Lack of genetic structure in greylag goose (*Anser anser*) populations along the European Atlantic flyway

Irene Pellegrino, Marco Cucco, Arne Follestad, Mathieu Boos

Greylag goose populations are steadily increasing in north-western Europe. Although individuals breeding in the Netherlands have been considered mainly sedentary birds, those from Scandinavia or northern Germany fly towards their winter quarters, namely over France as far as Spain. This study aimed to determine the genetic structure of these birds, and to evaluate how goose populations mix. We used mitochondrial DNA and microsatellites from individuals distributed throughout the European Atlantic flyway, from breeding sites in Norway and the Netherlands to stopover and wintering sites in northern and south-western France. The mtDNA marker (CR1 D-Loop, 288 bp sequence, 151 ind.) showed 24 different haplotypes. The genetic distances amongst individuals sampled in Norway, northern France and the Netherlands were low (range 0.012-0.013). Individuals in south-western France showed a slightly higher genetic distance compared to all other sampling areas (ranges 0.016-0.017). The NJ tree does not show evidence of any single clades grouping together all individuals from the same geographic area. Besides, individuals from each site are found in different branches. Bayesian clustering procedures on 14 microsatellites (169 individuals) did not detect any geographically distinct cluster, and a high genetic admixture was recorded in all studied areas except for the individuals from the breeding sites in Norway, which were genetically very close. Estimation of migration rates through Bayesian inference confirms the scenario for the current mixing of goose populations.

2 **Lack of genetic structure in greylag goose (*Anser anser*) populations along the**

3 **European Atlantic flyway**

5 ***Irene Pellegrino,*** *University of Piemonte Orientale, DISIT, Alessandria, Italy. Email:*

6 *irene.pellegrino@unipmn.it*

7 ***Marco Cucco***, *University of Piemonte Orientale, DISIT, Alessandria, Italy. Email: cucco@unipmn.it*

8 ***Arne Follestad,*** *Norsk Institutt for Naturfoskning Postboks 5685 Sluppen, 7485 Trondheim, Norway.*

9 *Email: arne.follestad@nina.no*

10 ***Mathieu Boos,*** *Naturaconst@, Research Agency in Applied Ecology, 67270 Wilshausen, France.*

11 *Email: mboos.naturaconst@free.fr*

13 **Corresponding author**

14 ***Marco Cucco***, *University of Piemonte Orientale, DISIT, viale Michel 11, 15121 Alessandria, Italy. Tel.*

15 *+30.0131360276, Email: cucco@unipmn.it*

Introduction

The greylag goose (*Anser anser*) is widespread throughout the Palearctic. In Europe, the main breeding populations are located in central and northern countries, and the species rarely breeds in Mediterranean areas (Cramp, 1977; Hagemeijer & Blair, 1997; BirdLife International, 2004). European populations show different patterns of movement. Although individuals breeding in Scotland and the Netherlands are considered sedentary birds (Delany & Scott, 2006), those from Scandinavia or central Europe fly longer distances, namely over France to Spain, with some individuals reaching North Africa (Fox et al., 2010; Nilsson et al., 2013). Icelandic breeders winter in Ireland and Britain, and greylags from Russia reach the regions bordering the eastern Mediterranean, Black and Caspian seas. Individuals with morphological characters ascribed to the oriental subspecies *rubirostris* have been observed on rare occasions in western Europe (Cramp, 1977).

Widescale movement patterns have been studied through the recapture or resighting of marked birds (coloured neck collars and leg rings, see Nilsson et al., 2013). These methods gave valuable information about the origin of birds that were found in moulting areas (Nilsson, Kahlert & Persson, 2001), flying, or staging in winter quarters. Birds from Sweden and Norway fly to Denmark and/or the Netherlands (SOVON, 1987; Persson, 1993; Andersson et al., 2001). Recent monitoring of greylags tagged with GPS devices in Norway show that approximately 30-50% can stay in the Netherlands during the whole wintering season, whereas others migrate to France or Spain. These geese all return to their previous breeding sites, thus showing a high breeding site fidelity (Boos et al., 2012, Boos unpublished data). According to Ramo et al. (2012), an increasing number of greylag geese winter at higher latitude. A noticeable effect of climatic changes probably explains this increasing tendency for geese to winter more closely to their breeding grounds.

The European Atlantic flyway covers a vast area stretching from northern France to Spain and Portugal, with arrivals from Scandinavia, Poland, Denmark and Germany (Fouquet et al., 2009). The situation in France is particularly complicated, because noticeable fluxes of geese coming from Northern

43   or Central Europe are found not only along the Atlantic flyway but also in other areas located in central

44   and south-eastern France. The departure areas of these birds have yet to be been fully determined, and the

45   timing of migration can probably differ depending on the origin of the populations (Fouquet, 1991;

46   Comolet-Tirman, 2009). Furthermore, the relative proportion of geese travelling to France and

47   originating from different countries may change over time (Pistorius, Follestad & Taylor, 2006; Pistorius

48   et al., 2007). However, data from neck-collared or ringed geese can be skewed by variations in the

49   marking and resighting efforts of the countries involved (Nilsson, 2007; Nilsson et al., 2013), and this

50   makes it difficult to fully define the composition of goose subpopulations migrating south from

51   observational data alone.

52   Genetics have become a useful tool in the study of migration and wintering patterns. Recent

53   studies on Anseriforms examined spatial structure along the flyways or in wintering zones, then

54   compared it to genetic data in breeding areas. In the king eider (*Somateria spectabilis*), strong site fidelity

55   to wintering areas and pair formation at wintering quarters indicated a population structure defined by

56   wintering rather than nest-site philopatry (Pearce et al., 2004). However, genetic analyses of mtDNA and

57   microsatellite alleles showed a lack of spatial genetic structure, suggesting the possible existence of flows

58   with homogenized gene frequencies. In the mallard (*Anas platyrhynchos*), single nucleotide markers were

59   used to investigate population structure on a continental scale throughout the northern hemisphere. This

60   genetic analysis found a general panmixia, suggesting that mallards form a single large, interbreeding

61   population (Kraus et al., 2013). The tufted duck (*Aythya fuligula*) shows high breeding site fidelity, but

62   mtDNA and microsatellite markers revealed an extensive population admixture on the wintering ground

63   (Liu et al., 2012, 2013). In the common pochard (*Aythya ferina*), genetic differentiation was observed

64   among Eurasian breeding populations, but no evidence of genetic structure was detected for pochards

65   sampled on European wintering grounds (Liu, Keller & Heckel, 2011).

66   Relatively few studies have investigated the genetic aspects of European geese of the genus

67   *Anser,* and the subject has not been thoroughly investigated at all in the greylag goose. Studies by

68   Ruokonen (2004, 2005) examined the genetic variability in two species of conservation concern, the

69   lesser white-fronted goose (*Anser erythropus*) and the pink-footed goose (*Anser brachyrhynchus*), and

70   investigated the phylogenetic relationship between seven *Anser* species (Ruokonen, Kvist & Lumme,

71   2000). A small amount of genetic differentiation between species has been observed in this genus

72   (Ruokonen, Kvist & Lumme, 2000; Johnsen et al., 2010).

73   Actually, mitochondrial DNA showed the presence of highly fragmented populations in two

74   species of conservation concern, the lesser white-fronted goose (*Anser erythropus,* Ruokonen et al., 2004)

75   and the pink-footed goose (*Anser brachyrhynchus,* Ruokonen, Aarvak & Madsen, 2005).

76   However, population genetics among species of geese has not been investigated to date. Here we used

77   both mitochondrial DNA and microsatellites to study the characteristics of greylag geese from two

78   breeding areas (the Netherlands and north-western Norway) and two wintering zones (northern France

79   and south-western France). This study investigates to what extent populations are genetically

80   differentiated. Genetic structure could have been increased by the fragmentation of breeding geese in

81   separated areas, or on the contrary, a limited genetic structure could have been developed by i) the

82   widespread practice of amateur breeding and selling of geese (Hagemeijer & Blair, 1997; Wang et al.,

83   2010), ii) the recent increase in the size of several populations (Klok et al., 2010), and iii) the habit of

84   European geese to rest several times during their flight toward their winter quarters (Fouquet, Schricke &

85   Fouque, 2009) at stopover sites where individuals from distant areas can admix and form pair bonds.

86   Knowledge of the genetic structure and diversity of greylag goose populations is a necessary

87   scientific basis to manage this emblematic species (Lorenz, 1966) and decide on appropriate action for its

88   conservation (Kampe-Persson, 2002) in the light of serious recent conflicts with agricultural and habitat

89   protection interests in most North European countries (Klok et al., 2010).

90

91   **Methods**

92

93   **Sample collection and DNA extraction**

94 We analyzed feather samples from 174 greylag geese (Appendix 1) collected over the European Atlantic

95 flyway (from two breeding grounds: in North and Western Norway, and six staging grounds: in the

96 Netherlands, northern France and south-western France; see Table 1 and Fig. 1). One additional

97 individual was collected in the Republic of Kalmykia in an area associated with the eastern *rubirostris*

98 subspecies (Cramp, 1977). During the 2010/2011 and 2011/2012 (including 1-10 February) wintering

99 seasons in France, goose feathers were obtained from greylag geese collected during the legal hunting

100 period in natural areas by hunters collaborating with the study. Samples from the Netherlands were

101 obtained on wild free-ranging geese collected in natural areas in the Zeeland region (near Rilland) by a

102 local hunter before the 20th of September in 2011 and 2013, i.e. before the arrival of geese breeding in

103 Norway or in Sweden (Nilsson, 2006, 2007; Boos pers. obs. based on GPS data). Samples from Norway

104 were obtained from birds that were collected during the spring and summer legal hunting seasons, or

105 from geese that were caught during the moulting period in 2010 and 2011 by A.F. for the Nordic Greylag

106 Goose Project, which studies the ecology of the Norwegian breeding goose population (Nilsson, 2007).

107 Feather calami were stored in ethanol at -20 °C, and total DNA was extracted using the commercial

108 NucleoSpin®Tissue kit (Macherey-Nagel, Düren, Germany). After extraction, genomic DNA was

109 stocked at -20 °C.

110

111 **Mitochondrial DNA sequencing**

112 Partial mitochondrial control region (CR1 D-Loop 288 bp) was amplified in 144 of the 174 individuals

113 (Appendix 1) using L180 (5'TGGTTATGCATATTCGTGCATAGA'3) and H466

114 (5'TTTCACGTGAGGAGTACGAC TAAT'3) primers (Ruokonen et al., 2000). PCR amplifications were

115 carried out in a Bio-Rad thermal cycler (Bio-Rad Laboratories Inc.; Hercules, California, USA). PCR

116 reaction was performed in a final volume of 25 μl containing 0.4 μl dNTPs (10 mM), 1 μl MgCl2 (25

117 mM), 0.3 μl of each primer (25 pmol/μl), 2.5 μl 10× buffer, 0.4 μl Taq polymerase (5 unit/μl; QIAGEN),

118 ddH2O and genomic DNA (20–100 ng/μl). The selected cycling profile included a 4 min preliminary

119 denaturation cycle at 94 °C followed by 32 denaturation, annealing and extension cycles (30 s at 94 °C,

120 30 s at 58 °C and 30 s at 72 °C, respectively) before a final extension of 7 min. Negative controls were

121 included for amplification procedures to detect contaminations.

122   The PCR product was purified using the EXO-SAP procedure with Exonuclease I (Exo;

123 Fermentas, Burlington, Canada) and Shrimp Alkaline Phosphatase (SAP; Fermentas, Burlington,

124 Canada). The purification cycle consisted of 30 min at 37 °C, then 15 min at 80 °C to deactivate the

125 enzymes followed by a 10 min cooling-down step at 4 °C. DNA concentration was determined after

126 electrophoresis in 1.8% agarose gels (TBE 1%) stained with ethidium bromide and visualized in a UV-

127 trans illuminator Gel Doc XR (Bio-Rad Laboratories Inc.; Hercules, California, USA) using the

128 Molecular Imager ChemiDoc XRS System and Quantity One software (Bio-Rad Laboratories Inc.;

129 Hercules, California, USA).

130   Sequencing was carried out at Macrogen Laboratories (Amsterdam, The Netherlands) in an ABI

131 3730xl Analyzer (Applied Biosystems).

132   Raw electropherograms were checked visually using FinchTV (Geospiza Inc.; Seattle, WA,

133 USA; http://www.geospiza.com), and sequences were aligned with ClustalW algorithm in BioEdit 7.05

134 (Hall, 1999). The haplotype network was calculated in Network 4.6 (Fluxus Technology Ltd; Clare,

135 Suffolk, England; fluxus-engineering.com) using the median joining procedure (MJ: Bandelt et al., 1999).

136 DnaSP version 5 (Librado & Rozas, 2009) was used to estimate mtDNA haplotype diversity (h),

137 nucleotide diversity (π) and the mean number of pairwise differences (k) in the sampled areas.

138 Demographic and/or spatial population expansion events were investigated using the mismatch

139 distribution implemented in DNASP v. 5. MEGA 5.0 (Tamura et al., 2011) was used to perform the

140 neighbour-joining method (NJ: Saitou and Nei, 1987), clustering pairwise Tamura-Nei's genetic

141 distances between haplotypes (TN93: Tamura and Nei, 1993). Support for the internodes in the NJ tree

142 was assessed by bootstrap percentages (BP: Felsenstein, 1988) after 1000 resampling steps. One sequence

143 of *Anser anser anser* (GenBank AF159962) from Finland and another of *Anser anser rubrirostris*

144 (GenBank AF159963) from Slimbridge Wetland Center, England, were included as reference sequences

145   in tree construction. A sequence of the lesser white-fronted goose (*Anser erythropus*, GenBank

146   AY072580) and the bean goose (*Anser fabalis*, GenBank AB551534) were used as outgroups.

147   Maximum likelihood (ML) and maximum parsimony (MP) trees were obtained through the

148   DNAML, CONSENSE, DNAPARS programmes in PHYLIP 3.67 (Felsenstein, 2005). Bootstrap values

149   were based on 1000 replicates, and the tree topologies were visualized with FigTree 1.3.1 (Rambaut,

150   2008). The best substitution model for molecular evolution was selected using the corrected Akaike

151   Information Criterion (AICc, Burnham and Anderson, 2004) in jModelTest (Posada, 2008). Maximum

152   likelihood bootstrap supports were estimated by performing 100 runs with 1000 bootstrap replicates.

153   The partition of mtDNA diversity within and among the sampled geographical populations were

154   investigated by running analyses of molecular variance (AMOVA, Excoffier et al., 1992) using Arlequin

155   3.3 (Excoffier & Lischer, 2010).

156

157   **Microsatellite genotyping**

158   A total of 169 of the 174 samples (Appendix 1) were genotyped by PCR amplification at 14

159   microsatellite loci (Ans02, Ans04, Ans07, Ans13, Ans17, Ans18, Ans21, Ans24, Ans25, Aalμ1b,

160   Aph12b, Aph19b, Smo7b, Hhiμ1b) that had previously been isolated and tested in *Anser anser* (Weiß et

161   al., 2008). We used PCR reactions, thermal profiles, fluorescent dye and multiplex sets, as indicated by

162   Weiß et al. (2008). Microsatellite genotyping was performed on an ABI Prism 3100 Genetic Analyzer

163   (Applied Biosystems) using the Macrogen Inc. GenScan service (Seoul, Korea). Negative controls were

164   included for amplification procedures. Results were analysed in GeneMapper v. 4.0 (Applied Biosystems,

165   Foster City, California).

166   Allele frequencies, standard diversity indices, observed heterozygosity (H_O) and expected

167   heterozygosity (H_E) for each locus and population were calculated in GENALEX v. 6 (Peakall & Smouse,

168   2006).

169   We performed a factorial correspondence analysis (FCA) of individual multilocus scores in

170   GENETIX 4.05 (Belkhir et al., 2004) to describe genetic clusters.

171   GENEPOP 3.4 (Raymond & Rousset, 1995; Rousset, 2008) was used to calculated departures from

172   the Hardy-Weinberg equilibrium (HWE) at each locus and within each population. Statistics were

173   computed with Markov chain parameters at default settings.

174   We used ARLEQUIN 3.5 (Excoffier & Lischer, 2010) to estimate the genetic variance within and

175   between populations through a hierarchical Analysis of Molecular Variance (AMOVA; Excoffier et al.,

176   1992).

177   The genetic structure of the sampled populations was computed using Bayesian clustering

178   procedures in STRUCTURE v. 2.3 (Pritchard, Stephens & Donnelly, 2000; Falush, Stephens & Pritchard,

179   2003), without prior information about the origin and under an admixed model. Analyses were performed

180   where K=1-10 with 50×10^5 iterations following a burn-in period of 50×10^4 iterations; all simulations were

181   independently replicated four times for each K. We explored the optimal value of K by plotting the

182   average estimated LnP(D) (Ln probability of the data) and using ΔK statistics (Evanno, Regnaut &

183   Goudet, 2005) calculated using STRUCTURE HARVESTER 0.6.93 (Earl & VonHoldt, 2012). CLUMPP v.

184   1.1.2 (Jakobsson & Rosenberg, 2007) and DISTRUCT v. 1.1 (Rosenberg, 2003) were used to align the

185   cluster membership coefficients of the five STRUCTURE runs and display the results.

186   We investigated the presence of bottleneck events with BOTTLENECK v. 1.2.02 software (Cornuet

187   & Luikart, 1996) for two models: the infinite alleles (IAM, Maruyama and Fuerst, 1985) and the two-

188   phase model (TPM, Di Rienzo et al., 1994).

189   Migration rate was estimated using the Bayesian inference approach implemented in BAYESASS

190   3.0.3 (Wilson and Rannala 2003). We performed 10 runs of 9×10^6 iterations with a burn-in of 10%, and a

191   sampling frequency of 200. Delta values were varied for all parameters, and resulted in acceptance rates

192   between 40% and 60% of the total iterations (Wilson & Rannala, 2003).

193   Finally, isolation by distance was tested via Mantel tests with GENEPOP (Raymond & Rousset,

194   1995; Rousset, 2008; see also Legendre and Fortin, 2010); FST and geographic distance were compared

195   using 1000 random permutations. The geographic distance connecting samples was represented by

196   Euclidean (linear geographic) distances computed in QGis (QGIS Development Team, 2014).

197

198 **Results**

199

200 *mtDNA*

201   The mtDNA marker sequences (CR1 D-Loop 288 bp) showed 24 haplotypes defined by ten

202 polymorphic sites and distributed in eight locations (Appendix 2). Among the 24 haplotypes found

203 (GenBank accession numbers ....-..........), 14 haplotypes were shared by 2-47 individuals. According

204 to the study areas, we found a total of eleven private haplotypes, the majority of which occurred in the

205 Landes (SW France) population (Appendix 2). The diversity indices for mtDNA revealed moderate levels

206 of genetic variation in the greylag goose in all sampled areas (Table 1). Haplotype diversity showed high

207 values in all groups (range 0.798-0.949) except in breeding areas in Finnmark (0.564 ± 0.13 SD) and

208 Vega, Norway (0.170 ± 0.10 SD).

209   The genetic distances recorded in Norway, Northern France and the Netherlands were low (range

210 0.012-0.013). A slightly higher genetic distance was observed in south-west France in comparison to all

211 other sampling sites (ranges 0.018-0.022), while the two breeding sites in Norway were genetically very

212 close (Table 2).

213   The NJ tree shows that clades are composed of a wide variety of different geese from different

214 areas. Individuals from each site were present in different branches (Fig. 2). Besides, none of the clades

215 grouped together individuals originating from the same areas. About half the individuals were grouped

216 together with the GenBank reference sequence relating to the *anser* subspecies, while the remainder were

217 either grouped with the sequence relating to the *rubrirostris* subspecies or differed clearly from both

218 subspecies. Very similar topologies were obtained from trees generated with other tree-building methods

219 (MP and ML; not shown).

220   The haplotype median-joining network (Fig. 2) was concordant with the phylogenetic tree

221 topology and did not reveal any geographic structures. The number of mutations separating the different

222 haplotypes was low (max = 10).

223   Whilst 88.85% of the total genetic variance shown in hierarchical AMOVA was within

224   populations, the remaining 11.15% occurred among populations. This indicates a small differentiation

225   between the sampled areas.

226   Non-significant raggedness indices indicated a good fit to a model of population expansion in all

227   sampled areas. Mismatch distribution results also suggested a population expansion in all areas except the

228   Gironde region (P = 0.044) and Finnmark (P = 0.042) (Appendix 4). Fu's FS value (Table 1) was only

229   significantly negative for the Oise region, and was consistent with a demographic expansion for all other

230   areas.

231

232   *Microsatellites*

233   Among the 15 microsatellites previously isolated by Weiß et al. (2008) only Ans26 was shown to

234   be monomorphic in all investigated individuals. The remaining 14 polymorphic microsatellite loci

235   showed 2-12 different alleles per locus ($n = 169$ individuals; Table 3).

236   Observed and expected heterozygosities were moderate, with similar values in each sampled

237   population (H_o ranging from 0.374 to 0.484 and H_e from 0.433 to 0.549). Geese from the Landes

238   wintering area exhibited the highest number of private alleles ($n = 3$, Appendix 3).

239   Genetic structure was visualized using factorial correspondence analysis (FCA) in each

240   population (Fig. 3). The plot shows an absence of phylogeographic structure in the different investigated

241   areas: individuals from different areas overlap, with the exclusion of four samples from Nord, one sample

242   from Finnmark, one from Oise and one from Gironde.

243   Significant departures from HWE, due to heterozygote deficit and related to positive Fis values,

244   were observed in all populations (Table 3, Appendix 3).

245   AMOVA analyses showed that 97.9% of the total genetic variance in geese was significantly

246   distributed within populations (p < 0.001), while only 2.1% was distributed among populations.

247   Overall fixation index FST from AMOVA was 0.02105, indicating a low differentiation between areas.

248   STRUCTURE analyses, performed without the use of prior information on sample locations,

249   showed a maximum ΔK at $K = 4$, while likelihood values reached a plateau at $K = 7$ (Fig. 4). Graphs

250   show no evidence of phylogeographic structure across sampled populations, whatever the K value. With

251     $K = 4$, only 23 individuals with individual q_i values were each assigned to a single cluster: two

252   individuals from Finnmark, three from Vega and one from Gironde were attributed to cluster 1; one

253   individual from Netherland, two from Finnmark, six from Vega, one from Charente Maritime, two from

254   Gironde and one from Landes were attributed to cluster 2; two individuals from Oise and two from

255   Landes were assigned to cluster 3. All other birds had a highly mixed genotype. In the case of $K = 7$, five

256   other individuals, one from Finnmark, Charente Maritime and Gironde and two from Vega, were assigned

257   to the same cluster with $q_i > 0.90$.

258   Bottleneck events tested under IAM revealed a significant excess of heterozygotes (evidence of a

259   recent bottleneck) in Nord, Landes and Oise populations (Wilcoxon sign-rank tests, all $P < 0.05$).

260   Analysis under TPM only confirmed a recent bottleneck event for the Nord population ($P < 0.05$).

261   BAYESASS detected a low migration rate among localities and a high proportion of local

262   individuals ($> 68\%$, Table 4), suggesting that the flows among different areas were limited. Indeed, the

263   analysis found a high proportion of local geese in six populations ($> 90\%$). In two cases, gene flow

264   appears to be strongly asymmetrical, with many birds moving from Charente Maritime to the Netherlands

265   ($20.7\% \pm 3.79$ SD) and from Oise to Gironde ($20.2\% \pm 3.44$ SD), but not in the opposite direction (1.5%

266   and 1.1% respectively).

267   The Mantel test calculated on geographic and genetic distances yielded a non-significant

268   correlation coefficient ($r = 0.107$; $P = 0.08$), suggesting that there is no strong relationship between

269   geographic and genetic distances.

270

271

272   **Discussion**

273

274   In this study we used a pool of 14 microsatellites isolated by Weiß et al. (2008) for greylag goose

275   parentage in the long-established goose population at Konrad Lorenz Research Station, Grünau, Austria

276   (Lorenz, 1966; Hirschenhauser, Möstl & Kotrschal, 1999). We found that these microsatellites can be

277   successfully employed for geese sampled in a wide range of localities along the European Atlantic

278   flyway. This is the first large scale study showing a moderate genetic variability of mtDNA and nuclear

279   DNA in all French wintering areas and in the Netherlands, with slightly lower mtDNA variability in the

280   Norwegian breeding sites. A moderate genetic variability in the greylag goose was already reported two

281   decades ago by Blaakmeer (1995), and has been found in other species of geese (*Anser erythropus*:

282   Ruokonen et al., 2004; Ruokonen et al., 2010; *Anser brachyrhynchus*: Ruokonen et al., 2005). Low

283   genetic variability also seems to be typical for other Anatidae species (*Aythya ferina*: Liu et al., 2011;

284   *Aythya fuligula*: Liu et al., 2012). Interestingly, our results show that the genotypes deviated from Hardy-

285   Weinberg expectation at eight loci, and in all study areas, were due to heterozygote deficiency. Besides,

286   the deficit of heterozygotes matched with positive FIS values. These results could be related to different

287   factors, such as population substructuring or recent population growth (Cornuet & Luikart, 1996).

288   Genetic distances between the different areas were low (range from 0.012 to 0.017) and the

289   hierarchical AMOVA showed genetic variance to mainly occur within populations. These findings could

290   be explained by a small differentiation between the sampled areas and a general admixture of greylag

291   goose populations in our Western European study region. However, it should be taken into account that

292   genetic divergences in geese are characteristically very low, with the lowest interspecific divergence

293   reported here for avian species (Ruokonen, Kvist & Lumme, 2000). The genetic tree shows that different

294   branches include individuals from each sampling area. No single branch exclusively grouped together

295   individuals originating from the same zone. Moreover, birds sampled in the western part of the breeding

296   range, traditionally ascribed to the *anser* subspecies, were not separated from birds collected in the

297   eastern part that were traditionally assigned to the *rubrirostris* subspecies (Kampe-Persson, 2002). Birds

298   from Iceland, Scotland and coastal Norway have been sometimes separated as a race, *sylvestris*, classified

299   in the *anser* group (Snow, Perrins & Cramp, 1998). Although Icelandic and Scottish birds were absent

300 from this study, individuals from the Norwegian west coast did not appear to be clearly distinct from

301 other European geese. Our present results slightly differ from the findings of Blaakmeer's preliminary

302 study (1995), which reported genetic differences between breeders in two Dutch sites in comparison to

303 breeding sites in South Sweden and Norway. However, Blaakmeer's analyses (1995) show significant

304 differences for only one of six minisatellites, in only two of the three Dutch areas studied.

305 Interestingly, the ANS19 sequence was recently found in the white Roman goose in Taiwan. This

306 race is widely bred for commercial purposes, and has been found to originate from the European species

307 (*Anser anser,* Wang et al., 2010). Our data confirm the presence of this sequence in Europe, particularly

308 in the breeding population of the Norwegian west coast.

309 Haplotypes ANS02, ANS08, ANS11, ANS23 and ANS24 were only found during the winter in

310 France, and were absent in Norway and the Netherlands: this result could indicate that some of the geese

311 arriving in France came from areas we did not sample on the breeding grounds. Ring recoveries and

312 resighting records indicate that these birds probably originated from North Germany, Poland, Denmark

313 and Sweden (Nilsson et al., 2013).

314 The haplotype network confirmed the tree configuration. There was no geographic pattern, and

315 the number of mutations separating the different nodes was very low. This confirms the low genetic

316 distance between our studied populations in the large north-western European population (as defined by

317 Delany and Scott, 2006), and may reflect the rapid population expansion (Aris-Brosou & Excoffier,

318 1996).

319 Data obtained from nuclear DNA by microsatellites were in accordance with findings from

320 mtDNA. As the mtDNA is uniparentally inherited whereas microsatellites are part of the biparentally

321 inherited nuclear DNA, a difference between the two genomes would have indicated the presence of sex-

322 biased dispersal (Fahey, Ricklefs & Dewoody, 2014). However, sex-biased dispersal seems to be unlikely

323 in greylag geese for three reasons: the family unit remains together at least until autumn migration, the

324 birds tend to pair before returning to the breeding grounds, and males and females have long-term pair

325 bonds (Rohwer & Anderson, 1988; Doherty et al., 2002). Sex-biased dispersal in birds is probably not a

326   species constant (Clarke et al., 1997). Within Anatidae in general, sex-biased dispersal was not detected

327   in several species (Doherty et al., 2002; Mabry et al., 2013), while it was found in some species such as

328   the white-fronted goose (*Anser erythropus,* Ruokonen et al., 2010), the common eider (*Somateria*

329   *mollissima,* Paulus & Tiedemann, 2003), and the spectacled eider (*Somateria fisheri,* Scribner et al.,

330   2001).

331   In our study of microsatellites, individuals from different geographic localities were found to be

332   combined in the Factorial Correspondence Analysis representation. Bayesian structure analysis resulted

333   in a best combination of four or seven groups, according to ΔK and LnP(D) methods respectively. As

334   seen in our previous analyses, no geographic clustering was observed inside these STRUCTURE groups.

335   Almost all individuals, with few exceptions, showed admixed genotype regardless of the number of

336   groups considered in the analysis.

337   The high mixing of genotypes and the lack of geographic structure among our studied

338   populations could be interpreted in the light of the data obtained through ringing activity and the

339   extensive neck-banding programme carried out in Scandinavia from 1984 to 2004 (Nilsson, 2007;

340   Voslamber, Knecht & Kleijn, 2010). Ring recoveries and visual observations showed that Scandinavian

341   geese breeding in different zones can admix not only in the moulting areas (Nilsson, Kahlert & Persson,

342   2001) but also along the European Atlantic flyway, i.e. in the Netherlands (Voslamber, Knecht & Kleijn,

343   2010), France (Fouquet, Schricke & Fouque, 2009; Nilsson et al., 2013) and Spain (Ramo et al., 2012)

344   where they can form pairbonds. Besides this Scandinavian data, the monitoring of collared and/or ringed

345   individuals performed in other European areas showed the presence of birds in France originating from

346   Germany, the Czech Republic and Poland. Populations that breed further east do not seem to reach

347   France in winter (Kampe-Persson, 2010). From these data it appears that the greylag geese that cross

348   France or winter there could result from a mixture of populations from different areas.

349   Our findings are somewhat unexpected if one assumes that the fragmentation of breeding

350   populations into separate areas during the first part of the last century (Hagemeijer & Blair, 1997;

351   Kampe-Persson, 2002), should have led to an increase in genetic structure. Moreover, birds breeding in

352 the Netherlands have recently become highly sedentary (Fox et al., 2010), and this may also have

353 contributed to the increase in genetic structure (Blaakmeer, 1995). However, a genetic panmixia could

354 have been promoted by the widespread amateur breeding and selling of geese, and the recent increase and

355 dispersal of several wild goose populations (Klok et al., 2010). In particular, geese with pink bills and

356 legs, most probably *rubirostris* subspecies, have been spreading in Europe over the last few decades;

357 their natural flyway toward wintering areas crosses other European countries (from Russia to Hungary,

358 the Balkan States and Italy) but does not reach France.

359   The breeding of geese is a widespread practice among amateurs, who can easily obtain both

360 goslings and adults with a grey wild appearance (B. and G. Vaschetti, pers. comm.). In some cases geese

361 were released as part of assisted restoring projects and are now indistinguishable from the wild

362 individuals (Kampe-Persson, 2010). Besides, birds with white plumage are common in breeding farms. In

363 Asia, white geese are mostly descendants of the swan goose (*Anser cygnoides*). Even if descendants of

364 *Anser anser* can also be found there, they are usually farmed in Europe (Wang et al., 2010). Although the

365 two species can hybridize in captivity, hybrids can be detected through karyotype (Shahin, Ata & Shnaf,

366 2014) or genotype examinations (Sun et al., 2014). The contribution of escaped white form geese to the

367 admixture observed in wild populations is probably low given the high assortative mating of wild greylag

368 geese, their long-term monogamous pair bonds, female-bonded clan structure, long parent-offspring

369 relationships, and elaborate patterns of mutual social support (Hirschenhauser et al., 2000; Kotrschal,

370 Scheiber & Hirschenhauser, 2010).

371   Our findings on the greylag goose genetic admixture are similar to those reported in the snow

372 goose (*Chen caerulescens,* Avise et al., 1992) and the barnacle goose (*Branta leucopsis,* Jonker et al.,

373 2013). Despite the high rate of site philopatry seen in the snow goose, which has also shown a high

374 increase in population over the last decades, mtDNA markers showed no clear distinctions between

375 nesting populations across species range (Avise et al., 1992). The barnacle goose recently changed its

376 migratory traditions, and new populations differing in migratory distance were observed. Genetic data

377 showed an admixture between all populations, despite the assumed traditions of migration within areas

378   and the presence of a newly established nonmigratory population in the Netherlands (Jonker et al., 2013).

379   A lack of genetic structure in wintering areas was also found in four species of Anatidae, namely the

380   common pochard, the mallard, the king eider and the tufted duck (Pearce et al., 2004; Liu, Keller &

381   Heckel, 2011, 2012; Kraus et al., 2013; Liu et al., 2013), and in the black-tailed godwit (*Limosa limosa,*

382   Lopes et al., 2013). The mixing of breeding populations in wintering areas is believed to be a common

383   phenomenon in birds, because the breeding ranges of most species are considerably larger than their

384   wintering ranges (Winker & Graves, 2008). However, migratory populations vary in the degree to which

385   individuals from distinct breeding localities mix on different sites. Therefore, to understand population

386   demographics and genetic diversification, it is crucial to pinpoint which populations mix on breeding and

387   wintering grounds (Chabot et al., 2012).

388   Our DNA-based estimates of migration during the wintering period indicated a low rate of

389   exchange between our sampled areas. In five of eight areas the vast majority of individuals (86-95%) did

390   not switch among the different zones, and a moderate exchange (about 20%) was only observed from

391   Charente Maritime to the Netherlands and from Oise to Gironde. These results seem to support the

392   hypothesis that the French wintering birds arrive from various areas, including zones that are not sampled

393   here (i.e. Germany, Poland), while the contribution of the Norwegian population represents only a portion

394   of the whole assemblage (Fouquet, Schricke & Fouque, 2009). This low exchange rate is also supported

395   by evidence of great changes in spatial ecology recorded in 28 GPS tagged western European greylag

396   geese, i.e. very low home ranges on wintering areas (<8.9±2.5 km²) compared to 2-5 fold values obtained

397   in staging areas during migratory and premigratory periods (Boos unpublished data).

398   Our data show evidence of genetic bottlenecks in just three groups under IAM, all located along

399   the same flyway (Nord, Oise and Landes), and in a single case (Nord) under TPM. The discrepancy

400   between IAM and TPM could be related to the different ability of the mutation models to detect

401   bottleneck events. Empirical data suggest that TPM is the most appropriate model for microsatellite loci

402   (Ellegren, 2000, 2004) while IAM results should be interpreted with caution (Cornuet & Luikart, 1996).

403   We did not observe any sign of bottlenecks in the breeding populations: this indicates that greylag geese

404 have never suffered any severe demographic reduction, even at the beginning of the past century when

405 the number of breeding individuals was low in several European areas (Kampe-Persson, 2002).

406 Greylag goose populations are steadily increasing in North-Western Europe (Kampe-Persson,

407 2002). The large number of geese in some areas is now in conflict with agricultural interests, since geese

408 not only forage in natural environments but also forage on crop fields, and claims for a need to control the

409 species are widespread (Klok et al., 2010). Our data suggest that the migratory geese harvested over

410 France (about 20000 geese/year, see Landry and Migot, 2000) show a relatively high diversity of origin.

411 From this result it is difficult to conclude if there is a strong impact on a specific breeding population.

412 Future studies could analyze other European breeding and wintering areas; this could clarify the status of

413 the different populations and subspecies on the continent (the main *Anser anser anser* and *A. a.*

414 *rubrirostris*, as well as the *sylvestris* forms from Iceland, Scotland and Norway), and help to build an

415 effective international management strategy for this migratory species (Chabot et al., 2012).

416

417 **Acknowledgements** – Our thanks to all the hunters who provided feather samples from France, to

418 Wilmer Remijnse who provided feather samples from Netherlands, and to Paul Shimmings for his field

419 assistance in Norway. We are grateful to Bruno and Gabriella Vaschetti which provided useful

420 information on geese farming, the International Union for Conservation of Nature IUCN which provided

421 an updated distribution map, and to Joanna Lignot-Munro for language editing.

422
423
424 **References**
425

426 Andersson Å, Follestad A, Nilsson L, Persson H. 2001. Migration patterns of Nordic Greylag Geese *Anser anser*.
427   *Ornis Svecica* 11:19–58.

428 Aris-Brosou S, Excoffier L. 1996. The impact of population expansion and mutation rate heterogeneity on DNA
429   sequence polymorphism. *Molecular Biology and Evolution* 13:494–504.

430 Avise JC, Alisauskas RT, Nelson WS, Ankney CD. 1992. Matriarchal population genetic structure in an avian
431   species with female natal philopatry. *Evolution* 46:1084–1096.

432   Bandelt H-J, Forster P, Rohl A. 1999. Median-Joining networks for inferring intraspecific phylogenies. *Molecular*
433   *biology and evolution* 16:37–48.

434   Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 2004. *GENETIX 4.05, logiciel sous Windows TM pour la*
435   *génétique des populations*. Montpellier, France.: Laboratoire Genome, Populations, Interactions, CNRS UMR
436   5000, Université de Montpellier II.

437   BirdLife International B. 2004. *Birds in Europe, population estimates, trends and conservation status*. Cambridge.

438   Blaakmeer K. 1995. *Genetic variation within the West European population of the Greylag goose* (Anser anser).
439   Groningen.

440   Boos M, Schricke V, Green AJ, Shimmings P, Lefranc H, Amat JA, Ramo C, Follestad A. 2012. Migration of
441   Greylag Geese tagged in Norway and Spain using GPS devices: first results from a new joint European
442   Research Program. In: *14th Meeting of the Goose Specialist Group*. 8. Steinkjer, Norway.

443   Burnham KP, Anderson DR. 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection.
444   *Sociological Methods & Research* 33:261–304.

445   Chabot AA, Hobson KA, Van Wilgenburg SL, McQuat GJ, Lougheed SC. 2012. Advances in linking wintering
446   migrant birds to their breeding-ground origins using combined analyses of genetic and stable isotope markers.
447   *PloS one* 7:e43627.

448   Clarke AL, Saether B-E, Roskaft E. 1997. Sex biases in avian dispersal: a reappraisal. *Oikos* 79:429-438.

449   Comolet-Tirman J. 2009. *L'Oie cendrée Anser anser en France et en Europe. Dynamique de population, statuts de*
450   *conservation, voies de migration et dates de migration prénuptiale*. Service du Patrimoine Naturel
451   Département Ecologie et Gestion de la Biodiversité.

452   Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population
453   bottlenecks from allele frequency data. *Genetics* 144:2001–14.

454   Cramp S. 1977. *Handbook of the birds of Europe, the Middle East, and North Africa. The birds of the Western*
455   *Palearctic, vol. 1*. Oxford: Oxford University Press.

456   Delany S, Scott D. 2006. *Waterbird population estimates, Fourth Edition*. Wageningen, the Netherlands.: Wetland
457   International.

458   Doherty PF, Nichols JD, Tautin J, Voelzer JF, Smith GW, Benning DS, Bentley VR, Bidwell JK, Bollinger KS,
459   Brazda AR, Buelna EK, Goldsberry JR, King RJ, Roetker FH, Solberg JW, Thorpe PP, Wortham JS, Service
460   USG, Wildlife P, Road BF. 2002. Sources of variation in breeding-ground fidelity of mallards (*Anas*
461   *platyrhynchos*). *Behavioral Ecology* 13:543–550.

462   Earl DA, VonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE
463   output and implementing the Evanno method. *Conservation Genetics Resources* 4:359–361.

464   Ellegren H. 2000. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in*
465   *Genetics* 16:551–558.

466   Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. *Nature reviews. Genetics* 5:435–45.

467 Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software
468   STRUCTURE: a simulation study. *Molecular Ecology* 14:2611–2620.

469 Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics
470   analyses under Linux and Windows. *Molecular Ecology Resources* 10:564–567.

471 Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among
472   DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.

473 Fahey AL, Ricklefs RE, Dewoody JA. 2014. DNA-based approaches for evaluating historical demography in
474   terrestrial vertebrates. *Biological Journal of the Linnean Society* 112:367–386.

475 Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked
476   loci and correlated allele frequencies. *Genetics* 164:1567–1587.

477 Felsenstein J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics*
478   22:521–565.

479 Felsenstein J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6.

480 Fouquet M. 1991. Migration et hivernage de l'oie cendrée (*Anser anser*) en France. Rôle et importance du Centre-
481   Ouest. *Oiseau et la Revue Francaise d'Ornithologie* 61:111–130.

482 Fouquet M, Schricke V, Fouque C. 2009. Greylag Geese *Anser anser* depart earlier in spring: an analysis of goose
483   migration from western France over the years 1980 – 2005. *Wildfowl* 59:143–151.

484 Fox AD, Ebbinge BS, Mitchell C, Heinicke T, Aarvak T, Colhoun K, Clausen P, Dereliev S, Faragó S, Koffijberg
485   K, Kruckenberg H, Pihl S, Jeugd HVANDER. 2010. Current estimates of goose population sizes in western
486   Europe, a gap analysis and an assessment of trends. *Ornis Svecica* 20:115–127.

487 Hagemeijer EJM, Blair MJ. 1997. *The EBCC Atlas of European Breeding Birds: Their Distribution and Abundance.*
488   London: T. & A. D. Poyser.

489 Hall T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows
490   95/98/NT. *Nucleic Acids Symposium Series* 41:95–98.

491 Hirschenhauser K, Mo E, Bernard Ã, John W, Kurt D. 2000. Endocrine and behavioural responses of male Greylag
492   Geese *Anser anser* to pairbond challenges during the reproductive season. *Ethology* 106:63–77.

493 Hirschenhauser K, Möstl E, Kotrschal K. 1999. Seasonal patterns of sex steroids determined from feces in different
494   social categories of Greylag geese (*Anser anser*). *General and comparative endocrinology* 114:67–79.

495 IUCN 2015. The IUCN Red List of threatened species. Version 2015.1. *http://www.iucnredlist.org*. Downloaded on
496   11 May 2015.

497 Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label
498   switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806.

499 Johnsen A, Rindal E, Ericson PGP, Zuccon D, Kerr KCR, Stoeckle MY, Lifjeld JT. 2010. DNA barcoding of
500   Scandinavian birds reveals divergent lineages in trans-Atlantic species. *Journal of Ornithology* 151:565–578.

501 Jonker RM, Kraus RHS, Zhang Q, van Hooft P, Larsson K, van der Jeugd HP, Kurvers RHJM, van Wieren SE,
502   Loonen MJJE, Crooijmans RPMA, Ydenberg RC, Groenen MAM, Prins HHT. 2013. Genetic consequences
503   of breaking migratory traditions in barnacle geese *Branta leucopsis*. *Molecular ecology* 22:5835–5847.

504 Kampe-Persson H. 2002. *Anser anser* Greylag Goose. *BWP Update* 4:181–216.

505 Kampe-Persson H. 2010. Naturalised geese in Europe. *Ornis Svecica* 20:155–173.

506 Klok C, Schekkerman H, Willems F, Ebbinge B, van Turnhout C. 2010. Analysis of population development and
507   effectiveness of management in resident greylag geese *Anser anser* in the Netherlands. *Animal Biology*
508   60:373–393.

509 Kotrschal K, Scheiber IBR, Hirschenhauser K. 2010. Individual performance in complex social systems: the greylag
510   goose example. In: Kappeler P ed. *Animal Behaviour: Evolution and Mechanisms*. Berlin: Springer Verlag,
511   121–148.

512 Kraus RHS, van Hooft P, Megens H-J, Tsvey A, Fokin SY, Ydenberg RC, Prins HHT. 2013. Global lack of flyway
513   structure in a cosmopolitan bird revealed by a genome wide survey of single nucleotide polymorphisms.
514   *Molecular Ecology* 22:41–55.

515 Landry P, Migot P. 2000. Enquête nationale sur les tableaux de chasse à tir: saison 1998/1999. *Faune Sauvage*
516   *Cahiers techniques n° 251*.

517 Legendre P, FortinM-J. 2010. Comparison of the Mantel test and alternative approaches for detecting complex
518   multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources* 10:831–844.

519 Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.
520   *Bioinformatics* 25:1451–1452.

521 Liu S, Li Y, Lu J, Su X, Tang M, Zhang R, Zhou L, Zhou C, Yang Q, Ji Y, Yu DW, Zhou X. 2013. SOAPBarcode:
522   revealing arthropod biodiversity through assembly of Illumina shotgun sequences of PCR amplicons. *Methods*
523   *in Ecology and Evolution* 4:1142–1150.

524 Liu Y, Keller I, Heckel G. 2011. Range-wide genetic population structure of common pochard (*Aythya ferina*): a
525   potentially important vector of highly pathogenic avian influenza viruses. *Ecology and Evolution* 1:529–545.

526 Liu Y, Keller I, Heckel G. 2012. Breeding site fidelity and winter admixture in a long-distance migrant, the tufted
527   duck (*Aythya fuligula*). *Heredity* 109:108–116.

528 Liu Y, Keller I, Heckel G. 2013. Temporal genetic structure and relatedness in the Tufted Duck *Aythya fuligula*
529   suggests limited kin association in winter. *Ibis* 155:499–507.

530 Lopes RJ, Alves JA, Gill JA, Gunnarsson TG, Hooijmeijer JCEW, Lourenco PM, Masero JA, Piersma T, Potts PM,
531   Rabacal B, Reis S, Sanchez-Guzman JM, Santiago-Quesada F, Villegas A. 2013. Do different subspecies of
532   Black-tailed Godwit *Limosa limosa* overlap in Iberian wintering and staging areas? Validation with genetic
533   markers. *Journal of Ornithology* 154:35–40.

534 Lorenz KZ. 1966. The triumph ceremony of the Greylag Goose, *Anser anser* . *Philosophical Transactions of the*
535   *Royal Society B: Biological Sciences* 251:477–477.

536 Mabry KE, Shelley EL, Davis KE, Blumstein DT, Van Vuren DH. 2013. Social mating system and sex-biased
537   dispersal in mammals and birds: a phylogenetic analysis. *PloS one* 8:e57980.

538 Maruyama T, Fuerst PA. 1985. Population bottlenecks and non equilibrium models in population genetics. II.
539   Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* 111:675–689.

540 Nilsson L. 2007. The Nordic Greylag Goose (*Anser anser*) Project. *Aves* 44:177 – 184.

541 Nilsson L, Follestad A, Guillemain M, Schricke V, Voslamber B. 2013. France as a staging and wintering area for
542   Greylag Geese *Anser anser*. *Wildfowl* 63:24–39.

543 Nilsson L, Kahlert J, Persson H. 2001. Moult and moult migration of Greylag Geese *Anser anser* from a population
544   in Scania, south Sweden. *Bird Study* 48:129–138.

545 Paulus KB, Tiedemann R. 2003. Ten polymorphic autosomal microsatellite loci for the Eider duck *Somateria*
546   *mollissima* and their cross-species applicability among waterfowl species (Anatidae). *Molecular Ecology*
547   *Notes* 3:250–252.

548 Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and
549   research. *Molecular Ecology Notes* 6:288–295.

550 Pearce JM, Talbot SL, Pierson BJ, Petersen MR, Scribner KT, Dickson DL, Mosbech A. 2004. Lack of spatial
551   genetic structure among nesting and wintering King Eiders. *The Condor* 106:229–240.

552 Persson H. 1993. Arrival patterns of Greylag Geese *Anser anser* in the Guadalquivir Marismas. *Wildfowl* 44:19–23.

553 Pistorius PA., Follestad A, Nilsson L, Taylor FE. 2007. A demographic comparison of two Nordic populations of
554   Greylag Geese *Anser anser*. *Ibis* 149:553–563.

555 Pistorius PA, Follestad A, Taylor FE. 2006. Declining winter survival and fitness implications associated with
556   latitudinal distribution in Norwegian Greylag Geese *Anser anser*. *Ibis* 148:114–125.

557 Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular biology and evolution* 25:1253–1256.

558 Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data.
559   *Genetics* 155:945–959.

560 QGIS Development Team T. 2014. QGIS Geographic Information System. *Open Source Geospatial Foundation*
561   *Project*.

562 Rambaut A. 2008. FigTree: Tree Figure Drawing Tool Version 1.3. 1 2006–2009. *Institute of Evolutionary Biology,*
563   *University of Edinburgh. http://tree.bio.ed.ac.uk.*

564 Ramo C, Amat JA, Calderón J, Gómez-Crespo E, Navedo JG, Green AJ, Jubete F, Masero JA, Palacios J,
565   Rodríguez-Alonso M, Boos M, Schricke V. 2012. Distribution and population trends of wintering Greylag
566   Geese in Spain. In: *14th Meeting of the Goose Specialist Group*. 28.

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

569 Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB. 1994. Mutational processes of simple-
570   sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United*
571   *States of America* 91:3166–3170.

572   Rohwer FC, Anderson MG. 1988. Female-biased philopatry, monogamy, and the timing of pair formation in
573   migratory waterfowl. In: Johnston E, Richard F eds. *Current Ornithology v.5*. Boston: Springer US, 187–221.

574   Rosenberg NA. 2003. Distruct: a program for the graphical display of population structure. *Molecular Ecology*
575   *Notes* 4:137–138.

576   Rousset F. 2008. GENEPOP'007: a complete reimplementation of the GENEPOP software for Windows and Linux.
577   *Molecular ecology resources* 8:103–106.

578   Ruokonen M, Kvist L, Aarvak T, Markkola J, Morozov V, Øien IJ, Jr EES, Tolvanen P. 2004. Population genetic
579   structure and conservation of the lesser white-fronted goose *Anser erythropus*. *Conservation Genetics*:501–
580   512.

581   Ruokonen M, Aarvak T, Chesser RK, Lundqvist A-C, Merilä J. 2010. Temporal increase in mtDNA diversity in a
582   declining population. *Molecular Ecology* 19:2408–2417.

583   Ruokonen M, Aarvak T, Madsen J. 2005. Colonization history of the high-arctic pink-footed goose *Anser
584   brachyrhynchus*. *Molecular Ecology* 14:171–8.

585   Ruokonen M, Kvist L, Lumme J. 2000. Close relatedness between mitochondrial DNA from seven *Anser* goose
586   species. *Journal of Evolutionary Biology* 13:532–540.

587   Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees.
588   *Molecular biology and evolution* 4:406–425.

589   Scribner KT, Petersen MR, Fields RL, Talbot SL, Pearce JM, Chesser RK. 2001. Sex-biased gene flow in
590   Spectacled Eiders (Anatidae): inferences from molecular markers with contrasting modes of inheritance.
591   *Evolution* 55:2105–2115.

592   Shahin AAB, Ata ATM, Shnaf ASMA. 2014. Karyotype and C-banding pattern of the domestic geese *Anser anser*
593   populations (Aves: Anatidae) in Egypt. *Folia biologica* 62:49–58.

594   Snow DW, Perrins CM, Cramp S. 1998. *The Complete Birds of the Western Palaearctic: On CD-ROM*. Oxford:
595   Oxford University Press.

596   SOVON. 1987. *Atlas van de Nederlandse vogels*. Nederland, Leiden.: SOVON Vogelonderzoek Nederland,
597   Arnhem NL.

598   Sun J, Zhang S, He D-Q, Chen S-Y, Duan Z-Y, Yao Y-G, Liu Y-P. 2014. Matrilineal genetic structure of domestic
599   geese. *Journal of Poultry Science* 51:130.

600   Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics
601   Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular
602   Biology and Evolution* 28:2731–2739.

603   Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial
604   DNA in humans and chimpanzees. *Molecular biology and evolution* 10:512–526.

605   Voslamber B, Knecht E, Kleijn D. 2010. Dutch Greylag Geese *Anser anser*: migrants or residents? *Ornis Svecica*
606   20:207–214..

607 Wang CM, Way TD, Chang YC, Yen NT, Hu CL, Nien PC, Jea YS, Chen LR, Kao JY. 2010. The origin of the
608   white Roman goose. *Biochemical genetics* 48:938–43.

609 Weiß BM, Poggemann K, Olek K, Foerster K, Hirschenhauser K. 2008. Isolation and characterization of
610   microsatellite marker loci in the greylag goose (*Anser anser*). *Molecular Ecology Resources* 8:1411–1413.

611 Wilson GA, Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*
612   163:1177–1191.

613 Winker K, Graves GR. 2008. Genetic structure of breeding and wintering populations of Swainson's Warbler. *The*
614   *Wilson Journal of Ornithology* 120:433–445.

615

617
618 **Supplemental Information**
619
620
621 **Appendix 1** - List of samples used in this study, with ID, population, sample provenance, collecting date,
622   haplotype assignment and whether the sample was genotyped with microsatellites (STRs).
623
624 **Appendix 2 -** Table of mtDNA haplotypes found in *Anser anser* individuals.
625
626 **Appendix 3 –** Summary of genetic variation at 14 microsatellite loci in sampled populations. Na = No. of
627   different alleles; Ne = No. of effective alleles. Ho = observed heterozygosity; He = expected
628   heterozygosity; F = fixation index; HWE = Hardy-Weinberg equilibrium.
629
630 **Appendix 4** - Distributions of pairwise differences (mismatch distribution) among mtDNA haplotypes
631   for overall dataset and each sampled areas.
632
633
634
635
636   **Figure Captions**
637
638
639 **Figure 1** – Main *Anser anser* flyways from breeding (red) to wintering (blue) areas (modified from
640   IUCN, 2015). Pie charts indicate the proportion of different haplotypes (mtDNA) found in each
641   sampled population. Colours are identical to those used in the haplotype network (Fig. 2), and
642   haplotypes found in one area alone are the same colour.
643
644 **Figure 2** - Left: median-joining haplotype network. Areas of circles represent different sampled mtDNA
645   haplotypes in proportion to their frequencies. Distances between haplotypes are proportional to
646   the number of base differences. Colours match those utilized in Fig. 1, and haplotypes found in
647   one area alone are the same colour. Right: Neighbour-joining tree based on 280 bp of CR1.
648   Sampled areas are labelled with abbreviations: NF, Northern France (Oise and Nord); SWF,
649   South-Western France (Gironde, Charente Maritime and Landes); Nor, Norway; Neth,
650   Netherlands. Numbers below branches indicate bootstrap values; only values above 50% are
651   shown, most of the clades are supported by low bootstrap values.
652
653 **Figure 3** - Factorial Correspondence Analysis (FCA) of microsatellites data. Outliers concern individuals
654   from Nord (4 ind.), Oise, Gironde and Finnmark (1 ind.) populations.
655
656 **Figure 4** - Estimated population structure in Greylag Goose sampled populations. Each vertical line
657   represents one individual and each colour represents a single cluster.
658

Map of Anser in Europe

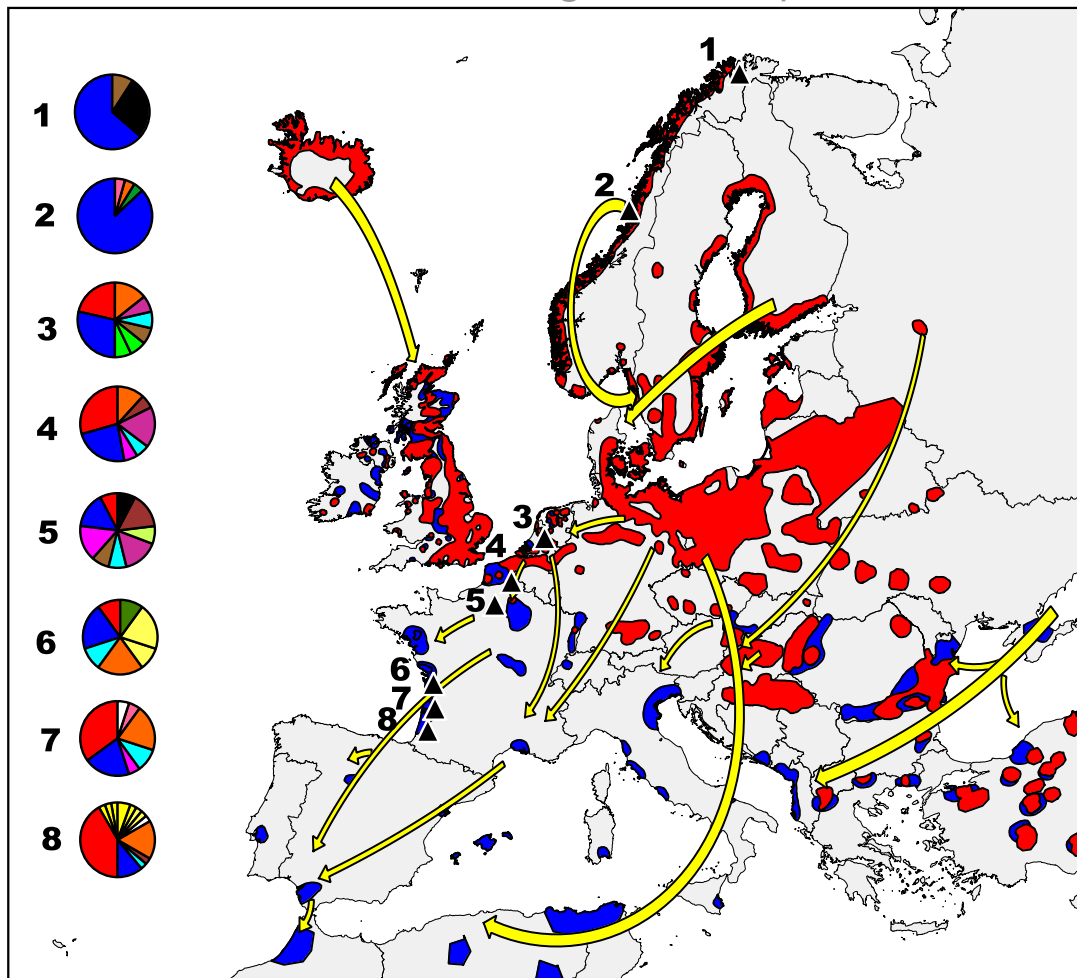
Figure 1 - Main *Anser anser* flyways from breeding (red) to wintering (blue) areas (modified from IUCN, 2015). Pie charts indicate the proportion of different haplotypes (mtDNA) found in each sampled population. Colours are identical to those used in the haplotype network (Fig. 2), and haplotypes found in one area alone are the same colour.

2

Image of network

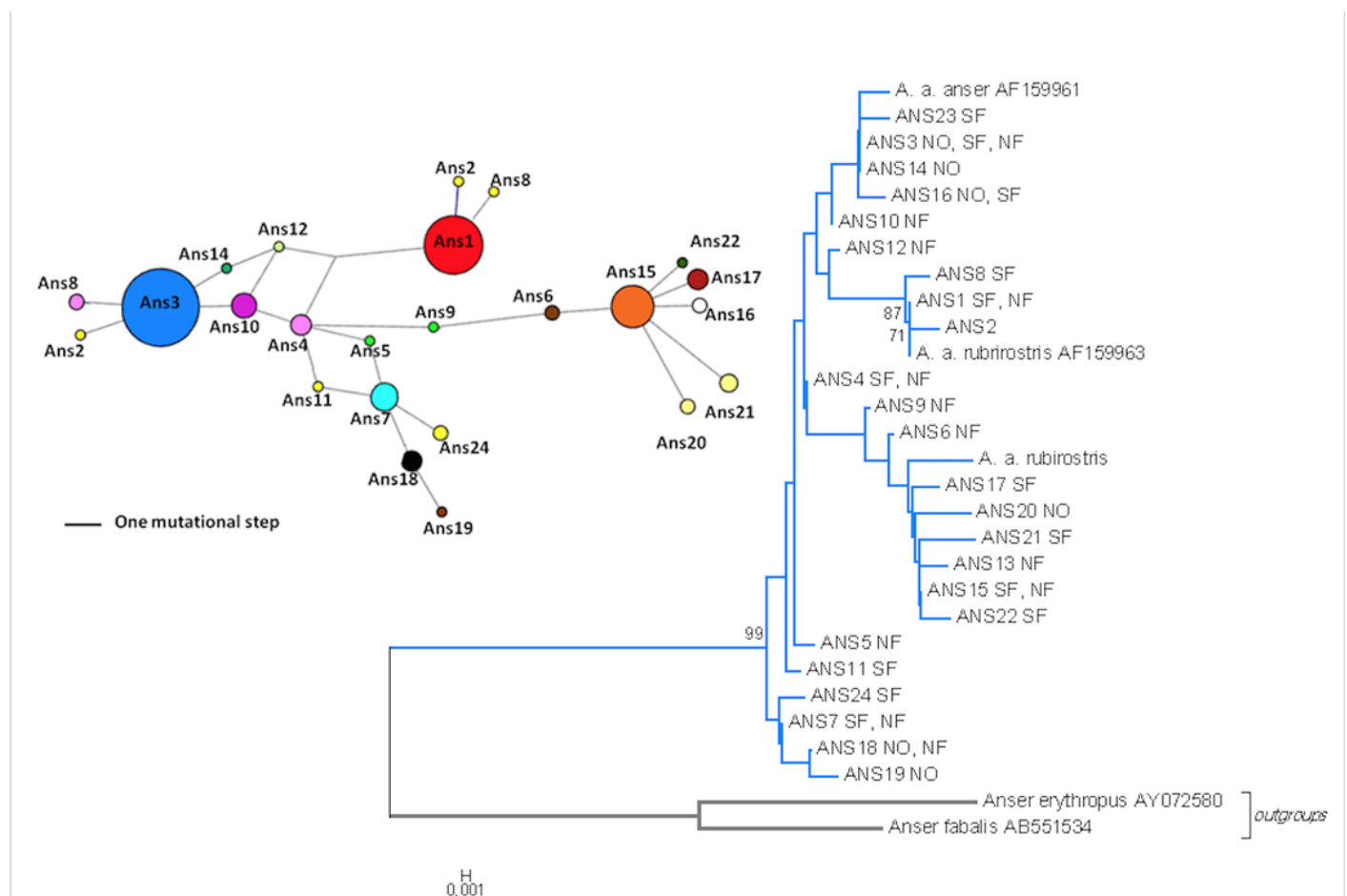
Figure 2 - Left: median-joining haplotype network. Areas of circles represent different sampled mtDNA haplotypes in proportion to their frequencies. Distances between haplotypes are proportional to the number of base differences. Colours match those utilized in Fig. 1, and haplotypes found in one area alone are the same colour. Right: Neighbour-joining tree based on 280 bp of CR1. Sampled areas are labelled with abbreviations: NF, Northern France (Oise and Nord); SWF, South-Western France (Gironde, Charente Maritime and Landes); Nor, Norway; Neth, Netherlands. Numbers below branches indicate bootstrap values; only values above 50% are shown, most of the clades are supported by low bootstrap values.

Graph of FCA

Figure 3 - Factorial Correspondence Analysis (FCA) of microsatellites data. Outliers concern individuals from Nord (4 ind.), Oise, Gironde and Finnmark (1 ind.) populations.

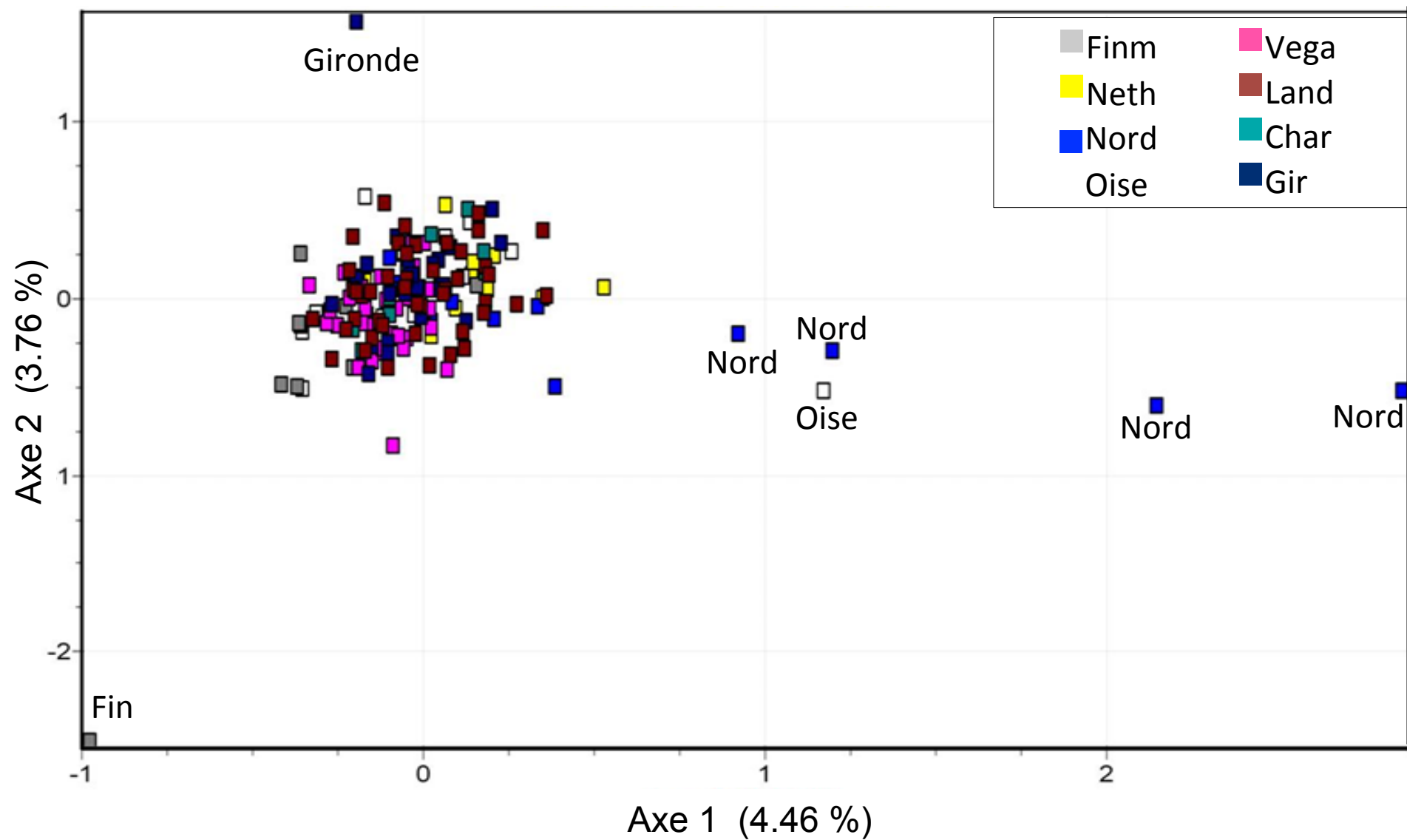



Figure 4 **(on next page)**

Graph of Structure

Figure 4 - Estimated population structure in Greylag Goose sampled populations. Each vertical line represents one individual and each colour represents a single cluster

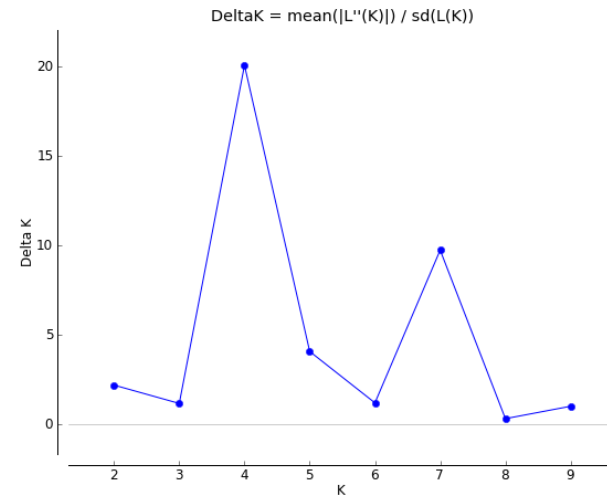
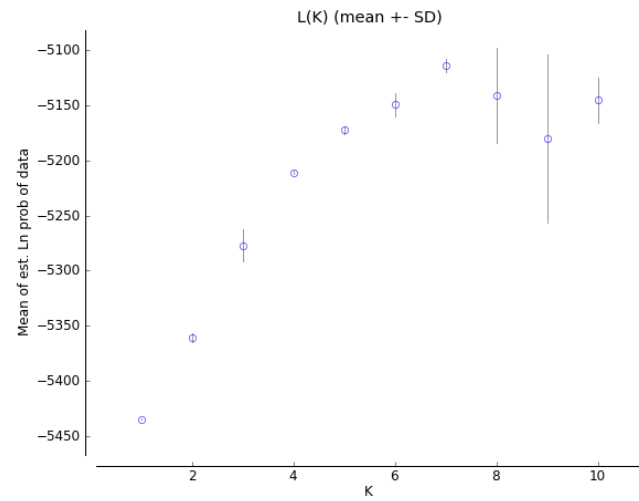
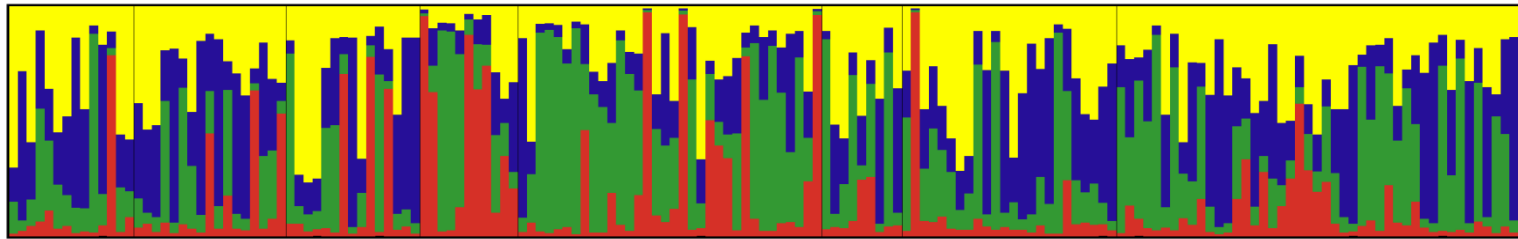
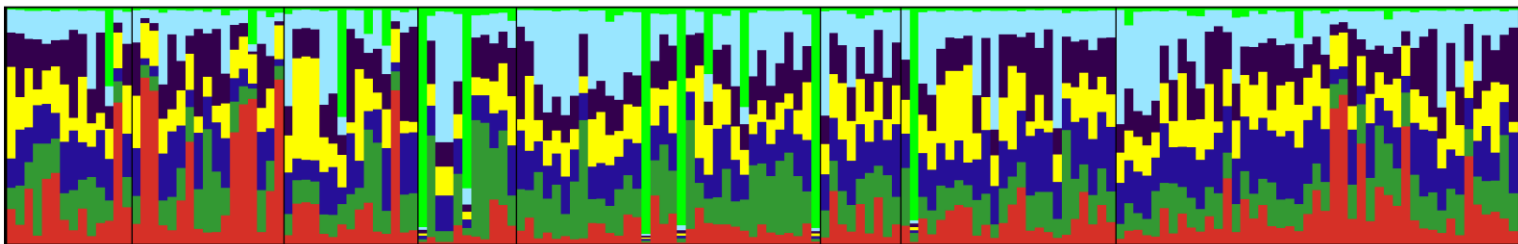



Table 1

Table 1 - Genetic variability of mtDNA CR1 D-Loop. h = haplotype diversity, π = nucleotide diversity, D = Tajima's D, FS = Fu's FS. Standard errors are showed in brackets

	Population	Season	N	Polymorphic sites	Hapl.	Private hapl.	*h*	π	K	D	FS
South-western France	Charente-Maritime	Autumn-winter	10	16	7	1	0.933 (0.06)	0.01790 (0.00254)	5.156	-0.409	-0.664
	Gironde	Autumn-winter	20	12	7	0	0.821 (0.06)	0.01519 (0.0012)	4.374	1.052	1.190
	Landes	Autumn-winter	36	19	13	5	0.798 (0.06)	0.0161 (0.0013)	4.633	0.0376	-0.360
North-France	Nord	Autumn-winter	17	10	6	0	0.833 (0.06)	0.01054 (0.0016)	3.025	0.014	0.433
	Oise	Autumn-winter	13	12	9		0.949 (0.042)	0.01309 (0.0020)	3.769	-0.103	-2.691*
Northern Europe	Norway Finnmark	Late spring	11	6	3	1	0.564 (0.13)	0.00947 (0.002)	2.727	1.31175	3.038
	Norway Vega	Late spring	23	7	3	1	0.170 (0.10)	0.0021 (0.0016)	0.609	-2.147*	0,270
	Netherlands	Autumn	14	11	8	2	0.890 (0.006)	0.01244 (0.0013)	3.571	0.541	-1.112
All Samples			144	20	24	-	0.823 (0.022)	0.01331 (0.0006)	3.819	-1.419	-4.515*

Table 2

Table 2 – Tamura Nei genetic distance assessed by mtDNA.

	SW France - Landes	SW France - Gironde	N France - Nord	N France Oise	Netherlands	Norway - Finnmark	Norway - Vega	SW France Charente M
SW France - Landes		0.004	0.004	0.004	0.004	0.005	0.006	0.005
SW France - Gironde	0.016		0.004	0.004	0.004	0.004	0.005	0.005
N France - Nord	0.015	0.014		0.004	0.004	0.004	0.004	0.005
N France Oise	0.016	0.015	0.013		0.004	0.004	0.004	0.005
Netherlands	0.015	0.014	0.013	0.013		0.004	0.004	0.005
Norway - Finnmark	0.018	0.016	0.014	0.013	0.014		0.003	0.006
Norway - Vega	0.017	0.014	0.012	0.012	0.012	0.008		0.006
SW France Charente M	0.020	0.019	0.019	0.018	0.018	0.022	0.019	

Table 3

Table 3 - Summary of genetic variation at 14 microsatellite loci. N = number of individuals; Na = No. of different alleles; Ne = No. of effective alleles; Ho = observed heterozygosity; He = expected heterozygosity.

	Population	N	Na	Ne	Ho	He	HWE (P)
South-western France	Charente-Maritime	9	8.931 (0.071)	2.468 (0.340)	0.483 (0.082)	0.512 (0.051)	<0.001
	Gironde	24	4.286 (0.633)	2.575 (0.329)	0.462 (0.063)	0.504 (0.065)	<0.001
	Landes	45	4.643 (0.684)	2.625 (0.344)	0.484 (0.048)	0.542 (0.049)	<0.001
North France	Nord	17	4.000 (0.584)	2.680 (0.332)	0.483 (0.072)	0.549 (0.056)	<0.001
	Oise	15	3.857 (0.573)	2.702 (0.403)	0.385 (0.061)	0.527 (0.063)	<0.001
North Europe	Norway Finnmark	11	3.357 (0.372)	2.137 (0.266)	0.374 (0.072)	0.433 (0.065)	<0.01
	Norway Vega	34	4.571 (0.661)	2.485 (0.312)	0.476 (0.060)	0.529 (0.048)	<0.001
	Netherlands	14	3.929 (0.549)	2.352 (0.305)	0.447 (0.063)	0.493 (0.052)	<0.001
All Samples		169	4.027 (0.200)	2.503 (0.115)	0.449 (0.023)	0.511 (0.020)	<0.001

Table 4

Table 4 – Mean estimated number of migrants between breeding and wintering sites as calculated with BayesAss (standard deviations in parentheses). Values on the diagonal (in bold) represent the estimated proportion of resident individuals in each population.

Migration from	Migration into							
	Netherlands	NF- Nord	NF- Oise	NO-Finnmark	NO-Vega	SWF-Charente Maritime	SWF-Gironde	SWF-Landes
Netherlands	**0.8904** *(0.0338)*	0.0167 *(0.0160)*	0.0151 *(0.0147)*	0.0152 *(0.0144)*	0.0160 *(0.0154)*	0.0152 *(0.0146)*	0.0157 *(0.0151)*	0.0157 *(0.0148)*
NF- Nord	0.0145 *(0.0139)*	**0.9009** *(0.0316)*	0.0135 *(0.0130)*	0.0140 *(0.0135)*	0.0144 *(0.0138)*	0.0133 *(0.0129)*	0.0149 *(0.0143)*	0.0145 *(0.0141)*
NF- Oise	0.0142 *(0.0135)*	0.0460 *(0.0238)*	**0.6820** *(0.0146)*	0.0144 *(0.0139)*	0.0138 *(0.0132)*	0.0136 *(0.0132)*	0.2023 *(0.0344)*	0.0137 *(0.0133)*
NO- Finnmark	0.0191 *(0.0183)*	0.0212 *(0.0198)*	0.0174 *(0.0165)*	**0.8619** *(0.0394)*	0.0215 *(0.0210)*	0.0177 *(0.0164)*	0.0219 *(0.0211)*	0.0194 *(0.0184)*
NO-Vega	0.0085 *(0.0081)*	0.0078 *(0.0078)*	0.0078 *(0.0078)*	0.0081 *(0.0079)*	**0.9439** *(0.0192)*	0.0079 *(0.0077)*	0.0081 *(0.0079)*	0.0081 *(0.0077)*
SWF- Charente Maritime	0.2072 *(0.0379)*	0.0175 *(0.0167)*	0.0173 *(0.0165)*	0.0176 *(0.0166)*	0.0181 *(0.0174)*	**0.6872** *(0.0192)*	0.0177 *(0.0168)*	0.0174 *(0.0166)*
SWF- Gironde	0.0114 *(0.0112)*	0.0107 *(0.0105)*	0.0111 *(0.0102)*	0.0199 *(0.0160)*	0.0109 *(0.0104)*	0.0110 *(0.0107)*	**0.9142** *(0.0270)*	0.0109 *(0.0105)*
SWF- Landes	0.0069 *(0.0069)*	0.0068 *(0.0067)*	0.0063 *(0.0063)*	0.0067 *(0.0066)*	0.0065 *(0.0064)*	0.0063 *(0.0063)*	0.0068 *(0.0067)*	**0.9538** *(0.0164)*