

Stepwise colonization of the Andes by Ruddy Ducks and the evolution of novel β -globin variants

V. MUÑOZ-FUENTES,*† M. CORTÁZAR-CHINARRO,† M. LOZANO-JARAMILLO‡ and K. G. MCCracken§

*Estación Biológica de Doñana-CSIC, Avda. Américo Vespucio s/n, 41092 Sevilla, Spain, †Department of Population and Conservation Biology, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden, ‡Laboratorio de Biología Evolutiva de Vertebrados, Departamento de Ciencias Biológicas, Universidad de Los Andes, Apartado 4976, Bogotá, Colombia, §Department of Biology and Wildlife, Institute of Arctic Biology, University of Alaska Museum, University of Alaska Fairbanks, Fairbanks, Alaska 99775, USA

Abstract

Andean uplift played a key role in Neotropical bird diversification, yet past dispersal and genetic adaptation to high-altitude environments remain little understood. Here we use multilocus population genetics to study population history and historical demographic processes in the ruddy duck (*Oxyura jamaicensis*), a stiff-tailed diving duck comprising three subspecies distributed from Canada to Tierra del Fuego and inhabiting wetlands from sea level to 4500 m in the Andes. We sequenced the mitochondrial DNA, four autosomal introns and three haemoglobin genes (α^A , α^D , β^A) and used isolation-with-migration (IM) models to study gene flow between North America and South America, and between the tropical and southern Andes. Our analyses indicated that ruddy ducks dispersed first from North America to the tropical Andes, then from the tropical Andes to the southern Andes. While no nonsynonymous substitutions were found in either α globin gene, three amino acid substitutions were observed in the β^A globin. Based on phylogenetic reconstruction and power analysis, the first β^A substitution, found in all Andean individuals, was acquired when ruddy ducks dispersed from low altitude in North America to high altitude in the tropical Andes, whereas the two additional substitutions occurred more recently, when ruddy ducks dispersed from high altitude in the tropical Andes to low altitude in the southern Andes. This stepwise colonization pattern accompanied by polarized β^A globin amino acid replacements suggest that ruddy ducks first acclimatized or adapted to the Andean highlands and then again to the lowlands. In addition, ruddy ducks colonized the Andean highlands via a less common route as compared to other waterbird species that colonized the Andes northwards from the southern cone of South America.

Keywords: coalescent, haemoglobin, high-altitude, hypoxia, introns, isolation with migration, mtDNA, Patagonia, stiff-tailed ducks, waterfowl

Received 25 June 2012; revision received 18 September 2012; accepted 2 October 2012

Introduction

A lasting question in evolutionary biology is to identify factors that have led to the extraordinary diversity of tropical organisms (Dobzhansky 1950; Pianka 1966; Janzen 1967; Haffer 1969; Terborgh 1971). Mountain ranges

have played key roles generating environmental heterogeneity by creating new biotic and abiotic gradients, which in combination with climatic oscillations have had a profound effect on tropical species diversity. In South America, the Andean uplift created abundant new high-altitude habitats and fragmented both low- and high-altitude habitats. The Andes uplifted relatively recently in geological time (Gregory-Wodzicki 2000; Barke & Lamb 2006) and were strongly glaciated during

Correspondence: Violeta Muñoz-Fuentes, Fax: +34-954621125; E-mail: Violeta.Munoz@ebd.csic.es

the Pleistocene (Abbott *et al.* 2003; Paduano *et al.* 2003; Singer *et al.* 2004; Chepstow-Lusty *et al.* 2005). While some species persisted through glacial cycles in ice-free refugia (Paduano *et al.* 2003; Hanselman *et al.* 2011), others probably colonized deglaciated areas recently. Birds have been of particular interest because of their extraordinary diversity in the Neotropics as more bird species occur here than in any other region of the planet (Stotz *et al.* 1996). Neotropical birds, consequently, have featured in numerous studies that have examined their evolutionary history as well as colonization and dispersal patterns to and from the Andes (Vuilleumier 1969, 1986; Vuilleumier & Simberloff 1980; Remsen 1984; Bates & Zink 1994). The effect of the Andes as an isolating barrier on the patterns of divergence between lineages occupying the eastern and western margins of these cordillera have been of interest (Chapman 1917; Cracraft & Prum 1988; Brumfield & Capparella 1996; Cheviron *et al.* 2005; Miller *et al.* 2008), as has the role of high-elevation habitats driving species diversity (Bates & Zink 1994; Weir 2006; Brumfield & Edwards 2007; Cadena *et al.* 2007; Weir *et al.* 2008). More recently, the effects of local adaptation and the balance between selection and migration for genes or traits related to high-altitude function have been examined across elevational gradients (Cheviron & Brumfield 2009; McCracken *et al.* 2009a,b; Milá *et al.* 2009; Bulgarella *et al.* 2012; Wilson *et al.* 2012).

High-altitude habitats present environmental challenges, the most significant of which are colder temperatures, aridity and hypoxia (Beall 2007). High-altitude habitats are thus metabolically demanding, and moving between low- and high-altitude habitats is expected to be challenging for most organisms. Hypoxia is among the most important factors influencing survival and affects a wide range of traits related to metabolism and development (Hochachka 1986; Monge & León-Velarde 1991; Carey 1994; Beall *et al.* 2004; Bennett *et al.* 2008; Cheviron *et al.* 2012). For example, birds from low-altitude environments experience low egg viability at high-altitude (Visschedijk *et al.* 1980; Monge & León-Velarde 1991; Gou *et al.* 2005; Zhang *et al.* 2007), and numerous studies have revealed examples of genetic adaptations in high-altitude species (Storz *et al.* 2007; Simonson *et al.* 2010; Yi *et al.* 2010; Scott *et al.* 2011). However, how populations came to colonize high-altitude regions has been less well studied. Notably, the historical sequence by which specific genetic polymorphisms evolved can shed light on the colonization patterns of species.

For the waterbird species that are common in the Andes, Fjeldsá (1985) concluded that colonization of high-altitude regions proceeded south to north for most species, from lower altitude, steppe-like areas in southern South America to higher altitude in the puna or

páramo (the upper parts of the central or northern Andes, respectively, between the tree line and the permanent snow line, from approximately 3000 to 5000 m elevation; see also Vuilleumier 1986). However, species diversity in the region is high, and it is likely that colonization occurred in different directions from various geographic regions (Chesser 2000, 2004; Burns & Naoki 2004; Brumfield & Edwards 2007). Fjeldsá (1985) further noted that the fauna of the northern Andes, particularly in Colombia, received immigrants from North America and Central America (Fjeldsá 1994; Kattan *et al.* 2004; Cuervo *et al.* 2008). Sedano & Burns (2010) concluded that the northern Andes have been an important source of lineages for other regions, with numerous dispersals to the central Andes, as well as to low-altitude habitats on either side of the Andes (Fjeldsá & Rahbek 2006).

Here we use the isolation-with-migration (IM) model (Hey & Nielsen 2004) to test predictions about the colonization of North America and South America by the ruddy duck (*Oxyura jamaicensis*), a stiff-tailed diving duck naturally endemic to the Americas and comprised of three subspecies distributed from Canada to Tierra del Fuego and inhabiting wetlands from sea level to 4500 m in the Andes (Johnsgard & Carbonell 1996; Brua 2001; Fig. 1). The nominate subspecies *O. j. jamaicensis* inhabits North America from southern Canada to Central America and the Caribbean. Males of this subspecies are distinguished from the two South American subspecies by their uniform white cheek patches and contrasting black crowns. In South America, two subspecies of ruddy ducks inhabit wetlands confined to the Andes and precordillera. *Oxyura j. ferruginea* occurs in both high- and low-altitude wetlands from Ecuador south to Tierra del Fuego. Males of this subspecies are distinguished from *O. j. jamaicensis* by their all-black head. The third subspecies *O. j. andina* is endemic to the Colombian highlands. Males of this subspecies show variable cheek colour that ranges from completely white (e.g. *O. j. jamaicensis*) to completely black (e.g. *O. j. ferruginea*; Adams & Slavid 1984; Fjeldsá 1986). The Colombian subspecies *O. j. andina* thus represents an intermediate plumage between *O. j. jamaicensis* and *O. j. ferruginea*. Although there are no band recoveries confirming that North American ruddy ducks migrate farther south than Mexico, Guatemala or the Caribbean (Bird Banding Laboratory (Patuxent Wildlife Research Center, U.S. Geological Survey, Laurel, Maryland), Fjeldsá (1986) suggested that this may not necessarily have been the case historically and that *O. j. andina* represents a hybrid population. Ruddy ducks have excellent dispersal capabilities and, while they are migratory in North America, the South American subspecies are either resident or wander locally through the Andes. In contrast to ruddy ducks from North

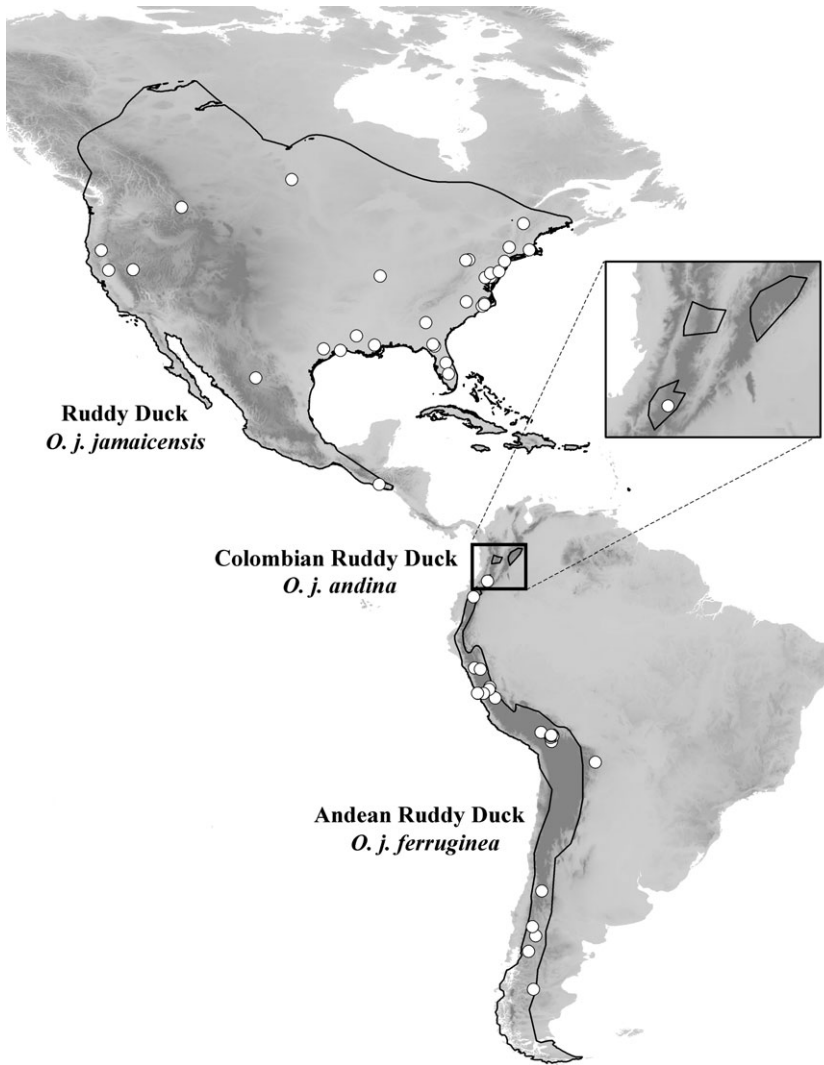


Fig. 1 Geographic range of the three ruddy duck subspecies and sampling localities in this study.

America (*jamaicensis*) that are predominantly found at low elevation sites, ruddy ducks from the tropical Andes are found at high altitude. The Andes are higher in the north than in the south and, south of latitude 32°S near the Aconcagua Massif, they abruptly lose height and transition to Patagonia (southernmost region of the Andes). Consequently, ruddy ducks inhabit an elevational range that is highest in the north and lowest in the south. With the exception of populations on the west coast of Peru, most ruddy ducks at northern latitudes in South America (*ferruginea* and *andina*) occur at high elevations (2500–4500 m), whereas most ruddy ducks in the southern Andes of Chile and Argentina (*ferruginea*) occur at lower elevations (<2500 m).

In a previously published study, McCracken & Sorenson (2005) suggested two alternative hypotheses that might explain the colonization of the Americas by ruddy ducks. If the ancestral population evolved in South America, ruddy ducks might have colonized

North America from South America. Within the context of this hypothesis (hereafter referred to as south to north colonization), it is thus possible that ruddy ducks colonized high-altitude habitats in the tropical Andes from the southern cone of South America, as appears to be the case for other waterfowl species such as dabbling ducks (Fjelds  1985; McCracken *et al.* 2009a,b; Wilson *et al.* 2012). Alternatively, ruddy ducks might have first evolved in North America and colonized South America from North America, arriving first to the high-altitude regions in the tropical Andes, and later colonizing low-altitude sites in the southern Andes in a stepwise fashion, following the longitudinal axis of the cordillera, north to south, from high to low altitudes (hereafter referred to as 'top-down', north to south colonization). From the physiological perspective, this hypothesis is intriguing because it would first require the need to acclimatize or adapt from a low-altitude environment (North America) to a high-altitude

environment (the tropical Andes), and then again to a lower-altitude environment (the southern Andes) similar to the ancestral environment in North America. Additionally, it is possible that ruddy ducks dispersing from North America colonized the tropical and southern Andes simultaneously.

To test these hypotheses, we assembled previously existing mtDNA data sets obtained throughout the range of the ruddy ducks (McCracken & Sorenson 2005; Muñoz-Fuentes *et al.* 2006) and added sequences from new specimens from different elevation zones in South America. Additionally, we also sequenced three haemoglobin genes (α^A , α^D , β^A) as well as four intron loci from ruddy ducks collected in North America and at both low- and high-altitude sites in the tropical and southern Andes. In contrast to control region mtDNA and the four introns, which are likely to be effectively neutral in the context of high-altitude function, the globin genes are potential targets of selection related to high-altitude respiration. In a study comparing the α^A and β^A haemoglobin subunit genes across major waterfowl lineages, a pattern of parallel evolution was found in highland birds, and in the case of ruddy ducks ($n = 18$) three β^A -chain haemoglobin amino acid polymorphisms differed in frequency between populations sampled at both low- and high-altitude sites in the Andes (McCracken *et al.* 2009c). To infer the historical direction of colonization (to/from South America and to/from low- and high-altitude sites within the Andes), we fitted isolation-with-migration (IM) coalescent models to the data as developed by Hey & Nielsen (2004, 2007) and Hey (2010). Using a variety of different three- and two-population models including five putatively neutral loci and three candidate loci, we calculated the rates of historical gene flow, population size parameters and the timing of divergence, and compare the latter to other waterfowl species living in the same Andean habitats. Finally, we reconstructed the historical sequence in which β^A globin amino acid substitutions evolved in low- and high-altitude populations of ruddy ducks using the inferred genealogies.

Material and methods

Sample collection

We collected samples from 111 ruddy ducks from North America ($n = 69$) and South America ($n = 42$; Fig. 1). Within South America, we sampled five *O. j. andina* from Colombia and 37 *O. j. ferruginea* from Ecuador, Peru, Bolivia and Argentina. Locality information was available for only two of the five museum specimens from Colombia, collected in the Puracé region, Departamento de Cauca. As for the *ferruginea*, 25 were collected

from high-altitude sites (>2200 m) in the tropical Andes (Ecuador, Peru, Bolivia), four were collected from low-altitude sites in the tropical Andes (875 m, Bolivia, $n = 2$, and 0 m on the coast of Peru, $n = 2$), and eight were collected at low-altitude sites (<2100 m) in the southern Andes (Argentina). Tissues were obtained from a mix of freshly collected specimens and study skins dating back to 1940 and classified to subspecies using published morphological criteria. Voucher information including geo-referenced localities are available in GenBank (accession numbers AY747742–747751, AY747756–747778, AM084943–084997, JX910949–911251).

DNA sequencing

Genomic DNA was extracted from muscle tissue or from skin and/or feathers from museum specimens using DNeasy Tissue Kits (Qiagen, Valencia, California). In the case of feathers, 30 μ l of 100 mg/ml dithiothreitol (DTT) was added to the digestion buffer. We sequenced a 575-, 900-, or 1191-bp fragment from the 5' end of the mtDNA control region for 91 recently collected specimens and a shorter 301-bp fragment for 20 museum specimens. Among these, 80 mtDNA sequences were already published, and we used previously published PCR primers and protocols for the other specimens (McCracken & Sorenson 2005; Muñoz-Fuentes *et al.* 2005, 2006). The complete coding sequences of three haemoglobin genes (α^A , α^D , β^A) that compose the definitive major (HbA) and minor (HbD) haemoglobin isoforms were PCR amplified and sequenced using primers and protocols described in McCracken *et al.* (2009b, 2010). Four additional autosomal loci (Table 2) comprising intron and adjacent exon sequences were sequenced using primers and PCR protocols described by Muñoz-Fuentes *et al.* (2007) and McCracken *et al.* (2009b). Sequences were edited using Sequencher v.4.6 (Gene Codes, Ann Arbor, MI) and aligned by eye using Se-Al v.2.0a11 (Rambaut 1996).

Population genetic analysis

To infer the gametic phase of nuclear sequences with two or more heterozygous sites, we used PHASE v.2.1, which implements a Bayesian algorithm to infer haplotypes from biallelic sequences while incorporating recombination and decay of linkage disequilibrium with genetic distance (Stephens *et al.* 2001). We used the default software settings followed by 1000 burn-in and 1000 sampling iterations ($-\times 10$), repeating the run five times for each locus from different random number seeds ($-\times 5$). The PHASE analysis resulted in high posterior allele pair probabilities for most loci; all of the allele pairs were resolved for the α^D globin, ODC1 and

FGB with $P > 0.99$. For PCK1, six individuals had allele pairs with $P = 0.88$, and the α^A globin had one allele pair with $P = 0.51$, the latter resulting from a single autapomorphy.

Detectable levels of recombination were observed in two loci, GRIN1 and the β^A globin, as indicated by the four-gamete test (Hudson & Kaplan 1985) in DnaSP v.5.10 (Librado & Rozas 2009). GRIN1 had one pair of potentially recombining sites resulting in six allele pairs with $P = 0.88$ – 0.99 and two pairs with $P = 0.50$ (see below for details describing how we treated this locus in the IM analyses). The β^A globin showed patterns consistent with high rates of recombination across the entire 1562 bp sequence (see also McCracken *et al.* 2009a,b; Bulgarella *et al.* 2012). To further analyse the β^A globin, we restricted the PHASE analysis to a subset of the full sequence including the first 625 bp on the 5' end, which contained the three observed amino acid polymorphisms in exons 1 and 2, and for which there was one pair of potentially recombining sites as indicated by the four-gamete test. Seven individuals from North America had allele pairs with $P = 0.42$ – 0.94 , and one individual had an allele pair with $P = 0.29$ in this analysis.

To illustrate genealogical relationships among haplotypes within each locus we constructed unrooted parsimony networks with TCS v.1.2.1 (Clement *et al.* 2000). For the mtDNA control region network we used the first 301 bp on the 5' end of the sequences to include all the individuals because the mtDNA sequences were shorter for 20 museum specimens (out of a total of 111). In addition, we also constructed networks for the individuals with longer sequences (575 bp, $n = 91$; 897 bp, $n = 35$), but they did not provide any information that was not evident from the shorter sequences (figures not shown). Because the sample sizes differed significantly between the northern and southern Andes, we calculated allelic richness standardized to the smallest sample size ($n = 14$ alleles for nuclear DNA and $n = 8$ haplotypes for mtDNA) using the online Rarefaction Calculator (<http://www2.biology.ualberta.ca/jbrzusto/rarefact.php>).

To test for differentiation among ruddy duck populations sampled from North America, the tropical Andes and southern Andes, we calculated Φ_{ST} using the Tamura & Nei (1993) model and Fisher's (1922) exact test in Arlequin v.3.5.1.2 (Excoffier & Lischer 2010). We also calculated Fu's F_s (Fu 1997), which evaluates the site-frequency spectrum of segregating sites. Negative values of Fu's F_s can result when there is an excess of low-frequency polymorphisms, which may result from rapid population expansion or a selective sweep acting on linked polymorphisms. Conversely, a positive value for this test statistic can be indicative of population

decline. To control for type I error in multiple comparisons we used the method of Benjamini & Hochberg (1995; also see Verhoeven *et al.* 2005). We used the software STRUCTURE (Pritchard *et al.* 2000) to test whether individuals could be accurately assigned to their sampling location (North America, tropical Andes, southern Andes) based on their eight-locus genotypes. We used the admixture model ($\alpha = 1$) with independent allele frequencies ($\lambda = 1$) and $K = 3$. The algorithm was run for 100 000 steps following a burn-in of 10 000 steps. Individuals that possessed missing data at more than one locus were excluded from this analysis ($n = 35$ for which no more than one locus was missing).

Isolation-with-migration analyses

Isolation-with-migration coalescent models can be used to understand the evolutionary processes that occur during population divergence or speciation. To infer rates of gene flow and other population parameters between North America and South America and between the tropical and southern Andes populations, we analysed our data using the isolation-with-migration (IM) coalescent model as implemented in the programmes IM (Hey & Nielsen 2004) and IMA2 (Hey 2010). These programmes, together with IMA (Hey & Nielsen 2007), fit an isolation model with migration to genetic data obtained from closely related populations or species. The softwares are not equivalent. For example, IM and IMA are limited to two populations, whereas IMA2 can analyse up to ten populations. However, as the number of populations fitted in IMA2 increases, the number of parameters scales almost exponentially and, consequently, more data (loci) are needed to infer meaningful estimates (Hey 2010). Therefore, we decided to use both a three-population model as implemented in IMA2, and a two-population model, IM. IM (unlike IMA) can accommodate changing population sizes and measure the proportion of individuals in the ancestral population that contributed to found each of the two derived populations.

The i model estimates demographic parameters, including the effective population size parameter ($\Theta = 4N_e\mu$) for each descendant population and the ancestral population(s), the gene flow between each pair of populations (M) and the divergence time (t). Because parameter estimates in coalescent genealogy samplers like IM are scaled to the mutation rate, the gene flow parameter (M) reflects the rate by which new alleles (or substitutions) are introduced to the population by immigration relative to mutation ($M = m/\mu$). The more biologically meaningful parameter corresponding to the number of effective immigrants per generation ($4N_e m$) was extracted by

multiplying Θ ($4N_e\mu$) and M (m/μ). The divergence time parameter t corresponds to the raw number of nucleotide substitutions, that is, tree height of the genealogy following divergence, whereas the time in years since two populations diverged (T) is calculated as $T = \mu t$. For each parameter, the peak value and 95% (IMa2) or 90% (IM) highest posterior density (HPD) are reported. Finally, the splitting parameter (s ; included in IM but not in IMa or IMa2) measures the fraction of individuals in the ancestral population that contributed to one population, whereas $1 - s$ is the fraction that contributed to the other population, at the time of the split (Hey 2005). In this respect, $s = 0$ or 1 indicates that one population was colonized by a small number (i.e. <1%) of individuals dispersing from the other population, whereas intermediate values of s would indicate that the ancestral population contributed more equally to the derived populations.

Given the additional number of nodes in the three-population tree, IMa2 had to estimate thirteen parameters, whereas IM had to estimate seven. Therefore, for the three-population model fitted in IMa2, we included all eight loci to maximize the amount of loci. We used the population-level phylogeny corresponding to: North America (tropical Andes, southern Andes). For the two-population IM analyses, we performed three types of analysis contrasting: (i) North America vs. South America with the tropical and southern Andes combined, (ii) North America vs. the tropical Andes and North America vs. the southern Andes, and (iii) the tropical vs. the southern Andes.

For the two-population IM analyses contrasting North American and South American populations, we treated the mtDNA, the β^A globin, and the other six nuclear loci separately. While the mtDNA is maternally inherited and the inheritance scalar (0.25 vs. 1.0) can be incorporated into IM analyses to accommodate this property of the mtDNA, evidence indicates that most species of waterfowl exhibit male-biased dispersal (Anderson *et al.* 1992; Sonsthagen *et al.* 2009; Peters *et al.* 2012a,b). Additionally, the three nonsynonymous substitutions we observed in the β^A globin of ruddy ducks make this locus a potential candidate gene possibly subject to differential selection. There was thus pre-existing rationale for treating the mtDNA and the β^A globin separately.

For the two-population IM analysis contrasting ruddy ducks in the tropical and southern Andes (North America excluded), we analysed all eight loci combined and then the mtDNA separately because the sample size and allelic diversity were limited in the southern Andes population (there were only eight mtDNA haplotypes and fourteen for the nuclear loci). Table S1 (Supporting Information) lists each three-population (IMa2) and two-population

(IM) analysis that was performed and the parameter estimates obtained.

Since the IM model does not incorporate recombination, and the inclusion of recombining sites can bias estimates of Θ and time since divergence upwards (Strasburg & Rieseberg 2010), we excluded potentially recombining blocks of sequence from our analyses. For GRIN1, we excluded the first 60 bp and truncated the sequence to positions 61–308. For the β^A globin, we used IMgc (Woerner *et al.* 2007) to selectively remove sequences from the 625-bp data to obtain a set of allele sequences consistent with no recombination, as validated by the four-gamete test. This resulted in removing five allele sequences from the North American population exhibiting potential recombination break points between sites 103 and 268.

We used the Hasegawa-Kishino-Yano model (Hasegawa *et al.* 1985) of nucleotide substitution for the mtDNA and the infinite alleles model (Kimura 1969) for nuclear loci. To account for differences in N_e due to maternal and bi-parental inheritance, we used an inheritance scalar of 0.25 for the mtDNA and 1.0 for the nuclear DNA. Analyses were first run with wide priors to explore the sensitivity of parameter estimates to different upper bounds. Then, analyses were conducted with uniform priors that encompassed the full posterior distribution of each parameter from the preliminary runs. We used a burn-in period of 500 000 steps. All runs included 20 chains with a geometric heating scheme from different random number seeds. Autocorrelation was monitored during the run, to ensure that parameter estimates converged, and all runs were allowed to proceed until the effective sample size (ESS) was $\gg 100$ for each parameter.

Results

Population genetic structure

Ruddy duck populations were significantly differentiated between North America and South America and between the tropical and southern Andes. Overall Φ_{ST} for each of the eight loci ranged from 0.11 to 0.78 (Table 1) and remained significant after correcting for multiple tests (Benjamini & Hochberg 1995). Populations that were not significantly differentiated included North America and the tropical Andes for the α^D globin; North America and the southern Andes for FGB; and the tropical and southern Andes for FGB, GRIN1 and PCK1. FGB exhibited the lowest overall Φ_{ST} , whereas the β^A globin had the highest. The mtDNA control region was highly and similarly differentiated for all pairwise comparisons ($\Phi_{ST} = 0.30$ – 0.41), and so was the β^A globin ($\Phi_{ST} = 0.70$ – 0.98). For the four

Table 1 Overall and pairwise population Φ_{ST} values for eight loci between three populations of ruddy ducks

	mtDNA	FGB	GRIN1	ODC1	PCK1	α^A	α^D	β^A
North America vs. Tropical Andes	0.38	0.11	0.39	0.55	0.27	0.28	0.05	0.70
North America vs. Southern Andes	0.41	0.08	0.37	0.79	0.20	0.44	0.75	0.79
Tropical Andes vs. Southern Andes	0.30	0.00	0.00	0.32	0.18	0.43	0.84	0.98
Overall	0.39	0.11	0.39	0.64	0.25	0.35	0.59	0.78

Overall geometric mean $\Phi_{ST} = 0.38$.

Values in bold are significant at the 0.05 level after applying the false discovery rate method for multiple comparisons (Benjamini & Hochberg 1995). The α globins (α^A or α^D) had silent substitutions, whereas three amino acid substitutions were observed in the β^A globin.

intron loci, higher Φ_{ST} values were observed between North America and South America than between the tropical and southern Andes, as the individuals from the southern Andes all possessed a single allele at each nuclear intron locus that was the same as the most frequent allele in the northern Andes. The α globins, however, had smaller Φ_{ST} values between North America and the tropical Andes than between the tropical and southern Andes, as North America and the tropical Andes shared two alleles, whereas the southern Andes had a private allele.

STRUCTURE, using the full eight loci and a subset of individuals for which no more than one locus was missing ($n = 35$), assigned all individual ruddy ducks to the population from which they were sampled with $P > 0.96$.

Nuclear intron and mtDNA alleles/haplotypes

For the nuclear intron loci, alleles from the North American and South American ruddy duck populations differed by 1–2 substitutions (Fig. 2). For FGB a single indel differentiated alleles from North America and South America, but no variation was observed within South America. For ODC1 there were two alleles in South America, both of which occurred in the tropical Andes, and one of which (the most common) was also found in the southern Andes. One copy of the most common ODC1 allele was also found in North America. For PCK1 the most common allele was found in North America and the tropical and southern Andes, but a private allele occurred in the tropical Andes. GRIN1 had three alleles in South America, one was a common allele that also occurred at low frequency in North America, the other was a private allele found only in the tropical Andes, and the remaining allele occurred in one copy in the tropical Andes but was common in North America. Thus, in all cases, more alleles were found in North America than in the tropical or southern Andes. In North America, the number of different alleles ranged from 3 to 8, of which 2–7 were private.

In the tropical Andes, 1–3 alleles were found, of which none or one was private, and in the southern Andes only one allele was found, which was the most frequent allele in the tropical Andes. Allelic richness standardized to the smallest sample size was greater in North America than in South America and in the tropical Andes than in the southern Andes (Table 2).

Similar patterns were observed for the mtDNA, but haplotype diversity was greater (Fig. 3). We found 19 haplotypes in North America and 10 in South America. Out of the nine haplotypes found in the tropical Andes, five were private. In the southern Andes, only one haplotype was observed, which was private. Of the five haplotypes from Colombia, one was identical to the most common haplotype in the tropical Andes and the other four were identical to three haplotypes found in North America; the two individuals that shared the same haplotype were sampled in the same region (Puracé). Haplotypes from two individuals (KGM 1317, 1318) from Cerro de Pasco, central highlands of Peru, were identical to the second most common North American haplotype. Finally, North America had an excess of rare singleton mtDNA haplotypes. Fu's F_s was highly significant and negative in North America ($F_s = -13.04$, $P < 0.00001$), suggesting that this population experienced a recent population expansion (see also Muñoz-Fuentes *et al.* 2006). Fu's F_s was marginally significant in the tropical Andes ($F_s = -2.89$, $P = 0.047$).

Haemoglobin alleles

The two α -globin genes did not possess nonsynonymous substitutions and exhibited patterns like other nuclear loci (Fig. 2). There were eight α^A globin alleles, four present in South America and six in North America. Two of the South American alleles were shared (in opposite proportions) with the two most common North American alleles. There were four private alleles in North America, one in the tropical Andes and one in the southern Andes. Ruddy ducks in the southern Andes possessed a private α^A globin allele that differed

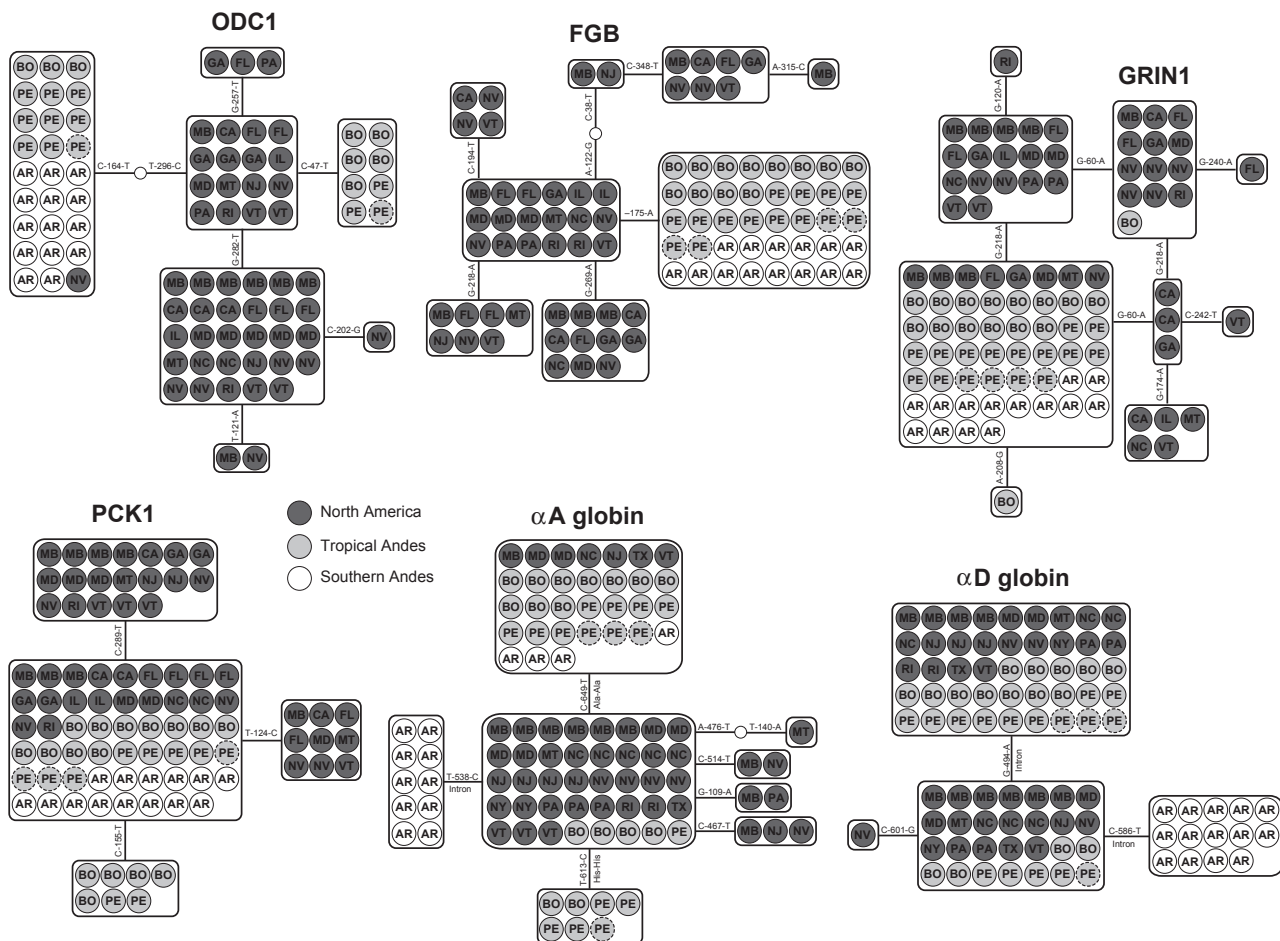


Fig. 2 Allelic networks for six nuclear loci. Most ruddy ducks from the tropical Andes (PE = Peru, BO = Bolivia) were collected at elevations >2200 m, whereas all ruddy ducks from the southern Andes (AR = Argentina) were collected at elevations <2100 m. Two ruddy ducks from the tropical Andes, indicated by dotted lines, were collected from low-altitude sites on the coast of Peru (0 m). Abbreviations for North American localities follow standard U.S. and Canadian postal codes.

Table 2 Sequence length, number of alleles sampled and allelic richness corrected for unequal sample size (\pm SD) for each locus

Locus	Length (bp)	Allelic richness				Allelic richness			
		<i>n</i>	North America	<i>n</i>	South America	<i>n</i>	Tropical Andes	<i>n</i>	Southern Andes
mtDNA control region (mtDNA)	301	69	17.5 ± 1.00	42	9	34	4.1 ± 1.00	8	1
β fibrinogen intron 7 (FGB)	396	50	6.4 ± 0.60	40	1	26	1.0 ± 0.00	14	1
N-methyl-D-aspartate-1-glutamate receptor intron 11 (GRIN1)	308	48	7.9 ± 0.30	46	3	32	1.9 ± 0.70	14	1
ornithine decarboxylase intron 5 (ODC1)	345	52	5.1 ± 0.70	34	2	20	2.0 ± 0.00	14	1
phosphoenolpyruvate carboxykinase intron 9 (PCK1)	316	48	3.0 ± 0.00	40	3	26	2.0 ± 0.03	14	1
α ^A haemoglobin	670	50	5.9 ± 0.30	46	4	32	2.9 ± 0.20	14	2
α ^D haemoglobin	867	42	3	46	3.0 ± 0.0	32	2.0 ± 0.00	14	1
β ^A haemoglobin	625*	42	15	46	2.9 ± 0.3	32	1.4 ± 0.50	14	1

*First 625 bp of the β^A haemoglobin, excluding the start codon, used for analyses (see text).

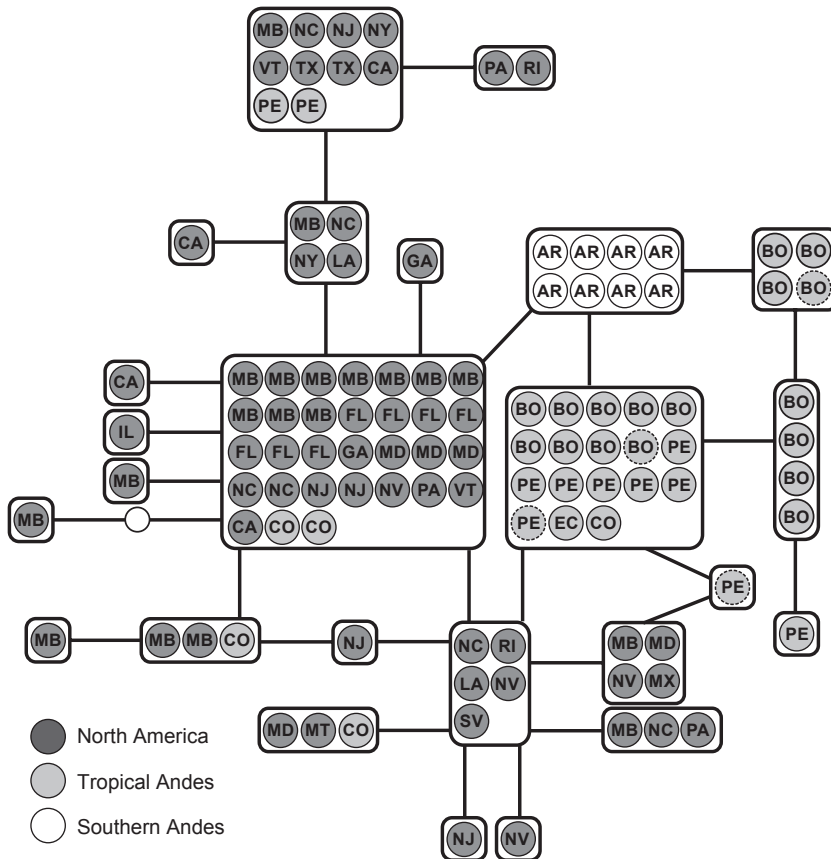


Fig. 3 Haplotype network for the mtDNA control region. Most ruddy ducks from the northern Andes (CO = Colombia, EC = Ecuador, PE = Peru, BO = Bolivia) were collected at elevations >2200 m, whereas all ruddy ducks from the southern Andes (AR = Argentina) were collected at elevations <2100 m. Four ruddy ducks from the tropical Andes, indicated by dotted lines, were collected from low-altitude sites in Bolivia (875 m; $n = 2$) and on the coast of Peru (0 m; $n = 2$). Abbreviations for North American localities follow standard U.S. and Canadian postal codes.

from the most common allele by a single base. The α^D globin possessed four alleles (Fig. 2). Two were shared between North America and the tropical Andes in opposite proportions, one was private and fixed for the southern Andean population, and the fourth was a private allele in North America.

In contrast to the α globins, the β^A globin possessed three nonsynonymous substitutions (Fig. 4). The North American population was fixed for Gly-13/Leu-14/Thr-69, the tropical Andean population differed by a single amino acid substitution and was fixed for Gly-13/Leu-14/Ser-69, and the southern Andean population differed by yet two more amino acid substitutions and was fixed Ser-13/Ile-14/Ser-69. The β^A globin exhibited a high level of variation, with 18 different alleles reconstructed from the first 625 bp on the 5' end of the sequence. Fifteen alleles occurred in North America, two in the tropical Andes and one in the southern Andes. No alleles were shared between North America and South America or between the tropical and southern Andes.

Effective population size

The effective population size parameter Θ was approximately ten times greater in North America than in the tropical Andes (three-population IMA2 analysis; Fig. 5)

or in South America (two-population IM analyses contrasting North America and South America, tropical and southern Andes combined; Fig. 6a–c). Ancestral effective population size estimates indicated that the South American populations have probably contracted or remained stable, whereas the North American population has probably expanded since population divergence (Figs 5, 6a). The Θ estimate for the southern Andes peaked sharply at zero (Figs 5, 6d) because ruddy ducks in the southern Andes possessed only a single allele at each locus, except for the α^A globin, for which there were two alleles.

Gene flow and population splitting

The IMA2 analysis (eight loci combined) provided limited information about the rate of gene flow, but was informative about the direction of gene flow. In this analysis, gene flow from North America to the tropical Andes was positive and greater than zero (Fig. S1a, Supporting Information). Gene flow in the opposite direction from the tropical Andes to North America or from the southern Andes to North America peaked sharply at zero (Fig. S1b, Supporting Information), as did gene flow from the southern to the tropical Andes (Fig. S1c, Supporting Information). Gene flow from

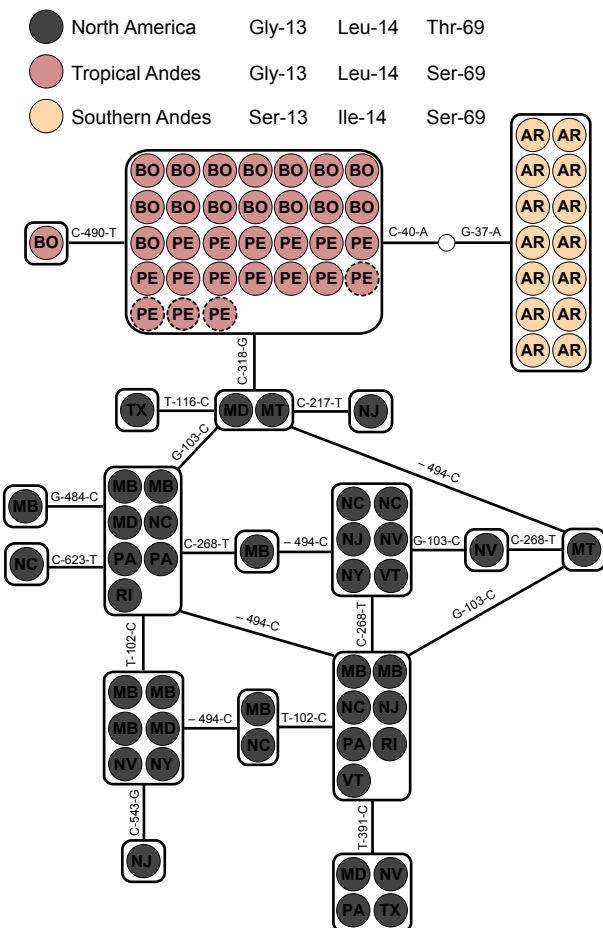


Fig. 4 Allelic network for the first 625 bp on the 5' end of the β^A globin. Colours correspond to the amino acid composition of each allele resulting from three nonsynonymous substitutions. Most ruddy ducks from the tropical Andes (PE = Peru, BO = Bolivia) were collected at elevations >2200 m, whereas all ruddy ducks from the southern Andes (AR = Argentina) were collected at elevations <2100 m. Two ruddy ducks from the tropical Andes, indicated by dotted lines, were collected from low-altitude sites on the coast of Peru (0 m). Abbreviations for North American localities follow standard U.S. and Canadian postal codes.

North America to the southern Andes and from the tropical Andes to the southern Andes could be zero or greater; the posterior distribution was flat (Fig. S1b, c, Supporting Information).

The two-population IM analyses provided more information about gene flow (Fig. 7). The six-locus nuclear DNA data set (excluding mtDNA and the β^A globin) indicated that gene flow between North America and South America has been about four times greater into South America ($m/\mu = 1.09$) than into North America ($m/\mu = 0.25$), although the posterior distributions overlapped broadly and the 90% HPDs confidence intervals overlapped zero (Fig. 7a). The mtDNA gene flow was

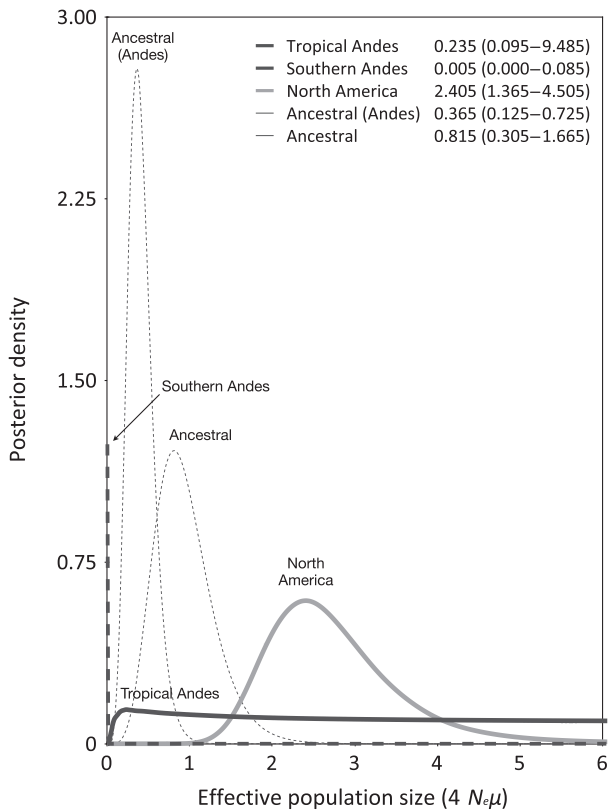


Fig. 5 Effective population size parameter estimates from the three-population IMA2 analysis of eight loci combined (95% HPD in parentheses).

asymmetric with positive (nonzero) gene flow into South America ($m/\mu = 1.23$), but zero gene flow into North America (Fig. 7b). Based on the small effective size of the recipient population in South America ($\Theta = 0.15$, 90% HPD = 0.04–0.35 for nuclear DNA and $\Theta = 3.32$, 90% HPD = 1.07–9.11 for mtDNA) it is likely that $4N_e\mu$ has been less than one effective immigrant for nuclear DNA but about four effective immigrants per generation for mtDNA. The splitting parameter (s) was not informative for the six-locus nuclear DNA data set (excluding mtDNA and the β^A globin) as no distinct peak was observed. However, the mtDNA splitting parameter peaked sharply at 1.0 for the mtDNA indicating that South America was colonized by ruddy ducks dispersing southwards from North America and that the South American population made no contribution to mtDNA in North America. In other words, <1% of the individuals in the ancestral North American population left to found the South American population (Fig. S2, Supporting Information).

To determine whether the tropical and southern Andes were colonized simultaneously or whether colonization occurred in steps (North America → tropical

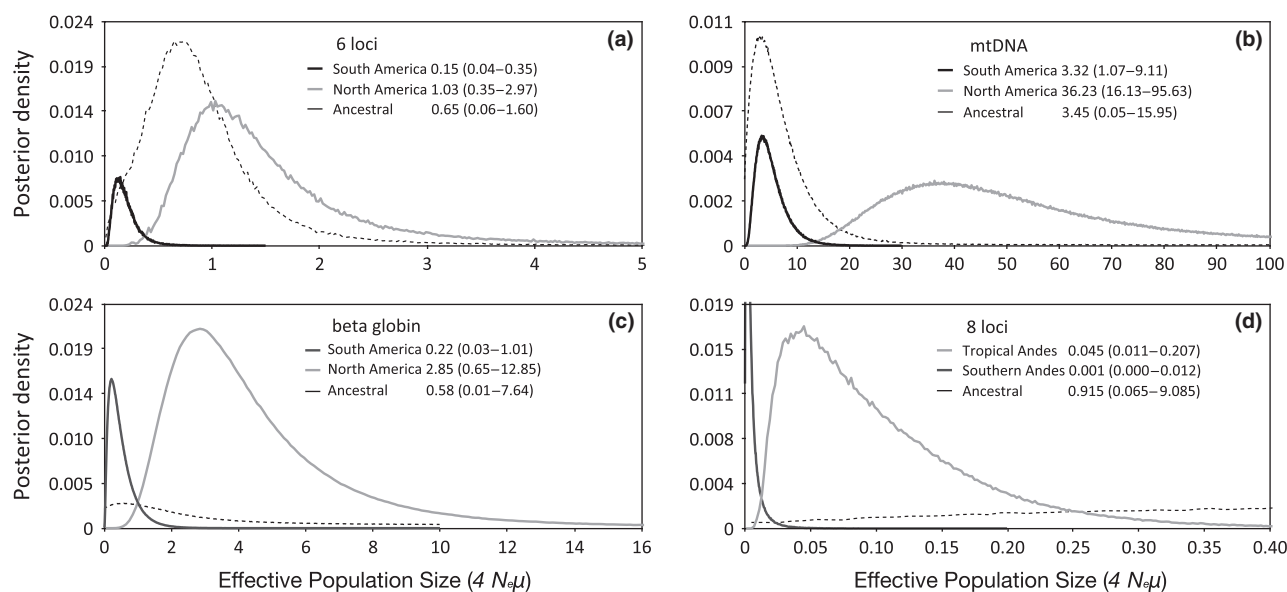


Fig. 6 Effective population size parameter estimates from the two-population IM analyses of (a–c) North America vs. South America and (d) tropical vs. southern Andes (90% HPDs in parentheses).

Andes → southern Andes), we also examined the paired gene flow estimates between North America and the tropical Andes and between North America and the southern Andes using two-population analysis in IM. For the six nuclear loci, the peak estimate of gene flow from North America to the tropical Andes was positive ($m/\mu = 1.27$; Fig. 7c), whereas the posterior distribution for gene flow from North America to the southern Andes was nearly flat, indicating that this parameter could not be estimated (Fig. 7d); thus, it is equally likely that gene flow from North America to the southern Andes could have been positive or zero. The mtDNA indicated that gene flow from North America to the tropical Andes has been positive ($m/\mu = 1.33$; Fig. 7e), whereas gene flow from North America to the southern Andes has probably been zero ($m/\mu = 0.00$; Fig. 7f). Using all eight loci combined in an additional analysis of the two Andean populations, high gene flow ($m/\mu = 17.95$) was observed from the tropical Andes into the southern Andes, whereas gene flow from the southern to the tropical Andes peaked sharply at zero (Fig. 7g). The mtDNA splitting parameter indicated that the southern Andes was colonized by ruddy ducks dispersing from the tropical Andes and that less than 1% of the ancestral tropical Andean population left to found the southern Andean population (Fig. S2, Supporting Information).

Finally, gene flow for the β^A globin peaked sharply at zero in both directions for all pairwise estimates, indicating that migration of alleles for this particular locus has been completely restricted between North America

and South America (Fig. 7h) and between the tropical and southern Andes (figure not shown but same as Fig. 7h).

Time since divergence

It is difficult to ascertain precisely the time in years at which ruddy ducks diverged between North America and South America or between the tropical and southern Andes. Substitution rates have been shown to be time dependent (Ho *et al.* 2005), with significantly higher rates over the recent evolutionary time periods in which these ruddy duck populations likely diverged. Nonetheless, it is useful to compare relative differences in the t parameter between ruddy ducks and low- and high-altitude populations of other Andean duck species using the same sets of genes.

For the IMA2 and IM comparisons contrasting ruddy duck populations in the tropical and southern Andes, time since divergence peaked sharply at or near zero, likely due to the fact that ruddy ducks in the southern Andes possessed only a single allele at all but one locus. Between North America and South America, $t/site$ peaked at 0.00049529 and 0.00029732 in the IMA2 (8-locus) and IM (6-locus) analyses, respectively (Fig. 8). While the posterior distributions from these analyses overlapped closely, the upper tail from the six-locus IM analysis was flat, suggesting that the upper bound may be overestimated. Intriguingly, we found that $t/site$ between North America and South America was similar to the $t/site$ obtained by Bulgarella *et al.* (2012;

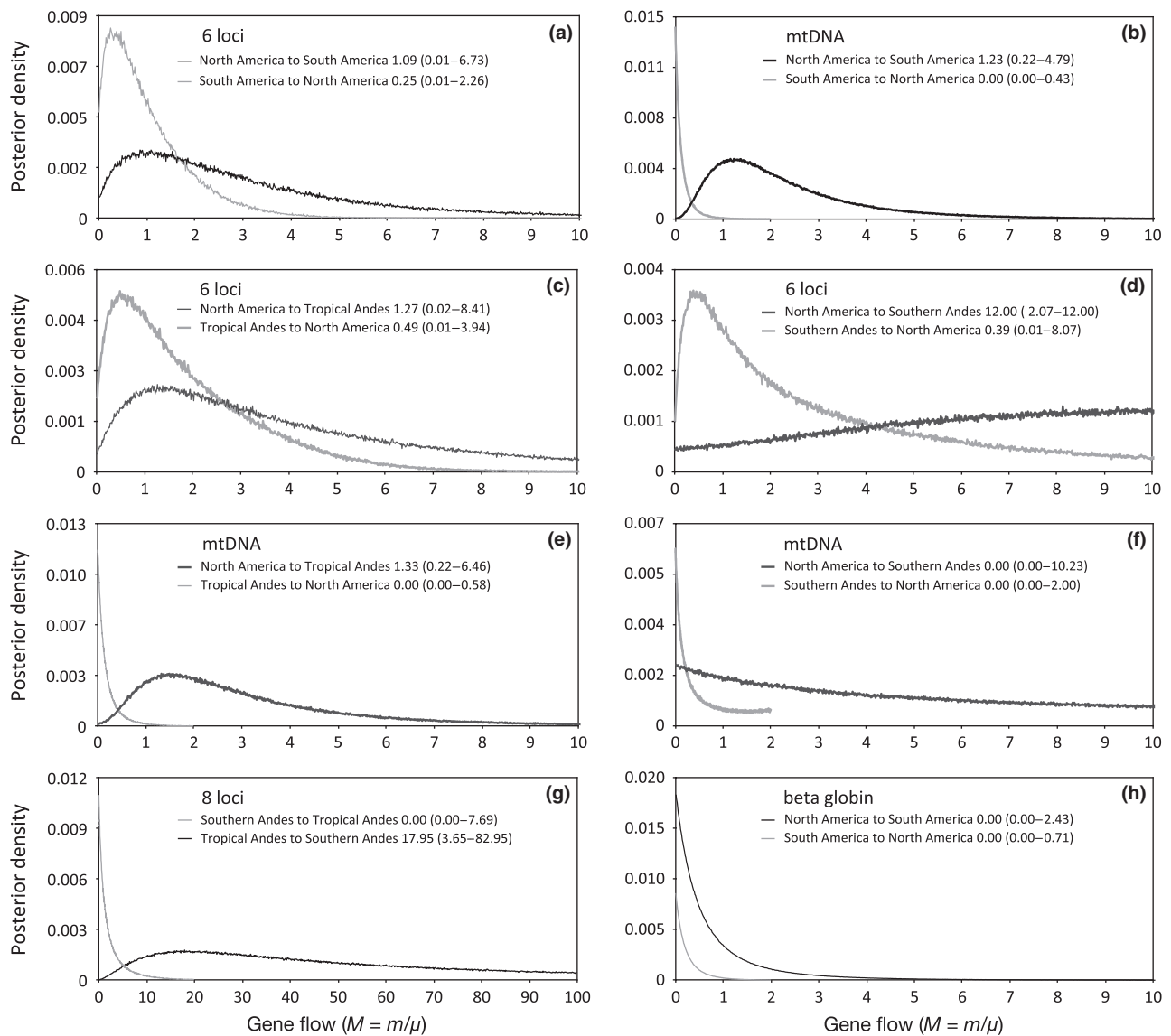


Fig. 7 Gene flow estimates from the two-population IM analyses of (a–c) North America vs. South America, (d) tropical vs. southern Andes, (e–f) North America vs. tropical Andes, and (g–h) North America vs. southern Andes (90% HPDs in parentheses).

$t/site = 0.000276102$) for an overlapping set of the same loci contrasting low- and high-altitude populations of crested ducks within the Andes. Furthermore, $t/site$, as estimated in the eight-locus IMA2 analysis, between North American and South American ruddy ducks was about 12.5 times greater ($t/site = 0.00049529$, 95% HPD = 0.00021625–0.00091849; Fig. 8) than the upper 95% HPD between the tropical and southern Andes ($t/site = 0.00000233$, 95% HPD = 0.0–0.00003953). The $t/site$ for the mtDNA control region was about five times deeper than for the six-locus nuclear analysis excluding the mtDNA. Finally, the $t/site$ for the β^A globin was bimodal with distinct peaks at 0.00034240 and 0.00100160 (Fig. 8).

Discussion

Historical population expansion from North America into South America

North American ruddy ducks probably experienced a historical population expansion as indicated by nuclear and mtDNA coalescent estimators of their current and ancestral effective population sizes (Θ) and the significant excess of low-frequency mtDNA polymorphisms (F_s) (see also Muñoz-Fuentes *et al.* 2006). In contrast, the South American population harbours an order of magnitude less genetic diversity and has probably been either stable or declined since population divergence.

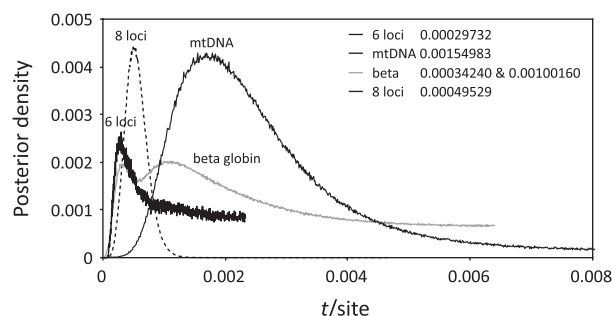


Fig. 8 Divergence ($t/site$) estimates for North America vs. South America.

These findings (effective population sizes, evidence for expansion of the North American population and greater diversity in North America as compared to South America) support the colonization of South America by ruddy ducks from North America. Moreover, two additional lines of evidence further support this inference. First, the splitting parameter s , which incorporates exponential population growth, peaked sharply at 1.0 in the mtDNA analysis indicating contribution of the North American population to South America but not the other way around (nuclear genes were uninformative). Second, gene flow from North America to South America was positive for all pairwise comparisons except the β^A globin, whereas gene flow from South America into North America was indistinguishable from zero.

Based on the data it is plausible that colonization of tropical South America coincided with the historical expansion of the ruddy duck population in North America. The time in years may be difficult to infer because substitution rates have been shown to be elevated in the short term (Shapiro *et al.* 2004; Ho *et al.* 2007; Saarma *et al.* 2007) and decline asymptotically over time scales encompassing most of the ruddy duck population history (Ho *et al.* 2005; Ho & Larson 2006). However, comparing $t/site$ to other duck species using an over-lapping set of nuclear loci including four of the same introns revealed a close correspondence between $t/site$ for ruddy ducks (North America vs. South America; $t/site = 0.000297320$) and low- and high-altitude populations of crested ducks (*Lophonetta specularioides*, $t/site = 0.000276102$ for tropical vs. southern Andes; Bulgarella *et al.* 2012). Furthermore, $t/site$ for ruddy ducks was about ten times deeper than for South Georgia pintails (*Anas g. georgica*, $t/site = 0.0000272$; K. G. McCracken, unpubl. data), which probably diverged from the continental South American pintail population at or near the LGM. Comparison to other duck species suggests that divergence between North American and South American ruddy duck populations thus likely

occurred well within or prior to the Late Pleistocene (11.7–126 kya) and far in advance of the last glacial maximum (LGM; 26.5–19 kya). While no band recoveries have been obtained from ruddy ducks farther south than Mexico, Guatemala or the Caribbean (Bird Banding Laboratory, Patuxent Wildlife Research Center, U. S. Geological Survey, Laurel, Maryland), this may not have been the case during the Pleistocene during periods of repeated glacial advance and retreat (Fjelds  1986). The timing of the inferred colonization event may thus coincide with periods when waterfowl species were probably expanding or contracting their range globally, perhaps in response to glacial cycles. Additionally, the large-scale drainage of wetlands near Bogota has caused the extinction of endemic species and affected migratory birds (Asociaci n Bogotana de Ornitiolog a 2000), such as North American waterfowl that were once wintering here. Thus, it is further plausible that gene flow between North America and South America was greater in the past and has now diminished.

Seasonal migratory behaviour of North American ruddy ducks is thus one probable factor contributing to the colonization of South America. It is likely that ruddy ducks occasionally migrate as far south as northern South America, and even into the southern cone of South America, as numerous species of North American waterfowl do so routinely (Botero & Rusch 1988). For example, KGM observed a ruddy duck with the North American male head-plumage pattern (Fig. S3, Supporting information) at Laguna de la Alumbreira, in Antofagasta de la Sierra, Catamarca, Argentina in 2005. The individual was too distant to collect, so it never became clear whether this particular individual was genotypically North American or South American, or in the case of the latter a rare plumage variant, such as those that are typical in Colombia (Adams & Slavid 1984; Fjelds  1986). Although the level of gene flow between North America and South America has probably been less than or only slightly greater than one effective migrant per generation, occasional movements, accompanied with breeding, have probably contributed to the shared haplotypes between North America, the tropical Andes and the southern Andes, and between the Colombian and the North American populations.

Finally, in a previous study McCracken & Sorenson (2005) considered *O. j. jamaicensis*, *O. j. andina*, and *O. j. ferruginea* to be subspecies in agreement with previous taxonomic studies. Wider geographic sampling presented here revealed the existence of shared mtDNA haplotypes between North America and South America in the tropical Andes, but not the southern Andes, as well as evidence for recent or ongoing introgression from North America into South America.

In this respect, the STRUCTURE analyses indicated that the three populations are well differentiated, as all individuals assigned to their sampling population with high probability ($P > 0.96$). The taxonomic implications are that while *O. j. jamaicensis* is paraphyletic and the South American population of *O. j. ferruginea* is nested within *O. j. jamaicensis* (McCracken & Sorenson 2005); there is sufficient evidence of historical and ongoing gene flow to treat all populations of ruddy ducks as a single biological species.

Stepwise colonization of the Andes

Three lines of evidence indicate that ruddy ducks did not colonize the northern and southern Andes simultaneously but rather stepwise, first from North America to the tropical Andes and then from high altitude in the tropical Andes to low altitude in the southern Andes (i.e. 'top-down', north to south, supporting one of the two biogeographic hypotheses proposed by McCracken & Sorenson 2005; see Introduction). First, as mentioned above, gene flow was higher from North America to South America than *vice versa*, and the mtDNA splitting parameter indicated that ruddy ducks colonized South America. Second, gene flow was higher from the tropical Andes to the southern Andes than *vice versa*, and the mtDNA splitting parameter indicated that ruddy ducks colonized the southern Andes from the tropical Andes. The third line of evidence is the order in which the β^A globin substitutions occurred. Ruddy ducks from North America possess Gly-13/Leu-14/Thr-69, which is the ancestral genotype shared by five other stiff-tailed duck species (McCracken *et al.* 2009c). The tropical Andean population differs from the ancestral genotype by a single nonsynonymous substitution at position 69; Gly-13/Leu-14/Ser-69 was found in all tropical Andean individuals. The southern Andean population differs by two additional nonsynonymous substitutions at tandem positions and has the Ser-13/Ile-14/Ser-69 genotype. The sequence by which these three substitutions evolved thus required a minimum of two evolutionary steps comprising three substitutions; the first substitution occurred when ruddy ducks colonized the tropical Andes and the second two substitutions occurred when ruddy ducks colonized the southern Andes. Power analysis shows that sequencing 21 ruddy duck genotypes from the North American population gives a 95% chance of detecting all β^A globin alleles occurring with a frequency >20% (Gregorius 1980). It is therefore possible but unlikely that Gly-13/Leu-14/Ser-69 represents an ancestral allele that exists at low frequency in the North American population.

Finally, we can conclude that time since divergence between tropical and southern Andes ruddy ducks was

much more recent than between North American and South American ruddy ducks, and that it differed by about one order of magnitude. Divergence between ruddy duck populations in the tropical and southern Andes could easily date to the end of the Pleistocene (12 kya) or be more recent. Alternatively, ruddy ducks could have colonized the southern Andes at an earlier date but experienced a bottleneck or extinction followed by recolonization.

Violation of coalescent model assumptions

IM and IMa2 implement the isolation-with-migration model and have a number of assumptions that, if violated, may affect parameter estimates. Because at least one of the assumptions might be violated in any natural system, Strasburg & Rieseberg (2010) examined the robustness of the IM model to violations of these assumptions. Two such assumptions are of concern in our system: presence of population structure, because of the ruddy duck comparisons from North America vs. South America, in which the tropical Andes and the southern Andes populations are significantly differentiated; and gene flow with a third unsampled population, because of the two-population IM comparisons involving at least one of the Andean populations. In the case of hidden population structure within a species, Strasburg & Rieseberg (2010) concluded that even high levels of population structure had little effect on parameter estimates. In the case of a third unsampled population, moderate levels of gene flow produced modest biases. In addition, inferences from the two-population models in IM were consistent with those from IMa2, in which all three populations were analysed together (see Table S1, Supporting Information).

Evolution of new β^A -globin genotypes

Using isolation-with-migration models and phylogenetic reconstruction of the evolution of β^A -globin variants (Fig. 4; see above), we have shown that the evolution of novel genotypic variants in ruddy ducks probably coincided with their historical range expansion from North America into South America followed by stepwise colonization of the Andes. In doing so, ruddy ducks probably acquired Ser- β^A 69 when they colonized high-altitude habitats in the tropical Andes, and then Ser- β^A 13 and Ile- β^A 14 when they colonized lower-altitude habitats in the southern Andes. This stepwise pattern of dispersal to new habitats that differ in ambient pressure would have required to adjust first to high-altitude hypoxia in the tropical Andes, and later again to normoxia in the lowlands of the southern Andes.

Numerous studies have shown that haemoglobins play an important role in mitigating the effects of high-altitude hypoxia (Perutz 1983; Jessen *et al.* 1991; Weber *et al.* 1993; Storz *et al.* 2007, 2009; Weber 2007). In the greylag goose (*Anser anser*) haemoglobin crystal structure (Liang *et al.* 2001), the three β^A -globin substitutions that we observed in ruddy ducks occur at external, solvent accessible positions on the A helix ($\beta 13$, $\beta 14$) and E helix ($\beta 69$). The first substitution (Thr \rightarrow Ser- $\beta^A 69$ on the E helix) results in the exchange of two polar amino acids that are biochemically similar; they each possess a hydroxyl. The second pair of substitutions (Gly \rightarrow Ser- $\beta^A 13$ and Leu \rightarrow Ile- $\beta^A 14$ at adjacent sites on the E helix) results in a nonpolar to polar change (with a larger R-group) plus the exchange of two nonpolar amino acids with similar biochemical properties. Intriguingly, Ser- $\beta^A 13$ also was observed to have evolved independently in high-altitude populations of yellow-billed pintails (*Anas georgica*) and speckled teal (*A. flavirostris*). No variation at this site is known to occur in any of the other seven Andean species of waterfowl studied (McCracken *et al.* 2009c).

The mechanistic effects of the observed β^A globin substitutions and whether they are neutral or beneficial in ruddy ducks remains to be determined. Using the data presented here it would be difficult to use a statistical outlier approach (e.g. Beaumont 2005; Storz 2005) to infer whether these particular haemoglobin substitutions have been targets of selection or linked to positively selected sites. The Φ_{ST} for the β^A globin, although empirically the highest among eight loci in this study, is not much greater than Φ_{ST} for some other loci. Therefore, inter-locus contrasts and/or simulations such as those employed in a variety of other recent studies in which duck globin genes generally exhibited more dramatic shifts in allele frequency than reference loci (McCracken *et al.* 2009a,b; Bulgarella *et al.* 2012; Wilson *et al.* 2012) would therefore likely be of limited utility for ruddy ducks. A final determination in the case of ruddy ducks will require measuring haemoglobin-O₂ affinity of each of the observed genotypic variants, or whether the substitutions we observed influence other properties of haemoglobin such as their solubility or epistatic interactions with other haemoglobin isoforms.

Conclusion

In a recent study of Neotropical birds, Weir (2006) found that while speciation rates in lowland habitats declined throughout the Pleistocene, speciation in the Andean highlands increased during this time period, presumably in response to the onset of glacial cycles followed by deglaciation. Dispersal from lowland habitats to the highlands was furthermore found to be about twice as

common as dispersal from the highlands to lowland habitats. In this context, the colonization of South America by ruddy ducks probably coincided with climate fluctuations and glacial cycles that caused ruddy ducks to expand their range from North America into the tropical Andes. While the exact timing and causes are unknown, these events could have been precipitated by population expansion into recently deglaciated habitats (in North America and South America) or a shifting of the North American ruddy duck wintering range to more southerly latitudes. Within South America, ruddy ducks colonized the Andes via a less common route as compared to numerous other waterbird species that likely colonized the Andes northwards from the southern cone of South America (Fjelds  1985; Vuilleumier 1986; McCracken *et al.* 2009a; Wilson *et al.* 2012). Given the *t/site* estimates we observed between the tropical and southern Andes, it is likely that ruddy ducks only recently expanded their range into the southern Andes. It is possible that this event coincided with late stages of deglaciation in Patagonia (Rabassa & Clapperton 1990; Rabassa *et al.* 2005; Turner *et al.* 2005).

To complete the biogeographical analysis of this species, the Colombian ruddy duck population would be of further interest, but sampling of this region has been limited for various reasons. Out of the five individuals studied, four had mtDNA haplotypes that were identical to haplotypes from North America, whereas the other haplotype was identical to those of other tropical Andean individuals. These data indicate that populations in the northernmost regions of the tropical Andes may be more highly admixed than elsewhere in South America, in agreement with the hypothesis that the Colombian ruddy duck is an intergrade or hybrid subspecies of the North American and the Andean ruddy duck (Todd 1979; Fjelds  1986; McCracken & Sorenson 2005).

Finally, this study illustrates how a multilocus population genetic approach accompanied by phylogenetic reconstruction of the historical sequence by which potentially functional genetic polymorphisms evolved allows to test for hypotheses of colonization and distribution of species in a biogeographic context.

Acknowledgements

We thank Rosario Acero, Hugo Asensio, Chris Barger, Bob Brua, Mariana Bulgarella, Ra l Clarke, Sergio Goldfeder, Isabel G mez, Alejandro Gonz lez, Adri n Gorrindo, Michael Grogan, Gordon Jarrell, Cecilia Kopuchian, Dario Lijtmaer, Pamela McCracken, Kazuya Naoki, Jeff Peters, Carmen Quiroga, Daniel Ramadori, Karina Ram rez, Elisa Ruiz, Gustavo Sol , Pablo Tubaro, Thomas Valqui, Alejandro del Valle, and Robert Wilson for their assistance. Field sampling was made possible by the Argentine provincial governments of Mendoza, Neuqu n,

Chubut, and Santa Cruz, Colección Boliviana de Fauna, and Instituto Nacional de Recursos Naturales (INRENA) de Peru. Shawn Houston provided high-performance computing through the University of Alaska's Life Science Informatics Portal. Logistical support was provided by Laboratorio de Sistemas de Información Geográfica y Teledetección, Estación Biológica de Doñana-CSIC (LAST-EBD). VMF was supported by a Spanish National Research Council (CSIC) JAE position and by Grupo de Excelencia en Ecología Evolutiva, Comportamiento y Conservación de la Biodiversidad, Generalitat de Catalunya, Spain. Funding to KGM was provided by the National Science Foundation (DEB-0444748 and IOS-0949439).

References

- Abbott MB, Wolfe BB, Wolfe AP *et al.* (2003) Holocene paleohydrology and glacial history of the central Andes using multiproxy lake sediment studies. *Palaeogeography Palaeoclimatology Palaeoecology*, **194**, 123–138.
- Adams J, Slavid ER (1984) Cheek plumage in the Colombian ruddy duck *Oxyura jamaicensis*. *Ibis*, **126**, 405–407.
- Anderson MG, Rhymmer JM, Rohwer FC (1992) Philopatry, dispersal, and the genetic structure of waterfowl populations. In: *Ecology and management of breeding waterfowl* (eds Batt BDJ, Afton AD, Anderson MG *et al.*), pp. 365–395. University of Minnesota Press, Minneapolis.
- Asociación Bogotana de Ornitología (2000) *Aves de la Sabana de Bogotá, Guía de Campo*. ABO, CAR, Bogotá.
- Barke R, Lamb S (2006) Late Cenozoic uplift of the Eastern Cordillera, Bolivian Andes. *Earth and Planetary Science Letters*, **249**, 350–367.
- Bates J, Zink RM (1994) Evolution into the Andes: molecular evidence for species relationships in the genus *Leptopogon*. *Auk*, **111**, 507–515.
- Beall CM (2007) Two routes to functional adaptation: tibetan and Andean high-altitude natives. *Proceedings of the National Academy of Sciences, USA*, **104**, 8655–8660.
- Beall CM, Song K, Elston RC, Goldstein MC (2004) Higher offspring survival among Tibetan women with high oxygen saturation genotypes residing at 4,000 m. *Proceedings of the National Academy of Sciences, USA*, **101**, 14300–14304.
- Beaumont MA (2005) Adaptation and speciation: what can *Fst* tell us? *Trends in Ecology and Evolution*, **20**, 435–440.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, **57**, 289–300.
- Bennett A, Sain SR, Vargas E, Moore LG (2008) Evidence that parent-of-origin affects birth-weight reductions at high altitude. *American Journal of Human Biology*, **20**, 592–597.
- Botero JE, Rusch DH (1988) Recoveries of North American waterfowl in the Neotropics. In: *Waterfowl in winter* (ed Weller MW), pp. 469–482. University of Minnesota Press, Minneapolis, Minnesota.
- Brua RB (2001) Ruddy duck (*Oxyura jamaicensis*). In: *Birds of North America*, No. 696. (eds Poole A, Gill A) B, Philadelphia, USA.
- Brumfield RT, Capparella AP (1996) Historical diversification of birds in northwestern South America: a molecular perspective on the role of vicariant events. *Evolution*, **50**, 1607–1624.
- Brumfield RT, Edwards SV (2007) Evolution into and out of the Andes: a Bayesian analysis of historical diversification in *Thammodon antshrikes*. *Evolution*, **61**, 346–367.
- Bulgarella M, Peters JL, Kopuchian C, Valqui T, Wilson R (2012) Multilocus coalescent analysis of haemoglobin differentiation between low- and high-altitude populations of crested ducks (*Lophonetta specularioides*). *Molecular Ecology*, **21**, 350–368.
- Burns KJ, Naoki K (2004) Molecular phylogenetics and biogeography of Neotropical tanagers in the genus *Tangara*. *Molecular Phylogenetics and Evolution*, **32**, 838–854.
- Cadena CD, Klicka J, Ricklefs RE (2007) Evolutionary differentiation in the Neotropical montane region: molecular phylogenetics and phylogeography of Buarremon brush-finches (Aves, Emberizidae). *Molecular Phylogenetics and Evolution*, **44**, 993–1016.
- Carey C (1994) Structural and physiological differences between montane and lowland avian eggs and embryos. *Journal of BioScience*, **19**, 429–440.
- Chapman FM (1917) The distribution of bird-life in Colombia. *Bulletin of American Museum of Natural History*, **36**, 1–728.
- Chepstow-Lusty A, Bush MB, Frogley MR, Baker PA, Fritz SC, Aronson J (2005) Vegetation and climate change on the Bolivian Altiplano between 108,000 and 18,000 yr ago. *Quaternary Research*, **63**, 90–98.
- Chesser RT (2000) Evolution in the high Andes: the phylogenetics of *Muscisaxicola* ground-tyrants. *Molecular Phylogenetics and Evolution*, **15**, 369–380.
- Chesser RT (2004) Systematics, evolution, and biogeography of the South American ovenbird genus *Cinclodes*. *Auk*, **121**, 752–766.
- Cheviron ZA, Brumfield RT (2009) Migration-selection balance and local adaptation of mitochondrial haplotypes in Rufous-collared Sparrows (*Zonotrichia capensis*) along an elevational gradient. *Evolution*, **63**, 1593–1605.
- Cheviron ZA, Hackett SJ, Capparella AP (2005) Complex evolutionary history of a Neotropical lowland forest bird (*Lepidothrix coronata*) and its implications for historical hypotheses of the origin of Neotropical avian diversity. *Molecular Phylogenetics and Evolution*, **36**, 338–357.
- Cheviron ZA, Bachman GC, Connaty AD, McClelland GB, Storz JF (2012) Regulatory changes contribute to the adaptive enhancement of thermogenic capacity in high-altitude deer mice. *Proceedings of the National Academy of Sciences, USA*, **109**, 8635–8640.
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.
- Cracraft J, Prum RO (1988) Patterns and processes of diversification: speciation and historical congruence in some Neotropical birds. *Evolution*, **42**, 603–620.
- Cuervo AM, Pulgarín PC, Calderón D (2008) New distributional bird data from the Cordillera Central of the Colombian Andes, with implications for the biogeography of northwestern South America. *The Condor*, **110**, 526–537.
- Dobzhansky T (1950) Evolution in the tropics. *American Scientist*, **38**, 209–221.
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

- Fisher RA (1922) On the interpretation of χ^2 from contingency tables, and the calculation of P . *Journal of the Royal Statistical Society*, **85**, 87–94.
- Fjeldså J (1985) Origin, evolution and status of the avifauna of Andean wetlands. *Ornithological Monographs*, **36**, 85–112.
- Fjeldså J (1986) Color variation in the ruddy duck (*Oxyura jamaicensis andina*). *Wilson Bulletin*, **98**, 592–594.
- Fjeldså J (1994) Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodiversity and Conservation*, **3**, 207–226.
- Fjeldså J, Rahbek C (2006) Diversification of tanagers, a species rich bird group, from lowlands to montane regions of South America. *Integrative and Comparative Biology*, **46**, 72–81.
- Fu Y-X (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.
- Gou X, Li N, Lian L, Yan D, Zhang H, Wu C (2005) Hypoxia adaptation and hemoglobin mutation in Tibet chick embryo. *Science in China Series C: Life Sciences*, **48**, 616–623.
- Gregorius HR (1980) The probability of losing an allele when diploid genotypes are sampled. *Biometrics*, **36**, 643–652.
- Gregory-Wodzicki KM (2000) Uplift history of the Central and Northern Andes: a review. *Geological Society of America Bulletin*, **112**, 1091–1105.
- Haffer J (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131–137.
- Hanselman JA, Bush MB, Gosling WD, Collins A, Knox C (2011) A 370,000-year record of vegetation and fire history around Lake Titicaca (Bolivia/Peru). *Palaeogeography, Palaeoclimatology, Palaeoecology*, **305**, 201–214.
- Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.
- 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 (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905–920.
- 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.
- Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences, USA*, **104**, 2785–2790.
- Ho SYW, Larson G (2006) Molecular clocks: when times are a-changin'. *Trends in Ecology and Evolution*, **22**, 79–83.
- Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561–1568.
- Ho SYH, Shapiro B, Phillips MJ, Cooper A, Drummond AJ (2007) Evidence for time dependency of molecular rate estimates. *Systematic Biology*, **56**, 515–522.
- Hochachka PW (1986) Defense strategies against hypoxia and hypothermia. *Science*, **17**, 234–241.
- 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.
- Janzen DH (1967) Why mountain passes are higher in the tropics. *American Naturalist*, **101**, 233–249.
- 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, USA*, **88**, 6519–6522.
- Johnsgard PA, Carbonell M (1996) *Ruddy ducks and other stiff-tails: their behavior and biology*. University of Oklahoma Press, Norman, Oklahoma.
- Kattan GH, Franco P, Rojas V, Morales G (2004) Biological diversification in a complex region: a spatial analysis of faunistic diversity and biogeography of the Andes of Colombia. *Journal of Biogeography*, **31**, 1829–1839.
- Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, **61**, 893–903.
- Liang YH, Liu XZ, Liu SH, Lu GY (2001) 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.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- 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*)? *Systematic 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. *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 β A hemoglobin subunit of yellow-billed pintails in the Andes. *Molecular Biology and Evolution*, **26**, 815–827.
- McCracken KG, Barger CP, Bulgarella M *et al.* (2009c) Parallel evolution in the major haemoglobin 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 (α A/ β A) and HbD (α D/ β A) 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.
- Milá B, Wayne RK, Fitze P, Smith TB (2009) Divergence with gene flow and fine-scale phylogeographical structure in the wedge-billed woodcreeper, *Glyphorhynchus spirurus*, a Neotropical rainforest bird. *Molecular Ecology*, **18**, 2979–2995.
- Miller MJ, Bermingham E, Klicka J (2008) Out of Amazonia again and again: episodic crossing of the Andes promotes diversification in a lowland forest flycatcher. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 1133–1142.
- Monge C, León-Velarde F (1991) Physiological adaptation to high altitude: oxygen transport in mammals and birds. *Physiological Reviews*, **71**, 1135–1172.
- Muñoz-Fuentes V, Green AJ, Negro JJ, Sorenson MD (2005) Population structure and loss of genetic diversity in the endangered white-headed duck, *Oxyura leucocephala*. *Conservation Genetics*, **6**, 999–1015.

- Muñoz-Fuentes V, Green AJ, Sorenson MD, Negro JJ, Vilà C (2006) The ruddy duck *Oxyura jamaicensis* in Europe: natural colonisation or human introduction? *Molecular Ecology*, **15**, 1441–1453.
- Muñoz-Fuentes V, Vilà C, Green AJ, Negro JJ, Sorenson MD (2007) Hybridization between white-headed ducks and introduced ruddy ducks in Spain. *Molecular Ecology*, **16**, 629–638.
- Paduano GM, Bush MB, Baker PA, Fritz SC, Seltzer GO (2003) A vegetation and fire history of Lake Titicaca since the Last Glacial Maximum. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **194**, 259–279.
- Perutz MF (1983) Species adaptation in a protein molecule. *Molecular Biology and Evolution*, **1**, 1–28.
- Peters JL, Thieman K, Pearce JM (2012a) Behavioral vs. molecular sources of conflict between nuclear and mitochondrial DNA: the role of male-biased dispersal in a Holarctic sea duck. *Molecular Ecology*, **21**, 3562–3575.
- Peters JL, Roberts TE, Winker K, McCracken KG (2012b) Heterogeneity in genetic diversity among non-coding loci fails to fit neutral coalescent models of population history. *PLoS ONE*, **7**, e31972.
- Pianka ER (1966) Latitudinal gradients in species diversity: a review of concepts. *American Naturalist*, **100**, 33–46.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Rabassa J, Clapperton CM (1990) Quaternary glaciations of the Southern Andes. *Quaternary Science Reviews*, **9**, 153–174.
- Rabassa J, Coronato AM, Salemmé M (2005) Chronology of the Late Cenozoic Patagonian glaciations and their correlation with biostratigraphic units of the Pampean region (Argentina). *Journal of South American Earth Sciences*, **20**, 81–103.
- Rambaut A (1996) *Se-Al: Sequence Alignment Editor*. Department of Zoology, University of Oxford, Oxford, UK.
- Remsen JV Jr (1984) High incidence of “leap frog” pattern of geographic variation in Andean birds: implications for the speciation process. *Science*, **224**, 171–173.
- Saarma U, Ho SY, Pybus OG *et al.* (2007) Mitogenetic structure of brown bears (*Ursus arctos* L.) in northeastern Europe and a new time frame for the formation of European brown bear lineages. *Molecular Ecology*, **16**, 401–413.
- Scott GR, Schulte PM, Egginton S, Scott ALM, Richards JG, Milsom WK (2011) Molecular evolution of cytochrome c oxidase underlies high-altitude adaptation in the bar-headed goose. *Molecular Biology and Evolution*, **28**, 351–363.
- Sedano RE, Burns KJ (2010) Are the Northern Andes a species pump for Neotropical birds? Phylogenetics and biogeography of a clade of Neotropical tanagers (Aves: Thraupini). *Journal of Biogeography*, **37**, 325–343.
- Shapiro B, Drummond AJ, Rambaut A *et al.* (2004) Rise and fall of Beringian steppe bison. *Science*, **306**, 1561–1565.
- Simonson TS, Yang Y, Huff CD (2010) Genetic evidence for high-altitude adaptation in Tibet. *Science*, **329**, 72–75.
- Singer BS, Ackert RP Jr, Guillou H (2004) 40Ar/39Ar and K-Ar chronology of Pleistocene glaciations in Patagonia. *Geological Society of America Bulletin*, **116**, 434–450.
- Sonsthagen SA, Talbot SL, Lancot RB, Scribner K, McCracken KG (2009) Hierarchical spatial genetic structure of Common Eiders (*Somateria mollissima*) breeding along a migratory corridor. *Auk*, **126**, 744–754.
- Stephens M, Smith NJ, Donnelly 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, 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 (2009) Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. *Proceedings of the National Academy of Sciences, USA*, **106**, 14450–14455.
- Stotz DF, Fitzpatrick JW, Parker TA III, Moskovits DK (1996) *Neotropical birds: ecology and conservation*. University of Chicago Press, Chicago.
- Strasburg JL, Rieseberg LH (2010) How robust are “Isolation with Migration” analyses to violations of the IM model? A simulation study. *Molecular Biology and Evolution*, **27**, 297–310.
- 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.
- Terborgh J (1971) Distribution on environmental gradients: theory and a preliminary interpretation of distributional patterns in the avifauna of the Cordillera Vilcabamba, Peru. *Ecology*, **52**, 23–40.
- Todd FS (1979) *Waterfowl: ducks, geese, and swans of the world*. Sea World Press, San Diego, California.
- Turner KJ, Fogwill CJ, McCulloch RD, Sugden DE (2005) Deglaciation of the eastern flank of the North Patagonia Icefield and associated continental-scale lake diversions. *Geografiska Annaler*, **87A**, 363–374.
- Verhoeven KJF, Simonsen KL, McIntyre LM (2005) Implementing false discovery rate control: increasing your power. *Oikos*, **108**, 643–647.
- Visschedijk AHJ, Rahn A, Ar H, Piiper J (1980) The independent effects of atmospheric pressure and oxygen partial pressure on gas exchange of the chicken embryo. *Respiratory Physiology and Neurobiology*, **39**, 33–44.
- Vuilleumier F (1969) Pleistocene speciation in birds living in the high Andes. *Nature*, **223**, 1179–1180.
- Vuilleumier F (1986) Origins of the tropical avifaunas of the high Andes. In: *High altitude tropical biogeography* (eds Vuilleumier F & Monasterio M), pp. 586–622. Oxford University Press, New York.
- Vuilleumier F, Simberloff D (1980) Ecology versus history as determinants of patchy and insular distributions in high Andean birds. *Evolutionary Biology*, **12**, 235–379.
- Weber RE (2007) High-altitude adaptations in vertebrate hemoglobins. *Respiratory Physiology and Neurobiology*, **158**, 132–142.
- Weber RE, Jessen TH, Malte H, Tame J (1993) Mutant hemoglobins ($\alpha 119$ -Ala and $\beta 55$ -Ser) – functions related to high-altitude respiration in geese. *Journal of Applied Physiology*, **75**, 2646–2655.
- Weir JT (2006) Divergent timing and patterns of species accumulation in lowland and highland Neotropical birds. *Evolution*, **60**, 842–855.
- Weir JT, Bermingham E, Miller MJ, Klicka J, Gonzalez MA (2008) Phylogeography of a morphologically diverse Neotropical montane species, the common bush-tanager (*Chlorospingus ophthalmicus*). *Molecular Phylogenetics and Evolution*, **47**, 650–664.
- Wilson RE, Peters JL, McCracken KG (2012) Genetic and phenotypic divergence between low- and high-altitude

populations of two recently diverged Cinnamon Teal subspecies. *Evolution*, in press, doi:10.1111/j.1558-5646.2012.01740.x.

Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851–1853.

Yi X, Liang Y, Huerta-Sanchez E (2010) Sequencing of 50 human exomes reveals adaptation to high altitude. *Science*, **329**, 75–78.

Zhang LF, Lian LS, Zhao CJ, Li JY, Bao HG, Wu C (2007) Expression pattern of HIF1 α mRNA in brain, heart and liver tissues of Tibet chicken embryos in hypoxia revealed with quantitative real-time PCR. *Animal*, **1**, 1467–1471.

Conceived and designed the study, coordinated sample collection, prepared DNA: V.M.F. and K.G.M. Performed the experiments: V.M.F., M.C.C. and K.G.M. Analysed the data: V.M.F., M.C.C., M.L.J. and K.G.M. Contributed reagents /materials / analysis tools: V.M.F. and K.G.M. Wrote the paper: V.M.F. and K.G.M. All authors read and approved the paper.

Data accessibility

DNA sequences: GenBank accessions numbers AY747742–747751, AY747756–747778, AM084943–084997,

JX910949–911251. Individual sequence information and other data files (e.g. NEXUS; IM, Structure and Arlequin input files): Dryad accession doi: 10.5061/dryad.jh4dr.

Supporting information

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

Fig. S1 Gene flow parameter estimates from the three-population IMA2 analysis (95% HPDs in parentheses).

Fig. S2 Splitting parameter estimates for (a) North America vs. South America and (b) Tropical vs. Southern Andes.

Fig. S3 Photograph of ruddy duck with North American head-plumage pattern at Laguna de la Alumbreira, in Antofagasta de la Sierra, Catamarca, Argentina on 4 November 2005.

Table S1 IMA2 and IM output of parameter estimates and their 95% and 90% credible intervals, respectively.