UPPSALA
UNIVERSITET

Genetic diversity and differentiation in Hazel grouse (*Bonasa bonasia*)

A comparison between populations at an expanding range in the French Alps and long-term stable populations in northern Sweden

Jani Rózsa

Degree project in biology, Master of science (2 years), 2011
Examensarbete i biologi 45 hp till masterexamen, 2011
Biology Education Centre and Department of Population Biology and Conservation Biology, Uppsala University
Supervisors: Jacob Höglund and Tanja Strand

Summary

Range expansions - as opposed to range contractions and habitat fragmentation – has seen rather recent interest in population genetics and may have some counter intuitive effects on diversity, such as gradients in allele frequencies and loss of diversity. Hazel grouse (*Bonasa bonasia*) is one of the smallest members of the grouse (*Tetraonidae*) and one of few expanding ranges of its distribution occurs in a north to south axis in the French Alps. In this study, a comparison is made of both neutral (microsatellite) and adaptive (MHC) genetic diversity and differentiation in Hazel grouse, comparing three subpopulations along the range expansion axis in France, and two subpopulations in Sweden that have a continuous stable distribution. It is asked (i) whether patterns of diversity and differentiation follows any theoretical expectations of range expansions, (ii) if patterns observed in neutral genetic diversity at microsatellite loci correspond to patterns observed in adaptive genetic diversity at MHC class IIB loci, and (iii) what conservation management decisions, in the context of establishing management units (MUs), can be made from the above. While the Swedish subpopulations had uniform neutral diversity indices and showed lower than expected isolation-by-distance, the French subpopulations showed a gradient-like decline in both neutral and adaptive genetic diversity and an increasing isolation-by-distance along the range expansion axis.

Table of contents

Introduction

Anthropogenic pressure by extensive land use has resulted in large-scale changes to landscapes and has been an ever increasing contributor to biodiversity loss since the industrial revolution, mainly through the effects of habitat fragmentation and loss (Fahrig 2003, Opdam & Wascher 2004). Sedentary species with limited dispersal capabilities are particularly vulnerable to fragmentation as this reduces landscape permeability, in many cases creating a change from continuous species distributions to a meta-population structure; with local colonization-extinction (source-sink) dynamics between suitable habitat patches (Opdam & Wascher 2004, Hanski 1991). These dynamics are governed by local recruitment, local deaths, and migration to and from habitat patches. As connectivity and habitat cohesion decreases, species in these habitat patches suffer from increased risk of local extinction through environmental and demographic stochasticity (Willi *et al.* 2006). Ultimately, this degradation may lead to range-wide extinction.

While the demographic impacts of habitat fragmentation have been subjected to extensive theoretical and to some extent empirical inquiry, less is known about how genetic factors contribute to the subsistence or demise of the colonization-extinction dynamics of meta-populations and species as a whole. Current debate has focused on the potential of genetic effects to accentuate demographic declines, and the possibility of species extinction as a consequence of the depletion of genetic diversity (Spielman *et al.* 2004, Höglund 2009).

At the other end of the spectrum (as opposed to range contractions and their often subsequent reduction in genetic diversity), range expansions and its ensuing population growth linked to the increased available habitat may have some counter intuitive effects on genetic diversity. Some theoretical work in the form of simulations has been done on this subject (Excoffier 2004, Excoffier *et al.* 2009), which allows for testable predictions in natural systems experiencing range expansion. It is predicted that under a stepping-stone colonization pattern, gene frequency gradients, serial bottlenecks and progressively increased population differentiation from the original deme are likely outcomes.

Hazel grouse (*Bonasa bonasia*) is one of the smallest members of the grouse family (*Tetraonidae*) and thus has limited dispersal capabilities and narrow habitat requirements. It is dependent on dense old-growth or managed coniferous-deciduous forest (Storch 2000). According to IUCN (the International Union for Conservation of Nature, 1996) it is classified as Lower Risk (least concern) overall, but red-listed in some central and southern European countries including France. Its distribution range is mainly continuous across large parts of Eastern Europe while declines and range contractions have occurred during the last century over large parts of central and Western Europe (Figure 1). Many remaining populations are scattered and small.

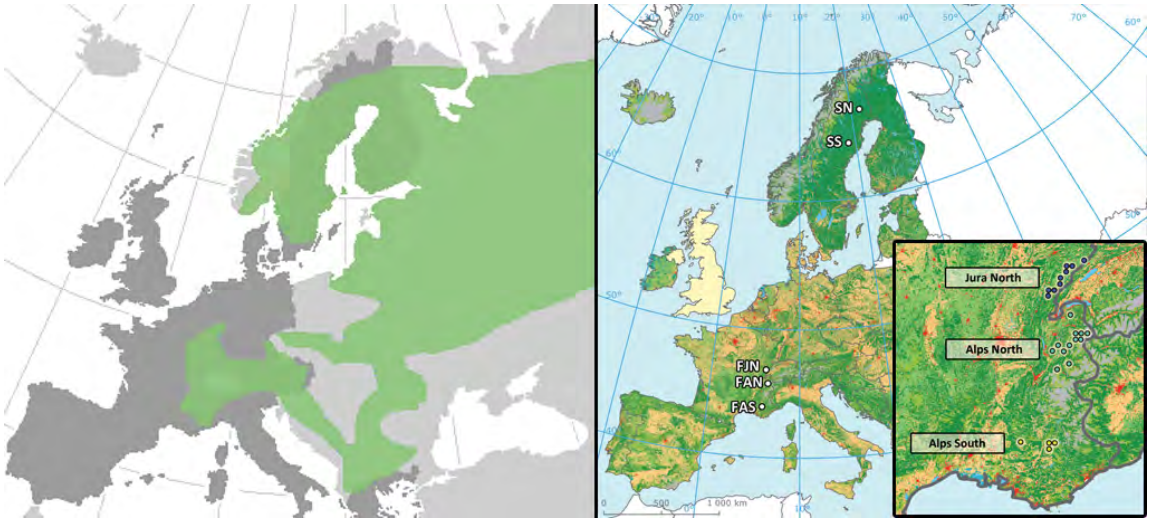

Figure 1. Hazel grouse (*Bonasa bonasia*) current distribution pattern in Europe (left map) and sampling locations (SN: Sweden North, SS: Sweden South, FJN: France Jura North, FAN: France Alps North, FAS: France Alps South) of this study (right map, folded right map: red, urban areas; yellow, arable land; green, forested land and semi-natural vegetation). Left map (Copyright holder: European Union 1995-2011), right map modified from EEA 2010 (Copyright holder: European Environment Agency), reproduction authorised.

In southeastern France along the French Alps one of few expanding ranges of Hazel grouse exist, mainly due to a few decades of land use change; abandonment of grazing of pastures leading to spontaneous reforestation in the area (Montadert and Léonard 2003). A research program has been established to monitor population dynamics along the range expansion axis from Jura North to the expansion front in Alps South. The northern Jura population (FJN) is estimated to be long-term demographically stable, albeit at a low density. The southern Alp population (FAS) is rather recently established in the 1950s and has reached a high density in the 1970s, and is assumed to have a positive intrinsic growth rate due to high survival and a slight net emigration to nearby populations (Montadert and Léonard 2003). Investigating the genetic structure of these populations may contribute to the elucidation of conservation management decisions already made from the observed population dynamic patterns and in ascertaining criteria for the establishment of evolutionarily significant units (ESUs) and management units (MUs) (Moritz 1994) along the range expansion axis of Hazel grouse.

Many molecular marker types have been tailored for these purposes. Traditionally, neutral markers such as allozyme and microsatellite loci have been the most widely used to estimate population-level genetic diversity (Segelbacher *et al.* 2000, Höglund 2009). Microsatellite markers are non-coding, codominant, have relatively high mutation rate, and evolve by step-wise mutation making them ideal for most population genetic studies.

Although neutral genetic diversity may provide evidence of historical demographic processes (population connectivity and subdivision, past reductions in gene flow, genetic drift and inbreeding), it is often an insufficient proxy for selection and the future adaptive evolutionary potential of populations in a changing environment. As neutral markers are non-coding they have no direct effect on individual fitness (Holderegger *et al.* 2006). Neutral genetic diversity on the population level is in a large part only governed by mutation, recombination, gene flow and genetic drift, and is only affected by selection in the few cases when in linkage

disequilibrium with loci under selection or when otherwise affected on a genome-wide level (DeWoody and DeWoody 2005). Conversely, adaptive genetic diversity constitutes all the coding parts of the genome and is on the population level affected by all neutral processes as well as selection, which in itself is a mechanistically complex process. In meta-analyses of population genetic studies, few general conclusions have been made about the relationship between neutral and adaptive genetic diversity (Reed and Frankham 2003, Holderegger *et al.* 2006) and it may thus be of informative value to make a distinction between the two and monitor them both in natural systems.

When assessing the total genetic effects on adaptive traits, local adaptation and population viability linked to individual fitness, a quantitative trait approach has traditionally been applied. However, new methods have emerged, focusing entirely on molecular techniques and a few candidate genes with direct links to fitness. One such molecular marker that has seen extensive scrutiny during the last decades is the codominant genes of the major histocompatibility complex (MHC) (Höglund 2009). The *Mhc* genes codes for molecules involved in the adaptive immune response to pathogens, where they bind antigen peptides to be presented to T and B cells. There are two distinct classes of MHC molecules: Class I is present in all nucleated cells presenting peptides to cytotoxic T cells, while Class II is involved in the specific immune cells such as macrophages, making them specific for certain pathogens, and also allowing for communication with B cells and their production of antibodies. The peptide binding regions of the MHC molecule and its corresponding coding sequences have shown remarkably high polymorphism; among the highest in the entire vertrebrate genome with up to a hundred different alleles per locus. Although mammals have a large *Mhc* genomic region with several hundred loci, the avian *Mhc* is much smaller, constituting only a dozen loci and has thus been termed 'the minimal essential' *Mhc*. There is mounting evidence that the high allelic richness observed in these genes is maintained by strong balancing selection, for which heterozygote and rare-allele advantage have been the proposed mechanisms (Höglund 2009, Ekblom *et al.* 2010). An individual that has a higher proportion of polymorphic rather than monomorphic *Mhc* loci is likely to be better suited to cope with a wider array of parasites (heterozygote advantage), or an individual that bears a rare allele is likely to be better at coping with new strains of parasites (rare-allele advantage). Studying the *Mhc* in concert with neutral genetic markers can provide important insights to the development of conservation management plans as well as in estimating population viability.

The aim of this study was to estimate the relative genetic diversity and differentiation of Hazel grouse (*Bonasa bonasia*) along a range expansion axis in the French Alps, as compared with a long-term stable population in northern Sweden. More specifically, it was asked (i) if patterns of diversity and differentiation followed any theoretical expectations of range expansions, i.e. in the form of isolation-by-distance, bottlenecking and inbreeding (explained further in the following section), (ii) if patterns observed in neutral genetic diversity at microsatellite loci corresponded to patterns observed in adaptive genetic diversity at MHC class IIB loci, and (iii) what conservation management decisions, in the context of establishing MUs, could be made from the above.

Materials and methods

Sampling and DNA extraction

Here I used georeferenced tissue samples of Hazel grouse in order to compare neutral and adaptive genetic diversity and population genetic differentiation on both the inter- and intra-regional scale (Table 1), using one set of samples from a Swedish region, and one set of samples from a French region. Samples were genotyped and matched for both molecular marker types as far as possible, and were grouped into five putative subpopulations based on the geographic location where the samples were originally taken. Consideration was taken to what was currently known about migration and gene flow patterns in the two regions. In Sweden, two poorly differentiated genetic clusters exist; one to the north, and one to the south, with presumed little barriers to gene flow between them (Sahlsten *et al.* 2008). Thus, two population subdivisions were determined according to this scheme. In France, the colonization pattern of Hazel grouse has been studied with radio tracking methods. Mean post-juvenile dispersal distances have been estimated to range 2-4 km, but some males have been observed to disperse up to 24.9 km (Montadert and Léonard 2006). Given this information, the French samples were grouped into three putative subpopulations with a minimum distance of 80 km apart, to ensure that all samples within subpopulations were likely to be subject to only moderate amounts of contemporary gene flow, whereas samples between clusters would have little or no contemporary gene flow.

Table 1. Hazel grouse populations, subpopulations and number of samples genotyped per molecular marker type.

			Number of samples genotyped	
Region	**Code**	**Year**	**Microsatellites**	**MHC**
Sweden	S	1981 & 1982	111	105
France	F	1990-2009	131	141
Locality				
Sweden North	SN	1981 & 1982	55	51
Sweden South	SS	1981 & 1982	56	54
France Jura North	FJN	1990-2009	28	32
France Alps North	FAN	1990-2009	20	20
France Alps South	FAS	1990-2009	83	82

In France, 143 feather samples were collected from individuals captured in hunting bags as well as feathers lost from predated carcasses. Here, the samples were referenced by the precision of a 10x10 km European grid FID as well as by the corresponding governmental administrative department of the area. The sampling was conducted over several consecutive years, ranging from 1990 to 2009. In Sweden, 102 samples previously collected in the years 1981 and 1982 of a study by Sahlsten *et al.* (2008) were used. These consisted of feathers and foot pad scrape from individuals captured in hunting bags. Here, the samples were referenced according to exact coordinates, as well as by the corresponding governmental administrative department of the area. Only a subset of the samples from both regions was determined by sex, and consequently the sex ratio is not known.

DNA for all samples was extracted using a high salt purification protocol, and to ensure consistent DNA concentration and a minimum of organic and non-organic contamination this was measured using a NanoDrop Fluorospectrometer (Thermo Fisher Scientific Inc.). Samples containing more than 20 ng/µl were regarded as good, and constituted the majority of samples, however a few samples containing as low as 10 ng/µl was also used with slightly higher volume in subsequent PCRs. The 260/280 ratios were consistently over 1.7 with few exceptions and thus the purity was regarded as generally good. See Appendix 1 for DNA extraction protocol.

As noted above, part of the aim of this study was to draw management and conservation status conclusions about Hazel grouse in a human-modified heterogeneous landscape in the French Alps. Thus, although all samples were genotyped and analyzed, the reader should mainly focus on the results of the French populations, while keeping in mind that the Swedish populations can be viewed as essentially a genetic reference of a long term stable, naturally continuous reference landscape (Sahlsten *et al.* 2008), to which the French Alps populations can be compared.

PCR and genotyping

Two separate PCR protocols were followed, since the two markers used in this study, microsatellites and MHC respectively, were genotyped using two distinct methods. These are briefly summarized below. See Appendix 2 for complete PCR protocols.

Microsatellite genotyping

Eleven microsatellite markers (ADL230, ADL257, ADL142, ADL184, BG15, BG16, BG18, TUT1, TUT2, TUT3, TUT4) were amplified by polymerase chain reaction (PCR) and genotyped following Segelbacher *et al.* 2000, with the modification of amplifying markers in concert in three separate primer multiplex reactions. Primer pairs with similar annealing temperatures (48, 54, 60 ºC respectively) were multiplexed. In order to allow for the separation of primer dye set standards between each locus typed, further grouping of reactions were done. In total, five different reactions were performed for all samples.

Scoring errors are methodological problems that lead to false genotypes being recorded, the true genotype actually being different. Several potential scoring errors may arise when genotyping microsatellite markers, having large scale effects which introduce biases in the data, affecting the conclusions drawn about the data and its reproducibility (Bonin *et al.* 2004, Dewoody *et al.* 2006, Van Oosterhout *et al.* 2006, Björklund 2005, Carlsson 2008). Large allelic dropout, stuttering and null alleles are well documented problems likely to occur; mainly being consequences of the technical limitations of PCR and capillary gel electrophoresis, rather than human error. However, these problems can, and must, be mitigated before subsequent analyses, and I will return to how this is done.

Large allelic dropout occurs when the PCR template of one allele fails to amplify during the reaction, thus increasing overall homozygosity. Null alleles give a similar increase in homozygosity and are the result of mutations in the flanking regions of a microsatellite allele, where primers anneal. Failure of primer annealing will occur and thus amplification failure of

the allele then follows. A heterozygote with one null allele is then shown as a homozygote, while a homozygote for a null allele is impossible to genotype, and is thus part of the ungenotyped parts of the data set. This is likely to be prevalent when directly transferring primers between taxa, and is thus of particular relevance to this study. Stuttering is the phenomenon where there may be signs of inflated number of alleles, due to scoring errors of alleles that on the one hand have short repeat lengths and on the other very similar allele lengths, thus making them hard to discern.

As stated above, these three phenomena are known to cause large and systematic biases in homozygosity, allele number and allele length. If left untreated, this may thus affect many of the standard diversity and population structure indices, which are dependent on correct frequency and haplotype data. Consequently, when working with large multi-locus data sets, as in this study, it is always of importance to minimize problems arising from this. One way to counteract scoring errors in microsatellite genotyping is to perform the typing more than once per sample, over several replicated reactions. This was not possible in its entirety in this study due to time constraints; however a subset of samples chosen blindly were genotyped twice to control for possible large scale PCR artefacts between samples.

Another method important to consider when dealing with scoring errors, is to estimate the actual error rates. Hence, this enables cleaning up the data set before any analysis is conducted, and also for detecting any eventual biases in the results. Here the program MicroChecker 2.2.3 (Van Oosterhout *et al.* 2004) was used for this purpose. For every putative subpopulation, null allele frequencies were estimated, and tests for large allelic drop-out and stuttering were made. Two loci (ADL230, ADL257) were excluded from further analysis due to some severely high null allele estimates (Table 2), no subpopulations showed large allelic dropout, and only one locus showed signs of slight stuttering (BG15). In total, nine loci were included in subsequent analyses.

As shown in Table 2, a few loci included in analyses did show signs of high prevalence of null alleles in some subpopulations, but as the estimations are based on measures of heterozygote deficiency, some null alleles can be expected entirely as a consequence of true population genetic processes such as inbreeding. Significance level of the null allele estimation was 0.05 and values at or around this value can be considered zero.

Table 2. Null allele estimates. Values in bold was considered high, values in bold italic severely high and the loci severely affected excluded (ADL 230, ADL 257).

Locus	Subpopulation				
	SN	**SS**	**FJN**	**FAN**	**FAS**
ADL230	***0.23***	***0.21***	***0.12***	*0*	*0*
ADL257	***0.37***	***0.28***	***0.34***	*0*	*0.06*
ADL142	**0.16**	0	0.08	0	0
ADL184	0	0	0	0	0
BG15	0	0.0063	0.0042	0.07	0
BG16	0.05	0.1	**0.13**	**0.16**	0.13
BG18	0	0.05	0	0.03	0.05
TUT1	0.024	**0.16**	**0.21**	**0.15**	0.13
TUT3	0	0.09	0.12	0.07	0.08
TUT2	0	0	0	0	0.01
TUT4	0	0	0	0	0

MHC genotyping - RSCA

Several methods have been developed in order to genotype MHC genes, where among the most widely used are direct cloning and sequencing, pyrosequencing, single strand conformation polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE) and reference strand-mediated conformational analysis (RSCA). The MHC genomic region constitutes a vast number of highly polymorphic genes, which not only are very similar in sequence between alleles at the same locus, but also showing high similarity between loci since many of the genes have evolved by gene duplication and gene conversion (Chaves *et al.* 2010). This high polymorphism renders direct cloning and sequencing of MHC genes in large part both economically unfeasible and time-consuming. Here, the RSCA method is used, originally developed by Arguello *et al.* (1998) and further developed by Strand and Höglund (in press.) for genotyping of MHC class IIB (*BLB*) loci in non-model species such as in the grouse family of birds (*Tetraonidae*).

RSCA uses fluorescently labeled reference sequences (FLRs) cloned from known *BLB* alleles, in a denatured single-strand state, to bind to unknown single-strand sequences of *BLB* alleles amplified by polymerase chain reaction, thus forming a heteroduplex of the two single strands. As a result of mismatches in sequence between the reference allele and the unknown allele, a conformational structure is formed and when run in capillary gel electrophoresis each allele to be genotyped gives a separate migration value in the gel.

A primer pair (RNA F1a (5'-GACAGCGAAGTGGGGAAATA-3') and RNA R1a (5'-CGCTCCTCTGCACCGTGA-3')) originally designed to target the flanking regions of the ~160 base-pair exon 2 in Black grouse (*Tetrao tetrix*) *BLB* region was used to amplify the corresponding *BLB* sequences in Hazel grouse. PCR products were confirmed of the correct length on polyacrylamide gels. Reference strands (FLRs) were amplified from known clones of Black grouse *BLB* genes using a modified RNA F1a primer, labelled with a fluorescent dye. In total, four different FLRs were created, to allow for a multi-dimensional separation of

alleles; giving a higher resolution of the true allelic differentiation than if only one FLR would be used.

The *BLB* exon 2 in Hazel grouse is likely to vary in copy number repeats between individuals, as has been shown in Black grouse (Strand *et al.* 2007), which might have up to two or even three *BLB* copies. None of the copies are pseudogenes; all are expressed and functional. However, one of the limitations of the RSCA is that the primer pair is not designed to make a distinction between loci, and thus amplifies all copies of the *BLB* locus. The RSCA did indeed find up to eleven alleles in one individual, which suggests that there is on the one hand more than one locus amplified in the PCR, and on the other also genotyped alleles that have arisen as artefacts from differential binding of specific FLRs to the allele. It was impossible to discriminate between these two outcomes. To compensate for this, a binomial probability test was performed, to remove one in a pair of genotyped alleles that did show significant association in the data set. The data for MHC is thus based entirely on allele frequencies and not complete diploid haplotypes.

While this will lead to biased estimates of absolute allele numbers and heterozygosity of MHC class IIB diversity, the estimates are still informative in a relative subpopulation-by-subpopulation comparison context. Genotyping errors of this sort are likely to arise randomly and evenly across loci regardless of the origin of samples.

Analysis of microsatellite data

Initially, a test for linkage disequilibrium (LD) between each pair of loci was conducted, both globally and for each subpopulation in Genepop 1.2 (Raymond & Rousset 1995). LD is the non-random association of alleles at two or more loci and may arise due to close proximity of loci on the same chromosome thus limiting their recombination rate, or due to population genetic processes such as reduced recombination in populations with small effective sizes (McVean 2002). The test for LD was done to check whether that any of the loci was in physical linkage, as this would bias any parameter of genetic diversity that treats loci as independent replicates, and also to investigate whether any of the subpopulations showed any consistent patterns of LD between loci due to population bottlenecking.

Furthermore, tests for recent population bottlenecks and deviations from Hardy-Weinberg distribution (heterozygote excess and deficit) were done in the program Bottleneck 1.2.02. (Cornuet & Luikart 1996) using the TPM model with 95% step-wise mutations with a variance of 5%. An observed mode shift from an L-shaped distribution of allele sizes (Wilcoxon test) would provide evidence for a recent severe population bottleneck.

Hereon after, analyses and results can be divided into the two broad categories: population genetic diversity and population genetic differentiation respectively, of which there are several types of parameters.

Neutral population genetic diversity indices

To assess neutral genetic diversity between the putative subpopulations the standard diversity indices:

- H_E (unbiased expected heterozygosity / Nei's unbiased gene diversity) (Nei 1972),
- H_O (observed heterozygosity),
- F_{IS} (Wright's inbreeding coefficient) (Wright 1943) and
- AR (allelic richness)

were calculated in FSTAT 2.9.3.2 (Goudet 1995). The output of the diversity indices in FSTAT is given by locus and to achieve a multi-locus mean, jack-knifed estimates over all loci were obtained using SYSTAT.

The unbiased expected heterozygosity H_E is the heterozygosity expected under Hardy-Weinberg equilibrium and is defined as:

$$H_E = 1 - \sum_{i=1}^{\#alleles} p_i^2 \qquad \text{(Nei 1972)} \qquad (1)$$

where p_i is the frequency of the $i{:}th$ allele.

Wright's inbreeding coefficient F_{IS} is a measure of either a) a subpopulation's departure from panmixia under an island model (in other words, subpopulation structure) or b) random mating of close relatives which causes inbreeding (Wright 1943). The latter is of interest here. F_{IS} has been defined as the ratio of the average individual level variance in multi-locus heterozygosity, to the total subpopulation variance in multi-locus heterozygosity:

$$F_{IS} = \frac{H_E - H_O}{H_E} \qquad \text{(Wright 1943)} \qquad (2)$$

Allelic richness is the rarified mean number of alleles, meaning that it is normalized to the smallest sample size. Tests for significant differences in the standard diversity indices among groups of populations were performed using FSTAT 2.9.3.2 with 1000 permutations and a one-tailed test.

Neutral population genetic differentiation indices

To examine population structure and differentiation, global and pair-wise F_{ST} for all possible subpopulation pairs (Weir & Cockerham 1983) were estimated in Arlequin 3.5.1.2 (Excoffier and Lischer 2010). F_{ST} is defined as:

$$F_{ST} = \frac{H_T - H_S}{H_T} \qquad \text{(Weir \& Cockerham 1983)} \qquad (3)$$

where H_T is the expected heterozygosity of the total population and H_S is the expected heterozygosity of a subpopulation.

Each putative subpopulation was included in F_{ST} estimates as stated above in a global calculation with no predefined structure, but also in a calculation where the Sweden and France subpopulations were distinguished in a hierarchical structure of two regional groups

and five subpopulations in total. This allowed for significance testing of differences between regions and subpopulations as well as a way to assess the relative contribution of each pair-wise differentiation estimate to the global mean. Pair-wise F_{ST} was estimated for all possible subpopulation pairs.

An equivalent to F_{ST} was also used, Jost's differentiation index D_{EST} (Jost 2008). This estimate is based on the assumption that the effective number of alleles, rather than their expected heterozygosity, in each subpopulation provide a sufficient partitioning of diversity (Meirmans and Hedrick 2011). Calculation of D_{EST} was done in the software SPADE (Chao *et al.* 2008).

Significance testing for population genetic differentiation was done by a multi-locus hierarchical analysis of molecular variance (AMOVA) as a weighted average over loci, also in Arlequin 3.5.1.2. Tests were performed at three organizational levels: between groups (Sweden/France, F_{CT}), among subpopulations within groups (F_{SC}) and among individuals within subpopulations (F_{IS}). Since both the pair-wise F_{ST} and AMOVA tests constitute multiple comparisons, a Bonferroni correction of significance levels was performed to reduce the likelihood of obtaining type I errors.

Subdivided, demographically stable populations that exchange migrants and which have a continuous, but not panmictic, distribution are expected to have dispersal and gene flow reduced by increasing distance, and thus has differentiation and F_{ST} positively correlated with distance. This has been termed isolation-by-distance (IBD) (Wright 1943) and deviations from the linear model can be indicative of either high gene flow or extreme isolation – or, of more relevance in this study, signs of the dynamics of range expansions and contractions. Testing for correlation and potential outliers was done by plotting the natural logarithm of the distance between subpopulation pairs on their respective pair-wise F_{ST} values. A logarithmic scale of distance was used since the colonization axis of Hazel grouse in the French Alps cannot be considered perfectly linear, but rather two-dimensional in geographical structure (Montadert and Leonard 2003), and log-transformation has been shown appropriate by simulation for species distributions that resemble two-dimensional rather than one-dimensional geographical structure (Rousset 1997), and correspondingly a transformation of pair-wise F_{ST} values were done (Slatkin's linearized F_{ST}, $F_{ST}/(1-F_{ST})$, Slatkin 1995).

Distances were estimated in kilometers using Google Maps, measuring from the center of each sampled putative subpopulation. Both straight-line Euclidean distances and a point-by-point estimate of a potential terrestrial migration pathway between subpopulations were used. The latter could, after inspection, be considered more biologically relevant than the former, and these IBD estimates are presented in the results. To test if pair-wise F_{ST} were indeed spatially auto-correlated, Mantel tests and partial Mantel tests were performed in Arlequin with 1000 bootstraps.

An additional method of detecting subpopulation structure involves factorial component analysis (FCA). This was performed using the software Genetix 4.05.2 (Belkir *et al.* 2000). The analysis partitions the total multilocus genotypic variation in the data set into two

independent factors and accordingly assigns every individual genotype into a two-dimensional space. This provides further explanatory power to differentiation indices and isolation by distance models.

Analysis of MHC data

Adaptive population genetic diversity indices

To assess the relative diversity at MHC class IIB loci the total number of alleles, the number of unique (private) alleles, mean number of alleles per individual and the gene diversity $\hat{H}$ (Nei 1987) was calculated for each putative subpopulation in SYSTAT.

The gene diversity $\hat{H}$ is an equivalent to the unbiased expected heterozygosity H_E for diploid data and is estimated from the available gene copies and allele frequencies according to the following formula:

$$\hat{H} = \frac{n}{n-1}\left(1 - \sum_{i=1}^{k} p_i^2\right) \qquad \text{(Nei 1987)} \qquad (4)$$

where n is the number of gene copies in the sample, k is the number of haplotypes and p_i is the sample frequency of the $i{:}th$ haplotype. Since no discrimination could be made in the RSCA based on haplotype, k is here treated as 1.

Adaptive population genetic differentiation indices

To assess population differentiation at MHC class IIB loci, the F_{ST}-equivalent D_{EST} (Jost 2008) was used as for microsatellite data, and was similarly calculated in the software SPADE (Chao *et al.* 2008). This allowed for a direct comparison of population differentiation between neutral and adaptive molecular markers and to assess their relative magnitude. To assess whether adaptive pair-wise differentiation followed a similar isolation by distance trend as neutral genetic differentiation, a correlation between the differentiation index D_{EST} for microsatellite and MHC was done by a Pearson product-moment correlation.

An additional way of visualizing MHC differentiation is to plot allele frequencies of each *BLB* allele according to their prevalence in each subpopulation. Allele frequencies were estimated as the total number of individuals in a subpopulation carrying a particular allele divided by sample size. This is robust to different sample sizes similar to allelic richness, as well as it provides a way to visualize how allelic prevalence changes between subpopulations, similar to the FCA plot.

Control for yearly variation in sampling scheme

Since the French sampling had been conducted over a total of nineteen consecutive years, a control for yearly variation in sampling was performed. This was done by grouping samples according to year, instead of geographic location, and then checking for population differentiation in the data, as was done with the location-based data in Arlequin 3.5.1.2. If there is significant yearly variation in the sampling, one would expect more structure (ie. higher pair-wise F_{ST} values) for samples taken far apart in time, than those closely spaced in time. However, differences in pair-wise F_{ST} were either relatively small, or non-significant, when grouped according to year (range: pair-wise F_{ST} = 0.021-0.072; p-values = 0.003-

0.560). The same was done with MHC data, and similarly, only small differences in D_{EST} were found (range: pair-wise D_{EST} = 0.0-0.0; global D_{EST} 95% CI = 0.000-0.016), suggesting that there is relatively small variation in the data due to sampling scheme and that the French sample provides a sufficient snapshot of genetic variation, despite being sampled over several years.

Results

There was no evidence of any consistent linkage disequilibrium between loci. Significant deviations from Hardy-Weinberg conditions were observed for subpopulations FAS ($p = 0.00977$) and FJN ($p = 0.00195$) after Bonferroni correction ($p = 0.01$). Subpopulation SN was almost significant at ($p = 0.027$). No mode shift from L-shaped distributions of allele sizes indicative of severe population bottlenecks were observed.

Neutral population genetic diversity - Microsatellites

Allelic richness ranged from 3.48 (France Alps South) to 6.24 (France Jura North), expected heterozygosity ranged from 0.44 to 0.64 between the same subpopulations, observed heterozygosity ranged from 0.40 to 0.59 also between the same subpopulations and Wright's inbreeding coefficient ranged from 0.070 (Sweden North, France Jura North) to 0.099 (France Alps North) (Table 3).

Table 3. Standard microsatellite genetic diversity indices of all sampled subpopulations. n, sample size; AR, allelic richness; H_E, expected heterozygosity; H_O, observed heteroygosity; F_{IS}, inbreeding coefficient.

Subpopulation	Code	n	AR	H_E	H_O	F_{IS}
Sweden North	SN	55	4.76	0.58	0.54	0.070
Sweden South	SS	56	4.73	0.54	0.53	0.008
France Jura North	FJN	28	6.24	0.64	0.59	0.070
France Alps North	FAN	20	4.40	0.53	0.48	0.099
France Alps South	FAS	83	3.48	0.44	0.40	0.077

No significant differences in AR, H_E, H_O or F_{IS} could be observed, neither at between-regional or between-subpopulation comparisons ($P > 0.15$). However, it is important to note that the high p-values may be indicative of a lack of statistical power in the performed exact test, since the units of permutation is the subpopulations *per se* (here being as low as $n = 5$), not the individual samples. The exact test for differences in diversity measures used here might therefore be especially prone to type II errors. Given larger replication of subpopulations in each region, clarification of any significant differences observed in Table 3 might be possible.

Neutral population genetic differentiation - Microsatellites

Global microsatellite fixation index was high at 0.284 (95% CI = 0.190-0.373) and pair-wise F_{ST} ranged from 0.035 (Sweden North-Sweden South) to 0.376 (Sweden South-Alps South). All F_{ST} values were highly significant ($P < 0.0001$) after 1000 bootstraps and Bonferroni correction. Similarly, global microsatellite differentiation index D_{EST} was high at 0.297 (s.e. = 0.014, 95% CI = 0.269-0.324) and pair-wise D_{EST} ranged from 0.044 (Sweden North-Sweden South) to 0.565 (Sweden South-Alps South) (Table 4).

Table 4. Microsatellite pair-wise F_{ST} (above diagonal, Weir & Cockerham 1984) and differentiation index D_{EST} (below diagonal, Jost 2008) for all population pairs. All F_{ST} values were highly significant (P < 0.0001) after 1000 bootstraps and Bonferroni correction.

		Sweden		France		
		North	**South**	**Jura North**	**Alps North**	**Alps South**
Sweden	North		0.035	0.170	0.234	0.360
	South	0.044		0.197	0.257	0.376
France	Jura North	0.322	0.350		0.045	0.140
	Alps North	0.372	0.374	0.056		0.117
	Alps South	0.564	0.565	0.170	0.108	

A significant isolation-by-distance signature was found when performing Mantel tests of spatial auto-correlation of genetic distances, using both Slatkin's linearized pair-wise F_{ST} and differentiation index D_{EST} ($r_{M\,(FST)} = 0.76$ p = 0.008, $r_{M\,(DEST)} = 0.89$ p = 0.006) (Figure 2.). When the Swedish outlier (SN-SS) was removed, a slightly stronger correlation was achieved, however with reduced power (higher, but still significant, p-value) ($r_{M(FST)} = 0.78$ p = 0.035, $r_{M\,(DEST)} = 0.93$ p = 0.042). Although it would have been informative to perform further Mantel tests, excluding further potential outliers, this was not possible due to lack of power when data were removed.

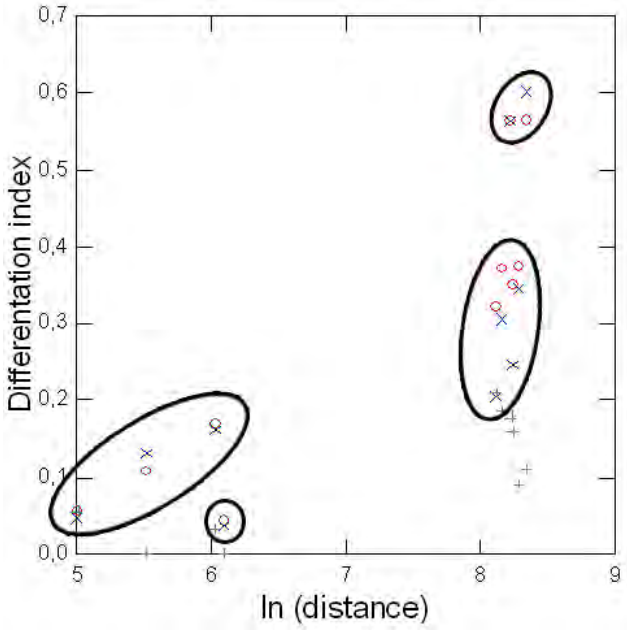

Figure 2. Isolation-by-distance plot with Slatkin's linearized pair-wise F_{ST} (blue crosses) and the differentiation index D_{EST} (red circles) as a function of the natural logarithm of distance (km). The left cluster is population pairs of France (FJN, FAN, FAS), the lower right cluster is population pairs of Sweden (SN, SS) and France (FJN, FAN), the upper right cluster is population pairs of Sweden (SN, SS) and France (FAS). Single lower point is population pair of Sweden (SN, SS). Also, uncircled points (green plusses) are the differentiation index D_{EST} for MHC.

The FCA plot (Figure 3) for two factors showed that the Swedish subpopulations were not differentiated among themselves but clearly in a cluster to the left, while France Jura North and Alps North were undifferentiated to the right, leaving France Alps South in a separate cluster to the far right. The tight formation of Alps South into a separate cluster is in accordance with the lower diversity seen in this subpopulation, as well as with the highest pair-wise differentiation indices which are paired with Alps South. Similarly, the observed conformity of the Jura North and Alps North clusters are in accordance with their respective diversity and differentiation measures. The x-axis explained 7.63 % of the observed variation while the y-axis explained 3.81 % of the observed variation.

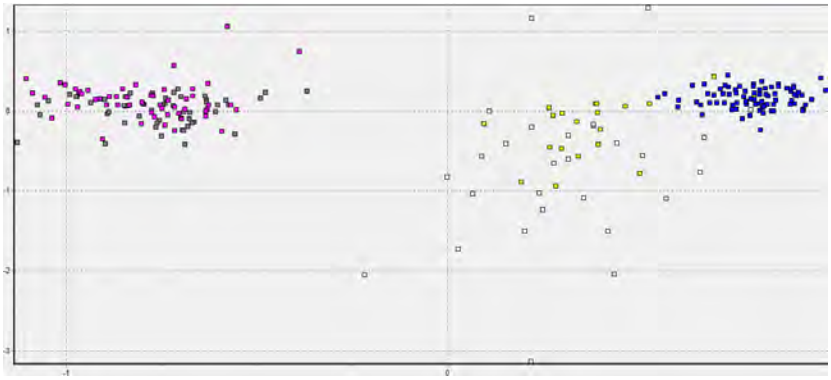

Figure 3. Two-factorial component analysis (FCA) for all genotyped microsatellites. For visualization purposes 4 factor FCA is shown. Grey points (SN), violet points (SS), white points (FJN), yellow points (FAN), blue points (FAS).

Adaptive population genetic diversity - MHC

Total number of alleles per subpopulation ranged from 7 (France Alps North) to 11 (Sweden North, Sweden South), private alleles were only observed in the two Swedish subpopulations and in France Jura North (1-2 private alleles), mean number of alleles per individual ranged from 3.13 (France Jura North) to 5.11 (Sweden South) and gene diversity ranged from 0.76 (France Jura North) to 0.87 (Sweden South) (Table 5).

Table 5. MHC genetic diversity indices of all sampled subpopulations. n, sample size; $\hat{H}$, gene diversity.

Subpopulation	Code	n	Total no. of alleles	Private alleles	Mean no. of alleles per ind.	$\hat{H}$
Sweden North	SN	51	11	2	4.80 (3-8)	0.86
Sweden South	SS	54	11	1	5.11 (2-9)	0.87
France Jura North	FJN	32	8	1	3.13 (3-7)	0.76
France Alps North	FAN	20	7	0	4.90 (4-7)	0.84
France Alps South	FAS	82	8	0	4.52 (3-7)	0.81

Adaptive population genetic differentiation - MHC

Global MHC fixation index was moderate at 0.100 (95% CI = 0.059-0.141, s.e = 0.021 after 200 bootstraps) and pair-wise D_{EST} ranged from 0.000 (Sweden North-Sweden South, France Alps North-France Alps South) to 0.210 (Sweden North-France Jura North) with standard error ranging from 0.019 to 0.050 after 200 bootstraps (Table 6).

Table 6. MHC pair-wise differentiation index D_{EST} (Jost 2008) for all population pairs. Negative values can be considered zero as they are due to sampling variation.

		Sweden		**France**		
		North	**South**	**Jura North**	**Alps North**	**Alps South**
Sweden	North	0				
	South	-0.003	0			
France	Jura North	0.210	0.159	0		
	Alps North	0.185	0.091	0.055	0	
	Alps South	0.178	0.110	0.032	-0.015	0

The Pearson product-moment correlation between pair-wise differentiation index D_{EST} for microsatellites and MHC respectively (Figure 4) showed a strong correlation (Pearson's r = 0.73, p = 0.018). Few putative outliers can be observed, although the Sweden North-France Jura North pair at D_{EST} = 0.210 might be one. This slightly higher than expected differentiation is in accordance with the lowest adaptive population genetic diversity and heterozygosity observed for France Jura North.

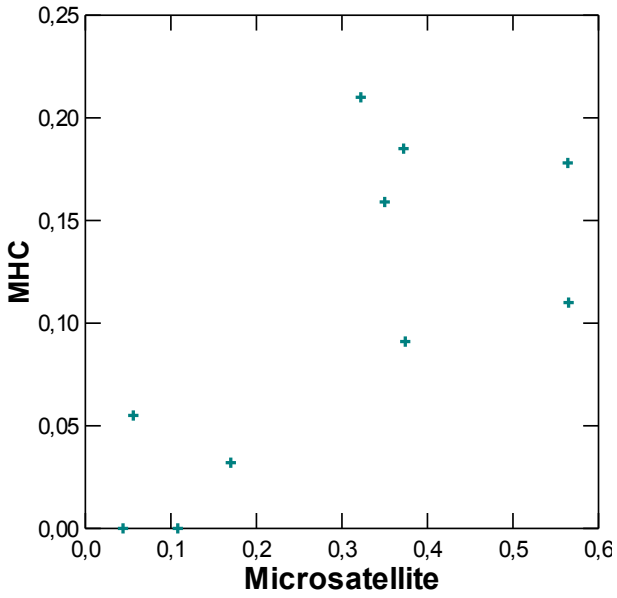

Figure 4. Pearson product-moment correlation between pair-wise differentiation index D_{EST} for microsatellites and MHC respectively (Pearson's r = 0.73, p = 0.018).

When plotting allele frequencies per subpopulation of each putative *BLB* allele (Figure 5) it is possible to discern which alleles that have been maintained over different geographical regions. The most striking observation is that the France Jura North subpopulation, in accordance with its relatively lower diversity measures and slightly higher differentiation, has up to seven *BLB* alleles (BLB05, BLB07, BLB03, BLB04, BLB01, BLB12, BLB06) that exist in lower frequencies here than in the other French subpopulations. Four alleles (BLB04, BLB01, BLB12, BLB06) show signs of a similar gradient-like decline in diversity between FAN and FAS subpopulations as is shown for neutral diversity. Also, two unique alleles at high frequencies (BLB02, BLB10) for the two Swedish subpopulations can be observed.

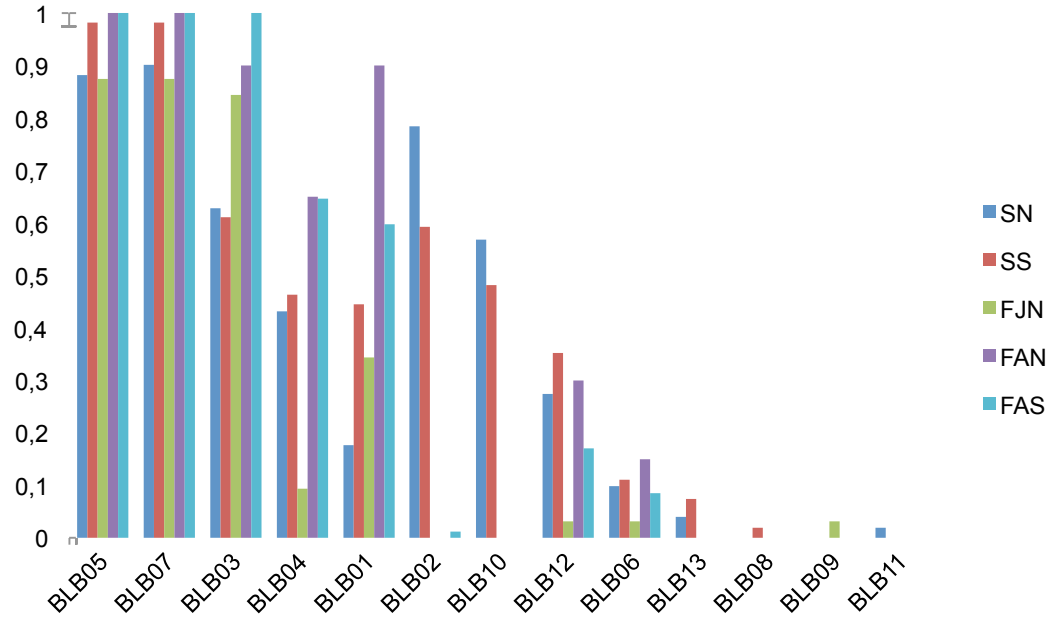

Figure 5. MHC class IIB (*BLB*) allele frequencies per allele and subpopulation.

Discussion

Evidence from both neutral and adaptive genetic diversity measures in this study of Hazel grouse, support that the subpopulation at the range expansion front (France Alps South) is highly isolated from the main distribution range, has been subjected to a gradient-like decline in neutral genetic diversity and become highly differentiated since migrating southwards and becoming established in the 1950s.

For microsatellite data, France Jura North harbors the highest observed neutral diversity for most diversity measures, France Alps North have slightly lower, while France Alps South have the lowest observed allelic richness and expected heterozygosity of all subpopulations, while at the same time does not differ from the other in its inbreeding coefficient (Table 3). Similarly for MHC data and adaptive diversity, France Alps South corresponds to this weakly declining trend; at least when compared to France Alps North and the Swedish reference populations (Table 5).

The observed gradient in diversity can be considered slight for both neutral and adaptive genetic diversity, especially considering that no clear differences in the inbreeding coefficient is observed. This rather slight loss suggests, as shown by spatially explicit simulation of range expansions (Excoffier *et al.* 2009) that the subpopulations of Hazel grouse in the French Alps have maintained relatively large effective population sizes during their migration along the Alp axis. As loss of both neutral and adaptive diversity during range expansion is related to number of migrants, density and effective population size, a higher rate of loss would likely have been observed if migration had been hampered by a low population size. Similarly, loss of diversity at MHC class IIB loci is related to the strength of selection and amount of genetic drift (which is related to effective population size). A loss of adaptive diversity due to drift can be expected if drift is significantly counteracting selection, which also plausibly is not observed in this system, although further testing of this hypothesis might elucidate if this might be so.

Furthermore, the observed gradient in diversity is not unexpected. The colonization pattern along the axis is likely to resemble a stepping-stone structure, where only adjacent subpopulations exchange migrants, which may create frequent population bottlenecks - or more specifically, serial founder effects - along the range expansion axis (Frankham *et al.* 2009, Excoffier *et al.* 2009).

It can be argued that the isolation-by-distance model (Figure 2) supports these findings. Under the assumption of a continuous, two-dimensional but not panmictic distribution, linearity under isolation by distance can be expected as shown by simulation (Rousset 1997, Sokal and Wartenberg 1983), and France Alps South (when paired with the Swedish reference populations) is seemingly an outlier deviating from linearity, having a higher differentiation than under model predictions, using both pair-wise F_{ST} and D_{EST} (upper right cluster in Figure 2). Considering that Hazel grouse in France Alps South is rather isolated at the range expansion front, being newly established and ever since acting as a source population, a slightly higher differentiation than at continuous distribution can be expected (Rousset 1997). The fact that there also is an observed deviation from mutation-drift

equilibrium in this subpopulation further consolidates France Alps South as being subjected to the above-mentioned processes.

The SN-SS pair and the FAN-SS|FJN-SS pairs are arguably even more pronounced outliers in Figure 2, showing lower than expected differentiation. While the Swedish pair can be expected to have high rates of gene flow leading to lower differentiation than expected by distance (Sahlsten *et al.* 2008) this is not a feasible explanation for the lower differentiation seen between the northern French subpopulations and the Swedish; current gene flow should not occur between these pairs due to the substantial distance and the distribution pattern of Hazel grouse in Europe. This suggests that processes other than contemporary gene flow between Sweden and France, such as population demographic history (migration, gene flow and admixture) in northern France (and possibly at even larger spatial and temporal scales than considered here) have caused this higher similarity of diversity in the northernmost French subpopulations. This reasoning is also supported by the higher heterogeneity and lack of differentiation between clusters of France Jura North and Alps North in the FCA plot (Figure 3).

Also, the FCA plot shows France Alps South as a clearly differentiated cluster to the far right (Figure 3), and this population has the highest pair-wise differentiation estimates (Table 4), supporting the notion that it is relatively isolated, has a positive intrinsic growth rate and acts as a source population since at least the 1970s (Montadert and Léonard 2003). This is also a likely explanation why no recent population bottlenecks were observed in any subpopulation. Range expansion of Hazel grouse along the Alps axis, and its initial deleterious effects on diversity, is arguably no longer very recent.

It is important not to neglect potential effects of unsampled 'ghost' populations when discussing population genetic processes such as migration and gene flow at limited spatial and temporal scales (Slatkin 2005). For obvious reasons, it is seldom possible to sample all existing populations and monitor their change over time, but their effects on the actual sampled populations are inevitable when species distributions resemble an infinite-island model with migration and gene flow. This bears particular relevance in this study.

There is empirical evidence from several species that shows that historical processes such as a long term dynamic of complete isolation followed by admixture of divergent lineages may create a 'melting pot' or 'hotspot' of intraspecific genetic diversity (Petit *et al.* 2003, Canestrelli *et al.* 2010) and the relatively higher neutral genetic diversity of Hazel grouse in the northernmost French subpopulations may be indicative of this (Table 3, France Jura North). Surprisingly, the adaptive genetic diversity for this subpopulation shows the exact converse; relatively lower diversity as compared to the other subpopulations, but also the highest differentiation as measured by D_{EST} (Table 5, Figure 4). It is possible that at the larger central European scale, a selective regime more directional in nature, favoring fewer haplotypes more specific to certain pathogens, has over time shaped a more uniform and less diverse arsenal of MHC class IIB genes in Hazel grouse. At the same time, having concomitant admixture of these individuals from the many scattered populations around France, may thus have created the previously mentioned genetic hotspot of neutral (but not

adaptive) genetic diversity in France Jura North. The higher prevalence of MHC class IIB alleles in the two southernmost French subpopulations may, since the range expansion during the 1950s, have experienced a broader diversifying selection, driving them to diverge in this regard.

Although the 'genetic hotspot' of neutral diversity may seem likely, given the comprehensive microsatellite genotyping done in this study, the observed low adaptive diversity is enigmatic. An alternative scenario of MHC class IIB diversity in Hazel grouse for central Europe would be that there exists a higher diversity overall than in the French Alps, due to a broad diversifying balancing selection, rather than narrow directional selection. This is not unlikely since the habitat across large parts of Europe is highly fragmented and subjected to a broad range of anthropogenic pressures at different geographical scales which might lead to differential selection pressures in certain regions. This scenario would provide a higher overall MHC diversity in the case of migration and admixture into a subpopulation, as is purportedly present in France Jura North.

Given the relatively low sampling in this subpopulation and the genotyping methods used at MHC class IIB loci, PCR artifacts and scoring errors in the MHC data cannot be ruled out. To further investigate this, a phylogeographical study of mtDNA genealogies may be applied to the currently sampled populations, as well as increasing the sampling of populations further north in France and to the parts of central Europe for microsatellite and MHC data. This may provide sufficient sampling to test the hypothesis and possibly achieving supporting evidence for the here proposed mechanisms behind the current rather enigmatic pattern observed in France Jura North.

Conclusions

The patterns observed in genetic diversity along the range expansion axis of the French Alps is by and large in accordance with what is previously known about population dynamics of Hazel grouse in the region. Both neutral and adaptive diversity has seen a gradient-like decrease - however slight for adaptive diversity - along the axis.

The diversity and differentiation of the France Jura North (FJN) subpopulation seems to somewhat contradict what is currently known about this subpopulation. The high neutral diversity is indicative of high levels of admixture of individuals, possibly due to the long-term low density of individuals in this fragmented subpopulation. Alternatively, admixture may be due to migration and gene flow into the subpopulation from parts of Central Europe at larger scales than studied here. The low adaptive diversity of France Jura North may be a consequence of the same admixture process, providing inflow of more directionally selected MHC from larger parts of Central Europe. This scenario assumes a uniform selective regime for Central Europe. Alternatively, this pattern at adaptive genetic markers is the consequence of PCR artefacts and genotyping errors, while selective regimes actually may be more diversified across Europe. Further studies are thus needed to elucidate this.

As far as the two southernmost French subpopulations are concerned (France Alps North and France Alps South), they indicate a decline in both neutral and adaptive genetic diversity, with Alps South showing the most pronounced reduction in diversity as a consequence of the

serial founder effects along the range expansion axis. However, since the loss in adaptive diversity is slight, it has arguably maintained a viable population size since being established, which also is in accordance with the observation that France Alps South experiences a positive intrinsic growth rate and acts as a source population since a few decades.

This underlines that conservation management should focus on separate actions for the southernmost Alps South and the northernmost Jura North subpopulations. Their current differences in neutral and adaptive genetic diversity as well as population demographic processes are likely to be substantial. Considering the two southernmost subpopulations Alps North and Alps South, there is indication of low contemporary gene flow between them, as they have relatively high neutral differentiation (pair-wise $F_{ST} = 0.117$). Furthermore, there is no significant adaptive differentiation (MHC $D_{EST} = -0.015$), and consequently these two subpopulations can be considered as uniform units. They only differ in neutral and adaptive diversity due to the serial founder effects during the range expansion, otherwise being uniform in genetic structure. Thus, this might be a sufficient criterion for treating these two subpopulations as a management unit (MU). The loss of diversity due to the observed founder effect in France Alps South may be counteracted by translocating individuals from France Alps North to France Alps South if its population trend would become negative.

Acknowledgements

I would like to thank my two supervisors Jacob Höglund and Tanja Strand for being so supportive and for providing invaluable help throughout the project. A special thanks goes to the two very friendly lab technicians Gunilla Engström and Reija Dufva for keeping the laboratory in such a nice, workable condition and atmosphere. Also deserving mention are all those population geneticists with exceptional programming skills who have created the gazillion computer programs used in this study.

References

Arguello J R, Little A M, *et al*. 1998. High resolution HLA class I typing by reference strand mediated conformation analysis (RSCA). *Tissue Antigens* **52**: 57-66.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P. 2004. How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13**: 3261-3273.

Belkir K, Borsa P, Chiki L, Raufaste N, Bonhomme F. 2000. GENETIX 4.05, Logiciel sous Windows pour la Génétique des Populations. Laboratorie Génome, Populations, Interactions, CNRS UMR. 5000, Université de Montpellier II, Montpellier, France.

Björklund M. 2005. A method for adjusting allele frequencies in the case of microsatellite allele drop-out. *Molecular Ecology Notes* **5**: 676-679.

Canestrelli D, Aloise G, Cecchetti S, Nascetti G. 2010. Birth of a hotspot of intraspecific genetic diversity: notes from the underground. *Molecular Ecology* **19**: 5432-5451.

Carlsson J. 2008. Effects of microsatellite null alleles on assignment testing. *Journal of Heredity* **99**: 616-623.

Chaves L D, Faile G M, Krueth S B, Hendrickson J A, Reed K M. 2010. Haplotype variation, recombination, and gene conversion within the turkey *MHC-B* locus. *Immunogenetics* **62**: 465-477.

Chao A, Jost L, Chiang S C, Jiang Y H, Chazdon R. 2008. A two-stage probabilistic approach to multiple-community similarity indices. *Biometrics* **64**: 1178-1186.

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

Dewoody J, Nason J D, Hipkins V D. 2006. Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes* **6**: 951-957.

Dewoody Y D, Dewoody J A. 2005. On the estimation of genome-wide heterozygosity using molecular markers. *Journal of Heredity* **96**: 85-88.

Ekblom R, Sæther S A, Fiske P, Kålås J A, Höglund J. 2010. Balancing selection, sexual selection and geographic structure in MHC genes of Great Snipe. *Genetica* **138**: 453-461.

Excoffier L. 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology* **13**: 853-864.

Excoffier L, Foll M, Petit R J. 2009. Genetic consequences of range expansions. *The Annual Review of Ecology, Evolution and Systematics* **40**: 481-501.

Excoffier L, Lischer H E L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**: 564-567.

Fahrig L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution and Systematics* **34**: 487-515.

Frankham R, Ballou J D, Briscoe D A. 2009. Introduction to conservation genetics. *Cambridge University Press,* Cambridge.

Goudet J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**: 485-486.

Hanski I. 1991. Metapopulation dynamics: brief history and conceptual domain. *Biological Journal of the Linnean Society* **42**: 3-16.

Holderegger R, Kamm K, Gugerli F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology* **21**: 797-807.

Höglund J. 2009. Evolutionary conservation genetics. *Oxford University Press*, Oxford, UK.

Jost L. 2008. G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* **17**: 4015-4026.

Lacy R C. 1987. Loss of genetic diversity from managed populations: Interacting effects of drift, mutation, immigration, selection and population subdivision. *Conservation Biology* **1**: 143-158.

McVean G A T. 2002. A genealogical interpretation of linkage disequilibrium. *Genetics* **162**: 987-991.

Meirmans P G, Hedrick P W. 2011. Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources* **11**: 5-18.

Montadert M, Léonard P. 2003. Survival in an expanding hazel grouse *Bonasa bonasia* population in the southeastern French Alps. *Wildlife Biology* **9**: 357-364.

Montadert M, Léonard P. 2006. Post-juvenile dispersal of Hazel Grouse *Bonasa bonasia* in an expanding population of the southeastern French Alps. *Ibis* **148**: 1-13.

Moritz C. 1994. Defining 'Evolutionary Significant Units' for conservation. *TREE* **9**: 373-375.

Nei M. 1972. Genetic distance between populations. *The American Naturalist* **106**: 283-292.

Nei M. 1987. Molecular Evolutionary Genetics. *Columbia University Press*, New York, NY, USA.

Opdam P, Wascher, D. 2004. Climate change meets habitat fragmentation: linking landscape and biogeographical scale levels in research and conservation. *Biological Conservation* **117**: 285-297.

Petit R J, Aguinagalde I, Beaulieu J L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Marti J P, Rendell S, Vendramin G G. 2003. Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* **300**: 1563-1565.

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

Reed D H, Frankham R. 2003. Correlation between fitness and genetic diversity. *Conservation Biology* **17**: 230-237.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**: 1219-1228.

Sahlsten J, Thörngren H, Höglund J. 2008. Inference of hazel grouse population structure using multilocus data: a landscape genetic approach. *Heredity* **101**: 475-482

Segelbacher G, Paxton R, Steinbrueck G, Trontelj P, Storch I. 2000. Characterisation of microsatellites in capercaillie (*Tetrao urogallus*) (AVES). *Molecular Ecology* **9**: 1934-1935.

Slatkin M. 2005. Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology* **14**: 67-73.

Sokal R R, Wartenberg D E. 1983. A test of spatial autocorrelation analysis using an isolation-by-distance model. *Genetics* **105**: 219-237.

Spielman D, Brook B W, Frankham R. 2004. Most species are not driven to extinction before genetic factors impact them. *PNAS* **101**: 15261-15264.

Strand T, Westerdahl H, Höglund J, Alatalo RV, Siitari H. 2007. The Mhc class II of the Black grouse (Tetrao tetrix) consists of low numbers of B and Y genes with variable diversity and expression. *Immunogenetics* **59**: 725-734.

Strand T, Höglund J. 2008. Multiple Black grouse MHC loci genotyped using 96-well RSCA. *In preparation.*

Storch I. 2000. Grouse status survey and conservation action plan 2000-2004. *WPA/BirdLife/SSC Grouse Specialist Group. IUCN, Gland. Switzerland and Cambridge, UK.*

Van Oosterhout C, Hutchinson W F, Wills D P M, Shipley P. 2004. Micro-Checker: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**: 535-538.

Van Oosterhout C, Weetman D, Hutchinson W F. 2006. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes* **6**: 255-256.

Weir B S, Cockerham C C. 1984. Estimating F-Statistics for the analysis of population structure. *Evolution* **38**: 1358 – 1370.

Wright S. 1943. Isolation by distance. *Genetics* **28**: 114.

Appendix 1

DNA extraction protocol

For one (1) sample of tissue (foot pad scrape) of approx. 10-30 mg, use:

- 350 µl SET-buffer (0.15 M NaCl, 0.05 M Tris, 1 mM EDTA pH 8.0)
- 12.5 µl Proteinase K (10 mg/ml)
- 15.5 µl SDS (25%)
- 300 µl NaCl (6M, saturated)
- 150 µl Tris (0.01 M, pH 8.0)
- 750 µl 99.5 % EtOH
- 100 µl TE-buffer (pH 7.6)

Cut a thin strip of foot pad scrape and dissolve in a 1.5 ml Eppendorf-tube of 350 µl SET-buffer, 12.5 µl ProtK, and 15.5 µl SDS.

Incubate for 2 h in 55 ºC or 1 h in 60 ºC. Check after approx. 30 min that tissue is dissolving. If not, add additional ProtK.

Add 300 µl NaCl, vortex 10-20 seconds.

Centrifuge 10 min at 13000 rpm.

Transfer 600 µl of the supernatant to a new tube. (The previous tube and remaining pellet can be discarded when extraction and purification is confirmed.)

Add 150 µl Tris. Mix thoroughly by flipping tube over a few times.

Add 1 volume (750 µl) of 99.5% EtOH (-20 ºC). Mix.

Let DNA precipitate overnight at -20 ºC or 2 h at -70 ºC.

Centrifuge 15 min at 10700 rpm.

Discard the supernatant and wash with 1 ml of 70% EtOH (-20 ºC).

Centrifuge 10 min at 10700 rpm.

Discard the supernatant and dry the pellet overnight (or with SpeedVac for 30 min).

Dissolve the sample in 100 µl TE-buffer.

Measure DNA concentration and contamination in NanoDrop and store sample at -20 ºC long-term, or at 4 ºC short-term.

Appendix 2

Microsatellite PCR protocol

Table 7. Microsatellite PCR reactions.

Locus	Dye	Multiplex	PCR-program	Cycles
ADL230	FAM	1	T: 48 ºC	40
ADL257	HEX	1	T: 48 ºC	40
ADL142	HEX	2a	T: 54 ºC	40
ADL184	NED	2a	T: 54 ºC	40
				40
BG15	FAM	2b	T: 54 ºC	40
BG16	HEX	2b	T: 54 ºC	40
BG18	NED	2b	T: 54 ºC	40
TUT1	NED	3a	T: 60 ºC	40
TUT3	FAM	3a	T: 60 ºC	40
TUT2	HEX	3b	T: 60 ºC	40
TUT4	FAM	3b	T: 60 ºC	40

Table 8. Multiplex PCR-kit.

Component	Volume / reaction	x100 reactions
Reaction mix	5 µl	500 µl
PCR-mix	1 µl	100 µl
10x primer mix	1 µl	100 µl
Rnase-free water	1 µl	100 µl
Q-solution	1 µl	100 µl
		8 µl in each well
Template DNA	2 µl	2 µl
Total volume	10 µl	10 µl

Multiplex PCR program

Initial activation step, 1 cycle: T: 95 ºC, 15 min. / Anneal and extension, 40 cycles: T: 94 ºC, 30 sec. T: anneal, 90 sec. T: 72 ºC, 60 sec. / Final extension, 1 cycle: T: 60 ºC, 30 min. / Stand-by, eternity: 4 ºC.

MHC PCR protocol

Construction of FLRs

Table 9. FLR PCR reaction volumes.

Concentration	Volume (µl)		Mix (µl)
	10	10*buff utan MgCl2	**100**
10mM (2,5mM av varje)	6	dNTP's	**60**
10uM	5	*RNA F flour 1a*	**50**
10uM	0.5	*RNA R 1a*	**5**
50mM	6	MgCl2	**60**
5u/ul	0.6	Biotaq	**6**
	0	tillsatser	**0**
	1	DNA	
tot-H20	29.1		
	70.9	ddH20	**709**
totV	100.0		
Mix per sample:	1		

FLR PCR program

Initial activation step, 1 cycle: T: 94 ºC, 15 min. / Anneal and extension, 30 cycles: T: 94 ºC, 60 sec. T: anneal: 64.9 ºC, 30 sec. T: 72 ºC, 30 sec. / Final extension, 1 cycle: T: 72 ºC, 10 min. / Stand-by, eternity: 4 ºC.

MHC PCR program

Initial activation step, 1 cycle: T: 94 ºC, 15 min. / Anneal and extension, 35 cycles: T: 94 ºC, 60 sec. T: anneal: 64.9 ºC, 30 sec. T: 72 ºC, 30 sec. / Final extension, 1 cycle: T: 72 ºC, 10 min. / Stand-by, eternity: 4 ºC.

Table 10. MHC PCR reaction volumes.

Concentration	Volume (µl)		Mix (µl)
	2.5	10*buff utan MgCl2	**70**
10mM (2,5mM av varje)	1.5	dNTP's	**42**
10uM	1.2	*RNA F 1a*	**33.6**
10uM	1.2	*RNA R 1a*	**33.6**
50mM	1.5	MgCl2	**42**
5u/ul	0.15	Biotaq	**4.2**
	0	tillsatser	**0**
	1	DNA	
tot-H20	9.05		
	15.95	ddH20	**446.6**
totV	25.0		
Mix per sample:	24		

Hybridization PCR program

Initial activation step, 1 cycle: T: 95 ºC, 10 min. / Anneal of single strands, 1 cycle: T: 55 ºC, 15 min. / Final anneal, 1 cycle: T: 4 ºC, 15 min. / Stand-by, eternity: 4 ºC.

Dilute each well (containing hybrids of FLR and MHC alleles) with 4 µl ddH_2O.

Prepare a MegaBace plate with 0.45 µl ET400R size standard and 7.55 µl of ddH_2O per well (make a mix first for the entire plate).

Transfer 2 µl of the hybridization solution to the MegaBace plate max. a few hours before capillary gel electrophoresis.