



Adaptive introgression of the beta-globin cluster in two Andean waterfowl

Allie M. Graham¹ · Jeffrey L. Peters² · Robert E. Wilson³ · Violeta Muñoz-Fuentes⁴ · Andy J. Green⁵ · Daniel A. Dorfsman⁶ · Thomas H. Valqui^{7,8} · Kevin Winker⁹ · Kevin G. McCracken^{6,7,9,10,11}

Received: 14 October 2020 / Revised: 4 April 2021 / Accepted: 8 April 2021
© The Author(s), under exclusive licence to The Genetics Society 2021

Abstract

Introgression of beneficial alleles has emerged as an important avenue for genetic adaptation in both plant and animal populations. In vertebrates, adaptation to hypoxic high-altitude environments involves the coordination of multiple molecular and cellular mechanisms, including selection on the hypoxia-inducible factor (HIF) pathway and the blood-O₂ transport protein hemoglobin (Hb). In two Andean duck species, a striking DNA sequence similarity reflecting identity by descent is present across the ~20 kb β -globin cluster including both embryonic (*HBE*) and adult (*HBB*) paralogs, though it was yet untested whether this is due to independent parallel evolution or adaptive introgression. In this study, we find that identical amino acid substitutions in the β -globin cluster that increase Hb-O₂ affinity have likely resulted from historical interbreeding between high-altitude populations of two different distantly-related species. We examined the direction of introgression and discovered that the species with a deeper mtDNA divergence that colonized high altitude earlier in history (*Anas flavirostris*) transferred adaptive genetic variation to the species with a shallower divergence (*A. georgica*) that likely colonized high altitude more recently possibly following a range shift into a novel environment. As a consequence, the species that received these β -globin variants through hybridization might have adapted to hypoxic conditions in the high-altitude environment more quickly through acquiring beneficial alleles from the standing, hybrid-origin variation, leading to faster evolution.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41437-021-00437-6>.

✉ Allie M. Graham
graham.allie@gmail.com

✉ Kevin G. McCracken
kevin.g.mccracken@gmail.com

¹ Eccles Institute for Human Genetics, University of Utah, Salt Lake City, UT, USA

² Department of Biological Sciences, Wright State University, Dayton, OH, USA

³ School of Natural Resources, University of Nebraska–Lincoln, Lincoln, NE, USA

⁴ European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge, UK

⁵ Department of Wetland Ecology, Estación Biológica de Doñana, EBD-CSIC, Sevilla, Spain

Introduction

Genetic variation, either from mutation or standing variation, is crucial for adapting to novel environments; however, adaptation from standing genetic variation is likely to

⁶ Human Genetics and Genomics, Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA

⁷ Centro de Ornitología y Biodiversidad (CORBIDI), Surco, Lima, Perú

⁸ Universidad Nacional Agraria, La Molina, Perú

⁹ University of Alaska Museum and Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK, USA

¹⁰ Department of Biology, University of Miami, Coral Gables, FL, USA

¹¹ Department of Marine Biology and Ecology, Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, FL, USA

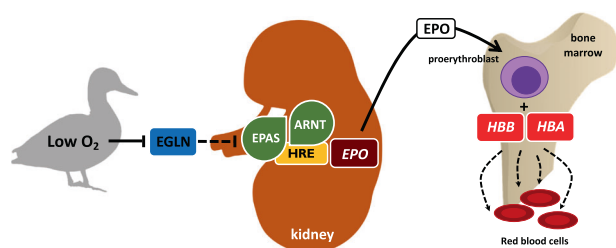


Fig. 1 A simplified schematic showing how HIF-pathway-induced expression of erythropoietin (EPO) in the kidney stimulates expression of red blood cells and hemoglobin (Hb). The *HBA* and *HBB* genes are expressed in the bone marrow and other tissues during proerythroblast development (Gassmann and Muckenthaler 2015).

lead to faster evolution compared to waiting for de novo mutations (Barrett and Schluter 2008; Hermisson and Pennings 2005). Such standing variation can originate from within the same population, from other populations, or even different species, through adaptive introgression (i.e., standing introgression variation; SI, *sensu* Jagoda et al. 2017). In animals, numerous examples have highlighted the importance of adaptive introgression in providing alleles important for a wide variety of environmental pressures, including resistance to insecticides (Clarkson et al. 2014; Song et al. 2011), rodenticides (Liu et al. 2015; Norris et al. 2015), pollution (Oziolor et al. 2019), high-altitude adaptation (Huerta-Sánchez et al. 2014), seasonal camouflage (Jones et al. 2018), aposematic mimicry (Pardo-Diaz et al. 2012), and beak size (Grant and Grant 2016). In plants, this phenomenon is even more widespread (Morjan and Rieseberg 2004; Rieseberg 2011; Suarez-Gonzalez et al. 2018; Whitney et al. 2015).

Hemoglobin (Hb) has been a flagship molecular marker in studies of adaptation to high altitude (Storz 2019). The globin-gene families provide a source of variation from which adaptation has occurred both through standing variation and new mutations (Galen et al. 2015; Hoffmann et al. 2012; Natarajan et al. 2016; Projecto-Garcia et al. 2013; Storz 2016a). At high altitudes, the low partial pressure of oxygen (P_{O_2}) results in a reduction in O_2 saturation of arterial blood (Beall 2001). Waterfowl and other bird species (Natarajan et al. 2015; Projecto-Garcia et al. 2013) have featured prominently in studies of Hb function and evolution in response to colonization of high-altitude regions. Specific amino-acid substitutions have been identified that increase O_2 -binding affinity of the HbA (major) isoform by affecting α/β -subunit interactions and sensitivity to allosteric cofactors (Perutz 1983; Weber et al. 1993). In most birds, there is a consistent trend toward increased Hb- O_2 affinity in the high-altitude lineages (Storz 2016b), a pattern that has been experimentally validated using both naturally purified Hb and recombinant Hb (Natarajan et al. 2015; Projecto-Garcia et al. 2013). In addition, a high degree of parallel amino-acid

substitution is also observed in the α - and β -subunit genes in many avian lineages, suggestive of a strong role for selection on standing variation in addition to the role of de novo mutation (Natarajan et al. 2015). Finally, recent work has implicated genes in the hypoxia-inducible factor (HIF) pathway (Fig. 1) as being important for high-altitude adaptation in some of the same species and other high-altitude species (Cho et al. 2013; Pamerter et al. 2020; Qiu et al. 2012; Qu et al. 2013; Scheinfeldt and Tishkoff 2010; Wang et al. 2014). In waterfowl, for example, variants in the same exon of *EPAS1* (*HIF-2 α*) show a striking pattern of parallel evolution qualitatively similar to that seen in Hb (Graham and McCracken 2019). However, it is unknown to what degree any of these variants were due to parallel evolution involving de novo mutations or due to adaptive introgression.

Here we use statistical analyses of genomic DNA sequence data to evaluate evidence for adaptive introgression from standing interspecific variation in the β -globin and HIF-pathway genes (31 genes in all) of two species of Andean ducks, the speckled teal (*Anas flavirostris*) and yellow-billed pintail (*A. georgica*), which are widespread across southern South America and each have resident populations at high altitude. Ultimately, we show that linked nonsynonymous genetic variations in two β -globin genes (*HBE*, *HBB*) have likely introgressed between high-altitude populations of these two species. We furthermore show that this introgression likely occurred into the high-altitude yellow-billed pintail population from the speckled teal population following a possible range shift into a novel environment by the former. This potentially enabled the yellow-billed pintail to become established in the Andes mountains after the speckled teal population had already occupied high altitudes for some time prior and for evolution to proceed faster than it would have had yellow-billed pintails had to wait for advantageous *de novo* mutations. By contrast and despite strong evidence of parallel evolution, introgression of *EPAS1* and *EGLN1* or other loci in the HIF pathway was not supported by our analysis. Studies of adaptive introgression between these two species and the consequences thereof deserve study at much wider genomic scale and fine-scale chromosome mapping to elucidate the genomic architecture of these past hybridization event(s).

Methods

Study organisms

Our study focuses on the speckled teal and the yellow-billed pintail, the two most common waterfowl (family Anatidae) species in South America. These species are sympatric

throughout their range and are superficially similar, each with brown plumage that is similar between the sexes (i.e., monochromatic), and each with a similar yellow bill, yet the species are significantly divergent in body size (e.g., averaging 412 vs. 776 g for male teal versus pintail respectively; (Kear and Hulme 2005). However, these two species are not sister taxa and belong to different clades that are not particularly closely related within the genus *Anas*. Uncorrected mtDNA divergence between species is 5.2% (Johnson and Sorenson 1999), corresponding to ~1 million or more years ago divergence time based on based on 4.8×10^{-8} substitutions/site/year reported by Peters et al. (2005) for the avian mtDNA control region. Furthermore, there is a substantially deeper divergence between mtDNA lineages of low- and high-altitude populations of speckled teal compared to yellow-billed pintail (~3 vs. 0%; Supplementary Fig. 1), suggesting that speckled teal (for which the low- and high-altitude populations are classified as different subspecies) have occupied high-altitude habitats for a much longer period of time (Graham et al. 2018). In contrast, pintail populations do not show significantly different mtDNA haplotype frequencies between low- and high-altitude populations, are classified as the same subspecies, and therefore probably colonized high-altitude regions more recently and/or have maintained much higher levels of gene flow between low and high-altitude sites (McCracken et al. 2009c).

Specimen collection and DNA extraction

A total of 40 individual vouchered specimens were used for this study from these two Andean duck species. We used 20 individuals of each species, the same used by Graham and McCracken (2019) in a previous study of the HIF pathway. For the speckled teal, individuals from low-altitude populations are the *A. f. flavirostris* subspecies ($n = 10$; elevation range 77–860 m) and from high-altitude are the *A. f. oxyptera* subspecies ($n = 10$; elevation range 3211–4405 m). For the yellow-billed pintail, individuals from both populations are classified taxonomically as *A. g. spinicauda*. A total of 20 yellow-billed pintails were collected from low- ($n = 10$; elevation range 292–914 m) and high-altitude sites ($n = 10$; elevation range 3332–4070 m). Maps depicting the provenance of these specimens and accompanying georeferenced locality data were published previously by McCracken et al. (2009a, 2009b, 2009c) and Graham et al. (2018). Genomic DNA was extracted from tissue using the DNeasy Tissue Kit (Qiagen, Valencia, California, USA) following manufacturer's protocols. Tissues are deposited in the Genomic Resources Collection at the University of Alaska Museum, and vouchered specimens reside at the University of Alaska Museum and Colección Boliviana de Fauna in La Paz, Bolivia.

Target-enrichment sequencing, sequence assembly, and annotation

We used in-solution target capture prior to next-generation sequencing to selectively enrich libraries for the complete α - and β -globin cluster genes in combination with capture of 26 genes from the HIF pathway (Gnirke et al. 2009). For details on the genes chosen and the methods in subsequent assembly for the HIF-pathway genes see Graham and McCracken (2019). Briefly, the HIF-pathway genes utilized in this analysis were chosen for being either (1) genes of interest in previous studies of humans using whole-genome scans looking to identify genetic adaptations to high-altitude or hypoxic environments, (2) part of the canonical HIF pathway, either as a known downstream target, or a part of the repression machinery of the pathway, and/or (3) part of similar transcription factor families. Rationale for sequencing the complete α - and β -globin clusters, and linked embryonic genes in particular, followed Natarajan et al. (2015)'s experimental characterization of the genetic basis of increased Hb-O₂ affinity in these same species of Andean ducks and the preliminary finding that identical by descent β -globin alleles were present across this linkage group.

For the Hb, these sequence capture targets comprised linkage groups spanning 17,851 bp (α) and 42,936 bp (β) respectively, including all of the known genes and pseudogenes in the α - and β -globin clusters, expressed for both adult and embryonic Hb isoforms. The intronic and exonic Hb regions for each of the α subunits (*HBZ* = LOC101800713; *HBA1* = LOC101800520; *HBA2* = LOC101800332) and β subunits (*HBE* = LOC101798290; *HBB* = LOC101798111) were then annotated using genomic features available through the mallard (*Anas platyrhynchos*) genome on NCBI (BGI_duck_1.0, GCA_000355885.1). In the case of the β -globin, our 43 kb sequence also included two olfactory receptor genes (*OR51G2*, *OR51L1*) flanking the 5' and 3' ends of the 20 kb β -globin cluster.

All steps of the sequence capture process were performed by MYcroarray (now Arbor Biosciences; Ann Arbor, MI). Custom MYbaits[®] biotinylated ssRNA target capture baitsets were designed. Specifically, a custom set of 210 (α -globin) and 579 (β -globin) 120mer probes at 2X tiling density was designed from the mallard genome. The DNA samples were sonicated, size-selected, and converted to Illumina libraries using dual indexes (P5, P7) during library amplification. These libraries were then enriched using the MYbaits[®] v3.0 procedure (Gnirke et al. 2009), quantified using qPCR, and then pooled for sequencing. Sequencing was performed on an Illumina Hi-Seq v4 platform paired end (100 bp) with a 250–300 bp insert size. Sequences were received demultiplexed by individual, with adapters trimmed and quality

filtered (>Q30). Additionally, sequences were trimmed of adapter sequences (*fastq-clipper*), then filtered by length and quality using the FASTX-Toolkit v. 0.0.13 package (Gordon and Hannon 2010).

The globin complexes were assembled, and variants described using two different methods—one producing contiguous sequences, and another producing a variant file (vcf). First, a reference-guided assembly was performed with Bowtie2 (Langmead and Salzberg 2012) against the α -globin (NW_004678373, scaffold 2065) and β -globin (NW_004682656, scaffold 6035) sequences of the mallard. Thus, the annotations were pulled from NCBI's most recent version of this assembly at the time (July 2015). Consensus sequences were then extracted for each individual and aligned separately by species using MAFFT (Katoh et al. 2002). The intronic and exonic Hb regions for each of the α subunit ($\alpha^T = \text{LOC101800713}$; $\alpha^D = \text{LOC101800520}$; $\alpha^A = \text{LOC101800332}$) and β subunit ($\beta^E = \text{LOC101798290}$; $\beta^A = \text{LOC101798111}$) clusters were then annotated using genomic features available through the mallard genome on NCBI. Second, a custom pipeline (<https://github.com/amgraham07>) was created to remove orphan sequences and assemble sequences against the α -globin and β -globin sequences of the mallard reference, using BWA v0.7.15 (Li and Durbin 2009). The Samtools package v1.3.1, including BCFtools v.1.3.1 (Li et al. 2009) was then used to create a VCF file and to provide assembly statistics (i.e., bp-by-bp coverage). These programs used in the pipeline called SNP variants against the mallard reference, including indels (insertion/deletion); however, indels were excluded in the final dataset as were sites with missing data.

For the HIF-pathway genes and Hb clusters, variant call format (VCF) files were created for downstream analyses. These files included all four populations comprising speckled teal (low and high) and yellow-billed pintail (low and high). All sites containing indels and missing data were removed, generating a total of 35,515 variable positions for the 26 HIF-pathway genes, 636 variable positions for the 17.9 kb-long α -globin cluster, and 2931 variable positions for the 42.9 kb-long β -globin cluster including adjacent olfactory receptor genes.

RAD-Seq reference data

Previously generated RAD-Seq (Restriction Site Associated DNA Sequencing) data for the same samples were used to function as a “backdrop”, i.e., genomic background, for analyses for the α - and β -globin clusters and HIF-pathway enriched SNP data. The single-digest RAD-Seq data were generated via FLORAGENEX (www.floragenex.com) using methods described in Natarajan et al. (2015) and Graham et al. (2018). Ultimately, the SNP output comprising 47,731 SNPs for speckled teal and 49,670 SNPs for

yellow-billed pintail was then incorporated into the final Hb and HIF-pathway SNP dataset and used for downstream population genetic analyses (see next section).

F_{ST} for identifying outlier and introgressed loci

To identify outlier loci we calculated F_{ST} , which measures population genetic differentiation due to genetic structure and the corresponding reduction in heterozygosity caused by population subdivision and/or directional selection. Whereas the latter, directional selection, is expected to have locus-specific effects related to individual gene function, the former affects genome-wide patterns because population subdivision is a demographic process (Peters et al. 2012). We used Weir and Cockerham (1984)'s F_{ST} calculated locus-by-locus (SNP-by-SNP) in VCFtools (Danecek et al. 2011). F_{ST} was first calculated between low- and high-altitude populations within species to identify conventional outliers potentially evolving by directional selection in high-altitude populations. This was done for the complete α - and β - globin clusters new to this study and also the 26 HIF-pathway genes previously surveyed by Graham and McCracken (2019; e.g., see Figs. 2, 3). The RAD-Seq data were used as a genome-wide backdrop to assess significance of outliers in the case of the α - and β - globin clusters. This analysis had already been performed previously for the HIF-pathway genes

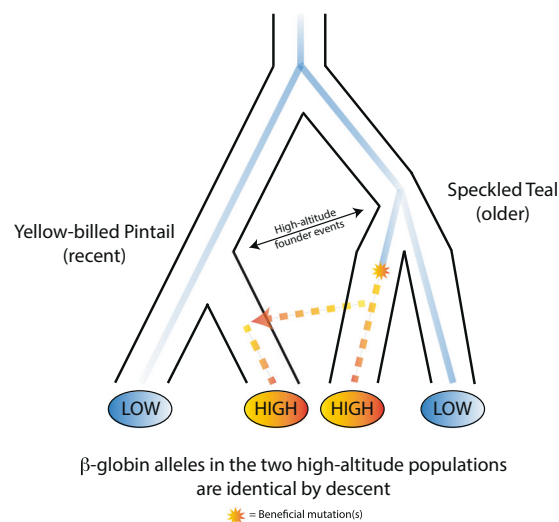


Fig. 2 Coalescent model illustrating how a more recently established population of yellow-billed pintails at high altitude likely acquired beneficial β -globin allele(s) from an older, more established population of speckled teal at high altitude. This example of adaptive introgression presumably involved a founder event by yellow-billed pintails followed by hybridization and backcrossing to the parental population and a selective sweep on newly-acquired beneficial β -globin alleles, which appear identical by descent (IBD) across the entire β -globin cluster. A selective sweep likely occurred earlier the original speckled teal β -globin prior to yellow-billed pintail becoming established at high altitude.

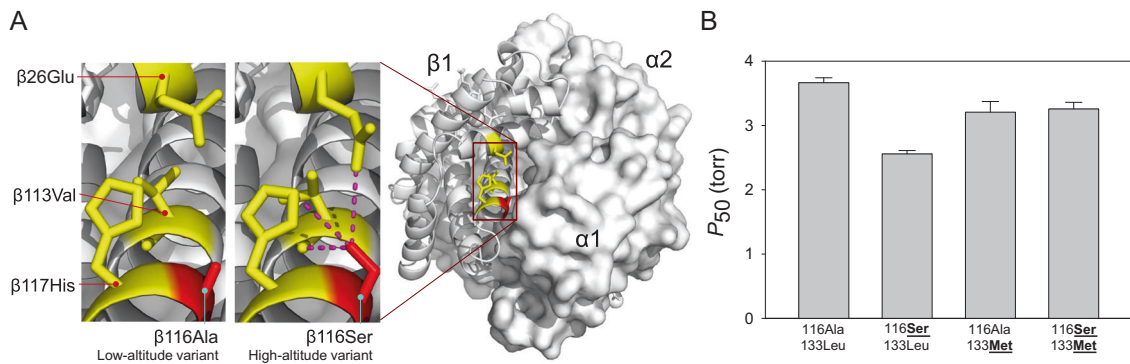


Fig. 3 Function properties of the observed β -globin amino acid substitutions in the HbA isoform of speckled teal and yellow-billed pintail. **(A)** Three-dimensional homology model of HbA showing the position of Ala β^A 116Ser and **(B)** independent and joint effects of Ala β^A 116Ser and Leu β^A 116Met on intrinsic hemoglobin P_{50} (y-axis) in the absence of allosteric effectors (figures modified and reprinted from Natarajan et al. 2015 under the Creative Commons Attribution (CC BY) 4.0 License applied by PLoS journals). The largest increase

in Hb-O₂ affinity (i.e., lower P_{50}) occurs in the presence of Ala β^A 116Ser and absence of Leu β^A 116Met, whereas a moderate increase in Hb-O₂ affinity is caused by substitution of the two amino acids combined. The derived non-ancestral amino acids present in high-altitude populations of speckled teal and yellow-billed pintail are underlined in bold, whereas ancestral amino acids present in the low-altitude populations are not underlined in plain text. Numbers refer to position in the β^A -chain (*HBB*) polypeptide subunit.

using the log-posterior odds of directional selection (Graham and McCracken 2019), whereas the present analysis of the globin clusters has utilized a quantile approach to identifying extreme outliers.

Next, we calculated F_{ST} unconventionally, comparing not within species, but between species to demarcate genomic regions appearing identical by descent. This was done by calculating F_{ST} between the high-altitude population of speckled teal and the high-altitude population of yellow-billed pintail, and between the low-altitude population of speckled teal and the low-altitude population of yellow-billed pintail. Corresponding Manhattan plots were produced for each pairwise interspecific comparison, including the complete α - and β -globin clusters and the 26 HIF-pathway genes. The rationale for this analysis is that pairwise F_{ST} comparisons between species as genetically divergent as speckled teal and yellow-billed pintail (i.e., 5.2% in mtDNA) should reveal a relatively uniform distribution of high F_{ST} values reflecting many fixed or nearly fixed differences. However, the expectation for recently introgressed loci is that these regions will show a lower F_{ST} closer to zero, reflecting identity by descent (IBD) and therefore qualitatively different from regions of the genome representing non-introgressed loci. If especially little time has elapsed for new fixed differences to accumulate since introgression and a strong selective sweep occurred, such loci should also exhibit low nucleotide diversity (π) around the selected region with high linkage disequilibrium (LD) and slower LD decay.

In such examples, IBD thus will be reflected by “valleys” of low F_{ST} as opposed to “peaks” *sensu* Feder et al. (2012). We predicted that these “valleys” should be revealed between comparisons of the introgressed high-altitude populations only, and isolated to only those targeted loci

that have actually introgressed, whereas non-introgressed regions of the genome should exhibit uniformly high F_{ST} and patterns of π and LD predicted above. Finally, we also predicted there should be correspondence between regions identified by this method of analysis and regions identified as introgressed in the ABBA–BABA analysis (see below). These analyses were performed for separate partitions of the data, including the α - and β -globin clusters and 26 HIF-pathway genes and contrasted with the genome-wide backdrop provided by the RAD-Seq data (Graham and McCracken 2019). VCF files were created separately to be fed into these analyses.

Tests for introgression (ABBA–BABA)

In conjunction with the F_{ST} -based approaches, we used the ABBA–BABA method to obtain Patterson’s D -statistic to test if a pattern of shared variation between groups can be better explained by hybridization than incomplete lineage sorting (ILS; Durand et al. 2011). Specifically, this method entails calculating the ratio of ABBA- to BABA-like four-taxa tree topologies, where the true tree topology is reflected by AABB, and ABBA and BABA reflect alternative topologies produced by ILS or introgression (Supplementary Fig. 2). In cases where introgression prevails, the ratio of ABBA:BABA \neq 1, whereas in cases of ILS the expectation is that ABBA:BABA = 1 because ILS is a stochastic process subject to the influence of genetic drift resulting in relatively evenly distributed alternative tree topologies. Finally, a positive D -statistic signifies an excess ABBA sites, whereas a negative D -statistic signifies an excess of BABA sites or vice versa. In these analyses, a significant excess of ABBA sites corresponding to a positive D -statistic was used to infer introgression between

high-altitude speckled teal and high-altitude yellow-billed pintail populations.

We performed the four-taxa D test in the software package Dsuite (Malinsky et al. 2021) using standard ABBA–BABA gene trees comprising four clades corresponding to each low- and high-altitude population (Supplementary Fig. 2). While the ABBA–BABA test is usually depicted in the context of a fully pectinate topology (((YP-H,YP-L),ST-H),ST-L), the correct topology for these taxa comprises two reciprocally monophyletic sister clades ((YP-H,YP-L),(ST-H,ST-L)) (Fig. 2). However, there is no reason to assume either topology to perform the introgression tests because the frequencies of ABBA vs. BABA sites are topology independent (i.e., calculated from the SNPs themselves, not reconstructed on internal branches), and a priori knowledge of the true topology is not required (Malinsky et al. 2021). Thus, here we use the ABBA–BABA test despite not having a pectinate topology since the differences in ABBA vs. BABA patterns are informative of introgression events under a topology with two reciprocally monophyletic sister clades. Using the Dtrios subprogram in Dsuite, we performed 20 jackknife replicates to calculate the D -statistic and record the number of ABBA and BABA allocated SNPs for each locus; n.b., averaging Dtrios output over the 20 jackknife replicates resulted in non-integer ABBA and BABA values. The z -scores were used to determine significant introgression using a conservative cutoff of $z = 3.3$ ($P < 0.001$) as recommend by Durand et al. (2011) and Ottenburghs et al. (2017), followed by adjusting α for multiple tests using the false discovery rate of Benjamini and Hochberg (1995). Because the D -statistic can have a large variance when applied to small genomic regions and can give inflated values in regions of low diversity, we also calculated global fD (Martin et al. 2015) using Dinvestigate for each locus, and also along a sliding window for the β -globin cluster. This was performed for (a) the full α - and β -globin clusters, (b) all 26 genes from the HIF pathway (Graham and McCracken 2019), and finally (c) the composite RAD-Seq comprising the ~2 million bp of RAD-clusters shared between speckled teal and yellow-billed pintail.

Inferring the direction of gene flow using the coalescent

To determine directionality of interspecific gene flow for the putatively introgressed β -globin locus, we performed Isolation-with-Migration coalescent analyses in the software IMA2 (Hey 2011). We used the 4-population model (Fig. 2) defined by two reciprocally monophyletic populations of speckled teal (ST-Low, ST-High) and two populations of yellow-billed pintail (YP-Low, YP-High). The following

parameters were used for population size (θ_{ST-H} , θ_{ST-L} , θ_{YP-H} , θ_{YP-L} , θ_{ST-ANC} , θ_{YP-ANC} , θ_{ANC}), bidirectional migration rates ($M_{ST-L>H}$, $M_{ST-H>L}$, $M_{YP-L>H}$, $M_{YP-H>L}$, $M_{ST-L>YP-L}$, $M_{YP-L>ST-L}$, $M_{ST-H>YP-H}$, and $M_{YP-H>ST-H}$, etc.), and time since divergence (t_0 , t_1 , and t_2 respectively, where t_0 is divergence between yellow-billed pintail populations, t_1 is time since divergence between speckled teal populations, and t_3 is divergence between species).

This analysis was executed on a 291 bp fully-phased capillary-sequenced segment of the β^A -globin (*HBB*) gene near the 3' end of the coding region incorporating 28 SNPs (21 parsimony informative) and the amino-acid substitutions Ala β 116Ser and Leu β 133Met previously characterized experimentally by Natarajan et al. (2015). The joint and independent effects of these amino-acid substitutions on intrinsic Hb-O₂ affinity are shown in Fig. 3 as illustrated by Natarajan et al. (2015). Phasing of this *HBB* sequence was performed in the software PHASE (Stephens et al. 2001), in conjunction with the 4-gamete test, as implemented in IMgc (Woerner et al. 2007) to identify a minimally non-recombining block of sequence comprised of ≤ 5 possible recombination events and haplotypes from each of the four low- and high-altitude populations (McCracken et al. 2009a). To illustrate these haplotype relationships, we constructed an allelic network for the 3' end of the β^A -globin using the median-joining algorithm in NETWORK 4.6 (Bandelt et al. 1999). Due to the short length of the RAD-Seqs, and small number of SNPs per locus in this RAD-Seq dataset, coalescent analysis was not performed for the RAD-Seq data. Instead, we used the corresponding sequences located on different chromosomes from the coalescent analyses published by McCracken et al. (2009a), which included fully-phased allele sequences from 5-intronic loci with no detectable recombination (see McCracken et al. (2009a) for the corresponding GenBank accession numbers and capillary sequencing methods).

IMA2 was run first with wide uninformative priors. Analyses were then conducted with more narrow uniform priors that encompassed the posterior distributions from the preliminary runs (e.g., $\theta = 10$ –50, $M = 100$, and $t = 2$ –5). The Markov chain Monte Carlo was run five times independently for as many as 10 million steps in the case of intronic DNA sequences, sampling the posterior distribution every 150 steps, with a burn-in of 500,000 steps. All runs included 20 chains with a geometric heating scheme. Thus, we were able to obtain estimates of bidirectional migration rates between all pairs of low- and high-altitude populations, the effective population size parameter ($\theta = 4N_e\mu$) for each extant and ancestral population, and time since divergence using these different data sets.

Table 1 Tests for introgression across the α - and β -globin gene clusters, 26 HIF-pathway genes, and genome-wide RAD-Seq analysis of 81,957 SNPs.

Locus	Dinvestigate		Dtrios					F_{ST} within species		F_{ST} between species	
	D	μD	D_{trios}	ABBA	BABA	z -score	P value	Speckled teal	Yellow-billed pintail	Low vs. Low	High vs. High
ACE	-0.013	-0.012	0.022	13.1	12.6	1.287	0.099	0.079	0.016	0.364	0.380
ANGPT1	0.026	0.014	0.003	25.6	25.4	0.048	0.481	0.059	0.016	0.568	0.585
ANGPT2	0.011	0.003	0.006	5.2	5.2	0.210	0.417	0.060	0.000	0.557	0.622
ARNT	0.178	0.009	0.047	0.1	0.1	0.220	0.413	0.190	-0.018	0.567	0.585
ARNT2	-0.008	0.005	0.004	25.2	25.0	0.266	0.395	0.057	-0.003	0.499	0.501
ARNTL	0.374	0.050	0.292	0.2	0.1	2.333	0.010	0.112	0.128	0.745	0.654
CLOCK	-0.040	-0.003	0.060	0.8	0.7	0.490	0.312	0.134	0.018	0.666	0.773
CUL2	-0.152	-0.010	0.017	1.4	1.4	0.183	0.428	0.082	0.012	0.507	0.554
<i>EGLN1</i>	<i>0.052</i>	<i>0.005</i>	<i>0.023</i>	<i>12.7</i>	<i>12.1</i>	<i>0.186</i>	<i>0.426</i>	<i>0.596*</i>	<i>0.403*</i>	<i>0.589</i>	<i>0.879</i>
EGLN3	-0.128	-0.017	0.086	1.5	1.3	1.334	0.091	0.266	-0.058	0.681	0.698
<i>EPAS1</i> (<i>HIF2α</i>)	<i>0.064</i>	<i>0.006</i>	<i>0.063</i>	<i>19.1</i>	<i>16.8</i>	<i>0.851</i>	<i>0.197</i>	<i>0.508*</i>	<i>0.444*</i>	<i>0.525</i>	<i>0.865</i>
HIF1A	0.096	0.017	0.015	1.5	1.4	0.115	0.454	0.189	0.193	0.527	0.530
HSP90AA1	0.476	0.009	0.348	0.1	0.0	0.951	0.171	0.182	0.106	0.708	0.705
MTOR	0.093	0.038	0.052	21.4	19.3	0.854	0.197	0.082	0.019	0.476	0.465
NOS1	-0.046	-0.041	0.005	32.6	32.2	0.422	0.336	0.068	0.024	0.302	0.376
NOS2	-0.200	0.127	0.049	2.7	2.4	0.640	0.261	0.042	0.142	0.622	0.569
P4HA1	0.025	0.021	0.016	15.3	14.9	0.887	0.188	0.051	0.005	0.446	0.466
P4HA2	0.081	0.137	0.040	20.2	18.7	2.286	0.011	0.090	-0.012	0.372	0.269
P4HA3	0.075	0.101	0.033	6.8	6.4	0.666	0.253	0.027	0.032	0.252	0.294
PDHA1	-0.303	-1.746	0.051	6.2	5.6	0.854	0.196	0.030	0.039	0.484	0.554
PPARA (α)	0.088	0.014	0.042	1.1	1.0	0.000	1.000	0.179	0.022	0.878	0.883
PPARD (δ)	0.042	0.041	0.013	4.0	3.9	0.408	0.342	0.042	0.017	0.215	0.284
PPARG (γ)	0.000	0.000	0.030	0.2	0.2	0.095	0.462	0.092	-0.018	0.778	0.764
TCEB3	0.037	0.029	0.022	5.6	5.4	0.896	0.185	0.032	-0.017	0.283	0.367
THRB	-0.054	-0.024	0.017	12.4	12.0	1.093	0.137	0.062	0.017	0.485	0.493
VEGF	-0.018	-0.006	0.012	4.5	4.4	0.177	0.430	0.264	0.029	0.469	0.533
β-globin cluster	0.775	0.683	0.203	43.7	28.9	4.586	0.000002	0.425*	0.363*	0.163	0.277
α -globin cluster	-0.038	-0.005	0.006	3.4	3.3	0.057	0.477	0.338*	0.003	0.558	0.664
RAD-Seq	0.002	0.001	0.006	581.3	574.0	2.442	0.007	0.062	0.008	0.410	0.473

We used D -statistics (ABBA–BABA) to test if shared polymorphism between non-sister taxa (high-altitude speckled teal and high-altitude yellow-billed pintail) signifies adaptive introgression, or alternatively incomplete lineage sorting (ILS). The null expectation for equal ABBA:BABA proportions under ILS is that $D = 0$, whereas a significantly positive or negative D corresponds to introgression. Results are derived from 20 jackknife replicates in Dsuite (Malinsky et al. 2021) using the Dinvestigate and Dtrios programs showing the number of SNPs supporting either ABBA or BABA patterns; the averaging Dtrios performs over the 20 replicates results in non-integer ABBA and BABA values. Only the β -globin cluster shown in bold text has a z -score >3.3 standard deviations corresponding to a P value for two-tailed test <0.001 , with more ABBA patterns than expected by chance, reflecting introgression between the high-altitude populations. At right is the average pairwise F_{ST} for each locus, between populations within each species (low- vs. high-altitude) and between each species (low altitude vs. low altitude and high altitude vs. high altitude, respectively). The rows given in italic denote divergently-selected candidate loci with high F_{ST} outliers marked with asterisks (between populations within species) relative to background F_{ST} from the genome-wide analysis.

Signature of selective sweep and LD on introgressed loci

To further examine evidence for a selective sweep on introgressed alleles in the high-altitude population, we calculated three additional population genetic parameters,

including π , the allele-frequency spectrum for derived alleles (dAFS), and decay of LD. Following a selective sweep on introgressed adaptive variation in the recipient high-altitude population, we predicted that π would be lower, the dAFS should be more skewed toward low- and high-frequency polymorphisms, and LD should be higher

Table 2 Putatively-convergent amino-acid replacements in embryonic *HBE* (ϵ) globin gene of high-altitude speckled teal and yellow-billed pintail populations.

Full seq. pos.	Coding pos.	AA pos.	High AA	Low AA	F_{ST} Speckled teal	F_{ST} Yellow-billed pintail
66	40	14	Gly*	Ser ^a	1.000	0.778
69	43	15	Leu	Ile ^a	1.000	0.778
75	49	17	Gly*	Ser ^a	1.000	0.733
1468	376	126	Thr*	Ala ^a	1.000	0.753

Shown are (1) base-pair position on the full sequence (including introns and exons), (2) base-pair position on the concatenated exon sequence, (3) amino-acid position on the polypeptide sequence, (4) identities of amino acids segregating at high frequency in each population, and (5) F_{ST} between low- and high-altitude populations of each species. Asterisks further denote a change in amino-acid polarity.

^aIndicates the ancestral amino acid also present in the chicken genome.

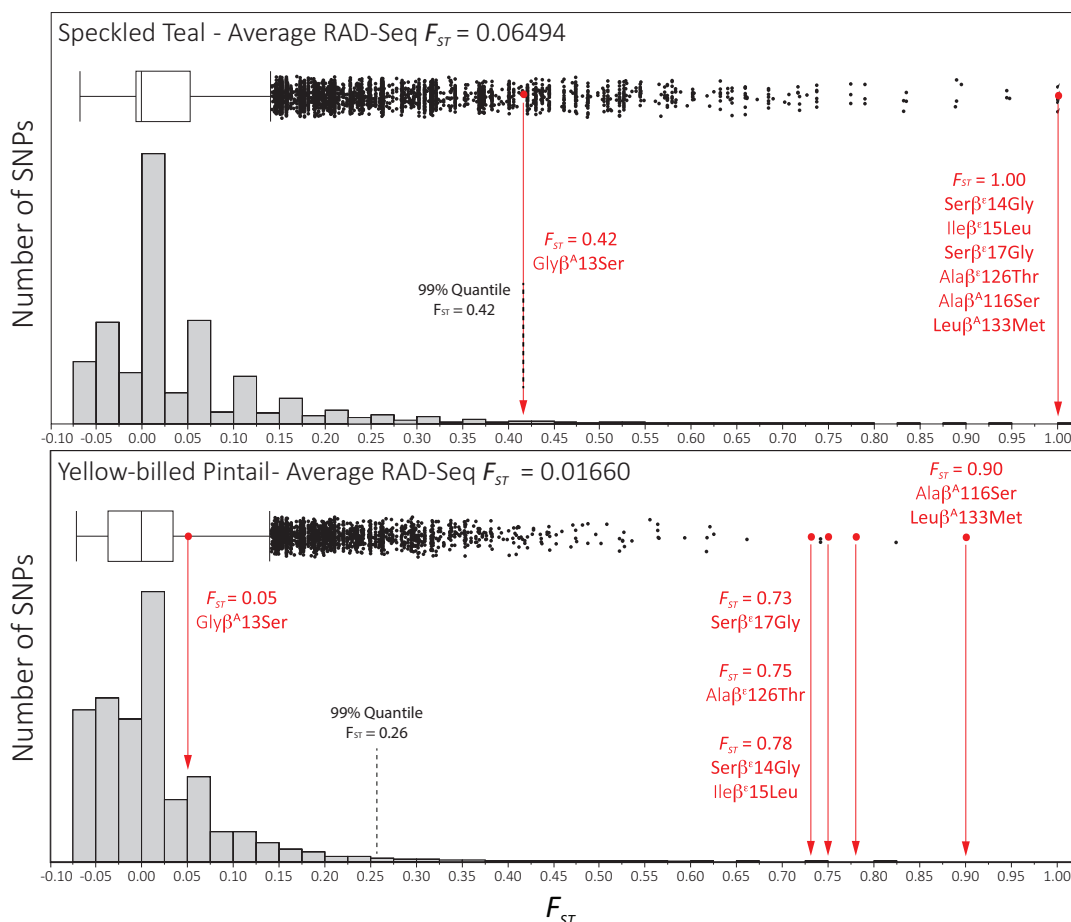


Fig. 4 Amino-acid polymorphisms in the adult β^A and embryonic ϵ -globin genes (*HBB* and *HBE*, shown in red) are extreme outliers in genomic scans of allele-frequency differentiation between low- and high-altitude populations of speckled teal and yellow-billed pintail. The histograms and box plots depict genome-wide

distributions of SNP-specific F_{ST} and 99% upper quantiles based on the RAD-Seq survey of genomic polymorphism (speckled teal = 47,731 SNPs, yellow-billed pintails = 49,670 SNPs). Arrows indicate extreme outliers resulting in the labeled amino-acid replacements.

whereas LD decay should be slower. For the β -globin cluster, π and LD decay were calculated, whereas π , dAFS, and LD decay were calculated for *EPAS1* and *EGLN1*. The dAFS was also calculated for the RAD-Seq for a genome-wide comparison outside the selected regions. VCFtools (Danecek et al. 2011) was used to calculate π and the dAFS. PopLDdecay (Zhang et al. 2019) was used to plot LD decay.

Results

α - and β -globin assembly statistics

Coverage was calculated for both the Bowtie2 (contiguous sequence assembly) and BWA (VCF pipeline) assemblies. For Bowtie2, the average coverage for the α -globin cluster was 1,022 \times with an average of 110,779 reads per individual

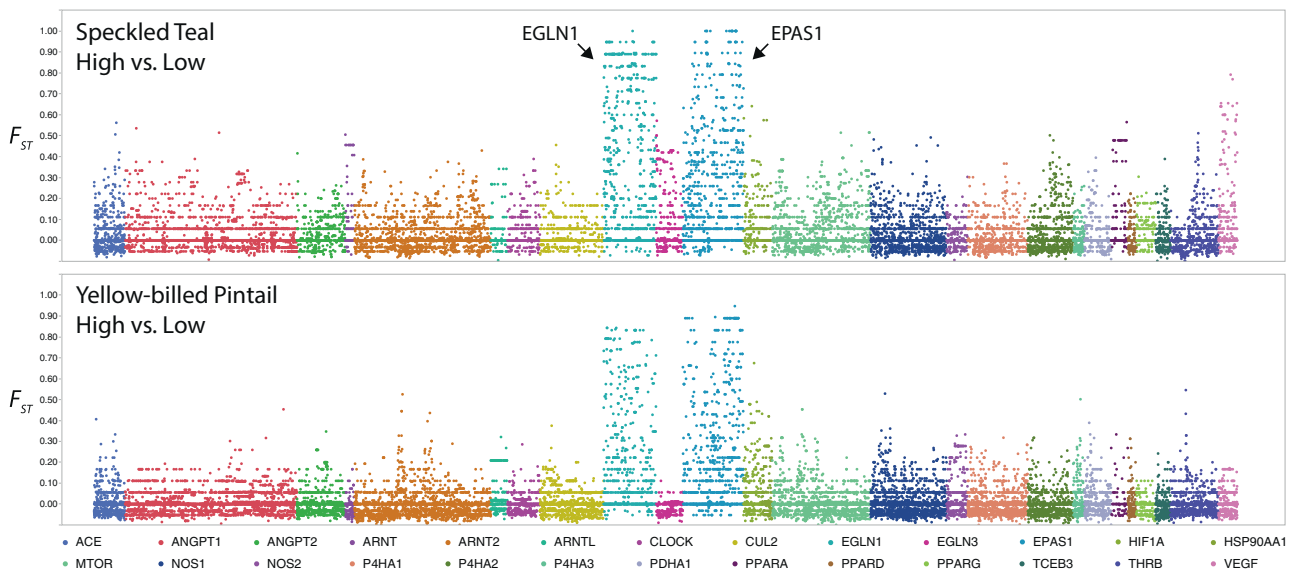


Fig. 5 Manhattan Plots showing locus-by-locus F_{ST} calculated pairwise for all SNPs for the 26 HIF-pathway genes between low- vs. high-altitude populations of speckled teal (within species comparison) and low- vs. high-altitude populations of yellow-billed

pintail (within species comparison). The HIF-pathway outliers *EPAS1* and *EGLN1* identified by Graham and McCracken (2019) are shown with arrows. Individual gene identities are labeled with corresponding colors below the plots.

assembly, containing an average 49.6% GC content. The average coverage for the β -globin cluster was 275 \times with an average of 120,586 reads per individual assembly, containing an average 49.11% GC content. For BWA, across the two species, the per base-pair coverage was 279 \times covering 92.5% of the β -globin cluster, and 525 \times covering 97.5% of the α -globin cluster. Assembly statistics for the 26-HIF pathway genes were published previously (Graham and McCracken 2019).

Introgression of the β globin

To test for introgression, D -statistics were calculated for (a) the α - and β -globin clusters, (b) the 26 genes from the HIF-pathway, two of which had shown prior evidence for convergent evolution in these same species (Graham and McCracken 2019), and (c) the combined RAD-Seq clusters comprising 81,957 SNPs. We found strong evidence supporting adaptive introgression in the β -globin genes ($D = 0.78$, $\hat{\mu}D = 0.68$, 43.7 ABBA vs. 28.9 BABA sites; $z = 4.586$, $P < 0.000002$) after correcting α for multiple tests using the false discovery rate (Benjamini and Hochberg 1995), but not in any of the α -globin genes ($D = -0.04$, $\hat{\mu}D = -0.005$, 3.4 ABBA vs. 3.3 BABA sites; $z = 0.057$, $P < 0.477$; Table 1). Of the 26 genes in the HIF Pathway, D -statistics were not significant at the $z = 3.3$ (uncorrected $P < 0.001$) level for any locus (Table 1). *EGLN1* ($D = 0.05$, $\hat{\mu}D = 0.005$, 12.7 ABBA vs. 12.1 BABA sites) and *EPAS1* ($D = 0.06$, $\hat{\mu}D = 0.006$, 19.1 ABBA vs. 16.8 BABA sites), which are on the same chromosome, had z -scores < 0.86 , corresponding to uncorrected P values > 0.19 , thus providing no evidence that

high-altitude sequences in these two species are introgressed (Table 1). While some other HIF-pathway loci had elevated D -statistics (*ARNTL*, $D = 0.37$; *HSP90AA1*, $D = 0.48$), no loci had uncorrected z -scores > 2.58 ($P > 0.01$), and $\hat{\mu}D$ was low or very low for all HIF-pathway loci. Finally, there was no significant evidence of genome-wide introgression between the two species in the pooled RAD-Seq clusters ($D = 0.002$, $\hat{\mu}D = 0.001$, 581.3 ABBA vs. 574.0 BABA sites; Table 1).

High F_{ST} outlier loci

In the embryonic *HBE* (ϵ) gene of yellow-billed pintail and speckled teal there are four novel amino-acid changes that have evolved convergently in high-altitude populations of speckled teal and yellow-billed pintail. These convergent changes include Ser ϵ 14Gly, Ile ϵ 15Leu, Ser ϵ 17Gly, and Ala ϵ 126Thr (Table 2). Compared to the background F_{ST} between lowland and highland populations in these two species ($F_{ST} = 0.06494$ and 0.01660, respectively), F_{ST} for these ϵ -globin polymorphisms ($F_{ST} = 0.733$ –1.000) falls far outside the 99% quantiles for each species (Fig. 4), ranking them among the most extreme outliers for these populations, a pattern qualitatively similar to the *HBB* (β^A) subunit amino acid changes experimentally shown to increase Hb-O₂ affinity (Fig. 3).

In the embryonic *HBE* (ϵ) globin subunit, three of the four nonsynonymous changes resulted in a difference of polarity in their respective side-chains: two involve a change from a polar, hydrophilic amino acid (Ser) to an aliphatic amino acid (Gly), whereas the other is from

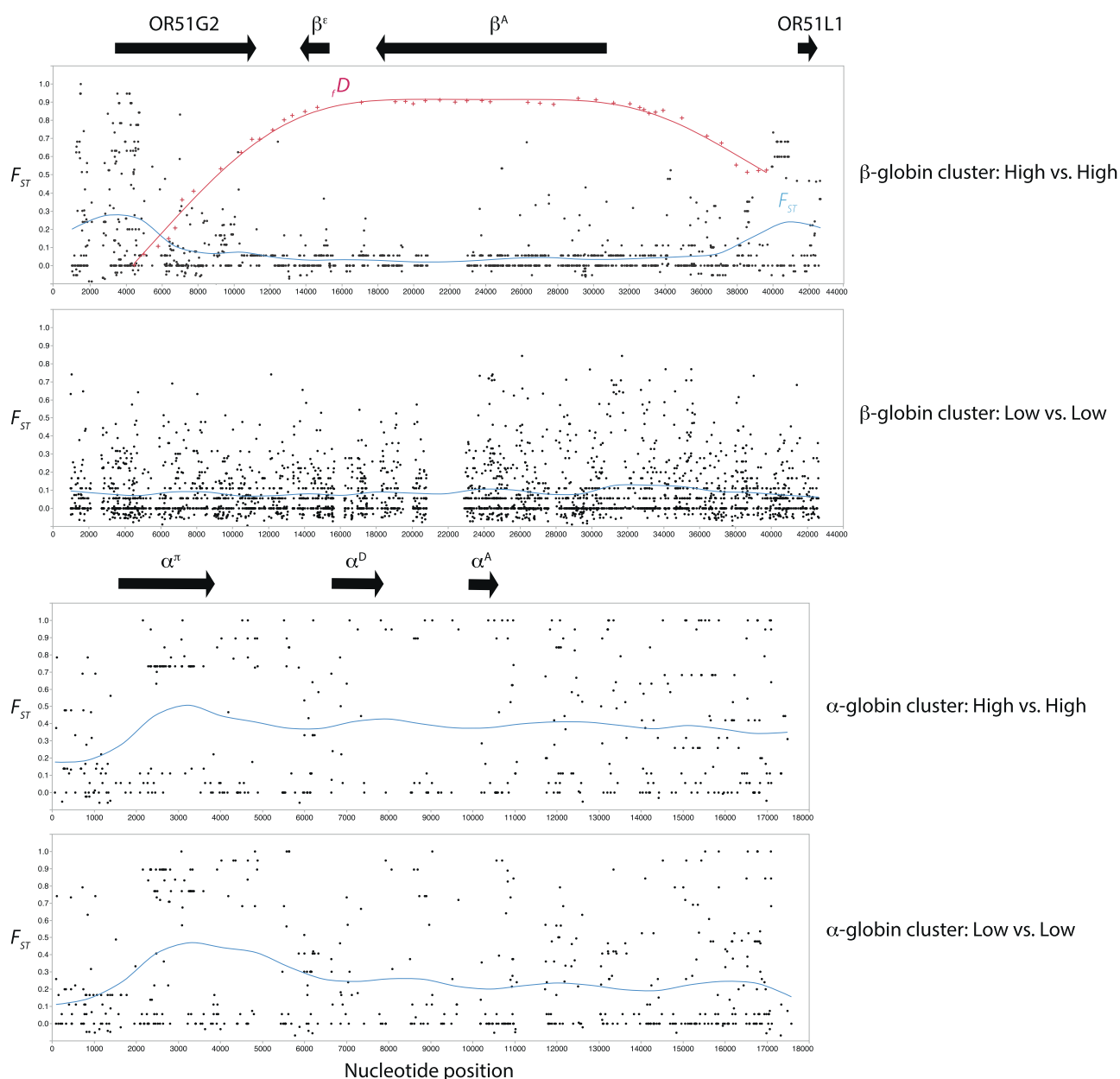


Fig. 6 Manhattan Plots showing locus-by-locus F_{ST} calculated pairwise for all SNPs for the α - and β -globin clusters between high-vs-high (between species comparisons) and low-vs-low (between-species comparisons) populations of speckled teal vs. yellow-billed pintail (bottom and top, respectively). The high vs. high comparison of the β -globin cluster differs qualitatively from the low vs. low comparison, as it also does from both α -globin comparisons. Specifically, the high-vs.-high β -globin shows a “valley” of low F_{ST} values centered on the β^A - and ϵ -globin genes. The J_D -statistic calculated in Dsuite (Malinsky et al. 2021) and shown in red furthermore shows a broad peak spanning most of the β -globin cluster

and corresponding to the region of introgressed sequence. The two graphs at the bottom for the α -globin cluster do not differ qualitatively from each other, as each display long tracts of high F_{ST} values across the linkage group. Finally, the olfactory receptor genes $OR51G2$ and $OR51L1$ flanking the 5' and 3' ends of the β -globin cluster exhibit high F_{ST} compared to the intervening low F_{ST} β -globin genes in the high-vs.-high comparison, but this pattern is not evident in the low-vs.-low comparison. The smoothed lines illustrating average J_D and F_{ST} were calculated using sliding windows and default parameters in JMP v14. Positions of α - and β -globin genes on the scaffolds are shown at the top of each set of panels.

an aliphatic, hydrophobic amino acid (Ala) to a polar, hydrophilic amino acid (Thr). For speckled teal, all four sites are fixed for different amino acids between low- and high-altitude populations ($F_{ST} = 1.000$), whereas allele frequencies are highly divergent but not fixed in yellow-billed

pintail ($F_{ST} = 0.733\text{--}0.778$) (Table 2; Fig. 4). The first three of these at positions 14, 15, and 17 occur in the first helical turn of the first α subunit, and the last one at position 126 in the last helical turn. These amino-acid substitutions were located on the human embryonic protein structure

(Protein Data Bank PDBe > 1a9w). No bird embryonic structure is currently available to study these on an avian background. Among the 26 genes in the HIF pathway, only *EPASI* and *EGLN1* exhibit unusually high F_{ST} ($F_{ST} = 0.715\text{--}0.910$), qualitatively dissimilar to any other HIF-pathway locus or the genome-wide background F_{ST} , and the pattern for these two loci is similar, involving outliers in the same *EPASI* exon (exon 12), for both speckled teal and yellow-billed pintail (Fig. 5; see Figs. 2, 3 in Graham and McCracken 2019). The two genes are both located on the same chromosome of the duck genome (Chr3: 37,562,797–37,679,832 and Chr3: 42,539,456–42,573,466 respectively), about 5 Mb apart and with the undifferentiated gene *VEGFA* (Chr3: 33,313,820–33,336,832) also analyzed in this study ~4 Mb to the other side of *EPASI* relative to *EGLN1*.

Low F_{ST} outlier loci

Our between-species F_{ST} scan for low-value valleys revealed that the β -globin cluster shows a pronounced low- F_{ST} valley between the high-altitude populations of each species corresponding to the same region of the sequence with elevated d , but not between the low-altitude populations of each species (Fig. 6). This pattern spans all adult and embryonic β -globin genes in the cluster and is clearly demarcated by high F_{ST} regions of sequence in the two olfactory receptor genes (*OR51G2* and *OR51L1*) flanking the 5' and 3' ends of the β -globin cluster. By comparison the α -globin cluster, which does not exhibit D -statistics characteristic of introgression, shows no such pattern (Fig. 6). In the α -globin cluster, high F_{ST} polymorphisms reflecting many fixed differences are uniformly distributed across the entire α cluster for the comparisons of high- vs. high-altitude populations and low- vs. low-altitude populations. The β -globin therefore appears identical by descent because it is introgressed between high-altitude populations of the two species, whereas the α -globin cluster does not.

By contrast, we found no patterns of low F_{ST} similar to the β -globin cluster for either *EPASI* or *EGLN1* or the other HIF-pathway loci. For both the high- vs. high-altitude and low- vs. low-altitude population comparisons, many fixed differences were apparent across all 26 loci, including *EPASI* and *EGLN1* (Supplementary Fig. 3A, B). Therefore, there is no qualitative pattern consistent with IBD, as confirmed by the ABBA–BABA D -statistics. However, the 3' half of *EPASI* and all of *EGLN1* exhibit disproportionately fewer loci with intermediate F_{ST} values ($0.25 < F_{ST} < 0.75$) than other HIF-pathway loci (Supplementary Fig. 3C; see also Supplementary Fig. 4). This pattern is apparent in the high- vs. high-altitude pairwise F_{ST} comparison of these two loci but not the low- vs. low-altitude pairwise F_{ST} comparison.

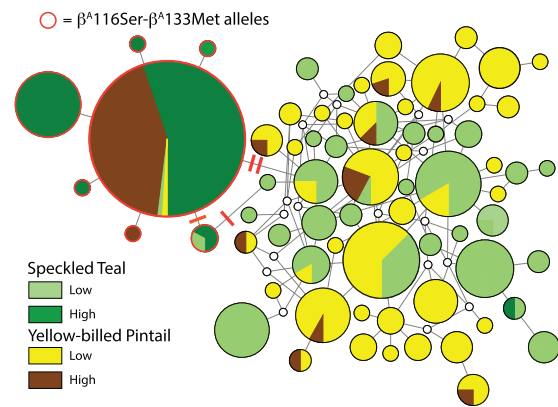


Fig. 7 Haplotype network for the β -globin cluster showing the 3' coding region of capillary-sequenced *HBB* (β^A), including 291 bp spanning part of intron 2 and all of exon 3 including the Ala β^A116 Ser and Leu β^A116 Met amino-acid substitutions previously characterized experimentally by Natarajan et al. (2015). Bold edged circles represent haplotypes with putatively adaptive amino-acid substitutions appearing identical by descent (IBD) between high-altitude populations of the two duck species. Six β^A116 Ser- β^A116 Met haplotypes were identified. One haplotype (at the center) is common, and the other five are singletons or occur at low frequency and radiate from the central common haplotype in a star-like pattern. Five of these rare β^A116 Ser- β^A116 Met haplotypes are present in speckled teal, whereas only one is present in yellow-billed pintail, consistent with the coalescent model data indicating that speckled teal has an older ancestry in the highlands.

Identity by descent (IBD)

The allelic network we constructed for the β -globin cluster represents 291 bp from the 3' end of the *HBB* (β^A) globin gene (Fig. 7). It includes part of intron 2 and all of exon 3 including the Ala β^A116 Ser and Leu β^A116 Met amino-acid substitutions determined by Natarajan et al. (2015) to increase intrinsic Hb-O₂ affinity (Fig. 3). It clearly illustrates a pattern of IBD, reflecting a selective sweep on high-altitude populations sharing the same haplotypes at approximately equal frequencies. One β^A116 Ser- β^A116 Met haplotype constituting 99.3% of high-altitude speckled teal and 89.2% of high-altitude yellow-billed pintail predominates in the highlands. Six other β^A116 Ser- β^A116 Met haplotypes occur, but they exist at low frequencies (4 are singletons in our dataset) in the highlands and differ by a single nonsynonymous substitution, radiating from the central common haplotype in a star-like pattern (Fig. 7). Interestingly, five of these rare β^A116 Ser- β^A116 Met haplotypes are present in speckled teal only, whereas only one is present in yellow-billed pintail. This ratio of singletons between species is consistent with speckled teal having older ancestry in the highlands than the yellow-billed pintail.

Direction of the β -globin introgression

In the allelic network for the 291 bp from the 3' end of the *HBB* (β^A) globin gene, many yellow-billed pintails at high altitude share haplotypes with yellow-billed pintails at low altitude, whereas few speckled teal at high altitude share haplotypes with speckled teal at low altitude (Fig. 7). This pattern suggests that yellow-billed pintails exhibit higher gene flow between low and high altitude than speckled teal, a pattern which was corroborated previously and is consistent with the highly volant characteristics of this species (see McCracken et al. 2009c). We also know from depth of mtDNA divergence (~3 vs. 0%; see also Supplementary Fig. 1) that the high-altitude speckled teal population (a separate subspecies, and phenotypically distinct from the low-altitude subspecies) probably was established at high altitude earlier and for much longer than the yellow-billed pintail population (not a separate subspecies, and not phenotypically distinct).

Using the same 291 bp capillary-sequenced alleles from the 3' end of the *HBB* (β^A) globin shown in Fig. 7, we conducted 4-population coalescent analyses in IMA2 (Hey 2011) for the 3' *HBB* (β^A) globin. For comparison to putatively “non-introgressed” loci reflecting the genomic background, we used fully-phased sequence data from five intronic loci sampled from different chromosomes (McCracken et al. 2009a). Our analysis reveals that while there was no evidence of interspecific gene flow in the intronic loci analyzed by McCracken et al. (2009c), the 3'-end of the *HBB* (β^A) globin introgressed into the high-altitude yellow-billed pintail population from high-altitude speckled teal, but not the reverse (Supplementary Table 1). As expected, no nonzero estimates of gene flow were found in any other direction for all possible pairwise comparisons of migration rates between species. In sum, our finding strongly confirms that yellow-billed pintail acquired its present-day $\beta^A116\text{Ser}$ and $\text{Leu}\beta^A116\text{Met}$ -containing globin variants through adaptive introgression. Furthermore, it was the more genetically and morphologically divergent species that probably colonized high altitude earlier in history (speckled teal) that transferred putatively adaptive genetic variation to the species that probably colonized high altitude more recently (yellow-billed pintail).

Signature of selective sweep and LD on introgressed loci

Contrary to our prediction that nucleotide diversity (π) would be lower in high-altitude yellow-billed pintail than high-altitude speckled teal, we did not find this to be the case for the β -globin cluster for which π was similar between species and actually slightly higher in the high-altitude yellow-billed pintail population (Supplementary Fig. 5).

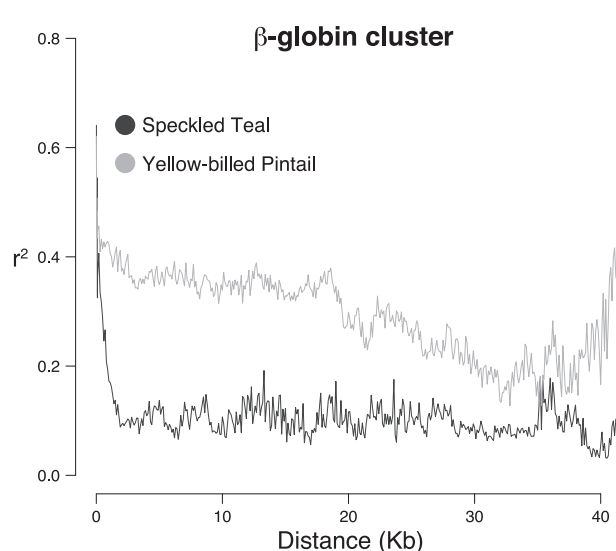


Fig. 8 Decay of linkage disequilibrium (LD) plotted for the high-altitude populations of speckled teal and yellow-billed pintail, respectively, across the β -globin cluster. Following a selective sweep in both species at different times in history, linkage disequilibrium has persisted to the present in yellow-billed pintail and falls off slower compared to the older high-altitude speckled teal population.

However, nucleotide diversity was lower in high-altitude yellow-billed pintail for the two HIF-pathway *EPAS1* and *EGLN1*. As predicted, the dAFS of *EPAS1* and *EGLN1* were more skewed in high-altitude yellow-billed than high-altitude speckled teal (Supplementary Fig. 6). Whereas yellow-billed pintail dAFS was qualitatively different from the RAD-Seq dAFS, the speckled teal dAFS was not, suggesting a more recent selective sweep in yellow-billed pintail. Finally, β -globin LD was higher and LD decay was slower for the high-altitude yellow-billed pintail population than for the high-altitude speckled teal population (Fig. 8). A similar LD pattern also occurred in *EPAS1*, but not in *EGLN1* (Supplementary Fig. 7).

Discussion

In this study, we present a comprehensive analysis of the two Hb gene clusters and 26 HIF-pathway genes of two Andean duck species. This includes previously unexamined embryonic Hb genes (*HBE*, *HBZ*), and the *EPAS1* and *EGLN1* genes, the latter which have been a prominent focus of human genetic studies. Our results reveal previously undescribed adaptive introgression of the β -globin cluster between the two high-altitude populations of speckled teal and yellow-billed pintail.

Adaptive introgression has been identified across plants and animals as an important source of genetic variation allowing organisms to quickly adapt to environmental

pressures (Goulet et al. 2017; Hedrick 2013; Racimo et al. 2015). One prominent example is in Tibetan populations, where adaptation to hypoxia at high-altitude in the HIF-pathway was acquired through interbreeding of ancient human populations with Denisovans (Huerta-Sánchez et al. 2014; Jeong et al. 2014). Similar patterns were seen in Tibetan mastiffs and cattle, which showed adaptive introgression of the HIF pathway (Li et al. 2014; Miao et al. 2017; Wu et al. 2018). There is also now evidence that adaptive introgression involving the β globin and increased Hb-O₂ binding affinity in the Tibetan mastiff resulted from a legacy of selective interbreeding and hybridization with the Tibetan wolf (*Canis lupus laniger*) (Signore et al. 2019). The HIF-pathway genes (*EPASI*, *EGLN*) and Hb are notable because of their key role in mediating the core physiological response to changes in O₂ availability (Storz and Cheviron 2016), and are frequent targets for selection in high-altitude living (Hanaoka et al. 2012; Li et al. 2014; Qiu et al. 2012; Qu et al. 2013; Scheinfeldt and Tishkoff 2010; Simonson et al. 2012; Wang et al. 2014). Often, genetic variation in *EPASI* and *EGLN1* are associated with changes in Hb concentration, among other physiological traits (Ai et al. 2014; Beall et al. 2010; Lorenzo et al. 2014; Peng et al. 2017; Xiang et al. 2013; Yi et al. 2010). Our study shows exciting new evidence of adaptive introgression in this key cluster of genes influencing Hb-O₂ affinity and therefore likely arterial-O₂ saturation (e.g., Tate et al. 2017), and it is only the second documented example of adaptive introgression involving high-altitude hemoglobin variants (see also Signore et al. 2019).

Our results also suggest additional potential mechanisms of adaptations to high altitude through embryonic Hb in speckled teal and yellow-billed pintail. Based on currently published structure of embryonic human Hb, the locations of amino acid changes found in this study are peripheral to the locations of metal/heme binding. However, this does not preclude the possibility of those amino-acid replacements having functional effects on Hb-O₂ binding affinity, as even slight changes resulting in small structural differences such as the elimination of a single H-bond can cause profound changes in intrinsic Hb-O₂ affinity (Natarajan et al. 2015). Embryonic Hb thus may be under even stronger selection for increased O₂-binding at altitude because of egg-shell impermeability, which in numerous species has evolved to maximize O₂ availability to the developing embryo at the expense of water loss (Carey et al. 1994; Carey et al. 1989). Further surveys of embryonic Hb are thus warranted but potentially more difficult to undertake than for adult-expressed Hbs, as nests, eggs, and young offspring of some high-altitude species can be difficult to locate and must be sampled at known developmental intervals to purify the embryonic Hb isoforms.

Qualitative differences between the β -globin cluster and HIF-pathway

Our analysis revealed strong qualitative differences between the β -globin cluster and HIF-pathway genes. The ABBA-BABA analysis strongly supported introgression of the β -globin cluster but not *EPASI* or *EGLN1*, which nonetheless are under strong directional selection in these taxa (Table 1). When we calculated pairwise F_{ST} between high-altitude populations of both species (Supplementary Figs. 3, 4), a pronounced F_{ST} “valley” reflecting IBD was apparent for the β -globin cluster in the high-altitude yellow-billed pintail population, but no such F_{ST} “valley” was observed for *EPASI* and *EGLN1*. In the case of the β -globin cluster, detection of this F_{ST} “valley” was revealed by sequencing well outside the β -globin cluster into adjacent flanking olfactory receptor genes that do not exhibit the same signature of introgression (i.e., not IBD). Our sliding window ABBA-BABA analysis further revealed that f_D , which is more robust to regions of low genomic diversity, is inversely related to F_{ST} in this region (Fig. 6).

Nucleotide diversity (Supplementary Fig. 5) was lower and the dAFS (Supplementary Fig. 6) was more skewed in *EPASI* and *EGLN1* of the high-altitude population of yellow-billed pintail compared to the high-altitude population of speckled teal, despite the two species having similar census population sizes and therefore possibly subject to similar magnitudes of genetic drift (yellow-billed pintail is probably more influenced by gene flow between the lowlands and highlands). Therefore, these findings may be consistent with a more recent selective sweep on high-altitude yellow-billed pintail (as opposed to high-altitude speckled teal), followed by the fixation of rare variants. Like the β -globin cluster, this conclusion is also reinforced by higher LD and slower LD decay in the *EPASI* of high-altitude yellow-billed pintail (Supplementary Fig. 7). Most evidence for the β -globin cluster and at least some evidence for *EPASI* and *EGLN1* thus point towards a more recently diverged yellow-billed pintail population, and in the case of the β -globin cluster, the recipient of adaptive genetic variation from the older diverged speckled teal population (Fig. 2).

In sum, the strong qualitative differences we observed between the β -globin cluster and HIF-pathway genes suggest introgression of the β -globin cluster followed by a selective sweep on mutations increasingly Hb-O₂ affinity (Fig. 3) may have occurred recently in the history of this pair of species. This is consistent with the remarkable IBD between β -globin alleles from the two high-altitude populations (Fig. 7), as well as pronounced LD present in yellow-billed pintails (Fig. 8). For the β -globin cluster, mutation and recombination have so far done little to erase the signature of a selective sweep on this locus in high-altitude yellow-billed pintails.

Selection in the face of countervailing gene flow

One important confounding factor in our analysis is gene flow between low- and high-altitude populations of each species. It is a general rule that local adaptation occurs in the face of countervailing gene flow when the selection coefficient $s > m$ the migration rate (Slatkin 1987). This is because selection is locus specific, whereas population subdivision and migration affect polymorphisms genome wide (Peters et al. 2012; Rieseberg and Burke 2001; Wu 2001). In a previous study, it was found that gene flow into the high-altitude population of yellow-billed pintails from low-altitude yellow-billed pintails was an order of magnitude greater than gene flow into the high-altitude population of speckled teal from low-altitude speckled teal (McCracken et al. 2009a). This gene flow has predictable effects, principally contributing to augmented N_e . This might be one factor explaining why we found similar π between high-altitude yellow-billed pintails and high-altitude speckled teal for the β -globin cluster (Supplementary Fig. 5). Many yellow-billed pintails at high altitude share β -globin haplotypes with yellow-billed pintails at low altitude, whereas few speckled teal at high altitude share haplotypes with speckled teal at low altitude (Fig. 7). Numerous yellow-billed pintails were heterozygous for $\beta^A116\text{Ser}$ and $\beta^A116\text{Met}$, and most were sampled at the periphery of the geographic range of the high-altitude yellow-billed pintail population (McCracken et al. 2009c). No individuals homozygous for the low-altitude ancestral alleles $\beta^A116\text{Ala}$ and $\beta^A116\text{Leu}$ were found at high-altitude sites in any of these extensive surveys (McCracken et al. 2009a; McCracken et al. 2009b; McCracken et al. 2009c), suggesting that fitness of β -globin heterozygotes may be relatively high at altitude due to codominant expression of Hb, and that such heterozygous individuals may not be adversely affected at low altitude either.

Propensity to hybridize

Interspecific hybridization is common in plants and animals, especially in waterfowl (Anatidae), which hybridize extensively (Johnsgard 1960; Randler 2008; Tubaro and Lijtmaer 2002). A principal factor contributing to hybridization in waterfowl occurs when one species is rare and the other is common in sympatry (Randler 2002)—otherwise known as the Hubbs Principle, or “Desperation Hypothesis” (Hubbs 1955), which predicts that with skewed mating opportunities the rarer species is more likely to mate with heterospecifics. Here we report adaptive introgression resulting from interspecific hybridization and backcrossing between two waterfowl species that presently coexist in broad sympatry and mixed flocks throughout the Andean highlands of South America.

This is not the first documented hybridization between these species. In previously documented examples of hybridization between these two species on remote oceanic islands, one species was abundant and the other was extremely rare (McCracken and Wilson 2011; McCracken et al. 2013). Our finding of adaptive introgression in the Andean highlands raises the question about what a hybridization event leading to adaptive introgression might have looked like. If in the past, speckled teal were previously already established and abundant in the Andean highlands, whereas yellow-billed pintails were occasional sojourners to high altitude, scarcity of conspecific mating opportunities could possibly have contributed to hybridization and backcrossing between these two species following the range shift into a novel environment. Subsequent to adaptive introgression, yellow-billed pintails, unlike speckled teal, would not have had to wait for adaptive de novo mutations to enable successful colonization of high-altitude environments in the wake of such a founder event, but rather acquired beneficial alleles from the standing variation of another species that experienced a selective sweep at the same locus much earlier in its history, leading to faster evolution. This study of two Andean waterfowl thus illustrates an example of how adaptive introgression may be a frequently overlooked mechanism that species use to enhance their evolvability and adapt to a changing environment while still maintaining species integrity.

Conclusions

Introgression through hybridization is an important mechanism for genetic adaptation in plants and animals. In vertebrates, adaptation to hypoxic high-altitude conditions involves coordination of multiple molecular and cellular mechanisms, including selection on the HIF pathway and the blood- O_2 transport protein Hb. Genetic variation within these genes has frequently been found to be a target of selection, in multiple organisms, and recent evidence indicates adaptive introgression played a role in Tibetan people, which exhibit high heritability for both $\% \text{O}_2$ saturation and Hb concentration. In this study, striking DNA sequence similarity and strong LD reflecting IBD in two Andean duck species is present across the entire β -globin cluster including both embryonic (*HBE*) and adult (*HBB*) paralogs. However, this was not the case for *EPAS1* and or *EGLN1*. In the case of the β -globin, these similarities likely resulted from historical interbreeding between high-altitude populations of two different distantly-related species, and possibly at more than one time in history under similar selective influences coinciding with past founder events. We furthermore examined the direction of introgression and discovered that in the β globin, the speckled teal, the species

with a deeper mtDNA divergence that first colonized high altitude earlier in history transferred adaptive variation to the yellow-billed pintail, the species with a shallower divergence that likely colonized high altitude more recently (not the reverse). Thus, the yellow-billed pintail received these variants through hybridization and may not have waited for de novo mutations to adapt to the high-altitude environment, but rather acquired beneficial alleles from the standing variation, albeit via another species with a similar genetic background, leading to faster evolution.

Data availability

Parsed Illumina reads are available through NCBI SRA (PRJEB11624, PRJNA508951). NCBI nucleotide accessions for the mtDNA are FJ618397-FJ618512, JN223305-JN223375, and MG520106-MG520175. NCBI nucleotide accessions for the 3' *HBB* and intronic loci are FJ617703–FJ618396, GQ269874–GQ270616, and GQ272063–GQ272202. VCF files are available in Dryad (dryad.kd08516, dryad.bnzs7h4b4). Scripts associated with analysis are available on GitHub (<https://github.com/amgraham07>).

Acknowledgements We thank the many people including Pablo Tubaro and Cecilia Kopuchian and provincial and federal governments in Argentina, Bolivia, and Peru who assisted us with fieldwork for many years. This work was supported in part by the high-performance computing and data storage resources operated by both the Research Computing Systems Group (formerly Life Science Informatics/Arctic Regional Supercomputing) at the University of Alaska Fairbanks, Geophysical Institute, and the core facilities at Center for Genome Research and Biocomputing at Oregon State University. We thank Jay Storz, Roy Weber, Chandru Natarajan, Joana Projecto-Garcia, Angela Fago, and Hideaki Moriyama for many helpful discussions that led to a much stronger paper and for their pioneering work and insight into avian Hb function. Jay Storz and three anonymous reviewers provided helpful comments on the manuscript.

Funding Funding was provided to KGM by Alaska EPSCoR (NSF EPS-0092040, EPS-0346770), the National Science Foundation (DEB-0444748 and IOS-0949439), Frank M. Chapman Fund at the American Museum of Natural History, and the James A. Kushlan Endowment for Waterbird Biology and Conservation from the University of Miami. AMG was supported by a University of Miami Graduate School Maytag Fellowship, and by a National Science Foundation Postdoctoral Research Fellowship in Biology (DBI-1812103).

Author contributions AMG, JLP, and KGM designed the study; VMF, AJG, and KGM provided funding; AMG, KGM, REW, and THV performed the field research and/or generated the sequence data; AMG, JLP, DAD, KW, and KGM analyzed the data and wrote the paper. All authors commented on the paper.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Ethical approval Specimen collections were carried out under protocols approved by the University of Alaska-Fairbanks Institutional Animal Care & Use Committee (IACUC).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Ai H, Yang B, Li J, Xie X, Chen H, Ren J (2014) Population history and genomic signatures for high-altitude adaptation in Tibetan pigs. *BMC Genom* 15(1):834
- Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16(1):37–48
- Barrett RD, Schluter D (2008) Adaptation from standing genetic variation. *Trends Ecol Evol* 23(1):38–44
- Beall CM (2001) Adaptations to altitude: a current assessment. *Annu Rev Anthropol* 30(1):423–456
- Beall CM, Cavalleri GL, Deng L, Elston RC, Gao Y, Knight J et al. (2010) Natural selection on *EPAS1* (*HIF2 α*) associated with low hemoglobin concentration in Tibetan highlanders. *PNAS* 107(25):11459–11464
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodol)* 57(1):289–300
- Carey C, Leon-Velarde F, Dunin-Borkowski O, Monge C (1989) Shell conductance, daily water loss, and water content of Puna teal eggs. *Physiol Zool* 62(1):83–95
- Carey C, Dunin-Borkowski O, León-Velarde F, Espinoza D, Monge C (1994) Gas exchange and blood gases of Puna teal (*Anas versicolor puna*) embryos in the Peruvian Andes. *J Comp Physiol B* 163(8):649–656
- Cho YS, Hu L, Hou H, Lee H, Xu J, Kwon S et al. (2013) The tiger genome and comparative analysis with lion and snow leopard genomes. *Nat Commun* 4:2433
- Clarkson CS, Weetman D, Essandoh J, Yawson AE, Maslen G, Manske M et al. (2014) Adaptive introgression between *Anopheles* sibling species eliminates a major genomic island but not reproductive isolation. *Nat Commun* 5:4248
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA et al. (2011) The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158
- Durand EY, Patterson N, Reich D, Slatkin M (2011) Testing for ancient admixture between closely related populations. *Mol Biol Evol* 28(8):2239–2252
- Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends Genet* 28(7):342–350
- Galen SC, Natarajan C, Moriyama H, Weber RE, Fago A, Benham PM, Chavez AN, Cheviron ZA, Storz JF, Witt CC (2015) Contribution of a mutational hot spot to hemoglobin adaptation in high-altitude Andean house wrens. *Proceedings of the National Academy of Sciences* 112(45):13958–13963
- Gassmann M, Muckenthaler MU (2015) Adaptation of iron requirement to hypoxic conditions at high altitude. *J Appl Physiol* 119(12):1432–1440
- Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W et al. (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 27(2):182–189
- Gordon A, Hannon G (2010). Fastx-toolkit. Computer program distributed by the author, website http://hannonlabshedu/fastx_toolkit/index.html. Accessed 2014–2015
- Goulet BE, Roda F, Hopkins R (2017) Hybridization in plants: old ideas, new techniques. *Plant Physiol* 173(1):65–78

- Graham AM, McCracken KG (2019) Convergent evolution on the hypoxia-inducible factor (HIF) pathway genes *EGLN1* and *EPAS1* in high-altitude ducks. *Heredity* 122(6):819–832
- Graham AM, Lavretsky P, Muñoz-Fuentes V, Green AJ, Wilson RE, McCracken KG (2018) Migration-selection balance drives genetic differentiation in genes associated with high-altitude function in the speckled teal (*Anas flavirostris*) in the Andes. *Genome Biol Evol* 10(1):14–32
- Grant PR, Grant BR (2016) Introgressive hybridization and natural selection in Darwin's finches. *Biol J Linn Soc* 117(4):812–822
- Hanaoka M, Droma Y, Basnyat B, Ito M, Kobayashi N, Katsuyama Y et al. (2012) Genetic variants in *EPAS1* contribute to adaptation to high-altitude hypoxia in Sherpas. *PLoS ONE* 7(12):e50566
- Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol Ecol* 22(18):4606–4618
- Hermisson J, Pennings PS (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169(4):2335–2352
- Hey J (2011) Documentation for IMA2. Rutgers Uni, New Brunswick, NJ
- Hoffmann FG, Opazo JC, Storz JF (2012) Whole-genome duplications spurred the functional diversification of the globin gene superfamily in vertebrates. *Mol Biol Evol* 29(1):303–312
- Hubbs CL (1955) Hybridization between fish species in nature. *Syst Zool* 4(1):1–20
- Huerta-Sánchez E, Jin X, Bianba Z, Peter BM, Vinckenbosch N, Liang Y et al. (2014) Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512(7513):194–197
- Jagoda E, Lawson DJ, Wall JD, Lambert D, Muller C, Westaway M et al. (2017) Disentangling immediate adaptive introgression from selection on standing introgressed variation in humans. *Mol Biol Evol* 35(3):623–630
- Jeong C, Alkorta-Aranburu G, Basnyat B, Neupane M, Witonsky DB, Pritchard JK et al. (2014) Admixture facilitates genetic adaptations to high altitude in Tibet. *Nat Commun* 5:3281
- Johngard PA (1960) Hybridization in the Anatidae and its taxonomic implications. *Condor* 62(1):25–33
- Johnson KP, Sorenson MD (1999) Phylogeny and biogeography of dabbling ducks (genus: *Anas*): a comparison of molecular and morphological evidence. *Auk* 116(3):792–805
- Jones MR, Mills LS, Alves PC, Callahan CM, Alves JM, Lafferty DJ et al. (2018) Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. *Science* 360(6395):1355–1358
- Katoh K, Misawa K, Kuma KI, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30(14):3059–3066
- Kear J, Hulme M (2005) Ducks, geese and swans: species accounts (*Cairina to Mergus*), Vol 2. Oxford University Press, Oxford, UK
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Meth* 9(4):357–359
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14):1754–1760
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079
- Li Y, Wu D-D, Boyko AR, Wang G-D, Wu S-F, Irwin DM et al. (2014) Population variation revealed high-altitude adaptation of Tibetan mastiffs. *Mol Biol Evol* 31(5):1200–1205
- Liu KJ, Steinberg E, Yozzo A, Song Y, Kohn MH, Nakhleh L (2015) Interspecific introgressive origin of genomic diversity in the house mouse. *PNAS* 112(1):196–201
- Lorenzo FR, Huff C, Myllymäki M, Olenchock B, Swierczek S, Tashi T et al. (2014) A genetic mechanism for Tibetan high-altitude adaptation. *Nat Genet* 46(9):951–956
- Malinsky M, Matschiner M, Svoldal H (2021) Dsuite-fast D-statistics and related admixture evidence from VCF files. *Mol Ecol Resour* 21(2):584–595
- Martin SH, Davey JW, Jiggins CD (2015) Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Mol Biol Evol* 32(1):244–257
- McCracken KG, Wilson RE (2011) Gene flow and hybridization between numerically imbalanced populations of two duck species in the Falkland Islands. *PLoS ONE* 6(8):e23173
- McCracken KG, Wilson RE, Martin AR (2013) Gene flow and hybridization between numerically imbalanced populations of two duck species on the subantarctic island of South Georgia. *PLoS ONE* 8(12):e82664
- McCracken KG, Barger CP, Bulgarella M, Johnson KP, Sonsthagen SA, Trucco J et al. (2009b) Parallel evolution in the major haemoglobin genes of eight species of Andean waterfowl. *Mol Ecol* 18(19):3992–4005
- McCracken KG, Barger CP, Bulgarella M, Johnson KP, Kuhner MK, Moore AV et al. (2009a) Signatures of high-altitude adaptation in the major hemoglobin of five species of Andean dabbling ducks. *Am Nat* 174(5):631–650
- McCracken KG, Bulgarella M, Johnson KP, Kuhner MK, Trucco J, Valqui TH et al. (2009c) 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. *Mol Biol Evol* 26(4):815–827
- Miao B, Wang Z, Li Y (2017) Genomic analysis reveals hypoxia adaptation in the Tibetan mastiff by introgression of the gray wolf from the Tibetan Plateau. *Mol Biol Evol* 34(3):734–743
- Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Mol Ecol* 13(6):1341–1356
- Natarajan C, Projecto-Garcia J, Moriyama H, Weber RE, Muñoz-Fuentes V, Green AJ et al. (2015) Convergent evolution of hemoglobin function in high-altitude Andean waterfowl involves limited parallelism at the molecular sequence level. *PLoS Genet* 11(12):e1005681
- Natarajan C, Hoffmann FG, Weber RE, Fago A, Witt CC, Storz JF (2016) Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science* 354(6310):336–339
- Norris LC, Main BJ, Lee Y, Collier TC, Fofana A, Cornel AJ et al. (2015) Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. *PNAS* 112(3):815–820
- Ottenburghs J, Megens H-J, Kraus RH, Van Hooft P, van Wieren SE, Crooijmans RP et al. (2017) A history of hybrids? Genomic patterns of introgression in the true geese. *BMC Evol Biol* 17(1):1–14
- Oziolor EM, Reid NM, Yair S, Lee KM, VerPloeg SG, Bruns PC et al. (2019) Adaptive introgression enables evolutionary rescue from extreme environmental pollution. *Science* 364(6439):455–457
- Pamenter ME, Hall JE, Tanabe Y, Simonson TS (2020) Cross-species insights into genomic adaptations to hypoxia. *Front Genet* 11:743
- Pardo-Diaz C, Salazar C, Baxter SW, Merot C, Figueiredo-Ready W, Joron M et al. (2012) Adaptive introgression across species boundaries in *Heliconius* butterflies. *PLoS Genet* 8(6):e1002752
- Peng Y, Cui C, He Y, Zhang H, Yang D, Zhang Q et al. (2017) Down-regulation of *EPAS1* transcription and genetic adaptation of Tibetans to high-altitude hypoxia. *Mol Biol Evol* 34(4):818–830
- Perutz MF (1983) Species adaptation in a protein molecule. *Mol Biol Evol* 1(1):1–28
- Peters JL, Gretes W, Omland KE (2005) Late Pleistocene divergence between eastern and western populations of wood ducks (*Aix sponsa*) inferred by the ‘isolation with migration’ coalescent method. *Mol Ecol* 14(11):3407–3418

- Peters JL, Roberts TE, Winker K, McCracken KG (2012) Heterogeneity in genetic diversity among non-coding loci fails to fit neutral coalescent models of population history. *PLoS ONE* 7(2):e31972
- Projecto-Garcia J, Natarajan C, Moriyama H, Weber RE, Fago A, Cheviron ZA et al. (2013) Repeated elevational transitions in hemoglobin function during the evolution of Andean hummingbirds. *PNAS* 110(51):20669–20674
- Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z et al. (2012) The yak genome and adaptation to life at high altitude. *Nat Genet* 44(8):946–949
- Qu Y, Zhao H, Han N, Zhou G, Song G, Gao B et al. (2013) Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan Plateau. *Nat Commun* 4(1):1–9
- Racimo F, Sankararaman S, Nielsen R, Huerta-Sánchez E (2015) Evidence for archaic adaptive introgression in humans. *Nat Rev Genet* 16(6):359–371
- Randler C (2002) Avian hybridization, mixed pairing and female choice. *Anim Behav* 63(1):103–119
- Randler C (2008) Hybrid wildfowl in Central Europe—an overview. *Waterbirds* 31(1):143–146
- Rieseberg L (2011) Adaptive introgression: the seeds of resistance. *Curr Biol* 21(15):R581–R583
- Rieseberg LH, Burke J (2001) A genic view of species integration. *J Evol Biol* 14(6):883–886
- Scheinfeldt LB, Tishkoff SA (2010) Living the high life: high-altitude adaptation. *Genome Biol* 11(9):133
- Signore AV, Yang Y-Z, Yang Q-Y, Qin G, Moriyama H, Ge R-L et al. (2019) Adaptive changes in hemoglobin function in high-altitude Tibetan canids were derived via gene conversion and introgression. *Mol Biol Evol* 36(10):2227–2237
- Simonson TS, McClain DA, Jorde LB, Prchal JT (2012) Genetic determinants of Tibetan high-altitude adaptation. *Hum Genet* 131(4):527–533
- Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* 236(4803):787–792
- Song Y, Endepols S, Klemann N, Richter D, Matuschka F-R, Shih C-H et al. (2011) Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Curr Biol* 21(15):1296–1301
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68(4):978–989
- Storz JF (2016b) Hemoglobin–oxygen affinity in high-altitude vertebrates: is there evidence for an adaptive trend? *J Exp Biol* 219(20):3190–3203
- Storz JF (2016a) Causes of molecular convergence and parallelism in protein evolution. *Nat Rev Genet* 17(4):239–250
- Storz JF, Cheviron ZA (2016) Functional genomic insights into regulatory mechanisms of high-altitude adaptation. *Adv Exp Med Biol* 903:113–128
- Storz JF (2019) Hemoglobin: insights into protein structure, function, and evolution. Oxford University Press, Oxford, UK
- Suarez-Gonzalez A, Lexer C, Cronk QC (2018) Adaptive introgression: a plant perspective. *Bio Lett* 14(3):20170688
- Tate KB, Ivy CM, Velotta JP, Storz JF, McClelland GB, Cheviron ZA, Scott GR (2017) Circulatory mechanisms underlying adaptive increases in thermogenic capacity in high-altitude deer mice. *The Journal of Experimental Biology* 220(20):3616–3620
- Tubaro PL, Lijtmaer DA (2002) Hybridization patterns and the evolution of reproductive isolation in ducks. *Biol J Linn Soc* 77(2):193–200
- Wang G-D, Fan R-X, Zhai W, Liu F, Wang L, Zhong L et al. (2014) Genetic convergence in the adaptation of dogs and humans to the high-altitude environment of the Tibetan Plateau. *Genome Biol Evol* 6(8):2122–2128
- 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. *J Appl Physiol* 75(6):2646–2655
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38(6):1358–1370
- Whitney KD, Broman KW, Kane NC, Hovick SM, Randell RA, Rieseberg LH (2015) Quantitative trait locus mapping identifies candidate alleles involved in adaptive introgression and range expansion in a wild sunflower. *Mol Ecol* 24(9):2194–2211
- Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics* 23(14):1851–1853
- Wu CI (2001) The genic view of the process of speciation. *J Evol Biol* 14(6):851–865
- Wu D-D, Ding X-D, Wang S, Wójcik JM, Zhang Y, Tokarska M et al. (2018) Pervasive introgression facilitated domestication and adaptation in the *Bos species* complex. *Nat Ecol Evol* 2(7):1139–1145
- Xiang K, Peng Y, Yang Z, Zhang X, Cui C, Zhang H et al. (2013) Identification of a Tibetan-specific mutation in the hypoxic gene *EGLN1* and its contribution to high-altitude adaptation. *Mol Biol Evol* 30(8):1889–1898
- Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, Pool JE et al. (2010) Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329(5987):75–78
- Zhang C, Dong S-S, Xu J-Y, He W-M, Yang T-L (2019) PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35(10):1786–1788