Nikoleta Karaiskou[1], Alex Ball[2], Konstantinos Gkagkavouzis[3], ARISTOTELIS MOULISTANOS[3], Konstantinos Kalaentzis[3], Muhammad Ghazali[2], Styliani Minoudi[3], David Parish[4], Evaggelos Chatzinikos[5], Dimitrios Kiousis[5], Nikolaos Manios[6], Jenny Dunn[7], and Alexandros Triantafyllidis[3]

[1]Aristotle University of Thessaloniki Faculty of Sciences
[2]Royal Zoological Society of Scotland
[3]Aristotle University of Thessaloniki School of Biology
[4]Game and Wildlife Conservation Trust
[5]4th Hunting Federation of Sterea Hellas
[6]Forestry Authority of Trikala
[7]Keele University School of Life Sciences

October 23, 2024

Abstract

The grey partridge population has experienced significant declines across Europe, largely due to agricultural intensification and loss of habitat, leading to conservation actions such as Red-listing in the UK and hunting bans in Greece. The genetics of Balkan and Scottish populations remain largely unexplored; genetic analyses are essential to evaluate the impact of past restocking efforts on wild populations, as breeding between released and wild-living partridges may complicate recovery efforts. In this study, we sample wild and farmed individuals of grey partridge from the Balkans (Greece, North Macedonia) and the United Kingdom (UK) and employ 2300 SNPs, eight microsatellites and two mitochondrial markers to investigate the genetic structure and diversity of their populations and the impact of past restocking activities. We reveal a clear distinction between two clades, an Eastern and a Western, as in previous studies, with wild birds from Greece and the UK classified to each clade respectively. However, birds from North Macedonia belonged to either clade, suggesting a contact zone between the two or a genetic legacy of past restocking practices. The captive stock in Greece and the UK is clearly of Western origin, with minor introgression of the Eastern clade being detected. Finally, an informative SNP marker panel is presented that accurately assigns each individual to either the Eastern or Western clade and will serve as a valuable tool for monitoring population structure, guiding conservation efforts, and assessing the impact of restocking activities on grey partridge populations.

Authors:

Καραισκου Ν.[1,2*], Βαλλ Α.[3*], Γκαγκαουζις Κ.[1,2], Μουλιστανος Α.[1,2], Καλαεντζις Κ.[1,4], Γηαζαλι Μ.[3], Μινουδι Σ.[1,2], Παριση Δ.[5], ϗατζινικος Ε.[6], Κιουσις Δ.[6], Μανιος Ν.[7], Δυνν Θ.[8], Τριανταφψλλιδις Α.[1,2]

1 Department of Genetics, Development & Molecular Biology, School of Biology, Faculty of Science, Aristotle University of Thessaloniki, Greece

2 Genomics and Epigenomics Translational Research (GENeTres), Center for Interdisciplinary Research and Innovation (CIRI-AUTH), Balkan Center, Thessaloniki, Greece

3 RZSS WildGenes Laboratory, Royal Zoological Society of Scotland, Edinburgh, United Kingdom

4 Hydrobiological Station of Rhodes, Hellenic Centre for Marine Research, Rhodes, Greece

6 4th Hunting Federation of Sterea Hellas, Athens, Greece

not-yet-known not-yet-known

not-yet-known

unknown

7 Forestry Authority of Trikala, Greece

8 University of Lincoln, United Kingdom

***equal contribution**

Correspondence Author: nikolbio@bio.auth.gr

not-yet-known not-yet-known

not-yet-known

unknown

Abstract

The grey partridge population has experienced significant declines across Europe, largely due to agricultural intensification and loss of habitat, leading to conservation actions such as Red-listing in the UK and hunting bans in Greece. The genetics of Balkan and Scottish populations remain largely unexplored; genetic analyses are essential to evaluate the impact of past re-stocking efforts on wild populations, as breeding between released and wild-living partridges may complicate recovery efforts. In this study, we sample wild and farmed individuals of grey partridge from the Balkans (Greece, North Macedonia) and the United Kingdom (UK) and employ 2300 SNPs, eight microsatellites and two mitochondrial markers to investigate the genetic structure and diversity of their populations and the impact of past restocking activities. We reveal a clear distinction between two clades, an Eastern and a Western, as in previous studies, with wild birds from Greece and the UK classified to each clade respectively. However, birds from North Macedonia belonged to either clade, suggesting a contact zone between the two or a genetic legacy of past restocking practices. The captive stock in Greece and the UK is clearly of Western origin, with minor introgression of the Eastern clade being detected. Finally, an informative SNP marker panel is presented that accurately assigns each individual to either the Eastern or Western clade and will serve as a valuable tool for monitoring population structure, guiding conservation efforts, and assessing the impact of restocking activities on grey partridge populations.

Keywords: grey partridge, restocking practices, phylogeography, population structure

Introduction

When undertaking any reintroduction program, it is important to consider the geographic origin between released individuals and the original population. Introgressive hybridization between nonlocal translocated organisms and local wild populations can have unforeseen genetic effects on the indigenous population since it potentially affects local adaptation in the native wild population when pairing between translocated and native individuals takes place (Randi, 2008).

Although the grey partridge *(Perdix perdix)* is considered to be common and not threatened on a global scale, its populations appear to be declining in intensively farmed areas most likely due to a reduction in breeding sites and food stocks (Ewald, Sotherton and Aebischer, 2020). In Europe, since the beginning of the 20th century, there has been a drastic decline in grey partridge populations in all of the 31 countries it is found (Kuijper, Oosterveld and Wymenga, 2009). For example, its breeding abundance fell by 92% from 1967 to 2022 across the UK probably due to the intensification of agriculture (BirdTrends 2023,

www.bto.org/birdtrends). For this reason, the grey partridge has been Red-listed as a Bird of Conservation Concern and the British government launched a major programme to help partridge recovery in the UK nearly 20 years ago (Aebischer and Ewald, 2012). Population shrinkage is also observed in southern populations of grey partridge like Greece where, mainly from 1950 and onwards, there has been a drastic population decline (Thomaides, C. & Papageorgiou, 1992). For this reason, the hunting of grey partridge has been prohibited in Greece since 1984. Releases in Greece have also taken place with farmed individuals imported mainly from countries in the Balkans and Northern Europe of unknown provenance. Several studies have clearly shown that the survival of hand-reared birds is poor after their release into the wild (Liukkonen, 2006). It has been suggested that differences in the genetic adaptation of separate subspecies (*P. p. lucida* , *P. p. perdix*) to different climatic conditions of their original range could be a partial explanation for the failed introductions (Liukkonen, 2006). Therefore, it is of utmost importance that a genetic analysis of wild populations is carried out.

Genetic provenance of individuals is a key assessment criterion in the IUCN guidelines for reintroductions and translocations (IUCN 2013) as both inbreeding and outbreeding depression are potential concerns. Some genetic information is available for European populations of grey partridge; phylogeographical analysis of 227 birds from several localities based on mitochondrial DNA sequencing revealed a distinction of populations in two major clades (western and eastern) in agreement with the subspecies taxonomy (Liukkonen, 2006). However, little is known about the genetics of range-edge populations, such as those in Scotland and the Southern Balkans, Greece and North Macedonia. No individuals have been sampled so far from Scotland while a genetic study with less than 10 individuals from Greece revealed that the Greek population belongs to the eastern group (Liukkonen-Anttila *et al.* , 2002; Liukkonen, Kvist and Mykrä, 2012).

We use genetic analyses to study a Northern and a Southern wild and farmed European population of grey partridge, comparing the provenance of each. Our results will complement efforts to manage and support wild native populations by providing important insights into the status of these populations, namely on the levels of genetic diversity and the impact of past restocking activities. This could have implications for the sourcing of released birds as breeding between released stocked and wild partridges may complicate the recovery of partridge populations.

not-yet-known not-yet-known

not-yet-known

unknown

Material and Methods

Sample collection

A total of 256 samples (blood, muscle tissue, feathers or eggs) were collected from the UK, Greece and North Macedonia and used in the present analysis (Figure 1). Specifically, 61 (five sampling sites) and 56 samples from wild and farmed individuals respectively, were collected from the UK, 74 (five sampling sites) and 26 samples from wild and farmed individuals respectively, were collected from Greece and 16 samples from wild individuals (two sampling sites) were provided from North Macedonia (Table S1_ Appendix). A representative sample set from European populations previously published by (Liukkonen-Anttila *et al.* , 2002) was also provided by **the** University of Oulu (12 Finnish and five Greek individuals) and they were included in our analysis. Furthermore, six mixed-ancestry individuals, i.e. offsprings of mixed pairing between individuals of Western and Eastern origin, were analysed.

DNA Extraction

DNA was successfully extracted from all blood/muscle tissue and feather samples either following the DNA extraction protocol by Hillis et al. (1996) based on CTAB reagent for feather and muscle tissue samples or using the QIAamp DNA Mini Kit on blood samples using the manufacturer's protocol. DNA extraction methods were also developed for use with eggs from captive hens. We obtained DNA by locating and

dissecting the perivitelline layer from the egg yolk, rinsing it in phosphate-buffered saline solution and using the QIAamp DNA Investigator Kit to extract DNA. The manufacturer's 'Bone and teeth' extraction protocol was followed due to the high calcium content.

DNA sequencing

Primer pairs LPPGLU (forward) - H414 (reverse) designed to amplify a fragment of the mitochondrial control region (Liukkonen-Anttila *et al.* , 2002) and L14578 (forward) - H16065 (reverse) of the cytochrome b region (Crowe *et al.* , 2006), were used. Although successful in the largest proportion of the samples, DNA extraction from feather samples are often of lower DNA quantity and quality. For these samples, primers amplifying smaller sections of these two loci were designed. The longer successful sequences amplified from the blood/tissue samples were aligned to locate conserved regions for the primer design. Primer3 (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) was used to obtain potential primer pairs amplifying a product of roughly 400bp. The primers designed for control region were PER_CR_Frag2F (TGGTTATGCTCGACGTACCA) and PER_CR_Frag2R (ATTCCCCATACACGCAAACC) and those for Cytb were PER_CytB_Frag2F (TGATGAAACTTCGGCTCCCT) and PER_CytB_Frag2R (GGTCGGGTTGTCAACTGAGA).

The total volume of the polymerase chain reaction of the control region was 30 µl in which 100 ng of genomic DNA was amplified, using 1 unit of Qiagen Taq polymerase, 0.2 mM dNTPs, 1 pmol of each primer, 2 mM $MgCl_2$, 3 µl of 10 X Reaction Buffer and 0.3 µl of 100 X BSA. Thermal cycling amplification conditions were as follows: initial denaturation at 94°C for 2 min, followed by 33 cycles of strand denaturation at 94 °C for 1 min, annealing at 54 °C for 1 min and primer extension at 72°C for 1.5 min and a final 5 min elongation time at 72°C.

For cytochrome b, 100 ng of genomic DNA was amplified in 25 µl total volume of polymerase chain reaction, using 1 unit of Qiagen Taq polymerase, 0.2 mM dNTPs, 1 pmol of each primer, 1.5 mM $MgCl_2$, 3 µl of 10 X Reaction Buffer and 0.3 µl of 100 X BSA. Thermal cycling amplification conditions were as follows: initial denaturation at 94 °C for 5 min, followed by 35 cycles of strand denaturation at 94 °C for 1 min, annealing at 52 °C for 1 min and primer extension at 72°C for 2 min and a final 8 min elongation time at 72°C. The PCR products were purified using the Nucleospin Extra kit (Macherey-Nagel, Duren, Germany) and sequenced by Sanger method using an ABI 3730XL DNA Analyzer.

The shorter control region and cytb fragments were run using the same PCR conditions. Reaction volumes were 10 µl, containing 7 µl HotStart PCR master mix (Thermofisher Scientific), 1 µl forward primer (10µM), 1 µl reverse primer (10µM) and 1 µl template DNA at 5ng/µl concentration. Thermal cycling amplification conditions were as follows: initial denaturation at 95°C for 5 min, followed by 40 cycles of strand denaturation at 95 °C for 30 sec, annealing at 55 °C for 30 sec and primer extension at 72 °C for 1 min and a final 10 min elongation time at 72 °C.

DNA Sequence statistical analysis

A total of 224 sequences of the control region (CR) (Genbank Accession numbers: PQ431575 - PQ431798) and 228 sequences of the cytochrome B gene (*cytb*) (Genbank Accession numbers: PQ295559 - PQ295786) were aligned using ClustalW (Clustal, 1994) and manually trimmed, to have the same size, considering the polymorphic sites detected, as well as the sequencing quality, using the Geneious R10 v10.2.3 software. Data from the present analysis were combined with sequences of *P. perdix* retrieved from GenBank. More specifically, sequences of the main eastern (AF115404) and western (AF115405) haplotype (Liukkonen-Anttila *et al.* , 2002), as well as 43 more sequences (PopSet: 51449199) of the sub-region CR1 (Liukkonen, 2006) were acquired. For the analysis of*cytb* , the sequences AF02879 (Kimball and Braun, 2014), EU839466 and GU214276 (Bao *et al.* , 2010) were recovered.

A median-joining network (Bandelt, Forster and Röhl, 1999) was constructed using the software Network 5.0.0 and the frequencies of the sequences (Fluxus Technology) assuming equal weights for all mutations and setting the genetic distance parameter e to zero to restrict the choice of feasible links in the final network.

Genetic variability indices of grey partridge populations, i.e. the number of distinct haplotypes (h), haplotype diversity (Hd) and nucleotide diversity (π) values, were estimated using the combined data set of both CR and cytb with DnaSP 5 (Librado and Rozas, 2009) software.

Microsatellite statistical analysis

Observed (H_o) and expected (H_e) heterozygosity values for each locus were calculated using GENEPOP 4.0 (Rousset, 2008). Deviation from Hardy-Weinberg (HW) equilibrium was tested using Fisher's exact tests (Rousset 2008) with the same software. FSTAT 2.9.3.2 (Goudet, 1995) was used to compute inbreeding coefficient values (F_{IS}) and allelic richness (A_R). CERVUS 3.0.3 (Kalinowski, Taper and Marshall, 2007) was used to evaluate polymorphic information content (PIC), null allele probability and the mean number of alleles for each locus. Fst values between pairs of groups were calculated using Fstat 2.9 software (Goudet, 1995).

The genetic structure of grey partridge was investigated using a Bayesian clustering method. STRUCTURE 2.3 (Pritchard, Stephens and Donnelly, 2000) was used to infer the number of genetic clusters (K). The log-likelihoods of our data set (ln Pr(X|K)) were estimated for different numbers of genetic clusters using an admixture ancestry model based on 100,000 burn-in steps followed by 1,000,000 MCMC replicates. We utilized a method developed by (Evanno, Regnaut and Goudet, 2005) to determine the number of populations present, based on the second-order rate of change in the log probability of the data (ΔK) among 20 runs of each assumed K using the web-based utility "Harvest" (http://taylor0.biology.ucla.edu/struct_harvest).

Molecular sex identification

For the molecular sex identification in the grey partridge, conserved primers were used, namely the 1237L - 1272H pair (Kahn, John & Quinn, 1998) which delineates an intron that varies in size between the two sex chromosomes in most bird species. The total volume of the polymerase chain reaction was 10 µl in which 100 ng of genomic DNA was amplified, using 0.5 unit of Qiagen Taq polymerase, 0.2 mM dNTPs, 1 pmol of each primer, 1.5 mM $MgCl_2$ and 1 µl of 10 X Reaction Buffer. Thermal cycling amplification conditions were as follows: initial denaturation at 95 °C for 5 min, followed by 34 cycles of strand denaturation at 94 °C for 45 sec, annealing at 56 °C for 45 sec and primer extension at 72°C for 1 min and a final 7 min elongation time at 72°C.

ddRAD Library preparation

DNA quality was assessed via agarose gel electrophoresis on a 1% gel, and only nondegraded DNA (as judged by a tight high molecular weight band against a lambda standard) was selected for the library preparation stage. DNA was quantified using a Qubit Broad Range dsDNA Assay (ThermoFisher Scientific) according to the manufacturer's instructions and normalized to c. 7 ng/µl.

A total of 2 ddRAD libraries were constructed according to a modified protocol of the original (Peterson *et al.* , 2012) methodology. This is described in detail elsewhere (Brown *et al.* , 2016; Manousaki *et al.* , 2016). An inter and intra-library replicate sample was included during the construction of each library. Briefly, for each library individual genomic DNA (21 ng per sample) was restriction digested by *Sbf* I and *Sph* I, and then Illumina-specific sequencing adaptors (P1 & P2), each with unique combinatorial inline barcodes, were ligated to fragment ends. The samples were then pooled and size selected (400–700 bp fragments) by gel electrophoresis, PCR amplified (15 cycles) and the resultant amplicons (ddRAD library) were purified and quantified. The combinatorial inline barcodes (five or seven bases long) included in the P1 and P2 adaptors allowed each sample replicate to be demultiplexed post-sequencing. Each ddRAD library was sequenced on the Illumina MiSeq Platform (a single paired-end run; v2 chemistry, 2×150 bases).

ddRAD SNP calling

The sequences were quality assessed using FastQC (Andrews, 2010) and the reads were demultiplexed by barcode using the "*process_radtags* " module (default parameters) of the Stacks v1.48 bioinformatics pipeline (Catchen *et al.* , 2013). This module also filtered out low-quality reads. The retained reads, now missing

variable length barcodes, were then trimmed to a standard 148 bases in length. For each individual, matching forward and reverse reads were then concatenated into a single longer "artificial" read using a custom Perl script (Genbank BioProject Accession number: PRJNA1173582). Individuals that had <100,000 reads were removed from further processing.

The individual data were then processed using the "*denovo_map.pl* " module of stacks (-m 10 -M 2 -n 0) to assemble and create a catalogue of genetic loci contained in the data. The Stacks scripts "*export_sql.pl* " and "*populations* " and filtering steps were used to retain all loci that fulfilled the following criteria: a) contained exactly one SNP (in the concatenated forward and reverse reads) to avoid physically linked markers and ensure conserved sequence surrounding the target SNP to facilitate primer design; b) contained exactly two alleles c) had a read depth of [?]10 reads per individual to maximize the likelihood of the SNP being real.

ddRAD statistical analysis

A genotype file was created containing the SNPs genotypes, called from the ddRAD data, for 87 individuals and 8,236 SNPs. All the SNPs were checked and were biallelic, and we proceeded with the filtering of the dataset. The PLINK toolbox (Purcell *et al.* , 2007) was used to remove loci that were not genotyped in [?]80% of the individuals. Additionally, a dataset containing three full-sib families (parents and two offsprings) was used to check the mendelian inheritance of the alleles.

The genetic diversity indices as well as a Principal Component Analysis (PCA), based on Nei's genetic distance (Nei, Tajima and Tateno, 1983) were calculated using GENEALEX 6.502 software (Peakall and Smouse, 2006). Population pairwise FST was calculated using Arlequin 3.5 (Excoffier and Lischer, 2010). The genetic structure of grey partridge was investigated using STRUCTURE 2.3 (Pritchard, Stephens and Donnelly, 2000). The log likelihoods of our data set [ln Pr(X|K)] were estimated for different numbers of genetic clusters using an admixture ancestry model based on 100,000 burn-in steps followed by 1,000,000 Markov chain Monte Carlo (MCMC) replicates. We utilized a method developed by (Evanno, Regnaut and Goudet, 2005) to determine the number of populations present, based on the second-order rate of change in the log probability of the data (ΔK) among 20 runs of each assumed K using the web-based utility "Harvest" (http://taylor0.biology.ucla.edu/struct_harvest). We then evaluated the membership coefficient value (q) for each individual.

Analysis for the identification of an informative SNP marker panel for allowing population assignment and identification of non-native stocks for captive rearing was performed based on the "Informativeness of assignment" method, as implemented in TRES software (Kavakiotis *et al.* , 2015). Furthermore, assignment tests were performed according to Rannala and Mountain method (Rannala and Mountain, 1997) in GENECLASS2 software (Piry *et al.* , 2004).

Results

MtDNA analysis

Aligned sequences of a 381 bp section of the control region (CR) (n=224 individuals) and a 370 bp section of the cytochrome B gene (*cytb*) (n=228 individuals) were obtained. The CR was more polymorphic revealing 21 haplotypes compared to cytb, where only five haplotypes were detected. A total of 204 individuals have been sequenced at both regions and the mitochondrial diversity estimates of the combined sequences (751 bp in total) were summarized in Table 1. A total of 24 haplotypes (h) were observed, with 11 found within the wild UK birds and six within the wild Greek birds. Our dataset revealed that there was greater mitochondrial diversity within the wild UK birds (haplotype diversity = 0.776) than within the wild Greek birds (haplotype diversity = 0.439). This pattern was also mirrored within the captive stocks of both countries, with the UK captive stocks (hd=0.786) having a higher diversity than the Greek captive stocks (hd=0.409). Surprisingly, the wild-caught birds from North Macedonia had the highest haplotype diversity (hd=0.833). The three sampling locations with the highest haplotype diversity were all from different countries, East Lothian (hd=0.868; Scotland), Kumanovo (hd=0.822; North Macedonia) and Kozani (hd=0.757; Greece). This is contrasted with the two lowest estimates, also from different countries, Fife (hd=0.182; Scotland) and Drama

(hd=0; Greece).

The haplotype networks revealed and confirmed the eastern and western lineages in both the cytb and CR sequences (Figures 2 and 3). The inclusion of the haplotypes by (Liukkonen-Anttila *et al.* , 2002) in our network of the control region sequences, verified the structuring of our samples into the Western and Eastern lineages since all wild birds in the UK had a western mitochondrial haplotype while all wild birds in Greece showed an Eastern haplotype. The two populations from North Macedonia had a mix of birds as indicated by the networks, based on the CR data, most of the birds had an eastern haplotype (69%, n=11) and some possessed a western haplotype (31%, n=5).

The analysis also revealed that 17 new haplotypes (eight western and nine Eastern) exist within our control-region dataset that were not observed by the previous study. The most striking observation is the clustering of all Greek captive individuals within the Western lineage, showing higher relatedness with the wild birds from the UK rather than those in Greece.

Microsatellite analysis

In total, 106 grey partridge samples from the UK and 118 samples from Greece/North Macedonia were used for microsatellite analysis at eight loci. All samples that were successfully genotyped in more than six loci were included in the global analysis. Thus, a dataset was created with 224 individuals, and they were split into five different groups for further analysis: "UK Wild" including all individuals from Scotland and England, "UK Captive", "Greek/North Macedonia Wild", "Greek Captive" referring to Greek farmed individuals and "Hybrids" first generation offspring of Greek captive individuals paired with wild individuals. Eighty alleles were found across the eight microsatellite loci. The most variable loci were MNT412 and MNT12 while the least polymorphic were MNT404, MNT45 and Aru1A (Table S4_Appendix).

Genetic diversity indices were estimated for the five different groups of samples (Table 2). Heterozygosity values were moderate for all groups ranging from 0.438 to 0.631 with captive stocks as expected showing higher inbreeding coefficient levels due to higher levels of inbreeding. None of the populations met the criteria for HWE at a significance level of P[?]0.05. Further analysis revealed that most loci were out of HWE probably due to null alleles (frequencies 15-20%). Pairwise F_{ST} values range from 0.069 (UK Captive-Greek Captive) to 0.2 (Hybrids-UK Wild).

Population genetic structure was investigated on the entire global dataset and then separately for the wild birds in the UK, and Greece and North Macedonia. In the global analysis, there was a clear separation into two clusters, geographically into Western and Eastern lineages (Figure 4). The wild birds of North Macedonia and Greece clearly cluster separately from the wild birds of the UK. However, Greek captive stocks are more closely aligned with the -UK-Wild birds than any of the wild Greek birds. The microsatellites also reveal fine-scale information compared to mitochondrial data, including the large number of individuals within the captive UK stock as well as Greek stock that appear to be descended from birds of both Western and Eastern origin, but with a larger proportion of Western origins. Indeed, even in the wild birds of both regions there is evidence of hybrid individuals, one in the UK and approximate five in Greece/North Macedonia.

No genetic structure was found within the wild UK birds (Figure S1_Appendix). Regarding the population structure of wild birds in Greece and North Macedonia, three main geographical clusters were revealed by the analysis (Figure S2_Appendix). One genetic cluster consisted of the birds from Drama in north-eastern Greece, the second genetic cluster included the birds within North Macedonia and the third cluster including mostly birds from Grevena. Birds in Grevena and the three remaining populations (Kozani, Thessaloniki and Xanthi) exhibited a mix of the three genetic clusters, with the most south-westerly population (Grevena) showing slight distinction, i.e. exhibiting less genetic similarity to the birds from Drama than the other populations.

not-yet-known not-yet-known

not-yet-known

unknown

Discussion

Phylogeography of grey partridge

The phylogeographic data produced by (Liukkonen-Anttila *et al.* , 2002) were used as a basis for comparisons with our dataset. We sequenced an overlapping section of the mitochondrial control region that (Liukkonen-Anttila *et al.* , 2002) sequenced in individuals sampled across 17 Eurasian countries. One of their main findings was that there was an Eastern/Western split in mitochondrial haplotypes within Europe, suggesting that grey partridge had colonized Europe from two separate glacial refugia. However, only a few samples had been collected in Greece and none within Scotland. In our analysis, we have included a large number of data from these two countries as well as North Macedonia and we found that all wild birds in the UK had a western mitochondrial haplotype. Also consistent with (Liukkonen-Anttila*et al.* , 2002), all studied wild birds in Greece belong to Eastern lineage. The fact that the two populations from North Macedonia had a mix of birds, matches the situation in neighboring Bulgaria, perhaps suggesting that these regions of the Balkans are contact zones between the eastern and western lineages (Liukkonen-Anttila *et al.* , 2002). Alternatively, birds of Western origin may have been released for hunting. However, more research focusing on the Balkans is required to distinguish between these two scenarios.

By comparing the mitochondrial haplotypes to those previously published, we can also gain some limited insight into the origins of the lineages. Most of the wild birds in the UK had the main western haplotype (W1). This occurs throughout Europe in wild and captive stocks (Liukkonen-Anttila *et al.* , 2002; Andersen and Kahlert, 2012). However, the two other main haplotypes in our UK samples (W17 and W4) had also been previously identified by (Liukkonen-Anttila *et al.* , 2002) (Figure 3). The haplotype W17 was found in this study in captive birds from Greece and both captive and wild from the UK. A Danish study concluded that this haplotype originates from Bohemia (Czechia) and is found in European captive stocks and Danish populations that were stocked with Bohemian birds between 1895 and 1934 (Andersen and Kahlert, 2012). The captive origin of this haplotype is consistent with our data. The third haplotype that is most common in the UK samples is W4, this haplotype is highly likely to be of French origin based on the study by (Liukkonen-Anttila *et al.* , 2002). The other haplotypes revealed in our British samples are novel, however, the majority are very similar to the bohemian haplotype (W17). Neither of the two haplotypes that were unique to British wild birds in the previous study was found among our samples. As far as the origin of lineages in Greece, the main eastern haplotype (E1) that has been previously described in Finland, Bulgaria and Ireland (Liukkonen-Anttila *et al.* , 2002), was the most frequent and widespread haplotype in Greece and it was probably expanded to the area from the Balkan refugia. Five out of seven haplotypes detected in Greece were novel with haplotype E22 and E20 being the second and third most abundant, respectively. North Macedonia consisted of the two most frequent haplotype in Europe (E1 and W1), haplotype E16 with a probable Finish origin and four new haplotypes.

The mitochondrial data is, therefore, of limited use in population assignment; we can accurately use it to identify the presence of Eastern and Western lineages but without a much larger database of historical samples, it is currently difficult to identify the geographic origins of each haplotype, as the majority of British wild birds had haplotypes that are found in captive stocks.

The power of different types of markers to detect population structure

When comparing the relative performance of microsatellites and RADseq there was the expected discrepancy of higher heterozygosity values in the microsatellite compared to the RADseq data, which reflect the high mutational rate of microsatellites. Overall, both markers were able to uncover genetic structuring. Based on the small but widely geographically spread, number of sampling sites in the UK, both microsatellite and SNP data agree that there is a lack of deep phylogeographic structure within the wild British samples. Considering the larger distance between sampling sites compared to Greek sampling sites, for example, more than 500 km between Fife and Essex, this was unexpected. This result is suggestive of a large, homogenous population within the UK, but considering the low dispersal of grey partridge, this points towards large-scale

human-mediated translocations. However, sampling from a much larger number of sites with known histories of management and partridge releasing are required to confirm this finding more generally. On the contrary, both molecular markers suggested similar trends concerning phylogeographic structure in Greek populations, with SNP data being somewhat more conclusive, although fewer individuals were included in the analysis. The most distinct was the population of Drama, followed by the population of Grevena while the other three populations (Kozani, Thessaloniki and Xanthi) revealed more mixed patterns and greater similarity with North Macedonia.

To be able to confidently assign individuals to either a Western or Eastern origin, a marker panel of 15 SNPs has been identified, based on the ddRAD libraries. Only wild individuals were used for this panel development, however not a single SNP was at fixation in both populations, i.e. only occurred in the Eastern individuals, while completely absent from the Western individuals (see Table S5_Appendix for a summary of the SNP frequencies). These may reflect either the recent divergence of the two lineages from the common ancestor or could be the result of past introduction events that affected the allele frequency. However, this panel of 15 SNPs could be combined with a mitochondrial DNA test for the identification of non-native stocks within Greece or within the UK.

Estimate effect of introgressive hybridization on wild populations

All three datasets (mitochondrial, microsatellites and SNPs) show that captive stocks in both the UK and Greece have been predominantly derived from birds of western origin (*Perdix perdix perdix*). However, there is also evidence of Eastern genotypes within both these captive populations, signalling that birds of the eastern lineage (*Perdix perdix lucida*) have also been used for captive breeding. However, considering that in both the UK and Greece, individuals from captive stocks are released each year, very little evidence of hybrids within the wild individuals has been found; one hybrid individual was identified in the UK and five in Greece/North Macedonia based on microsatellites (Figure 4). Although they provide insight, the use of just eight microsatellites has limitations. As can be seen with the known hybrid individuals, only three out of the six can be identified as hybrids using this microsatellite data. However, even when considering the robust SNP data, there is no evidence of F1 hybrids in the wild (Figure 7), although the microsatellite data indicates that there may be one wild F1 hybrid in the UK dataset (Figure 4).

Obviously, Greece would be predicted to have a larger number of hybrids, considering the captive stocks being released are of Western origin, however, there is very little evidence to suggest that hybridization has occurred within the wild Greek birds. This contrasts with neighboring North Macedonia where all our genetic datasets (mitochondrial, microsatellite and SNPs) indicate that individuals of both eastern and western genealogies are present in the wild population. However, as already discussed, whether this is a natural transition zone between*P. p. perdix* and *P. p. lucida* or the result of captive releases requires further investigation. The differences between levels of hybridization in the captive stocks compared to the wild individuals in both the UK and Greece are intriguing. Several reasons could explain this: 1) Released birds do not survive and therefore do not contribute to the wild gene pool, 2) Released hybrid individuals do not survive and contribute to the wild gene pool, 3) Hybridization within the captive stocks is a relatively recent phenomenon and not many hybrids have been released yet, 4) The numbers of released birds have been very small compared to the wild population. We argue that the last point is unlikely, as the grey partridge has been the most important European game bird for centuries and there is historical evidence that stocking and translocations have been conducted throughout Europe as far back as the 1560s (Andersen and Kahlert, 2012). The other common name of the grey partridge (the Hungarian Partridge) is based on the widespread exporting of captive stock from Hungary across Europe and to North America during the 20th century (Browne, Buner and Aebischer, 2009). Estonian and French birds were more recently imported into Ireland for captive breeding between 2002-2005 (Buckley *et al.* , 2012). This widespread sharing of game bird stock across Europe is of routine and historical occurrence. However, the geographical origins of captive stock are rarely known. Regarding the third point, it does seem that captive populations are mainly of Western origin. This has been the finding of our study in both the UK and Greece but also of the two previous genetic studies in Finland and Denmark (Liukkonen-Anttila*et al.* , 2002; Andersen and

Kahlert, 2012) It is, therefore, feasible that hybridization is of limited success or has not occurred until recently. However, regarding the two first points, multiple studies have found that released birds have very low survival (Brittas*et al.* , 1992; Mallord *et al.* , 2024). This is the most likely explanation for Greece, where releases of captive birds have occurred since the 1980s, but without evidence of western genotypes in the wild populations yet. However, contrasting with Greece, in the UK we failed to find evidence of genetic structure between wild populations within the UK. These wild populations also share genetic similarities to the captive population based on mitochondrial (Figure 3) and microsatellite data (Figure S3_Appendix). Considering the low dispersal of Grey Partridge, this is indicative of widespread historic translocation of birds and/or the survival of captive releases.

not-yet-known not-yet-known

not-yet-known

unknown

Literature cited

Aebischer, N. J. and Ewald, J. A. (2012) 'The grey partridge in the uk: Population status, research, policy and prospects', *Animal Biodiversity and Conservation* .Andersen, L. W. and Kahlert, J. (2012) 'Genetic indications of translocated and stocked grey partridges (Perdix perdix): Does the indigenous Danish grey partridge still exist?',*Biological Journal of the Linnean Society* . doi: 10.1111/j.1095-8312.2011.01833.x.Andrews, S. (2010) 'FastQC - A quality control tool for high throughput sequence data.', *Available at http://www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc/* . doi: citeulike-article-id:11583827.Bandelt, H. J., Forster, P. and Röhl, A. (1999) 'Median-joining networks for inferring intraspecific phylogenies', *Molecular Biology and Evolution* . doi: 10.1093/oxfordjournals.molbev.a026036.Bao, X. kang *et al.* (2010) 'The phylogenetic position and speciation dynamics of the genus Perdix (Phasianidae, Galliformes)', *Molecular Phylogenetics and Evolution* . doi: 10.1016/j.ympev.2010.03.038.Bech, N. *et al.* (2010) 'Transferability of microsatellite markers among economically and ecologically important galliform birds', *Genetics and Molecular Research* , 9(2), pp. 1121–1129. doi: 10.4238/vol9-2gmr760.Brittas, R.*et al.* (1992) 'Survival and Breeding Success of Reared and Wild Ring-Necked Pheasants in Sweden', *The Journal of Wildlife Management* , 56(2), p. 368. doi: 10.2307/3808836.Brown, J. K. *et al.* (2016) 'Mapping the sex determination locus in the hāpuku (Polyprion oxygeneios) using ddRAD sequencing', *BMC Genomics* , 17(1), p. 448. doi: 10.1186/s12864-016-2773-4.Browne, S. J., Buner, F. and Aebischer, N. J. (2009) 'A review of gray Partridge Restocking in the UK and its Implications for the UK Biodiversity Action Plan', in*Gamebird 2006* .Buckley, K. *et al.* (2012) 'Every partridge counts, successful techniques used in the captive conservation breeding programme for wild grey partridge in Ireland', *Animal Biodiversity and Conservation* , 35(2), pp. 387–393. doi: 10.32800/abc.2012.35.0387.Catchen, J. *et al.* (2013) 'Stacks: an analysis tool set for population genomics', *Molecular Ecology* , 22(11), pp. 3124–3140. doi: 10.1111/mec.12354.Clustal, W. (1994) 'improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice Thompson, Julie D.; Higgins, Desmond G.; Gibson, Toby J',*Nucleic Acids Research* .Crowe, T. M. *et al.* (2006) 'Phylogenetics, biogeography and classification of, and character evolution in, gamebirds (Aves: Galliformes): effects of character exclusion, data partitioning and missing data', *Cladistics* . doi: 10.1111/j.1096-0031.2006.00120.x.Evanno, G., Regnaut, S. and Goudet, J. (2005) 'Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study', *Molecular Ecology* , 14(8), pp. 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x.Ewald, J. A., Sotherton, N. W. and Aebischer, N. J. (2020) 'Research Into Practice: Gray Partridge (Perdix perdix) Restoration in Southern England', *Frontiers in Ecology and Evolution* , 8. doi: 10.3389/fevo.2020.517500.Excoffier, L. and 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(3), pp. 564–567. doi: 10.1111/j.1755-0998.2010.02847.x.Ferrero, M. E. *et al.* (2007) 'Sixteen new polymorphic microsatellite markers isolated for red-legged partridge (Alectoris rufa) and related species', *Molecular Ecology Notes* , 7(6), pp. 1349–1351. doi: 10.1111/j.1471-8286.2007.01877.x.Goudet, J. (1995) 'FSTAT (Version 1.2): A Computer Program to Calculate F-statistics', *Journal of Heredity* , 86(6), pp. 485–486. doi: 10.1093/ox-

fordjournals.jhered.a111627.St. John, J. and Quinn, T. W. (1998) 'Chromosome-Specific Intron Size Differences in the Avian CHD Gene Provide an Efficient Method for Sex Identification in Birds',*The Auk* . doi: 10.2307/4089527.Kalinowski, S. T., Taper, M. L. and Marshall, T. C. (2007) 'Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment', *Molecular Ecology* , 16(5), pp. 1099–1106. doi: 10.1111/j.1365-294X.2007.03089.x.Kavakiotis, I. *et al.* (2015) 'TRES: Identification of Discriminatory and Informative SNPs from Population Genomic Data', *Journal of Heredity* , 106(5), pp. 672–676. doi: 10.1093/jhered/esv044.Kimball, R. T. and Braun, E. L. (2014) 'Does more sequence data improve estimates of galliformphylogeny? Analyses of a rapid radiation using a complete data matrix',*PeerJ* . doi: 10.7717/peerj.361.Kuijper, D. P. J., Oosterveld, E. and Wymenga, E. (2009) 'Decline and potential recovery of the European grey partridge (Perdix perdix) population—a review', *European Journal of Wildlife Research* , 55(5), pp. 455–463. doi: 10.1007/s10344-009-0311-2.Librado, P. and Rozas, J. (2009) 'DnaSP v5: A software for comprehensive analysis of DNA polymorphism data',*Bioinformatics* . doi: 10.1093/bioinformatics/btp187.Liukkonen-Anttila, T. *et al.* (2002) 'Mitochondrial DNA variation and the phylogeography of the grey partridge (Perdix perdix) in Europe: From pleistocene history to present day populations', *Journal of Evolutionary Biology* . doi: 10.1046/j.1420-9101.2002.00460.x.Liukkonen, T. (2006) 'Finnish native grey partridge (Perdix perdix) population differs clearly in mitochondrial DNA from the farm stock used for releases', *Annales Zoologici Fennici* .Liukkonen, T., Kvist, L. and Mykrä, S. (2012) 'Microsatellite markers show distinctiveness of released and wild grey partridges in Finland', *Animal Biodiversity and Conservation* .Mallord, J. W. *et al.* (2024) 'Survival rates of wild and released White-rumped Vultures (Gyps bengalensis), and their implications for conservation of vultures in Nepal', *Ibis* , 166(3), pp. 971–985. doi: 10.1111/ibi.13303.Manousaki, T. *et al.* (2016) ' Exploring a Nonmodel Teleost Genome Through RAD Sequencing—Linkage Mapping in Common Pandora, Pagellus erythrinus and Comparative Genomic Analysis ', *G3: Genes|Genomes|Genetics* , 6(3), pp. 509–519. doi: 10.1534/g3.115.023432.Nei, M., Tajima, F. and Tateno, Y. (1983) 'Accuracy of estimated phylogenetic trees from molecular data',*Journal of Molecular Evolution* , 19(2), pp. 153–170. doi: 10.1007/BF02300753.Peakall, R. and Smouse, P. E. (2006) 'GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research', *Molecular Ecology Notes* , 6(1), pp. 288–295. doi: 10.1111/j.1471-8286.2005.01155.x.Peterson, B. K. *et al.* (2012) 'Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species', *PLoS ONE* , 7(5). doi: 10.1371/journal.pone.0037135.Piry, S. *et al.* (2004) 'GENECLASS2: a software for genetic assignment and first-generation migrant detection.', *The Journal of heredity* , 95(6), pp. 536–9. doi: 10.1093/jhered/esh074.Pritchard, J. K., Stephens, M. and Donnelly, P. (2000) 'Inference of Population Structure Using Multilocus Genotype Data', *Genetics* , 155(2), pp. 945–959. doi: 10.1093/genetics/155.2.945.Purcell, S. *et al.* (2007) 'PLINK: a tool set for whole-genome association and population-based linkage analyses.', *American journal of human genetics* , 81(3), pp. 559–75. doi: 10.1086/519795.Randi, E. (2008) 'Detecting hybridization between wild species and their domesticated relatives', *Molecular Ecology* . doi: 10.1111/j.1365-294X.2007.03417.x.Rannala, B. and Mountain, J. L. (1997) 'Detecting immigration by using multilocus genotypes', *Proceedings of the National Academy of Sciences of the United States of America* . doi: 10.1073/pnas.94.17.9197.Rousset, F. (2008) 'GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux', *Molecular Ecology Resources* , 8(1), pp. 103–106. doi: 10.1111/j.1471-8286.2007.01931.x.Thomaides, C. & Papageorgiou, N. (1992) 'Nesting biology and habitat use of the Grey Partridge (Perdix perdix) in Northern Greece', in Birkan, M., Potts, G.R., Aebischer, N.J. & Dowell, S. D. (ed.) *Gibier Faune Sauvage* , pp. 443–446.

Table 1. Summary of mitochondrial diversity estimates based on both gene regions (751bp). Estimates have only been calculated for locations where [?] 10 individuals were sampled.

Region	Locations	No. individuals	H	Hd	Hd(SD)	π	S
UK_wild	ALL(n=5)	42	11	0.776	0.053	0.00168	10
	Fife	11	2	0.182	0.144	0.00024	1
	East Lothian	26	10	0.868	0.036	0.00238	9
Greece_wild	ALL(n=5)	73	6	0.439	0.069	0.00107	8
	Drama	30	1	0	0	0	0

Region	Locations	No. individuals	H	Hd	Hd(SD)	π	S
	Grevena	13	3	0.603	0.088	0.00113	3
	Kozani	17	5	0.757	0.063	0.00239	7
North Macedonia_wild	ALL(n=2)	13	6	0.833	0.071	0.0095	21
	Kumanovo	10	5	0.822	0.097	0.00936	21
Greece_captive	ALL(n=2)	29	3	0.409	0.102	0.00076	2
UK_captive	ALL(n=5)	44	6	0.786	0.026	0.00161	5
ALL	20+	201*	24	0.838	0.014	0.0118	37

Η: Νο. οφ ηαπλοτψπες, Ηδ: Ηαπλοτψπε διερσιτψ, Ηδ(ΣΔ): Ηδ Στανδαρδ δειατιον, π: νυχλεοτιδε διερσιτψ, Σ: Νο. αριαβλε σιτες

* three Finnish samples analysed in both genes were not included in the analysis

Table 2. *Genetic diversity indices per group of samples analysed for 8 microsatellite markers (first column) and 2300 SNPs (second column).*

Sample origin	N	A_r	H_e	H_O	F_{is}	P_{HW}
UK Wild	57/16	3.38/1.20	0.59/0.13	0.55/0.12	0.073/0.035	0.0621/0.9
UK Captive	49/3	3.64/1.15	0.63/0.09	0.55/0.12	0.130/-0.319	0/0
Greek/North Macedonia Wild	95/43	3.69/1.22	0.57/0.140	0.46/0.12	0.086/0.054	0/0.9
Greek Captive	17/14	2.90/1.21	0.55/0.128	0.51/0.12	0.188/-0.016	0.0003/0.9
Hybrids	6/5	3.07/1.19	0.47/0.122	0.42/0.16	0.04/-0.175	0.7491/0.005

Ν: νυμβερ οφ αναλψσεδ ινδιυδυαλς, Αρ: μεαν αλλελις ριχηνεσς, ΦΙΣ: ινβρεεδινγ χοεφφιχιεντ, Ηε: εξπεχτεδ ηετεροζψγοσιτψ, Ηο: οβσερεδ ηετεροζψγοσιτψ, ΠΗ

Figure 1. Sampling locations of wild birds in the UK, Greece and North Macedonia.

Figure 2. Haplotype network of the Cyt-B sequences including three *Perdix perdix* Genbank sequences. Each circle represents a distinct haplotype or maternal lineage and the markers on the lines between them represent the number of base-pair differences between each lineage.

Figure 3. Haplotype network of the CR sequences including the Eastern (E) and Western (W) haplotypes of Liukkonen *et al.,* 2002. Each circle represents a distinct haplotype or maternal lineage and the markers on the lines between them represent the number of base pairs between each lineage. The new haplotypes identified in this study are numbered in red.

Figure 4. A genetic Structure plot of the global microsatellite database, consisting of 8 microsatellites. Each bar represents an individual bird and the proportion of green colour represents the proportion of "western" genetic ancestry, and the proportion of red represents the proportion of "eastern" genetic ancestry.

Figure 5 : Principal Component Analysis (PCA) of the 81 individuals based on 2300 SNPs.

Figure 6. Genetic Structure plot of the wild-caught Greek and North Macedonia birds based on 2300 SNPs. Each bar represents an individual bird and the proportion of each colour represents the proportion of each of five genetic ancestries. The locations refer to those shown in Figure 1.

Figure 7. A genetic structure plot focusing on the levels of hybridisation in the sampled Grey partridges using 2300 SNPs. Each bar represents an individual bird and the proportion of orange colour represents the proportion of "western" genetic ancestry, and the proportion of blue represents the proportion of "eastern" genetic ancestry.

Data Availability Statement

A total of 224 sequences of the control region (CR) (Genbank Accession numbers: PQ431575 - PQ431798) and 228 sequences of the cytochrome B gene (cytb) (Genbank Accession numbers: PQ295559 - PQ295786), was produced in this work and are was deposited in a public database.

Competing Interests Statement

The authors declare no conflicts of interest.

Acknowledgements

The A.G. Leventis Foundation funded this work, and it was a collaborative effort between the Royal Zoological Society of Scotland's (RZSS) WildGenes laboratory and the Aristotle University of Thessaloniki (AUTH), Greece. We would like to thank the Game and Wildlife Conservation Trust for the collection of samples. We would also like to thank Rosie Lennon and Keith Chalmers-Watson for the provision of samples in the UK. Dr Helen Senn and Dr Gill Murray-Dickson have been instrumental in driving this project forward and Jennifer Kaden has tirelessly worked alongside them. In Greece, samples were provided by the Greek Hunter's Association. We also extend thanks to the wider RZSS team including Jenny Kaden, Adam Naylor and Donna Brown.

APPENDIX

	Sample ID	CR haplotype	Cyt-B haplotype	No Msat loci	ddRAD – GenBank Accession number
1	PER001	W1	CBW3	8	
2	PER002	W1	CBW1	8	
3	PER003	W1	CBW1	8	
4	PER004		CBW1	8	
5	PER005	W1	CBW1	8	
6	PER006	W1	CBW1	8	SAMN44308948
7	PER007	W1		8	
8	PER008	W1		8	
9	PER009.1	W1	CBW1	8	
10	PER010.1	W1	CBW1	8	
11	PER011.1	W1	CBW1	8	
12	PER012.1	W1	CBW1	8	
13	PER013.1		CBW1	8	
14	PER014.1	W34	CBW1	8	
15	PER015	W31	CBW1	8	
16	PER016	W1	CBW1	8	
17	PER017	W36	CBW1		
18	PER018		CBW2	8	
19	PER019	W1	CBW1		
20	PER021	W1	CBW1	8	
21	PER022	W17	CBW1	8	
22	PER023		CBW1		
23	PER024	W17		8	
24	PER026	W1			
25	PER027.1			8	SAMN44308949
26	PER030.1	W1	CBW2	7	SAMN44308950
27	PER031.1	W1	CBW2	8	SAMN44308951
28	PER037.1	W1	CBW2		SAMN44308952
29	PER041.1	W1	CBW2		SAMN44308953
30	PER042.1	W1	CBW2	8	SAMN44308954
31	PER043.1	W1	CBW2	8	SAMN44308955
32	PER044.1				SAMN44308956

	Sample ID	CR haplotype	Cyt-B haplotype	No Msat loci	ddRAD – GenBank Accession number
33	PER049.1				SAMN44308957
34	PER050.1	W1	CBW2		
35	PER052.1	W1	CBW2	8	SAMN44308958
36	PER053.1	W1	CBW2	8	SAMN44308959
37	PER054.1	W1	CBW2		SAMN44308960
38	PER061.1				SAMN44308961
39	PER063.1	W1	CBW2	8	SAMN44308962
40	PER064.1	W1	CBW2	8	SAMN44308963
41	PER065.1	W1	CBW2		
42	PER073	W1	CBW2	8	
43	PER074	W1	CBW2	8	
44	PER075	W1	CBW2	8	
45	PER076.1	E1	CBE1	8	SAMN44308964
46	PER077.1	E1	CBE1	8	SAMN44308965
47	PER078.1	E1	CBE1	8	SAMN44308966
48	PER079.1	E1	CBE1	7	SAMN44308967
49	PER080		CBE1		
50	PER081		CBE1		
51	PER083	E3	CBE1		
52	PER085	E16	CBE1		
53	PER086		CBE1		
54	PER087		CBE1		
55	PER088		CBE1		
56	PER090		CBE1		
57	PER091		CBE1		
58	PER092		CBE1		
59	PER094		CBE1		
60	PER095	E17	CBE1		
61	PER096	E1	CBE1	8	
62	PER097	E1	CBE1	8	
63	PER098	E1	CBE1	8	
64	PER099	E1	CBE1	8	SAMN44308968
65	PER100	E1	CBE1	8	SAMN44308969
66	PER101	W33	CBW2	8	SAMN44308970
67	PER102	W38	CBW2	8	SAMN44308971
68	PER104		CBW1	8	SAMN44308972
69	PER105	W17	CBW1	8	SAMN44308973
70	PER106	W37	CBW1	8	
71	PER107	W1	CBW2	8	SAMN44308974
72	PER109.1			8	
73	PER110.1	E1	CBE1	8	
74	PER111			7	
75	PER112			7	
76	PER113.1	E20	CBE1	8	
77	PER114.1	E1	CBE1	8	SAMN44308975
78	PER115.1	E20	CBE1	8	SAMN44308976
79	PER116	E1	CBE1	8	
80	PER117.1	E1	CBE1	8	SAMN44308977
81	PER118	E22	CBE1	8	SAMN44308978
82	PER119	E22	CBE1	8	SAMN44308979

	Sample ID	CR haplotype	Cyt-B haplotype	No Msat loci	ddRAD – GenBank Accession number
83	PER120	E22		8	SAMN44308980
84	PER121	E1	CBE1	8	SAMN44308981
85	PER122	E22	CBE2	8	SAMN44308982
86	PER123	E1	CBE1	8	SAMN44308983
87	PER124	E23	CBE1	8	SAMN44308984
88	PER125	E24	CBE1	8	SAMN44308985
89	PER126	E22	CBE1	8	SAMN44308986
90	PER127.1	E22	CBE1	8	SAMN44308987
91	PER128.1	E22	CBE2	8	
92	PER129.1	E22	CBE1	8	
93	PER130	E1	CBE1	6	SAMN44308988
94	PER131	E22	CBE2	8	
95	PER132	E20	CBE1	8	SAMN44308989
96	PER133	E20	CBE1	8	SAMN44308990
97	PER134	E20	CBE1	8	SAMN44308991
98	PER135	E20	CBE1	8	SAMN44308992
99	PER136	E1	CBE1	8	SAMN44308993
100	PER137	W1	CBW1	8	SAMN44308994
101	PER138	W1	CBW1	8	SAMN44308995
102	PER139	W17	CBW1	8	SAMN44308996
103	PER140	W17	CBW1	8	
104	PER141	W1	CBW1	6	SAMN44308997
105	PER142	W17	CBW1	8	SAMN44308998
106	PER143	W1	CBW1	8	SAMN44308999
107	PER144	W1	CBW2	8	
108	PER145		CBW2	8	
109	PER146	W1	CBW1	8	
110	PER147	W1		8	SAMN44309000
111	PER148	W1	CBW2	8	
112	PER149	W1	CBW2	8	
113	PER150	W17	CBW1	8	
114	PER151	W1	CBW2	8	
115	PER152	W1	CBW1	8	
116	PER153		CBW1	8	
117	PER154	W1	CBW2	8	
118	PER155	W4	CBW1	8	
119	PER156	W1	CBW1	8	
120	PER157	W4	CBW1	8	
121	PER158	W17	CBW1	8	
122	PER159	W4		8	
123	PER160	W4	CBW1	8	
124	PER161		CBW2	8	
125	PER162	W17		8	
126	PER163	W4	CBW1	8	
127	PER164	W17	CBW1	8	
128	PER165	W4	CBW1	8	
129	PER166	W4	CBW1	8	
130	PER168.1	W1	CBW2	8	SAMN44309001
131	PER169.1	W1	CBW2	8	SAMN44309002
132	PER170.1	W1	CBW2	8	SAMN44309003

	Sample ID	CR haplotype	Cyt-B haplotype	No Msat loci	ddRAD – GenBank Accession number
133	PER171.1	W1	CBW2	6	
134	PER172.1	W1	CBW2	8	SAMN44309004
135	PER173.1	W1	CBW2	8	SAMN44309005
136	PER174			8	
137	PER175.1	W1	CBW2	8	
138	PER177.1	W1		8	
139	PER178.1	W1	CBW2	8	
140	PER179	W1	CBW2	8	
141	PER180		CBW2	8	
142	PER181	W1	CBW1	8	
143	PER182	W1	CBW2	8	
144	PER183	W1	CBW2	8	
145	PER184	W1	CBW2	8	
146	PER185	W1	CBW1	8	SAMN44309006
147	PER186	W1		8	SAMN44309007
148	PER187		CBW2	8	
149	PER188	W1	CBW1	8	SAMN44309008
150	PER189.1	E1	CBE1	8	
151	PER190.1	E1	CBE1	8	
152	PER191	E1	CBE1	8	
153	PER192	E1	CBE1	8	
154	PER193	E1	CBE1	8	SAMN44309009
155	PER194	E1	CBE1	8	SAMN44309010
156	PER195	E1	CBE1	8	SAMN44309011
157	PER196	E1	CBE1	8	
158	PER197	E1	CBE1	8	SAMN44309012
159	PER198	E1	CBE1	8	SAMN44309013
160	PER199	E1	CBE1	8	
161	PER200	E1	CBE1	8	SAMN44309014
162	PER201	E1	CBE1	8	SAMN44309015
163	PER202.1	E1	CBE1	8	
164	PER203.1	E1	CBE1	8	
165	PER204.1	E1	CBE1	7	
166	PER205	E1	CBE1	7	
167	PER206	E1	CBE1	7	
168	PER207	E1	CBE1	8	
169	PER208	E1	CBE1	8	
170	PER209.1	E1	CBE1	8	
171	PER210.1	E1	CBE1	8	
172	PER211	E1	CBE1	8	SAMN44309016
173	PER212	E1	CBE1	8	
174	PER213	E1	CBE1	8	
175	PER214	E1	CBE1	8	SAMN44309017
176	PER215	E18		8	SAMN44309018
177	PER216	E22	CBE1	8	
178	PER217	E22	CBE2	8	
179	PER218	W1	CBW1	8	SAMN44309019
180	PER219	W1	CBW1	8	SAMN44309020
181	PER220	W1	CBW1	8	SAMN44309021
182	PER221	W1		8	SAMN44309022

	Sample ID	CR haplotype	Cyt-B haplotype	No Msat loci	ddRAD – GenBank Accession number
183	PER222	W1		8	SAMN44309023
184	PER223	W1		8	SAMN44309024
185	PER224	W1		8	SAMN44309025
186	PER225	W1	CBW1	8	SAMN44309026
187	PER226	E20	CBE1	7	
188	PER227	E1	CBE1	8	SAMN44309027
189	PER228	E1	CBE1	8	
190	PER229	E1	CBE1	8	
191	PER230	E1	CBE1	8	
192	PER231	E1	CBE1	8	
193	PER232	E1	CBE1	7	
194	PER233	E1	CBE1	8	
195	PER234	E1	CBE1	8	
196	PER235	W1	CBW2	8	
197	PER236	W1	CBW1	7	
198	PER237	W1	CBW1	8	
199	PER238	W4	CBW1		
200	PER239		CBW1		
201	PER240		CBW1		
202	PER248	W4	CBW1	6	
203	PER249	W1			
204	PER250	W1	CBW1	8	
205	PER252	W1	CBW1	6	
206	PER253	W1	CBW2	6	
207	PER254	W1	CBW1	8	
208	PER255	W35	CBW1	8	
209	PER256	W31	CBW1	8	
210	PER257	W1	CBW1	8	
211	PER258	W35	CBW1	8	
212	PER259	W1	CBW1	8	
213	PER260	W17	CBW1	6	
214	PER261	W4	CBW1	8	
215	PER262	W17	CBW1	8	
216	PER263		CBW1		
217	PER264	W17	CBW1	8	
218	PER265		CBW1	6	
219	PER266	W17	CBW1	8	
220	PER267	W17	CBW1	8	
221	PER268	W1	CBW1		
222	PER269	W17	CBW1	8	
223	PER270	W17	CBW1	8	
224	PER271	W4	CBW1	8	
225	PER272	E1	CBE1	6	
226	PER273	E1	CBE1	6	SAMN44309028
227	PER274	E1		6	
228	PER275	E1	CBE1	6	
229	PER276	E21	CBE1	7	SAMN44309029
230	PER277	W1	CBW1	7	
231	PER278	E19	CBE1	8	
232	PER279	W1	CBW1	8	

	Sample ID	CR haplotype	Cyt-B haplotype	No Msat loci	ddRAD – GenBank Accession number
233	PER280	E16	CBE1	8	SAMN44309030
234	PER281	E16	CBE1	8	SAMN44309031
235	PER282	E1	CBE1	7	SAMN44309032
236	PER283	E16	CBE1	7	
237	PER284	E1	CBE1	7	
238	PER285	E16	CBE1	8	
239	PER286	W32	CBW1	8	SAMN44309033
240	PER287	E1	CBE1	7	
241	PER288	W1		7	
242	PER289	E18		7	
243	PER290	W1		6	
244	PER291	E1	CBE1	7	SAMN44309034
245	PER294	W1	CBW2	7	
246	PER295		CBW1	8	
247	PER296	W4	CBW1	8	
248	PER297	W1	CBW1	8	
249	PER298	W1	CBW2	8	
250	PER299	W4	CBW1	6	
251	PER300	W4	CBW1	8	
252	PER301	W4	CBW1	8	
253	PER302	W1	CBW2		
254	PER303	W4	CBW1	8	
255	PER304	W1	CBW1	8	
256	PER305	W1	CBW2	8	

		Primer Sequence	Fragment size	Final concentration	Refe
Panel A	**MNT412**	F: CCCATGTGAGCAGTGAATTG R: GTCATCACAGTGGAGGATCG	235-273	0.33 pmol/µl	(Bech
	Aru1G4	F : CTGCAGTCACACAAGGCTAC R: AGTGGGTCAAGGATGAGTGG	140-166	0.16 pmol/µl	(Ferr
	Aru1A1	F: GGAAGCCAGATGAACCAAGG R: ATGCATGCGTGGAGGCTGAG	203-227	0.16 pmol/µl	(Ferr
	MNT45	F: ACATGGAGGCAGAGAACCTC R: TGTCAGCCTGAATGTTTCCTC	107-113	0.33 pmol/µl	(Bech
Panel B	**MNT12**	F: AGGTGTTTTTGGGCAGTCTC R: TGCAAGCACCATCTGCTAAG	140-180	0.25 pmol/µl	(Bech
	MNT408	F: GTGTCCCTGCCACACTACAG R: GGGAATTTGCTCCAACTGAC	219-237	0.33 pmol/µl	(Bech
	MNT404	F: AACCAGCTCTGGAGATACCG R: GGACTGCAAGGACAACATCC	250-254	0.16 pmol/µl	(Bech
	MNT477	F: TTCACCACGCTCATTCAAAG R:TCCAAAATGTGACTAGATGATAAAGTG	229-263	0.25 pmol/µl	(Bech

Table S3: PCR amplification conditions of multiplex PCR reaction for microsatellite analysis

Steps	Temperature	Time	Cycles
Initial denaturation	95°C	15min	

Steps	**Temperature**	**Time**	**Cycles**
Denaturation	94°C	30sec	
Annealing	60°C-50°C	45sec	19
Extension	72°C	60sec	
Denaturation	94°C	30sec	
Annealing	50°C	45sec	20
Extension	72°C	60sec	
Final Extension	72°C	20min	

Table S4. Microsatellite loci analysed. Number of alleles (A), allelic size range in base pairs (R), expected and observed heterozygosity (He, Ho), probability value for Hardy-Weinberg tests (P_{HW}), polymorphic information content (PIC), null alleles per locus (F), sample size (N), mean number of alleles (Amean) and allelic richness (A_R) for all samples.

	MNT412	Aru1G4	Aru1A1	MNT45	MNT12	MNT404	MNT477	MNT408
A	18	14	3	4	27	3	12	10
R	235-273	140-166	203-227	107-113	140-180	250-254	229-263	227-237
He	0.893	0,704	0.394	0.474	0.886	0.338	0.701	0.721
HO	0,741	0,582	0.400	0.366	0.556	0.214	0.445	0.593
PHW	***	0.005	0,862	0.0062	***	***	***	0.005
PIC	0.882	0,666	0.359	0.3767	0.876	0.290	0.666	0.665
F	0,0942	0.0955	-0.0103	0.1283	0,233	0,219	0,228	0.099

Table S5. Allele and genotype frequencies for the top-scored 15 SNPs.

not-yet-known not-yet-known not-yet-known unknown

	Genotyped	1st Allele	2nd Allele	Genotypes		
SNP6107_124		freq: 2	freq: 4	22	44	24
East Wild	37	0.014	0.986	0	0.973	0.027
West_Wild	16	0.969	0.031	0.938	0	0.062
Total	53	0.302	0.698	0.283	0.679	0.038
SNP14657_117		freq: 2	freq: 4	44	22	42
East Wild	38	1	0	0	1	0
West_Wild	14	0.071	0.929	0.857	0	0.143
Total	52	0.75	0.25	0.231	0.731	0.038
SNP3226_123		freq: 2	freq: 4	44	22	42
East Wild	39	1	0	0	1	0
West_Wild	16	0.125	0.875	0.75	0	0.25
Total	55	0.745	0.255	0.218	0.709	0.073
SNP3650_6		freq: 2	freq: 4	44	22	42
East Wild	40	1	0	0	1	0
West_Wild	16	0.219	0.781	0.562	0	0.438
Total	56	0.777	0.223	0.161	0.714	0.125
SNP2343_45		freq: 1	freq: 3	33	11	31
East Wild	42	1	0	0	1	0
West_Wild	16	0.25	0.75	0.562	0.062	0.375
Total	58	0.793	0.207	0.155	0.741	0.103

	Genotyped	1st Allele	2nd Allele	Genotypes		
SNP501_30		freq: 2	freq: 3	22	33	23
East Wild	40	0.013	0.988	0	0.975	0.025
West_Wild	16	0.781	0.219	0.625	0.062	0.312
Total	56	0.232	0.768	0.179	0.714	0.107
SNP1284_111		freq: 2	freq: 3	33	22	32
East Wild	41	0.963	0.037	0	0.927	0.073
West_Wild	16	0.188	0.812	0.75	0.125	0.125
Total	57	0.746	0.254	0.211	0.702	0.088
SNP6462_91		freq: 2	freq: 4	44	22	24
East Wild	42	0.714	0.286	0.071	0.5	0.429
West_Wild	16	0	1	1	0	0
Total	58	0.517	0.483	0.328	0.362	0.31
SNP1204_116		freq: 3	freq: 4	44	33	43
East Wild	40	0.787	0.212	0.075	0.65	0.275
West_Wild	16	0.031	0.969	0.938	0	0.062
Total	56	0.571	0.429	0.321	0.464	0.214
SNP5462_40		freq: 1	freq: 2	11	22	21
East Wild	34	0.294	0.706	0.147	0.559	0.294
West_Wild	14	1	0	1	0	0
Total	48	0.5	0.5	0.396	0.396	0.208
SNP4231_43		freq: 1	freq: 3	11	33	31
East Wild	37	0.108	0.892	0.054	0.838	0.108
West_Wild	12	0.875	0.125	0.75	0	0.25
Total	49	0.296	0.704	0.224	0.633	0.143
SNP1401_14		freq: 1	freq: 3	11	33	31
East Wild	40	1	0	1	0	0
West_Wild	16	0.312	0.688	0.188	0.562	0.25
Total	56	0.804	0.196	0.768	0.161	0.071
SNP7369_110		freq: 1	freq: 3	33	11	13
East Wild	38	0	1	1	0	0
West_Wild	14	0.643	0.357	0.214	0.5	0.286
Total	52	0.173	0.827	0.788	0.135	0.077
SNP7300_126		freq: 1	freq: 3	11	33	13
East Wild	37	0	1	0	1	0
West_Wild	13	0.615	0.385	0.385	0.154	0.462
Total	50	0.16	0.84	0.1	0.78	0.12
SNP14163_31		freq: 3	freq: 4	33	44	34
East Wild	37	0	1	0	1	0
West_Wild	14	0.607	0.393	0.357	0.143	0.5
Total	51	0.167	0.833	0.098	0.765	0.137