# Ukrainian Black Grouse Tetrao tetrix: Genetic Diversity and Population Structure

Authors: Pavlovska, Mariia, and Höglund, Jacob

Source: Wildlife Biology, 21(6) : 283-293

Published By: Nordic Board for Wildlife Research

URL: https://doi.org/10.2981/wlb.00093

**Wildlife Biology 21: 283–293, 2015**
doi: 10.2981/wlb.00093

Subject Editor: Sara Oyler-McCance. Editor-in-Chief: Ilse Storch. Accepted 31 January 2015

# Ukrainian black grouse *Tetrao tetrix*: genetic diversity and population structure

**Mariia Pavlovska and Jacob Höglund**

*M. Pavlovska (mawwwa88@gmail.com), CISCoS Ecolab, National Pedagogical Dragomanov University, Pirogova Str. 9, UA-01601 Kyiv, Ukraine. – J. Höglund, Dept. of Ecology and Genetics, Uppsala Univ., Norbyvägen 18D, SE-75236 Uppsala, Sweden*

We studied neutral (microsatellite) and adaptive (MHC) genetic diversity of two Ukrainian populations of black grouse *Tetrao tetrix* to fill a gap in the data on the diversity of European populations of this species. We found that Ukrainian populations are more diverse than recently fragmented central European ones and also highly differentiated from these populations. Both studied populations in Ukraine, the northern and the Carpathian, did not show any signs of recent bottleneck events. The population structure among Ukrainian black grouse was more pronounced for neutral variation than for adaptive MHC diversity, suggesting that a common selective force has led to balancing selection shaping the MHC diversity. The MHC differentiation among the two studied regions was still high (Dest = 0.454), which could be a sign of local adaptation as a response to shifting balance in space. Thus, we suggest that the northern and the Carpathian black grouse populations should be treated as separate Management Units (MU). The black grouse population in the Carpathian Mountains appeared to be more diverse than the one in the north in terms of neutral genetic variation, whereas in terms of adaptive variation the two populations vary in diversity depending on the method of scoring diversity. Therefore, we suggest that the Carpathian Mountains could have been a refuge for the black grouse during the last glaciation.

In the past conservation genetic studies have documented neutral genetic variation using microsatellite markers (Jehle and Arntzen 2002, Moss et al. 2003, Oliveira et al. 2006). Such markers can be used to infer population structure and to study demographic history, inbreeding, and genetic drift. In contrast, Major histocompatibility complex (MHC) genes represent functional and expressed adaptive variation shaped by natural selection. MHC genes code for cell surface glycoproteins, which launch the adaptive immune response against pathogens in all jawed vertebrates studied so far. MHC molecules present antigens on cell surfaces, which is followed by antibody production or by destruction of the antigen-presenting cells (reviewed by Ekblom et al. 2007). Selection on MHC genes is believed to be driven by pathogens and/or selective mate choice (Sommer 2005, Piertney and Oliver 2006). The spatial and temporal pattern of functional variation reflects the possibility for a population to adapt to novel conditions (Höglund 2009). In addition, it has recently been shown that MHC variation is more sensitive to bottleneck events, as the loss of MHC polymorphism is 15% greater than that of neutral genetic diversity in case of the rapid reduction of population size (Sutton et al. 2011). Therefore, it is highly important to study both neutral and the adaptive genetic variation to get a full insight into a species pattern of diversity.

The present study focused on neutral and adaptive genetic diversity of black grouse *Tetrao tetrix* in Ukraine. Black grouse is characterized by a wide distribution range – from Britain to eastern Siberia with a northern limit in Norway (70°N) and a southern limit in Kyrgyzstan and North Korea (40°N). It is considered as Least Concern according to the IUCN Red List of Species. Some populations of this species, however, are fragmented and shrinking in size in western and central Europe and are listed in national Red-data books (Storch 2007).

There are three populations of black grouse in Ukraine – in the north, in the west (Rostochia) and in the Carpathian Mountains (Fesenko and Bokotei 2002) (Fig. 1). The northern population is distributed throughout the whole Polissia region (from Volyn to the Chernigiv region). It is considered to be continuous and expanding in central Polissia (Zhytomyr region) (Panov et al. 2002), yet there is evidence that the species is declining in numbers in the northwestern part of Ukraine (Gorban' and Mateichyk 2009). The Carpathian population is believed to be isolated and characterized by a negative population trend (according to the Red-data book of Ukraine, Akimov 2009). The total size of Ukrainian population of black grouse was estimated at roughly 13 000 individuals in 2007 (according to the Red-data book of Ukraine, Akimov 2009).

Our objective was to estimate both neutral and adaptive genetic variation in two Ukrainian populations of black grouse – the northern and the Carpathian one. We compared genetic diversity in a continuous (northern) population connected to the main range of the species with an isolated (Carpathian) one to formulate management suggestions for

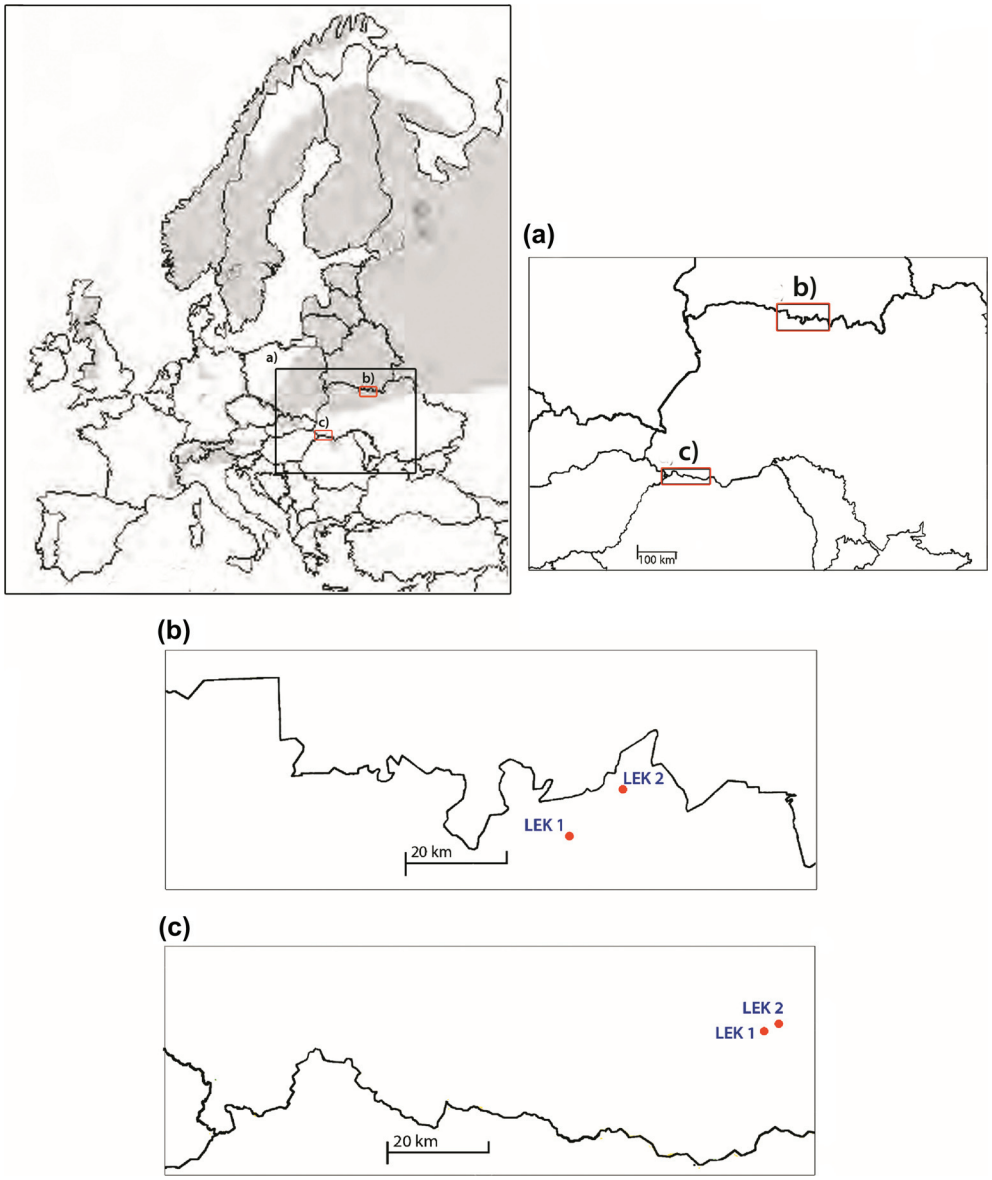


Figure 1. European black grouse distribution (modified from Storch 2007) and areas sampled: (a) Ukrainian black grouse populations sampled; (b) the northern population sampled; (c) the Carpathian population sampled.

the black grouse in the Ukraine. We compared neutral and adaptive genetic diversity of the Ukranian populations with others in Europe.

## Methods

### Material, sampling and DNA extraction

Moulted black grouse feathers were sampled during the lekking period in spring 2011 in two geographic regions in Ukraine – in the Carpathian Mountains (C) and in northern Polissia (N) (Fig. 1). The feathers were collected non-invasively at four lek sites (two in each region), as black grouse hunting is illegal in most regions in Ukraine. All the samples were put individually in paper envelopes, labeled by place and date of collection and stored dry for 3–6 months.

DNA was extracted using a QIAGEN DNeasy tissue kit. The standard manufacturer's protocol was optimized in order to receive higher yield and better quality of DNA. The modifications to the extraction procedure included: 1) extended lysis time (24 h) (Corrales and Höglund 2012), 2) double volume of K-proteinase (40 µl) and ATL-buffer (360 µl), 3) double volume of AL-buffer and ethanol (400 µl of each), 4) adding cold ethanol. The DNA yield and quality was measured using a NanoDrop spectrophotometer and varied between 0.1 ng $\mu l^{-1}$ and 230.2 ng $\mu l^{-1}$. However, the DNA quantity and quality from moulted feathers is usually low and is believed to be hard to measure, therefore we have drawn conclusions about the DNA quality based on the amplification success (Gebhardt et al. 2009). Black grouse feathers have been previously shown to be a reliable source both for microsatellite and MHC variation analyses, and those of comparable quality have been used in studies

together with blood and tissue samples (Strand and Höglund 2011, Strand et al. 2012). If freshly collected, grouse feathers are a reliable source for genotyping (Johansson et al. 2012) and we only used freshly molted feathers in this study.

## Assessment of neutral variation

### *Genotyping*

A QIAGEN Multiplex PCR Kit was used for the amplification of 10 microsatellite loci – ADL230, ADL 257, ADL 142, ADL 184, BG 15, BG 18, TUT1, TUT2, TUT3, TUT4 following previously published protocols (Segelbacher et al. 2000, Piertney and Höglund 2001, Sahlsten et al. 2008, Corrales and Höglund 2012). Briefly, microsatellite primer pairs were organized in three multiplexes according to their annealing temperatures (48°C, 54°C and 60°C). The reaction mix was prepared in 10 µl volume following the standard manufacturer's protocol. PCR amplifications were performed in a BIOER thermal cycler. Negative controls were included to test for false amplifications. In order to avoid contamination DNA extraction, pre-PCR and post-PCR pipetting were done in separate rooms. In order to minimize scoring errors (due to allelic dropout, stuttering and the presence of null alleles) each sample was amplified in three separate PCR reactions, which were replicates of the same extract. PCR products were run on a automatic sequencer and Megabace Genetic Profiler ver. 1.2 was used to score the allele sizes for each locus. Based on these three runs the genotypes were assigned as follows: an individual was considered to be homozygous/heterozygous at a certain locus if this genotype was confirmed in at least two reactions.

### *Data analysis*

Prior to conducting the initial analysis, the data set was examined for scoring errors. The rates of allelic dropout and null allele frequencies for each locus were estimated with the help of Gimlet (Valière 2002) and Genepop 1.2 software (Raymond and Rousset 1995) respectively. The ADL 257 locus was excluded from the further analysis due to high frequency of null alleles in the northern population (Table 1). Some of the other loci showed signs of the presence of null alleles as well, but they were still included in the analysis, as the frequency of the putative scoring error was comparatively low and was present in only one of the two regions.

The data set was cleared of repeated genotypes as all the samples were collected non-invasively. Some of them could belong to the same individual and, therefore, the same individual could have been typed several times. The Gimlet software was applied to find the matching pairs of samples. The samples were considered to be identical if there was at most one mismatch for one allele at one locus (Straka et al. 2012). The data set was also pruned from the point of the amplification success. Samples that did not amplify at more than a half of all loci used were excluded from further analysis. The final data set included 106 unique individuals (75 – from the N population and 31 – from C) (Table 2), genotyped for 9 loci.

Prior to calculating diversity measures, tests for linkage disequilibrium and deviations from Hardy–Weinberg equilibrium were conducted in Genepop 1.2 (Raymond and Rousset 1995). The two regions were examined for evidence of any recent bottleneck events with the help of the Bottleneck 1.2.02 software (Cornuet and Luikart 1997). The TPM mutation model with 88% step-wise mutations (Ps = 0.88) and a variance of 10% was used to test for recent bottlenecks, as recommended by Straka et al. (2012).

There is evidence that Bottleneck 1.2.02 can fail to predict bottleneck events if the effective size of the reduced population is too small and a new mutation-drift equilibrium is set (Larsson et al. 2008). Therefore, the MRatio approach was also used to detect possible recent reduction of population sizes (M is the mean ratio of the number of alleles to the range in allele size). The rationale of this method is that the total number of alleles ($k$) is reduced faster than the range in allele size ($r$) when a population experiences a bottleneck event (Garza and Williamson 2001). Thus, the M is expected to be lower in recently reduced populations than in those that are in equilibrium. In general, it is claimed that M is greater than 0.82 in populations that have not experienced a recent bottleneck event, and is lower than 0.7 in groups that suffered a rapid reduction in size (Garza and Williamson 2001).

In order to test for the recent bottleneck in the northern and in the Carpathian populations of black grouse, M was calculated in M_Pval program (Garza and Williamson 2001). Then an equilibrium M (Meq) and the Critical M (Mc) were estimated in CriticalM program with the proportion of one-step mutations Ps = 0.88, the average size of non one-step mutations $\Delta g = 2.8$, and with 10 000 simulation (Garza and Williamson 2001). The value of $\theta = 4$ Neµ was varied over three values – 0.1, 1 and 10 as recommended by Straka et al. (2012) in order to account for the various possible effective population size prior to the bottleneck event. The M value was then compared to the threshold values and the Meq was aligned to the Mc to draw conclusions about the recent reduction in the population sizes.

We calculated observed and expected heterozygosities ($H_o$ and $H_e$) and the number of alleles per each locus using Microsatellite Toolkit (Park 2001). Heterozygosity expected at Hardy–Weinberg equilibrium ($H_e$) was calculated according to Nei (1972). Allelic richness (AR) and Wright's inbreeding coefficient ($F_{is}$) (Wright 1943) were calculated for each locus using FSTAT 2.9.3.2 (Goudet 1995). AR is a mea-

Table 1. Number of samples (N), inbreeding coefficients (Fis) and microsatellite diversity statistics ($H_o$ – observed heterozygosity, $H_e$ – expected heterozygosity and AR – allelic richness) calculated separately for populations from 4 leks.

| Region | n | Lek | Fis | 95% CI | He | He SD | Ho | Ho SD | AR |
|---|---|---|---|---|---|---|---|---|---|
| Northern | 67 | Lek 1 | −0.082 | (−0.166, −0.036) | 0.690 | 0.030 | 0.746 | 0.019 | 3.140 |
| Northern | 8 | Lek 2 | −0.065 | (−0.363, −0.065) | 0.771 | 0.063 | 0.814 | 0.057 | 3.467 |
| Carpathian | 16 | Lek 1 | 0.151 | (−0.021, 0.262) | 0.855 | 0.044 | 0.735 | 0.043 | 4.455 |
| Carpathian | 15 | Lek 2 | −0.102 | (−0.348, −0.074) | 0.722 | 0.036 | 0.801 | 0.038 | 3.216 |

Table 2. MHC diversity calculated for the two Ukrainian populations of black grouse.

| Population | n | No. of alleles | No. of distinct alleles/ total no. of alleles | No. of alleles/ sample size | No. of alleles/ind | No. of private alleles | $\hat{H}$ | Theta k | 95% confidence interval | π | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Northern | 36 | 16 | 0.17 | 0.44 | 2.19 | 8 | 0.85 | 5.35 | [3.00. 9.18] | 14.10 | 7.07 |
| Carpathian | 20 | 9 | 0.32 | 0.45 | 1.2 | 1 | 0.83 | 4.19 | [1.88. 9.01] | 12.51 | 6.48 |

sure of the number of alleles in the population weighted by sample size. $F_{is}$ measures the standardized deviation of observed from expected heterozygosity which can be caused by mating among close relatives and/or hidden population structure. An ANOVA was performed in the R software (<www.r-projekt.org/>) in order to estimate the significance of the differences in diversity measures between the two populations.

The structure and differentiation of the populations was examined in ARLEQUIN (Excoffier et al. 2005) by calculating the global $F_{st}$ and conducting multi-locus hierarchical analysis of molecular variance (AMOVA). Multilocus $F_{st}$ ($G_{st}$) estimates the extent of population subdivision (Weir and Cockerham 1984). It has been argued that $F_{st}$ and $G_{st}$ indexes of differentiation can be unreliable and imprecise, when the diversity of the studied populations (the number of unique alleles) is high and gene flow is significant (Jost 2008, Leng and Zhang 2011). In order to control for such an effect, the Dest index was also used (Jost 2008). Dest was calculated in the SPADE software (Chao and Shen 2010). Population structure was visualized with GenAlEx 6.5 (Peakall and Smouse 2012) by performing a factorial correspondence analysis (FCA).

We also studied population structure of Ukrainian black grouse using Bayesian clustering implemented in the STRUCTURE software. The program finds population structure under the assumptions of Hardy–Weinberg equilibrium and linkage equilibrium within populations and searches for the groups of individuals that are not in disequilibrium (Pritchard 2000, Falush et al. 2003). Both admixture and non-admixture models were tested and the admixture model was chosen as the more realistic one. The parameters used were: undefined population structure, correlated allele frequencies among populations, length of burning period = 10 000, number of MCMC replicates after burn-in = 50 000 (Evanno et al. 2005). We tested 10 possible numbers of groups (K = 1–10), each trial was run for 20 times. Structure harvester was used to visualize the STRUCTURE output and to find the most probable ΔK. ΔK was used instead of L(K) value, as it is more reliable and accurate according to Evanno et al. (2005). Finally, the results for the most probable ΔK were generalized in CLUMPP (Jakobsson and Rosenberg 2007) and visualized in DISTRUCT software (Rosenberg 2004).

## Adaptive genetic diversity assessment using RSCA

Reference strand-mediated conformational analysis (RSCA) was applied to genotype the birds at two paralogous MHC Class IIB (BLB) loci, which were recently shown to be expressed in black grouse (Strand et al. 2013). MHC BLB loci are characterized by high dN/dS ratio, are highly polymorphic and under the influence of balancing selection (Strand et al. 2007, 2013, Wang et al. 2012). Therefore sampling these particular markers should give a considerable insight into the pattern of genetic diversity under balancing selection.

The main advantage of RSCA in this study was its proven effectiveness both for amplification of DNA from both blood and tissue samples and from molted feathers (Strand and Höglund 2011).

We typed 56 samples (36 from the northern population and 20 from the Carpathian Mountains) with RSCA. The main criterion for including a sample was their amplification success during the microsatellite genotyping - most samples included amplified in at least 7 loci out of 9.

We followed the protocol, developed for the black grouse by Strand and Höglund (2011), using the primer pair (RNA F1a (5′GACAGCGAAGTGGGGAAATA3′) and RNA R1a (5′CGCTCCTCTGCACCGTGA3′)). This yields the PCR-products starting at the 108th base pair in exon 2 and stopping at the 270th (the last) bp of exon 2, thus covering 46% of the exon and about 2/3 of the peptide binding sites) (Strand et al. 2012). All the PCR products were confirmed on agarose gels. The RSCA was run on an automatic sequencer and afterwards alleles were scored in the Megabace Fragment Profiler ver. 1.2.

Our RSCA provided information from multiple loci as PCR products from paralogous genes were amplified simultaneously. So, the data on adaptive genetic diversity in this study is based on frequency estimates of the alleles and not on the diploid haplotypes (unlike the data on microsatellite diversity). This is a consequence of the fact that both MHC BLB loci in black grouse are subjected to duplication and homogenizing concerted evolution through intergenetic exchange between the BLB1 and the BLB2 loci (Strand et al. 2013).

## Data analysis and sequencing

The resulting MHC alleles were compared to those listed in the RSCA library constructed by Strand and Höglund (2011) following the same scoring rules. An individual allele was confirmed if it was detected in three out of four fluorescent labelled reference strands (FLRs) (for strong single peaks) or in four out of four FLRs (for weaker peaks). PCR artifacts are important to consider in MHC studies. In order to decrease the probability of PCR artifacts only alleles present in two independent PCR reactions were regarded as confirmed and included into subsequent analysis (Strand and Höglund 2011, Strand et al. 2012). All samples were genotyped twice to receive higher resolution and precision. Three new alleles were detected, which had not been previously included in the RSCA library. In order to confirm these unique variants, the three new alleles were cloned and sequenced.

Four individuals were chosen, which had the highest amplification success with both microsatellites and the

RSCA markers and showed these new alleles in two independent PCR reactions. PCR products were cloned using the TOPO TA cloning kit, 9–20 positive clones per individual were picked and send for sequencing.

The sequences were analyzed in Codon Code Aligner ver. 3.7.1 and compared to the alleles listed in the RSCA library. The reference sequences were downloaded from the GenBank (Benson et al. 2005). If the analyzed sequence included one or more unique motifs and was present in at least three clones, it was scored as the new one. It is needed to estimate the new alleles' RSCA score for contributing to the existent RSCA library. Yet, this is beyond the scope of our study and should be planned for the subsequent research.

The diversity measures calculated for the MHC data were: the total number of alleles per population, the number of alleles per sample size, the number of distinct alleles per total number of alleles found in a population, mean number of alleles per individual, the number of private alleles in each population, theta k – the index of allelic richness (estimated from infinite-allele equilibrium between expected number of alleles, the sample size and theta parameter), $\pi$ – nucleotide diversity (mean number of pairwise differences between all pairs of alleles in a population) and $\hat{H}$ – gene diversity (Nei 1987). Theta k and $\pi$ were calculated in ARLEQUIN (Excoffier et al. 2005). $\hat{H}$ was computed in Microsoft Excel using the following formula:

$$\hat{H} = \frac{n}{n-1}\left(1 - \sum_{i=1}^{k} p_i^2\right)$$

where $n$ is the total number of gene copies in the sample, $k$ is the number of haplotypes and $p_i$ is the frequency of the ith haplotype (Nei 1987). Since no haplotype could be discerned in the RSCA, $k$ was considered to be equal to the number of alleles in each population. Thus, the allele frequency was calculated as occurrence (number) of certain allele/total occurrence (number) of all alleles in a population. The differentiation in adaptive diversity of the Carpathian and the northern population was estimated using $F_{st}$, calculated in ARLEQUIN (Excoffier et al. 2005) and Dest, calculated in SPADE (Chao and Shen 2010).

## Results

### Neutral genetic diversity and differentiation

Signs of linkage disequilibrium (LD) were present among ADL 230 and ADL 142. In order to test, whether this could influence the estimation of the standard diversity measures and population differentiation parameters, separate tests were conducted, including and excluding these loci from the analysis. As the results did not differ, putative LD between those loci did not influence the estimation of the population neutral diversity measures. Moreover, it has been previously shown that ADL 230 and ADL 142 are not physically linked (Groenen et al 2000, Sahlsten et al. 2008). Thus, the signs of linkage disequilibrium between them could rather be attributed to specific population structure.

The rates of allelic dropout were estimated in the Gimlet software and were 0.000 for all loci except for ADL 184, BG 15 and TUT 1, for which it equaled 0.093, 0.060 and 0.080 respectively.

Both N and C showed signs of deviation from Hardy–Weinberg equilibrium at certain loci. Yet, globally, the N population did not deviate (Fis = –0.045, 95% CI (–0.107; -0.002)), whereas the C did (Fis = 0.111, 95% CI (0.001; 0.165)) (Supplementary material Appendix 1). The two populations did not deviate from the expected L-shaped distribution and displayed a high probability for heterozygosity deficiency ($p = 0.00293$ for N and $p = 0.00488$ for C), suggesting they did not experience any severe bottlenecks in the past.

The MRatio estimated from the number of alleles and their size range was 1.47 for N and 1.03 for C. Both values are higher than the threshold 0.7 given in Garza and Williamson (2001). The Meq was higher than the Mc in both populations, in all the replicates, regardless of the different values for $\theta$. This supports the results of the Bottleneck and indicates that both the northern and Carpathian populations are unlikely to have experienced any recent reduction in population size.

Both observed and expected heterozygosities were higher for C than for N ($0.75 \pm 0.02/0.71 \pm 0.03$ and $0.87 \pm 0.01$ $/0.78 \pm 0.03$, Ho SD / He $\pm$ SD, respectively). This difference was even more pronounced and highly significant in term of Allelic richness (9.18 in C compared to 6.44 in N, ANOVA, $F = 9.81$, $p = 0.0064$). Additionally, Wilcoxon tests revealed the studied populations were significantly different for allelic richness ($p = 0.004$), but not for heterozygosity. The difference in inbreeding coefficients was significant ($p = 0.004$), which could possibly be attributed to a Wahlund effect in the Carpathian sample.

An AMOVA showed that the most significant variation was within populations (84.6%) but the amount of variation among N and C (15.4%) was significant. The significant variation among populations was supported by the global microsatellite fixation index $F_{st}$, which was estimated at 0.148 and was significant ($p < 0.001$ after 110 permutations). Dest 0.662 with a standard error of 0.024 after 1000 bootstrap replications. An FCA plot for the first two factors showed that Ukrainian black grouse population was differentiated into three separate clusters: a northern one and two groups from the Carpathian Mountains (Fig. 2).

Analyses with STRUCTURE illustrated the presence of both global and regional differentiation in two populations (Fig. 3). The highest change in likelihood was for the presence of four clusters ($\Delta K = 35.745$), whereas two clusters also had high change in likelihood ($\Delta K = 30.423$) (Supplementary material Appendix 1). Therefore, we present both variants. When compared to the geographical origin of the samples, the two clusters coincided with the data from the Carpathian and the northern region; and the four clusters, to a lesser extent, corresponded to the different leks in each region. The two clusters approach illustrated that there were some individuals from N assigned to C. However, the higher resolution with four clusters indicated that these individuals belong to one of the leks in N, and not in C region.

Taking into account the presence of regional structure in the northern and in the Carpathian populations of black grouse, additional analyses were conducted and neutral diversity indexes for each lek sampled were calculated separately

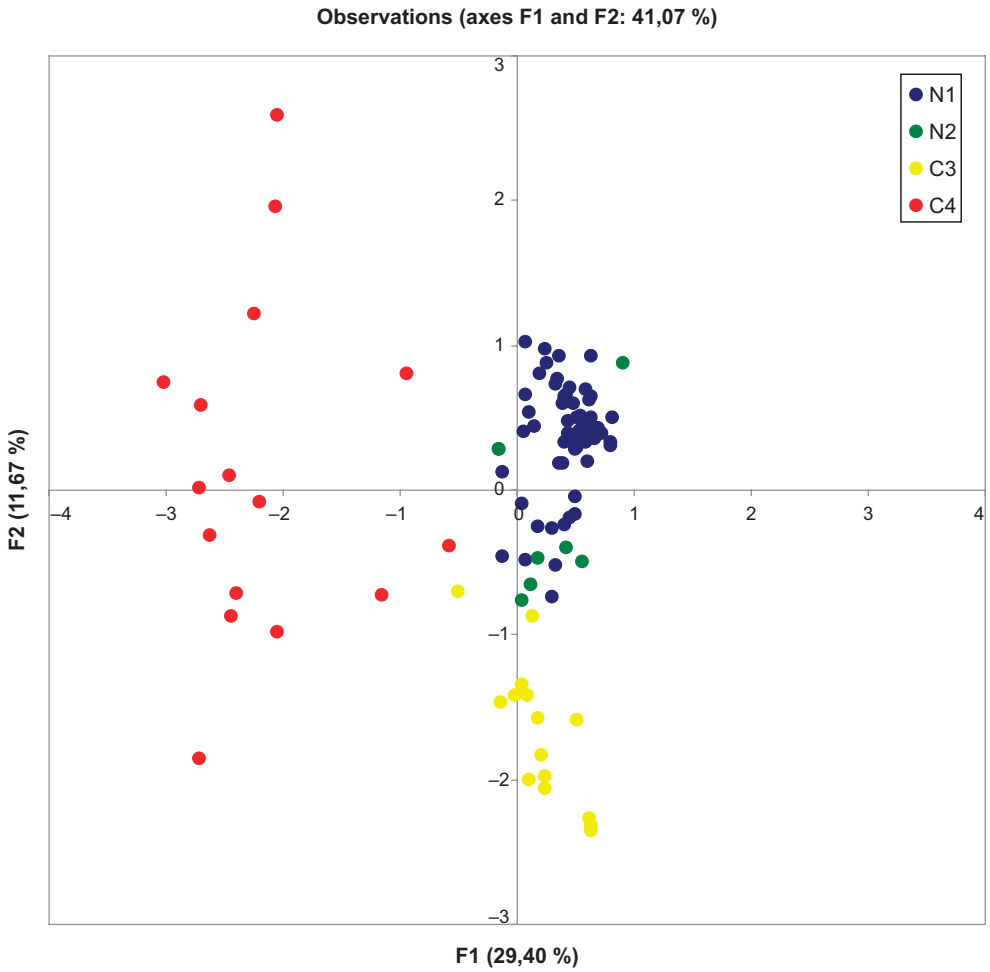


Figure 2. Two-factorial correspondence analysis for all genotyped microsatellites in two populations of black grouse in Ukraine (red and yellow dots – Carpathian population, green and blue dots – northern population). The first axis on the plot explains 29.4% of the total variation and the second axes – 11.7%.

(Table 1). The $F_{st}$-values calculated between Lek 1 and Lek 2 from C ($F_{st} = 0.227$) are higher than $F_{st}$-values among 2 leks sampled in N ($F_{st} = 0.078$). Similarly, the Dest index of differentiation between the two C leks (Dest = 0.871, CI (0.815, 0.927) is higher than between the two N leks (Dest = 0.366, CI (0.235, 0.497).

Only one (Lek 1) of the two C leks sampled was characterized by a positive Fis value (Table 3). Yet, the confidence interval for this value is wide. The only significant difference between leks was found for allelic richness ($F = 4.2072$, $p = 0.014$). Subsequent Tukey HSD tests revealed that two pairs of populations: Lek 1 (N) and Lek 3 (C) ($p = 0.019$),

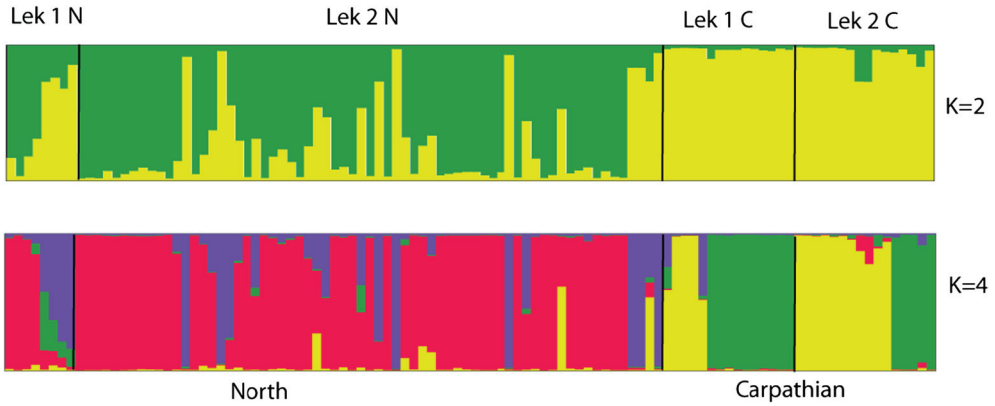


Figure 3. Bayesian clustering for the northern (N) and the Carpathian (C) populations of black grouse assuming two and four clusters.

Table 3. p-values from Welch two-sample t-tests for significant differences in microsatellite allelic richness among European populations of black grouse. Above the diagonal t-values, p-values below the diagonal.

| | Present Europe | Historical Europe | UA_ North | UA_ Carpathian |
|---|---|---|---|---|
| Present Europe | | 1.1913 | –3.1369 | –7.9805 |
| Historical Europe | 0.2477 | | –2.7329 | –7.5644 |
| UA_North | 0.02799 | 0.04614 | | 2.4524 |
| UA_Carpathian | 0.0003719 | 0.0006857 | 0.04238 | |

and Lek 3 (C) and Lek 4 (C) ($p = 0.03$) significantly differed for AR.

## Adaptive genetic diversity and differentiation

The total number of BLB-MHC alleles found in N was 16 and for C we found 9 (Table 1). The N population harbored eight private alleles – Tete-BLB*03, Tete-BLB*04, Tete-BLB*06, Tete-BLB*09, Tete-BLB*11, Tete-BLB*21, Tete-BLB*22 and Tete-BLB*29, whereas the C population had one private allele – BLB25. Yet, we cannot rule out that this private allele distribution can be attributed to the limited sample size in the Carpathian region. The two populations varied in MHC diversity patterns: the N population was more diverse than the C in terms of gene diversity ($\hat{H}$), the number of alleles, allelic richness, alleles/individual and nucleotide diversity. The Carpathian population, on the other hand, revealed a higher number of distinct alleles/total number of alleles found (Table 2).

We found three previously undescribed alleles – Tete-BLB*35, Tete-BLB*36 and Tete-BLB*37 (Supplementary material Appendix 1). Tete-BLB*35 and Tete-BLB*36 were shared among the two regions, but Tete-BLB*37 was found in the N population only. The N population was more diverse than C, with higher gene diversity ($\hat{H}$) and more alleles in total. However, the C population had more alleles/sample size and more distinct alleles/total number of alleles found in a population (Table 2). The global $F_{st}$ was 0.046 but not significantly different from zero after 1023 permutations ($p = 0.054$), though close to significance. The Dest was 0.454 with the standard error of 0.132 and the 95% confidence interval (0.183, 0.694).

## Comparative analysis of genetic diversity and differentiation in Ukrainian and other European populations

The levels of heterozygosity, allelic richness, nucleotide diversity and MHC allelic richness in both Ukrainian populations of black grouse corresponded to those previously found in Finland (Jyväskylä), Norway (Kristiansand), Sweden (Jämtland), Latvia and Austria (Alps) (data from Strand and Höglund (2011) and Segelbacher et al. (unpubl.)). Many of these samples have been scored in the same lab using the same equipment as the present and if not these were standardized typing a few reference samples on the different machine. While we can not completely rule out errors induced by different runs and scorers, we tried to keep these differences at a minimum. Both of the studied Ukrainian populations had high levels of allelic richness and were characterized by high heterozygosities compared to the other European populations studied in Segelbacher et al. (unpubl.) (Fig. 4, Table 3).

A similar pattern was identified by ANOVA – the N and C populations were significantly different from the other European populations of black grouse for allelic richness, but not for the other parameters of neutral genetic diversity. This was consistent both when the two Ukrainian populations were pooled and compared to the pooled European populations ($F = 31.655$, $p < 0.0001$), and when all the populations were compared as separate units ($F = 19.703$, $p < 0.0001$) (Fig. 4).

All pairwise $F_{st}$ indexes were significant ($p < 0.001$ after 110 permutations) and ranging from 0.027 between the historical German and Dutch populations and 0.492 between the northern Ukrainian and the historical Danish populations (Table 4). The pattern of high genetic differentiation between the Black grouse populations in Ukraine

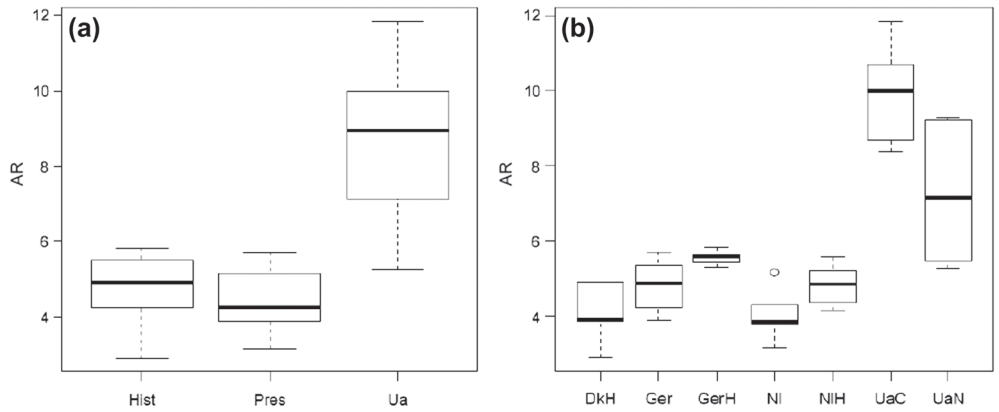


Figure 4. The differences in microsatellite allelic richness of European populations of black grouse: (a) the populations pooled in three categories – the present European (Pres), the historic European (Hist) and Ukrainian (Ua): (b) the populations considered as separate units: Danish historical (DkH), German present (Ger), German historical (GerH), Dutch present (Nl), Dutch historical (NlH), Carpathian Ukrainian (UaC) and northern Ukrainian (UaN). Boxes represent median (thick line through boxes) and 75% (upper and lower end of boxes) and 95% (upper and lower hinges) interquartile ranges. Circles are outside values.

Table 4. Neutral genetic differentiation between European population of black grouse (Dk_H – Danish historical, LH - German present, Ger_H - German historical, SH - Dutch present, Nl_H - Dutch historical, C - Carpathian Ukrainian and N - northern Ukrainian ($F_{st}$ values are given below the diagonal and Dest – above the diagonal; Significant $F_{st}$ and Dest for Ukrainian black grouse populations are given in bold).

| | Dk_H | LH | Ger_H | SH | Nl_H | N | C |
|---|---|---|---|---|---|---|---|
| Dk_H | | 0.299 | 0.244 | 0.545 | 0.391 | **0.985** | **0.969** |
| LH | 0.136 | | 0.213 | 0.290 | 0.229 | **0.967** | **0.947** |
| Ger_H | 0.088 | 0.067 | | 0.399 | 0.115 | **0.985** | **0.955** |
| SH | 0.309 | 0.135 | 0.158 | | 0.353 | **0.950** | **0.921** |
| Nl_H | 0.172 | 0.082 | 0.027 | 0.152 | | **0.983** | **0.968** |
| N | **0.492** | **0.377** | **0.328** | **0.411** | **0.353** | | 0.613 |
| C | **0.295** | **0.192** | **0.142** | **0.231** | **0.167** | 0.098 | |

and in the other European countries appeared to be even more pronounced, when pairwise Jost Dest was estimated. The Dest values between the Ukrainian populations and the European ones exceeded the Dest between both contemporary and historic European populations of black grouse. The lowest differentiation was again between the two historical European populations (0.115) and the highest – between the historic German and Danish populations and the northern Ukrainian population (0.985).

## Discussion

This study was aimed at discovering the patterns of neutral and adaptive genetic diversity of Ukrainian black grouse. Unbiased and standardized data from the entire the distribution of endangered species is important for efficient allocation of management efforts. However, the challenge in getting this type of data from eastern European non-EU countries (e.g. Ukraine, Belarus, Georgia and the European part of Russia) is still daunting because of a lack of research infrastructure and resources in this region. As eastern European countries are part of the same zoogeographical region as the rest of Europe (Holt et al. 2013) and share continuous populations, it is urgent to fill the gap in the data for many species, thereby facilitating comparative analysis of European species' distributions and diversity.

Microsatellite diversity in the Carpathian population was higher than in the northern one. The two regions were not significantly different in terms of the standard measures of MHC diversity (π, theta k), except for the total number of MHC alleles and the number of distinct MHC alleles adjusted by the total number of alleles found in a population. We detected twice as many alleles for the northern population, than for the Carpathian one. Yet, adjusting this measure to the sample size and to the total number of alleles found in a population yielded a different result, illustrating that the black grouse from the Carpathian Mountains could be, in fact, more diverse than those from the north.

A possible explanation for this pattern may be that the Carpathian region is closer to the south, which was free from the ice during the last glaciations period (20 000 YA), and could have been a refuge for the black grouse. The Carpathian basin has been shown to be a refuge for several species including European viper *Vipera berus* and moor frog *Rana arvalis*, and it is listed as an example of a core area for the distribution of continental species (Schmitt 2007). Similarly, the brown bear *Ursus arctos* is considered to have spread to the north from both an Iberian and a Carpathian/Caucasian refuge (Hewitt 2000). The role of the Carpathian basin as a source of the post-glaciation colonization of northern Europe is thought to be underestimated due to the lack of reliable genetic data from this region (Hewitt 2000).

If the Carpathian population of black grouse is the source for the populations further north, the comparatively low diversity in the latter could be explained by an expansion-contraction model of dispersal (Taberlet et al. 1998). According to this model the colonization process is characterized by a number of extinction events in the northern populations due to decreased temperatures and a series of the northwards expansions of the species range during the warm periods. The model implies population bottlenecks occurred that would reduce genetic diversity in the northern populations (Taberlet et al. 1998). This process might be especially significant in cases of the species with specific habitat requirements (Schmitt 2007), such as the black grouse. However, testing of the refuge hypothesis and distinguishing between expansion–contraction model and founder effect requires additional research including mtDNA data.

There was no evidence for recent bottleneck events either in the Carpathian or the northern populations. This result is reasonable for the northern population as it is consistent with the fact that it did not deviate from Hardy–Weinberg equilibrium. However, the Carpathian population appeared to deviate from Hardy–Weinberg equilibrium (Fis = 0.111), which could in itself be a sign of small population size. In this case the Bottleneck program might have failed to indicate the recent bottleneck event, as a new mutation-drift equilibrium may have been set. A similar scenario has been described for the Dutch black grouse population (Larsson et al. 2008). However, the fact that the M_P_val and Critical M programs did not find any signs of recent size reductions indicates that the deviation from Hardy–Weinberg equilibrium could have been caused by some other reason, such as scoring errors (the presence of null alleles, allelic dropout). Another reason for deviation from HW equilibrium observed for Carpathian population could be a Wahlund effect (Selkoe and Toonen 2006).

The analysis using the STRUCTURE software revealed both regional and the local differentiation of the Ukrainian populations of black grouse. The total population was first separated into two large clusters – the northern one and the Carpathian one, both of which were further split into two groups. The Carpathian population was more clearly differentiated than the northern one, as the genotypes from the two clusters did not show signs of admixture. Moreover, the two leks in Carpathian region were shown to be highly differentiated (both in $F_{st}$ and Dest), and significantly different in their allelic richness from each another. This supports the above discussed idea about the higher differentiation and more significant isolation of black grouse subpopulations in the Carpathian Mountains, and could have contributed to the diversity pattern observed. Moreover, such structure illustrates that the northern and the Carpathian populations differ in their genetic composition and should both be the subjected to conservation measures if the goal is to preserve the highest possible genetic diversity of Ukrainian black grouse populations.

Both the $F_{st}$ and the Dest indexes supported genetic subdivision in neutral diversity in the two populations studied and weak gene flow between them. High values of both these indexes may reflect strong and early differentiation between the northern and Carpathian populations (Leng and Zhang 2011). Yet, investigating the patterns of adaptive diversity and differentiation (such as for MHC) may be more important for conservation than studying the neutral variation, as the latter does not necessarily reflect the local adaptation (Miller et al. 2010, Ekblom et al. 2007). Microsatellite diversity has been argued to be a poor surrogate for MHC variation, as their mutation mechanisms differ to a large extent (Miller et al. 2010).

Similar patterns of low MHC structure compared to the neutral markers has been shown for red grouse *Lagopus lagopus scoticus* (Piertney 2003), Trinidadian guppy *Poecilia reticulate* (Fraser et al. 2009) and Chinook salmon *Oncorchynchus tshawytscha* (Evans et al. 2010). The results of the present study are in line with those reported for the European black grouse in which higher differentiation in neutral diversity than in MHC was found (Strand et al. 2012). It has been claimed that selection for functionally similar MHC-alleles has been uniform in black grouse across Europe, which can be the outcome of the previous wider distribution of this species and a consequent higher gene flow between the populations (Strand et al. 2012). Studies on other species have described opposite pattern of population structure: more differentiation in MHC diversity than for neutral markers – great snipe *Gallinago media* (Ekblom et al. 2007), Atlantic salmon *Salmo salar* (Landry and Bernatchez 2001) and *Tuatara sphenodon* spp. (Miller et al. 2010).

t should be stressed that there is some structure in adaptive variation between the northern and the Carpathian population, even though MHC differentiation is less than neutral differentiation. This is implicated by the high Dest value for the MHC and by the fact that there were private alleles discovered in both populations. This pattern can potentially indicate some local shifting in selection pressures in the two regions (Piertney 2003). However, in order to claim that there are significant differences in the allelic composition of the MHC of the northern and the Carpathian black grouse populations, larger sample sizes need to be examined. The differences in sample size between the two populations could have resulted in an underestimation of allelic diversity in the Carpathian black grouse. Making any conclusions about local adaptations based on our data only would be speculative and requires further analysis and more sophisticated information on local selection pressures. Yet, such data on MHC and neutral population diversity and structure allows for population genetic studies.

Comparative analysis between Ukrainian and other European populations of black grouse illustrated that both the Carpathian and the northern population resembled large continuous European populations in terms of diversity. This outcome seems to be reasonable for the northern population, but it is unexpected for the Carpathian one, as the topography of Carpathian Mountains should result in isolation and restricted gene flow. Yet, the fact that the population from Alps is also genetically diverse, shows that this can be the case for the other mountain populations. The Carpathian population may have the same (or even higher) genetic diversity than large continuous European populations of black grouse because it may have been a refuge during the last glaciation period.

We have shown that both Ukrainian populations are highly differentiated from the other European populations. Historical populations from Germany and the Netherlands (sampled during the period of 1887–1968) had $F_{st}$ values that were more similar to Ukrainian populations' than were the contemporary. This is reasonable, as the historical populations used for comparison were characterized with higher genetic diversity than the present ones (Segelbacher et al. unpubl.), and, therefore, are expected to be more similar to the northern and the Carpathian population. The Dest values did not support this, yet the proportional differences in Dest between Ukrainian and European historical and Ukrainian and European contemporary populations were negligible. In any case, it is clear so far that there is strong and longstanding differentiation between the Ukrainian and the European populations of black grouse.

## Conclusions

This is the first study examining genetic diversity of black grouse in Ukraine. Our work provides insight into both adaptive and neutral genetic diversity of the northern and Carpathian populations of black grouse, and suggests that these regions should be treated as two separate management units. Future research of the population of black grouse in the north of Ukraine (Polissia region) should focus on more extensive sampling of this area, including its western and eastern boarders, as the diversity pattern can be different there. The Carpathian population requires more sampling effort as well, and a larger number of isolated leks should be included in the subsequent analysis to capture the whole diversity pattern of this region.

*Acknowledgements* – We are grateful to Tanja Strand, Eleanor Jones, Yvonne Meyer-Lucht, Leanne Faulks, Gunilla Engström, Reija Dufva, Biao Wang and Rado Kozma for invaluable help during all stages of the project. Special thanks go to Ievgen Dyky, Oleksandr Zakletskyi, Andrii Bokotei and Oleksandr Bokotei for getting the samples from Ukraine and to Ievgen Olefirenko for his encouragement and support.

## References

Akimov, I. A. 2009. Red-data book of Ukraine. Fauna Edited. – Globalconsulting, Kyiv, Ukraine.

Benson, D. A. et al. 2005. GenBank. – Nucleic Acids Res. 33: D34–D38.

Chao, A. and Shen, T. –J. 2010. Program SPADE. – <http://chao.stat.nthu.edu.tw>.

Cornuet, J. M. and Luikart, G. 1997. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. – Genetics 144: 2001–2014.

Corrales, C. and Höglund, J. 2012. Maintenance of gene flow by female-biased dispersal of black grouse *Tetrao tetrix* in northern Sweden. – J. Ornithol. 153: 1127–1139.

Ekblom, R. et al. 2007. Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). – Mol. Ecol. 16: 1439–1451.

Evanno, G. et al. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. – Mol. Ecol. 14: 2611–2620.

Evans, M. L. et al. 2010. MHC genetic structure and divergence across populations of Chinook salmon (*Oncorhynchus tshawytscha*). – Heredity 104: 449–459.
Excoffier, L. et al. 2005. Arlequin (ver. 3.0): an integrated software package for population genetics data analysis. – Evol. Bioinf. 1: 47–50.
Falush, D. et al. 2003. Inference of populations structure using multlocus genotype data: linked loci and correlated allele frequencis. – Genetics 164: 1567–1587.
Fraser, B. A. et al. 2009. Selection at the MHC class IIB locus across guppy *(Poecilia reticulata)* populations. – Heredity 104: 155–167.
Fesenko, G. and Bokotei, A. 2002. Birds of Ukrainian fauna: field key book. – Kyiv.
Garza, J. C. and Williamson, E. G. 2001. Detection of reduction in population size using data from microsatellite loci. – Mol. Ecol. 10: 305–318.
Gebhardt, K. et al. 2009. Molted feathers from clay licks in Peru provide DNA for three large macaws (*Ara ararauna*, *A. chloropterus*, and *A. macao*). – J. Field Ornithol. 80: 183–192.
Gorban', I. and Mateichyk, V. 2009. Nesting birds of the peat bogs of Shazkii National park in 2009. – Proc. Conf. "State and biodiversity of the ecosystems of Shazkii National Park.
Goudet J. 1995. fstat ver. 1.2: a computer program to calculate F-statistics. – J. Hered. 86: 485–486.
Groenen, M. A. et al. 2000. A consensus linkage map of the chicken genome. – Genome Resour. 10: 137–147.
Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. – Nature 405: 907–913.
Höglund, J. 2009. Evolutionary conservation genetics. – Oxford Univ. Press.
Holt, B. G. et al. 2013. An update to Wallace's Zoogeographic Regions of the World. – Science 339: 74–77.
Jakobsson, M. and Rosenberg, N. A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. – Bioinformation 23: 1801–1806.
Jehle, R. and Arntzen, J. W. 2002. Microsatellite markers in amphibian conservation genetics: a review. – Herpetol. J. 12: 1–9.
Johansson, M. et al. 2012. Amplification success of multilocus genotypes from feathers left for varying times in the field. – Ibis 154: 15–20.
Jost, L. 2008. Gst and its relatives do not measure differentiation. – Mol. Ecol. 17: 4015–4026.
Landry, C. and Bernatchez, L. 2001. Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). – Mol. Ecol. 10: 2525–2539.
Larsson, J. K. et al. 2008. Genetic impoverishment of the last black grouse (*Tetrao tetrix*) population in the Netherlands: detectable only with a reference from the past. – Mol. Ecol. 17: 1897–1904.
Leng, L. and Zhang, D. X. 2011. Measuring population differentiation using GST or D? A simulation study with microsatellite DNA markers under a finite island model and nonequilibrium conditions. – Mol. Ecol. 20: 2494–2509.
Miller, H. et al. 2010. Genetic diversity and differentiation at MHC genes in island populations of tuatara (*Sphenodon spp.*). – Mol. Ecol. 19: 3894–3908.
Moss, R. et al. 2003. The use and abuse of microsatellite DNA markers in conservation biology. – Wildl. Biol. 9: 243–250.
Nei, M. 1972. Genetic distance between populations. – Am. Nat. 106: 283–292.
Nei, M. 1987. Molecular evolutionary genetics. – Columbia Univ. Press.
Oliveira, E. J. et al. 2006. Origin, evolution and genome distribution of microsatellites. – Genet. Mol. Biol. 29: 294–307.
Panov, G. et al. 2002. Black grouse number and possible reasons of its changes in the northern part of Zhytomyr region. – Berkut 11: 173–180.
Park, S. D. E. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. – PhD thesis, Univ. of Dublin.
Peakall, R. and Smouse, P. E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. – Bioinformatics 28: 2537–2539.
Piertney, S. B. 2003. Major histocompatibility complex B-LB gene variation in red grouse *Lagopus Lagopus scoticus*. – Wildl. Biol. 9: 251–259.
Piertney, S. B. and Höglund, J. 2001. Polymorphic microsatellite DNA markers in black grouse (*Tetrao tetrix*). – Mol. Ecol. Notes 1: 303–304.
Piertney, S. B. and Oliver, M.-K. 2006. The evolutionary ecology of the Major Histocompatability Complex. – Heredity 96: 7–21.
Pritchard, J. K. 2000. Inference of population structure using multilocus genotype data. – Genetics 155: 945–959.
Raymond, M. and Rousset, F. 1995. GENEPOP (ver. 1.2): population genetics software for exact tests and ecumenicism. – J. Hered. 86: 248–249.
Rosenberg, N. A. 2004. Distruct: a program for the graphical display of population structure. – Mol. Ecol. Notes 4: 137–138.
Sahlsten, J. et al. 2008. Inference of hazel grouse population structure using multilocus data: a landscape genetic approach. – Heredity 101: 475–482.
Schmitt, T. 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. –Front. Zool. 4: 11.
Segelbacher, G. et al. 2000. Characterisation of microsatellites in capercaillie (*Tetrao urogallus*) (AVES). – Mol. Ecol. 9: 1934–1935.
Selkoe, K. A. and Toonen, R. J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. – Ecol. Lett. 9: 615–629.
Sommer, S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. – Front. Zool. 2: 16.
Storch, I. 2007. Grouse status survey and conservation action plan 2006–2010. WPA/BirdLife/SSC Grouse Specialist Group. – IUCN, Gland, Switzerland and Cambridge, UK and the World Pheasant Association, Reading, UK.
Straka, M. et al. 2012. Microsatellite diversity and structure of Carpathian brown bears (*Ursus arctos*): consequences of human caused fragmentation. – Conserv. Genet. 13: 153–164.
Strand, T. and Höglund, J. 2011. Genotyping of black grouse MHC class II B using reference strand-mediated conformational analysis (RSCA). – BMC Res. Notes 4: 183.
Strand, T. et al. 2007. Diversity and variable expression of Mhc class II B and Y genes in black grouse (*Tetrao tetrix*). – Immunogenetics 59: 725–734.
Strand, T. et al. 2012. Can balancing selection on MHC loci counteract genetic drift in small fragmented populations of black grouse? – Ecol. Evol. 2: 341–353.
Strand, T. et al. 2013. Evolutionary history of black grouse BLB1 and BLB2 revealed through single locus sequence-based typing. – BMC Genetics 14: 29.
Sutton, J. T. et al. 2011. Disentangling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. – Mol. Ecol. 20: 4408–4420.

Taberlet, P. et al. 1998. Comparative phylogeography and post glacial colonization routes in Europe. – Mol. Ecol. 7: 453–464.
Valière, N. 2002. Gimlet: a computer program for analyzing genetic individual identification data. – Mol. Ecol. Notes 2: 377–379.

Supplementary material (available online as Appendix wlb. 00093 at <www.wildlifebiology.org/readers/appendix>). Appendix 1.

Wang, B et al. 2012. Transcriptome sequencing of black grouse (Tetrao tetrix) for immune gene discovery and microsatellite development. – Open Biol. 2: 120054.
Weir, B. S. and Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. – Evolution 38: 1358–1370.
Wright, S. 1943. Isolation by distance. – Genetics 28: 114–138.