Cent. Eur. J. Biol. • 7(4) • 2012 • 759-775
DOI: 10.2478/s11535-012-0051-2

VERSITA

Central European Journal of **Biology**

Population genetics of the hazel hen *Bonasa bonasia* in Poland assessed with non-invasive samples

Research Article

Robert Rutkowski[1,*], Marek Keller[†,2], Patrycja Jagołkowska[1]

[1]*Department of Molecular and Biometrical Techniques, Museum and Institute of Zoology PAS, 00-679 Warsaw, Poland*

[2]*Department of Forest Zoology and Game Management, Faculty of Forestry, Warsaw University of Agriculture, 02-776 Warsaw, Poland*

Received 02 November 2011; Accepted 25 April 2012

Abstract: Despite a severe decrease in the number of hazel hens during the 20th century, nowadays this grouse species is rather common in the forests of Northeastern and Southern Poland. We used mitochondrial control region and microsatellite markers to examine the genetic variability of Polish populations of hazel hens. We used non-invasively collected faeces to estimate genetic variability within populations, genetic differentiation among populations as well as genetic differentiation between two regions inhabited by two different subspecies of hazel hens. Our results confirm the usefulness of DNA from faeces to obtain reliable information on the population genetics of hazel hens. We found a rather high level of genetic variability in the Polish population. Genetic variability was higher in birds from continuous forests in the South of the country than in birds from fragmented habitats in the Northeast. Genetic differentiation was higher among subpopulations from Northeastern Poland. Additionally, both classes of molecular markers suggested the presence of two distinct genetic groups of birds, corresponding to previously described subspecies. We conclude that the genetic variability of the Polish hazel hen population has been influenced by habitat fragmentation and the history of the population during its post-glacial colonization of Poland from different glacial refugia.

Keywords: *Hazel hen • Bonasa bonasia • Cross-species amplification • Microsatellites • Non-invasive sampling • Genetic differentiation*

1. Introduction

The hazel hen *Bonasa bonasia* is a grouse species found in boreal forests and mountain regions in the Western Palearctic. The habitat requirements of the species are strongly interlinked with the abundance of spruce *Picea abies*, larch *Larix spp.* and fir *Abies alba* in the tree stands, as well as dense undergrowth, consisting of hazel *Corylus*, birch *Betula,* alder *Alnus,* and mountain ash *Sorbus* [1-3]. During the 20th century, the distribution of hazel hens in Western and Central Europe was drastically reduced, and many local populations heavily decreased in numbers or even vanished. However, in Poland, the species has recently shown signs of recovery in some areas which were formerly abandoned or poorly inhabited [4]. Nowadays, the hazel hen is a rather common species in large forests in Northeastern and Southern parts of Poland, not only in natural or semi-natural forests, but also in intensively managed forests. Current changes in the distribution and abundance of the hazel hen in Poland are described in Tomiałojć and Stawarczyk [5] and Bonczar [6].

Previous hypotheses regarding the taxonomy the hazel hen (identification of subspecies) were based on morphological differentiation (body size, details of plumage coloration), and hence cannot be treated as conclusive. Morphological features, as a rule, have a clinal variation; as a consequence, there are contradictory opinions in the literature about even basic questions regarding subspecies boundaries [3,7,8].

* *E-mail: robertrut@miiz.waw.pl,* † *deceased*

It is believed that *B. b. bonasia* occurs in Northern Europe, Scandinavia, Belarus and Northeastern Poland, whereas Southern Poland (Carpathians) is inhabited by *B. b. rupestris* [1,9]. Additionally, according to some authors, hazel hens found in mountains in Poland belong to the subspecies *styriaca* [8]. Birds from Western Poland (Sudety mountains) are even sometimes classified as *B. b. rhenana* (different from the Carpathian birds - *rupestris*) and *B. b. bonasia* from the Westernmost parts of its range (in Eastern Poland) are classified as *B. b. volgensis* [10].

The hazel hen is not as well studied as other forest grouse (such as capercaillie *Tetrao urogallus* and the black grouse *Tetrao tetrix*), perhaps due to its lower risk of extinction [11]. To our knowledge, all genetic studies on hazel hens performed until now have focused on the phylogeography of the species [12,13] and population genetics in continuous habitat [14]. Phylogenetic studies based on mitochondrial DNA have identified high genetic variability in a population from Hokkaido (Japan) and a clear separation between this population and birds from continental Asia [12,13]. A population study based on nuclear microsatellite markers has also shown high within-population genetic variability and rather high gene flow among subpopulations in Sweden [14], despite the low dispersal ability of this species.

The relatively high numbers of hazel hens living in Polish forests make it possible to study their genetic variability as well as the genetic differentiation among several isolated subpopulations. Estimates of genetic differentiation will enable not only the establishment of subspecies' boundaries, but also provide answers to more general questions such as the role of forest isolation in gene flow between particular subpopulations of the species [15].

Blood or tissue samples for population genetic studies of birds are often difficult or even impossible to collect in sufficient numbers. Thus, non-invasive sampling may provide an alternative solution, particularly in the case of species which are rare, elusive or in small densities. In many bird species, faeces can be easily recognized and sampled in large quantities. Moreover, advances in DNA-extraction technology and PCR based methods of DNA amplification make faeces a very useful source of genetic material [e.g. 16-18]. On the other hand, some problems exist with DNA obtained from samples collected non-invasively, especially from faeces. For example, DNA extracted from such biological material is often degraded, and PCR inhibitors are present in the extract. Both these factors make it difficult to amplify genetic markers efficiently [19]. Additional problems occur with microsatellite genotyping. As the amount of DNA extracted from faeces is usually small, incorrect genotypes can arise because of so called allelic drop-out (ADO) - non-amplification of an allele and estimation of a heterozygous genotype as homozygous and false alleles (FA) - interpretation of a homozygous genotype as heterozygous [20]. Many different strategies have been proposed to counter such errors; however, the majority of these sulotions are based on multiple multilocus genotyping from which a single consensus genotype is inferred [e.g. 20-22]. Recently, a real-time PCR technique has also been applied to evaluate the usefulness of non-invasive samples for microsatellite genotyping [e.g. 23,24]. However, since these approaches are costly (both in terms of time and money), a pilot study should be conducted [20] to determine whether population genetic analysis based on non-invasive sources of DNA are altogether appropriate. The pilot study should also help to estimate genotyping error rates and, in turn, predict the minimal number of times each sample must be amplified before accepting its genotype at a particular locus [25-27].

The aim of our study was to estimate the usefulness of non-invasively collected faeces in studies of the subpopulations genetics of *Bonasa bonasia.* Our aim was also to perform an evaluation of the genetic differentiation between populations from Northeastern and Southern Poland, described by many authors as two different subspecies. We analyzed two classes of molecular markers: mitochondrial control region and microsatellites. In the case of microsatellites, we also made an attempt to estimate genotyping errors arising due to the use of template DNA extracted from faeces.

2. Experimental Procedures

2.1 Sample collection

Faeces from hazel hens (total *N*=180) were collected in two regions of Poland: Northeastern Poland, including samples from the Romincka Forest (RF), Borecka Forest (BorF), Piska Forest (PF), Augustowska Forest (AF), Białowieża Forest (BiF) and Knyszyńska Forest (KF), and Southern Poland, including samples from the Bieszczady (BM), Beskid Sądecki (BSM), Gorce (GM) and Tatra (TM) mountains (Table 1; Figure 1). The RF, BorF, and PF sites in Northeastern Poland are isolated, situated several tens of kilometers from each other, whereas KF and BiF are in close proximity. AF and KF were separated only by the Biebrza river valley at approximate distance of 80 km. All the forests in Southern Poland were located in successive mountain ranges; thus, the distribution hazel hens in this region can be regarded as continuous.

Since the territories of hazel hen pairs are relatively small (approximate radius of 200 m^2, [3]) and to avoid repeated measurements on the same individuals, only

Locality (code and name)		Region	Sample size
AF	Augustowska Forest	Northeastern Poland	12
BiF	Białowieska Forest		22
BorF	Borecka Forest		26
KF	Knyszyńska Forest		22
PF	Piska Forest		12
RF	Romincka Forest		26
BM	Bieszczady Mountains	Southern Poland	17
BSM	Beskid Sądecki		15
GM	Gorce Mountains		9
TM	Tatra Mountains		19

Table 1. Information on sampling locations and sample sizes for *B. bonasia* in Poland. The sample size indicated is the number of faeces samples collected

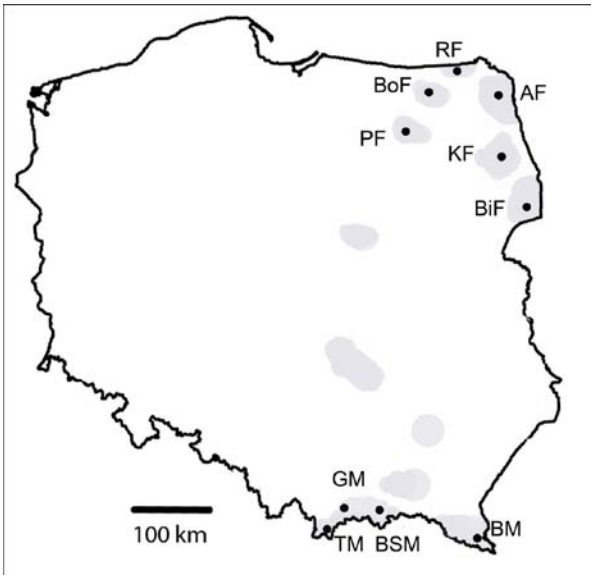

Figure 1. Sampling locations (see abbreviations in Table 1) and other forests (in grey) inhabited by the hazel hen in Poland.

one faecal sample was collected from each putative territory. The hazel hens excrete droppings in the form of small piles and it is unlikely that droppings from different individuals get mixed in a single pile. Therefore, we considered each pile as a faecal sample from a single individual. Samples were collected over a three-year period (2005–2007) in winter (January–March), when they are easy to find. Samples were dried immediately after collection, transported to the laboratory within a few days and stored frozen in sterile, glass tubes.

2.2 DNA isolation

DNA extraction was performed according to the methodology described by Regnaut *et al*. [28]. To avoid contamination, the isolation process was conducted in a room exposed to ultraviolet light overnight. DNA isolation was performed as 18 independent cycles. In each of the cycle, DNA was extracted from sets of 10 samples, including a "blind sample", which contained all reagents without any faecal matter.

2.3 Amplification of the control region and selection of extracts for microsatellite analysis

To assess the feasibility of using non-invasively collected faeces for describing the molecular population genetics of *Bonasa bonasia,* we used a pilot study as proposed by Taberlet *et al*. [20]. Briefly, we amplified the fragment of mitochondrial control region in the extracts obtained. The success of the amplification procedure was estimated using 1.5% gel electrophoresis – the presence of a visible product of the expected size was treated as a successful amplification. These samples were then used to amplify microsatellite markers

To amplify mitochondrial DNA, two PCR primers were designed (BonB5, BonB6) based on complete control region sequences deposited in GenBank (AF532418, AF532419, AF532420, AJ297155) using the on-line software Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). These primers amplified the 474 base pair fragment of the IA domain of the mitochondrial control region. Primer BonB5 (5'-CCAAGGACTACGGCTTGAAA-3') was complementary to the fragment of tRNA; and BonB6 (5'-GGTGTAGGGGGAAAGAATGG-3') was complementary to the fragment of the domain IB - hypervariable part of domain I in galliforms control region.

The PCR reaction was performed under the following conditions – initial denaturation: 94°C/3 min.; 30 cycles: 94°C/45 s; 55°C/45 s; 72°C/45 s; final elongation: 72°C/5 min.

The reaction mix (25 µl) contained 5-6 µl of DNA extract, 12.5 µl of REDTaq PCR ReadyMix (Sigma), 7.5 µl of water and 10 pmol of each primer.

Sequencing of PCR products was performed using a Beckman Coulter CEQ8000 (Comesa, Warsaw, Poland) capillary sequencer. Firstly, the PCR product was purified using Clean-Up kit (A&A Biotechnology, Gdańsk, Poland). Then, the purified PCR product was used as a template for the sequencing reaction: 90ºC/20 s; 50ºC/20 s; 60ºC/4 min. The reaction mix (20 µl) contained 10 µl of purified PCR product, 4 µl of CEQ DTSC Quick Start Kit (Comesa), 1.6 pmol of primer and 4 µl of water.

Sequences of light and heavy strands were aligned using BioEdit [29] and deposited in GenBank (accession numbers: JN935902–JN935922).

2.4 Microsatellite analysis

To our knowledge, no species-specific microsatellite markers have ever been developed for the hazel hen. In a previous study [14], 12 microsatellites from closely related grouse species (the capercaillie and the black grouse), were cross-amplified using PCR primers described by Segelbacher *et al*. [30] and Piertney and Höglund [31]. We used 6 loci tested by Sahlsten *et al*. [14] (TUT1, TUT2, TUT3, TUT4, Bg15 and Bg18), and the locus TTT1 (source species: *T. tetrix*, [32]) as well as loci Bg6, Bg10, Bg14, Bg19 and Bg4 (source species: *T. tetrix*, [31]). We focused on tetra-nucleotide repeats, as they are easy to analyze and because microsatellites containing longer units (e.g. tetra-nucleotide repeats) seems to evolve faster than those containing shorter units (e.g. di-nucleotide repeats), thus might have shown higher polymorphism within populations [33].

Microsatellite markers were amplified using PCR. Reactions were performed in 25 µl of reaction mix, which contained: 10 pmol of each primer, 5 µl of DNA extract, 12.5 µl of REDTaq PCR ReadyMix (Sigma) and 7.5 µl of PCR grade water (Sigma) and 0.5 µl of Anti-inhibitor (DNA Gdańsk, Poland). Forward primers were labelled fluorescently on their 5' ends with one of the following labels: Dye2, Dye3 or Dye4 (ProOligo). Prior to PCR, all reagents, tubes and pipettes were exposed to ultraviolet light for 15 minutes. The following PCR profiles were used in a Techne Touchgene thermocycler: initial denaturation – 3 min. at 94°C; followed by 35 cycles of 30 s at 94°C, 45 s in 55°C, 45 s in 72°C; and one cycle of 30 s at 94°C, 1 min. at 50–60°C, 5 min. at 72°C.

Negative PCR controls were always included for each set of reactions. No amplification product was found in any negative controls after electrophoresis in agarose gels and analysis in an automatic sequencer.

The length of the amplified fragments was estimated using a CEQ8000 Beckman Coulter automatic sequencer. The data were analyzed using the Beckman Coulter Fragment Analysis Software.

2 4.1 Estimation of genotyping errors

Loci that were successfully amplified and appeared to be polymorphic in a set of 118 analyzed samples (8 loci, Table S2) were used to estimate the probability of ADO and FA. We performed six independent amplifications for each locus in 46 randomly selected samples. After the analysis in an automatic sequencer, the genotypes obtained were used to construct consensus genotypes and to estimate error rates arising from ADO and FA with the software GIMLET [34]. Then, we used the software GEMINI [26], which uses 10 000 simulations to estimate the number of reactions needed to achieve a high reliability that the genotype obtained is correct.

2.5 Estimation of genetic variability and genetic differentiation between populations

2.5.1 Mitochondrial DNA

To estimate the genetic variability in each subpopulation, the number of haplotypes (*H*), nucleotide (π) and haplotype (*h*) diversity [35], based on the observed number of segregating sites, were calculated using DNAsp v. 4.0 [36]. The software Arlequin v3.5.1.2 [37] was used to estimate Fu's F_S statistic [38]. This value is often used as an indicator of a recent change in population size because it is sensitive to departures from the demographic equilibrium. The mismatch distributions were calculated for each locality and for two regions separately, as an additional test for demographic equilibrium [39] using Arlequin. The fit between the observed and expected distributions was evaluated by the sum of the square deviations from the mean (SSD, [40]). The significance of SSD (P_{SSD}) was assessed by parametric bootstrap re-sampling (1000 simulations), as implemented in Arlequin. Using the same software, we also computed the raggedness index *r* [41] and the probability of its significance (*Pr*).

Pairwise differentiation among populations was calculated with the same program using F_{ST}, which was calculated using haplotype frequencies. The significance of the observed variance components was tested by a nonparametric permutation method [42].

To analyse relationships among haplotypes, we created a minimum spanning network using Arelquin and a Neighbour-Joining tree using MEGA4 [43].

2.5.2 Microsatellites

The polymorphism of microsatellite loci was estimated on a few different levels. Firstly, we assessed allelic diversity (*A*), allelic richness (*R*, [44]), observed heterozygosity (H_O) and unbiased expected heterozygosity (H_E) [45] for the whole data set. The fixation index (F_{IS}) for each locus was calculated and its significance was tested using a randomization procedure and a Bonferroni correction for multiple comparisons. These analyses were performed using GenAlEx version 6.0 [46] and FSTAT version 2.9.3 [47]. Genotypic linkage disequilibrium between all pairs of loci and deviations from the Hardy-Weinberg equilibrium were evaluated using Genepop Web version 4.0.10 [48,49]. Following this procedure, the analyses were performed for two regions separately and for each of the subpopulations.

Genetic differentiation between subpopulations was estimated using F_{ST}. Overall F_{ST} [50] and pairwise F_{ST}

were obtained with FSTAT. The 95% confidence intervals for the overall F_{ST} were also estimated in FSTAT.

The significance of differences between mean values of allelic richness (calculated across separate subpopulations within each of the regions), F_{IS}, F_{ST}, and observed and expected heterozygosity in Northeastern and Southern Poland were tested using the permutation procedure as implemented in FSTAT.

The Bayesian-clustering method (STRUCTURE version 2; [51]) was used to examine how well the predefined "subpopulations" corresponded to genetic groups (K). Structure was run 10 times for each user-defined K (1–10), with an initial burn-in of 50,000, and 100,000 iterations of the total data set. The admixture model of ancestry and the correlated model of allele frequencies were used. The proportion of membership of each predefined population within each genetic group was estimated for K with the highest likelihood. Sampling location was not used as prior information. Then, we examined ΔK statistics that identify the largest change in the estimates of K produced by STRUCTURE, as ΔK may provide a more realistic estimation of K than estimations based on likelihoods [52].

We used STRUCTURE HARVESTER to visualize STRUCTURE results [53]. Then, we applied CLUMPP [54] to average the multiple runs given by STRUCTURE and correct for label switching. The output from CLUMPP was visualized with DISTRUCT v 1.1 [55].

Finally, for both measures of pairwise genetic differentiation (F_{ST} from mitochondrial haplotypes and microsatellite genotypes), we analyzed the correlation between genetic and geographic distances using the Mantel test implemented in Arlequin.

3. Results

3.1 Mitochondrial DNA

Successful amplification of the mitochondrial control region was obtained in 118 samples of hazel hen faeces (65% of the total sample size). The 474 base pair fragment was amplified using primers BonB5 and BonB6, but only the 350 base pair long fragment was clearly readable for all 118 samples. The variable sites of the fragment are shown in Figure 2. 21 haplotypes were found in the population analyzed, including eight unique haplotypes (found in only one subpopulation). There was a strict geographical division among haplotypes: one group of haplotypes occurred only in Northeastern Poland, the other only in Southern Poland (Table 2). Comparing both regions, we found similar genetic variability in terms of the number of haplotypes and the haplotype diversity but not in nucleotide diversity, which was clearly lower in sites from Southern Poland (Table 3). The lowest number of haplotypes and the lowest haplotype diversity were found in PF (the only population with a single haplotype) and RF subpopulations from Northeastern Poland; the highest number of haplotypes and haplotype diversity were found in TM from Southern Poland. Nucleotide diversity was the lowest in RF. The mismatch distribution and tests of neutrality suggested a population expansion for birds in both Northeastern Poland and Southern Poland (Table 3). The sums of squared deviations (SSD) of the mismatch distribution were not statistically significant, indicating that the curves fitted the sudden expansion model tested. The raggedness index was low for both regions, which indicates a smooth distribution, as is expected in cases where there is population expansion. The measure of Fu's *F*s statistic revealed a value significantly different from zero in Southern Poland, suggesting a recent population expansion in this region.

The analysis of pairwise genetic differentiation indicated a higher F_{ST} among populations from Northeastern Poland (range: 0.17–0.92), than among populations in the South of the country (-0.01–0.44). In the Northeastern region, eight of the 15 pairwise F_{ST} considerably exceeded the value

```
          1111111111111111222
          0123555666678899014
          5222036678908490948 1
          ---------------------
H1        GTCCACTTCTCTCTCTCAAT
H3        .C.......CTC.....G..
H4        .C.......C.C........
H9        A...................
H10       ........A..C........
H11       A....T..............
H12       .C...T...C......TG..
H13       .CT..T...C......TG..
H14       .C...T...C......T...
H17       A............C......
H18       .CTT-T...C......TG..
H20       ............T.......
H21       ...................C
H23       .C....C..C.......G..
H24       .C...T.....C......G.
H26       ..............C.G..
H27       .CT.-T...C......TG..
H29       .C.....C.C.C.....G..
H30       .C.......C.C...C.G..
H31       .CT..T...C....T..G..
H33       .C.....C.C.C........
```

Figure 2. Variable sites in fragments of the mitochondrial control region of hazel hens analyzed. Dashes indicate a missing nucleotide and dots indicate similarity among sequences. Haplotype names are given on the left.

	Northeastern Poland						Southern Poland			
Haplotype	AF	BiF	BorF	KF	PF	RF	BM	BSM	GM	TM
H3		+		+						
H4	+		+	+		+				
H17		+								
H18	+	+	+							
H20		+								
H21		+	+		+	+				
H23			+							
H24			+							
H27				+						
H29				+						
H33	+									
H1									+	
H9							+			
H10								+	+	+
H11									+	+
H12							+	+		+
H13							+	+		+
H14							+			
H26										+
H30								+		
H31								+		+

Table 2. Occurrence of mtDNA haplotypes in Poland. Unique haplotypes are shown in bold.

	n	H	*h*	π	SSD	P (SSD)	r	P (r)	Fs	P(Fs)
AF	8	3	0.603	0.01179	0.20372	0.00000	0.35587	0.05000	3.20019	0.944
BiF	12	5	0.864	0.01323	0.05391	0.05000	0.12374	0.15000	1.80117	0.809
BorF	18	5	0.529	0.01106	0.37225	0.00000	0.22359	1.00000	2.46811	0.863
KF	14	4	0.758	0.01119	0.05131	0.25000	0.12438	0.45000	2.97680	0.910
PF	10	1	0	0	0.00000	0.00000	0.00000	0.00000	0	–
RF	15	2	0.143	0.00132	0.02967	0.00000	0.77551	0.80000	0.90290	0.506
Northeastern Poland	77	11	0.836	0.01120	0.00972	0.78000	0.01863	0.91000	0.74566	0.637
BM	11	4	0.600	0.00213	0.02385	0.30000	0.18750	0.45000	-1.89165	0.017
BSM	12	5	0.788	0.00397	0.00804	0.35000	0.10514	0.35000	-1.27025	0.109
GM	7	3	0.667	0.00500	0.39002	0.00000	0.20862	1.00000	0.66806	0.638
TM	11	6	0.873	0.00651	0.00708	0.75000	0.08562	0.60000	-1.86685	0.064
Southern Poland	41	10	0.841	0.00522	0.00550	0.24000	0.08806	0.17000	-3.53994	0.024

Table 3. Comparison of the genetic variability, measured by polymorphism in the mtDNA control region, and mismatch distribution (sum of square deviation SSD and raggedness index r) and Fu's neutrality (Fs) test results for Polish populations of *B. bonasia*. The significant values are shown in bold. *n* - number of samples for which the sequence of mtDNA was obtained; H – number of haplotypes; *h* – haplotype diversity; π – nucleotide diversity; SSD - sum of squared deviations; p(SSD) - probability of observing a less good fit between the model and the observed distribution by chance; r - raggedness index; P(r) - probability of getting higher value by chance; Fs - Fu's F_S statistics; p(Fs) - randomization probability test of Fs significance

of 0.30; in Southern Poland, only one comparison gave a F_{ST} value higher than 0.30. In the Southern region, only 3 out of 6 pairwise comparisons were significant, whereas all F_{ST} values were significant in Northeastern Poland (Table 4). Between populations from different regions, F_{ST} values ranged from 0.14 to 0.79 and the majority exceeded 0.30 (17 among 24 pairwise comparisons).

The minimum spanning network did not indicate an unambiguous relationship among haplotypes and showed multiple connections (Figure 3). However, the haplotypes seemed to be divided into three groups. One group consisted of Southern haplotypes exclusively, whereas Northeastern haplotypes were divided into two groups. Haplotype H1, found in the GM population (Table 2), was connected to haplotypes H3 and H4, which are frequent in Northeastern Poland. The NJ phylogenetic tree produced a similar picture of the relationships among haplotypes (Figure 1S).

3.2 Microsatellite loci

3.2.1 Cross-species amplification

Among the 12 loci tested, 10 (83%) were successfully amplified (Table S2). No amplification products were found in BG14 and BG19, despite using a wide range of different annealing temperatures.

Eight loci (representing 80% of the loci successfully amplified and 66% of all the loci analyzed) showed substantial polymorphism with 7 to 17 alleles (Table S2). No linkage disequilibrium was found among loci. The majority of loci were not in Hardy-Weinberg equilibrium, and, in six out of eight cases, F_{IS} values were significantly higher than zero, indicating a heterozygote deficit. A significant F_{IS} was also found for the total data set (F_{IS}=0.190, P<0.01).

3.2.2 Estimation of genotypic errors

Both types of errors, ADO and FA, were detected only in one locus — BG6 (Table S1). Genotyping errors

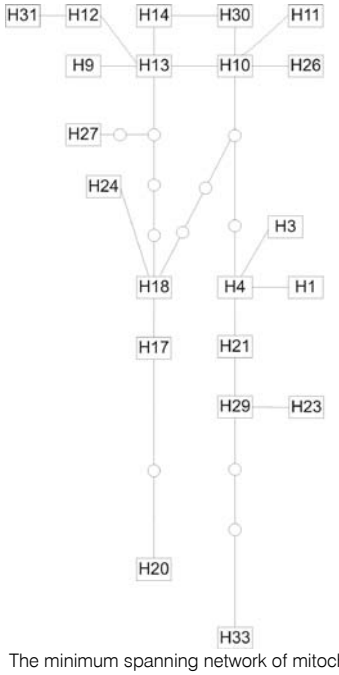

Figure 3. The minimum spanning network of mitochondrial control region haplotypes identified in 118 samples from hazel hens in Northeastern and Southern Poland. Names of the haplotypes are given inside circles. Each line between two circles represents a one point mutation. Small, empty circles indicate missing haplotypes.

	AF	BiF	BorF	PF	KF	RF	BM	BSM	GM	TM
AF		**0.23825**	**0.38554**	**0.73887**	**0.20252**	**0.15893**	**0.40519**	**0.29378**	**0.43268**	**0.25060**
BiF	0.0367		**0.29771**	**0.48638**	**0.17571**	**0.52914**	**0.28345**	**0.18939**	**0.30091**	**0.14724**
BorF	0.0693	0.0365		**0.65239**	**0.34096**	**0.59270**	**0.43318**	**0.33893**	**0.45607**	**0.30235**
PF	**0.1091**	**0.0839**	**0.0927**		**0.58860**	**0.91988**	**0.73377**	**0.59365**	**0.79301**	**0.56364**
KF	0.0810	0.0451	**0.0703**	**0.1230**		**0.43753**	**0.32008**	**0.22736**	**0.33853**	**0.18686**
RF	**0.1023**	**0.0959**	**0.0687**	**0.0862**	**0.1102**		**0.67529**	**0.55124**	**0.72127**	**0.52054**
BM	**0.2123**	**0.2058**	**0.1755**	**0.2223**	**0.2083**	**0.2686**		0.12302	**0.44398**	0.08245
BSM	**0.2038**	**0.1937**	**0.1741**	**0.2050**	**0.1892**	**0.2668**	0.0003		**0.23784**	-0.01443
GM	**0.2082**	**0.2174**	**0.1721**	**0.2290**	**0.1975**	**0.2709**	0.0402	**0.0587**		**0.22402**
TM	**0.2261**	**0.2133**	**0.1948**	**0.2368**	**0.2193**	**0.2874**	0.0060	0.0290	0.0445	

Table 4. Among subpopulations genetic differentiation (pairwise F_{ST}) for *B. bonasia* population, estimated using mtDNA control region haplotypes (above diagonal) and using microsatellite markers (below diagonal). The significant values are shown in bold. Grey background - genetic differentiation between populations from Northeastern and Southern Poland. The significance of mtDNA differentiation was tested with 110 permutations. The overall F_{ST} for microsatellites was 0.149 [95%CI 0.102–0.203]. Significant values after Bonferroni correction (Bonferroni corrected P-value, based on 900 permutations, at α=0.05 was 0.00111) are shown in bold.

occurred only in two types of genotypes: 202/310 and 206/310 (alleles in base pairs), and had a relatively low probability of occurrence, with ADO errors being significantly more probable.

The analysis of the percentage of correct identifications of a genotype as a function of the number of typing repetitions indicated that two PCR repetitions was the optimal number of reactions for determining correct genotypes. The success rate after 3 or more PCR repetitions decreased because additional PCRs resulted in more instances of ADO and FA errors and a greater occurrence of ambiguous genotypes. Hence, all 118 samples were analysed twice for each locus, and consensus microsatellite genotypes were used in the analyses of genetic variability and genetic differentiation.

3.2.3 Estimation of genetic variability and genetic differentiation

Despite the lower number of samples analyzed from this region, we found higher genetic variability in Southern Poland compared to the Northeastern. The mean allelic richness and heterozygosity were higher in Southern Poland (Table 5). Moreover, in the Northeastern region, 5 loci were not in Hardy-Weinberg equilibrium and the overall F_{IS} was significantly greater than zero. In the Southern region only one locus (TUT1) exhibited significant heterozygote deficiency, and the overall F_{IS} was not significantly higher than zero.

When comparing the genetic variability within particular subpopulations (Table 6), we found that the mean allelic richness was generally higher in subpopulations from the Southern region. Two subpopulations from the Northeastern region were not in Hardy-Weinberg equilibrium (BorF and PF), and one of them (BorF) had an F_{IS} value significantly higher than zero. In Southern Poland, only GM was not in equilibrium. The lowest values of allelic richness and heterozygosity were found in RF from Northeastern Poland, whereas the highest values were found in BS from Southern Poland. A comparison of the mean within-population genetic variability between the two groups of subpopulations (Northeastern and Southern) indicated significantly higher allelic richness in the South than in the Northeast (4.442 *vs* 3.615, P=0.007, two-tailed test, 1000 permutations) as well as higher observed heterozygosity in the South (0.636 *vs* 0.503, P=0.015, two-tailed test, 1000 permutation).

We found moderate but highly significant genetic differentiation between Northeastern and Southern

Region	Locus	*A*	*R*	H_O	H_E	*HWE*	F_{IS}
Northeastern Poland	TUT1	6	5.72	0.468	0.671	<0.001	0.309*
	TUT2	3	2.52	0.091	0.111	0.2226	0.185
	TUT3	8	7.41	0.662	0.777	0.0369	0.154
	TUT4	7	5.79	0.273	0.340	0.0378	0.204
	TTT1	12	10.38	0.724	0.830	0.0434	0.134
	Bg6	12	9.92	0.684	0.757	0.0022	0.102
	Bg18	8	6.82	0.513	0.522	0.2197	0.023
	Bg4	8	7.01	0.610	0.647	0.1889	0.063
	Mean	8.00	6.94	0.503	0.582	<0.001	0.141*
Southern Poland	TUT1	6	5.98	0.220	0.408	0.0031	0.472*
	TUT2	9	9.00	0.780	0.827	0.2924	0.068
	TUT3	8	7.98	0.805	0.791	0.3085	-0.005
	TUT4	4	4.00	0.317	0.302	0.7211	-0.038
	TTT1	8	7.95	0.732	0.783	0.7178	0.078
	Bg6	12	12.00	0.925	0.871	0.7811	-0.049
	Bg18	7	7.00	0.756	0.765	0.2701	0.024
	Bg4	4	4.00	0.550	0.549	0.0743	0.011
	Mean	7.25	7.24	0.636	0.662	0.0544	0.052

Table 5. The characterization of polymorphic microsatellites in two groups of *B. bonasia*: Northeastern Poland (n=77) and Southern Poland (n=41). *A* – number of alleles; *R* – allelic richness; H_O – heterozygosity observed; H_E – heterozygosity expected; HWE – *P*-values for HWE exact test for heterozygote deficiency/excess; F_{IS} – fixation index; * – significant F_{IS} value after Bonferroni correction (Bonferroni corrected P-value at α=0.05 was 0.00313).

		n	*A*	*R*	H_O	H_E	*HWE*	F_{IS}
Northeastern Poland	AF	8	3.63	3.52	0.438	0.500	0.0512	0.192
	BiF	12	4.88	3.96	0.549	0.522	0.5293	-0.008
	BorF	18	4.88	3.85	0.479	0.592	0.0014	0.217*
	KF	14	4.88	4.04	0.625	0.563	0.9948	-0.074
	PF	10	3.75	3.41	0.475	0.493	0.0019	0.089
	RF	15	3.50	2.91	0.433	0.439	0.6161	0.048
Southern Poland	BM	11	5.13	4.58	0.618	0.610	0.1325	0.035
	BS	12	5.75	4.77	0.667	0.641	0.8087	0.004
	GM	7	4.00	4.00	0.589	0.647	0.0068	0.165
	TM	11	5.00	4.42	0.648	0.604	0.4103	-0.025

Table 6. The characterization of genetic polymorphism of 10 subpopulations of *B. bonasia*. The mean values of genetic variability indicators for 8 polymorphic loci are shown. *n* - number of samples analysed for microsatellites; *A* – number of alleles; *R* – allelic richness; H_O – heterozygosity observed; H_E – heterozygosity expected; HWE – *P*-values for HWE exact test for heterozygote deficiency/excess; F_{IS} – fixation index; * – significant F_{IS} value after Bonferroni correction (Bonferroni corrected P-value at α=0.05 was 0.00063)

Poland (F_{ST}=0.1628; P<0.01). In the majority of cases, we found significant genetic differentiation among subpopulations (Table 4). Genetic differentiation was significant for all comparisons among subpopulations from different regions. The pairwise F_{ST} was also generally higher than 0.19 (ranging from 0.17 to 0.28). The highest values of F_{ST} were between Southern subpopulations and birds from PF and RF in the Northeastern region. In Southern Poland, we found significant genetic differentiation only between BSM and GM; all pairwise F_{ST} were very low (range: 0.0003 to 0.06). In Northeastern Poland, the comparison of all sites with PF and RF indicated significant genetic differentiation (F_{ST} ranged from 0.07 to 0.12), whereas differentiation among AF, BiF, BorF and KF was generally small and insignificant (F_{ST}=0.04–0.07).

In the STRUCTURE analyses, we found the highest mean likelihood (averaged over 10 runs) for K=6 (Figure 4A). Bayesian assignment of individuals to these clades indicated complete isolation between the Northeastern and Southern populations (Figure 4B). The analysis suggested the presence of two genetic clusters in the Southern region (clusters: V and VI). The majority of individuals belonged to cluster V. However, in every subpopulations we also found individuals of mixed ancestry, so the genetic groups were not restricted to particular subpopulations. In four populations from Northeastern Poland (AF, BiF, BorF and KF), we found individuals of mixed ancestry (mainly clusters I and II, Figure 4B). Almost all individuals from the PF population were assigned to a single cluster (IV), which was absent from other Northeastern subpopulations. Individuals from RF were also assigned to a single cluster (III), but individuals from this cluster were also present in AF and BorF (Figure 4B).

The pattern of mean likelihood distribution (increasing and then leveling off, Figure 5A) suggested that the subpopulations examined may represent a 'hierarchical island model' (a few islands within different archipelagos) [52]. In such a case, the number of clusters detected by STRUCTURE is usually higher than the real number. However, another statistic, based on the second order rate of change of the likelihood, ΔK, usually detects the real number of clusters [52]. Indeed, the analysis based on ΔK suggested two genetic clusters, which corresponded to the two regions of Poland we investigated (Figure 5A and B): all individuals from the Northeast were assigned to the first genetic group and birds from the South to the other. We assume that this division more realistically reflects the genetic structure of Polish population of hazel hens.

The Mantel test indicated a strong and highly significant correlation between genetic and geographic distances for microsatellite markers (r=0.912642, *P*=0.003) and a marginally non-significant correlation for mitochondrial DNA (r=0.168368, *P*=0.077).

4. Discussion

4.1 Usefulness of non-invasive samples and estimation of genotypic errors

Despite their undisputed attractiveness, non-invasively collected samples usually provide DNA extracts characterized by a low concentration and quality of target DNA. Microsatellite typing has proven to be sensitive to such problems. As such, population genetics inferred from non-invasive samples including faeces may not be an appropriate approach [20,56]. However, our study shows that droppings form hazel hens were suitable to

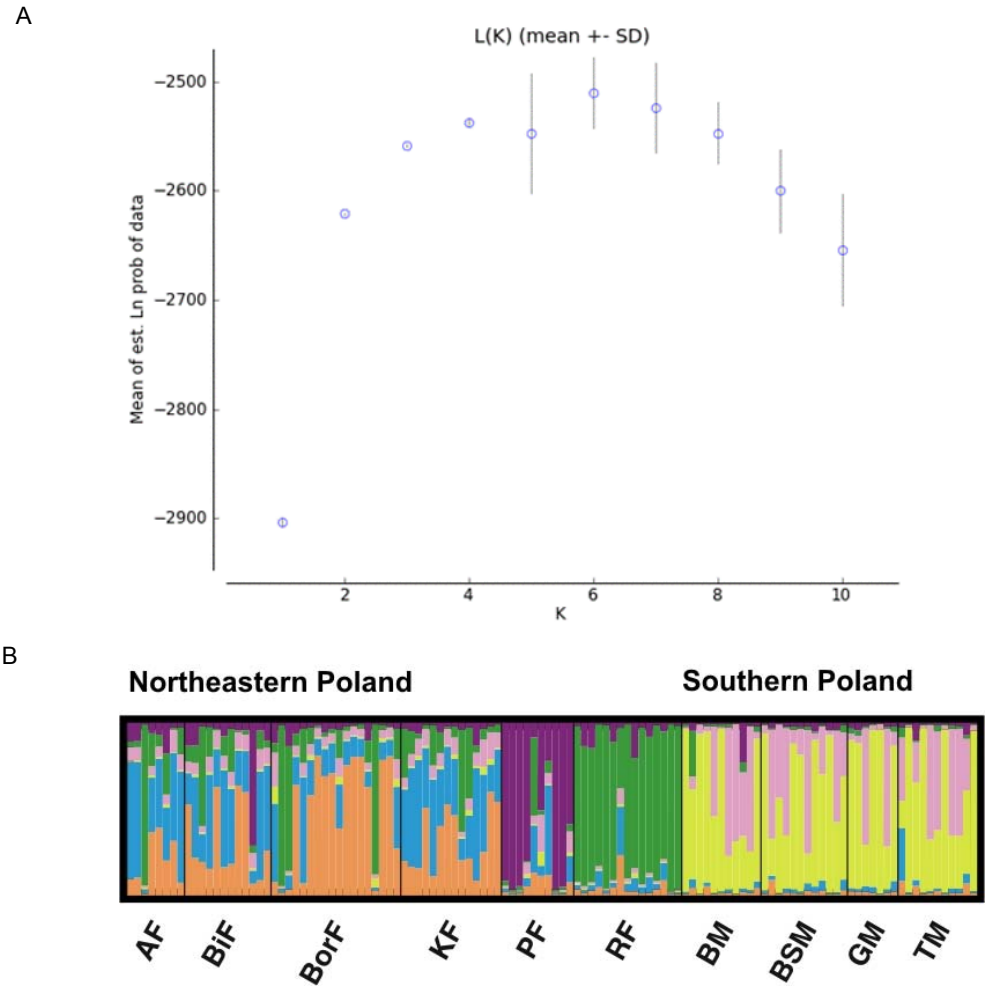

Figure 4. (A) STRUCTURE analysis of 118 individuals of *B. bonasia* from 10 subpopulations. The scatter plot shows the estimated likelihood of each number of inferred genetic clusters. The optimal number of clusters (K=6) was determined by the highest log-likelihood value and low variance for 10 independent iterations. (B) Bayesian assignment of individuals to 6 genetic clades (I - blue; II - brown, III - green; IV - violet; V - yellow; VI - pink). Each bar represents the estimated posterior probability of each individual bird belonging to each of the inferred clusters. Solid black lines define the boundaries between the populations used in the analysis.

extract DNA. We successfully amplified and sequenced mtDNA fragments in 65% of the samples collected, and the strategy proposed by Taberlet *et al*. [20] appeared to be very appropriate in our case: pre-selection of extracts based on the results from mtDNA amplifications resulted in successful microsatellite genotyping in all samples. Thus, we can recommend this method as a simple, cost-effective and fast procedure to select hazel hen droppings for microsatellite analysis. Moreover, ADO and FA errors were found only in one locus. These problems appeared only in two types of genotypes, characterized by large differences in allele length. This observation is in concordance with other studies of genotyping errors with non-invasive samples. For example, shorter DNA fragments (alleles) are thought to be more consistently amplified than longer ones during PCR (the short allele dominance hypothesis, [57]). This hypothesis was recently confirmed by a meta-analysis of published data [55]. In our study, genotypes of the BG6 locus in which we had identified an ADO error consisted of alleles that differed in length by 100 base pairs. According to Broquet *et al*. [58], such a differences in microsatellite fragment length would decrease amplification success from 80% to 71%, and increase the rate of ADO errors from 20% to 26%. The BG6 locus was also the most polymorphic among all loci tested, both in terms of number of alleles and in terms of heterozygosity. Again, as stated by Hoffman and Amos [59], the variability of microsatellite loci appeared to have a small but significant positive effect on the risk of ADO errors.

The rate of genotyping errors identified in our study (mean ADO per locus less than 2% and FA of 0.5%) seemed to be very low compared to other studies that

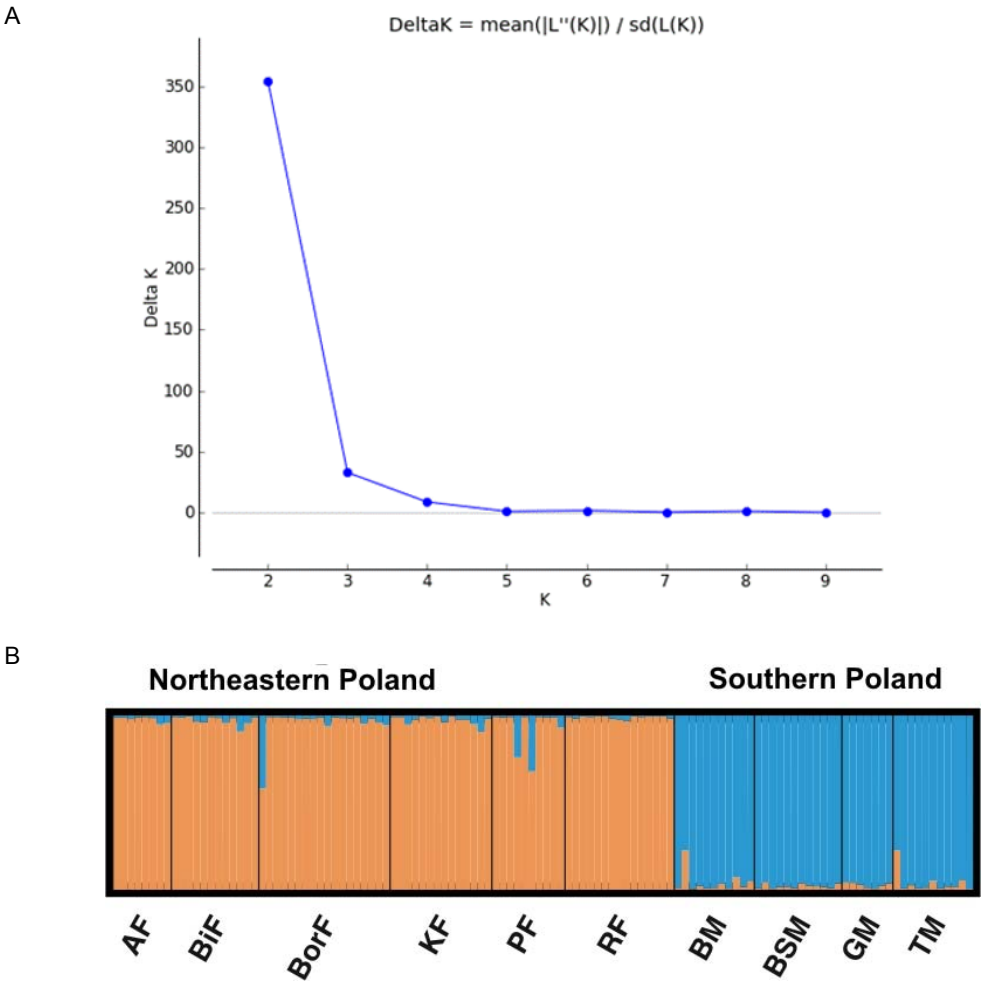

Figure 5. (A) ΔK values for K 2 through 9 (ΔK cannot be calculated for the first and final values in a series). (B) Bayesian assignment of individuals to two genetic groups, indicated by ΔK analysis. The separate bars represent the estimated posterior probability of each individual bird belonging to each of the inferred clusters (I – brown, II – blue). Solid black lines define the boundaries between the subpopulations used in the analysis.

used non-invasive sampling methods. ADO and FA error rates in these studies ranged between 2.5 [60] – 53% [19] and 1.2% [61] – 20% [62], respectively, although values of ADO and FA errors close or equal to zero are also reported [22,63]. In a study of microsatellite genotyping errors amplified on a DNA template extracted from the faeces of a member of the grouse family – capercaillie - ADO and FA errors occurred in the majority of the 11 loci tested (only one of them was free from FA errors). However, individual error rates were smaller than the ones we found in the BG6 locus, as well as the mean value for all eleven loci [28]. These findings suggest that grouse droppings could be particularly useful as a source of DNA that can be collected non-invasively if an appropriate method of isolation is employed. In our study, we used the same modification of standard extraction procedures as Regnaut *et al*. [28]. Another possible reason explaining the high reliability of genotypes obtained from hazel hen faeces in our study is seasonality. As mentioned in the 'Experimental procedure' section, we collected samples during the winter. Snow and low temperatures keep faeces very cold or even frozen, and hence significantly reduce DNA degradation. Similarly, Maudet *el al*. [64] found that faecal samples of Alpine ibex (*Capra ibex*) and Corsican mouflon (*Ovis gmelini*) collected in winter were much more useful for microsatellite genotyping and less prone to genotyping errors than samples collected in the spring. Moreover, hazel hens consume only plant material during the winter. While it has been suggested that most genotyping errors occur when working with faeces that contain plant remnants [65], plant components were not found to negatively impact amplification of faeces-extracted DNA in a previous

study [58]. Additionally, Idaghdour *et al*. [17] recommend collecting bird faeces for genetic analysis when insect prey are less prevalent in the environment as they are a source of powerful PCR inhibitors.

4.2 Estimation of genetic variability and genetic differentiation between populations

The estimated genetic variability of the Polish hazel hen population of based on microsatellite markers was very similar to that of the Swedish population [14]. In Sweden, the average number of alleles in polymorphic loci was 10.5 *versus* 9.87 in the Polish population. In contrast, the observed heterozygosity was higher in Poland than in Sweden (0.549 *vs* 0.466, respectively). Similarly, mitochondrial markers also suggested a high level of genetic variability in Poland. A study based on the same mitochondrial region in Finnish capercaillie, considered to be the most numerous population in Europe, indicated similar values of haplotype and nucleotide diversity as those measured in Polish hazel hens: within subspecies, haplotype diversity ranged from 0.670 to 0.860 and nucleotide diversity from 0.00283 to 0.00421 [66]. In Polish hazel hens, we found that within-region haplotype diversity was 0.840 and nucleotide diversity ranged from 0.005-0.01. This suggests that the Polish population of hazel hens has retained a substantial level of genetic variability, despite a population decrease in the 20th century. In contrast, the Polish population of capercaillie, which is extremely endangered and still decreasing in numbers, has much lower microsatellite indicators of genetic variability than the hazel hen, especially in terms of heterozygosity [67].

Our study provides useful information on the genetic variability of hazel hen subpopulations that differ in their level of isolation, from strongly isolated forests (RF, BorF and PF) to isolated but closely neighboring forests (AF, BiF and KF) and continuous mountain forests. We also describe patterns of genetic differentiation among populations differing in their taxonomy: birds from Northeastern Poland are often classified as *B. b. bonasia,* whereas populations from the South of the country are believed to belong to *B. b. rupestris*.

Our results obtained with molecular markers clearly indicated that the distribution of the genetic variability within Polish hazel hens was affected by two factors, which are discussed below.

The first factor was the level of spatial isolation of forest habitats. Molecular indicators of genetic variability were generally higher in subpopulations from continuous habitat than in subpopulations from fragmented habitats, supporting observations for other sedentary grouse species [68,69]. Habitat fragmentation has a much stronger effect on decreases in genetic variability than other forms of human disturbance, such as hunting [70]. Indeed, the allelic richness, heterozygosity, and polymorphism of the mitochondrial control region were lowest in the most spatially isolated subpopulations PF and RF. Additionally, the PF population was rather small in size compared to the other subpopulations studied, and contained only one mitochondrial haplotype. Such a low genetic diversity suggests that genetic drift may have a stronger effect in this subpopulations compared to other isolated subpopulations in which hazel hens were more common [5,6; M. Keller unpublished data]. Additionally, the hazel hen is a species with limited dispersal ability. Therefore, the continuity of forest habitat has a major effect on gene flow between subpopulations. The isolation of forest patches by open spaces such as agricultural lands, for example, could almost entirely impede the movement of individuals between isolated subpopulations. Previous work has shown that birds do not migrate between forests separated by farmlands and meadows more than 100–160 km apart [71]. Reduced dispersal may therefore explain the patterns of genetic differentiation observed in our study. Genetic differentiation among populations was very low in Southern Poland: in terms of microsatellite markers, the values of F_{ST} observed are comparable with values found in the continuous populations of other European grouse species such as the capercaillie [69,72]. In Northeastern Poland, despite higher levels of population differentiation (indicated by higher F_{ST} and F_{IS} values), we found a higher nucleotide diversity among haplotypes, suggesting that a group of haplotypes from this region differs more in nucleotide composition than haplotypes from the South of the country. We presume that this result indicates a random elimination of some haplotypes due to genetic drift, as supported by the minimum spanning network. However, in the case of separated, but closely spaced forests from Northeastern Poland (AF, BiF and KF), genetic differentiation was also rather low, although slightly higher than in Southern Poland. This suggests that the effects of genetic drift may be lower in closely located forests, due to the occasional migration of individuals.

The second factor affecting the distribution of the genetic variability in Polish hazel hens could be the evolutionary history of the population. Our results support the division of the Polish population into the subspecies *Bonas bonasia bonasia* and *B. b. rupestris*. Our molecular analyses suggest that two genetic groups are present in Poland, which correspond to the ranges of known subspecies: there were no shared mitochondrial haplotypes among the two groups, measures of genetic differentiation among populations of the different subspecies obtained with both classes of markers were

always much higher than measures of differentiation within subspecies, and, finally, microsatellite analyses suggest that individuals from the Northeast and South of Poland belong to separate genetic groups. Indeed, birds from the South and from the Northeast of the country also differ morphologically (for example, in coloration and body mass) [1; M. Keller unpublished data]. The presence of two distinct genetic groups in the hazel hen population from Sweden was also suggested, and explained by two different routes of post-glacial recolonization of the Scandinavian Peninsula [14]. Such a pattern could also be possible in the case of the Polish population, as many studies have confirmed the colonization of Northern Europe from Southern and Eastern refugia during glaciation [e.g. 73,74]. However, in contrast to the genetic differentiation observed in hazel hens from Sweden, genetic differences between Northeastern and Southern birds in Poland were much higher (F_{ST}=0.005 *vs* 0.16 in Sweden and Poland, respectively). This suggests that the division of the Polish population might be more long-lasting. A strong genetic structure in the Polish population corresponding to geographical regions is also surprising given the mitochondrial phylogeography of the capercaillie [66,75,76], which shows that boundaries of morphological subspecies correspond rather weakly to genetic differentiation.

Our analysis indicated an isolation-by-distance pattern, both for microsatellites and mitochondrial markers. However, we presume that this result is an outcome of patterns of differentiation between geographical regions: the highest differentiation was found among localities from different regions, which are the most distant locations. The correlation between genetic and geographical distance was much higher in the case of microsatellites than for the mitochondrial control region. This is probably due to the faster elimination of haplotypes by genetic drift in isolated subpopulations from the Northeast. Indeed, genetic differentiation estimated with mitochondrial markers was rather high in this region despite quite a low genetic distance among sampling localities shown by microsatellites. The other explanation for this pattern is the lower dispersal ability of females, as some studies have suggested that males can cover a larger distance than females [77].

In conclusion, our results suggest that: (i) non-invasive sampling of faeces is a useful tool to study the genetic structure of of hazel hen population, (ii) genetic variability and genetic differentiation in this species is strongly influenced by the level of habitat fragmentation, and (ii) two distinct genetic groups exist in the Polish population, which correspond to geographical regions/ subspecies boundaries.

Acknowledgements

This study was funded by the Polish State Committee for Scientific Research (grant no.: 2P06L01127). The authors kindly thank all those involved in the collection of samples. We also thank an anonymous reviewer for helpful comments on a previous version of this manuscript.

References

[1] Cramp S., Simmons K.E.L., (Eds.), Handbook of the Birds of Europe, the Middle East, and North Africa: The Birds of the Western Palearctic, Volume II, Hawks to Bustards, Oxford University Press, 1980

[2] Bergmann H.-H., Klaus S., Distribution, status and limiting factors of hazel grouse in central Europe, particularly in Germany, Gibier Faune Sauvage, 1994, 11, 5-32

[3] Bergmann H.-H., Klaus S., Műller F., Scherzinger W., Swenson J.E., Wiesner J., Hazel hen [Die Haselhűhner], Westarp Wissenschaften, Magdeburg, Germany, 1996, (in German)

[4] Różycki A.Ł., Keller M., Buczek T., Numbers and habitat preferences of the Hazel Grouse Bonasa bonasia in the Lasy Parczewskie forest [Liczebność i preferencje siedliskowe jarząbka Bonasa bonasia w Lasach Parczewskich], Notatki Ornitologiczne, 2007, 48, 151-162, (in Polish)

[5] Tomiałojć L., Stawarczyk T., The avifauna of Poland. Distribution, numbers and trends [Awifauna Polski. Rozmieszczenie, liczebność i zmiany], Wrocław, 2003 (in Polish)

[6] Bonczar Z., Hazel Hen Bonasa bonasia [Jarząbek Bonasa bonasia], In: Sikora A., Rohde Z., Gromadzki M., Neubauer G., Chylarecki P. (Eds.), The atlas of breeding birds in Poland 1985-2004 [Atlas rozmieszczenia ptaków lęgowych Polski 1985–2004], Bogucki Wydawnictwo Naukowe, Poznań, 2007, (in Polish)

[7] Johansen H., The avifauna of Western Siberia - Otides to Gallus [Die Vogelfauna Westsibiriens - Otides bis Gallus], J. Ornitol., 1961, 102, 237-269 (in German)

[8] Glutz von Blotzheim U.N., Bauer K., Bezzel E., Handbook of the birds of Central Europe [Handbuch der Vögel Mitteleuropas], Vol. 5, Galliformes und

Gruiformes, 2nd ed., Aula-Verlag, Wiesbaden, 1994 (in German)
[9] de Juana E., Family Tetraonidae, In: del Hoyo J., Elliott A., Sargatal J. (Eds.), Handbook of the Birds of the World, Volume 2, New World Vultures to Guineafowl, Lynx Edicions, Barcelona, 1994
[10] Dementev G.P, Gladkov N.A. (Eds.), Birds of USSR [Pticy Sovetskovo Sojuza], Sovetskaja Nauka, Moskva, 1951 (in Russian)
[11] Storch I., Status Survey and Conservation Action Plan 2000-2004: Grouse, IUCN, Gland, Switzerland and Cambridge, UK and World Pheasant Association, Reading, UK, 2000
[12] Baba Y., Fujimaki Y., Klaus S., Butorina O., Drovetski S., Koike H., Molecular population phylogeny of Hazel Grouse Bonasa bonasia in East Asia inferred from mitochondrial control-region sequences, Wildlife Biol., 2002, 8, 251-259
[13] Baba Y., Klaus S., Sun Y-H., Fujimaki Y. Molecular phylogeny and population history of the Chinese grouse and the hazel grouse, Bulletin of the Graduate School of Social and Cultural Studies, Kyushu University, 2005, 11, 77-82.
[14] Sahlsten J., Thorngren H., Hoglund J., Inference of Hazel Grouse population structure using multilocus data: a landscape genetic approach, Heredity, 2008, 101, 475–482
[15] Swenson J.E., Social organization of hazel grouse and ecological factors influencing it, PhD thesis, University of Alberta, Edmonton, USA, 1991
[16] Segelbacher G., Steinbrück G., Bird faeces for sex identification and microsatellite analysis, Vogelwarte, 2001, 41, 139-142
[17] Idaghdour Y., Broderick D., Korrida A., Faeces as a source of DNA for molecular studies in a threatened population of great bustards, Conserv. Genet., 2003, 4, 789-792
[18] Beja-Pereira A., Oliveira R., Alves P.C., Schwartz M.K., Luikart G., Advancing ecological understandings through technological transformations in noninvasive genetics, Mol. Ecol. Resou., 2009, 9, 1279-1301
[19] Taberlet P., Griffin S., Goossens B., Questiau S., Manceau V., Escaravage N., et al., Reliable genotyping of samples with very low DNA quantities using PCR, Nucleic Acids Res., 1996, 24, 3189-3194
[20] Taberlet P., Waits L.P., Luikart G., Noninvasive genetic sampling: look before you leap, Trends Ecol. Evolut., 1999, 14, 323-327
[21] Bellemain E., Taberlet P., Improved non invasive genotyping method: application to brown bear (Ursus arctos) faeces, Mol. Ecol. Notes, 2004, 4, 519-522
[22] Piggott M.P., Bellemain E., Taberlet P., Taylor A.C, A multiplex pre-amplification method that significantly improves microsatellite amplification and error rates for faecal DNA in limiting conditions, Conserv. Genet., 2004, 5: 417-420
[23] Morin P.A., Chambers K.E., Boesch C., Vigilant L., Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (Pan troglodytes verus), Mol. Ecol., 2001, 10, 1835-1844
[24] Campbell N.R., Narum S.R, Quantitative PCR assessment of microsatellite and SNP genotyping with variable quality DNA extracts, Conserv. Genet., 2009, 10, 779-784
[25] Miller C.R., Joyce P., Waits L.P, Assessing allelic dropout and genotype reliability using maximum likelihood, Genetics, 2002, 160, 357-366
[26] Valière N., Berthier P., Mouchiroud D., Pontier D., GEMINI: software for testing the effects of genotyping errors and multitubes approach for individual identification, Mol. Ecol. Notes, 2002, 2, 83-86
[27] Valière N., Bonefant C., Toigo C., Luikart G., Gaillard J-M., Klein F., Importance of a pilot study for non-invasive genetic sampling: genotyping errors and population size estimation in red deer, Conserv. Genet., 2007, 8, 69-78
[28] Regnaut S., Lucas F.S., Fumagalli L., DNA degradation in avian faecal samples and feasibility of non-invasive genetic studies of threatened capercaille populations, Conserv. Genet., 2006, 7, 449-453
[29] Hall T.A., BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp. Ser., 1999, 41, 95-98
[30] Segelbacher G., Paxton R.J., Steinbrueck G., Trontelj P., Storch I., Characterisation of microsatellites in capercaillie (Tetrao urogallus) (AVES), Mol. Ecol., 2000, 9, 1934-1935
[31] Piertney S.B., Höglund J., Polymorphic microsatellite markers in black grouse (Tetrao tetrix), Mol. Ecol. Notes, 2001, 1, 303-304
[32] Caizergues A., Dubois S., Mondor G., Rasplus J-F., Isolation and characterization of microsatellite loci in black grouse (Tetrao tetrix), Mol. Ecol. Notes, 2001, 1, 36-38
[33] Chambers, G.K., MacAvoy, E.S., Microsatellites: consensus and controversy, Comp. Biochem. Phys. B, 2000, 126, 455-476
[34] Valière N., Gimlet: a computer program for analysing genetic individual identification data, Mol. Ecol. Resour., 2002, 2, 377-379

[35] Nei M., Molecular Evolutionary Genetics, Columbia University Press, New York, 1987
[36] Rozas J., Sánchez-DelBarrio J.C., Messeguer X., Rozas R., DnaSP, DNA polymorphism analyses by the coalescent and other methods, Bioinformatics, 2003, 19, 2496-2497
[37] Excoffier L., Lischer H.E.L., Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour., 2010, 10, 564-567
[38] Fu Y-X., Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection, Genetics, 1997, 147, 915-925
[39] Rogers A.R., Harpending H., Population growth makes waves in the distribution of pairwise genetic differences, Mol. Biol. Evol., 1992, 9, 552-569
[40] Schneider S., Excoffier L., Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA, Genetics, 1999, 152, 1079-1089
[41] Harpending H.C., Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution, Hum. Biol., 1994, 66, 591-60
[42] Excoffier L., Smouse P.E., Quattro J.M., Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data, Genetics, 1992, 131, 479-491
[43] Tamura K., Dudley J., Nei M., Kumar S., MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0, Mol. Biol. Evol., 2007, 24, 1596-1599
[44] Petit R.J., El Mousadik A., Pons O., Identifying populations for conservation on the basis of genetic markers, Conserv. Biol., 1998, 12: 844–855
[45] Nei M., Roychoudhury A.K., Sampling variances of heterozygosity and genetic distance, Genetics, 1974, 76, 379-390
[46] Paekal R., Smouse P.E., GenAlEx V6: Genetic Analysis in Excel. Population genetic software for teaching and research, Mol. Eco. Not., 2006, 6, 288-295
[47] Goudet J., Fstat version 1.2: a computer program to calculate F statistics, J. Hered., 1995, 6, 485-486
[48] Raymond M., Rousset F., GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism, J. Hered., 1995, 86, 248-249
[49] Rousset F., Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux, Mol. Ecol. Resour., 2008, 8, 103-106
[50] Weir B.S., Cockerham C.C, Estimating F-statistics for the analysis of population structure, Evolution, 1984, 38, 1358-1370
[51] Pritchard J.K., Stephens M., Donnelly P., Inference of population structure using multilocus genotype data, Genetics, 2000, 155, 945-959
[52] Evanno G., Regnaut S., Goudet J., Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, Mol. Ecol., 2005, 14, 2611-2620
[53] Earl D.A., von Holdt B.M., STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method, Conserv. Gen. Res, 2012, 4, 359-361
[54] Jakobsson M., Rosenberg N.A., CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure, Bioinformatics, 2007, 23, 1801-1806
[55] Rosenberg N.A., DISTRUCT, a program for the graphical display of population structure, Mol. Ecol. Not., 2004, 4, 137-138
[56] Piggott M.P., Taylor A.C., Remote collection of animal DNA and its applications in conservation management and under standing the population biology of rare and cryptic species, Wildlife Res., 2003, 30, 1-13
[57] Wattier R., Engel C.R., Saumitou-Laprade P., Valero M., Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in Gracilaria gracilis (Rhodophyta), Mol. Ecol., 1998, 7, 1569-1573
[58] Broquet T., Ménard N., Petit E., Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effect on amplification success and genotyping error rates, Conserv. Genet., 2007, 8, 249-260
[59] Hoffman J.L., Amos W., Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion, Mol. Ecol., 2005, 14, 599-612
[60] Kohn M.H., York E.C., Kamradt D.A., Haught G., Sauvajot R.M., Wayne R.K., Estimating population size by genotyping faeces, Proc. R. Soc. Lond. [Biol], 1999, 266, 657-663
[61] Bayes M., Smith K., Alberts S., Altmann J., Bruford M., Testing the reliability of microsatellites typing from faecal DNA in the savannah baboon, Conserv. Genet., 2000, 1, 173-176

[62] Lathuillière M., Ménard N., Gautier-Hion A., Crouau-Roy B., Testing the reliability of noninvasive genetic sampling by comparing analyses of blood and fecal samples in Barbary macaques (Macaca sylvanus), Am. J. Primatol., 2001, 55, 151-158

[63] Lucchini V., Fabbri E., Marucco F., Ricci S, Boitani L., Randi E., Noninvasive molecular tracking of colonizing wolf (Canis lupus) packs in the western Italian Alps, Mol. Ecol., 2002, 11, 857-868

[64] Maudet C., Luikart G., Dubray D., von Hardenberg A., Taberlet P., Low genotyping error rates in wild ungulate faeces sampled in winter, Mol. Ecol. Notes, 2004, 4, 772-775

[65] Monteiro L., Bonnemaison D., Vekris A., Petry K.G., Bonnet J., Vidal R., et al., Complex polysaccharides as PCR inhibitors in feces: Helicobacter pylori model, J. Clin. Microbiol., 1997, 35, 995-998

[66] Liukkonen-Antilla T., Rätti O., Kvist L., Helle P., Orell M., Lack of genetic structuring and subspecies differentiation in the capercaillie (Tetrao urogallus) in Finland, Ann. Zool. Fennici, 2004, 41, 619-633

[67] Rutkowski R., Niewęgłowski H., Dziedzic R., Kmieć M., Goździewski J., Genetic variability of Polish population of the Capercaillie Tetrao urogallus, Acta Ornithologica, 2005, 40, 27-34

[68] Caizergues A., Ratti O., Helle P., Rotelli L., Ellison L., Rasplus J.-F., Population genetic structure of male black grouse (Tetrao tetrix L.) in fragmented vs. continuous landscapes, Mol. Ecol., 2003, 12, 2297-2305

[69] Segelbacher G., Hoglund J., Storch I., From connectivity to isolation: genetic consequences of population fragmentation in Capercaillie across Europe, Mol. Ecol., 2003, 12, 1773-1780

[70] DiBattista J.D., Patterns of genetic variation in anthropogenically impacted populations, Conserv. Genet., 2008, 9, 141-156

[71] Åberg J., Jansson G., Swenson J. E., Angelstam P., The effect of matrix on the occurrence of hazel grouse (Bonasa bonasia) in isolated habitat fragments, Oecologia, 1995, 103, 265-269

[72] Segelbacher G., Storch I., Capercaillie in the Alps:genetic evidence of metapopulation structure and population decline, Mol. Ecol., 2002, 11, 1669-1677

[73] Taberlet P., Fumagalli L., Wust-Saucy A.-G., Cosson J.-F., Comparative phylogeography and postglacial colonization routes in Europe, Mol. Ecol., 1998, 7, 453-464

[74] Marmi J., López-Giráldez F., Macdonald D.W., Calafell F., Zholnerovskaya E., Domingo-Roura X., Mitochondrial DNA reveals a strong phylogeographic structure in the badger across Eurasia, Mol. Ecol., 2006, 15, 1007-1020

[75] Segelbacher G., Piertney S., Phylogeography of the European Capercaillie (Tetrao urogallus) and its implications for conservation, J. Ornithol., 2007, 148, 269-274

[76] Duriez O., Sachet J.M., Menonni E., Pidancier N., Miqulet C., Taberlet P., Phyleography of the capercaillie in Eurasia: what is the conservation status in the Pyrenees and Cantabrian Mounts? Conserv. Genet., 2007, 8, 513-526

[77] Montadert M., Leonard P., Post-juvenile dispersal of Hazel Grouse (Bonasa bonasia) in an expanding population of the Southeastern French Alps, Ibis, 2006, 148, 1-13

Appendix

	TUT1	TUT2	TUT3	TUT4	TTT1	BG4	BG6	BG18
n PCR	276	276	276	276	276	276	276	276
ADO	0.00	0.00	0.00	0.00	0.00	0.00	**0.129**	0.00
FA	0.00	0.00	0.00	0.00	0.00	0.00	**0.035**	0.00

Table S1. Per locus error rates. n PCR – total number of PCRs, ADO – frequency of allelic drop-out, FA – false allele rates.

Locus	Size range	*A*	H_O	H_E	*HWE*	F_{IS}
TUT1	179–207	7	0.381	0.698	<0.001	0.457*
TUT2	171–207	10	0.331	0.486	<0.001	0.323*
TUT3	192–228	9	0.712	0.794	0.3035	0.108
TUT4	130–166	7	0.288	0.327	0.0806	0.124
TTT1	202–250	12	0.726	0.834	0.0278	0.133*
Bg6	198–322	17	0.767	0.856	<0.001	0.108*
Bg10	186	1	–	–	–	–
Bg15	201	1	–	–	–	–
Bg18	129–161	9	0.598	0.685	<0.001	0.131*
Bg4	172–200	8	0.590	0.714	<0.001	0.179*
Mean		8.1	0.549	0.674	<0.001	0.190*

Table S2. The characterization of microsatellite loci amplified in the total sample of *B. bonasa* (n=118). Size range – size of fragments obtained during PCR, including flanking region; *A* – number of alleles; H_O – heterozygosity observed; H_E – heterozygosity expected; HWE – *P*-values for HWE exact test for heterozygote deficiency/excess, F_{IS} – fixation index. * – significant F_{IS} value after Bonferroni correction (Bonferroni corrected P-value after 160 randomization at α=0.05 was 0.00625).

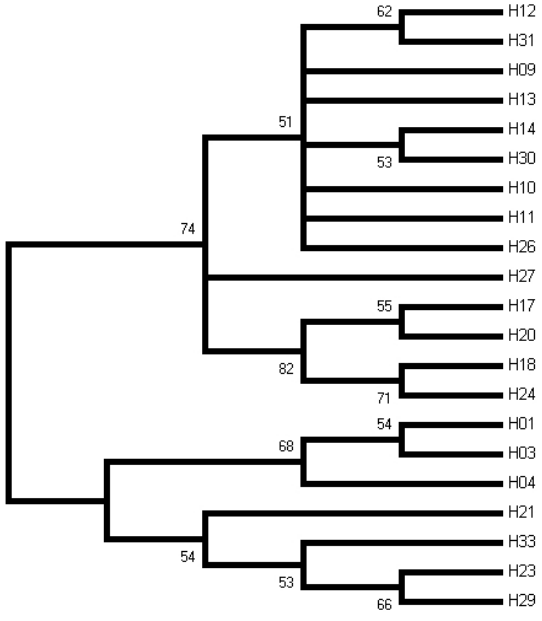

Figure S1. Neighbour-joining tree computed with mitochondrial haplotypes of hazel hens from Northeastern and Southern Poland. Numbers at the knots represents bootstrap values (1000 replicates).