Genetic variation within the endangered Irish red grouse (*Lagopus lagopus scoticus*)

Magnus Johansson

Degree project in biology, Master of science (2 years), 2010
Examensarbete i biologi 30 hp till masterexamen, 2010
Biology Education Centre and Department of Population Biology and Conservation Biology, Uppsala University
Supervisors: Prof. Jacob Höglund and Dr. Barry J. McMahon
Summary

The Irish red grouse (*Lagopus lagopus scoticus*) is believed to exist in low numbers and as such it faces the risk of extinction, as fitness might be reduced through inbreeding and genetic drift, along with the ability to cope with environmental changes is hampered. In this thesis 19 microsatellite markers were used to investigate the genetic status of 89 georeferenced samples of Irish red grouse. These populations roughly consist of individuals who reside in the following areas: Cork, Munster, Northwest and Wicklow, although some admixture was detected. The genetic variation within Ireland is low in comparison to Scottish red grouse and comparable to some threatened European grouse populations. Pair-wise FST-values within Ireland ranged between 0.019 and 0.040 and the overall average of allelic richness was 5.5. Effective population size in the Munster area was estimated to be 62 individuals. Wicklow was the most variable population with an AR-value of 5.4 alleles / locus. Local (Munster) neighborhood size was estimated to 31 individuals corresponding to a dispersal distance of 31km. The observed structure within Ireland’s red grouse is fairly weak. Fragmentation and destruction of habitats need to be prevented in order to preserve the red grouse as one of the few species considered to be an Irish sub-species and as an appreciated game bird.
Introduction

There are two reasons why genetics may influence the survival of a species: 1) lower fitness due to inbreeding and genetic drift (random loss of alleles) and 2) inability to evolve in a changing environment due to loss of genetic variation (Frankham et al. 2007, Höglund 2009). Inbreeding is bound to take place in small natural populations and especially among species where suitable habitats have become fragmented (Hartl & Clark 1997). As inbreeding (and genetic drift) increases the level of genetic variability decreases. This, in turn, affects the viability of new individuals, which perhaps gives rise to even more inbred offspring which have a higher mortality and lower reproduction success. This concept is known as the extinction vortex and was first put forward by Gilpin & Soulé (1986). Habitat destruction and exploitation will work as a synergic force and hasten the effects of the extinction vortex. When genetic variation is lost within a population two major ways can restore it, new mutations and gene flow (the moving of individuals between populations) (Frankham et al. 2007). Since mutation is a slow acting force on the genome, facilitating gene flow is the only feasible conservation action.

During the last decade the use of genetics has increased our understanding of how the genetic profile of a species or population is geographically distributed and its relation to increased extinction risk (Pertoldi et al. 2007). It also brings insight to the finer spatial processes and how the environment facilitates the shaping of populations (Manel et al. 2003). Earlier research on Scottish red grouse has shown that their movement range can be limited by geographical features such as rivers (Piertney et al. 1998). Genetic structural patterns in Swedish hazel grouse (Bonansa bonansia) have shown how a presumed unrestricted gene flow might be an indication of historical reinvasion (Sahlsten et al. 2008). Another use is to correlate a putative population with a specific environmental type and might therefore be helpful in the identification of management units (MU) or evolutionary significant units (ESU) (Crandall et al. 2000). Different genetic markers are used in such studies, e.g. AFLP and microsatellites. Microsatellites, or short tandem repeat polymorphisms (STRP), are repeated sequences of 2 – 6 base pairs (Hartl & Clark 1997) and are considered to be neutral, i.e. mutations in these regions do not affect the individual’s fitness and is therefore not a part of natural selection (Kimura 1968). Neutral markers are well suited to landscape genetics since the changes are not influenced by natural selection but rather show population differentiation.

The red grouse (Lagopus lagopus scoticus) is one of four bird species that is considered to be endemic to Ireland and the only grouse species (Dempsey & O’Clery 2002). With an estimated dwindling number of 4,200 birds scattered around the Irish bogs and moors, it is also listed as a threatened species (Cummins et al. in preparation). As a game bird, and one of Ireland’s native species, it has long held a high status and several forces are now trying to help preserve the red grouse on Ireland. Although research has been carried out on the Irish red grouse since the early 1900’s (Allen et al. 2004), focus has more and more shifted towards the genetic status. The red grouse is a bird that depends on peatland covered by a fair amount of heather. 0.5% and 3% of the breeding birds observed on raised bogs and montane blanket bogs, respectively, is red grouse (Bracken et al. 2008). Red grouse is also closely tied to the presence and abundance of ling heather (Finnerty et al. 2007), Calluna vulgaris, for food and protection. Studies have shown that the diet can be made up of around 90% ling heather (Lance & Mahon 1974).

Since snow on Ireland and Scotland is relatively rare the red grouse is adapted to such an environment and does not molt into a white winter plumage, as its relative the Willow
Ptarmigan (*Lagopus lagopus*) does. A disputed, subtle difference between the Scottish and the Irish red grouse is tied to the different habitats they live in. Habitats in Scotland are often dominated by *C. vulgaris* which in turn gives the Scottish red grouse a darker color than the Irish ones (Hutchinson 1989) which lives in areas with more grass. This difference in plumage color is one reason that some scientists believe the Irish red grouse to be its own subspecies under *L. lagopus*, called *L. lagopus hibernicus* (Potapov 1985). Freeland *et al.* 2007 found no clear genetic differentiation between the Irish and the Scottish red grouse, at least none that would give the Irish birds the status of subspecies rather than a variation of a subspecies. Although, in their analysis only fairly short mitochondrial DNA, which is inherited maternally, were considered and therefore might not give an accurate description of the genetic relationship between the two areas.

A substantial body of research has been carried out on the Scottish red grouse concerning mating behavior, susceptibility to parasites and vulnerability to predators (see Piertney *et al.* 1998, Piertney *et al.* 2008, Thirgood & Redpath 2008). Although some of the work might be applicable to the Irish birds, there are huge differences in the densities of birds in the two countries. Where Ireland, according to a survey made during 2006 – 2008 (Cummins *et al.* in preparation) has 1.1 birds km$^{-2}$, Scotland will have close to 60 birds km$^{-2}$ or more in some areas (Thirgood *et al.* 2002). The reason behind this is that landowners in Scotland manage their land in order to maximize the production of red grouse (to receive income from hunting) (Tharme *et al.* 2001) while Irish landowners focus on planting forest and keeping livestock (which results in habitat destruction).

This paper is mainly focusing on the genetic status of red grouse within Ireland. Four questions have been asked:

1. Do Irish grouse show any signs of population genetic clustering?
2. Are there any patterns of genetic isolation by distance?
3. What is the effective population size and neighbourhood size of Irish grouse?
4. Have Irish grouse lost genetic variation in comparison to published records from other European grouse populations?

In addition to these questions, a brief comparison between Irish and Scottish red grouse has been made. This comparison is not designed to decide whether or not the Irish red grouse should be called *L. lagopus scoticus* or *L. lagopus hibernicus*, but rather provide insight into the problem.

**Materials & Methods**

**Sampling**

Through the efforts of many bird watchers, hunters and hikers, 147 feather samples of red grouse were collected all over Ireland between the years 2006 and 2009. Of these, 80 were from shot birds, 65 were collected feathers, one sample was taken from a stuffed bird (mounted in 1998) and one sample was missing a reference. Most samples were sent via mail to Dr. Barry John McMahon at the UCD School of Agriculture, Food Science and Veterinary Medicine, University College Dublin, Ireland. Feathers were stored at room temperature. Each sample was given an ID (1 – 146) and as much additional information as possible about the sampler and where the sample was collected. In some cases the contributor chose to be anonymous and provided only a vague reference to the location. Often an Irish map grid
Aberdeen, Scotland. The Scottish samples arrived as already extracted DNA and were Wicklow (W) and West Ireland (WT) (figure 1). For genetic comparisons an additional 28 putative populations were created and called: Cork (C), North West (NW), Munster (M), Wicklow (W) and West Ireland (WT) (figure 1). For genetic comparisons an additional 28 samples from Scottish red grouse were provided by Dr. Stuart Piertney, University of Aberdeen, Scotland. The Scottish samples arrived as already extracted DNA and were therefore not treated the same as the Irish samples (concerning DNA extractions).

![Map of Ireland showing sample locations](image)

**Figure 1.** Map over Ireland showing all the collected samples (left) and those successfully genotyped (right). In the left map, the five original putative populations are circled. In the right map the four clusters found by Structure are shown. From north to south in the left map is: Northwest (NW), West (WT), Wicklow (W), Munster (M) and Cork (C). From north to south in the right map is: Northwest, Wicklow, Munster and Cork. The map is used with permission from Urban Institute of Ireland, University College Dublin, Ireland.

**DNA extraction**

Following the DNA extraction protocol from Freeland *et al.* (2007) concerning museum samples, a small cut of the tip (2 – 5 mm long) of the feather was made using a single-use razor blade. The sample was then divided into two and placed in a sterile 2 ml Eppendorf tube along with two sterile steel balls. Samples were then frozen in liquid nitrogen and shaken for one minute at 2500rpm to crush the feather. Afterwards 180µl of QIAGEN DNeasy ATL buffer was added together with 20µl of proteinase K and 20µl 1M DTT. Samples were then mixed and incubated at 55°C for a minimum of 12h. After vortexing for 15 seconds 300µl of AL buffer and 1µl of carrier RNA was added. Incubating once again for 10 min at 70°C the samples were then mixed. After adding 300µl ethanol to each sample, the mixture was pipetted into DNeasy spin columns and spun down at 8000rpm for 1 minute. The flow-through was discarded and the column placed in a new tube. A volume of 500µl of buffer AW1 was added and the column centrifuged at 8000rpm for 1 minute. After discarding the flow-through, 500µl of buffer AW2 was added and the spin column centrifuged at 13 000rpm for 3 minutes. DNA was eluted (twice) from the spin columns by adding 100µl of AE buffer and centrifuged for 1 min at 8000rpm.
The DNA concentration in the first elution was measured with NanoDrop. Concentrations over 10ng/µl were regarded as good and were run through PCR without further process. Samples containing lower amount of DNA were re-extracted. If the 260/280 ratio was regarded good, the concentration of the sample was increased by mixing one part DNA with one part 5M ammonium acetate and two parts 98% ethanol. The sample was centrifuged for 10 minutes at 13 000 rpm. Afterward the sample was washed with 75% ethanol and centrifuged as before. Samples were dried at room temperature and stored in water. If the DNA concentration was regarded poor along with the 260/280 ratio, the sample was still included in further work, but no further re-extractions or altering of concentrations were made.

**PCR and genotyping**

The 19 microsatellite loci were selected based on previous studies done on grouse species (Piertney & Dallas 1997, Sahlsten et al. 2008) and amplified using PCR (Appendix 1). The markers were labeled with fluorescent dye (HEX, FAM and NED). Using QIAGEN Multiplex mix PCR were performed on all samples with the following master mix: 5µl of multiplex mix; 2µl of ddH₂O; 1µl of Q-solution; 1µl of primer mix and 1µl of DNA. Primer mix was made by mixing 10µl of each 20µM primer stock (forward and reverse) and then adding ddH₂O to a total volume of 100µl. PCR conditions for each multiplex were as follows: denaturation at 95°C for 15 min; 40 cycles of; 94°C for 30s: annealing for 90s and 72°C for 60s. Followed by a final extension for 30 minutes at 60°C. Annealing temperatures for each primer is found in appendix 1.

PCR products were analyzed on a MegaBACE 1000. Each sample was prepared for genotyping by mixing 2µl of diluted (10 times) PCR product, 7.8µl ddH₂O and 0.2µl size standard. The output was analyzed with the software Fragment Profiler (Fragment Profiler 1.2, Amercham Biosciences, 2003). Scoring of microsatellites was done automatically using a constructed peak filter but each score was verified manually.

**Data analysis**

Of the 147 Irish samples collected, 89 were genotyped with three or less missing genotypes. 57 samples were hence discarded and no further analysis was made including these. All but one of the 28 Scottish samples were successfully genotyped. All genotypes included in the data analysis have been verified at least twice. No identical genotypes were found.

**Population genetics**

To check for null alleles, stuttering and large allelic dropout, the dataset was run through MicroChecker 2.2.3 (Van Oosterhout et al. 2004). A null allele is generally defined as an allele that has not amplified during PCR. Stuttering is when there are slight changes in allele size and large allele dropout occurs when a large allele do not amplify as successfully as a small one. Each putative population was run separately. The only locus that consistently was found to contain null alleles was TUT4 (* in appendix 1) and this was therefore removed from all further analyses (giving a total of 18 loci). Microsatellite Toolkit (Park 2001) was used to create the input file to FSTAT and to get an initial handle on the data by performing summary statistics.

The calculation of Weir and Cockerham’s (1984) pair-wise Fₜₜ was done by considering all possible pairs of populations and assuming that they constitute the entire population in each calculation (Höglund 2009) and was done using Genetix 4.05.2 (Belkir et al. 2000). Allelic
richness (AR) across loci was obtained using FSTAT 2.9.3.2 (Goudet 2001). Allelic richness is the rarified number of alleles in a population (El Mousadik & Petit 1996), which means that is it normalized to the smallest complete sample number (here 11). This is done in order to be able to compare different sized populations since the number of alleles depends on the population size. Expected and observed frequencies of heterozygotes ($H_e$ and $H_o$, respectively), for all loci, were obtained using the software Genetix. Expected heterozygocity is calculated as according to Nei (1978). The $F_{IS}$-value seen in table 4 is calculated by taking:

$$F_{IS} = \frac{H_e - H_o}{H_e}$$  \hspace{1cm} (1)

Genetix was also used to create an AFC 2D plot. AFC 2D is a factorial component analysis and shows the multidimensional relationship between each individual genotype in a two dimensional plot. Effective population size ($N_e$) is defined as the size of an ideal population displaying the same rate of genetic drift and (or) inbreeding as the real population under study (Höglund 2009) and here it is calculated via the software LDNe (Waples 2006). Since the gathered genetic material constitutes a present “genetic snapshot” the program uses a method to estimate $N_e$ by inferring it from the observed levels of linkage disequilibrium.

**Spatial distribution**

Locations where the samples were collected can be seen in figure 1. Points belonging to the samples without a GPS-coordinate are arbitrarily placed within the area close to the location. Coordinates were written down in the UTM (Europe 29N) system.

Structure is a software that uses a model based approach of assigning individuals to a specific cluster (Pritchard et al. 2000). The admixture model is built around the assumption that each cluster (putative population) can be described using $X$, $Z$, $P$ and $Q$. $X$ denotes the sampled genotype of the different individuals, $Z$ and $P$ denotes the unknown populations of origin and the frequencies of alleles in all populations, respectively, and $Q$ denotes the admixture proportion for each individual. The model then works by introducing population structure in order to account for HWE or linkage disequilibrium (LD) and then tries to find clusters (Pritchard et al. 2000). When inferring genotypes to different clusters, Structure uses a Bayesian approach by constructing model priors for $Z$, $P$ and $Q$ and then computing the posterior distribution with the help of a Markov chain Monte Carlo (MCMC). When the values of $Z$, $P$ and $Q$ are known Structure estimates

$$Pr(K|X) \propto Pr(X|K) Pr(K)$$  \hspace{1cm} (2)

by substituting $Pr(X|K)$ with

$$Pr(X|K) \approx \exp \left( \frac{-\hat{\mu}^2 - \sigma^2}{2} \right)$$  \hspace{1cm} (3)

$\hat{\mu}$ and $\hat{\sigma}$ denotes the functions of $M$, $X$, $Z$, $P$ and $Q$ (Pritchard et al. 2000), $M$ being the number of genotypes. This constitutes an ad hoc way of solving the posterior distribution of $K$ and is riddled with assumptions that may not be valid but will in practice give sensible answers (Pritchard et al. 2000). As output Structure gives the value of $\ln P(X|K)$ and to infer the true $K$ one often uses the maximal value of $\ln P(X|K)$. Evanno et al. (2005) suggest a way
of inferring the true K with the help of \( \ln P(X|K) \) but adjusting for an increase in variance as K increases and looking at the modal value of \( \Delta K \) (figure 2).

When running Structure the samples were arranged, alphabetically, according to their putative populations seen in figure 1. A pilot analysis of ten runs with a burn-in of 20 000 and 100 000 iterations, while using the admixture model with and without loc prior, concluded that the most interesting K-value would lie between 2 – 5. A more thorough analysis was run with admixture with and without loc prior for 50 runs at K = 2 – 5 with a burn-in of 50 000 and 100 000 iterations. All files in the produced result folder were then zipped and run through Structure Harvester (http://taylor0.biology.ucla.edu/struct_harvest/) to obtain the graphs showing the \( \ln P(X|K) \) for each K (figure 1). When assigning an individual to a cluster in Structure the clusters do not appear in the same order each run. This phenomenon is known as “label switching” and is due to the fact that the order of the clusters is randomized. In order to make an average over several runs (50) the output files have to be aligned. CLUMPP is an abbreviation for CLUstering Matching and Permutation Program and is a program developed by Jacobsson & Rosenberg (2007) that does this automatically. Running Structure Harvester will produce ready-made CLUMPP input files for each K. The averaged Structure output is then visualized using the software Distruct (Rosenberg 2004).

In order to describe the relationship between relatedness and distance, the software SPAGeDi 1.3a (Hardy & Vekemans 2002) was used. When calculating relatedness over a distance it is appropriate to define distance classes. I used ten classes which are defined in the program in such a way that there are approximately the same numbers of pair-wise comparisons within each class (Hardy & Vekemans 2002). The relatedness coefficient (r) (Queller & Goodnight 1989) is used to indicate relationships within each distance class. Neighborhood size, NS, can be estimated via the slope and intercept of a regression line made over all distance classes, when the distance is log-transformed (Hardy & Vekemans 2002).

\[
NS \approx -\frac{1 - F}{b \log}
\]  

(4)

In the equation above, \( F \) is the mean jack-knifed inbreeding coefficient and in the output from SPAGeDi represented by the intercept, \( b \log \) is the slope of the regression with the distance ln-transformed. Within NS individuals are assumed to be more related to each other than outside. When NS is known together with the effective population density (\( D = N_e / \text{Area} \)), the mean axial dispersal distance (\( \sigma \)) can be inferred by

\[
\sigma = \sqrt{\frac{NS}{4\pi D}} = \sqrt{\frac{NS \cdot \text{Area}}{4\pi N_e}}
\]  

(5)

The area of the region where the sampling took place was estimated via ArcMap 9.2 by calculating the area of a polygon covering mountainous land.

**Results**

Only 61% of the Irish samples were successfully genotyped with three or less missing microsatellite loci. Of these 89(146) samples, 74(80) were from shot birds and 15(65) were from collected feathers with the number in brackets corresponding to the full dataset of 146 samples (one sample missing a reference).
**Variation within Ireland**

Processing 50 runs on \(K=1 - 5\) with a burn-in of 50,000 and 100,000 iterations in Structure Harvester produced the graphs shown in figure 2. On the left is the mean of \(\ln P(X|K)\) with standard deviations for each \(K\). On the right is the Evanno method showing \(\Delta K\). Both indicate \(K=4\) as the most likely value.

![Graphs showing mean of lnP(X|K) and Evanno method for K=1-5](image)

Figure 2. Summarized result from 50 Structure runs with 50,000 burn-in and 100,000 iterations in an admixture model with no loc prior. To the right is the \(\Delta K\) graph accordingly to the Evanno method.

Figure 3 shows the averaged Structure output from 50 runs on \(K=4\), after being run through CLUMPP and Distruct. Figure 1 shows the putative populations (left) before and the joined putative populations (right) based on the outcome of Structure. The putative populations Northwest and West Ireland are joined under the label Northwest. In the following analysis the four putative populations are used and not the original five. 18 samples are labeled as Cork (C), 19 as Munster (M), 15 as Northwest (NW) and 37 as Wicklow (W).

![Averaged Structure output from 50 runs with 50,000 burn-in and 100,000 iterations](image)

Figure 3. The averaged output from 50 Structure runs with 50,000 burn-in and 100,000 iterations, admixture model without loc prior, \(K=4\). The clusters are arranged in alphabetical order and accordingly to the five putative populations.
Another way to visualize the four different grouping is to make an AFC 2D plot (figure 4). Axis one and two represent the two major factors that contribute to the separation of Cork, Munster, Wicklow and the joined putative population Northwest. Together they explain 7.07% of the variation. Although tightly grouped together there is a separation among the geographic areas.

![AFC 2D plot showing the 89 samples colored after the four putative populations. Axis 1 and 2 constitutes the two major components and they explain 7.07% of the variation observed in the data set. Northwest and West Ireland are grouped under “Northwest”.

Figure 4. AFC 2D plot showing the 89 samples colored after the four putative populations. Axis 1 and 2 constitutes the two major components and they explain 7.07% of the variation observed in the data set. Northwest and West Ireland are grouped under “Northwest”.

Pair-wise $F_{ST}$-values (Weir & Cockerham 1984) for the four putative populations within Ireland are shown in table 2. Pair-wise $F_{ST}$-values ranged from 0.019 (Munster – Wicklow) to 0.040 (Munster – Northwest). The global $F_{ST}$-value for Ireland is 0.028 (95% CI = 0.016 – 0.039). Due to limitations in the dataset, the effective population size was only calculated for the Munster sampling area and is estimated to be 62 (95% CI= 33 – 247) individuals.

<table>
<thead>
<tr>
<th></th>
<th>Munster</th>
<th>Northwest</th>
<th>Wicklow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cork</td>
<td>0.038</td>
<td>0.037</td>
<td>0.027</td>
</tr>
<tr>
<td>Munster</td>
<td>0.040</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>Northwest</td>
<td></td>
<td>0.022</td>
<td></td>
</tr>
</tbody>
</table>

Minimum, maximum and mean number of alleles (A) and allelic richness (AR), over all loci for each putative population and overall, are found in table 3. Values for allelic richness within each putative population ranges from 1 – 10.5 alleles per locus, with averaged values ranging from 4.8 to 5.4 alleles per locus. Overall average value for AR is 5.5 alleles per locus. Unbiased $H_e$ ranged from 0.63 – 0.67 with the Wicklow cluster in the top. Wicklow also had the top value of $H_s$ (0.68) which ranged from 0.64 – 0.68. Only two reliable values for Ne were calculated, Munster and Wicklow, 62 and 151 respectively. Cork provided a negative value and Northwest had an upper 95% CI of infinity. An overall estimate of Ne was calculated by taking the harmonic mean of the four region estimates and multiplying it with four, giving 456 (95% CI = 47 – $\infty$).
Table 3. Summary of unbiased $H_e$ (Nei 1978), $H_o$ and average values of allelic richness (AR) for the four putative populations as well as overall values. AR-values are rarefied to a sample size of $N=11$. $N_e$-values are presented with Jack-knifed 95% CI. The overall $N_e$-value is calculated as four times the harmonic mean over all areas.

<table>
<thead>
<tr>
<th>Area</th>
<th>$H_e$</th>
<th>$H_o$</th>
<th>AR</th>
<th>$N_e$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cork</td>
<td>0.63</td>
<td>0.64</td>
<td>4.8</td>
<td>-785 (93.6 – $\infty$)</td>
</tr>
<tr>
<td>Munster</td>
<td>0.65</td>
<td>0.64</td>
<td>4.8</td>
<td>62 (33.6 – 248.8)</td>
</tr>
<tr>
<td>Northwest</td>
<td>0.64</td>
<td>0.66</td>
<td>5.1</td>
<td>74 (32.1 – $\infty$)</td>
</tr>
<tr>
<td>Wicklow</td>
<td>0.67</td>
<td>0.68</td>
<td>5.4</td>
<td>151 (87 – 465.7)</td>
</tr>
<tr>
<td>Overall</td>
<td>0.67</td>
<td>0.62</td>
<td>5.5</td>
<td>456 (47 – $\infty$)</td>
</tr>
</tbody>
</table>

Relatedness decreased with distance as shown in figure 5; note that all samples genotyped were included in the graph. Local (Munster) neighborhood size was estimated to 32 individuals using the values: $F = 0.198$ and $b_{log} = -0.025$ (intercept and slope, respectively, from a regression analysis made on relatedness data from Munster) and equation 4. Dispersal distance within the Munster sampling region is estimated to 5.6 km using these values: $NS = 32$, $N_e = 62$ and $Area = 756 \text{ km}^2$ in equation 5.

Figure 5. Relatedness graph built on the output from the software SPAGeDi 1.3a (Hardy & Vekemans 2002), showing how relatedness differs with distance. Based on all samples.

Table 4 is a compilation of values of allelic richness, $F_{is}$ and observed heterozygocity ($H_o$) derived from three articles and this paper, concerning the genetic status of endangered and not endangered European populations of black grouse ($Tetrao tetrix$) and capercaillie ($Tetrao urogallus$). As comparison, values from red grouse in Scotland are included. Scottish red grouse has the highest $F_{is}$- and $H_o$-values, while the Irish red grouse has among the highest values of allelic richness (table 4). Bear in mind that the AR values are not directly comparable.
Table 4. Comparisons of values of allelic richness (AR), Weir & Cockerham (1984) $F_{is}$ ($F_{is}$) and observed heterozygocity ($H_o$) found in threatened and not threatened European red grouse, black grouse (*Tetrao tetrix*) and capercaillie (*Tetrao urogallus*) populations. † denotes threatened populations. $F_{is}$ - and $H_o$-values for Scotland are averaged over all loci used in the article.

<table>
<thead>
<tr>
<th>Species</th>
<th>Location</th>
<th>AR</th>
<th>$F_{is}$</th>
<th>$H_o$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red grouse*</td>
<td>Ireland†</td>
<td>5.50</td>
<td>0.028</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Scotland</td>
<td>NA</td>
<td>0.119</td>
<td>0.78</td>
</tr>
<tr>
<td>Black grouse**</td>
<td>Netherlands†</td>
<td>3.61</td>
<td>0.025</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>Norway</td>
<td>5.90</td>
<td>-0.023</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Austria</td>
<td>6.35</td>
<td>-0.056</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>England†</td>
<td>4.23</td>
<td>0.108</td>
<td>0.52</td>
</tr>
<tr>
<td>Capercaillie***</td>
<td>Spain†</td>
<td>2.45</td>
<td>-0.022</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Pyrenees†</td>
<td>3.25</td>
<td>0.091</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Archangelsk (Russia)</td>
<td>4.68</td>
<td>-0.004</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Jaroslawl (Russia)</td>
<td>4.73</td>
<td>0.071</td>
<td>0.68</td>
</tr>
</tbody>
</table>

*For Ireland values see this study, for Scottish see Piertney et al. 1997  
** Larsson et al. 2008  
*** Rodríguez-Muñoz et al. (2007)

Variation between Ireland and Scotland

Figure 6 shows an AFC 2D plot between Ireland (blue) and Scotland (red). Here there is a clear separation with no overlap. The two major factorial components explain 7.22% of the variation observed in this dataset. The Irish samples are also more tightly clustered than the Scottish samples which indicate a higher variability in the Scottish red grouse. A pair-wise $F_{ST}$-value of Ireland vs. Scotland is estimated to be 0.068 (95% CI = 0.043 – 0.098).

Figure 6. AFC 2D plot of the 89 Irish (blue) and 27 Scottish (red) samples. Axis 1 and 2 constitutes the two major factorial components and they explain 7.22% of the variation in the data set.
Discussion

Molted feathers, if in good condition, are a good source of DNA (Segelbacher 2002, Bayard de Volo et al. 2008). If in bad condition, *i.e.* degraded (mainly through exposure to UV-light), the success rate of DNA extraction goes down, which may be a reason why only 61% of the Irish samples were successfully genotyped. This success rate is comparable to that found in an earlier study where the estimated percent of reliable genotypes from collected molted feathers was 50% (Segelbacher 2002). Approximately 45% of the samples were collected feathers (molted feathers), of these only a 23% were successfully genotyped. Re-extractions were not always a solution because of the lack of feathers; often only a single feather per individual had been collected. Another possible reason to the low extraction success rate is the amount of inhibitors present during PCR. The problem of inhibitors can often be circumvented by diluting the extracted DNA but will, in turn, also lower the concentration of DNA. For these samples the amount of DNA was already low and therefore this method proved difficult. Unfortunately most of the unsuccessfully genotyped samples were collected as single feathers from isolated populations. Due to this, the geographical coverage of the samples is limited, which is especially true in the northwestern and western areas of Ireland (figure 1).

An important aspect of population analysis on genetic data is the possibility of errors occurring during PCR. Three such errors are: null alleles, stuttering and large allele dropout. A null allele is generally defined as an allele that has not amplified during PCR. Stuttering is when there are slight changes in allele size and large allele dropout occurs when a large allele do not amplify as successfully as a small allele. Such errors in the PCR will create a bias in any analysis based on genotypic data. As small populations become isolated, more or less inbreeding is bound to occur. In addition as inbreeding decreases, the amount of heterozygotes one is expected to find an increase in the amount of homozygotes (Frankham et al. 2007). Therefore it is not surprising that MicroChecker finds some null alleles as it searches for excessive amounts of homozygotes. When I tested the loci separately for each of the four putative populations (found in a pilot study) and only found one locus that was consistently classified as containing null alleles (TUT4), I chose to remove that and include all other (18) loci.

In order for Structure to assign samples to clusters the model works on the assumption that there is HWE within populations and complete linkage equilibrium between loci within populations (Pritchard et al. 2000). When Structure then forces a population structure on the samples, it searches for deviations from Hardy-Weinberg equilibrium (HWE) and the presence of linkage disequilibrium (LD) and is hence able to place samples in clusters (Pritchard et al. 2000). When deciding on the proper value of K one has to consider the simplest model possible that will capture the major structure of the data (Pritchard et al. 2009). As seen in figure 2 the log-likelihood values show two local maxima, one for K = 1 and one for K = 4. This may indicate that there is a weak structuring and many samples are very similar. Clustering these four putative populations together will in an AFC 2D plot (figure 4) explain 7.07% of the variation. It is possible that there is more structuring among the samples, especially in the area around Wicklow. The output from CLUMPP (K = 4) suggested that Cork and Munster are in essence two groups while the additional grouping may occur among Northwest and West Ireland groups. The structure run shown in figure 3 was a run without using the loc prior information. Including this would give the prior probability distributions information about where the samples were taken in relation to each other, thus possibly improving the performance (Hubisz et al. 2009). When including this information the outcome was not altered, Structure still favored K = 4. The separation of samples into clusters as indicated by Structure show only a weak differentiation. To get around this
problem and more clearly identify possible management units (MU) among the red grouse on Ireland, more microsatellite markers should possibly be added. It is also important to include more samples from small populations, mainly in the west and northwest areas of Ireland. To ensure good DNA yields and easy genotyping, samples should not be from pick-ups rather blood samples or feathers from caught birds. Since the numbers are dwindling, non-invasive sampling is to be recommended. For this experienced handlers are highly recommended since grous are very nervous when handled.

The results of this genetic study show that the Irish red grouse have indeed a low genetic variability. A recent survey (Cummings et al. in preparation) estimated the red grouse population on Ireland to be in the range of 4,200 birds (95% CI = 3,800 – 4,700). The same study also estimated that the population on Ireland has experienced a 50% decrease in numbers over the last 40 years. The Irish red grouse are therefore rightly classified as threatened under the IUCN categories (ICUN 2001). Values of allelic richness, \( H_0 \) and \( H_e \) have been estimated for threatened populations of black grouse and capercaillie around Europe (table 4). It should be said that the AR-values in table 4 are not comparable in the sense that they are not rarefied to the same population size, but since the mean of these values are quickly reached they can be used to get an idea of the situation.

Worth noting is that the Wicklow group does not appear as a homogenous group. Several of the individuals are largely made up of genotypes perhaps better suited in either of the Munster or Northwest groups. When looking at the pair-wise F\( ST \)'s, pairs including Wicklow have lower F\( ST \)-values than population combinations excluding Wicklow. This suggests that red grouse from Wicklow are in general more diverse than birds from other areas of Ireland. The population in Wicklow is very likely the most viable population on Ireland, with the highest \( H_0 \) (0.68), allelic richness (5.4 alleles/locus) and \( N_e \) (151 individuals) of all putative populations sampled (table 3). In view of this and looking at the historical distribution maps (Cummins et al. in preparation) a possible conclusion is that the population at Wicklow constitutes something similar to a source population. As such it is the most suitable population to remove birds from and relocate them to suitable but unpopulated areas or strengthen other populations around Ireland.

There are some limitations to estimating the mean axial dispersal distance (\( \sigma \)), mainly the size of the sampling area which here is roughly estimated via the map seen in figure 1. Mountainous areas were identified and size estimated in ArcMap 9.2 using a simple polygon area calculation. As seen in equation 5, a wrongly estimated area will give a wrong estimation of mean dispersal distance and since the area can be many times higher than any other entry, this error can be large. It is also important to bear in mind that since there is structure among the sampled individuals, a separate \( N_e \) has to be obtained for each putative population. \( N_e \) is one of the most important values in conservation genetics and a low \( N_e \)-value is associated with accelerated depletion of genetic variability (Pertoldi et al. 2007). In two putative populations (Cork and Northwest) estimates of \( N_e \) provided infinity as upper limit for a 95% CI, indicative of a too small sample size. \( N_e/N \)-ratio in wild populations is estimated to be in the range of 0.10 (Frankham 1995). This estimate would for Irish red grouse around Munster give a population size of 62 \times 10^2 = 620 individuals which is not an unreal number given that there is an estimated total population size of 4200 birds. In the Wicklow area this would give a total population size of around 1500 birds. The overall estimate of \( N_e \) times 10 will give a value (4560) that falls within the 95% CI of the survey stated earlier. Although you can calculate the effective population size for a meta-population (harmonic mean of the sub-populations) it is not clear what this means as it is highly dependent on the size and direction of gene flow and
if the sub-populations are of the same size (Wang & Caballero 1999). Of data gathered in the Munster and Wicklow regions, only Munster provided a reliable estimate of neighborhood size, (positive). Hence only a roughly estimated dispersal distance for Munster is given. Grouse species are typically not long distance dispersers. Mean estimated dispersal distance for Swedish hazel grouse (Bonasa bonasia) is 1514m (Sahlsten et al. 2008) and for English black grouse the estimated mean dispersal distance during early spring was 5.8 km and during fall 10.5km (Warren & Baines 2002). In comparison the estimated dispersal distance of the Irish red grouse of 5.6 km is not unreasonable and together with evidence of dispersal limitations (Piertney et al. 1997), an indication that natural dispersal between the putative populations is unlikely. Although only a reliable neighborhood size for Munster could be calculated one can get an idea of the value via figure 5. Since the values are built around distance classes a regression made on the small distances will illustrate an average population neighborhood size.

When listed, the most common threats to grouse populations in Europe are habitat destruction and habitat fragmentation (Storch 2000). Many landowners on Ireland use the heather moors to farm timber and therefore destroy large amounts of suitable red grouse habitats and as red grouse and heather go hand in hand preserving and protecting these habitats should be a top priority. Also, in order to help the Irish red grouse population, it is suitable to start with facilitating gene flow between the Irish birds and managing predators, mainly corvids and mustelides. With an estimated population size around 4000 birds and genetic variation weakly comparable to some of the threatened grouse populations in Europe, the Irish red grouse faces a hard future if nothing is done.

As seen in the comprehensive summary of the taxonomic history of the Irish red grouse in Allen et al. (2004), there is much debate whether or not the Irish birds should constitute a separate subspecies of L. lagopus. When mtDNA was used to investigate the relationship between Ireland’s and Scotland’s red grouse populations no large differentiation was found (Freeland et al. 2007). Since they only compared a short section (ca 300 bp) of the control region and mtDNA is only inherited maternally, this does not provide sufficient evidence to rule out the possibility that the Irish red grouse is a separate subspecies of L. lagopus. Although there has been many recorded and suspected introductions since the early 1900’s (Allen et al. 2004) of Scottish red grouse into Ireland, the AFC 2D plot in figure 6 do not show any overlap between the Irish and Scottish samples. Visible is also the larger genetic variation among Scottish samples, mainly due to larger population size. This does not provide a definitive solution to the “L. lagopus hibernicus-question”, although it indicates that the Scottish birds may be unsuccessful in surviving and reproducing within the Irish populations, in concordance with findings made by Allen et al. 2004. It is therefore not recommended to introduce Scottish red grouse to Ireland.

Acknowledgements
I would like to thank my two supervisors Jacob Höglund and Barry John McMahon for valuable support and teaching me the ways of science. Both shall also receive special thanks for the week spent in Dublin and its surrounding. I only wish the weather would have been better.

Acknowledgements should defiantly also go out to the two friendly lab technicians, Gunilla Engström and Reija Dufva, who have always had a moment to answer questions, no matter how trivial. Others, who have always had time over for questions, support, ideas or just a
friendly fika in the coffee room of floor 2, are my fellow master students, PhDs and post-docs, researchers and professors. In short all. No one mentioned and no one forgotten.

This genetic study would not have been possible without the help of volunteers collecting feathers all over Ireland particularly by the Irish Grey Partridge Conservation Trust, National Association of Regional Games Councils of Ireland, the Irish National Park and Wildlife Service and BirdWatch Ireland. Funding was provided by The Native Species Conservation Committee of Dublin Zoo and Fota Wildlife Park. Thank you all.

Literature


Appendix 1.

Table 1. Compilation of the microsatellite markers used arranged in multiplexes. Size is given in base pairs (bp); spacing is the repeat number of the microsatellite, also in bp.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Dye</th>
<th>Size</th>
<th>Spacing</th>
<th>Annealing T(°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiplex 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADL 230</td>
<td>FAM</td>
<td>90 - 115</td>
<td>2</td>
<td>48</td>
</tr>
<tr>
<td>ADL 142</td>
<td>HEX</td>
<td>207 - 255</td>
<td>2</td>
<td>48</td>
</tr>
<tr>
<td>Multiplex 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADL 184</td>
<td>NED</td>
<td>110 - 160</td>
<td>2</td>
<td>54</td>
</tr>
<tr>
<td>BG 15</td>
<td>HEX</td>
<td>130 - 173</td>
<td>4</td>
<td>54</td>
</tr>
<tr>
<td>BG 16</td>
<td>NED</td>
<td>130 - 170</td>
<td>4</td>
<td>54</td>
</tr>
<tr>
<td>BG 18</td>
<td>FAM</td>
<td>110 - 180</td>
<td>4</td>
<td>54</td>
</tr>
<tr>
<td>Multiplex 3a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LE1098</td>
<td>NED</td>
<td>135 - 165</td>
<td>2</td>
<td>60</td>
</tr>
<tr>
<td>TUT2</td>
<td>HEX</td>
<td>135 - 175</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>TUT4*</td>
<td>FAM</td>
<td>157 - 215</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>Multiplex 3b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUT1</td>
<td>NED</td>
<td>178 - 220</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>TUT3</td>
<td>FAM</td>
<td>150 - 180</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>Multiplex 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLSD4</td>
<td>HEX</td>
<td>185 - 220</td>
<td>2</td>
<td>58</td>
</tr>
<tr>
<td>LLSD6</td>
<td>FAM</td>
<td>88 - 126</td>
<td>2</td>
<td>58</td>
</tr>
<tr>
<td>LLSD7</td>
<td>HEX</td>
<td>140 - 176</td>
<td>2</td>
<td>58</td>
</tr>
<tr>
<td>LLSD8</td>
<td>FAM</td>
<td>138 - 170</td>
<td>2</td>
<td>58</td>
</tr>
<tr>
<td>Multiplex 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLST1</td>
<td>NED</td>
<td>120 - 170</td>
<td>3</td>
<td>54</td>
</tr>
<tr>
<td>LLSD2</td>
<td>HEX</td>
<td>100 - 117</td>
<td>2</td>
<td>54</td>
</tr>
<tr>
<td>LLSD3</td>
<td>HEX</td>
<td>123 - 145</td>
<td>2</td>
<td>54</td>
</tr>
<tr>
<td>LLSD5</td>
<td>FAM</td>
<td>120 - 140</td>
<td>2</td>
<td>54</td>
</tr>
</tbody>
</table>

*The microsatellite found to contain null alleles