

## ORIGINAL ARTICLE

## Inference of hazel grouse population structure using multilocus data: a landscape genetic approach

J Sahlsten<sup>1</sup>, H Thörnigren<sup>2</sup> and J Höglund<sup>1</sup><sup>1</sup>Population Biology and Conservation Biology/Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden and <sup>2</sup>Animal Ecology/Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden

In conservation and management of species it is important to make inferences about gene flow, dispersal and population structure. In this study, we used 613 georeferenced tissue samples from hazel grouse (*Bonasa bonasia*) where each individual was genotyped at 12 microsatellite loci to make inference on population genetic structure, gene flow and dispersal in northern Sweden. Observed levels of genetic diversity suggest that Swedish hazel grouse do not suffer loss of genetic diversity compared with other grouse species. We found significant  $F_{IS}$  (deviation from Hardy–Weinberg expectations) over the entire sample using jack-knifed estimators over loci, which is most likely explained by a Wahlund effect. With the use of spatial autocorrelation methods, we detected significant isolation by distance among individuals. Neighbourhood size was estimated in

the order of 62–158 individuals corresponding to a dispersal distance of 950–1500 m. Using a spatial statistical model for landscape genetics to infer the number of populations and the spatial location of genetic discontinuities between these populations we found indications that Swedish hazel grouse are divided into a northern and a southern population. We could not find a sharp border between these two populations and none of the observed borders appeared to coincide with any potential geographical barriers. These results imply that gene flow appears somewhat unrestricted in the boreal taiga forests of northern Sweden and that the two populations of hazel grouse in Sweden may be explained by the post-glacial reinvasion history of the Scandinavian Peninsula.

*Heredity* (2008) **101**, 475–482; doi:10.1038/hdy.2008.94; published online 1 October 2008

**Keywords:** geneland; Monte Carlo Markov chain; Bayesian clustering; isolation by distance; genetic discontinuities

## Introduction

All species are subdivided, display population structure and are to some extent influenced by spatially heterogeneous landscapes. Spatial heterogeneity affects, among other things, dispersal and, consequently, gene flow (Holderegger *et al.*, 2006). A central issue is whether changes in landscape features create barriers to gene flow and as such induce population structure. In conservation and management of species, it is important that conservation or management units are defined on relevant biological criteria rather than being arbitrarily defined (Bonin *et al.*, 2007).

To understand dispersal dynamics at a local scale, it is important to understand the patterns and processes of gene flow at larger scales (Manel *et al.*, 2003). For example, in studies of patch occupancy dynamics, such as in metapopulation models, dispersal is a crucial factor (Taylor *et al.*, 1993; Hanski, 2001; Hanski and Ovaskainen, 2003). However, the available knowledge of dispersal is often sparse and becomes a limiting factor when modelling movement patterns or patch occupancy. At

large scales, traditional mark-recapture methods for estimating dispersal become unfeasible because impossibly large number of individuals need to be monitored. This problem may be overcome by using molecular genetic techniques (Waser and Strobeck, 1998). Landscape genetics has emerged from a combination of spatial statistics, molecular genetic techniques and landscape ecological theories (Manel *et al.*, 2003; Holderegger and Wagner, 2006). This approach uses individuals as the study unit and attempts to address whether geographic and environmental structures affect gene flow and genetic structure.

Classically, genetic studies have used *a priori* defined populations when studying gene flow. However, identifying populations in advance may be undesirable due to potential biases arising from unidentified migrants and cryptic spatial structure (Rousset, 1999; Sumner *et al.*, 2001; Manel *et al.*, 2003). Analytical tools using Bayesian clustering algorithms to detect population structure use individuals as the study unit and thus there is no need of prior knowledge of discrete populations in advance. In such approaches, it is possible to assign genotyped individuals to  $k$  populations, where  $k$  may be unknown (Pritchard *et al.*, 2000; Dawson and Belkhir, 2001). Treating allele frequency as a random variable, these methods find units defined by Hardy–Weinberg equilibrium, and any signal of linkage disequilibrium is assumed to be due to population structure rather than physical linkage (Pritchard *et al.*, 2000). Thus it is possible

Correspondence: Dr J Sahlsten, Population Biology and Conservation Biology/Evolutionary Biology Centre, Uppsala University, Norbyväg 18D, SE-752 36 Uppsala, Sweden.

E-mail: jonas.sahlsten@ebc.uu.se

Received 1 October 2007; revised 26 June 2008; accepted 11 July 2008; published online 1 October 2008

to make inferences about the genetic structure of a sample of individuals (Pritchard *et al.*, 2000). Unknown parameters are integrated out using Markov chain Monte Carlo (MCMC) methods. In a recent expansion of Bayesian detection of population structure, the spatial locations of genetic discontinuities between populations are supposed to be spatially organized through the so-called coloured Poisson–Voronoi tessellation (Ripley 1981; Dupanloup *et al.*, 2002; Guillot *et al.*, 2005b), and from thematic maps, land cover features associated with genetic discontinuities can be identified.

In continuous populations, neighbourhood size is a basic population entity (Wright 1943; Slatkin and Barton, 1989), and is usually defined as  $NS = 4\pi D\sigma^2$ , where  $D$  is the population density and  $\sigma$  is the mean axial square distance between related individuals, the  $\sigma$  parameter determines how much genetic differentiation increases with distance (Rousset, 2000; Sumner *et al.*, 2001). From the definition of neighbourhood size and given knowledge about effective density it is thus possible to estimate dispersal distance (Rousset, 2000). With theoretical models of isolation by distance it can be shown that kinship and a distance measure, called Rousset's  $a$  described in Rousset (2000), are expected to vary approximately linearly with the logarithm of distance (Rousset 1997; 2000). Thus, neighbourhood size can be approximated by using kinship and Rousset's  $a$  coefficients (Hardy and Vekemans, 2002). Simulation studies have shown that estimates of  $a$  are fairly robust to biases introduced by temporal changes in dispersal, density reduction and spatial expansions with constant density but less so to variation in mutation rate and density increases in the recent past (Leblois *et al.*, 2003; 2004).

Previous studies of a range of organisms have used a combination of the landscape genetic approach and genetic estimates of dispersal distance to infer levels of population structure and gene flow in wild animals (for example, Coulon *et al.*, 2006; Fontaine *et al.*, 2007; Janssens *et al.*, 2008). In this paper, we attempt a similar approach using data from the hazel grouse (*Bonasa bonasia*), which is a forest breeding bird species with poor dispersal ability and ecological features that indicate a wide-spread, but patchy distribution (Swenson, 1991a, b; Åberg *et al.*, 2000). In Sweden as well as throughout most of its range, the hazel grouse distribution is tied to the occurrence of continuous old-growth coniferous forest. Previous genetic studies have used a phylogeographic approach to detect broad patterns of genetic structure within the entire range of the species (Baba *et al.*, 2002). There are, however, hitherto no studies of genetic structure at the regional level, that is at intermediate geographic scales. The species is ecologically well studied, and direct estimates of dispersal using radio-telemetry are available (Swenson, 1991a, b; Swenson and Danielsen, 1995). Moreover, density measures and sex ratio estimates from a former study (Swenson, 1991a) allow the calculation of effective density to indirectly estimate dispersal distance of hazel grouse.

In this study, we used 12 microsatellite markers to study genetic differentiation, dispersal and population structure of hazel grouse at a regional scale. The aims of the present study were the following: (1) to quantify basic levels of genetic diversity in Swedish hazel grouse, specifically whether genetic distance between individuals increases with geographic distance; (2) to establish

the genetic neighbourhood size and estimate gene flow and hence effective dispersal, and we compare these estimates to what is known from ecological studies; and (3) to investigate the landscape genetic pattern of hazel grouse in northern Sweden. In doing this, we establish the number of populations (units in Hardy–Weinberg and linkage equilibrium) and study whether there are any landscape features that coincide with any genetic discontinuities. Hence this study is an important contribution to the general understanding of how geographic and environmental structures at large scale affect gene flow, dispersal patterns and population structure of a forest-breeding bird species sensitive to a heterogeneous landscape.

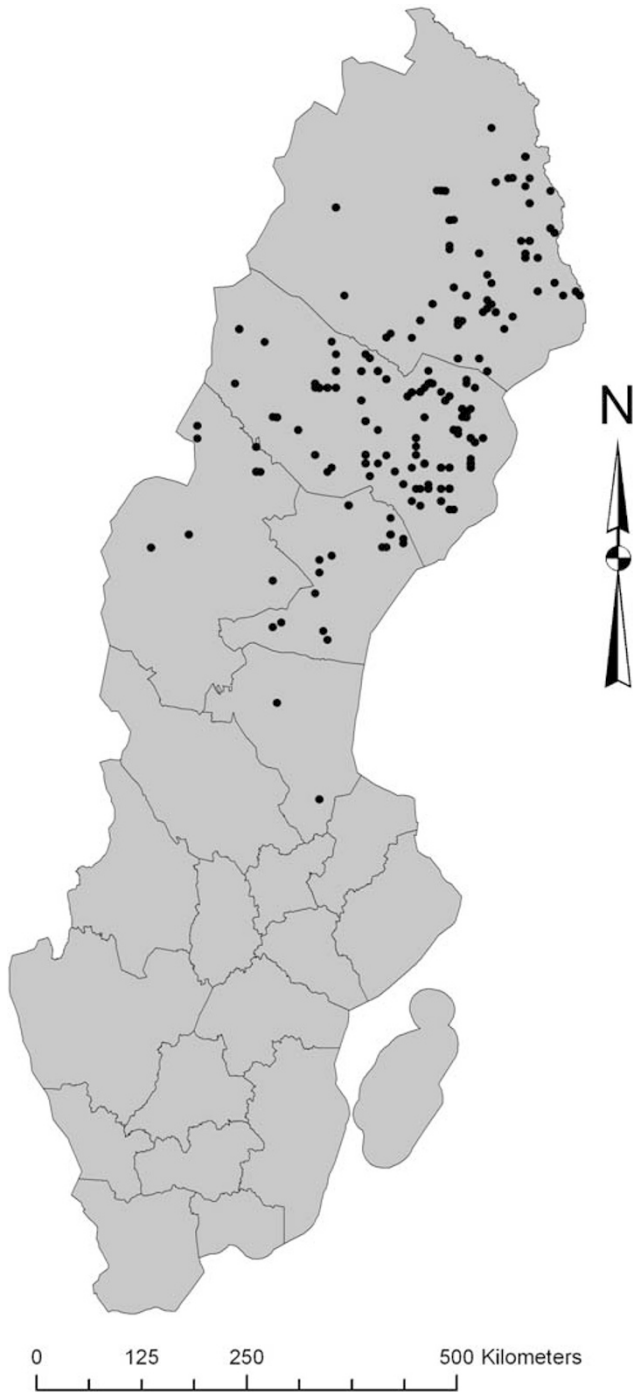
## Methods

### Georeferencing

To study genetic differences and population genetic structure at a regional scale, we used an available tissue sample collection covering the whole of northern Sweden (north of the river Dalälven), this collection being made and stored by the Swedish Museum of Natural History. The natural history museum collection consisted of hazel grouse wings collected by hunters during 1978–1986 (see Hörnfeldt (1978) for methods). From these wings, samples from 1981 and 1982 were selected because these years included wings from geographic locations that covered most of the distribution in northern Sweden (Figure 1). We genotyped each of the 613 georeferenced individuals in this collection at 12 microsatellite loci with the objective to quantify genetic diversity and to infer the number of populations and dispersal distance for hazel grouse in northern Sweden. The aim was to locate genetic boundaries and tie those to geographic structures. The sex ratio in the sample was not known, but is probably dominated by males because of the hunting technique in which a whistle pipe imitating male song is used. Males are more likely than females to respond to this stimulus (Swenson, 1991a, b). From the 613 hazel grouse wings, information on location in terms of coordinates or a name of the sampling locality was available to georeference the samples. One of the map systems (GSD-fastighetskartan) in Sweden divides the land cover into specific maps with  $5 \times 5$  km extents; thus using this system, the georeferencing was limited to a resolution of  $5 \times 5$  km. Using the available centre coordinates of each map, each wing sample was assigned a coordinate.

### Genotyping

DNA was extracted from the tissue samples by a high salt purification protocol. Genetic variation was determined at 12 microsatellite loci as described in Segelbacher *et al.* (2000) and Piertney and Höglund (2001). PCR conditions and temperature profiles followed the original publications; with slight modifications, the protocol was adjusted to allow amplification of the 12 loci in three multiplex reactions (Table 1). The markers used have been chosen and the multiplex protocol has been tested as to allow reliable genotyping of all tetraonid species and when we cannot completely rule the presence of null alleles and allelic drop out, such were kept at a minimum (J Höglund *et al.*, unpublished data). The post-PCR



**Figure 1** Distribution of the collected and georeferenced wing sample within northern Sweden (each point can contain more than one individual). Approximately 60°–68° N – 13°–20° E was covered. This is the main taiga region of Sweden, and compared with the southern part of Sweden, the forests are generally less fragmented.

products were prepared by diluting 1 µl post-PCR product with 9 µl ddH<sub>2</sub>O. Further, 2 µl of the diluted PCR product was dispensed in 8 µl of ET-ROX size marker and genotyped using MegaBase (Amersham Biosciences, Buckinghamshire, UK). Allele sizes and genotypes for each sample were scored using the software Fragment Profiler (Fragment Profiler 1.2, Amersham Biosciences, 2003). The scoring was

**Table 1** Microsatellite loci used in the study and peak filter settings used in Fragment Profiler to screen for alleles

| Marker name        | Dye name | Min. size | Max. size | Spacing |
|--------------------|----------|-----------|-----------|---------|
| <i>Multiplex 1</i> |          |           |           |         |
| ADL 230            | FAM      | 90        | 140       | 2       |
| ADL 257            | HEX      | 120       | 130       | 3       |
| ADL 142            | HEX      | 207       | 255       | 2       |
| <i>Multiplex 2</i> |          |           |           |         |
| ADL184             | NED      | 106       | 136       | 2       |
| BG 15              | HEX      | 106       | 150       | 4       |
| BG 16              | NED      | 190       | 318       | 4       |
| BG 18              | FAM      | 120       | 165       | 4       |
| <i>Multiplex 3</i> |          |           |           |         |
| LEI098             | NED      | 135       | 165       | 2       |
| TUT1               | NED      | 178       | 202       | 4       |
| TUT2               | HEX      | 160       | 215       | 4       |
| TUT3               | FAM      | 190       | 230       | 4       |
| TUT4               | FAM      | 110       | 200       | 4       |

Min. and max. size is given in nucleotides and repeat size in base pairs.

performed by implementation of a peak filter (Table 1). To check for discrepancies between peak filter results and genotype data, histograms of peak frequencies and bins were used to manually search for peaks indicating alleles.

#### Data analyses

Using the Excel add-in Micro-Satellite Toolkit (Park, 2001) the data set was checked to confirm that it did not contain non-numeric, non-integer or negative values. Descriptive genetic diversity was quantified as expected heterozygosity ( $H_E$ ), observed heterozygosity ( $H_O$ ), number of alleles at each loci and allele frequencies by using the Micro-Satellite Toolkit. Tests for deviations from Hardy–Weinberg equilibrium and linkage disequilibrium within and population differentiation among genetic clusters (see below) were performed using Genepop on the web (<http://genepop.curtin.edu.au/>). Confidence limits for  $F_{IS}$  were obtained with the jack-knifed estimates in Genetix (Belkhir *et al.*, 2000).

Spatial structure and isolation by distance between individuals were analysed with the software SPAGeDi v. 1.2 (Hardy and Vekemans, 2002). We used two estimators of genetic distance between individuals: kinship (Loiselle *et al.*, 1995) and  $a$  (Rousset, 2000). Kinship is expected to decrease as the geographic distance between individuals increases whereas the reverse is predicted for  $a$ . Spatial autocorrelation correlograms were obtained from SPAGeDi. We obtained variance estimates for each distance class by jack-knifing over loci.

Neighbourhood size was estimated using a kinship coefficient and with Rousset's  $a$  coefficient (Rousset, 1997, 2000). With theoretical models of isolation by distance it can be shown that kinship is expected to vary approximately linearly with the logarithm of distance; in addition, an estimate of neighbourhood size can be obtained using an estimate of Rousset's  $a$  (Rousset, 1997, 2000). These theoretical models were implemented here to approximate neighbourhood size among our samples. Using kinship and Rousset's  $a$  coefficients, approximation of neighbourhood size was made. Using kinship,

$NS \approx (1-F)/b\text{-log}$ , where  $b\text{-log}$  is the slope of regression of the kinship coefficients on log distance and  $F$  is a mean jack-knifed estimate of the inbreeding coefficient (Hardy and Vekemans, 2002). With Rousset's  $a$  coefficient the neighbourhood size was approximated by the inverse of slope of regression curve,  $NS \approx 1/b\text{-log}$ , where  $b\text{-log}$  is the slope of regression of  $a$  on log distance. Significance of the regression slopes were tested by 105 random permutations of individual locations (similar to a Mantel test) in SPAGeDi (Hardy and Vekemans, 2002).

Neighbourhood size is usually defined as  $NS = 4\pi D\sigma^2$ , where  $D$  is the population density and  $\sigma$  is the mean axial square distance; the  $\sigma$  parameter determines how much genetic differentiation increases with distance (Rousset, 2000; Sumner *et al.*, 2001). Thus, it is possible to estimate dispersal distance if density and neighbourhood size are known. With available data on male hazel grouse density and information that there were 40 percent more males in a study area located near Grimsö research station in south central Sweden (Swenson, 1991a,b), effective density could be estimated. Hence, with estimates of neighbourhood size and effective density, dispersal distances were estimated using both kinship and Rousset's  $a$ .

To calculate dispersal distance, we used the density estimates and sex ratio in Swenson (1991a,b) to calculate the effective populations size  $N_e$  using equation 1. To obtain effective density  $D_e$  we divided  $N_e$  with the size of the study area in Swenson (1991a,b).

$$N_e = \frac{4N_m * N_f}{N_m + N_f} \quad (1)$$

To determine population structure and the location of possible genetic boundaries, we used the R package Geneland (R, Development Core team, 2005; Guillot *et al.*, 2005a). This program infers population structure and genetic boundaries by using a Bayesian clustering method (Guillot *et al.*, 2005b, building on work by Pritchard *et al.*, 2000). Georeferenced multi-locus genotyped individuals of unknown origin were (probabilistically) assigned to a population by the use of a Bayesian cluster model, implemented using MCMC methods (Guillot *et al.*, 2005b). The assumption in this model is that each individual is a member of one population at Hardy-Weinberg equilibrium. Within populations, individuals are assumed to be randomly distributed and the linkage disequilibrium due to population structure may prevail between loci. The allele frequencies are assumed to follow independent Dirichlet distributions. Furthermore, populations are assumed to be organized by Poisson-Voronoi tessellation (Dupanloup *et al.*, 2002).

To implement the MCMC method, both priors and initial values for parameters have to be provided. For the number of populations, we assumed a uniform prior between 1 and 100. Each MCMC run was initialized in a state with 10 populations. As no iteration of the MCMC ever reached 100 populations, identical results would have been obtained with flat priors over larger ranges. With the parameters provided, the MCMC model was replicated 1 000 000 times with a thinning of 100 (each one-hundredth replication was stored for analysis). To account for uncertainty in the positioning of individuals, we used a prior additive noise blurring of coordinates in the MCMC model (Guillot *et al.*, 2005b). With the

estimated number of panmictic populations from the first run, the model was re-run with a fixed number of two populations, and Voronoi tessellation of observed genetic data resulted in maps of posterior probabilities of population membership. To conclude if any of the potential barriers (for example, rivers, mountains or main roads) coincided with genetic discontinuities, a thematic map (1:200 000) including the potential barriers was compared visually with the tessellation map, including genetic discontinuities.

We also used the model-based clustering algorithm implemented in Structure v. 2.1 (Pritchard *et al.*, 2000; Falush *et al.*, 2003) to find the most likely number of populations ( $k$ ). The burn-in period consisted of 100 000 replications, after which 1 000 000 MCMC iterations were run for a number of clusters from  $k = 1$  to  $k = 10$  under a model assuming admixture and allowing for correlation of allele frequencies between clusters. We adopted the approach suggested by Evanno *et al.* (2005) to calculate the most likely value of  $k$ .

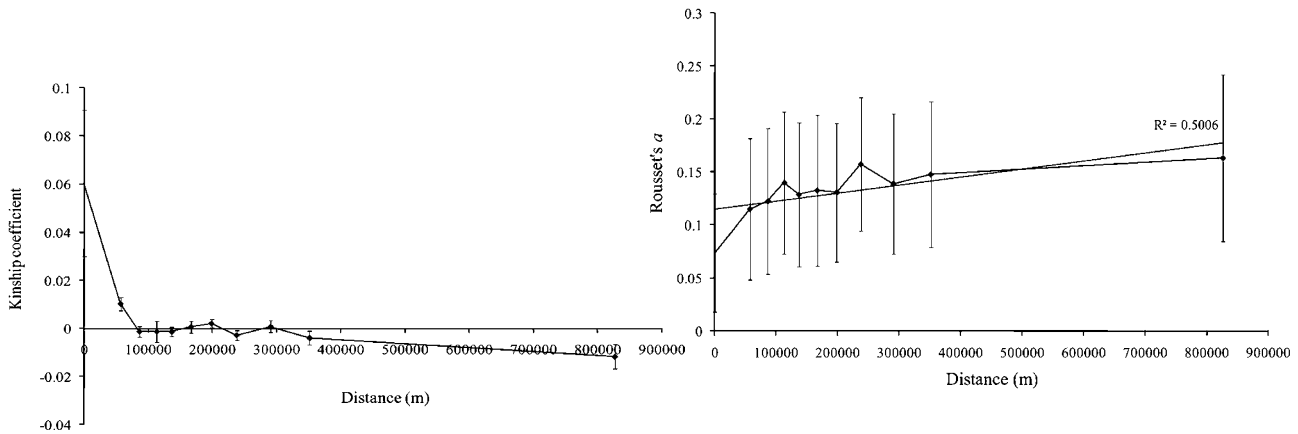
## Results

The analyses suggest that genetic diversity in Swedish hazel grouse is at the same magnitude as in several other grouse species. The primers, developed for black grouse, capercaillie and chicken, were suitable for hazel grouse as well. Using Fragment Profiler and Microsatellite Toolkit to screen the data it was concluded that two of the microsatellite makers (ADL257 and TUT2) were monomorphic among the sampled hazel grouse and they were consequently removed from the data set.

The expected unbiased heterozygosity for the Swedish hazel grouse population was  $0.561 \pm 0.066$  s.d. and the observed heterozygosity was  $0.466 \pm 0.007$  s.d., suggesting that the Swedish hazel grouse do not suffer from loss of genetic diversity compared with other grouse species. Using jack-knife estimators over loci, mean  $F_{IS}$  was estimated to  $0.1632 \pm 0.0545$  s.e. The average number of alleles was  $10.5 \pm 6.04$  s.d. We found evidence of two genetic clusters (see below).  $F_{IS}$  was significantly different from zero in both of the clusters (0.160 in the 'northern' and 0.216 in the 'southern' cluster) but no loci were found in significant linkage disequilibrium within clusters after Bonferroni correction. Population differentiation among the clusters was weak but significant ( $F_{ST} = 0.0052$ ,  $\chi^2 = 43.270$ , d.f. = 20,  $P = 0.0019$ ).

The kinship coefficient decreased with geographic distance and Rousset's  $a$  increased with distance (both significantly at  $P = 0.000$ , Figure 2), allowing the calculation of neighbourhood size and corresponding genetic dispersal distances.

The results from the calculation of the neighbourhood size give knowledge about effective density and thereby it was possible to estimate dispersal distances. Neighbourhood size determined by the kinship coefficient was 158.27 (min = 117.52, max = 261.63) and from Rousset's  $a$ , neighbourhood size was estimated to 62.85 (min = 39.78, max = 149.64). We estimated that the effective population size in the study area (195 hectares) of Swenson (1991a) was  $N_e = 10.72$  corresponding to a  $D_e = 5.5$ . Solving  $NS = 4\pi D_e \sigma^2$  for the mean axial parent-offspring dispersal distance,  $\sigma$  yielded an estimate of 1514 m (min = 1304, max = 1946) and 954 m (min = 759, max = 1472), respectively.



**Figure 2** Spatial autocorrelation correlogram for estimated kinship within 10 distance classes (left) and correlogram for the Rousset's  $a$  coefficient with 10 distance classes (right). Error bars represent s.e.

Using Geneland we found the number of populations to be  $k=2$ , with posterior probability very close to 1. Maps of the probability for population adherence of any pixel within northern Sweden suggested that one population had a more southern distribution and the other a more northern, but no sharp border between the populations was evident (Figure 3). We calculated the modal population for each pixel and this map also shows that there was no clear south–north subdivision between the two populations (Figure 3). From the comparison of the genetic discontinuities and a thematic map, no evidence of geographical barriers could be concluded. With Structure we also found that the most likely number of clusters ( $k$ ) was 2 (Table 2). However, the sampled individuals were mostly not assigned unambiguously to either cluster, and the posterior probability of cluster membership seemed to vary quite smoothly over individuals (Figure 4).

## Discussion

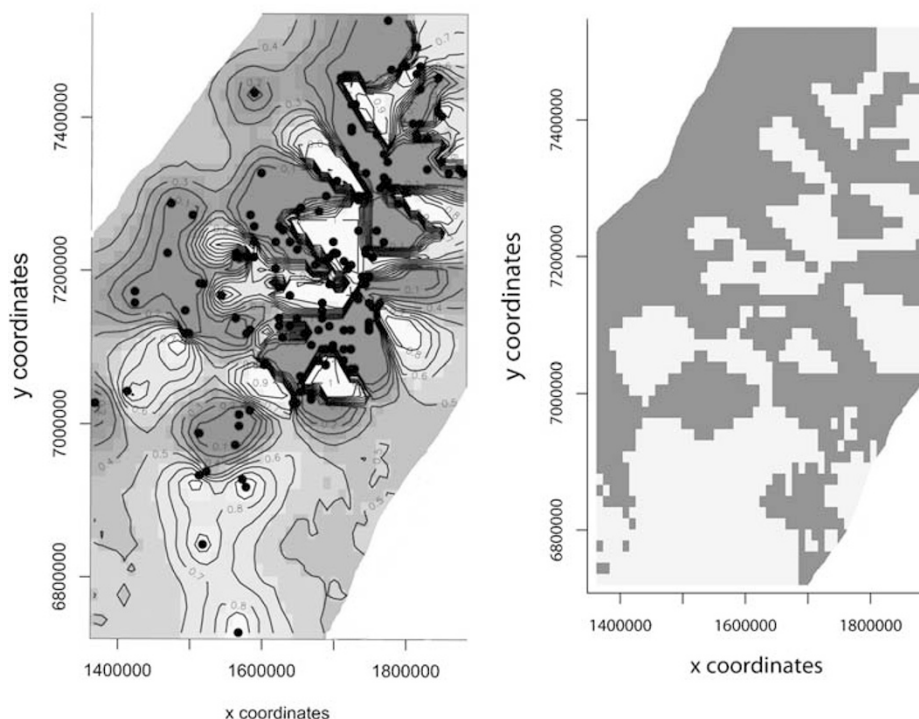
We found levels of genetic diversity in Swedish hazel grouse to be of the same magnitude as in other grouse species genotyped at many of the same microsatellite loci (capercaillie *Tetrao urogallus*: Segelbacher *et al.* (2003); black grouse *T. tetrix*: Höglund *et al.* (2007)). However, the levels of diversity are not strictly comparable between species as it is well known that not only amplification probability but also polymorphism decreases when loci developed in one species are cross-amplified in phylogenetically related species (Primmer *et al.*, 1996). The TUT primers were originally cloned in capercaillie (Segelbacher *et al.*, 2003) and the BG primers in black grouse (Piertney and Höglund, 2001). Hence, it may be expected that for this reason, the reported levels of microsatellite genetic diversity in capercaillie and black grouse may be higher and thus not ascribable to differing population processes among the species. Nevertheless, the levels of genetic diversity reported in this study suggest that Swedish hazel grouse have substantial levels of genetic diversity and do not suffer loss of genetic diversity compared with other grouse species.

We did find a significant  $F_{IS}$  (deviation from Hardy–Weinberg expectations) in the entire sample. This is most

likely explained by a Wahlund effect (Wahlund, 1928). When two population samples are lumped and analysed for departures from Hardy–Weinberg expectations as a single unit, the number of homozygotes become artificially increased because of the hidden population structure. As Geneland and Structure analyses found strong support for two genetic populations in Swedish hazel grouse, the significant  $F_{IS}$  may be partly explained by this. However, significant heterozygote deficiency remained within clusters after individuals were assigned to one of the clusters that argue against this being the sole explanation.

Kinship coefficients decreased and genetic distance among individuals increased with geographic distance. This allowed us to calculate neighbourhood size and corresponding genetic dispersal distances of roughly 900–1500 m per generation. These estimates are comparable with estimates from what is known from ecological studies. Juvenile dispersal in Fennoscandia has been found to be 800 m (Swenson, 1991b). In southeastern French Alps, an average dispersal distance of 4 km has been concluded for post-juvenile hazel (Montadert and Léonard, 2006). The longer distance in France may be because that two males dispersed 15 and 29.4 km, respectively. Hence, the median distance of 1.6 km found in the French study is more comparable with our results.

The landscape genetic pattern of hazel grouse in northern Sweden revealed evidence of two genetic populations in Sweden: one with a more northern distribution and one more southern. Similar patterns have been found in many studies of other animals (for example, bears: Taberlet *et al.* (1995); willow warblers: Bensch *et al.* (2002); shrews: Andersson, (2004)). The proposed explanation for this north–south divide is the post-glacial reinvasion of the Scandinavian Peninsula following the retreat of the inland ice approximately 10 000 years ago. The southern limit of the last large block of inland ice was situated in mid northern Sweden at an approximate latitude of 62° N. The Scandinavian Peninsula was thus re-colonized from two directions: one through a land bridge in the southwest, the other through Finland from the northeast. When the last inland ice finally melted away, previously separated populations came into secondary contact and in some



**Figure 3** Tessellation map of the inferred populations suggesting that one population has a more northern distribution and the other has a southern distribution (I). The modal population for each pixel shows the distribution of two populations, although no sharp border is evident (II).

**Table 2** Results of the Structure analysis and the calculations to infer the number of clusters ( $k$ )

| K  | $L(K)$   | $Var[Ln(K)]$ | $L'(K)$ | $L''(K)$ | $\Delta K$ |
|----|----------|--------------|---------|----------|------------|
| 1  | -10467   | 47.3         |         |          |            |
| 2  | -10323.1 | 569.9        | 143.9   | 1475     | 2.6        |
| 3  | -11654.2 | 3496.8       | 1331.1  | 2025.8   | 0.6        |
| 4  | -10959.5 | 2110.4       | 694.7   | 568.8    | 0.3        |
| 5  | -10833.6 | 2176.2       | 125.9   | 217.2    | 0.1        |
| 6  | -10924.9 | 2544.5       | 91.3    | 7.9      | 0          |
| 7  | -11008.3 | 2907.1       | 83.4    | 94.9     | 0          |
| 8  | -10996.8 | 3088.1       | 11.5    | 57.1     | 0          |
| 9  | -11042.4 | 3347.5       | 45.6    | 197.5    | 0.1        |
| 10 | -11285.5 | 3932.3       | 243.1   | 243.1    | 0.1        |

$L(k)$ ,  $L'(k)$ ,  $L''(k)$  and  $\Delta K$  are likelihoods and deviations thereof derived from simulations in Structure 1.2 and as suggested by Evanno *et al.* (2005). The number of clusters inferred in our case was based on choosing a high  $L(k)$  and  $\Delta K$ .

instances formed hybrid zones (see references above). In the case of hazel grouse, we could not detect a sharp hybrid zone. Rather, we interpret the data as evidence of quite substantial migration and clinal variation among two genetic populations.

Both Structure and Geneland analyses suggested two populations ( $k=2$ ) but the spatial autocorrelation analyses strongly hint at an isolation by distance pattern. Furthermore, when plotting population adherence sorted by  $Q$  (the most likely population for any individual), the pattern could be interpreted as clinal variation. Both Geneland and Structure may not perform so well when there is clinal variation, isolation by distance (Pritchard and Wen, 2004) and no sharp discontinuities between

populations. We interpret the mosaic pattern of the Voronoi tessellation as the result of this clinal variation. With a strong north–south cline,  $k=2$  is a more likely result than  $k=1$ . We suggest that the two populations were indeed separated for generations by the previous sheet of inland ice and that this period of separation allowed divergence of the populations. When the two populations came in to contact again, the populations became admixed. With a Voronoi tessellation, each pixel on a map is forced to belong to one of each of the detected genetic clusters, and without a hybrid zone or a geographical divide, the pattern appears as a mosaic.

We could not find evidence of any landscape features coinciding with genetic discontinuities. Rather it seems as if hazel grouse can disperse rather freely in the boreal taiga zone of Sweden. Earlier studies have indicated that the hazel grouse is a poor disperser (Swenson, 1991b; Swenson and Danielsen, 1995; Åberg *et al.*, 2000) and avoids open land (Sahlsten, manuscript). However, Montadert and Leonárd (2006) did find that radio-tagged hazel grouse could disperse over longer distances than previously thought (mean 4 km) and also over unsuitable habitats. Given that northern Sweden to a large extent is covered by unbroken forests suitable for hazel grouse, it is perhaps not surprising that we do find evidence of substantial dispersal and no sharp genetic boundaries between the populations. The main potential physical barriers to gene flow existing within the studied area are several moderately sized rivers flowing west to east from the Scandinavian mountains to the Baltic Sea. However, none of these appear to impose any barriers to gene flow in hazel grouse. This is somewhat surprising as previous



**Figure 4** Summary plot of estimates of  $Q$  (estimated membership coefficients for each individual in the two clusters) sorted by  $Q$  as inferred from the software Structure 2.1. Each individual is represented by a single vertical line broken into two coloured segments with lengths proportional to each of the inferred clusters.

studies have suggested that rivers in Scotland can prevent gene flow in red grouse (*Lagopus lagopus scoticus*, Pierrney *et al.*, 1998).

In summary, we found that there is evidence of a population structure reminiscent of what has been found in many other Scandinavian animals with a basic north–south divide. Genetic distance increased with geographic distance between individuals. However, we could not find any evidence that geographic and environmental structures affected gene flow and dispersal patterns for the forest-breeding hazel grouse. This may suggest that the boreal taiga region of northern Sweden is in general a good habitat to sustain hazel grouse.

## Acknowledgements

We thank Andreas Rudh for valuable comments on the manuscript, Maria Quintela for help with the Structure analyses and Kate St Onge for correcting the English. Birger Hörnfeldt initiated the collection of the wing samples, and Tjelvar Odsjö kindly allowed access to the collection. Financial support came from the Swedish Environmental Protection agency (SNV) and the Swedish Research Council (VR).

## References

- Andersson A-C (2004). Postglacial Population History of the Common Shrew (*Sorex araneus*) in Fennoscandia. *Acta Universitatis Uppsaliensis*: Uppsala.
- Åberg J, Swenson JE, Andrén H (2000). The dynamics of hazel grouse (*Bonasa bonasia* L.) occurrence in habitat fragments. *Can J Zool* **78**: 352–358.
- Baba Y, Fujimaki Y, Klaus S, Butorina O, Drovetskiy S, Koike H (2002). Molecular population phylogeny of the hazel Grouse *Bonasa bonasia* in east Asia inferred from mitochondrial control-region sequences. *Wildl Biol* **8**: 251–259.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2000). GENETIX 4.03, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UPR 9060, Université de Montpellier II, Montpellier (France).
- Bensch S, Kesson S, Irwin D (2002). The use of AFLP to find an informative SNP: genetic differences across a migratory divide in willow warblers. *Mol Ecol* **11**: 2359–2366.
- Bonin A, Nicole F, Pampanon F, Miaud C, Taberlet P (2007). Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conserv Biol* **21**: 697–708.
- Coulon A, Guillot G, Cosson J-F, Angibault JMA, Aulagnier S, Cargnelutti B *et al.* (2006). Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Mol Ecol* **15**: 1669–1679.
- Dawson K, Belkhir K (2001). A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genet Res* **78**: 59–77.
- Dupanloup I, Schneider S, Excoffier L (2002). A simulation annealing approach to define the genetic structure of populations. *Mol Ecol* **11**: 2571–2581.
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.
- Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.
- Fontaine MC, Baird SJE, Piry S, Ray N, Tolley KA, Duke S *et al.* (2007). Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in Old World waters. *BMC Biol* **5**: 30.
- Fragment Profiler version 1.2 (2003). Amersham Biosciences.
- Guillot G, Mortier F, Estoup A (2005a). GENELAND: a computer package for landscape genetics. *Mol Ecol Notes* **5**: 712–715.
- Guillot G, Estoup A, Mortier F, Cosson J-F (2005b). A spatial statistical model for landscape genetics. *Genetics* **170**: 1261–1280.
- Hanski I (2001). Spatially realistic theory of metapopulation ecology. *Naturwissenschaften* **88**: 372–381.
- Hanski I, Ovaskainen O (2003). Metapopulation theory for fragmented landscapes. *Theor Popul Biol* **64**: 119–127.
- Hardy OJ, Vekemans X (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**: 618–620.
- Holderegger R, Kamm U, Gugerli F (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecol* **21**: 797–807.
- Holderegger R, Wagner HH (2006). A brief guide to landscape genetics. *Landscape Ecol* **21**: 793–796.
- Höglund J, Larsson JK, Jansman HAH, Segelbacher G. (2007). Genetic variability in European black grouse (*Tetrao tetrix*). *Conserv Genet* **8**: 239–243.
- Hörnfeldt B (1978). Ålder- och könsfördelning bland skogshöns skjutna i Västerbottens län 1977/1978. *Viltnytt* **8**: 11–16.
- Janssens X, Fontaine MC, Michaux JR, Libois R, de Kermabon J, Defourny P *et al.* (2008). Genetic pattern of the recent recovery of European otters in southern France. *Ecography* **31**: 176–186.
- Leblois R, Estoup A, Rousset F (2003). Influence of mutational and sampling factors on the estimation of demographic parameters in a ‘continuous’ population under isolation by distance. *Mol Biol Evol* **20**: 491–502.
- Leblois R, Rousset F, Estoup A (2004). Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. *Genetics* **166**: 1081–1092.
- Loiselle AB, Sork VL, Nason J, Graham C (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* **82**: 1420–1425.
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol* **18**: 189–197.

- Montadert M, Leónard P (2006). Post-juvenile dispersal of hazel grouse *Bonasa bonasia* in an expanding population of the southeastern French Alps. *Ibis* **148**: 1–13.
- Park SDE (2001). Trypanotolerance in West African cattle and the population genetic effects of selection [PhD. thesis]. University of Dublin: Ireland.
- Piertney SB, Höglund J (2001). Polymorphic microsatellite DNA markers in Black grouse (*Tetrao tetrix*). *Mol Ecol Notes* **1**: 303–304.
- Piertney SB, Maccoll ADC, Bacon PJ, Dallas JF (1998). Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Mol Ecol* **7**: 1645–1654.
- Primmer CR, Møller AP, Ellegren H (1996). New microsatellites from the pied flycatcher *Ficedula hypoleuca* and the swallow *Hirundo rustica* genomes. *Hereditas* **124**: 281–283.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Pritchard JK, Wen W (2004). *Documentation for Structure software Version 2*. Department of Human Genetics, University of Chicago: IL.
- R Development Core Team (2005). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Ripley BD (1981). *Spatial statistics*. John Wiley and Sons: New York.
- Rousset F (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**: 1219–1228.
- Rousset F (1999). Genetic differentiation within and between habitats. *Genetics* **151**: 397–407.
- Rousset F (2000). Genetic differentiation between individuals. *J Evol Biol* **13**: 58–62.
- Segelbacher G, Paxton R, Steinbrueck G, Trontelj P, Storch I (2000). Characterisation of microsatellites in capercaillie (*Tetrao urugallus*) (AVES). *Mol Ecol* **9**: 1934–1935.
- Segelbacher G, Höglund J, Storch I (2003). From connectivity to isolation: genetic consequences of population fragmentation in capercaillie across Europe. *Mol Ecol* **12**: 1773–1780.
- Slatkin M, Barton NH (1989). A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **43**: 1349–1368.
- Sumner J, Rousset F, Estoup A, Moritz C (2001). Neighbourhood size, dispersal and density estimates in the prickly forest skink (*Gnypetoscincus queenslandiae*) using individual genetic and demographic methods. *Mol Ecol* **10**: 1917–1927.
- Swenson JE (1991a). Social organization of hazel grouse and ecological factors influencing it. Grimsö forskningstation S -730 91 Riddarhyttan Sweden.
- Swenson JE (1991b). Is the hazel grouse a poor disperser? *Proc Intern Cong Union Game Biol* **20**: 347–352.
- Swenson JE, Danielsen J (1995). Seasonal movements by hazel grouse in south-central Sweden. In: Jenkins D (ed), *Proceedings of the International Grouse Symposium on Grouse*, Vol. 6, World Pheasant Association: Reading, UK, pp. 37–40.
- Taberlet P, Swenson JE, Sandegren F, Bjärvall A (1995). Localization of a contact zone between two highly divergent mitochondrial DNA lineages of the brown bear (*Ursus arctos*) in Scandinavia. *Conserv Biol* **9**: 1255–1261.
- Taylor PD, Fahrig L, Henein K, Merriam G (1993). Connectivity is a vital element of landscape structure. *Oikos* **68**: 571–573.
- Wahlund S (1928). Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungslhre aus betrachtet. *Hereditas* **11**: 65–105.
- Waser PM, Strobeck C (1998). Genetic signatures of inter-population dispersal. *Trends Ecol Evol* **13**: 43–44.
- Wright S (1943). Isolation by distance. *Genetics* **28**: 114–138.