Gene Flow in the Face of Countervailing Selection: Adaptation to High-Altitude Hypoxia in the βA Hemoglobin Subunit of Yellow-Billed Pintails in the Andes

Kevin G. McCracken,† Mariana Bulgarella,*† Kevin P. Johnson,‡ Mary K. Kuhner,§ Jorge Trucco,|| Thomas H. Valqui,¶# Robert E. Wilson,*† and Jeffrey L. Peters*†*

*Department of Biology and Wildlife, Institute of Arctic Biology, University of Alaska Fairbanks; †University of Alaska Museum, University of Alaska Fairbanks; ‡Illinois Natural History Survey, Champaign, IL; §Department of Genome Sciences, University of Washington; ||Patagonia Outfitters, Perez 662, San Martin de los Andes, Neuquén, Argentina; ¶Centro de Ornitología y Biodiversidad (CORBIDI), Urbana Huertos de San Antonio, Surco, Lima, Perú; and #Museum of Natural Science, Louisiana State University

When populations become locally adapted to contrasting environments, alleles that have high fitness in only one environment may be quickly eliminated in populations adapted to other environments, such that gene flow is partly restricted. The stronger the selection, the more rapidly immigrant alleles of lower fitness will be eliminated from the population. However, gene flow may continue to occur at unlinked loci, and adaptive divergence can proceed in the face of countervailing gene flow if selection is strong relative to migration ($s > m$). We studied the population genetics of the major hemoglobin genes in yellow-billed pintails (*Anas georgica*) experiencing contrasting partial pressures of oxygen in the Andes of South America. High gene flow and weak population subdivision were evident at seven putatively neutral loci in different chromosomal linkage groups. In contrast, amino acid replacements (Ser-β13, Ser-β116, and Met-β133) in the βA hemoglobin subunit segregated by elevation between lowland and highland populations with significantly elevated F_{ST}. Migration rates for the βA subunit alleles were approximately 17–24 times smaller than for five unlinked reference loci, the αA hemoglobin subunit (which lacks amino acid replacements) and the mitochondrial DNA control region. The βA subunit alleles of yellow-billed pintails were half as likely to be transferred downslope, from the highlands to the lowlands, than in the opposite direction upslope. We hypothesize that migration between the lowlands and highlands is restricted by local adaptation, and the βA hemoglobin subunit is a likely target of selection related to high-altitude hypoxia; however, gene flow may be sufficiently high to retard divergence at most unlinked loci. Individuals homozygous for lowland alleles may have relatively little difficulty dispersing to the highlands initially but may experience long-term fitness reduction. Individuals homozygous for highland genotypes, in contrast, would be predicted to have difficulty dispersing to the lowlands if their hemoglobin alleles confer high oxygen affinity, predicted to result in chronic erythrocytosis at low elevation. Heterozygous individuals may have a dispersal advantage if their hemoglobin has a wider range of function due to the presence of multiple protein isoforms with a mixture of different oxygen affinities.

Introduction

Adaptation and population differentiation are genic processes (Wu 2001; Wu and Ting 2004). Individual genes or polymorphisms are the heritable units of adaptation, because recombination decouples linked genes, and different genes occur on different linkage groups. When populations become locally adapted to contrasting environments, alleles entering a new population may be eliminated by selection, such that gene flow is partly restricted. The stronger the selection, the more rapidly immigrant alleles of lower fitness will be eliminated from the population (Maynard Smith and Haigh 1974), thereby reducing effective migration rates and increasing the time to coalescence (Charlesworth et al. 1997). However, gene flow will continue to occur at unlinked loci until populations become reproductively isolated, and most of the genome may be undifferentiated if dispersal rates and gene flow levels are high. Identifying population subdivision under these circumstances is problematic because only one or a few positively selected loci may be responsible for adaptive divergence.

Several recent studies have identified compelling examples of the molecular basis of adaptation (Hoekstra et al. 2004; Colosimo et al. 2005; Storz et al. 2007); however, such studies are still relatively uncommon. It is difficult to study adaptive evolution at the genetic level when the loci that have responded to positive selection are unknown (Latta 2008). Identifying the signature of selection in one or a few genes is fundamentally difficult (Hedrick 1999), and linkage disequilibrium (LD) complicates efforts to identify differentially selected loci using outlier methods (Beaumont 2005; Storz 2005).

The genetic, ecological, and physiological factors that explain how gene flow persists in the face of countervailing selection have seldom been studied (but see Endler 1973; May et al. 1975; Hoekstra et al. 2004). Individuals with genotypes locally adapted to one environment may be maladapted to contrasting environments and experience reduced fitness, thus diminishing gene flow. Some genotypes may be relatively fit even in the wrong environment, whereas others may experience more significant difficulties (Bamshad and Wooding 2003). Gene flow can consequently be asymmetric, with some genotypes able to move more readily than others. Locally adapted populations experiencing strong selection on a well-defined adaptive trait in the face of high gene flow offer an unusually good opportunity to integrate population genetics and physiological genetics.

We studied the population genetics of the major hemoglobin genes in yellow-billed pintails (*Anas georgica*) inhabiting the Andes of South America. These ducks occur from sea level to 5,000-m elevation in the Andes and are abundant in the lowlands and highlands but are uncommon

Key words: altiplano, *Anas georgica*, Anatidae, ducks, elevation, migration, oxygen binding and delivery, Patagonia, population genetics, puna, South America, waterfowl.

E-mail: fnkgm@uaf.edu.

Mol. Biol. Evol. 26(4):815–827. 2009
doi:10.1093/molbev/msp007
Advance Access publication January 16, 2009

on the intermediate slopes of the high Andes because of the scarcity of wetlands at mid elevations. Hypoxia is debilitating at upper elevations (O_2 pressure is ~50% of sea level at 5,000 m), and the major hemoglobin protein is a likely target of selection, potentially providing a simple mechanism for local adaptation involving two codominant loci. Yellow-billed pintail populations likely exhibit some isolation by distance between the lowlands and highlands, but show frequent long-distance dispersal, consistent with elevated gene flow at many unlinked loci. We used coalescent models to measure gene flow between lowland and highland populations and compared the geographic distribution of major hemoglobin alleles with putatively neutral reference loci located on six other chromosomal linkage groups. Here, we explore these contrasting patterns using coalescent methods and simulated data sets and articulate a physiological genetic hypothesis that predicts how high levels of gene flow might be maintained in the face of countervailing selection on the βA subunit of the major hemoglobin.

High Oxygen Affinity Hemoglobins in Waterfowl

The major hemoglobin protein of adult birds consists of two αA subunits and two βA subunits, coded on chromosomes 14 and 1, respectively, in the chicken genome (Hillier et al. 2004). Reversible oxygen binding and delivery is achieved by small changes in the tertiary structure at the hemes and a large change in quaternary structure, resulting in a rotation and translation of one $\alpha\beta$ dimer relative to the other (Perutz 1983). The deoxy or tense (T-state) structure has a low affinity for oxygen and high affinity for heterotropic ligands: protons, chloride, organic phosphate, and CO_2. The oxy or relaxed (R-state) structure generally has a much lower affinity for these allosteric ligands, but a high affinity for oxygen. Heterotropic ligands lower the affinity for oxygen by stabilizing the T structure with salt bridges between the subunits. The principal allosteric effector in birds is inositolpentaphosphate (IPP), which binds to the central cavity between the N- and C-termini of the α and β subunits. Loosening constraints on the T structure increases the affinity for oxygen, as does changing the affinity of binding sites for organic phosphate and other allosteric effectors (Wang et al. 2000).

Waterfowl (Anseriformes) have featured prominently in studies of hemoglobin adaptation. The bar-headed goose (*Anser indicus*) breeds at >5,000-m elevation in Asia and migrates at >9,000 m (Perutz 1983). Oxygen loading is enhanced by Pro → Ala-α119. The smaller R-group side chain of Ala-α119 eliminates a van der Waals contact at an $\alpha^1\beta^1$ intersubunit contact, destabilizing the T structure and increasing oxygen affinity (Jessen et al. 1991; Weber et al. 1993). The bar-headed goose obtains 50% saturation (P_{50}) of the hemoglobin at much lower partial pressure of oxygen than do lowland species such as the graylag goose (*Anser anser*): 29.7 mmHg versus 39.5 mmHg, respectively (Petschow et al. 1977). The Andean goose (*Chloephaga melanoptera*), which is not a true goose (Anserinae) but belongs to a distantly related clade of goose-like ducks called sheldgeese endemic to South America (Livezey 1986), has evolved a similar mechanism to cope with chronic hypoxia in the Andes but with a different amino acid substitution on a different gene. In this species, Leu → Ser-β55 also results in a smaller R-group side chain that loosens the same $\alpha^1\beta^1$ intersubunit contact (Jessen et al. 1991; Weber et al. 1993). The Andean goose has a P_{50} of 33.9 mmHg (Hall et al. 1936).

These adaptive molecular changes have been studied at the mechanistic level using the crystal structures of the oxy and deoxy hemoglobins of bar-headed and graylag geese (Zhang et al. 1996; Wang et al. 2000; Liu et al. 2001; Liang, Hua, et al. 2001; Liang, Liu, et al. 2001). Nevertheless, little experimental effort or protein structure modeling of amino acid replacements has been focused on any other highland waterfowl species. Also, the population genetics of hemoglobins have not been studied previously for any high-elevation waterfowl.

Materials and Methods

Yellow-Billed Pintails

Yellow-billed pintails are one of the most abundant South American ducks. The subspecies *Anas georgica spinicauda*, which is the focus of this study, is widespread in the Andes from southern Colombia to Tierra del Fuego. Yellow-billed pintails inhabit high-altitude wetlands and puna grasslands of the altiplano and inter-Andean valleys to 5,000 m, whereas most lowland habitat occurs near sea level or below 1,500-m elevation in the southern Andes. Movements are yet to be studied, but lowland populations in southern South America are seasonally migratory, whereas populations in the high Andes probably are mostly nonmigratory. Yellow-billed pintails, nonetheless, are capable of long-distance dispersal, and individuals repeatedly have been observed on South Georgia Island, South Orkney and South Shetland Islands, and the Antarctic Peninsula. A distinctive endemic subspecies of yellow-billed pintail (*Anas georgica georgica*) occurs on South Georgia Island (Kear 2005). These broad-scale dispersal abilities are shared by the yellow-billed pintail's sister species, northern pintail (*Anas acuta*; Johnson and Sorenson 1999), which breeds farther north in the Arctic than any other waterfowl species. Related pintails also have colonized several other remote archipelagos in the sub-Antarctic (Kerguelen and Crozet Islands; *Anas eatoni*). Wide-ranging dispersal thus is a shared trait.

Specimen Collection

We collected yellow-billed pintails in Argentina, Bolivia, and Peru between 2001 and 2005. Sampling was conducted throughout the Andes in highland ($n = 51$) and lowland ($n = 65$) regions with contrasting partial pressures of oxygen (fig. 1). Yellow-billed pintails were collected at ≥3,063-m elevation in the central high Andes from Catamarca, Argentina, to Ancash, Peru, and at ≤1,809 m in Patagonia and central Argentina north to San Juan. One specimen was collected on the Pacific Coast of Peru, where pintails are uncommon. Voucher specimen information is available online (supplementary appendix 1, Supplementary Material online).

FIG. 1.—Yellow-billed pintail ($n = 116$) localities in South America. Black circles indicate highland localities ≥3,063 m; white circles indicate lowland localities ≤1,809 m.

Table 1
Genes Sequenced, Chromosomal Positions in the Chicken Genome, and Best-fit Nucleotide Substitution Models

Locus	Base Pairs Sequenced	Chicken Chromosome	Best-Fit Model
Ornithine decarboxylase intron 5	352	3	HKY + G
α-Enolase intron 8	314	21	K80 + G
β-Fibrinogen intron 7	246	4	HKY
N-Methyl-D-aspartate-1-glutamate receptor intron 11	328–330	17	TIM + I + G
Phosphoenolpyruvate carboxykinase intron 9	345–351	20	K80 + I
mtDNA control region	977–981	mtDNA	HKY + I
αA Hemoglobin submit	677	14	HKY + I + G[a]
βA Hemoglobin submit	1,579–1,582	1	TrN + I + G[b]

[a] Best-fit model for the 423-bp αA hemoglobin subunit amino acid coding sequence is K80 + I.

[b] Best-fit model for the 438-bp βA hemoglobin submit amino acid coding sequence is TrN + I + G; model for entire sequence not determined.

αA and βA Hemoglobin DNA Sequencing

DNA was isolated from muscle using DNeasy Tissue Kits (Qiagen, Valencia, CA). We sequenced the αA and βA hemoglobin subunit genes, which comprise the major hemoglobin protein isoform expressed in adult birds (table 1). Primers flanking the start and stop codons for each hemoglobin gene were designed de novo using duck, chicken, and other DNA sequences in GenBank (Reitman et al. 1993; Flint et al. 2001).

We used polymerase chain reaction (PCR) to obtain initial DNA sequences and design internal primers and additional primers flanking the αA and βA hemoglobin subunits. PCR was performed using AmpliTaq Gold PCR Master Mix (Applied Biosystems, Foster City, CA). Thermal cycling typically was as follows: 7 min preheat at 94 °C, followed by 45 cycles of 20 s at 94 °C, 20 s at 60–64 °C, 1 min at 72 °C, and a final extension of 7 min at 72 °C. Extension times varied (1–3 min) with the length of the fragment. The αA subunit was sequenced as a single fragment or with two overlapping fragments. The βA subunit was sequenced using nested or half-nested PCR. Product from an initial PCR spanning the complete βA-coding region was used as the template for a second PCR using multiple overlapping combinations of internal and end primers. The αA and βA hemoglobin primers are available online (supplementary appendix 2, Supplementary Material online).

To verify that both alleles were obtained for all fragments of each gene for each individual, we used a combination of widely overlapping primer pairs and carefully inspected all polymorphisms at overlapping nucleotide positions. Individuals that were homozygous for one or more fragments were resequenced using different overlapping primer pairs, and degenerate primers were designed specifically to amplify both alleles. We lowered PCR annealing temperature to 56 °C to decrease primer specificity.

PCR products were electrophoresed on agarose and purified using QIAquick Gel Extraction Kits (Qiagen). Forward and reverse strands of each product were cycle sequenced using BigDye Terminator Cycle Sequencing Kits, followed by electrophoresis on ABI 3100 automated DNA sequencers (Applied Biosystems). Sequences were edited using Sequencher 4.6 (Gene Codes, Ann Arbor, MI). Sequences that contained double peaks, indicating the presence of two alleles, were coded with IUPAC degeneracy codes and treated as polymorphisms. Indels were resolved by comparing the unambiguous 5′-ends of sequences with the 3′-ambiguous ends between forward and reverse strands (Peters et al. 2007). Gaps resulting in shifted peaks in the chromatograms, thus, enabled us to resolve short-length polymorphisms within the sequences and were treated as a fifth character state. All sequences were aligned by eye using the sequence alignment editor Se-Al 2.0a11 (Rambaut 2007). The αA and βA subunit sequences are deposited in GenBank (accession numbers FJ617587–FJ617816).

Autosomal Intron and Mitochondrial DNA (mtDNA) Control Region Sequencing

We also sequenced five autosomal introns located on different chromosomes in the chicken genome (Hillier et al. 2004) and the mtDNA control region (table 1). Intron

sequences ranged in length from 246 to 352 bp and included ornithine decarboxylase intron 5 (ODC1-5), alpha enolase intron 8 (ENO1-8), beta fibrinogen intron 7 (FGB-7), *N*-methyl-D-aspartate-1-glutamate receptor intron 11 (GRIN1-11), and phosphoenolpyruvate carboxykinase intron 9 (PCK1-9). Primers were developed specifically for ducks (supplementary appendix 2, Supplementary Material online), and all intron loci were chosen blind to levels of polymorphism. PCR was performed using the same thermal cycling protocols used for hemoglobins, except that only sequences from a single strand were collected for most individuals because the PCR templates were short. Sequences from the opposite strand were obtained for all individuals that possessed indels. We also sequenced a large segment of the mtDNA control region corresponding to positions 79−1250 in the chicken (*Gallus gallus*) (Desjardins and Morais 1990). Control region primers included L78 paired with H774 and L736 paired with H1251 or H1530 (Sorenson and Fleischer 1996; Sorenson et al. 1999; McCracken and Sorenson 2005). Intron and mtDNA control region sequences are deposited in GenBank (accession numbers FJ617817–FJ618512).

Gametic Phase of Allele Sequences

The gametic phases of individuals that were homozygous or possessed a single polymorphic position were unambiguous, and this constituted a large fraction of each diploid data set. We used a two-step procedure to determine the gametic phase of each sequence that was heterozygous at two or more nucleotide positions. We first analyzed the diploid consensus sequences of each individual using PHASE 2.1 (Stephens et al. 2001). PHASE uses a Bayesian method to infer haplotypes from diploid genotypic data with recombination and the decay of LD with distance. Each data set was analyzed using the default values (100 main iterations, 1 thinning interval, and 100 burn-in) followed by 1,000 main iterations and 1,000 burn-in (-X10 option) for the final iteration. The PHASE algorithm was run five times (-x5 option) from different starting points, selecting the result with the best overall goodness of fit. For individuals with allele pair probabilities <80%, we then designed allele-specific primers to amplify one allele but not the other (Bottema et al. 1993; Peters et al. 2005). The resulting haploid allele sequence was then subtracted from the diploid consensus sequence to obtain the gametic phase of the second haplotype. Each data set was then analyzed five more times using PHASE and the additional known allele sequences (-k option). PHASE analyses were performed for the complete αA hemoglobin sequences and the five autosomal introns. For the βA subunit, we used PHASE to infer the gametic phase of each allele by excluding the introns and analyzing only the protein-coding sequence of the three exons. We adopted this approach because the βA subunit was too long (1,579–1,582 bp) and sequenced with too many overlapping PCR fragments to carry out allele-specific priming across the entire gene. Start and stop codons and external sequences flanking the primer sites were excluded from all analyses. The gametic phases of 92.8% ($n = 752$) of the 810 individual autosomal sequences were identified experimentally or with >95% posterior probability, and 94.9% ($n = 769$) were identified with >90% posterior probability.

Population Genetic Analysis

We used Arlequin 3.11 (Excoffier et al. 2005) to calculate the number of polymorphic positions, number of alleles, nucleotide diversity (π), observed and expected heterozygosity, and test Hardy–Weinberg equilibrium. Allelic richness was standardized to the smallest sample size (102 alleles for autosomal loci and 51 haplotypes for mtDNA). We calculated overall F_{ST} and Φ_{ST} between each paired lowland and highland population, and F_{ST} for each segregating site. LD was evaluated using Fisher's (1922) exact test in Arlequin 3.11, for each locus separately and for all seven autosomal loci combined, and for lowland and highland populations combined and separately. Allelic networks showing reticulations among haplotypes were illustrated using the median joining algorithm in NETWORK 4.201 (Bandelt et al. 1999; Fluxus Technology, Ltd.). Hardy–Weinberg equilibrium and LD were also tested using only the amino acid allele sequences of the exons, with the two hemoglobin introns excluded from the analysis. We used the Bayesian Information Criterion (Schwarz 1978) with Modeltest 3.7 (Posada and Crandall 1998) and PAUP*4b10 (Swofford 2002) to determine the best-fit substitution model for each locus.

We used Structure 2.2 (Pritchard et al. 2000) to compute the probability of assignment to the lowland or highland populations and identify individuals with admixed lowland and highland genotypes. Structure analyses were conducted using the admixture model ($\alpha = 1$) with independent allele frequencies ($\lambda = 1$). The algorithm was run for 100,000 steps following a burn-in of 10,000 steps. Four analyses were performed, one including only the five autosomal introns and three additional analyses including the five introns plus mtDNA, αA subunit, and βA subunit coding sequences, added one at a time to each analysis comprising six loci total. No prior population information was used, and analyses were performed for one- and two-population models ($K = 1$ or 2).

We used the four-gamete test (Hudson and Kaplan 1985) in DnaSP 4.10 (Rozas et al. 2003) to calculate the minimum number of recombination events (R_M) for each locus. We used LAMARC 2.1 (Kuhner 2006) to calculate the recombination rate (r), which is the ratio (ρ/μ) between the per-site recombination and the per-site mutation rates (see below). A recombination rate equal to 1.0 means that recombination is equally likely to occur as mutation.

Selection Inferred from Differential Migration between Locally Adapted Populations

We used a coalescent model in LAMARC 2.1 (Kuhner 2006) to examine the balance between selection and migration by contrasting upslope and downslope migration between the αA and βA hemoglobin subunits and the five autosomal introns. The neutral parameters Θ ($4N_e\mu$), migration rates M (m/μ) between the lowlands and highlands, and recombination rates r (ρ/μ) were estimated for each locus.

We used Bayesian analyses with 1 million recorded genealogies sampled every 50 steps, with a burn-in of 100,000 (10%) genealogies. Priors were flat with the upper limit for Θ, M, and r set to 0.1, 10,000, and 10, respectively. We used the Felsenstein (1984) substitution model with empirical transition–transversion ratios and base frequencies calculated from the data. Analyses were repeated three times, each with increasing run length, to verify that parameter estimates converged within and among runs. Parameter estimates were evaluated by visually inspecting and comparing the 95% percentiles for the most probable estimates of the posterior distribution. The migration rate M was multiplied by Θ for each recipient population to calculate $4N_em$, the average number of effective migrants dispersing into each population per generation. Coalescent analyses were conducted separately for the αA subunit, βA subunit, and mtDNA. The five autosomal introns were combined under the assumption that they were unlinked by multiplying the likelihoods to obtain joint estimates.

LAMARC's migration rate estimates (M) are scaled to the mutation rate (m/μ), where m is the probability that a lineage immigrates per generation and μ is the mutation rate per site per generation. To calculate m for each autosomal locus, independently of the mutation rate, and compare these values between the αA and βA subunits and the reference loci, we multiplied M for each locus by the per-site substitution rate μ. Substitution rates were calibrated separately for each locus using the duck–goose split (Peters et al. 2007, 2008). The mean genetic distance between the snow goose (*Anser caerulescens*) sequence and resolved alleles from five species of South American dabbling ducks ($n = 516$ individuals; McCracken KG, unpublished data) was divided by the midpoint of the Oligocene (2×30.5 My) and multiplied by a generation time for *Anas* ducks of 3.2 years (Peters et al. 2008). The averaged substitution rate for the five autosomal introns was calculated as the mean. We also calculated N_e for lowland and highland populations by dividing Θ ($4N_e\mu$) by four times the mean substitution rate (μ) for the five autosomal introns.

Simulated Data Sets for the αA and βA Hemoglobin Subunits

We used ms (Hudson 2002) to simulate genetic data under drift–migration models of selective neutrality for the αA and βA subunits. We used the joint parameter estimates from LAMARC for the five autosomal introns to define population models that included population-size differences and asymmetrical migration rates between the lowlands and highlands. Because substitution rates differed among loci, we adjusted Θ for the αA and βA subunits using relative substitution rates. The substitution rates of the coding regions of the αA and βA subunits were 0.4 and 0.5 times the average rate for the introns, respectively. The Θ values in ms were scaled to the per-locus substitution rates, which were calculated by multiplying LAMARC's Θ by the length of each locus (l). We defined migration rates as $4N_em$ (calculation described above). We also included locus-specific recombination rates (ρ) for the αA and βA subunits, where $\rho = \Theta \times r \times (l - 1)$ and r is LAMARC's recombination rate. We simulated 1,000 data sets for αA and βA subunits, 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 with empirical values of Φ_{ST} for each locus.

Results

Weak Population Structure in Autosomal Introns, mtDNA, and αA Hemoglobin Subunit

The five autosomal introns and mtDNA control region showed little evidence of population subdivision between the lowlands and highlands (table 2, fig. 2). F_{ST} and Φ_{ST} were very small but significant for GRIN1-11 (0.02) and PCK1-9 (0.03), reflecting weak population differentiation. The maximum Φ_{ST} was 0.05 for mtDNA control region, but F_{ST} was nonsignificant. Two-population model Structure analyses ($K = 2$) of the five autosomal introns and five autosomal introns plus mtDNA revealed no evidence of subdivision (fig. 3*A* and *B*). Posterior assignment probabilities averaged 50.4% and 50.0%, respectively. No LD was observed between any of the five introns (Ps > 0.99) confirming that they are probably unlinked in this species.

The αA hemoglobin subunit, which contained only silent polymorphisms, showed no evidence of population subdivision (table 2, fig. 2). Neither F_{ST} nor Φ_{ST} was significant, and the observed value of Φ_{ST} fell within the 95% confidence limit of simulated αA subunit data (fig. 4*A*). Adding the αA subunit to the five autosomal introns resulted in no notable increase in posterior probability of population assignment in the Structure analysis (fig. 3*C*; 50.4% vs. 51.1%).

Elevated F_{ST} and Φ_{ST} on the βA Hemoglobin Subunit

The βA subunit displayed a strikingly different pattern. Overall F_{ST} (0.37) and Φ_{ST} (0.65) were high and significantly elevated between the lowlands and the highlands (table 2, fig. 2), with Φ_{ST} falling far outside the 95% confidence limit of simulated data (fig. 4*B*). Ten different βA subunit alleles were found in the highlands, whereas 21 alleles occurred in the lowlands (fig. 2). The average posterior probability of assignment using combined analysis of five introns and the βA subunit was 76.1% (±17.0% SD), consistent with a two-population model of subdivision segregating between the lowlands and highlands (fig. 3*D*). Number of polymorphic sites, allelic richness, and nucleotide diversity were lower in the highlands. Less heterozygosity than expected was found in the highlands but not in the lowlands.

Three amino acid replacements segregating at high frequency in the highlands (Gly → Ser-β13, Ala → Ser-β116, and Leu → Met-β133) contributed to elevated F_{ST} and Φ_{ST} on the βA subunit of yellow-billed pintails, and one additional amino acid replacement (Thr → Ser-β4) also was found in the highlands (table 3). Ser-β4 and Ser-β13 segregate independently on different alleles and occur at low frequency in the highlands but were not sampled in the lowlands. Ser-β4 was found in one

Table 2
Number of Polymorphic Positions, Alleles, Standardized Allelic Richness, Nucleotide Diversity, Θ, Heterozygosity, F_{ST}, and Φ_{ST}, for Five Unlinked Autosomal Introns, mtDNA Control Region, and the αA and βA Hemoglobin Subunits from Lowland and Highland Populations of Yellow-billed Pintails

Locus	Population	No. Polymorphic Positions	No. Alleles	Allelic Richness (±SD)	Nucleotide Diversity (π/site)	Θ	Θ (95%)	Observed Heterozygosity	Expected Heterozygosity	F_{ST}	Φ_{ST}
Ornithine decarboxylase	Lowland	24	16	14.8 ± 1.0	0.011713	0.012277	0.004535–0.021752	0.83	0.79	0.01	**0.02**
	Highland	16	13	13	0.009106	0.003901	0.001214–0.010230	0.73	0.75		
α Enolase	Lowland	12	13	12.7 ± 0.5	0.004662	0.004180	0.001313–0.011534	0.68	0.73	0.01	0.01
	Highland	11	13	13	0.004666	0.008171	0.002671–0.020918	**0.65**	0.76		
β Fibrinogen	Lowland	6	6	6.0 ± 0.1	0.004497	0.003282	0.000883–0.009240	0.78	0.68	0.00	–0.01
	Highland	5	5	5	0.004546	0.002149	0.000413–0.007392	0.69	0.68		
N-Methyl-D-aspartate-1-glutamate receptor	Lowland	31	34	31 ± 1.5	0.008993	0.020994	0.011945–0.034871	0.89	0.92	**0.02**	**0.02**
	Highland	22	23	23	0.008740	0.004489	**0.002169–0.010570**	0.92	0.91		
Phosphoenolpyruvate carboxykinase	Lowland	14	9	8.1 ± 0.8	0.001979	0.002574	0.000805–0.008152	0.52	0.53	**0.03**	**0.03**
	Highland	10	8	8	0.003563	0.004746	0.001464–0.011860	0.65	0.67		
Five introns—joint estimate	Lowland	—	—	—	—	0.007730	0.005997–0.012781	—	—	—	—
	Highland					0.004615	0.002703–0.007036	—	—	—	—
mtDNA control region	Lowland	41	39	32.0 ± 1.6	0.004139	0.022134	0.011544–0.046464	—	—	0.00	**0.05**
	Highland	28	25	25	0.003881	0.003364	**0.001089–0.009179**	—	—		
αA Hemoglobin-complete sequence	Lowland	27	40	35.8 ± 1.7	0.005145	0.012030	0.007211–0.022484	0.85	0.93	0.01	0.01
	Highland	19	25	25	0.004365	0.002513	**0.001340–0.005923**	0.86	0.90		
αA Hemoglobin-coding sequence	Lowland	11	14	12.6 ± 1.0	0.002472	0.005822	0.002346–0.012353	0.57	0.62	0.01	0.00
	Highland	5	7	7	0.002065	0.001364	0.000388–0.004433	0.61	0.60		
βA Hemoglobin-coding sequence	Lowland	14	21	18.8 ± 1.3	0.003962	0.007331	0.003881–0.012425	0.71	0.79	**0.37**	**0.65**
	Highland	9	10	10	0.002992	0.002081	0.000913–0.004700	**0.35**	0.41		

NOTE.—Sample sizes were $n = 65$ for the lowlands and $n = 51$ for the highlands. Bold numerals indicate a significant P-Value or no overlap in 95% estimates of Θ.

heterozygous individual, and Ser-β13 was found in eight heterozygous individuals (Supplementary Material online, fig. 1; $F_{ST} = 0.002$ and 0.08, respectively). Ser-β116 and Met-β133, in contrast, occur at high frequency in the highlands and segregate in complete LD on the same alleles (Supplementary Material online, fig. 1; $F_{ST} = 0.88$). F_{ST} for Ser-β116 and Met-β133 exceeded all reference-locus sites by a factor of 11. Maximum F_{ST} for any nucleotide position on the αA subunit or five autosomal introns, in contrast, was 0.08. LD was widely distributed across the βA subunit in the highlands, but not in the lowlands (Supplementary Material online, fig. 1). Most individuals (86.3%) collected in the highlands were homozygous for Ser-β116/Met-β133. Three (5.9%) were heterozygous, and four (7.8%) were homozygous for Ala-β116/Leu-β133. Overall heterozygosity of the amino acid sequence summed across the four segregating replacements was 21.6% in the highlands versus 0% in the lowlands. One individual homozygous for Ser-β116/Met-β133 was collected on the Pacific coast of Peru in Arequipa, but all other lowland individuals were homozygous for the Thr-β4/Gly-β13/Ala-β116/Leu-β133 allele.

Hardy–Weinberg tests for the βA subunit amino acid sequence were highly significant. More homozygous individuals were observed than expected when lowland and highland populations were analyzed separately (Ps = 0.008 and 0.0005, respectively; table 3). Many more homozygous individuals than expected (47.7 vs. 2.7%) were observed when the populations were pooled ($P < 0.00001$; table 3).

Coalescent Analyses—Recombination, Theta, and Migration

Coalescent analyses in LAMARC 2.1 indicate that the αA and βA hemoglobin subunits have experienced elevated recombination. Most probable estimates of the recombination rates for the αA and βA subunits were 1.6 and 3.8, respectively (Supplementary Material online, fig. 2). Most introns, in contrast, had either low recombination rates or too little polymorphism to allow recombination to be detected. The most probable recombination estimate for GRIN1-11 was 2.4, exceeding the αA subunit but less than for the βA subunit. Rank orders of R_M from the four-gamete tests matched the rank order of LAMARC estimates for most comparisons.

The most probable estimates of Θ were greater in the lowlands than the highlands for most reference loci, including the joint estimates, but 95% estimates overlapped between lowlands and highlands for most comparisons (table 2, Supplementary Material online, fig. 3). The αA and βA subunits had among the lowest Θ estimates in the highlands, but widely overlapping confidence intervals (CIs) with other loci. No overlap in 95% estimates of Θ was observed between the lowlands and highlands for the αA subunit, but the upper and lower tails of the βA subunit Θ estimates overlapped widely.

Upslope M estimates for all five introns, mtDNA, and the αA subunit fell on or broadly overlapped LAMARC's programmatic upper limit of $M = 10{,}000$. The joint

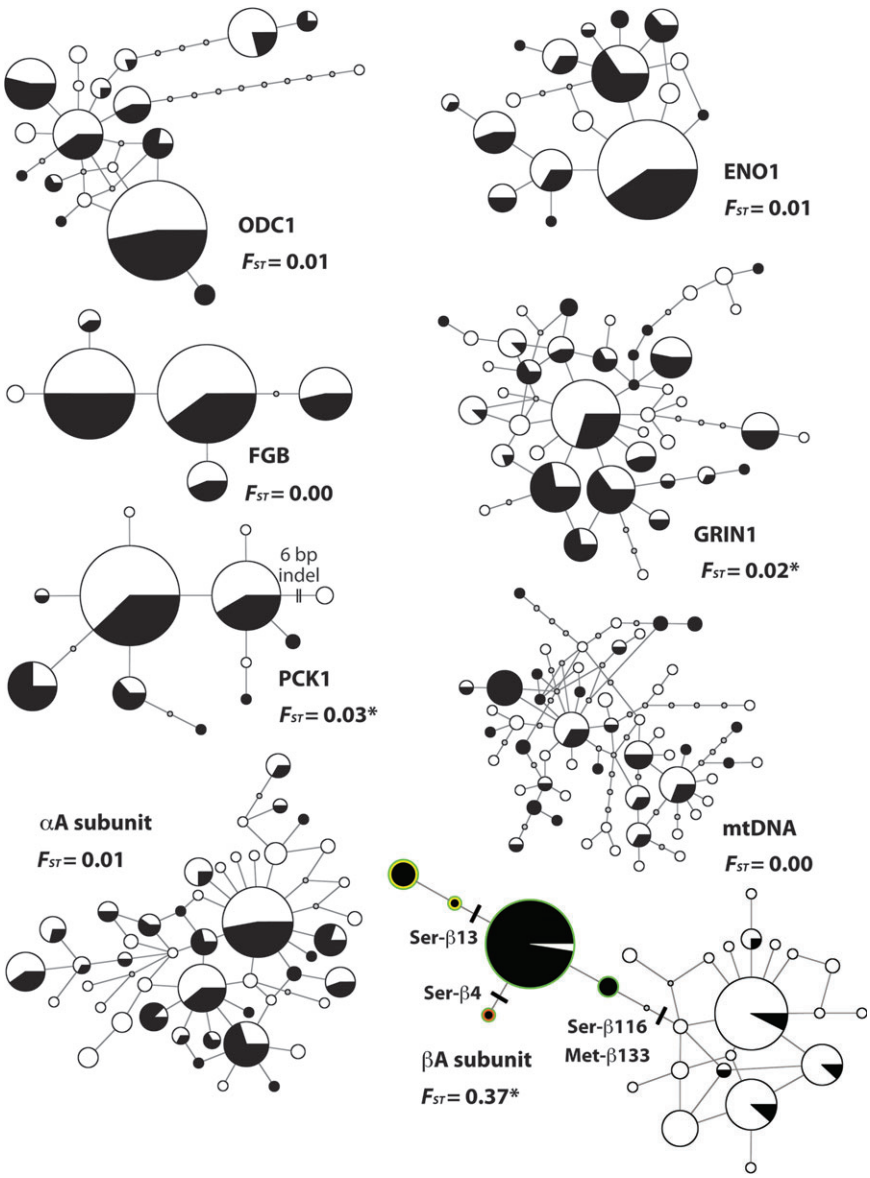

FIG. 2.—Allelic networks for eight loci. Alleles for yellow-billed pintails collected in the highlands are shown in black, and alleles collected in the lowlands are shown in white. Significant F_{ST} ($P < 0.05$) is indicated by an asterisk. βA subunit alleles segregating in the highlands are illustrated with grayscale circles showing the amino acid composition of each allele. Circle area is proportional to the number of shared alleles, and small gray circles indicate intermediate alleles not sampled.

estimate of upslope $4N_em$ for the five autosomal introns was 43.2 effective migrants per generation, but the true value probably is greater (fig. 5). Upslope $4N_em$ for the αA subunit and N_em for mtDNA were approximately similar, 24.6 and 31.8 effective migrants per generation, respectively. The upslope estimate of M (378; 95% 63–1,482) for the βA subunit, in contrast, was >24 times smaller than that for the five introns, with no overlap in 95% estimates. The most probable estimate of upslope $4N_em$ for the βA subunit was 0.8 effective migrants per generation.

Downslope M estimates were more variable among loci but also indicative of high levels of gene flow from the highlands to the lowlands. Downslope M was 3,308 (95% 2,103–4,943) for the five autosomal introns. The most probable joint estimate of downslope $4N_em$ was 25.6 effective migrants per generation (fig. 5). True values also may be greater, as ENO1-8, FGB-7, and PCK1-9 overlapped the upper boundary of the prior. Downslope $4N_em$ for the αA subunit and N_em for mtDNA were 35.2 and 42.3 effective migrants per generation, respectively. The downslope estimate of M (185; 95% CI 55–528) for the βA subunit was >17 times smaller than that for the five introns, with no overlap in 95% estimates. The most probable estimate of downslope $4N_em$ for the βA subunit was 1.4 effective migrants per generation.

In summary, the most probable estimates of M were greater upslope than downslope for all loci. Migration rate estimates for the five autosomal introns, αA subunit, and mtDNA were high in both directions, exceeding 24 effective migrants per generation for all loci. The βA subunit, in

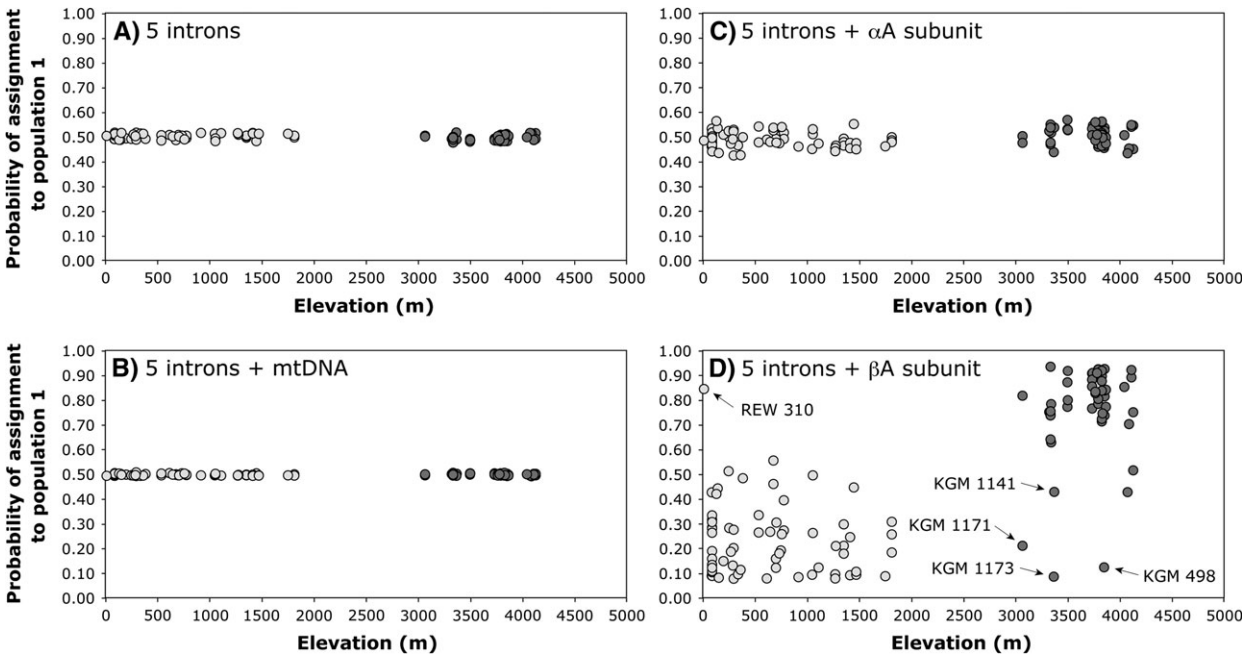

FIG. 3.—Structure 2.2 analysis showing posterior probability of assignment versus elevation for (*A*) five autosomal introns, (*B*) five introns and mtDNA control region, (*C*) five introns and the αA hemoglobin subunit, and (*D*) five introns and the *β*A hemoglobin subunit. Yellow-billed pintails collected in the highlands are shown with gray circles, and individuals from the lowlands are shown in white.

contrast, showed greatly restricted migration upslope and downslope, but the most probable estimates of M were twice as large upslope as downslope (fig. 5).

Substitution rates averaged for the five introns (3.72×10^{-9} substitutions/site/generation calculated using the duck–goose split) were 2-fold greater than for the *β*A subunit coding region (1.86×10^{-9} substitutions/site/generation), but this did not change the overall results. Most probable estimates of m were >49 times greater upslope and >35 times greater downslope for the five introns than for the *β*A subunit. The most probable estimates of m varied in the same direction as M (m/μ). The observed differences in M among loci were therefore not due to differences in mutation rates.

Discussion

Possible Effects of Ser-*β*4, Ser-*β*13, Ser-*β*116, and Met-*β*133

Oxygen dissociation curves have not yet been obtained for yellow-billed pintails, so the physiological effects of Ser-*β*4, Ser-*β*13, Ser-*β*116, and Met-*β*133 are yet to be determined. It is possible that these replacements have no effect on oxygen affinity, that their effects are relatively minor and inconsistent with the pattern we observed, or that hypoxia is not the causative agent of selection. The LD we observed, for example, could also occur in the absence of selection on the *β*A subunit if selection has acted on a closely linked locus, such as the *β*ε, *β*ρ, or *β*H subunits, which contribute to embryonic isoforms (Reitman et al. 1993; Weber 2007). Hypoxia nonetheless is among the most important factors affecting survival at high altitude, and the *β*A hemoglobin subunit is a likely target of selection.

Oxygen affinity can be enhanced by modifying affinity of the hemoglobin structure for oxygen directly or by changing the affinity for allosteric effectors such as IPP. *β*4 and *β*133 occur within van der Waals distance of IPP binding sites at the N- and C-termini of the *β*A subunit;

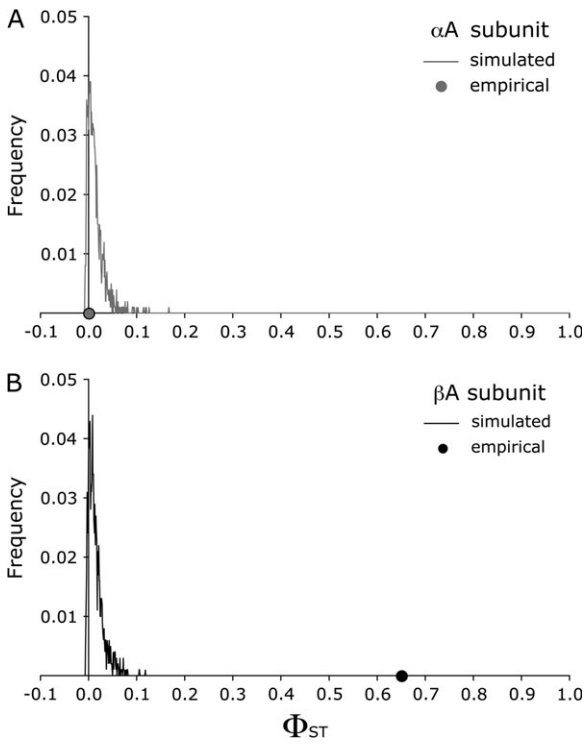

FIG. 4.—Empirical and simulated values of Φ_{ST} for 1,000 simulated data sets for the (*A*) αA hemoglobin subunit and (*B*) *β*A hemoglobin subunit.

Table 3
Genotypic Frequencies and Hardy–Weinberg Tests for the βA Hemoglobin Subunit of Lowland and Highland Yellow-billed Pintails

Genotype					Lowland		Highland		Overall	
					Observed	Expected	Observed[a]	Expected	Observed[a]	Expected
1	Thr4	Gly13	Ala116	Leu133	98.4% (62)	96.8%	7.8% (4)	1.2%	57.9% (66)	35.0%
	Thr4	Gly13	Ala116	Leu133						
2	*Thr4*	*Gly13*	*Ala116*	*Leu133*	—	—	—	0%	—	0.5%
	Ser4	*Gly13*	***Ser116***	***Met133***						
3	Thr4	Gly13	Ala116	Leu133	—	—	2.0% (1)	1.7%	0.9% (1)	4.1%
	Thr4	**Ser13**	**Ser116**	**Met133**						
4	Thr4	Gly13	Ala116	Leu133	—	3.1%	3.9% (2)	17.4%	1.8% (2)	43.6%
	Thr4	Gly13	**Ser116**	**Met133**						
5	***Ser4***	*Gly13*	***Ser116***	***Met133***	—	—	—	0%	—	0%
	Ser4	*Gly13*	***Ser116***	***Met133***						
6	***Ser4***	Gly13	***Ser116***	***Met133***	—	—	—	0.2%	—	0%
	Thr4	***Ser133***	***Ser116***	***Met133***						
7	**Ser4**	Gly13	**Ser116**	**Met133**	—	—	2.0% (1)	1.6%	0.9% (1)	0.3%
	Thr4	Gly13	**Ser116**	**Met133**						
8	*Thr4*	***Ser13***	***Ser116***	***Met133***	—	—	—	0.6%	—	0.1%
	Thr4	***Ser13***	***Ser116***	***Met133***						
9	Thr4	**Ser13**	**Ser116**	**Met133**	—	—	13.7% (7)	12.5%	6.1% (7)	2.6%
	Thr4	Gly13	**Ser116**	**Met133**						
10	Thr4	Gly13	**Ser116**	**Met133**	1.6% (1)	0%	70.6% (36)	64.6%	32.5% (37)	
	Thr4	Gly13	**Ser116**	**Met133**						

NOTE.—Amino acids that segregate in highland populations are shown in bold, and amino acids that segregate in lowland populations are shown in nonbold. Predicted genotypes not found in the sampled specimens are indicated in italics. The numbers of individuals with each genotype are shown in parentheses.

[a] Lowland, highland, and overall P values for Hardy–Weinberg tests equal 0.00794, 0.00045, and 0.00001, respectively.

β2 at 3.82 Å and β135 at 3.03 Å in graylag goose, respectively (Zhang et al. 1996; Wang et al. 2000; Liang, Hua, et al. 2001; Liang, Liu, et al. 2001). Ser-β116 occurs at an $\alpha^1\beta^1$ intersubunit contact site, and Ser-β13 occurs on the A helix of the βA heme pocket, two sites that have evolved parallel replacements in highland lineages of other Andean ducks (McCracken, KG, unpublished data). Ser-β4 and Met-β133 each evolved independently in two highland lineages, and Ser-β13 and Met-β133 are not known to occur in any non-Andean, lowland waterfowl species (McCracken, KG, unpublished data). Detailed physiological studies of the oxygen-binding effects of Ser-β4, Ser-β13, Ser-β116, and Met-β133 clearly are warranted.

Gene Flow in the Face of Countervailing Selection

Our analyses yielded two principal results that contribute to our understanding of major hemoglobin loci of yellow-billed pintails. First, upslope and downslope migration rates for the autosomal introns, mtDNA, and the αA subunit were very high in both directions, exceeding 24 effective migrants per generation, whereas the βA subunit alleles showed greatly restricted allelic migration rates, 17–24 times smaller than most probable joint estimates for reference loci. Incorporating estimates of the locus-specific mutation rates did not affect these conclusions.

Second, migration rates for all eight loci were greater upslope than downslope. LAMARC incorporates migration events that occurred during divergence and after the splitting period into a single estimate, and therefore, colonization could cause estimates of migration rates from the source to the derived population to be higher than in the opposite direction. Alternatively, asymmetrical migration might be a biological reality. Lowland pintail populations are at least partially migratory, whereas highland populations are more sedentary. If migratory behavior facilitates dispersal, then we might expect more lowland individuals to immigrate into the highlands than vice versa. However, dispersal should not be equated with gene flow (Pearce and Talbot 2006), and higher migration rates upslope than downslope might alternatively reflect differences in the fitness of lowland and highland individuals in mismatched environments. More specifically, individuals with lowland βA subunit genotypes might be more successful in the highlands than are highland individuals in the lowlands.

Hemoglobinopathies involving more than 100 variant high oxygen affinity hemoglobins have been described for humans (Charache et al. 1966; Huisman et al. 1998; Hardison et al. 2002). High-affinity hemoglobins are maladaptive in humans at low elevation because the hemoglobin is unable to efficiently unload oxygen at the capillaries, resulting in chronic tissue hypoxia (Wajcman and Galacteros 1996, 2005; David et al. 2002). Tissue hypoxia is compensated for by excess production of red blood cells (erythrocytosis), which increases the viscosity of the blood and adversely affects circulation and cardiac load. Lowlanders acclimate to hypoxia during the short term without high-affinity hemoglobin, and highlanders have adapted to high-altitude regions using other physiological mechanisms, including blunted ventilatory and pulmonary vascular responses, larger circulatory volumes, and increased red blood cell mass and O_2 carrying capacity (Hornbein and Schoene 2001; Hochachka and Somero 2002; Beall 2006).

The high-altitude physiology of yellow-billed pintails should be similar, with the exception that ducks fly and are potentially better or more rapid dispersers. The seven loci

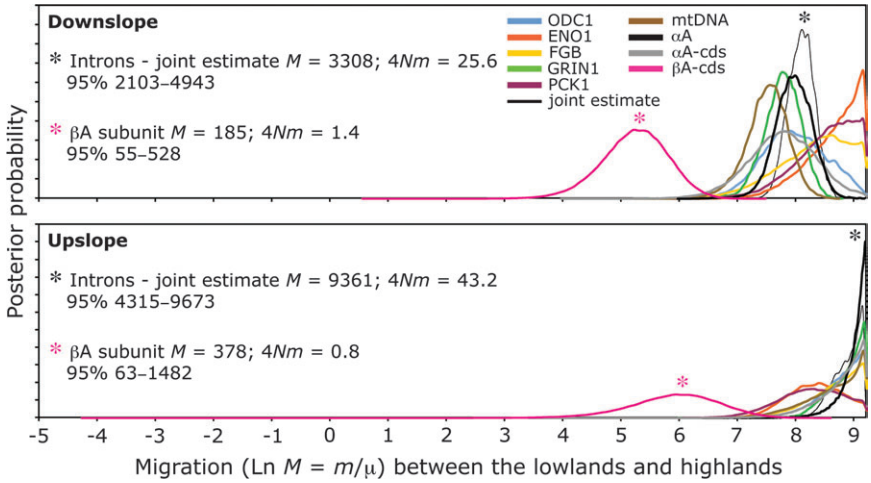

FIG. 5.—LAMARC analysis showing downslope and upslope migration estimates for yellow-billed pintails.

we sequenced, excluding the βA subunit, suggest that yellow-billed pintails exhibit high levels of gene flow between the lowlands and the highlands. Most yellow-billed pintails in the highlands probably are nonmigratory and locally adapted to the altiplano, in contrast to the lowlands, where the southern populations are seasonally migratory. If highland βA subunit alleles confer higher affinity for oxygen and this is advantageous in the highlands, highland fitness would be increased by restricted gene flow, and selective sweeps on beneficial βA subunit alleles would likely have occurred in the highlands. The same alleles would be predicted to be maladaptive in the lowlands, resulting in a stable polymorphism maintained by balancing selection and recurrent migration (Levene 1953; Maynard Smith 1970; Hedrick et al. 1976). Individuals with low-affinity genotypes might experience only slightly reduced fitness at high elevation, but over the long term would be predicted to show lower average fitness than individuals homozygous or heterozygous for high-affinity isoforms.

The spatial distribution of individual yellow-billed pintails possessing genotypes mismatched to their environment further suggests that heterozygous individuals may move between the lowlands and highlands with greater ease, but that homozygous individuals are potentially selected against, especially in the downslope direction. Eleven individuals heterozygous for lowland and highland βA subunit alleles were collected, and all were found in the highlands at widespread localities from Catamarca, Argentina to Ancash, Peru. Five mismatched homozygotes (i.e., individuals homozygous for highland alleles in the lowlands or vice versa) also were collected. Three pintails homozygous for the lowland allele Ala-β116/Leu-β133 (KGM 1141, 1171, 1173) were collected in the Catamarca highlands at the southern edge of the altiplano, and one was collected on Lake Titicaca (KGM 498) in the central and widest region of the altiplano. One individual homozygous for the highland allele Ser-β166/Met-β133 (REW 310) was collected at sea level on the Pacific coast of Peru in Arequipa. This individual was assigned to the highland population with posterior probability >0.88 in the Structure analysis, suggesting that it had dispersed from the highlands. Yellow-billed pintails are not common on the west slope of the Peruvian Andes, but individuals have been reported breeding on the southern coast in Arequipa, Moquegua, and Tacna (Pearson and Plenge 1974; Fjeldså and Krabbe 1990; Schulenberg et al. 2007); it is possible that this coastal population or others on the west slope are derived from the altiplano population. Four of the five mismatched homozygotes, thus, were collected in the highlands, and only one of 63 individuals collected in the lowlands was homozygous for a highland βA subunit genotype.

Selection on the βA hemoglobin subunit of yellow-billed pintail thus may be asymmetric. Individuals homozygous for lowland alleles (assuming the ancestral alleles confer low affinity) may have little difficulty dispersing to the highlands, at least initially, because they can acclimate to hypoxia by multiple other physiological pathways noted above. However, fewer individuals possessing homozygous highland genotypes would be expected to be able to disperse to the lowlands, if they have high-affinity isoforms, because high oxygen affinity hemoglobin would result in chronic erythrocytosis. Heterozygous yellow-billed pintails, in contrast, might have a dispersal advantage up or down the Andes if their major hemoglobin is a mixture of low affinity ($\alpha1\beta1^{L}/\alpha2\beta2^{L}$), high affinity ($\alpha1\beta1^{H}/\alpha2\beta2^{H}$), and hybrid affinity ($\alpha1\beta1^{L}/\alpha2\beta2^{H}$) isoforms capable of optimally binding and delivering oxygen at a wider range of different atmospheric pressures (Wajcman and Galacteros 2005; Weber 2007). Hypotheses presented here could be tested with an integrated study involving individual ducks with predetermined genotypes, complemented by biochemical assays of purified hemolysates to determine hemoglobin isoform ratios, oxygen affinities of the various genotypes, and intrinsic effects of each nonsynonymous substitution.

Assumptions and Limitations of the Coalescent Model

The coalescent model we used in LAMARC is likely oversimplistic because it assumes a stable long-term population structure and lacks parameters to describe population divergence. Our results suggest that lowland and

highland populations likely diverged recently, accompanied by colonization of the highlands from the southern lowlands, which is consistent with Fjeldså's (1985) finding that colonization of the high Andes occurred from the southern lowlands to the highlands for most taxa. Furthermore, migration rates for the reference loci may be too high to be adequately modeled by LAMARC. In this case, βA subunit migration rates might be more restricted than suggested by our analyses, but gene flow may be sufficiently high to retard divergence at most unlinked loci. Despite limitations and violations of underlying assumptions, we argue that the general inferences are sound. In the absence of selection, population history will affect all loci similarly, excepting the influence of stochastic processes. As such, any biases introduced by using an unrealistic model should cause similar discrepancies at all loci. Indeed, we found that parameters estimated from the five introns, αA subunit, and mtDNA, had broadly overlapping 95% CIs, consistent with those loci having experienced similar population histories. In contrast, the βA subunit consistently deviated from those general patterns, especially in estimates of migration rates. Regardless of possible deviations between the inferred model and the actual population history of yellow-billed pintails, our results strongly support a very different evolutionary history for the βA subunit than for the remaining loci.

Conclusion

Our study of yellow-billed pintails has general relevance to the study of population differentiation and adaptation. Gene flow is widely perceived to be a homogenizing evolutionary force, opposing population differentiation and speciation (Mayr 1963; Mills and Allendorf 1996; Coyne and Orr 2004). However, adaptation can proceed with high levels of countervailing gene flow when selection is strong relative to migration ($s > m$) in a large, spatially divergent population (Slatkin 1987; McKay and Latta 2002; Latta 2008). Effective population sizes (N_e) of the yellow-billed pintail were estimated to be 519,618 (95% CI 403,124–859,151) in the lowlands and 310,225 (95% CI 181,698–472,966) in the highlands, based on 3.72×10^{-9} substitutions/site/generation averaged for the five introns. These estimates are plausible given a continental population estimate of 1 million individuals (Wetlands International 2006). Using Wright's (1931) equation $F_{ST} = 1/(1 + 4N_em)$, m is approximately 1×10^{-5} given $F_{ST} = 0.01$. Although this estimate is only likely to be correct within a few orders of magnitude (Whitlock and McCauley 1999), m is probably several orders of magnitude lower than the selection coefficients (s) for the major hemoglobin of yellow-billed pintails, and at least three orders of magnitude lower than s for the duplicated α-globin subunit genes of *Peromyscus* (Storz and Kelly 2008). Comparing gene flow among putatively neutral and nonneutral candidate loci is a valuable way to identify loci that have evolved adaptively. The power to detect divergently selected loci will be maximized when selection is strong and migration is high because maladapted alleles are continually immigrating into the "wrong" environment, creating a migration load and contributing to reduced mean fitness (Storz and Kelly 2008). In such cases, migration decreases LD at unlinked loci, but LD is maintained at differentially selected loci, thus facilitating detection of genomic outliers. Yellow-billed pintails may be an unusually good model system for these reasons, possessing not only hemoglobin variants, but also potentially other traits that confer tolerance to hypoxia and the other harsh conditions of the altiplano.

Supplementary Material

Supplementary appendixes 1 and 2 are available at *Molecular Biology and Evolution online* (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank the many people and provincial and federal governments in Argentina, Peru, and Bolivia who assisted us with fieldwork for many years. Christopher Barger assisted with laboratory work, Andrew Moore wrote a program to parse the PHASE files, and Shawn Houston assisted with LAMARC analyses on the Life Science Informatics Computational Portal. Scott Edwards, John Pearce, Michael Sorenson, Jay Storz, Van Remsen, Kevin Winker, Christopher Witt, and three anonymous reviewers provided helpful discussions or comments on the manuscript. Funding was provided by the Institute of Arctic Biology, Alaska EPSCoR (NSF EPS-0092040 & EPS-0346770), grants from the Delta Waterfowl Foundation, Frank M. Chapman Fund at the American Museum of Natural History to R.E.W., and NSF (DEB-0444748) to K.G.M.

Resumen

Cuando poblaciones se adaptan localmente a ambientes contrastantes, los alelos que tienen un alto fitness en un ambiente podrían ser rápidamente eliminados en poblaciones adaptadas a otros ambientes, de forma que el flujo génico está parcialmente restricto. Cuanto mas fuerte es la selección, los alelos inmigrantes con menor fitness serán eliminados mas rápidamente de la población. Sin embargo, el flujo de genes podría continuar ocurriendo en loci no ligados y la divergencia adaptativa puede proceder en presencia de flujo génico que la contrarreste, en tanto que la fuerza selectiva sea mayor que la migración (s > m). Se estudió la genética de poblaciones de los genes de hemoglobina mayor en poblaciones de pato maicero (*Anas georgica*) experimentando contrastantes presiones parciales de oxígeno en los Andes de Sudamérica. Se evidenció alto flujo génico y poca subdivisión poblacional en siete loci que se presumen neutrales y están localizados en distintos grupos de cromosomas. En cambio, los reemplazos de aminoácidos (Ser-β13, Ser-β116, Met-β133) encontrados en la subunidad βA de la hemoglobina se segregaron por elevación entre las poblaciones de altura y las de las tierras bajas con un F_{ST} significativamente elevado. Las tasas de migración para los alelos de la subunidad βA fueron aproximadamente 17–24 veces menores que para los cinco loci no ligados de referencia, la subunidad de hemoglobina αA (que no

presenta reemplazo de aminoácidos) y la región de control mitocondrial. Los alelos de la subunidad βA de los patos maiceros tuvieron la mitad de probabilidad de ser transferidos cuesta abajo, desde los andes hacia las tierras bajas, que en la dirección opuesta (cuesta arriba). Se hipotetiza que la migración entre las tierras bajas y los andes está restricta por adaptación local y que la subunidad βA de la hemoglobina es un posible objetivo de la selección natural en relación a la hipoxia altoandina; sin embargo, el flujo de genes podría ser lo suficientemente alto como para retardar la divergencia en la mayoría de loci no ligados. Individuos homocigotas para los alelos de las tierras bajas tendrían inicialmente poca dificultad en dispersarse hacia las tierras altas, pero podrían experimentar reducción de fitness a largo plazo. Al contrario, se predice que los individuos homocigotas para los genotipos de altiplanicie tendrían dificultad dispersándose hacia las tierras bajas si los alelos de hemoglobina les confieren alta afinidad por el oxígeno, lo que resultaría en eritrocitosis crónica en regiones de baja elevación. Por su parte, los individuos heterocigotas tendrían una ventaja al dispersarse si su hemoglobina tiene un rango de funcionamiento mas amplio debido a la presencia de múltiples isoformas de la proteína con una mezcla de distintas afinidades por el oxígeno.

Literature Cited

Bamshad MJ, Wooding SP. 2003. Signatures of natural selection in the human genome. Nat Rev Genet. 4:99–111.

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 16:37–48.

Beall CM. 2006. Andean, Tibetan, and Ethiopian patterns of adaptation to high-altitude hypoxia. Integr. Comp Biol. 46:18–24.

Beaumont MA. 2005. Adaptation and speciation: what can F_{st} tell us? TREE. 20:435–440.

Bottema CDK, Sarkar G, Cassady JD, Ii S, Dutton CM, Sommer SS. 1993. Polymerase chain-reaction amplification of specific alleles: a general method of detection of mutations, polymorphisms, and haplotypes. Meth Enzymol. 218:388–402.

Charache S, Weatherall DJ, Clegg JB. 1966. Polycythemia associated with a hemoglobinopathy. J Clin Invest. 45:813–822.

Charlesworth B, Nordborg M, Charlesworth D. 1997. The effects of local selection, balanced polymorphism, and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet Res. 70:155–174.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science. 307:1928–1933.

Coyne JA, Orr HA. 2004. Speciation. Sunderland (MA): Sinauer Associates.

David O, Ivaldi G, Rabino-Massa E, Ricco G. 2002. Functional studies on nine different haemoglobins with high oxygen affinity. Acta Haematol. 108:132–138.

Desjardins P, Morais R. 1990. Sequence and gene organization of the chicken mitochondrial genome. J Mol Biol. 212:599–634.

Endler JA. 1973. Gene flow and population differentiation. Science. 179:243–250.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online. 1:47–50.

Felsenstein J. 1984. PHYLIP: phylogeny inference package ver. 2.6. Seattle: University of Washington.

Fisher RA. 1922. On the interpretation of $\chi 2$ from contingency tables, and the calculation of P.J. Roy. Statist Soc. 85: 87–94.

Fjeldså J. 1985. Origin, evolution, and status of the avifauna of Andean wetlands. Ornithol Monogr. 36:85–112.

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. Copenhagen, Denmark: Zoological Museum, University of Copenhagen.

Flint J, Tufarelli C, Peden J, et al. (14 co-authors). 2001. Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster. Hum Mol Genet. 10:371–382.

Higgs DR. 2001. Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster. Hum Mol Genet. 10:371–382.

Hall FG, Dill DB, Guzman Barron ES. 1936. Comparative physiology in high altitudes. J Cell Comp Physiol. 8:301–313.

Hardison RC, Chui DHK, Giardine B, Riemer C, Patrinos GP, Anagnou N, Miller W, Wajcman H. 2002. HbVar: a relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. Hum Mutat. 19:225–233.

Hedrick PW. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution. 53:313–318.

Hedrick PW, Ginevan ME, Ewing EP. 1976. Genetic polymorphism in heterogeneous environments. Ann Rev Ecol Syst. 7:1–32.

Hillier LW, Miller W, Birney E. (174 co-authors). 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: Oxford University Press.

Hoekstra HE, Drumm KE, Nachman MW. 2004. Ecological genetics of adaptive color polymorphism in pocket mice: geographic variation in selected and neutral genes. Evolution. 58:1329–1341.

Hornbein TF, Schoene RB. 2001. High altitude: an exploration of human adaptation. New York: Marcel Dekker, Inc.

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.

Huisman THJ, Carver MFH, Efremov GD. 1998. A syllabus of human hemoglobin variants, 2nd ed. Augusta (GA): The Sickle Cell Foundation.

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. Proc Natl Acad Sci USA. 88:6519–6522.

Johnson KP, Sorenson MD. 1999. Phylogeny and biogeography of dabbling ducks (genus: *Anas*): a comparison of molecular and morphological evidence. Auk. 116:792–805.

Kear J. 2005. Bird families of the world: ducks, geese, and swans. Oxford: Oxford University Press.

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

Latta RG. 2008. Conservation genetics as applied evolution: from genetic pattern to evolutionary process. Evol Appl. 1:84–4.

Levene H. 1953. Genetic equilibrium when more than one ecological niche is available. Am Nat. 87:331–333.

Liang YH, Hua ZQ, Liang X, Xu Q, Lu GY. 2001. The crystal structure of bar-headed goose hemoglobin in deoxy form: the allosteric mechanism of a hemoglobin species with high oxygen affinity. J Mol Biol. 313:123–37.
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 Crystall Sect D Biol Crystall. 57:1850–856.
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 Crystall Sect D Biol Crystall. 57:775–783.
Livezey BC. 1986. A phylogenetic analysis of recent Anseriform genera using morphological characters. Auk. 103:737–754.
May RM, Endler JA, McMurtrie RE. 1975. Gene frequency clines in the presence of selection opposed by gene flow. Am Nat. 109:659–676.
Maynard Smith J. 1970. Genetic polymorphism in a varied environment. Am Nat. 104:487–490.
Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favourable gene. Genet Res. 23:23–35.
Mayr E. 1963. Animal species and evolution. Cambridge (MA): Harvard University Press.
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*)? Syst Biol. 54:35–55.
McKay JK, Latta RG. 2002. Adaptive population divergence: markers, QTL and traits. TREE. 17:285–291.
Mills LS, Allendorf FW. 1996. The one-migrant-per-generation rule in conservation and management. Conserv Biol. 10:1509–1518.
Pearce JM, Talbot SL. 2006. Demography, genetics, and the value of mixed messages. Condor. 108:474–479.
Pearson DA, Plenge MA. 1974. Puna bird species on the coast of Peru. Auk. 91:626–631.
Perutz MF. 1983. Species adaptation in a protein molecule. Mol Biol Evol. 1:1–28.
Peters JL, McCracken KG, Zhuravlev YN, Lu Y, Wilson RE, Johnson KP, Omland KE. 2005. Phylogenetics of wigeons and allies (Anatidae: *anas*): the importance of sampling multiple loci and multiple individuals. Mol Phylogen Evol. 35:209–224.
Peters JL, Zhuravlev YN, Fefelov I, Humphries EM, Omland KE. 2008. Multilocus phylogeography of a Holarctic duck: colonization of North America from Eurasia by Gadwall (*Anas strepera*). Evolution. 62:1469–1483.
Peters JL, Zhuravlev Y, 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.
Petschow D, Wurdinger I, Baumann R, Duhm J, Braunitzer G, Bauer C. 1977. Cause of high blood O_2 affinity of animals living at high altitude. J Appl Physiol. 42:139–143.
Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics. 14:817–818.
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. Oxford: University of Oxford.
Reitman M, Grasso JA, Blumenthal R, Lewit P. 1993. Primary sequence, evolution, and repetitive elements of the *Gallus gallus* (chicken) β-globin cluster. Genomics. 18:616–626.
Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 19:2496–2497.
Schulenberg TS, Stotz DF, Lane DF, O'Neill JP, Parker TA III. 2007. Birds of Peru. Princeton (NJ): Princeton University Press.
Schwarz G. 1978. Estimating the dimension of a model. Ann Statist. 6:461–464.
Slatkin M. 1987. Gene flow and the geographic structure of natural populations. Science. 236:787–792.
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. Mol Phylogen Evol. 12:105–114.
Sorenson MD, Fleischer RC. 1996. Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. Proc Natl Acad Sci USA. 93: 15239–15243.
Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 68:978–989.
Storz JF. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. Mol Ecol. 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, Gering EJ, Moriyama H, Ferrand N, Monteiro B, Nachman MW. 2007. The molecular basis of high-altitude adaptation in deer mice. Plos Genet. 3:448–459.
Swofford DL. 2002. PAUP*: phylogenetic analysis using parsimony (and other methods) version 4. Sunderland (MA): Sinauer Associates.
Wajcman H, Galacteros F. 1996. Abnormal hemoglobins with high oxygen affinity and erythrocytosis. Hematol Cell Ther. 38:305–312.
Wajcman H, Galacteros F. 2005. Hemoglobins with high oxygen affinity leading to erythrocytosis. New variants and new concepts. Hemoglobin. 29:91–106.
Wang HC, Liang YH, Zhu JP, Lu GY. 2000. Crystallization and preliminary crystallographic studies of bar-headed goose fluoromethaemoglobin with inositol hexaphosphate. Acta Crystall Sect D Biol Crystall. 56:1183–1184.
Weber RE. 2007. High-altitude adaptations in vertebrate hemoglobins. Respir Physiol Neurobiol. 158:132–142.
Weber RE, Jessen TH, Malte H, Tame J. 1993. Mutant hemoglobins (α^{119}-Ala and β^{55}-Ser): functions related to high-altitude respiration in geese. J Appl Physiol. 75:2646–2655.
Wetlands International. 2006. Waterbird population estimates, 4th ed. Devon, UK: Wetlands International.
Whitlock MC, McCauley DE. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1$. Heredity. 82:117–125.
Wright S. 1931. Evolution in Mendelian populations. Genetics. 16:97–159.
Wu CI. 2001. The genic view of the process of speciation. J Evol Biol. 14:851–865.
Wu CI, Ting CT. 2004. Genes and speciation. Nat Rev Genet. 5:114–122.
Zhang J, Hua ZQ, Tame JRH, Lu GY, Zhang RJ, Gu X, C. 1996. The crystal structure of a high oxygen affinity species of haemoglobin (bar-headed goose haemoglobin in the oxy form). J Mol Biol. 255:484–493.

Scott Edwards, Associate Editor

Accepted January 3, 2009