabited by the population during the 70's is below the minimum area established by the IUCN to confer a population the status of endangered, it is essential that conservation measures be extended to that whole area

if any serious recovery plan is to be developed. Therefore, action to protect this population should be started urgently.

There are several important remarks for conservation that can be inferred from the phylogenetic and population genetics analyses. The Cantabrian capercaillie belongs to a group that is genetically distinct from those living beyond the Pyrenees, so non-local birds should never be used if any translocation is planned. Part of the Pyrenean population might be suitable for genetic exchange, but further research is essential before that possibility can be accepted. The low genetic variability and heterozygosity might be a consequence of population fragmentation and inbreeding, two important factors driving extinction processes (Brook et al. 2002; Reed 2004). Thus, the geographical distribution of genetic variability should be urgently assessed, and the population should be managed accordingly to minimize further allelic losses that could reduce the viability of the population (for instance, translocation of birds or eggs if inbreeding depression is detected). Identification of source and sink areas is essential in making decisions about any possible translocation. Action aiming to reduce the risks derived from genetic impoverishment will be a waste of resources unless the causes of decline are identified and corrected.

Acknowledgements We thank to all who provided feathers (Angel Nuño, Luis Fernández, J. Manuel Carral and L.A. Álvarez Usategui) and to Félix González Álvarez by supplying us with the figure maps. R. Rodríguez-Muñoz is supported by a fellowship from the Leverhulme Trust (F/00122/T). T. Tregenza is supported by a Royal Society Fellowship and the European Social Fund. G. Segelbacher received a fellowship from the Max Planck Society. P. Mirol is a fellow of the National Research Council in Argentina. We are grateful to two reviewers for valuable comments on an earlier version of this manuscript.

References

Alcobendas M, Dopazo H, Alberch P (1996) Geographic variation in allozymes of populations of *Salamandra salamandra* (Amphibia: Urodela) exhibiting distinct reproductive modes. J Evol Biol 9:83–102

Avise JC, Walker D (1998) Pleistocene phylogeographic effects on avian population and the speciation process. Proc R Soc B 265:457–463

Barraclough TG, Nee S (2001) Phylogenetics and speciation. Trends Ecol Evol 16:391–399

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier

Brook BW, Tonkyn DW, O'Grady JJ, Frankham R (2002) Contribution of inbreeding to extinction risk in threatened species. Conserv Ecol 6:16

Butlin RK, Ferris C, Gosalvez J, Hewitt GM, Ritchie MG (1992) Broad-scale mapping of a hybrid zone between subspecies of *Chorthippus parallelus* (orthoptera, acrididae). Ecol Entomol 17:359–362

Butlin RK, Hewitt GM (1985) A hybrid zone between *Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus* (Orthoptera, Acrididae) – behavioural characters. Biol J Linn Soc 26:287–299

Castroviejo J (1967) Eine neue auerhuhnrasse von der Iberisschen Halbinsel. J Ornithol 108:220–221

Castroviejo J (1975) El urogallo en España. Consejo Superior de Investigaciones Científicas, Madrid

Castroviejo J, Delibes M, García Dory MA, Garzón J, Junco E (1974) Censo de urogallos cantábricos (*Tetrao urogallus cantabricus*). Asturnatura 2:53–74

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

Cooper SJB, Ibrahim KM, Hewitt GM (1995) Postglacial expansion and genome subdivision in the European grasshopper *Chorthippus parallelus*. Mol Ecol 4:49–60

Cornuet J, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014

de Juana E (1994) Tetraonidae. In: del Hoyo J, Elliott A, Sargatal J (eds) Handbook of the birds of the world. New world vultures to guineafowl. Lynx Edicions, Barcelona, pp 376–411

Drovetski SV (2003) Plio-Pleistocene climatic oscilations, Hjolartic biogeography and speciation in an avian submamily. J Biogeogr 30:1173–1181

Emerson BC, Paradis E, Thébaud C (2001) Revealing the demographic histories of species using DNA sequences. Trends Ecol Evol 16:707–716

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

Fumihito A, Miyake T, Takada M, Ohno S, Kondo N (1995) The genetic link between the Chinese bamboo partridge (*Bombusicola thoracica*) and the chicken and junglefowls of the genus Gallus. Proc Natl Acad Sci USA 92:11053–11056

García Antón M, Franco Múgica F, Maldonado J, Morla Juaristi C, Sainz Ollero H (1997) New data concerning the evolution of the vegetation in Lillo pinewood (León, Spain). J Biogeogr 24:929–934

Gómez A, Lunt DH (2006) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N (eds) Phylogeography in Southern European refugia: evolutionary perspectives on the origins and conservation of european biodiversity, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 155–188

Goudet J (2001) FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices

Guillaume GP, Heulin B, Arrayago MJ, Bea A, Braña F (2000) Refuge areas and suture zones in the Pyrenean and Cantabrian regions: geographic variation of the female MPI sex-linked alleles among aviparous populations of the lizard Lacerta (Zootoca) vivipara. Ecography 23:3–10

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8:461–467

Hewitt G (2000) The genetic legacy of the quaternary ice ages. Nature 405:907–913

Hewitt GM (1993) Postglacial distribution and species substructure: lessons from pollen, insects and hybrid zones. In: Lees DR, Edwards D (eds) Evolutionary patterns and processes. Academic Press, London, pp 97–123

Hewitt GM (2001) Speciation, hybrid zones and phylogeography or seeing genes in space and time. Mol Ecol 10:537–549
Hewitt GM, Mason P, Nichols RA (1989) Sperm precedence and homogamy across a hybrid zone in the alpine grasshopper *Podisma pedestris*. Heredity 62:343–353
Hillis D, Huelsenbeck J, Cunningham C (1994) Application and accuracy of molecular phylogenies. Science 264:671–677
Höglund J, Larsson J, Jansman H, Segelbacher G (submitted) Effects of isolation and population size on genetic variability in European black grouse
Huelsenbeck J (1995) Performance of phylogenetic methods in simulation. Syst Biol 44:17–48
Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity 77:282–291
Kuhner M, Felsenstein J (1994) A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Mol Biol Evol 11:459–468
Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. Genetics 149:429–434
Liukkonen-Anttila T, Ratti O, Kvist L, Helle P, Orell M (2004) Lack of genetic structuring and subspecies differentiation in the capercaillie (*Tetrao urogallus*) in Finland. Ann Zool Fenn 41:619–633
Lucchini V, Höglund J, Klaus S, Swenson J, Randi E (2001) Historical biogeography and a mitochondrial DNA phylogeny of grouse and Ptarmigan. Mol Phylogenet Evol 20:149–162
Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered 89, 238–247
Martin PR, McKay JK (2004) Latitudinal variation in genetic divergence of populations and the potential for future speciation. Evolution 58:938–945
Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. Syst Biol 51:238–254
Nei M (1972) Genetic distance between populations. Am Nat 106:283–292
Obeso JR (2003) Evolución temporal de las poblaciones cantábricas. In: Obeso JR, Bañuelos MJ (eds) El urogallo (*Tetrao urogallus cantabricus*) en la Cordillera Cantábrica. Ministerio de Medio Ambiente, Madrid, pp 19–28
Pakkala T, Pellikka J, Linden H (2003) Capercaillie Tetrao urogallus – a good candidate for an umbrella species in taiga forests. Wildlife Biol 9:309–316
Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12:844–855
Piertney SB, Höglund J (2001) Polymorphic microsatellite DNA markers in Black Grouse (*Tetrao tetrix*). Mol Ecol 1:303–304
Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818
Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. Trends Ecol Evol 16:37–45
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
Quevedo M, Bañuelos MJ, Sáez O, Obeso JR (2006) Habitat selection by Cantabrian capercaillie at the edge of the species distribution. Wildlife Biol 12:269–278
Randi E, Lucchini V (1998) Organization and evolution of the mitochondrial DNA control region in the avian genus Alectoris. J Mol Evol 47:449–462
Reed DH (2004) Extinction risk in fragmented habitats. Anim Conserv 7:181–191
Rodríguez AE, Obeso JR (2000) Diet of the Cantabrian capercaillie: geographic variation and energetic content. Ardeola 47:77–83
Schneider S, Kueffer JM, Roessli D, Excoffier L (1997) Arlequin version 1.1: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneve, Switzerland
Segelbacher G (2002) Noninvasive genetic analysis in birds: testing reliability of feather samples. Mol Ecol Notes 2:367–369
Segelbacher G, Höglund J, Storch I (2003a) From connectivity to isolation: genetic consequences of population fragmentation in capercaillie across Europe. Mol Ecol 12:1773–1780
Segelbacher G, Paxton RJ, Steinbruck G, Trontelj P, Storch I (2000) Characterization of microsatellites in capercaillie Tetrao urogallus (AVES). Mol Ecol 9:1934–1935
Segelbacher G, Storch I (2002) Capercaillie in the Alps: genetic evidence of metapopulation structure and population decline. Mol Ecol 11:1669–1677
Segelbacher G, Storch I, Tomiuk J (2003b) Genetic evidence of capercaillie *Tetrao urogallus* dispersal sources and sinks in the Alps. Wildlife Biol 9:267–273
Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129:555–562
Storch I (2000) Grouse status survey and conservation action plan 2000–04, IUCN, Gland, Switzerland and Cambridge, UK and the World Pheasant Assoc., Reading, UK
Suter W, Graf RF, Hess R (2002) Capercaillie (*Tetrao urogallus*) and avian biodiversity: testing the umbrella-species concept. Conserv Biol 16:778–788
Swofford, DL (1998) Phylogenetic Analysis Using Parsimony. Version 4. Sinauer Associates, Sunderland, Massachusetts
Taberlet P, Fumagalli L, Wustsaucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. Mol Ecol 7:453–464
Thompson J, Gibson T, Plewniak F, Jeanmougin F, Higgins D (1997) The Clustal X windows interfase: flexible strategies for multiple sequence alignement aided by quality analysis tools. Nucleic Acids Res 24:4673–4680
Tregenza T (2002) Divergence and reproductive isolation in the early stages of speciation. Genetica 116:291–300