MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 350–368 doi: 10.1111/j.1365-294X.2011.05400.x

Multilocus coalescent analysis of haemoglobin differentiation between low- and high-altitude populations of crested ducks (*Lophonetta specularioides*)

MARIANA BULGARELLA,*[1] JEFFREY L. PETERS,† CECILIA KOPUCHIAN,‡ THOMAS VALQUI,§¶ ROBERT E. WILSON* and KEVIN G. McCRACKEN*

*Department of Biology and Wildlife, Institute of Arctic Biology, and University of Alaska Museum, University of Alaska Fairbanks, Fairbanks, AK 99775, USA, †Department of Biological Sciences, Wright State University, Dayton, OH 45435, USA, ‡División Ornitología, Museo Argentino de Ciencias Naturales Bernardino Rivadavia, Buenos Aires, Argentina, §CORBIDI, Santa Rita 117, Office 202, Lima, Peru, ¶Museum of Natural Science, LSU, Baton Rouge, LA 70801, USA

Abstract

Hypoxia is a key factor determining survival, and haemoglobins are targets of selection in species native to high-altitude regions. We studied population genetic structure and evaluated evidence for local adaptation in the crested duck (*Lophonetta specularioides*). Differentiation, gene flow and time since divergence between highland and lowland populations were assessed for three haemoglobin genes (α^A, α^D, β^A) and compared to seven reference loci (six autosomal introns and mtDNA). Four derived amino acid replacements were found in the globin genes that had elevated Φ_{ST} values between the Andean highlands and Patagonian lowlands. A single β^A-globin polymorphism at a site known to influence O_2 affinity was fixed for different alleles in the two populations, whereas three α^A- and α^D-globin polymorphisms exhibited high heterozygosity in the highlands but not in the lowlands. Coalescent analyses supported restricted gene flow for haemoglobin alleles and mitochondrial DNA but nonzero gene flow for the introns. Simulating genetic data under a drift–migration model of selective neutrality, the β^A-globin fell outside the 95% confidence limit of simulated data, suggesting that directional selection is maintaining different variants in the contrasting elevational environments, thereby restricting migration of β^A-globin alleles. The α^A- and α^D-globins, by contrast, did not differ from the simulated values, suggesting that variants in these genes are either selectively neutral, or that the effects of selection could not be differentiated from background levels of population structure and linkage disequilibrium. This study illustrates the combined effects of selection and population history on inferring levels of population divergence for a species distributed across an altitudinal gradient in which selection for hypoxia resistance has likely played an important role.

Keywords: Andes, gene flow, haemoglobin, hypoxia, Patagonia, South America, waterfowl

Received 28 July 2011; revision received 14 October 2011; accepted 21 October 2011

Introduction

Geographically widespread species often inhabit heterogeneous environments that can exert contrasting selective pressures in different populations. High-altitude environments in particular present a number of physiological challenges, including colder temperatures, increased desiccation and higher atmospheric radiation (Monge & León-Velarde 1991; Rezende *et al.* 2005). One of the most noticeable high-altitude effects is hypoxia. The partial pressure of oxygen (P_{O_2}) decreases with elevation by approximately 10% per 1000 m, and at elevations such as 4000 m in the Andean or Himalayan

Correspondence: Mariana Bulgarella, Fax: +64 6 350 5623; E-mail: M.Bulgarella@massey.ac.nz

[1]Current address: Ecology Group (PN 624), Institute of Natural Resources, Massey University, Private Bag 11-222, Palmerston North, 4442, New Zealand.

© 2011 Blackwell Publishing Ltd

plateaus, the P_{O_2} of inspired air is 60% of that at sea level (Beall 2007). This reduced P_{O_2} can make it more difficult to transport a sufficient oxygen supply to respiring tissues, which can exert contrasting selective pressures at different altitudes.

Animals vary in their capabilities to withstand hypobaric hypoxia resulting from the decrease in barometric pressure with increasing altitude. Among vertebrates, birds are particularly well adapted for maintaining oxygen supplies to the tissues. For example, house sparrows (*Passer domesticus*) remain alert and exhibit normal behaviour at atmospheric pressures equivalent to 6100 m, whereas mice become comatose (Tucker 1968). In addition, numerous species of birds undertake high-altitude flights or migrations. A Rüppell's griffon (*Gyps rueppellii*) collided with a jetliner at 11 300 m (Laybourne 1974), and bar-headed geese (*Anser indicus*) migrate over the Himalayas up to 9000 m (Swan 1961, 1970; Scott & Milsom 2007; Hawkes *et al.* 2011). Flying at high altitudes, where P_{O_2} is reduced, is particularly demanding, because O_2 consumption during flight can increase as much as 20-fold (Bernstein 1989).

How birds and other organisms have adapted to high-altitude regions continues to be an important question for evolutionary biologists and physiologists (Scott 2011). All species of birds possess crosscurrent arrangements of gas and blood flow that make the avian lung more efficient than the alveolar lungs of mammals (Piiper 1989; Scheid 1979). Bird muscle also possesses higher venous oxygen tension and increased densities of capillaries, mitochondria and myoglobin than mammalian muscle (Grubb 1981; Faraci 1991). In bar-headed geese, mitochondria and capillary spacing reduces the intracellular O_2 diffusion distance in flight muscle and results in increased aerobic capacity (Scott *et al.* 2009). Vertebrates residing at high elevations possess other traits such as higher red blood cell counts, larger lung volume and blunted hypoxic ventilatory response (Beall 2003). Furthermore, biochemical adaptations in O_2-binding proteins such as haemoglobin (Hb) may facilitate either O_2 uptake at the lungs or O_2 release at the tissues (Turek *et al.* 1973, 1978; Hardison 2001). Examples of physiological and biochemical changes in haemoglobin at high elevations include increases in Hb concentration in the blood, changes in Hb-O_2 affinity and changes in allosteric properties (Hochachka & Somero 2002; Storz *et al.* 2010).

Several recent studies have shown that genetically based adaptations that increase the O_2 affinity of haemoglobin play an important role in physiological adaptation to high-altitude hypoxia (Jessen *et al.* 1991; Weber 2007; Storz *et al.* 2010). Most vertebrate haemoglobins are tetrameric proteins composed of two α-polypeptide subunits and two homologous β subunits; each subunit binds O_2 reversibly at the haeme group (Perutz 1983, 1989). The α and β subunits are encoded by two different sets of duplicated genes, three of which are expressed at high levels in adult birds (α^D, α^A, β^A; Rowley & Ratcliffe 1988; Bulgarella *et al.* 2009). Adult birds thus possess two different isoforms, a major HbA (α^A/β^A) haemoglobin and a minor HbD (α^D/β^A) haemoglobin that share the same β-chain but differ in their O_2-binding properties owing to differences in the coding sequences of the α^A and α^D subunits (Weber *et al.* 1993; Weber & Fago 2004). In the Muscovy duck (*Cairina moschata*), the two adult α-globins are separated by 2.2 kb and encoded by the same DNA strand (Niessing *et al.* 1982). The HbD isoform has been shown to have higher O_2 affinity than HbA (Cirotto & Geraci 1975; Baumann *et al.* 1984; Riggs 1998; Knapp *et al.* 1999). Circulating red blood cells contain a mixture of Hb isoforms with different O_2-binding affinities (Hiebl *et al.* 1987, 1988; Weber *et al.* 1988). It is thus possible that regulatory adjustments altering the stoichiometric ratio of these two isoforms may modulate rates of O_2 flux in response to changes in metabolic demand or to different environments.

Haemoglobins have been shown to be under strong selection in high-altitude environments in a variety of taxa (Snyder 1981; Chappell *et al.* 1988; Storz *et al.* 2007, 2009; Storz & Kelly 2008). McCracken *et al.* (2009a) surveyed the α^A and β^A haemoglobin subunit genes (HBA2 and HBB) across major waterfowl lineages and found a pattern of parallel evolution in highland waterfowl. Identical nonsynonymous codon substitutions occurred more often than expected by chance under a neutral model of evolution. Numerous substitutions of similar amino acids were also observed within the same region of the protein, suggesting that adaptation to high-altitude hypoxia has resulted from unique but overlapping sets of one to five substitutions in the major haemoglobin of each lineage. While there is evidence that the haemoglobins of high-altitude waterfowl lineages have evolved in response to hypoxia (Hiebl *et al.* 1987; Braunitzer & Hiebl 1988; Hiebl & Braunitzer 1988; Jessen *et al.* 1991; Weber *et al.* 1993; Liu *et al.* 2001), population genetic studies of high-altitude bird species are rare. Studies have yet to examine patterns of polymorphisms in a coalescent framework that incorporates aspects of population history, which also might influence levels of divergence at neutral and candidate genes such as haemoglobin.

Here, we use a series of independent genetic contrasts to compare three haemoglobin genes (α^D, α^A, β^A) and seven autosomal and mitochondrial reference loci in the crested duck (*Lophonetta specularioides*), a waterfowl species endemic to the Andes that inhabits both high- and

© 2011 Blackwell Publishing Ltd

low-altitude habitats. This study also provides novel data on the α^D-globin gene, which constitutes the minor (HbD) haemoglobin isoform and is known to be expressed in a 3:1 ratio in adult crested ducks (Bulgarella *et al.* 2009). Using a coalescent model, incorporating divergence and gene flow, we estimated a population size parameter (Θ), rates of gene flow and time since divergence between lowland and highland populations. Compared to simple descriptive measures of population differentiation, this approach has several advantages because it incorporates genetic stochasticity by sampling from the range of genealogies that are consistent with the data (Nielsen & Wakeley 2001; Hey & Nielsen 2004). Furthermore, by simulating genetic data under a drift–migration model of selective neutrality, we then contrasted patterns of genetic differentiation for the three haemoglobin genes with parameter estimates obtained from the set of reference loci to test for deviations from the neutral model. This study emphasizes the importance of incorporating the combined effects of selection and population history on candidate genes and genomic levels of population divergence for comprehending the patterns of evolution of a species that is distributed across an altitudinal gradient influenced by hypoxia.

Materials and methods

Study species

Crested ducks are endemic to South America; they comprise two subspecies that occupy different elevational environments in the Andes. *Lophonetta s. specularioides* inhabits low-elevation habitats in the southern Andean regions of Patagonia and the Falkland Islands, whereas *L. s. alticola* occurs throughout the Andean highlands up to approximately 4800 m elevation from central Argentina to northern Peru (Fjeldså & Krabbe 1990; Schulenberg *et al.* 2007). The two subspecies can be distinguished morphologically. *L. s. specularioides* has a bright red iris, smaller body size (wing chord < 280 mm), and darker, more mottled plumage than *L. s. alticola*, which is larger (wing chord > 280 mm) and possesses a yellow-orange iris (Phillips 1922–1926; Johnsgard 1978; Kear 2005; Bulgarella *et al.* 2007). The two subspecies intergrade in a zone of intermediate elevational habitats in Mendoza, Argentina, and Talca, Chile (Navas & Bo 1998; Bulgarella *et al.* 2007). Because of their linear distribution along the Andean Cordillera and distinguishing traits between subspecies inhabiting lowland and highland environments, crested ducks are an excellent species for investigating how changes in the α^D, α^A and β^A haemoglobin genes might confer resistance to hypoxia in different elevational zones.

Specimen collection

The Andean Cordillera comprises the long chain of mountain ranges and volcanoes that parallel the western margin of South America. The central Andes, including the Altiplano, is the widest and, on average, highest part of the cordillera. It consists of a series of high plateaus and inter-montane valleys that extend from Cajamarca, Peru, to Catamarca, Argentina, and Copiapo, Chile. South of Catamarca, the Andes narrow, lose their mean elevation and gradually transition to lower elevation peaks of the southern Andes of Patagonia. As a result, most wetlands and grasslands that are suitable for waterfowl in the central Andes occur at elevations >3000 m, whereas most waterfowl habitat in the southern Andes occurs at elevations <1500 m.

We collected crested ducks throughout their range in the Andes from southern Argentina to northern Peru. Eighty specimens were collected, including 49 *L. s. alticola* from high-elevation sites in Peru, Bolivia and northwestern Argentina (3338–4611 m) and 23 *L. s. specularioides* from low-elevation sites in Patagonia and the Falkland Islands (0–934 m). Eight additional *L. s. alticola* specimens were collected from five intermediate elevation sites in Mendoza, Argentina (1522–2552 m, Fig. 1, Table S1, Supporting information). Each individual was identified to subspecies using previously published analyses of plumage and morphology (Bulgarella *et al.* 2007).

DNA sequencing

We sequenced the three adult haemoglobin genes (α^D, α^A, and β^A) and compared them to seven reference loci located on different chromosomal linkage groups in the chicken genome (Hillier *et al.* 2004). The complete coding regions of the α^A and β^A haemoglobin subunits (677 and 1581 bp, respectively) and the 5′ end of the α^D subunit (435 bp spanning exon 1, intron 1 and part of exon 2) were amplified using PCR and sequenced using standard protocols described by McCracken *et al.* (2009b). The mtDNA control region was sequenced using the overlapping primer pairs L78-H774 and L736-H1252 or H1530 (Sorenson & Fleischer 1996; Sorenson *et al.* 1999; McCracken & Sorenson 2005). Six intron sequences, ranging from 246 to 353 bp in length, were also sequenced, and those included ornithine decarboxylase intron 5 (ODC1), α enolase intron 8 (ENO1), β fibrinogen intron 7 (FGB), *N*-methyl-D-aspartate-1-glutamate receptor intron 11 (GRIN1), phosphoenolpyruvate carboxykinase intron 9 (PCK1) and lamin A intron 3 (LMNA). Primers for each locus were developed specifically for ducks, and all intron loci were chosen blind to

© 2011 Blackwell Publishing Ltd

Fig. 1 Geographic distribution and sampling locations of crested ducks in this study. The range for crested duck includes the coastal and inland habitats delimited by the contour line.

levels of polymorphism (McCracken & Sorenson 2005; McCracken *et al.* 2009b, 2010).

Sequences from opposite strands were reconciled and verified for accuracy using Sequencher v.4.7 (Gene Codes, Ann Arbor, MI, USA). Double peaks indicating the presence of two alleles were coded with IUPAC degeneracy codes and treated as polymorphisms. Indels were resolved by comparing forward and reverse strands, using the unambiguous 5′ end of one strand to edit the ambiguous 3′ end of the other strand to determine the length of the indel (Peters *et al.* 2007). Gaps resulting in shifted peaks in the chromatograms were coded as a fifth character state.

The gametic phases of sequences that were heterozygous at two or more nucleotide positions were determined using PHASE v.2.1.1 (Stephens *et al.* 2001). PHASE uses a Bayesian method to infer haplotypes from diploid genotypic data while incorporating recombination and the decay of linkage disequilibrium with distance. We first analysed the diploid consensus sequences of each individual using the default software values followed by 1000 burn-in and 1000 sampling iterations. The PHASE algorithm was run five times from different starting points, and the results with the best overall goodness of fit were selected. For individuals with low allele pair probabilities (<80%), we designed allele-specific primers to selectively amplify one allele. The resulting haploid allele sequence was then subtracted from the consensus sequence to obtain the gametic phase of the second allele. Each data set was analysed five more times using PHASE with the newly resolved alleles defined as 'known' alleles. PHASE analyses and allele-specific priming were performed for the complete α^A and α^D haemoglobin sequences and each of the six autosomal introns. For the β^A subunit, we inferred the gametic phase of a 596-bp segment spanning intron 1, exon 2 and part of intron 2 that had no detectable recombination as determined using the four-gamete test (Hudson & Kaplan 1985). Overall, the gametic phases of 97.9% ($n = 705$) of the 720 individual autosomal sequences were identified experimentally or with >95% posterior probability. All sequences were aligned by eye using SEAL v.2.0a11 (Rambaut 2007).

Analyses of population differentiation

To determine whether lowland and highland populations of crested ducks are genetically differentiated, we compared Φ_{ST} between *L. s. specularioides* and *L. s. alticola* for each locus. F_{ST} for each nonsynonymous (amino acid) polymorphism in the α^D, α^A and β^A haemoglobin subunits was also calculated. Additionally, we tested Hardy–Weinberg equilibrium to determine whether there was a significant excess or deficit of heterozygotes for any given locus, which might be interpreted as divergent selection acting on alleles that were overrepresented in one or the other population. All estimates of linkage disequilibrium, F_{ST} and Φ_{ST} were performed using the Tamura & Nei (1993) nucleotide substitution model in ARLEQUIN v.3.01 (Excoffier *et al.* 2005). Allelic richness standardized to the smallest sample size ($n = 46$ alleles) was calculated using the software RAREFAC (Petit *et al.* 1998) and then compared between highland and lowland populations using a paired *t*-test. The eight individuals from Mendoza (the area of overlap between the two subspecies; Bulgarella *et al.* 2007) were excluded from these analyses. Allelic networks for each locus included the Mendoza specimens and were calculated using the median-joining algorithm in NETWORK v.4.1 (Bandelt *et al.* 1999).

© 2011 Blackwell Publishing Ltd

We also used STRUCTURE v.2.2 (Pritchard *et al.* 2000) to examine population differentiation. STRUCTURE uses a Bayesian method to assign individuals to populations by maximizing Hardy–Weinberg equilibrium and minimizing linkage disequilibrium. We performed two STRUCTURE analyses; the first included only the six autosomal introns, and the second included the six introns plus the α^D, α^A, and β^A haemoglobin genes (nine loci in total, mtDNA was not included in the analysis; see below). The eight individuals from Mendoza were included in this analysis, and no a priori information describing specimen locality was provided. We used the admixture model with independent allele frequencies, 10 000 generations of burn-in and 20 000 generations of sampling. The optimum number of crested duck populations was identified using the Δk method (Evanno *et al.* 2005). To determine the number of populations (k), we estimated ln Pr(X | k) for $k = 2$ to $k = 5$. Using the value of k with the highest ln Pr(X | k), we assigned individuals to the inferred populations.

Coalescent analyses

We compared the three haemoglobin genes to the six autosomal intron loci using a two-population isolation-with-migration (IM, Hey & Nielsen 2004) analysis that allows for both divergence and gene flow. We estimated indices of effective population size ($\Theta = 4N_e\mu$), gene-flow rates (M), time since population divergence (t) and TMRCA (time since the most recent common ancestor) for each locus. The software IM implements a Bayesian MCMC method to fit the data to a coalescent model. We estimated six population parameters scaled to the neutral mutation rate, μ: Θ_H ($4N_e\mu$ for highland crested ducks), Θ_L ($4N_e\mu$ for lowland crested ducks), Θ_A ($4N_e\mu$ for the ancestral population at the time of divergence), t ($T\mu$, where T is the time since divergence in years before the present), M_L (m_L/μ, where m is the rate of immigration into the lowlands from the highlands) and M_H (m_H/μ, the rate of immigration into the highlands from the lowlands). The eight individuals collected from Mendoza were excluded from this analysis (see Results below).

Isolation-with-migration assumes that the loci are selectively neutral with no intralocus recombination. Therefore, we tested for recombination within each nuclear locus using a four-gamete test in DNASP v.4.10 (Rozas *et al.* 2003) and included sequence data from the largest independently segregating block consistent with no recombination. For the haemoglobin genes, we included the longest fragment consistent with no recombination that included all of the nonsynonymous amino acid replacements that we observed: α^D haemoglobin subunit positions 108–441, α^A haemoglobin subunit positions 1–204 and β^A subunit positions 140–735. GRIN1 and LMNA were truncated to the 5′-end positions 1–239 and positions 1–195, respectively. The remaining nuclear loci had no detectable recombination, and therefore, the full sequences were included in the analysis. Additionally, we verified the results of the four-gamete tests with an independent estimate of the overall recombination rate ($r = \rho/\mu$; where ρ is the per-site recombination rate and μ is the per-site mutation rate) for each locus using the software LAMARC v.2.1 (Kuhner 2006). The upper and lower limits for r were set to 0 and 10, respectively. A recombination rate of one means that recombination is equally likely to occur as mutation, whereas a recombination rate of zero means that no recombination was detectable.

For mtDNA, which is haploid and maternally inherited, we defined the inheritance scalar to be 0.25 and used the HKY (Hasegawa *et al.* 1985) model of mutation. For nuclear loci, we defined the scalar to be 1.0 because they are biparentally inherited, and we used an infinite-site model of substitution. IM was first run with wide priors to set appropriate upper bounds for each parameter. These runs were then repeated using uniform priors that encompassed the full posterior distribution of each parameter from the preliminary runs and were therefore assumed to be uninformative. However, estimates of t sometimes contained distinct peaks, but the tails were flat and did not approach zero. In these cases, we defined an upper bound based on preliminary runs by assuming that time since divergence could not be older than TMRCA (Peters *et al.* 2007). Averaging the posterior distribution of TMRCA for the six nuclear introns, we used the upper 90% highest posterior density (HPD) of TMRCA as the upper bound for t (upper bound = 0.85). The priors encompassed the full posterior distributions of all other parameters. We repeated runs with a burn-in of at least 200 000 steps. To assess convergence, we monitored autocorrelations and effective sample sizes (ESS) for each parameter throughout the run (Hey & Nielsen 2004). The run was continued until the smallest ESS was at least 100 (see Hey 2005). We ran the analysis a second time with a different random number seed to be sure that different runs converged on the same parameter estimates. All runs included 10 heated chains using a geometric heating scheme. Finally, to obtain an accurate estimate of TMRCA, we analysed each intron locus independently, because IM calculates a mutation rate scalar based on differences among loci for Θ. Therefore, IM will assume that a locus with a high level of polymorphisms has a higher substitution rate than lower diversity loci.

To convert IM parameter estimates to biologically informative values, estimates of generation time (G) and mutation rate (μ per locus) are necessary. We used the

© 2011 Blackwell Publishing Ltd

control region mutation rate calibrated by Peters *et al.* (2005) of 4.8×10^{-8} substitutions/site/year (s/s/y) for the mtDNA and calibrated rates for the three globin and the six intron loci based on the goose–duck split (see Peters *et al.* 2007, 2008). The mean genetic distance between the snow goose (*Anser caerulescens*) sequence and crested duck was divided by the mid-point of the Oligocene (2×30.5 myr). The geometric mean of substitution rates averaged for the six introns was 2.72×10^{-9} substitutions/site/year.

Simulated data sets

We simulated genetic data under a drift–migration model of selective neutrality using the software MS (Hudson 2002). Using the parameter estimates from IM for the six autosomal introns, we defined a population model that included empirical differences in the population size parameters (Θ_H and Θ_L) and gene-flow rates (M_L and M_H) between the lowlands and highlands, conditioned on the data. To account for uncertainty in the exact values of those parameters, we randomly sampled values from the posterior distributions of each parameter for 1000 replicates. The Θ values in MS were scaled to the per-locus substitution rates, calculated by multiplying IM's Θ by the ratio of the locus-specific length (l) to the geometric mean of the length of the introns used in IM. We also incorporated heterogeneous substitution rates by comparing uncorrected sequence divergence between crested ducks and snow geese (d; see above) and calculating relative substitution rates among loci, defined as d_i (sequence divergence for locus i) divided by the geometric mean of d for all loci. Locus-specific values of Θ were multiplied by the relative substitution rate. Locus-specific recombination rates (ρ) were incorporated, where $\rho = \Theta \times r \times (l - 1)$, Θ and r are LAMARC's estimates of $4N_e\mu$ and recombination rate, respectively. Uncertainty in ρ was incorporated by randomly sampling values for each replicate from the posterior distributions estimated in the LAMARC analysis. We simulated 1000 data sets, with each data set containing the same number of sequences as the empirical data. We then calculated the distribution of pairwise Φ_{ST} values expected under selective neutrality and compared that distribution to empirical values of Φ_{ST} for each locus.

Results

Population differentiation

We found strong evidence of population differentiation between highland and lowland populations of crested ducks. Φ_{ST} values for the six introns, three globins and mtDNA ranged from 0.02 to 0.87 and were significantly different between populations for eight of the ten loci (Table 1).

Most intron alleles, including the most common haplotypes, were shared between subspecies (Fig. 2B). Φ_{ST} values for the six autosomal introns varied from a minimum of 0.02 in GRIN1 to a maximum of 0.44 in ODC1. An average of 4.4 different alleles was expected per locus in a random sample of 46 alleles (AR_{46}) from the highlands compared to 3.8 different alleles from the lowlands, but allelic richness did not differ between the highlands and the lowlands among intron loci (paired

Table 1 Population genetic parameters estimates for ten loci from lowland and highland populations of crested ducks

Loci	Length (bp)	Chromosomal location[†]	No. of alleles[‡]	Allelic richness AR[‡]	Recombination rate (ρ/μ)	Φ_{ST}[§]
α^D-Haemoglobin (HBA1)	465	14	**6**/9	**3.9**/8	0.70	**0.33**
α^A-Haemoglobin (HBA2)	677	14	**12**/14	**8.6**/13	1.07	**0.23**
β^A-Haemoglobin (HBB)	596	1	**9**/4	**6.3**/3	0	**0.87**
α-Enolase intron 8 (ENO1)	314	21	**5**/6	**3.2**/5	0	**0.21**
β Fibrinogen intron 7 (FGB)	246	4	**2**/2	**0.7**/1	0	**0.13**
N-Methyl-D-aspartate-1-glutamate receptor intron 11 (GRIN1)	335	9	**11**/11	**8.5**/10	2.77	0.02
Lamin A intron 3 (LMNA)	280	1	**15**/6	**10.1**/5	1.10	**0.10**
Ornithine decarboxylase intron 5 (ODC1)	353	3	**4**/2	**2.7**/1	0	**0.44**
Phosphoenolpyruvate carboxykinase intron 9 (PCK1)	345	20	**2**/2	**1**/1	0	0.03
mtDNA control region (mtDNA)	980	MT	**17**/9	**10.6**/8	—	**0.85**

[†]Based on the chicken genome (Hillier *et al.* 2004).
[‡]Highland populations are shown in bold text. The individuals from Mendoza were excluded in all the calculations.
[§]Statistically significant Φ_{ST} values are in bold ($P < 0.001$).

© 2011 Blackwell Publishing Ltd

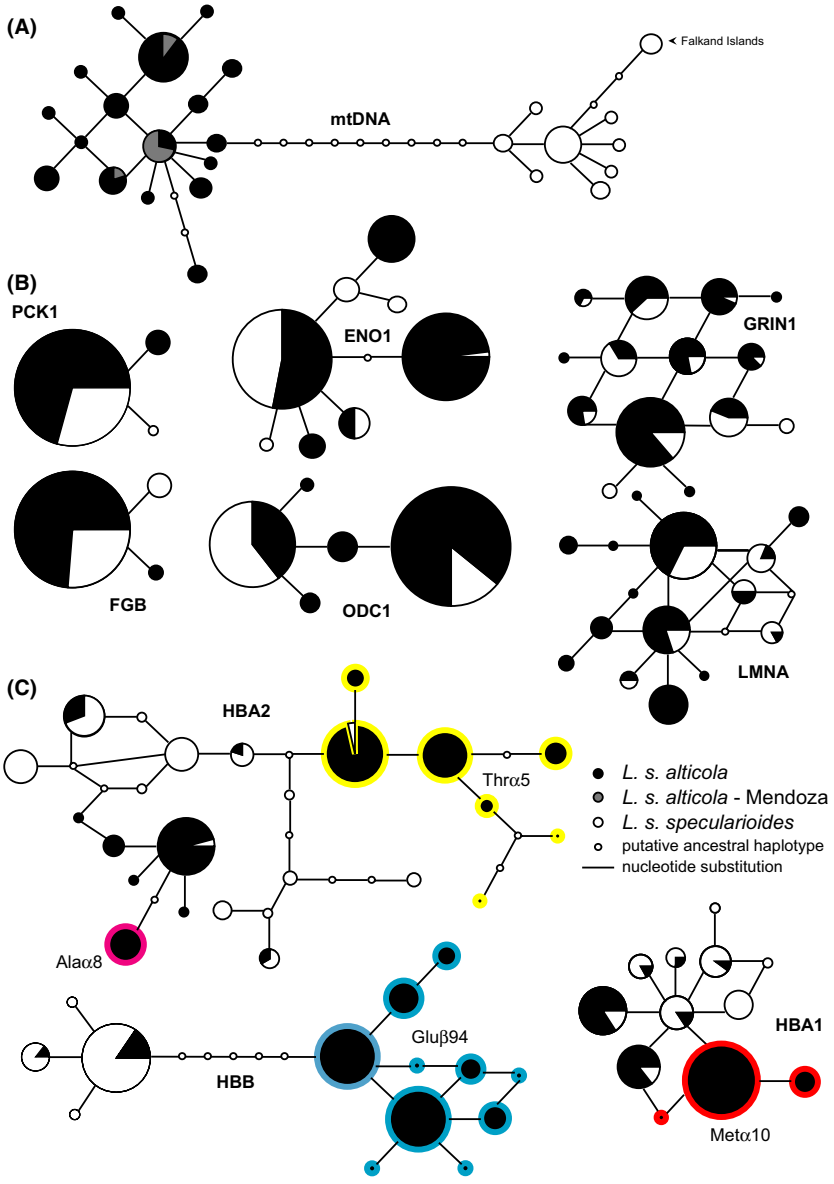

Fig. 2 Unrooted networks showing the relationships between (A) mtDNA haplotypes, (B) the six introns and (C) the three haemoglobin subunits. HBA1, HBA2 and HBB alleles possessing amino acids that occurred at high frequency in the highlands are shown with highlighted circles. The areas of the circles are proportional to the number of alleles found. The eight specimens from Mendoza were included in the networks (n = 80).

t-test, t = 0.51, d.f. = 5, P = 0.63; Table 1; differences among all 10 loci also were not significant; t = 0.06, d.f. = 9, P = 0.95). No linkage disequilibrium was found between the six reference loci, confirming that these loci are likely unlinked (Ps > 0.99).

Using the six introns, the STRUCTURE analysis with the maximum likelihood (lnL = −2189.0) and greatest value of Δk was for k = 2 populations. Excluding the eight Mendoza specimens, the mean posterior probabilities of assignment to the highland population (>3000 m) for *L. s. alticola* was 0.86 ± 0.18 and varied from 0.19 to 0.98. For *L. s. specularioides*, the mean posterior probability of assignment to the lowland population was 0.93 ± 0.06 and ranged from 0.77 to 0.99 (Fig. 3A). The eight individuals collected at 1522–2552 m of elevation and between 32.8 and 35.8°S in the province of Mendoza (shown in grey in Fig. 3) exhibited the full range of posterior probabilities to be assigned to the highland population (0.05–0.97).

Haemoglobin gene differentiation

The Φ_{ST} for the β^A haemoglobin subunit (0.87) was higher than for any other locus, and the Φ_{ST} values for α^A (0.23) and α^D (0.33) haemoglobins were higher than all intron loci, except ODC1 (Table 1). Four nonsynonymous amino acid substitutions were found in the three adult haemoglobin gene sequences. Two amino acid polymorphisms occurred in the α^A subunit (Thr/Ala-α^A5 and Ala/Thr-α^A8), one was in the α^D subunit

© 2011 Blackwell Publishing Ltd

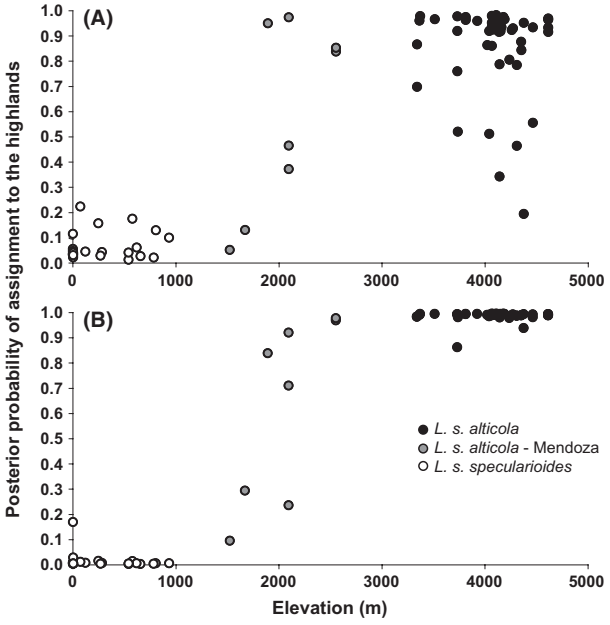

Fig. 3 Posterior probability of assignment to the highland population vs. elevation, based on Bayesian likelihood estimates for the 80 individuals in the study, for (A) the six autosomal introns, (B) the six autosomal introns and the three haemoglobin loci combined.

(Met/Thr-α^D10), and one was in the β^A subunit (Asp/Glu-β^A94, Fig. 4). All four exhibited significant genotype frequency differences between lowland and highland populations (Fig. 4). Pairwise F_{ST} values for these four nonsynonymous polymorphisms were 0.47, 0.07, 0.61 and 1.00, respectively.

The highland subspecies *L. s. alticola* exhibited high levels of heterozygosity on the α^A subunit. Excluding Mendoza, 55% (n = 27) were heterozygous for Thr/Ala-α^A5, whereas 31% (n = 15) were homozygous for Thr-α^A5, and 14% (n = 7) were homozygous for Ala-α^A5. In contrast, only one *L. s. specularioides* was heterozygous for Thr/Ala-α^A5; all others were homozygous for Ala-α^A5. Ala-α^A8 was found in eleven *L. s. alticola* individuals collected in the highlands; all other individuals from both the lowlands and the highlands were homozygous for Thr-α^A8. Finally, the Thr-α^A5 and Ala-α^A8 alleles did not occur on the same haplotype groups but were 1.5–2.2% divergent (uncorrected) in crested ducks and probably evolved from divergent ancestral sequences (Fig. 2; Table 2). Based on a previous study of haemoglobin polymorphism including most of the world's waterfowl species, Thr-α^A5 and Ala-α^A8 were both determined to be derived, whereas Ala-α^A5 and Thr-α^A8 are the ancestral amino acid residues found in other lowland species (McCracken *et al*. 2009c).

Highland crested duck also exhibited high levels of heterozygosity on the α^D subunit. Excluding Mendoza,

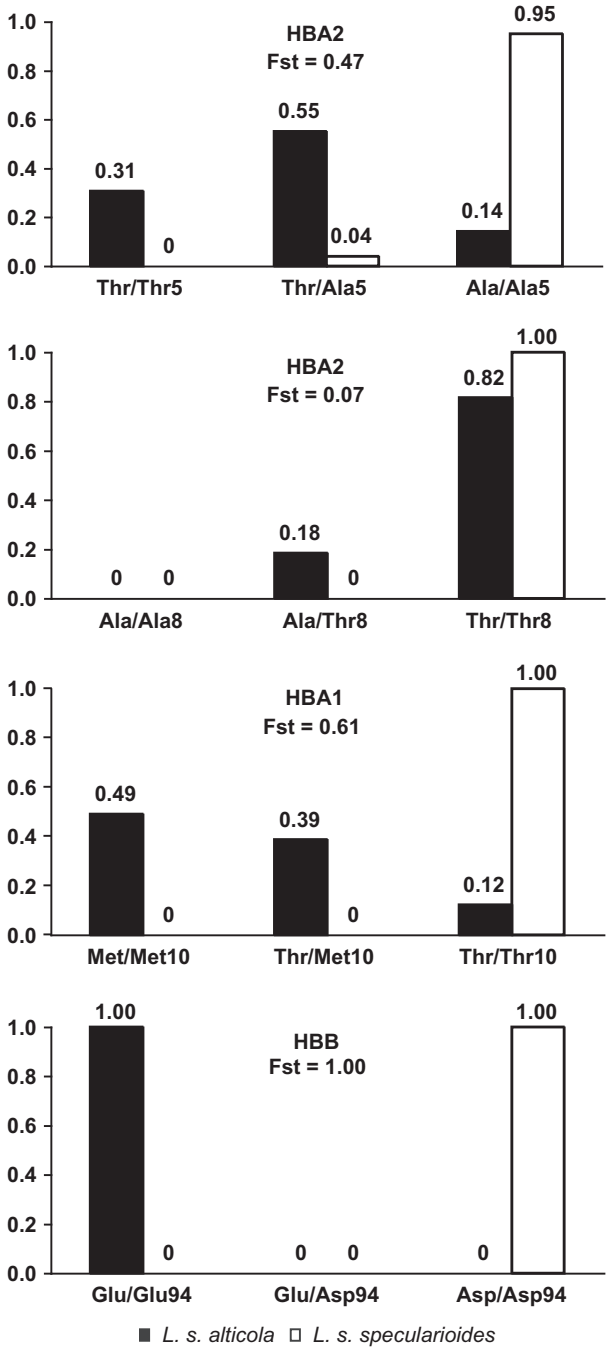

Fig. 4 Genotype frequencies for four amino acid replacements observed in highland and lowland crested ducks haemoglobin subunits: Thr/Ala-α^A5, Thr/Ala-α^A8, Met/Thr-α^D10 and Asp/Glu-β^A94.

39% (n = 19) were heterozygous for Met/Thr-α^D10, whereas 49% (n = 24) were homozygous for Met-α^D10, and 12% (n = 6) were homozygous for Thr-α^D10. All *L. s. specularioides* were homozygous for Thr-α^D10 (Fig. 4, Table 2); no Met-α^D10 alleles were found in the lowland population.

© 2011 Blackwell Publishing Ltd

Table 2 Haemoglobin genotypes observed in 80 crested ducks

HBA2 5	**HBA2** 8	**HBA1** 10	**HBB** 94	Highlands	Mendoza	Lowlands
Thr/Thr	Thr/Thr	**Met/Met**	**Glu/Glu**	3	—	—
Thr/Thr	Thr/Thr	**Met**/Thr	**Glu/Glu**	6	—	—
Thr/Thr	Thr/Thr	Thr/Thr	**Glu/Glu**	6	—	—
Thr/Ala	Thr/**Ala**	**Met/Met**	**Glu/Glu**	2	—	—
Thr/Ala	Thr/**Ala**	**Met**/Thr	**Glu/Glu**	5	—	—
Thr/Ala	Thr/Thr	**Met/Met**	**Glu/Glu**	13	—	—
Thr/Ala	Thr/Thr	**Met**/Thr	**Glu/Glu**	7	—	—
Ala/Ala	Thr/**Ala**	**Met/Met**	**Glu/Glu**	3	—	—
Ala/Ala	Thr/**Ala**	**Met**/Thr	**Glu/Glu**	1	—	—
Ala/Ala	Thr/Thr	**Met/Met**	**Glu/Glu**	3	—	—
Thr/Thr	Thr/Thr	**Met**/Thr	Asp/Asp	—	1	—
Thr/Thr	Thr/Thr	Thr/Thr	**Glu**/Asp	—	1	—
Thr/Thr	Thr/Thr	Thr/Thr	Asp/Asp	—	1	—
Thr/Ala	Thr/Thr	Thr/Thr	**Glu/Glu**	—	1	—
Ala/Ala	Thr/Thr	Thr/Thr	**Glu/Glu**	—	1	—
Thr/Ala	Thr/Thr	**Met**/Thr	**Glu**/Asp	—	2	—
Ala/Ala	Thr/Thr	Thr/Thr	**Glu**/Asp	—	1	—
Ala/Ala	Thr/Thr	Thr/Thr	Asp/Asp	—	—	22
Thr/Ala	Thr/Thr	Thr/Thr	Asp/Asp	—	—	1
Total number of individuals (*n*)				49	8	23

Highland alleles are shown in bold.

For the β^A subunit, highland and lowland crested ducks collected outside Mendoza were fixed for opposite genotypes, Glu-β^A94 or Asp-β^A94, respectively (Fig. 4, Table 2). Within Mendoza, four of the eight individuals were heterozygous for Asp/Glu-β^A94, and two each were homozygous for Glu-β^A94 and Asp-β^A94, respectively (Table 2). Glu-β^A94 is the derived genotype, and Asp-β^A94 is the ancestral genotype found in other waterfowl species (McCracken *et al.* 2009c).

The three haemoglobin genes thus exhibited two distinctly different patterns; 100% of highland crested ducks were homozygous for Glu-β^A94, whereas for the two α-globins, highland individuals were heterozygous for Thr/Ala-α^A5 (55%) or Met/Thr-α^D10 (39%). In contrast, the same three derived alleles (Thr-α^A5, Met-α^D10 and Glu-β^A94) were rare or absent in the lowland population.

Of the 81 possible combinations of alleles for each of the four amino acid substitutions found in the three adult globin subunits, 19 distinct genotypes were found. In the highlands, every individual's genotype contained at least two of the derived amino acid polymorphisms, with most individuals having at least three. In the lowlands, all but one individual possessed the ancestral genotype shared with other lowland waterfowl species (Table 2). Only in Mendoza did we find unique combinations of genotypes that were not sampled in either of the other two habitats (Table 2). Seven different unique genotypes were observed in eight individuals in this region, and it was the only place in which Asp/Glu-β^A94 heterozygotes occurred. Thus, the highest level of heterozygosity was found in the Mendoza subsample.

When haemoglobins were added to the STRUCTURE analysis including the six introns, population assignment probabilities improved notably (Fig. 3B). Excluding Mendoza, the mean probability of assigning *L. s. alticola* to the highland population was 0.99 ± 0.02 (range: 0.86–0.99), and the mean probability of assigning *L. s. specularioides* to the lowland population was 0.98 ± 0.03 (range: 0.83–0.99). Thus, all individuals were correctly assigned to subspecies (as determined using morphological characters; Bulgarella *et al.* 2007) with high probability. This result was not an artefact of including more loci in the analysis, because the same results were observed as with the six introns alone when we randomly resampled three intron loci for a total of nine loci. The eight individuals collected at intermediate elevation in Mendoza (shown in grey in Fig. 3), as in the introns-only analysis, showed intermediate posterior probabilities compared to the other individuals (range: 0.09–0.98), consistent with a transition zone occurring over short geographic and elevational distances in southern Mendoza.

© 2011 Blackwell Publishing Ltd

Finally, we tested Hardy–Weinberg (HW) equilibrium for the amino acid polymorphisms in the three haemoglobin genes. When highland and lowland populations were tested separately, all populations were found to be in HW equilibrium (Ps > 0.05). However, when highland and lowland populations were pooled, all HW tests were highly significant. More homozygotes and fewer heterozygotes were observed than expected for all three haemoglobin loci (Ps < 0.05). No linkage disequilibrium was observed between the α^A- and α^D-globin genes, which are separated by 2.2 kb in the Muscovy duck (Niessing *et al.* 1982). The average per cent amino acid sequence identity between α^A and the 5′–α^D sequence was 58.6%.

MtDNA differentiation

The mtDNA control region (Φ_{ST} = 0.85) yielded two reciprocally monophyletic clades corresponding to the lowland subspecies (*L. s. specularioides*) and the highland subspecies (*L. s. alticola*; Fig. 2A). All eight specimens collected in Mendoza, Argentina, possessed the highland *L. s. alticola* haplotypes. The three *L. s. specularioides* from the Falkland Islands possessed a unique haplotype that was three base pairs divergent and grouped with, but was not shared with, *L. s. specularioides* individuals from Argentina (Fig. 2A).

Coalescent analysis

Immigration rates between the highlands and the lowlands. The joint estimated scaled gene-flow rate for the six introns into the highlands (M_H = 8.11; 90% HPD_{adj} = 1.24–20.46) exceeded gene flow into the lowlands (M_L = 1.29; 90% HPD_{adj} = 0.01–15.11; Table 3, Fig. 5), although confidence intervals overlapped extensively. Following divergence, the joint estimate for the six introns indicated that there were on average approximately 2.6 migrants/generation ($4N_em$) into the highlands from the lowlands, and effectively no migration into the lowlands from the highlands ($4N_em$ = 0.3).

For the three globins, no migration was detected in any direction ($4Nm$ < 1), with the exception of the α^A haemoglobin gene, which showed a small ($4N_em$ = 0.90) but nonzero estimate of migration into the highland population from the lowland population. In contrast to the introns, the lower bound of the posterior probabilities for the three globins overlapped zero in both directions (Table 3, Fig. 5). This same pattern also occurred with mtDNA; no migration was detected in either direction, and the lower bound of the posterior distributions overlapped zero (Table 3).

Effective population size (N_e) and Theta (Θ) estimates. Based on joint Θ estimates for the six introns, we calculated an effective population size (N_e) of approximately 93 000 highland crested ducks (90% HPD_{adj} = 44 000–186 000 individuals) and 76 000 lowland crested ducks (90% HPD_{adj} = 29 000–183 000 individuals, Table 3). The N_e estimates for the highlands and lowlands were thus overlapping, and our estimates assume that lowland and highland crested ducks experienced similar substitution rates. The ancestral N_e was estimated to be approximately 41 000 individuals (90% HPD_{adj} = 500–900 000 individuals), suggesting that both populations have grown in size following divergence.

MtDNA estimates suggest a N_e of approximately 30 300 highland crested ducks (90% HPD_{adj} = 16 000–50 000 individuals) and 16 500 lowland crested ducks (90% HPD_{adj} = 7000–36 000 individuals) with an ancestral N_e of 19 500 individuals (90% HPD_{adj} = 58–101 000 individuals). Although estimates were broadly overlapping, the mean Θ was greater in the highland popula-

Table 3 IM estimates for the population size parameter theta ($\Theta = 4N_e\mu$), immigration rates ($M = m/\mu$) and TMRCA (t = TMRCA/μ) for six introns, three haemoglobin genes and the mtDNA control region

Locus	$\Theta_{alticola}$	$\Theta_{specularioides}$	$\Theta_{ancestral}$	M (m/μ)$_{alticola}$	M (m/μ) $_{specularioides}$	TMRCA/μ (BP) (TRMCA)
Six introns	0.32 (0.15–0.63)	0.26 (0.10–0.63)	0.14 (0.00–3.08)	8.11 (1.24–20.46)	1.29 (0.01-15.11)	**ENO1** 2 363 460 (0.6890) **FBG** 577 763 (0.2767) **LMNA** 2 995 129 (0.6456) **GRIN1** 2 271 997 (0.4069) **ODC1** 598 683 (0.2658) **PCK1** 1 715 825 (0.2550)
HBA1	0.05 (0.00–0.35)	0.79 (0.19–2.80)	—	2.19 (0.01–20.74)	0.01 (0.01–7.41)	1 969 545 (0.6038)
HBA2	0.29 (0.06–1.04)	0.63 (0.13–2.31)	—	3.16 (0.18–17.11)	0.56 (0.01–8.06)	3 415 066 (0.8568)
HBB	0.81 (0.29–1.89)	0.57 (0.13–1.68)	—	0.00 (0.00–0.55)	0.00 (0.00–0.82)	5 544 525 (2.5830)
mtDNA	18.31 (9.98–30.63)	9.98 (4.31–22.16)	—	0.00 (0.00–0.18)	0.00 (0.00–0.34)	119 494 (5.6210)

The 90% upper and lower posterior parameter estimates are shown in parentheses.

© 2011 Blackwell Publishing Ltd

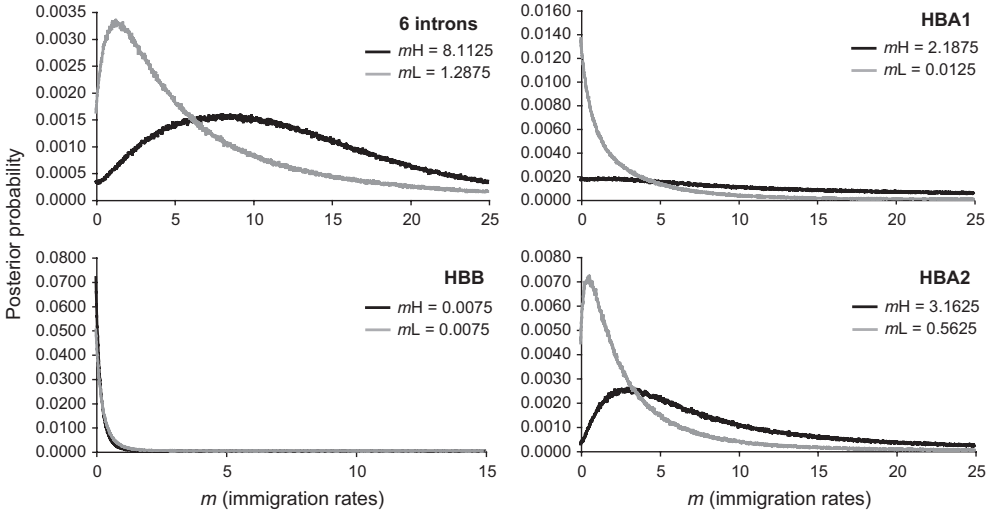

Fig. 5 Migration rates into the highlands (M_H) and into the lowlands (M_L) for the six introns and the three haemoglobin loci, respectively.

tion than the lowland for the six introns and for mtDNA.

Time since divergence and TMRCA. The mean values for TMRCA ($t \times \mu$) ranged between 0.25 and 0.68 for the six introns (Table 3, Fig. S1, Supporting information). TMRCA values for the globins were on average substantially larger, with 0.60 for HBA1, 0.86 for HBA2 and 2.58 for HBB. The mean TMRCA values for HBA2 and HBB thus exceeded the introns. MtDNA had the deepest TMRCA of all loci: 5.62 (Table 3). Converting the estimated TMRCA to time in years based on the per-locus substitution rates indicated that all three haemoglobin loci had deeper coalescent times than any of the introns (Fig. S1, Supporting information).

Based on the joint estimate for the six introns, time since divergence between highland and lowland populations was estimated to be 285 000 years ago. Nevertheless, the lower and upper bound of the 90% HPD_{adj} could not be estimated because the posterior distribution had a rising upper tail. Based on mtDNA, time since divergence was approximately 108 500 years ago (90% HPD_{adj} = 53 400–252 100).

Coalescent simulations. Despite incorporating uncertainty in our inferences of population history, the observed Φ_{ST} value for the β^A haemoglobin subunit fell outside the 95% confidence limit of the simulated data (Fig. 6). In contrast, Φ_{ST} values for all six introns were within the confidence limits, suggesting that inter-population differences were consistent with the inferred population history and neutral molecular evolution. Φ_{ST} values for the α^A and α^D haemoglobin subunits were both higher than the mean of the simulated values, but fell within the 95% confidence limits. Thus, we could not reject neutral evolution for these subunits.

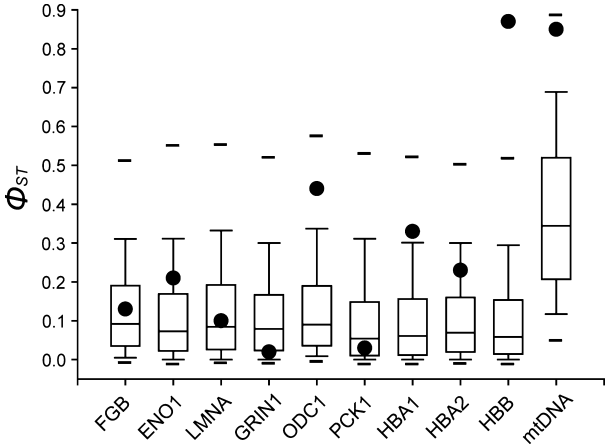

Fig. 6 Genetic differentiation between *Lophonetta s. alticola* and *L. s. specularioides* observed at ten independent loci (filled circles) and simulated under selective neutrality and the inferred isolation-with-migration model (box plots). The horizontal dashes indicate the 95% confidence interval calculated from the simulated data. Only HBB was more differentiated than expected under neutrality (only 0.2% of simulated values of Φ_{ST} were greater than or equal to the observed value). However, for mtDNA, only 3.4% of the simulated values were greater than or equal to the observed value.

Discussion

Contrasting differentiation patterns and gene flow

Analysis of multilocus genetic variation revealed a significant pattern of genetic differentiation between

© 2011 Blackwell Publishing Ltd

lowland and highland populations of crested ducks. The Φ_{ST} values for the six autosomal reference loci ranged from 0.02 to 0.44, and these values were consistent with the inferred model of population history and neutral evolution. Historical migration rates based on the joint estimate for the introns averaged 2.6 effective migrants/generation into the highlands from the lowlands, and the 90% confidence interval did not overlap zero. By contrast, no migration was observed into the lowland population from the highlands following population divergence.

The Φ_{ST} values for the three haemoglobin genes were generally greater (range 0.23–0.87), with the β^A haemoglobin gene showing the greatest differentiation of any locus in our 10-locus study. Whereas the observed Φ_{ST} values for HBA1 and HBA2 fell inside the 95% confidence limit in our simulations, the value for HBB did not, suggesting that this locus is not evolving neutrally. It is possible that the polymorphisms observed in the α^A- and α^D-globin are either selectively neutral, or the effects of selection could not be differentiated from background levels of population structure and linkage disequilibrium. Moderate levels of geographic structure, as those found in this study, can limit the detection of differentially selected loci (Beaumont 2005; Nosil *et al.* 2008, 2009; Storz 2005). Nonetheless, coalescent analyses for these three loci showed little evidence of allelic migration between highland and lowland populations, with the exception of the α^A haemoglobin gene, which showed a small ($4Nm = 0.90$) but nonzero estimate of migration into the highland population from the lowlands. Furthermore, both subunits generally had higher values of Φ_{ST} (albeit nonsignificantly) than simulated values.

Geographic analysis of gene flow thus documented a strong pattern of haemoglobin differentiation in the different environments that crested ducks inhabit. The mismatch in migration estimates between the globins and the autosomal introns likely reflect the more restricted movement of haemoglobin alleles along the elevational gradient relative to the neutral nuclear alleles. Four derived amino acid polymorphisms occurred at high frequencies in the highland population, whereas the lowland population generally lacked these alleles and was predominantly fixed for the ancestral haemoglobin alleles; a single individual possessed one Thr-α^A5 allele (Table 2).

Significance of haemoglobin amino acid replacements

Of the four amino acid replacements found in crested ducks, Glu-β^A94 is potentially important in influencing haemoglobin-O_2 affinity. Haemoglobins bind to allosteric effectors, such as inositolpentaphosphate (IPP), in both oxy and deoxy forms, with a higher affinity in the deoxy form (Rollema & Bauer 1979). Allosteric effectors lower Hb-O_2 affinity by strengthening salt bridges that favour the low-affinity T structure of haemoglobin (Perutz 1983). The positively charged groove at the entrance to the central cavity between the two β chains is the phosphate-binding site, and IPP binds directly to residues at β^A1, 2, 82, 104, 135, 139, 143, 144 and 146 (Zhang *et al.* 1996; Wang *et al.* 2000; Liang *et al.* 2001a,b; Liu *et al.* 2001). The Glu-β^A94 substitution lies adjacent to this binding cavity and in close proximity to IPP-binding sites at β^A144 and β^A146 (Liang *et al.* 2001a). In human haemoglobin, a salt bridge between the imidazole ring of the *N*-terminal His-β^A146 and Asp-β^A94 stabilizes the low-affinity deoxy structure and decreases haemoglobin-O_2 affinity (Perutz 1970; Shih *et al.* 1993). Bar-headed goose deoxy haemoglobin lacks this salt bridge, which weakens the Bohr effect and increases haemoglobin-O_2 affinity (Liang *et al.* 2001b). The presence or absence of this salt bridge in crested duck is yet unknown. However, McCracken *et al.* (2009c) found that high-altitude endemic populations of puna teal (*Anas puna*) are also fixed for Glu-β^A94, whereas their lowland sibling species silver teal (*Anas versicolor*) was fixed for Asp-β^A94, the same ancestral residue shared by lowland crested duck. Experiments examining sensitivity of Hb-O_2 binding to IPP and CO_2 would help determine the functional characteristics of this amino acid substitution in crested duck.

In contrast to the β^A subunit, which possessed only a single amino acid replacement, two derived substitutions were observed on the A helix of the α^A subunit of crested duck. Interestingly, highland populations of two other divergent Andean waterfowl species also possessed derived amino acid substitutions in this same structural domain of the HbA tetramer: Andean goose (*Chloephaga melanoptera*; Ala-α^A8) and cinnamon teal (*Anas cyanoptera*; Ser-α^A9; McCracken *et al.* 2009c). Amino acid residues α^A5, 8 and 9 each occupy external, solvent-accessible positions in the HbA structure (McCracken *et al.* 2010), and this region of the protein is known to undergo an important conformational change during the transition from the deoxy to the oxy state (Perutz 1990).

Finally, a novel fourth derived amino acid replacement was observed in crested duck in the same structural α domain of the HbD tetramer: Met-α^D10 (α^D10 corresponds to α^A11 in HbD and HbA, respectively, because of one amino acid difference in the length of the αD and αA chain subunits). The Andean goose also possesses a derived substitution in this region of the HbD protein (Leu-α^D9, McCracken *et al.* 2010). Thus, the A helix of the α^A and α^D-chains may be an important region of the haemoglobin protein involved in

© 2011 Blackwell Publishing Ltd

hypoxia resistance, a hypothesis that warrants further study. However, little physiological data are available for these highland species, and blood- or Hb-O_2 affinity as measured by p50 (the P_{O_2} at which blood or haemoglobin is half-saturated) have not been obtained for crested ducks, so the effects of Thr-α^A5, Ala-α^A8, Met-α^D10 and Glu-β^A94 are yet to be determined.

In sum, the general lack of evidence for gene flow for haemoglobin alleles between highland and lowland populations of crested ducks, as compared to the autosomal introns, and the biochemical properties of the amino acid substitutions are consistent with the effects of selection acting on these loci. Selection exerted by hypoxia at high elevations has been shown to be a strong determinant of haemoglobin haplotype frequencies in mice (Snyder 1981; Snyder *et al.* 1988; Storz *et al.* 2007). This is also likely to be the case in crested ducks. If highland haemoglobin alleles confer higher affinity for oxygen or decreased sensitivity to allosteric effectors such as IPP, then their effects are likely advantageous in the highlands, and individuals that possess such alleles would be expected to have higher fitness. Nonetheless, it should be noted that our study spans an extensive geographic scale and that the two main environments sampled differ in many parameters besides elevation, including differences in latitude, as well as factors that covary with elevation (i.e. temperature, desiccation, increased radiation). While it is probably unlikely that changes in haemoglobin structure and function relate to many of these variables, physiological experiments are necessary to clarify whether haemoglobin is a direct target of selection or the patterns we observed result from selection on other linked genes or traits that exhibit correlated responses to elevation.

Contrasting patterns of heterozygosity in the α- and β-globin genes and in intermediate elevational environments

One intriguing pattern we observed was the contrasting levels of heterozygosity in the α and β-globin subunits. The α subunits of highland individuals exhibited elevated heterozygosity not seen in the lowland ducks, whereas most highland individuals were homozygous for the β^A subunit (all of them if we exclude Mendoza), independently of their habitat. Numerous possible factors might contribute to these different patterns. A simple model of strong divergent selection favouring different homozygous genotypes appears to be consistent with the pattern observed for the β^A subunit. If the α^A and α^D polymorphisms we observed confer resistance to hypoxia, heterozygous individuals would be expected to have a mixture of different Hb isoforms circulating in their bloodstream with potentially different O_2-binding properties. Such variation in physiochemical characteristics might enhance the flexibility of an organism living in a variable elevation environment.

Additionally intriguing was the observation that the highest levels of α and/or β-globin heterozygosity were in Mendoza. Considering the four amino acid replacements, seven different unique haemoglobin genotypes were found in eight crested ducks from Mendoza, and these genotypes were not sampled elsewhere in either the lowlands or the highlands (Table 2). Mendoza, for example, was the only region where we sampled crested ducks that were heterozygous for Asp/Glu-β^A94. The results of the STRUCTURE analyses furthermore indicated the presence of a steep cline in allele frequencies in this region, whereby crested ducks sampled at higher elevations (~2500 m) in the north of the province of Mendoza were genotypically more similar to the highland population (e.g. Altiplano), and those sampled in the south of the province at lower elevations (~1500 m) were more genotypically similar to the lowland population in Patagonia (Fig. 3). This pattern was evident in both the STRUCTURE analysis including the six introns (Fig. 3A), and also in the analysis including six introns plus three haemoglobin loci (Fig. 3B).

The two subspecies of crested ducks (*Lophonetta s. specularioides/alticola*) intergrade morphologically in this region (Bulgarella *et al.* 2007), where the Andes abruptly lose their mean height. The population in Mendoza thus exists in intermediate elevational habitats that are bounded on either side by low- and high-elevation habitats. The partial pressure of O_2 in suitable habitat in this region may thus vary by as much as 20–25%, and waterfowl (especially crested ducks) are abundant and common in this region. Crested ducks in this region are furthermore known to descend seasonally to lower elevations when snow accumulates in the high cordillera (Young 2005). If individuals with such polymorphisms possess a mixture of different haemoglobin isoforms with different O_2-binding properties, as suggested by the high level of heterozygosity and presence of unique genotypes in this region, such individuals might have a dispersal advantage across mid-elevation regions, or in a seasonally variable elevational environment mediated by seasonal temperature changes and snow cover.

Ecogeographic features of population structure

We found strikingly concordant similarities between geographic variation in nuclear DNA and morphometrics reported in Bulgarella *et al.* (2007). In that study, crested ducks at higher elevations in the Andes (3000–4600 m) had larger body sizes than those inhabiting the coastal and inland lowlands of Patagonia. Consistent

© 2011 Blackwell Publishing Ltd

with those differences, we found reciprocally monophyletic lineages in mtDNA and diagnosable multilocus genotypes between highland and lowland populations. Furthermore, individuals from Mendoza were intermediate in size, and based on ten morphological measurements and plumage, seven of the eight Mendoza individuals were classified as *L. s. alticola*, and only one individual (KGM 1221; collected at 2093 m in southern Mendoza) was classified as *L. s. specularioides* (Bulgarella *et al.* 2007). Likewise, multilocus data indicated that Mendoza individuals carried a mix of highland and lowland genotypes (including individuals with admixed ancestry), suggesting that the two morphotypes interbreed within that region. These concordant results between mtDNA, nuclear DNA, and morphology indicate that crested ducks comprise at least two genetic populations, corresponding to subspecies designations, but that some gene flow occurs between them.

Coalescent analyses suggest that *L. s. alticola* and *L. s. specularioides* have been diverging for a hundred thousand years or more. At least two hypotheses can explain this divergence: allopatric divergence or parapatric divergence with ecological specialization. Crested ducks are continuously distributed from northern Peru through Tierra del Fuego, where they are abundant in both the arid puna grassland environment south to the province of Mendoza, Argentina (~1500 m of elevation), and in the lowlands from Chubut, Argentina, southwards throughout the Patagonian steppe (0–950 m). However, their densities appear to substantially decrease where the subspecies' ranges are adjacent in the low-elevation Argentine provinces of Neuquén and Río Negro (<1500 m). For example, no crested ducks were observed between Pino Hachado, Neuquén, and Tecka, Chubut, despite intensive collecting expeditions (Bulgarella & McCracken pers. obs.; see also McCracken *et al.* 2009a,b,c, 2010). This region is dominated by relict conifer *Araucaria araucana* forest and widespread *Nothofagus* spp. forest that are mostly unsuitable habitat for crested ducks. Thus, the two subspecies are mostly allopatric, but a low level of contact in the intervening habitat is likely.

Based on both morphology and nuclear DNA, the samples collected in Mendoza clearly demonstrate that the two subspecies intermix, at least at a narrow contact zone centred around 2000 m of elevation. Furthermore, coalescent analyses indicate that alleles move across this region. More specifically, we detected significant evidence of gene flow into the highlands from the lowlands, despite excluding the Mendoza samples from the analysis. Both of these observations indicate that the *Aracauria* and *Nothofagus* forest likely does not present a hard barrier for crested duck dispersal. Gene flow across this region could be facilitated through high-elevation occupancy above the timberline during the summer and in ice-free alpine lakes, or via the steppe and precordillera, particularly during winter. So while there are likely few or no crested ducks in the forest, there is abundant puna-like habitat in the precordillera and above the tree line in the talus slopes, making it possible for the ducks to inhabit this region. Nevertheless, caution is required when interpreting these results because coalescent analyses assume constant migration rates following divergence. Contemporary gene flow could be more restricted than historical gene flow (or vice versa) because present-day distributions may not reflect the past distributions or vegetational characteristics of the region. Indeed, the subspecies probably diverged prior to the last glacial maximum (*c.* 18 000 BP) and perhaps prior to the last major interglacial (*c.* 120 000 BP), and these periods of major climate change probably had a large effect on crested duck distributions, restricting gene flow between the different habitats.

As an alternative to this neutral, allopatric model of divergence, our data indicate that ecological specialization for different environments and adaptive divergence might have led to population divergence. First, we found evidence of restricted gene flow in the haemoglobins relative to the putatively neutral introns. These differences support the hypothesis that the haemoglobin variants we observed are adapted to the subspecies' respective environments (presumably to the partial pressure of O_2; see above for additional details) and that selection restricts or prevents the introgression of locally adapted alleles. Second, individuals in cooler and drier environments tended to be larger in size (Aldrich & James 1991; James 1991), and therefore, the larger body size of the highland subspecies *L. s. alticola* might be adaptive (Bulgarella *et al.* 2007). Although there is strong evidence of small amounts of gene flow, selection against hybrids could explain the overall divergence between highland and lowland crested ducks under a parapatric model of speciation.

Inferences of gene flow also differed between mtDNA and nuclear introns—whereas nuclear genotypes (and morphology) transitioned in the province of Mendoza, all individuals from Mendoza harboured 'highland' mtDNA haplotypes. Therefore, mtDNA haplotypes likely transitioned somewhere farther south in the provinces of Neuquén or Río Negro. Although the absence of 'lowland' mtDNA in Mendoza could be explained by our small sample size ($n = 8$), additional evidence suggests higher rates of gene flow for nuclear introns than for mtDNA. We obtained a nonzero estimate of migration from the lowlands to the highlands for the autosomal introns, but no detectable migration in either

© 2011 Blackwell Publishing Ltd

direction for the maternally inherited mtDNA (the posterior distributions of *m* peaked near zero). This difference between markers is consistent with male-biased dispersal for crested ducks, which could explain why we did not find 'lowland' haplotypes in Mendoza. Female philopatry coupled with low population densities in the *Araucaria* and *Nothofagus* forest might limit the dispersal of mtDNA haplotypes between regions. In contrast, if males disperse greater distances, then nuclear DNA might be more likely to move across this soft barrier.

Waterfowl, in general, are unusual among birds in that females show greater philopatry to natal and breeding areas than do males (Rohwer & Anderson 1988), and sex-biased dispersal has frequently been invoked as a possible explanation for low levels of mtDNA haplotype sharing among geographic regions (Scribner *et al.* 2001; Kulikova *et al.* 2005; Peters & Omland 2007; Sonsthagen *et al.* 2009). However, this paradigm is largely based on observations of Northern Hemisphere waterfowl, but Southern Hemisphere waterfowl, and crested ducks in particular, differ in a number of life history characteristics. Whereas most Northern waterfowl are migratory, crested ducks are resident (Phillips 1922–1926; Delacour 1954) or only partially migratory (Young 2005). Additionally, ducklings have been observed year-round (Breucker *et al.* 1989), and both sexes provide parental care for prolonged periods (Buitron & Nuechterlein 1989). Likewise, histological analysis has shown that male crested ducks possess fully developed germinal epithelium during all seasons (Breucker *et al.* 1989). Therefore, male dispersal might be more constrained in crested ducks than for Northern Hemisphere ducks. Genetic data may be particularly informative about dispersal patterns in crested ducks because little data on demographics and movements are currently available.

Conclusion

This study demonstrates that haemoglobin variants, especially the β^A-globin subunit, in crested ducks inhabiting lowland and highland environments are inconsistent with neutral evolution. Using a comparative framework, McCracken *et al.* (2009a) found evidence of parallel evolution in other species of Andean waterfowl that also supported this conclusion. However, crested ducks inhabiting the different environments are morphologically and genetically differentiated, and therefore, neutral genetic drift might have resulted in the observed haemoglobin differentiation. A complex model of population history that included gene flow and time since divergence demonstrated that the β^A-globin subunit deviated significantly from expected patterns despite incorporating uncertainty in the inferred population-level parameters. In contrast to the β^A haemoglobin, male-biased dispersal has likely played a role restricting gene flow in mtDNA between subspecies, which exhibited reciprocally monophyletic lineages. Furthermore, geographic structuring of genetic variation is expected to occur, on average, more rapidly for mitochondrial than for nuclear markers. This study exemplifies the need to incorporate and account for population history when using population genetics for testing the neutrality of candidate loci. In addition, the framework that we present is applicable to any species with populations inhabiting different ecological environments.

Acknowledgements

We thank Rosario Acero, Dylan Anderson, Yanina Arzamendia, Chris Barger, Raúl Cardón, Sonia Chavarra, Raúl Clarke, Adrián Contreras, Yolanda Davies, Alan Eagle, Sonia Felton, Alejandro del Valle, Víctor Hugo Eztellan, Sergio Goldfeder, Isabel Gomez, Alejandro González, Adrián Gorrindo, Ana Huespe, Gordon Jarrell, Kevin Johnson, Darío Lijtmaer, Pamela McCracken, Kazuya Naoki, Carmen Quiroga, Daniel Ramadori, Karina Ramírez, Sandra Rivera, Roberto Salinas, Gustavo Solá, Matias Soriano, and Pablo Tubaro. We also thank the following agencies for their assistance with collecting and export permits in: Argentina (Dirección de Fauna Silvestre – Secretaría de Ambiente y Desarrollo Sustentable de la República Argentina, Museo Argentino de Ciencias Naturales, Dirección de Fauna y Flora Silvestre del Chubut, Dirección de Fauna de Santa Cruz, Dirección de Recursos Naturales de la Subsecretaría del Ambiente de Catamarca, Dirección de Recursos Naturales Renovables de Mendoza), Bolivia (Colección Boliviana de Fauna), Perú (Instituto Nacional de Recursos Naturales del Perú), the Falkland Island government and Fitzroy Farm. Jay Storz, Kevin Winker and two anonymous reviewers provided helpful comments on the manuscript. Expedition and laboratory costs were funded by the Institute of Arctic Biology at the University of Alaska Fairbanks, Alaska EPSCoR (NSF EPS-0092040 & EPS-0346770), a grant from Frank M. Chapman Fund at the American Museum of Natural History to MB and NSF grants (DEB-0444748, DEB-0926162 & IOS-0949439) to KGM and JLP.

References

Aldrich JW, James FC (1991) Ecogeographic variation in the American Robin (*Turdus migratorius*). *The Auk*, **108**, 230–249.

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Baumann R, Goldbach E, Haller EA, Wright PG (1984) Organic phosphates increase the solubility of avian haemoglobin D and embryonic chicken haemoglobin. *Biochemistry*, **217**, 767–771.

Beall CM (2003) High-altitude adaptations. *Lancet*, **362**, S14–S15.

Beall CM (2007) Two routes to functional adaptation: Tibetan and Andean high-altitude natives. *Proceedings of the National*

© 2011 Blackwell Publishing Ltd

Academy of Sciences of the United States of America, **104**, 8655–8660.

Beaumont MA (2005) Adaptation and speciation: what can *F*st tell us? *Trends in Ecology and Evolution*, **20**, 435–440.

Bernstein MH (1989) Respiration by birds at high altitude and in flight. In:*Physiology of Cold Adaptation in Birds* (eds Bech C, Reinertsen RE), pp. 197–206. Plenum, New York.

Braunitzer G, Hiebl I (1988) Molecular aspects of high-altitude respiration in birds—Hemoglobins of the bar-headed goose (*Anser indicus*), the Andean goose (*Chloephaga melanoptera*) and Ruppell's Griffon vulture (*Gyps rueppelli*). *Naturwissenschaften*, **75**, 280–287.

Breucker H, Schafer E, Bustosobregon E (1989) Spermatogenesis of *Lophonetta specularioides* (Anatidae, Aves) in the High Plateau of the Southern Andes—a Light and Electron-Microscopic Study. *Microscopía Electrónica y Biología Celular*, **13**, 167–182.

Buitron D, Nuechterlein GL (1989) Male parental care of Patagonian Crested Ducks *Anas* (*Lophonetta*) *specularioides*. *Wildfowl*, **40**, 14–21.

Bulgarella M, Wilson RE, Kopuchian C, Valqui TH, McCracken KG (2007) Elevational variation in body size of Crested Ducks (*Lophonetta specularioides*) from the Central High Andes, Mendoza, and Patagonia. *Ornitologia Neotropical*, **18**, 587–602.

Bulgarella M, Stewart NC, Fedorov VB, Moore AV, McCracken KG (2009) Hemoglobin transcript abundance in a bone marrow cDNA library of Crested Ducks (*Lophonetta specularioides*) in the Peruvian High Andes. *The Auk*, **126**, 666–672.

Chappell MA, Hayes JP, Snyder LRG (1988) Hemoglobin polymorphisms in Deer Mice (*Peromyscus maniculatus*), physiology of beta-globin variants and alpha-globin recombinants. *Evolution*, **42**, 681–688.

Cirotto C, Geraci G (1975) Embryonic chicken hemoglobins. Studies on the oxygen equilibrium of two pure components. *Comparative Biochemistry and Physiology A*, **51**, 159–163.

Delacour J (1954) *The Waterfowl of the World*, vol. 1. Arco Publishing Company, Inc., New York.

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

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Faraci FM (1991) Adaptations to hypoxia in birds: how to fly high. *Annual Review of Physiology*, **53**, 59–70.

Fjeldså J, Krabbe N (1990) *Birds of the High Andes: A Manual to the Birds of the Temperate Zone of the Andes and Patagonia, South America*. Zoological Museum, University Copenhagen, Copenhagen, Denmark.

Grubb BR (1981) Blood flow and oxygen consumption in avian skeletal muscle during hypoxia. *Journal of Applied Physiology*, **50**, 450–455.

Hardison R (2001) *Organization, Evolution, and Regulation of the Globin Genes in Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management*. Cambridge University Press, Cambridge.

Hasegawa M, Kishino H, Yano T (1985) Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hawkes LA, Balachandran S, Batbayar N *et al.* (2011) The trans-Himalayan flights of bar-headed geese (*Anser indicus*). *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 9516–9519.

Hey J (2005) On the number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS Biology*, **3**, e193.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hiebl I, Braunitzer G (1988) Adaptation of the hemoglobins of bar-headed goose (*Anser indicus*), Andean goose (*Chloephaga melanoptera*) and Ruppell's griffon (*Gyps rueppellii*) to life under hypoxic conditions. *Journal Fur Ornithologie*, **129**, 217–226.

Hiebl I, Braunitzer G, Schneeganss D (1987) The primary structures of the major and minor hemoglobin components of adult Andean Goose (*Chloephaga melanoptera*, Anatidae): the mutation Leu → Ser in position 55 of the β-chains. *Biological Chemistry Hoppe-Seyler*, **368**, 1559–1569.

Hiebl I, Weber RE, Schneeganss D, Kosters J, Braunitzer G (1988) High-altitude respiration of birds. Structural adaptations in the major and minor hemoglobin components of adult Ruppell's Griffon (*Gyps rueppellii*, Aegypiinae)—A new molecular pattern for hypoxic tolerance. *Biological Chemistry Hoppe-Seyler*, **369**, 217–232.

Hillier LW, Miller W, Birney E *et al.* (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **432**, 695–716.

Hochachka PW, Somero GN (2002) *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. Oxford University Press, New York.

Hudson RR (2002) Generating samples under a Wright-Fisher neutral model. *Bioinformatics*, **18**, 337–338.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

James FC (1991) Complementary descriptive and experimental studies of clinal variation on birds. *American Zoologist*, **31**, 694–706.

Jessen TH, Weber RE, Fermi G, Tame J, Braunitzer G (1991) Adaptation of bird hemoglobins to high altitudes: demonstration of molecular mechanism by protein engineering. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 6519–6522.

Johnsgard PA (1978) *Ducks, Geese, and Swans of the World*. University Nebraska Press, Lincoln.

Kear J (2005) *Bird Families of the World: Ducks, Geese, and Swans*. Oxford University Press, New York.

Knapp JE, Oliveira MA, Xie Q, Ernst SR, Riggs AF, Hackert ML (1999) The structural and functional analysis of the hemoglobin D component from chicken. *Journal of Biological Chemistry*, **274**, 6411–6420.

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics*, **22**, 768–770.

Kulikova IV, Drovetski SV, Gibson DD *et al.* (2005) Phylogeography of the Mallard (*Anas platyrhynchos*): hybridization, dispersal, and lineage sorting contribute to complex geographic structure. *The Auk*, **122**, 949–965.

© 2011 Blackwell Publishing Ltd

Laybourne RC (1974) Collision between a vulture and an aircraft at an altitude of 37000 ft. *Wilson Bulletin*, **86**, 461–462.

Liang YH, Liu XZ, Liu SH, Lu GY (2001a) The structure of greylag goose oxy haemoglobin: the roles of four mutations compared with bar-headed goose haemoglobin. *Acta Crystallographica Section D, Biological Crystallography*, **57**, 1850–1856.

Liang Y, Hua Z, Liang X, Xu Q, Lu G (2001b) The crystal structure of bar-headed goose hemoglobin in deoxy form: the allosteric mechanism of a hemoglobin species with high oxygen affinity. *Journal of Molecular Biology*, **313**, 123–137.

Liu XZ, Li SL, Jing H, Liang YH, Hua ZQ, Lu GY (2001) Avian haemoglobins and structural basis of high affinity for oxygen: structure of bar-headed goose aquomet haemoglobin. *Acta Crystallographica Section D, Biological Crystallography*, **57**, 775–783.

McCracken KG, Sorenson MD (2005) Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (*Nomonyx-Oxyura*)? *Systematics Biology*, **54**, 35–55.

McCracken KG, Barger CP, Bulgarella M *et al.* (2009a) Signatures of high-altitude adaptation in the major hemoglobin of five species of Andean dabbling ducks. *The American Naturalist*, **174**, 631–650.

McCracken KG, Bulgarella M, Johnson KP *et al.* (2009b) Gene flow in the face of countervailing selection: adaptation to high-altitude hypoxia in the beta-A hemoglobin subunit of yellow-billed pintails. *Molecular Biology and Evolution*, **26**, 815–827.

McCracken KG, Barger CP, Bulgarella M *et al.* (2009c) Parallel evolution in the major hemoglobin genes of eight species of Andean waterfowl. *Molecular Ecology*, **18**, 3992–4005.

McCracken KG, Barger CP, Sorenson MD (2010) Phylogenetic and structural analysis of the HbA (αA/βA) and HbD (αD/βA) hemoglobin genes in two high-altitude waterfowl from the Himalayas and the Andes: bar-headed goose (*Anser indicus*) and Andean goose (*Chloephaga melanoptera*). *Molecular Phylogenetics and Evolution*, **56**, 649–658.

Monge C, León-Velarde F (1991) Physiological adaptation to high altitude: oxygen transport in mammals and birds. *Physiological Review*, **71**, 1135–1172.

Navas JR, Bo NA (1998) The geographical distribution of the subspecies of *Lophonetta specularioides* and *Merganetta armatta* (Anatidae) in the provinces of Mendoza and San Juan, Argentina. *Hornero*, **15**, 57–59.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Niessing J, Erbil C, Neubauer V (1982) The isolation and partial characterization of linked α-A and α-D globin genes from a duck DNA recombinant library. *Gene*, **18**, 187–191.

Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: isolation by adaptation and multiple roles for divergent selection. *Evolution*, **62**, 316–336.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375–402.

Perutz MF (1970) The Bohr effect and combination with organic phosphates. *Nature*, **228**, 734–739.

Perutz MF (1983) Species adaptation in a protein molecule. *Molecular Biology and Evolution*, **1**, 1–28.

Perutz MF (1989) Mechanisms of cooperativity and allosteric regulation in proteins. *Quarterly Review of Biophysics*, **22**, 139–236.

Perutz MF (1990) Mechanisms regulating the reactions of human hemoglobin with oxygen and carbon monoxide. *Annual Review of Physiology*, **52**, 1–25.

Peters JL, Omland KE (2007) Population structure and mitochondrial polyphyly in North American gadwalls (*Anas strepera*). *The Auk*, **124**, 444–462.

Peters JL, Gretes W, Omland KE (2005) Late Pleistocene divergence between eastern and western populations of Wood Ducks (*Aix sponsa*) inferred from the 'isolation-with-migration' coalescent method. *Molecular Ecology*, **14**, 3407–3418.

Peters JL, Zhuravlev YN, Fefelov I, Logie A, Omland KE (2007) Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between Gadwall and Falcated duck (*Anas spp.*). *Evolution*, **61**, 1992–2006.

Peters JL, Zhuravlev YN, Fefelov I, Humphries EM, Omland KE (2008) Multilocus phylogeography of a Holartic duck: colonization of North America from Eurasia by Gadwall (*Anas strepera*). *Evolution*, **62**, 1469–1483.

Petit RJ, Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Phillips JC (1922–1926) *A Natural History of Ducks*. Houghton Mifflin Company, Boston.

Piiper J (1989) Gas-exchange efficiency of fish gills and bird lungs. In: *Physiological Functions in Special Environments* (eds Paganelli CV, Farhi LE), pp. 159–171. Springer-Verlag, New York.

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

Rambaut A (2007) *Se-Al 2.0a11*. University of Oxford, Oxford, UK.

Rezende EL, Gomes FR, Ghalambor CK, Russell GA, Chappell MA (2005) An evolutionary frame of work to study physiological adaptation to high altitudes. *Revista Chilena de Historia Natural*, **78**, 323–336.

Riggs AF (1998) Self-association, cooperativity and supercooperativity of oxygen binding by hemoglobins. *Journal of Experimental Biology*, **201**, 1073–1084.

Rohwer FC, Anderson MG (1988) Female-biased philopatry, monogamy, and the timing of pair formation in migratory waterfowl. In: *Current Ornithology* (ed. Johnston RF), vol. 5, pp. 187–221. Plenum Press, New York.

Rollema HS, Bauer C (1979) The interaction of inositol pentaphosphate with the hemoglobins of highland and lowland geese. *Journal of Biological Chemistry*, **254**, 12038–12043.

Rowley AF, Ratcliffe NA (1988) *Vertebrate Blood Cells*. Cambridge University Press, Cambridge.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Scheid P (1979) Mechanisms of gas exchange in bird lungs. *Reviews of Physiology, Biochemistry and Pharmacology*, **86**, 137–186.

© 2011 Blackwell Publishing Ltd

Schulenberg TS, Stotz DF, Lane DF, O'Neill JP, Parker III TA (2007) *Birds of Peru*. Princeton University Press, Princeton.

Scott GR (2011) Elevated performance: the unique physiology of birds that fly at high altitudes. *Journal of Experimental Biology*, **214**, 2455–2462.

Scott GR, Milsom WK (2007) Control of breathing and adaptation to high altitude in the bar-headed goose. *American Journal of Physiology*, **293**, R379–R391.

Scott GR, Egginton S, Richards JG, Milsom WK (2009) Evolution of muscle phenotype for extreme high altitude flight in the bar-headed goose. *Proceedings of the Royal Society of London. Series B*, **276**, 3645.

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

Shih DT, Luisi BF, Miyazaki G, Perutz MF, Nagai K (1993) A mutagenic study of the allosteric linkage of His(HC3)146β in haemoglobin. *Journal of Molecular Biology*, **230**, 1291–1296.

Snyder LRG (1981) Deer mouse hemoglobins: is there genetic adaptation to high altitude?. *BioScience*, **31**, 299–304.

Snyder LRG, Hayes JP, Chappell MA (1988) Alpha-chain hemoglobin polymorphisms are correlated with altitude in deer mouse, *Peromyscus maniculatus*. *Evolution*, **42**, 689–697.

Sonsthagen SA, Talbot SL, Lanctot RB, Scribner KT, McCracken KG (2009) Hierarchical spatial genetic structure of Common Eiders (*Somateria mollissima*) breeding along a migratory corridor. *The Auk*, **126**, 744–754.

Sorenson MD, Fleischer RC (1996) Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 15239–15243.

Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP (1999) Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular Phylogenetics and Evolution*, **12**, 105–114.

Stephens M, Smith NJ, Donelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.

Storz JF, Kelly JK (2008) Effects of spatially varying selection on nucleotide diversity and linkage disequilibrium: insights from deer mouse globin genes. *Genetics*, **180**, 367–379.

Storz JF, Sabatino SJ, Hoffmann FG *et al.* (2007) The molecular basis of high-altitude adaptation in deer mice. *PLoS Genetics*, **3**, 448–459.

Storz JF, Runck AM, Sabatino SJ *et al.* (2009) Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 14450–14455.

Storz JF, Scott JR, Cheviron ZA (2010) Phenotypic plasticity and genetic adaptation to high-altitude hypoxia in vertebrates. *Journal of Experimental Biology*, **213**, 4125–4136.

Swan LW (1961) The ecology of the high Himalayas. *Scientific American*, **205**, 68–78.

Swan LW (1970) Goose of the Himalayas. *Natural History*, **70**, 68–75.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Tucker VA (1968) Respiratory physiology of house sparrows in relation to high-altitude flight. *Journal of Experimental Biology*, **48**, 55–66.

Turek Z, Kreuzer F, Hoofd LJC (1973) Advantage or disadvantage of a decrease of blood oxygen affinity for tissue oxygen supply at hypoxia: a theoretical study comparing man and rat. *Pflügers Archiv*, **342**, 185–187.

Turek Z, Kreuzer F, Ringnalda BEM (1978) Blood gases at several levels of oxygenation in rats with a left-shifted blood oxygen dissociation curve. *Pflügers Archiv*, **376**, 7–13.

Wang HC, Liang YH, Zhu JP, Lu GY (2000) Crystallization and preliminary crystallographic studies of bar-headed goose fluoromethaemoglobin with inositol hexaphosphate. *Acta Crystallographica Section D, Biological Crystallography*, **56**, 1183–1184.

Weber RE (2007) High-altitude adaptations in vertebrate hemoglobins. *Respiratory Physiology and Neurobiology*, **158**, 132–142.

Weber RE, Fago A (2004) Functional adaptation and its molecular basis in vertebrate hemoglobins, neuroglobins and cytoglobins. *Respiratory Physiology and Neurobiology*, **144**, 141–159.

Weber RE, Lalthantluanga R, Braunitzer G (1988) Functional characterization of fetal and adult yak hemoglobins: an oxygen binding cascade and its molecular basis. *Archives of Biochemistry and Biophysics*, **263**, 199–203.

Weber RE, Jessen TH, Malte H, Tame J (1993) Mutant hemoglobins (α119-Ala and β55-Ser): functions related to high-altitude respiration in geese. *Journal of Applied Physiology*, **75**, 2646–2655.

Young G (2005) Crested Duck *Lophonetta specularioides*. In: *Ducks, Geese and Swans* (ed. Kear J), vol. 2, pp. 484–486. Oxford University Press, New York.

Zhang J, Hua ZQ, Tame JRH, Lu GY, Zhang RJ, Gu XC (1996) The crystal structure of a high oxygen affinity species of haemoglobin (bar-headed goose haemoglobin in the oxy form). *Journal of Molecular Biology*, **255**, 484–493.

M.B. is a postdoctoral fellow with broad interests in population genetics, evolutionary biology, physiological adaptation and phylogeography. This manuscript resulted from her doctoral project at the University of Alaska Fairbanks. J.L.P. is assistant professor interested in population genetics, evolutionary ecology and molecular evolution. C.K. is a postdoctoral fellow in Argentina with interest in systematics, biogeography and evolution of Neotropical birds. T.V. studies the ecology and evolution of Neotropical birds. R.E.W. is interested in waterfowl, biogeography and ecological adaptation. K.G.M. studies population genetics and evolution of waterfowl, emphasizing on South American ducks and the processes by which they have adapted to the high Andes.

© 2011 Blackwell Publishing Ltd

Data accessibility

GenBank accession numbers GQ271065–GQ271144, GQ271804–GQ271883, GQ268964–GQ269043, GQ269044–GQ269123, GQ269124–GQ269203, GQ269204–GQ269283, GQ269284–GQ269363, HM063481–HM063503, JN833791–JN833847, JN833848–JN833927, JN833928–JN834007. See Table S2 (Supporting information) for detailed accession number information relating loci with specimens.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Locality and specimen information for the 80 crested duck included in this study.

Table S2 GenBank accession numbers for each locus and individual crested duck in this study.

Fig. S1 TMRCA ($t \times \mu$) estimates rescaled to the locus-specific mutation rate (μ), expressed in years before the present (BP).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2011 Blackwell Publishing Ltd