

RESEARCH

Open Access



# Conservation genetics of a resident population of Greylag Goose (*Anser anser*) at the southernmost limit of the species' distribution in Europe

Anastasios Bounas<sup>1</sup>, Giorgos Catsadorakis<sup>2</sup>, Annita Logotheti<sup>2</sup>, Berend Voslamber<sup>3</sup>, Antonios Magoulas<sup>4</sup> and Dimitris Tsaparis<sup>4\*</sup> 

## Abstract

**Background:** Conservation of small and isolated populations can be challenging since they are prone to loss of genetic diversity due to random genetic drift and inbreeding. Therefore, information from the assessment of genetic diversity and structure are needed for conservation programs to determine the appropriate management strategy for the populations. We investigated the levels of genetic variability in a resident Greylag Goose (*Anser anser*) population, the southernmost breeding population of the species in Europe and the sole viable population of any goose species in Greece.

**Methods:** A fragment of mtDNA Control Region and a panel of 11 microsatellite markers were used to search for any signs of genetic impoverishment and population substructure and to reveal the underlying processes through the identification of possible past demographic events.

**Results:** The population was found to be monomorphic in the amplified fragment of the mitochondrial Control Region, with all individuals sharing a single private haplotype. Analyses showed a lack of any population substructure indicating a panmictic population. Although the population seems to have experienced a strong and recent genetic bottleneck and exhibits a small effective population size, we did not find evidence of either extremely low levels of genetic diversity or inbreeding depression.

**Conclusions:** The recent demographic decline we detected and the combined influence of residency and anthropogenic factors have probably shaped the current genetic status. Our study population does not need emergency conservation actions but should be regarded as a discrete management unit. Future management strategies should focus on population and genetic monitoring and preventing further abundance declines that would increase the risk of genetic impoverishment.

**Keywords:** Conservation, Genetic variability, Inbreeding, Isolation, Management units, Microsatellites, mtDNA, Non-migratory population

\*Correspondence: tsaparis@hcmr.gr

<sup>4</sup> Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC), Hellenic Centre for Marine Research (HCMR), 71500 Heraklion, Greece  
Full list of author information is available at the end of the article



## Background

Significant demographic fluctuations and isolation can have major genetic consequences in wild populations. Partially or totally impeded connectivity through gene flow could eventually lead to increased genetic differentiation between populations (Frankham et al. 2002). However, conservation of small and isolated populations can be rather challenging as they are prone to loss of diversity due to random genetic drift and inbreeding (Allendorf and Luikart 2009). Such decreases in genetic variability compromise a species' ability to adapt to a changing environment (Amos and Balmford 2001), thus increasing its extinction risk due to stochastic events. Therefore, the assessment of genetic diversity and structure and the identification of the underlying processes that shape them are essential tasks providing valuable information towards the design and implementation of effective conservation strategies. For example, measuring connectivity and dispersal rates between populations can provide insight on the degree of their reproductive isolation and subsequently populations could be delineated as demographically independent units that should be monitored and managed separately (management units, MUs; Moritz 1994; Palsbøll et al. 2007). Thus, identifying populations or regions in need of localized conservation actions could be vital for the short-term management of the species in question.

One of the most severe threats for small, isolated populations is considered to be the decline in mean fitness (or mean phenotype) of some individuals as a consequence of systematic mating between related individuals termed as inbreeding depression (Hedrick and Kalinowski 2000). A common approach to examine the effects of inbreeding on various traits related to fitness is the testing of associations between heterozygosity of molecular markers and fitness components such as body size or breeding success (heterozygosity-fitness correlations, HFCs) on the individual level (Szulkin et al. 2010). The empirical observation of HFCs works under the assumption that genetic diversity at marker loci reflects genetic diversity at loci that affect trait variation (Grueber et al. 2008). Signs of inbreeding depression have been detected for populations of various taxa (Crnokrak and Roff 1999) and even for non-isolated populations (Höglund et al. 2002). As a result of philopatric behaviour, goose species have been previously reported the rise of HFCs and thus are in high risk of inbreeding depression (Harrison et al. 2011).

The Greylag Goose (*Anser anser*), is a waterfowl species, widely distributed through the Palearctic with the main breeding populations located in central and northern Europe (BirdLife International 2004). Most of the species populations are migratory, covering large distances to their wintering grounds in France and Spain

while some individuals reach North Africa (Fox et al. 2010). However, few non-migratory (i.e. resident) populations do exist (Scott and Rose 1996), one of them located in Prespa Lakes, in northwestern Greece (Catsadorakis 1997). The area hosts a small population of Greylag Goose consisting of 200–280 individuals that are geographically isolated, with the nearest breeding populations of the species located more than 400 km away (Catsadorakis et al. 2012). These characteristics render the Prespa population susceptible to the loss of genetic diversity and highlight the need for systematic research and special management. The population is considered to be of high conservation value since it is the sole viable population of any goose species in Greece as well as the southernmost population of this species in Europe (Catsadorakis et al. 2012).

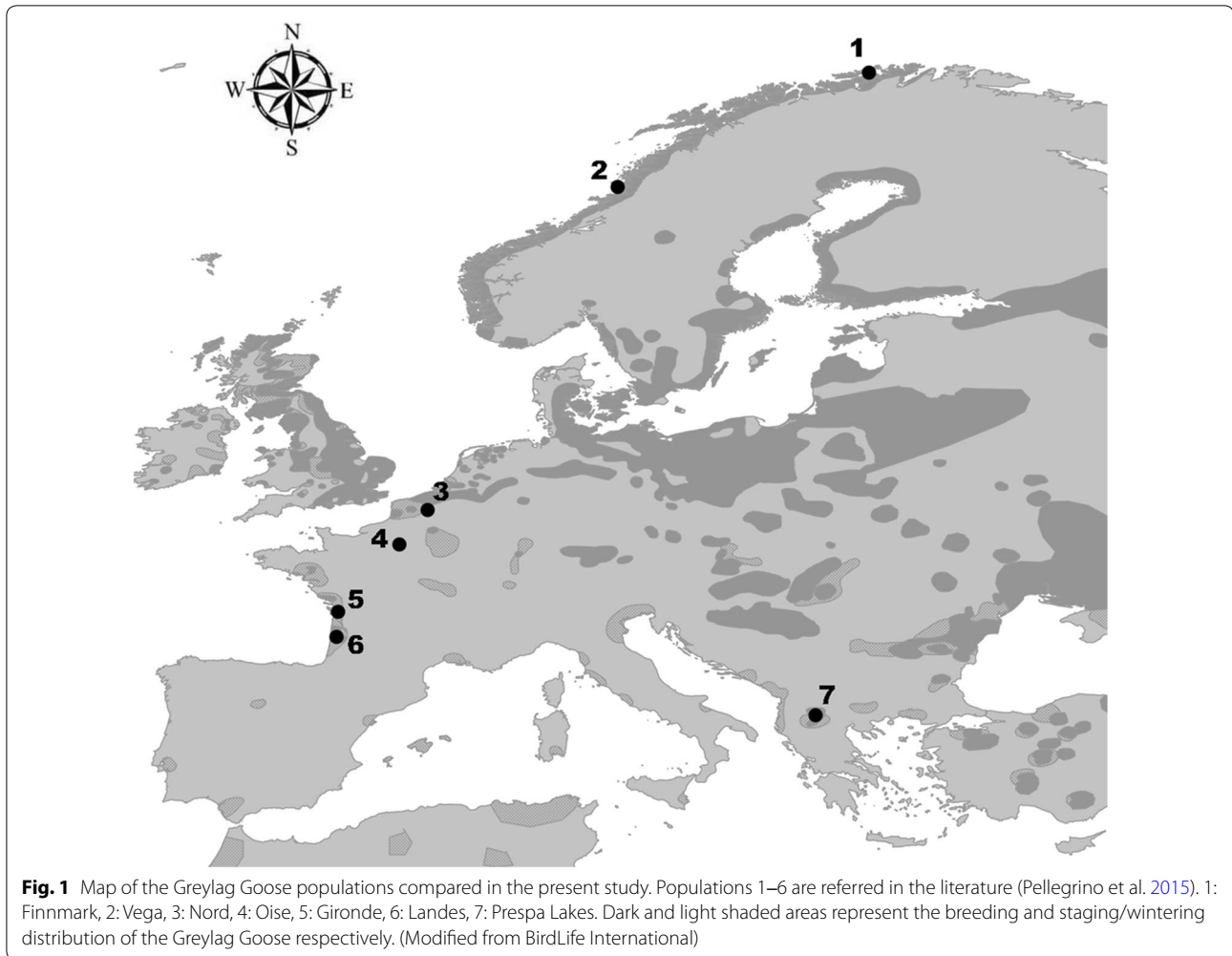
So far, a limited amount of studies have investigated the genetic aspects of Greylag Goose populations. Works by Heikkinen et al. (2015) and Pellegrino et al. (2015) examined the genetic variability of the species and its geographical distribution based on mtDNA and microsatellite markers. It seems that there is limited genetic structure and high genetic admixture of populations that can be attributed mainly to behavioural and anthropogenic factors. Admixture of individuals from distant areas and consequent formation of pair bonds in the wintering sites is a common behaviour for geese, while hybridization of wild and domesticated geese and reintroductions of Greylag Geese, are rather widespread practices in amateur goose husbandry (Heikkinen et al. 2015). However, not all goose populations are thoroughly studied, and the knowledge of genetic diversity and structure of small and isolated populations is necessary so that appropriate conservation actions can be designed.

In this study, we investigated the levels of genetic variability in the resident Greylag Goose population in Prespa Lakes using both a fragment of mtDNA Control Region and a panel of 11 microsatellite markers that have been previously used in genetic studies of the species. We aimed to search for any signs of genetic depletion, inbreeding depression or population substructure and to reveal the underlying processes through the identification of possible past demographic fluctuations. Such knowledge could be of crucial importance for in situ conservation programs to determine the appropriate management strategy for the population.

## Methods

### Sample collection

The study took place in the area of Prespa lakes, NW Greece (Fig. 1). Geese were captured using cannon nets between July and October during the years 2013–2015. On capture, birds were banded with a unique



combination of numbered neck collars and aluminum rings. For each banded bird the following measurements were taken: bill length, bill depth, tarsus length (all measured to the nearest 0.1 mm) and the length of the 9th primary (to the nearest 1 mm), according to standard procedures using calipers and a ruler (Svensson 1992). All birds were also weighed to the nearest 1 g. Sex and age of each individual were recorded in the field, and genetic material was extracted (blood and/or feathers). Blood samples were stored in blood storage cards (NucleoCards<sup>®</sup>) and feathers were stored in paper envelopes at room temperature until DNA extraction. In total, 83 samples were collected that included 52 blood samples, 30 feather samples and a single tissue sample, retrieved from an individual found dead (Additional file 1: Table S1).

#### Laboratory procedures

DNA from feathers was extracted using both the basal tip of the calamus and the blood clot from the superior

umbilicus of each sample (Horvath et al. 2005). All extractions were performed with the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's protocol. For mtDNA analyses we amplified a 394 bp fragment near the 5' end of the Control Region (CR), using primers L16642 and H411ANX (Ruokonen et al. 2000; Heikkinen et al. 2015) for 51 individuals. Each 12.5 µL PCR reaction contained 1 unit of Taq DNA polymerase (Finnzymes), 1× reaction buffer, 1.5 mM of MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.5 mM of each primer and 50 ng of DNA template. PCRs were carried out using the following profile: an initial denaturation step of 3 min at 95 °C, 30 cycles of 30 s at 94 °C, 30 s at 52 °C, 60 s at 72 °C, with a final extension step of 10 min at 72 °C. Sequencing of both strands was conducted using a Big Dye Terminator v. 3.1 Cycle Sequencing Kit (Applied Biosystems). Sequences were edited by eye in MEGA7 (Kumar et al. 2016) and aligned with CLUSTAL W (Thompson et al. 1994).

Genotypic data for all 83 samples were obtained for 11 polymorphic microsatellite loci developed specifically for

the Greylag Goose (Weiß et al. 2008; Additional file 1: Table S2). Loci were amplified in five multiplex reactions using forward 5'-fluorescent-labelled primers and the KAPA2G Fast Multiplex PCR Kit (Kapa Biosystems). Each 12.5  $\mu$ L reaction contained 2 pM of each primer and 1  $\times$  KAPA2G Mix and was carried out using the following profile: an initial denaturation step of 3 min at 95  $^{\circ}$ C, 30 cycles of 15 s at 95  $^{\circ}$ C, 30 s at 60  $^{\circ}$ C, 30 s at 72  $^{\circ}$ C, with a final extension step of 10 min at 72  $^{\circ}$ C. PCR products were separated and visualized using an ABI 3730xl capillary sequencer (Applied Biosystems) and genotypes were scored by eye with STRand v.2.4.59 (Toonen and Hughes 2001). Reliable genotypes from feather samples were obtained using the multiple tubes approach (Navidi et al. 1992) by performing three different PCR amplifications for each sample and consensus genotypes were retrieved using GIMLET (Valière 2002). Genotyping errors, due to null alleles and stuttering, were examined for all loci using Micro-Checker (Van Oosterhout et al. 2004).

### Genetic diversity

We counted the number of different CR haplotypes in the population from retrieved mtDNA sequences. To examine the phylogeographic position of the study population, we compared our results with previously published sequences from Greylag Goose (Heikkinen et al. 2015). Due to the presence of a single haplotype in all individuals it was not possible to proceed with further mtDNA diversity analyses.

Microsatellite loci were tested for deviations from Hardy–Weinberg proportions as well as for linkage disequilibrium (LD) using FSTAT 2.9.3.2 (Goudet 2002). Runs were implemented using 1000 randomizations and adjusting significance for multiple comparisons. To evaluate the presence of genetic substructure in the population, the Bayesian clustering software Structure 2.3.4 (Pritchard et al. 2000) was used to infer the number of genetically homogeneous clusters ( $K$ ) present in the dataset. We assumed the admixture ancestry model and correlated allele frequencies (Falush et al. 2003), and runs were set with a burn-in period of  $2 \times 10^5$  iterations followed by  $8 \times 10^5$  MCMC steps with 10 replicates for each  $K$  value (1–5). Structure runs were performed using Para-Structure Perl script (Lagnel 2015) at the HPC cluster of IMBBC. The most likely value of genetic clusters, was evaluated by calculating the posterior probability for each  $K$ . In addition, a factorial correspondence analysis (FCA) was performed in software GENETIX 4.03 (Belkhir 1999). The clustering of genotypes was visualized for different age classes (adult, young) and for each sex (male, female) using the package “ggplot2” (Wickham and Chang 2015) in R 3.4.3 (R Core Team 2017). The genetic variability of the population was evaluated by calculating

standard genetic diversity indices ( $A$ : number of alleles,  $A_e$ : number of effective alleles,  $H_o$ : observed and  $H_e$ : expected heterozygosity and  $F_{is}$ : inbreeding coefficient), using the program GenAlEx 6.5 (Peakall and Smouse 2012). To better interpret the above indices we compared them with those reported for six additional breeding and wintering Greylag Goose populations across Europe (Pellegrino et al. 2015; Fig. 1). Direct comparison with the population of Prespa Lakes was facilitated after recalculating the values from Pellegrino et al. (2015) only for the microsatellite loci we used in this study.

To check for signs of inbreeding depression in the study population, we tested for observations of HFCs. We used the R package “inbreedR” (Stoffel et al. 2016) to calculate  $g_2$ , an estimate of identity disequilibrium, which is a correlation in heterozygosity and/or homozygosity across loci (ID; David et al. 2007; Szulkin et al. 2010) and the effect of standardized multilocus heterozygosity (sMLH; Coltman et al. 1999) on three different traits (body size, bill shape and condition), using linear models (Additional file 1: Table S3). The overall body size for each individual was calculated by reducing three independent measurements (bill length, tarsus length, 9th primary length) to the first principal component (PC1) of a principal component analysis (PCA), according to Freeman and Jackson (1990). The first component extracted from the analysis accounted for 63.7% of total variance. Bill shape was measured by calculating the bill-index (bill depth/bill length ratio) and was used as a proxy to foraging efficiency (Gosler 1987). Finally, we regressed the mass of each individual on the PC1 (see above) and used the residual scores as a measure of condition.

### Demographic history

The contemporary effective population size ( $N_e$ ) along with 95% confidence intervals (CIs) was estimated using the bias-corrected version of the method based on linkage disequilibrium (Hill 1981; Waples 2006) as implemented in NeEstimator 2.01 (Do et al. 2014). In addition,  $N_e$  was estimated using a temporal method (Nei and Tajima 1981) implemented in the same software, using two sample cohorts collected in 2013 and 2015. Signs of recent changes in the effective population size were evaluated using the software Bottleneck 1.2 (Piry et al. 1999). A test for heterozygosity excess was performed using the Wilcoxon signed rank test running 1000 iterations and using the two-phase model (TPM; Di Rienzo et al. 1994). As the microsatellites we used consist of dinucleotide perfect repeats that may tend toward the infinite allele model (IAM; Cornuet and Luikart 1996), we fixed the proportions of the TPM in favour of the IAM (Cristescu et al. 2010) including 20% of the stepwise mutation model (SMM) and 80% of IAM. To complement the explicit test

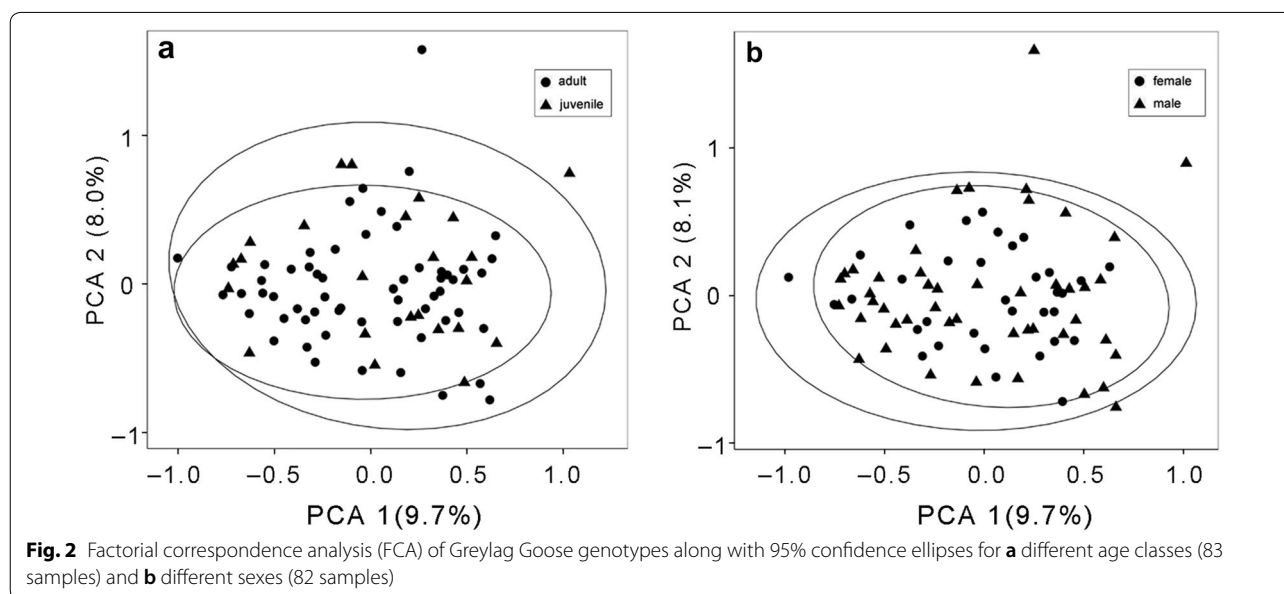
for population bottleneck, we calculated  $M$ , a ratio based on the number of alleles to the allelic size range (Garza and Williamson 2001) in Arlequin 3.5.1.3 (Excoffier and Lischer 2010).  $M$  is expected to be smaller in populations that have suffered a past decline than in populations that are in mutation-drift equilibrium. Finally, we explored the demographic history of the population using the program Migraine 0.4 (<http://kimura.univ-montp2.fr/~roussel/Migraine.htm>) via the newly developed model of a single population with past variations in population size (Leblois et al. 2014). Migraine was used to estimate the parameters  $\theta = 4N\mu$  and  $\theta_{anc} = 4N_{anc}\mu$ , where  $N$  and  $N_{anc}$  are the current and the ancestral effective population size (respectively,  $\mu$  is the mutation rate per locus per generation) and the time ( $D$ ) when the demographic change started ( $D = T/4N$ , where  $T$  is the time measured in generations). The magnitude of the putative demographic events was evaluated by calculating the value of the parameter  $N_{ratio} = N/N_{anc}$ , which is expected to be lower than one in the case of population contraction and higher than one in the case of population expansion. When the 95% CIs of the  $N_{ratio}$  point estimate did not include a value equal to one,  $N_{ratio}$  was considered significant (Leblois et al. 2014). Migraine runs were implemented with 2400 points, 2000–20,000 trees and 6 iterations.

## Results

### Genetic diversity

All 51 individuals sequenced were found to be monomorphic in the amplified fragment of the CR, providing a single haplotype. After comparison with published studies, our sequence was identified as the haplotype E7 that

was only found in birds from Prespa Lakes in Heikkinen et al. (2015). All microsatellite markers were found to be polymorphic in the population and the number of alleles per locus ranged from two (loci Aph19, Hhi $\mu$ 1b) to nine (locus Ans02; Additional file 1: Table S2). Two of the loci (Ans07, Ans18) were omitted from further analyses due to the presence of null alleles. The remaining nine loci did not deviate from Hardy–Weinberg proportions (adjusted  $p$ -value = 0.006) and no LD was detected between any of them. The Bayesian clustering method implemented in Structure proposed the presence of a single genetic cluster ( $K = 1$ ) in the dataset, suggesting the absence of any population substructure (Additional file 1: Table S4). Likewise, the FCA did not show any distinct clustering pattern of individuals, neither between age classes nor between sexes (Fig. 2), further enhancing the hypothesis of a rather panmictic population. However, two individuals appeared as outliers: both were males but belonged to different age classes. The population showed medium to low levels of genetic diversity, with a mean  $H_o$  of 0.45,  $F_{is}$  of 0.15 and the number of effective alleles ( $A_e$ ) estimated at 2.3. Genetic diversity indices for each locus are presented in Additional file 1: Table S5. Comparison of these estimates with published values of diversity for other European populations (Table 1), indicates that Prespa population exhibits similar levels of diversity with other breeding populations, but lower levels from populations that consist of individuals gathering at wintering sites (Table 1). For all individuals, standardized multilocus heterozygosity ranged from 0.25 to 1.72 whereas all three fitness traits (especially body size and condition) showed high variation (Additional file 1: Table S3). We did not





**Table 1 Estimates of genetic diversity of the Greylag Goose populations**

Code	Country/location	Site	<i>n</i>	<i>A</i>	<i>A<sub>e</sub></i>	<i>H<sub>o</sub></i>	<i>H<sub>e</sub></i>	<i>F<sub>is</sub></i>	
1	Norway	Finnmark	Breeding	11	4.0 ± 0.4	2.5 ± 0.3	0.481 ± 0.085	0.530 ± 0.072	0.104 ± 0.072
2		Vega	Breeding	34	5.0 ± 0.9	2.9 ± 0.4	0.595 ± 0.055	0.590 ± 0.054	-0.038 ± 0.078
3	France North	Nord	Wintering	17	4.9 ± 0.7	3.2 ± 0.4	0.599 ± 0.075	0.641 ± 0.046	0.084 ± 0.080
4		Oise	Wintering	15	4.3 ± 0.8	3.2 ± 0.6	0.437 ± 0.082	0.589 ± 0.079	0.246 ± 0.078
5	France South West	Gironde	Wintering	24	5.1 ± 0.8	3.2 ± 0.4	0.533 ± 0.062	0.626 ± 0.056	0.140 ± 0.062
6		Landes	Wintering	45	5.4 ± 0.9	3.2 ± 0.4	0.563 ± 0.050	0.628 ± 0.054	0.090 ± 0.047
7	Greece	Prespa	Breeding + wintering (resident population)	83	4.6 ± 0.9	2.3 ± 0.3	0.451 ± 0.066	0.528 ± 0.049	0.153 ± 0.087

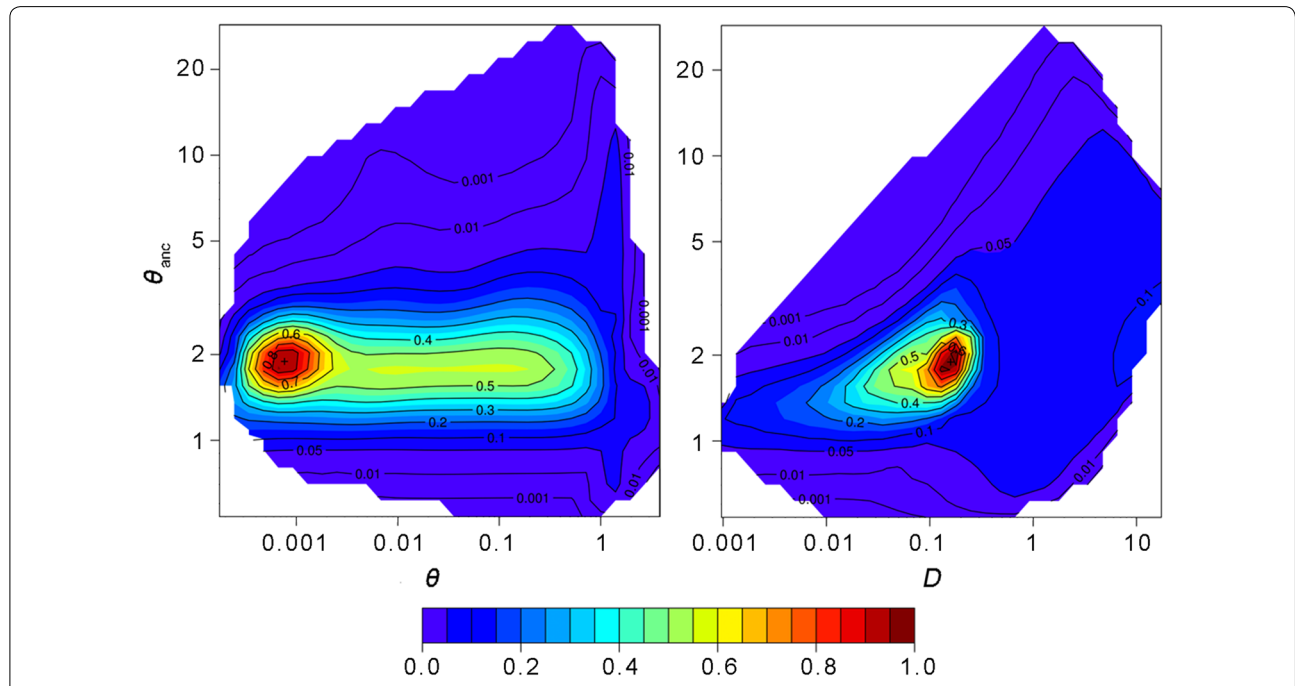
Values for populations 1–6 were recalculated from Pellegrino et al. (2015), only for the loci we used in this study. Number of genotyped individuals (*n*) and the type of site they were collected from, number of alleles per locus (*A*), number of effective alleles (*A<sub>e</sub>*), observed (*H<sub>o</sub>*) and expected (*H<sub>e</sub>*) heterozygosity and inbreeding coefficient (*F<sub>is</sub>*). Values are presented as mean ± standard error

find any evidence of identity disequilibrium in the population as *g<sub>2</sub>* was not significant (0.02, 95% CIs: -0.01 to 0.06, *p* = 0.12). As expected, since there is no variance in inbreeding, no HFCs emerged in our dataset: sMLH had no effect on body size (*p* = 0.79), condition (*p* = 0.06) or foraging efficiency (*p* = 0.26).

**Demographic history**

We found a low contemporary effective population size, estimated at 42.7 individuals (95% CIs: 26.9–75) based on the bias corrected linkage disequilibrium method, and 40.1 individuals (95% CIs: 23.3–77.1) based on the

temporal method. In addition, the Wilcoxon signed rank test conducted in Bottleneck software, detected signs of bottleneck for our study population (*p* = 0.006). *M* ratio value was 0.51, lower than the threshold value of 0.68 which according to Garza and Williamson (2001) suggests that the population has suffered a past bottleneck event. Although a very promising method, Migraine did not produce interpretable values but for few parameters and CIs were very broad (Fig. 3; Additional file 1: Table S6). Such unsatisfactory CI coverage is deemed to result when the population modeled has experienced a very recent and strong past contraction (Leblois et al.



**Fig. 3** Two-dimensional plots of likelihood ratio profiles for pairs of parameters inferred by Migraine. Ancestral and actual effective population size ( $\theta_{anc}$ ,  $\theta$ ), and timing of the demographic history events (*D*) are on the log scale. Point estimates and the 95% confidence intervals for all parameters are given in Additional file 1: Table S6

2014). This is supported by the  $N_{\text{ratio}}$  estimate, that was found to be extremely low and significantly lower than 1 (0.0005, 95% CIs:  $10^{-4}$ –0.85).

## Discussion

Our analysis showed an apparent lack of mtDNA diversity in the Greylag Goose population of Prespa Lakes in contrast to other populations across the western Palearctic that show significant levels of diversity (Heikkinen et al. 2015). Due to the large number of typed individuals ( $n=51$ ) it seems that the population is practically monomorphic in the CR fragment we examined. The small size and isolation of the population can justify at some level these findings, as the fixation of a single haplotype in the population could be attributed to the past demographic decline we detected coupled with reduced gene flow levels from other populations. Perhaps the most intriguing result regarding this single haplotype, is that it is private to the population of Prespa. Such finding further supports the isolation of the population while we could even consider the scenario of a relic population that settled in the area after the last glacial maximum (Hewitt 1999).

On the other hand, our results from microsatellite markers showed that the population retains similar genetic variability levels to those exhibited in other breeding populations of the species from Norway (Table 1). Despite the populations' small size it does not seem to experience extremely low levels of genetic diversity or apparent effects of inbreeding indicated by the absence of significant HFCs in the present study. Estimates of HFCs are expected to have low precision in cases of weak inbreeding (Szulkin et al. 2010) which seems to be the case in our study population. However, HFC studies could be also limited by sample size and heterozygosity estimates using microsatellite markers. Even if our sample size was adequate to provide robust results on the genetic diversity and structure of the population, it could produce large fluctuations of the effect sizes (Chapman et al. 2009). Additionally, genetic diversity estimated with small panel of microsatellite loci may not necessarily reflect genome-wide heterozygosity and is deemed to weakly reflect inbreeding depression (Slate et al. 2004). However, it has been proposed that it may actually be a better predictor of the phenotype (Forstmeier et al. 2012).

Avoidance of consanguineous mating could be facilitated by the dispersal of juveniles to other populations (Jonker et al. 2013), as juvenile geese are suggested to often embark on exploratory trips after the provision of parental care (Baker 1978). However, additional samples from the nearest breeding populations are needed to shed light on this hypothesis. A more plausible explanation though, comes from the understanding of gene flow between goose populations, with wintering sites

playing a major role since this is where the formation of pairs takes place (Rohwer and Anderson 1988). Studies on other Anseriformes such as the King Eider (*Somateria spectabilis*), have shown that strong site fidelity to wintering areas and pair formation at wintering quarters are actually the main factors defining the species' population structure (Pearce et al. 2004). Wintering sites host individuals from different populations thus it is expected that wintering populations will exhibit higher levels of genetic diversity, as also shown in our results. In the case of our study population, which is resident, some males from other populations could join the resident birds in the winter, form pairs, and due to the strong female natal philopatry in geese (van der Jeugd et al. 2002), subsequently breed there. Such movements could therefore facilitate gene flow at some level but were not identified from the analysis of the maternally inherited mtDNA.

Residency could be a main factor that shaped the populations' genetic structure (Willoughby et al. 2017). The migratory behaviour in geese is culturally transmitted through generations, with offspring travelling where their parents travel (van der Jeugd et al. 2002). Such traditions along with philopatry can limit gene flow between populations thus increasing their divergence rates (Friesen et al. 2007). A change in migratory traditions has been suggested to facilitate the emergence of new populations affecting the genetic structure of a species (Jonker et al. 2013). Decrease in migratory tendency has been suggested to reflect the response to alterations of the environment, such as changes in climate conditions or habitat quality (Møller et al. 2010; Newton 2010). However, the transition from migrancy to residency in the population of Prespa due to such factors, could only partially explain the strong and recent bottleneck experienced by the population. The influence of anthropogenic factors such as human induced mortality (Palacín et al. 2017) may have further shaped the current genetic patterns in the population.

Based on our findings, the Greylag Goose population of Prespa Lakes should be regarded as discrete management unit (MU sensu; Moritz 1994) and actions that focus in maintaining its genetic diversity should be implemented. Although we did not detect any signs of genetic erosion, the reduced number of effective alleles and the small effective population size could lead to inbreeding, especially if further population decline occurs. We should note that  $N_e$  estimates for our study population appear to be robust as the  $N_e/N_{\text{census}}$  ratio (14%) is within the range reported for vertebrates (Frankham 1995; Free-lund 2005). If we consider an effective population size of 50 as a threshold for avoidance of inbreeding depression (Frankham et al. 2002), it seems that without special conservation concern, the population could experience

signs of genetic impoverishment in the future that could further compromise the species' evolutionary potential. Towards the design of effective management strategies, genetic monitoring should be performed after few generations and additional data on the demographic parameters of the population (i.e. survival, breeding success) should be evaluated. Finally the extent of the population's isolation should be further investigated through the analysis of additional samples from the nearest breeding populations of the species.

## Conclusions

In this study we investigated the genetic status of a resident population of Greylag Goose (*Anser anser*). The population showed fixation of a single haplotype that could be attributed to the past strong demographic change we detected coupled with restricted gene flow from other populations. In addition, the fact that the specific haplotype was private further supports the isolation of the population. Analyses of multilocus genotypes showed a lack of any population substructure and small effective population size albeit no evidence of severe genetic impoverishment. We argue that isolation along with the influence of residency and anthropogenic factors may have shaped the current genetic patterns of the population. It is important that the population is treated as a separate management unit and conservation actions should be designed to avoid further abundance declines and future loss of genetic diversity that could compromise the species' evolutionary potential.

## Additional file

**Additional file 1: Table S1.** Individual characteristics and samples used in each analysis (mtDNA/msats). **Table S2.** Microsatellite loci and primers used in the present study. **Table S3.** Data used to examine HFCs for all individuals. **Table S4.** Summarizing table of the clustering result from the Bayesian analysis of population structure implemented in STRUCTURE software. **Table S5.** Estimates of genetic diversity of the Greylag Goose for each locus. Number of genotyped individuals ( $n$ ), number of alleles per locus ( $A$ ), number of effective alleles ( $A_e$ ), observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity and inbreeding coefficient ( $F_{is}$ ). **Table S6.** Output parameters of the Migraine analysis representing indicators of the ancestral and actual scaled effective population size ( $\theta_{anc}$ ,  $\theta$ ) and timing of the demographic history events ( $D$ ).

### Authors' contributions

GC, AL, AM and DT conceived and designed the project. GC, AL and BV collected the samples and data. AB and DT performed the lab work and analyzed the data. AB and DT drafted the manuscript with significant input from all other authors. All the authors read and approved the final manuscript.

### Author details

<sup>1</sup> Molecular Ecology and Conservation Genetics Lab, Department of Biological Applications and Technology, University of Ioannina, 45110 Ioannina, Greece.

<sup>2</sup> Society for the Protection of Prespa, Lemos Prespa, 53077 Agios Germanos, Greece. <sup>3</sup> Sovon Dutch Centre for Field Ornithology, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands. <sup>4</sup> Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC), Hellenic Centre for Marine Research (HCMR), 71500 Heraklion, Greece.

### Acknowledgements

We are especially indebted to Gerard Müskens (Wageningen Environmental Research) who was in charge of the catching efforts and took the samples. Special thanks need to go to the enthusiastic volunteers and friends, Jan Vegelin, Loes van den Bremer, and the late Hennie van den Brink who helped in the catching. The staff of the SPP Olga Alexandrou, Irene Koutseri, Myrsini Malakou, Haris Nikolaou, Lazaros Nikolaou also helped during the catching efforts. The support of the Management Body of the Prespa National Park and the help of local inhabitants are also acknowledged.

### Competing interests

The authors declare that they have no competing interests.

### Availability of data and materials

Data needed to replicate the analysis are included in this article as supporting information files. The raw genotype dataset analysed during the current study is available from the corresponding author upon request.

### Consent for publication

Not applicable.

### Ethics approval and consent to participate

This study was carried out in accordance with the European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes of the Council of Europe. Sampling was approved by the Hellenic Ministry of Environment and Energy (Permit Nos.: 68849/936 of 26/03/2012, 106322/517 of 13/03/2014, 128002/2280 of 04/08/2015).

### Funding

All the work of the Society for the Protection of Prespa is supported financially by the MAVA Foundation. This particular research was also partly funded through a grant by the Green Fund of the Hellenic Ministry of Environment and Energy, a grant by the Bodosaki Foundation and two grants by two groups of Dutch naturalists, led by the biologist Gerrit Jansen ("Nature Workshops Gerrit Jansen").

Received: 3 May 2018 Accepted: 13 December 2018

Published online: 21 December 2018

## References

- Allendorf FW, Luikart G. Conservation and the genetics of populations. New York: Wiley; 2009.
- Amos W, Balmford A. When does conservation genetics matter? *Heredity*. 2001;87:257–65.
- Baker RR. The evolutionary ecology of animal migration. New York: Holmes & Meier Publishers; 1978.
- Belkhir K. GENETIX, logiciel sous Windows™ pour la génétique des populations. 1999. <http://www.univmontp2fr/~genetix>. Accessed 24 Mar 2018.
- BirdLife International. Birds in Europe: population estimates, trends and conservation status. Cambridge: BirdLife Conservation Series; 2004.
- Catsadorakis G. The importance of Prespa National Park for breeding and wintering birds. *Hydrobiologia*. 1997;351:157–74.
- Catsadorakis G, Voslamber B, Logotheti A. First Greylag Geese *Anser anser rubrirostris* ringed in Greece. *Goose Bull*. 2012;15:28–31.
- Chapman J, Nakagawa S, Coltman D, Slate J, Sheldon B. A quantitative review of heterozygosity–fitness correlations in animal populations. *Mol Ecol*. 2009;18:2746–65.
- Coltman D, Pilkington J, Smith J, Pemberton J. Parasite mediated selection against inbred Soay sheep in a free-living island population. *Evolution*. 1999;53:1259–67.
- Cornuet JM, Luikart G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*. 1996;144:2001–14.



- Cristescu R, Sherwin WB, Handasyde K, Cahill V, Cooper DW. Detecting bottlenecks using BOTTLENECK 1.2.02 in wild populations: the importance of the microsatellite structure. *Conserv Genet*. 2010;11:1043–9.
- Crnokrak P, Roff DA. Inbreeding depression in the wild. *Heredity*. 1999;83:260–70.
- David P, Pujol B, Viard F, Castella V, Goudet J. Reliable selfing rate estimates from imperfect population genetic data. *Mol Ecol*. 2007;16:2474–87.
- Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB. Mutational processes of simple sequence repeat loci in human populations. *Proc Natl Acad Sci USA*. 1994;91:3166–70.
- Do C, Waples RS, Peel D, Macbeth G, Tillett BJ, Ovenden JR. NeEstimator v2: reimplementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Mol Ecol Res*. 2014;14:209–14.
- Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res*. 2010;10:564–7.
- Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*. 2003;164:1567–87.
- Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B. Heterozygosity–fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Mol Ecol*. 2012;21:3237–49.
- Fox AD, Ebbinge BS, Mitchell C, Heinicke T, Aarvak T, Colhoun K, Clausen P, Dereliev S, Faragó S, Koffijberg K. Current estimates of goose population sizes in western Europe, a gap analysis and assessment of trends. *Ornis Svec*. 2010;20:115–27.
- Frankham R. Effective population size/adult population size ratios in wildlife: a review. *Genet Res*. 1995;66:95–107.
- Frankham R, Briscoe DA, Ballou JD. Introduction to conservation genetics. Cambridge: Cambridge University Press; 2002.
- Freeland JR. Molecular ecology. New York: Wiley; 2005.
- Freeman S, Jackson WM. Univariate metrics are not adequate to measure avian body size. *Auk*. 1990;107:69–74.
- Friesen V, Burg T, McCoy K. Mechanisms of population differentiation in seabirds. *Mol Ecol*. 2007;16:1765–85.
- Garza JC, Williamson EG. Detection of reduction in population size using data from microsatellite loci. *Mol Ecol*. 2001;10:305–18.
- Gosler A. Pattern and process in the bill morphology of the Great Tit *Parus major*. *Ibis*. 1987;129:451–76.
- Goudet J. FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9.3.2. 2002. <http://www2.unil.ch/popgen/softwares/fstat.htm>. Accessed 20 Aug 2012.
- Grueber CE, Wallis GP, Jamieson IG. Heterozygosity–fitness correlations and their relevance to studies on inbreeding depression in threatened species. *Mol Ecol*. 2008;17:3978–84.
- Harrison XA, Bearhop S, Inger R, Colhoun K, Gudmundsson GA, Hodgson D, McElwaine G, Tregenza T. Heterozygosity–fitness correlations in a migratory bird: an analysis of inbreeding and single-locus effects. *Mol Ecol*. 2011;20:4786–95.
- Hedrick PW, Kalinowski ST. Inbreeding depression in conservation biology. *Annu Rev Ecol Evol Syst*. 2000;31:139–62.
- Heikkinen M, Ruokonen M, Alexander M, Aspi J, Pyhäjärvi T, Searle J. Relationship between wild greylag and European domestic geese based on mitochondrial DNA. *Anim Genet*. 2015;46:485–97.
- Hewitt GM. Post-glacial re-colonization of European biota. *Biol J Linn Soc*. 1999;68:87–112.
- Hill WG. Estimation of effective population size from data on linkage disequilibrium. *Genet Res*. 1981;38:209–16.
- Höglund J, Piertney SB, Alatalo RV, Lindell J, Lundberg A, Rintamäki PT. Inbreeding depression and male fitness in black grouse. *Proc R Soc B*. 2002;269:711–5.
- Horvath MB, Martinez-Cruz B, Negro JJ, Kalmar L, Godoy JA. An overlooked DNA source for non-invasive genetic analysis in birds. *J Avian Biol*. 2005;36:84–8.
- Jonker R, Kraus RH, Zhang Q, Hooff P, Larsson K, Jeugd H, Kurvers R, Wieren S, Loonen M, Crooijmans R. Genetic consequences of breaking migratory traditions in barnacle geese *Branta leucopsis*. *Mol Ecol*. 2013;22:5835–47.
- Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33:1870–4.
- Lagnel J. ParaStructure. 2015. <https://github.com/jacqueslagnel/parastucture>. Accessed 22 Mar 2018.
- Leblois R, Pudlo P, Néron J, Bertaux F, Reddy Beeravolu C, Vitalis R, Rousset F. Maximum-likelihood inference of population size contractions from microsatellite data. *Mol Biol Evol*. 2014;31:2805–23.
- Møller AP, Fiedler W, Berthold P. Effects of climate change on birds. Oxford: Oxford University Press; 2010.
- Moritz C. Defining 'evolutionarily significant units' for conservation. *Trends Ecol Evol*. 1994;9:373–5.
- Navidi W, Arnheim N, Waterman M. A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *Am J Hum Genet*. 1992;50:347.
- Nei M, Tajima F. Genetic drift and estimation of effective population size. *Genetics*. 1981;98:625–40.
- Newton I. The migration ecology of birds. London: Academic Press; 2010.
- Palacín C, Alonso JC, Martín CA, Alonso JA. Changes in bird-migration patterns associated with human-induced mortality. *Conserv Biol*. 2017;31:106–15.
- Palsbøll PJ, Berube M, Allendorf FW. Identification of management units using population genetic data. *Trends Ecol Evol*. 2007;22:11–6.
- Peakall R, Smouse PE. GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. 2012;28:2537–9.
- Pearce JM, Talbot SL, Pierson BJ, Petersen MR, Scribner KT, Dickson DL, Mosbech A. Lack of spatial genetic structure among nesting and wintering King Eiders. *Condor*. 2004;106:229–40.
- Pellegrino I, Cucco M, Follestad A, Boos M. Lack of genetic structure in greylag goose (*Anser anser*) populations along the European Atlantic flyway. *PeerJ*. 2015;3:e1161.
- Piry S, Luikart G, Cornuet J-M. Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *J Hered*. 1999;90:502–3.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–59.
- R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2017.
- Rohwer FC, Anderson MG. Female-biased philopatry, monogamy, and the timing of pair formation in migratory waterfowl. In: Johnston E, Richard F, editors. Current ornithology. Boston: Springer; 1988. p. 187–221.
- Ruokonen M, Kvist L, Lumme J. Close relatedness between mitochondrial DNA from seven Anser goose species. *J Evol Biol*. 2000;13:532–40.
- Scott DA, Rose PM. Atlas of Anatidae populations in Africa and western Eurasia. Ede: Wetlands International; 1996.
- Slate J, David P, Dodds K, Veenvliet B, Glass B, Broad T, McEwan J. Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity*. 2004;93:255–65.
- Stoffel MA, Esser M, Kardos M, Humble E, Nichols H, David P, Hoffman JI. InbreedR: an R package for the analysis of inbreeding based on genetic markers. *Methods Ecol Evol*. 2016;7:1331–9.
- Svensson L. Identification guide to European passerines. Thetford: British Trust for Ornithology; 1992.
- Szulkin M, Bierne N, David P. Heterozygosity–fitness correlations: a time for reappraisal. *Evolution*. 2010;64:1202–17.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22:4673–80.
- Toonen RJ, Hughes S. Increased throughput for fragment analysis on an ABI Prism® 377 automated sequencer using a membrane comb and STRand software. *Biotechniques*. 2001;31:1320–5.
- Valière N. GIMLET: a computer program for analysing genetic individual identification data. *Mol Ecol Notes*. 2002;2:377–9.
- van der Jeugd HP, van der Veen IT, Larsson K. Kin clustering in barnacle geese: familiarity or phenotype matching? *Behav Ecol*. 2002;13:786–90.
- Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes*. 2004;4:535–8.
- Waples RS. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conserv Genet*. 2006;7:167.
- Weiß BM, Poggemann K, Olek K, Foerster K, Hirschenhauser K. Isolation and characterization of microsatellite marker loci in the greylag goose (*Anser anser*). *Mol Ecol Res*. 2008;8:1411–3.

Wickham H, Chang W. Ggplot2: an implementation of the grammar of graphics. 2015. <http://CRAN.R-project.org/package=ggplot2>. Accessed 30 Mar 2018.

Willoughby JR, Sundaram M, Wijayawardena BK, Lamb MC, Kimble SJ, Ji Y, Fernandez NB, Antonides JD, Marra NJ, Dewoody JA. Biome and migratory

behaviour significantly influence vertebrate genetic diversity. *Biol J Linn Soc.* 2017;121:446–57.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

