

Original Research Article

Andean and Tibetan Patterns of Adaptation to High Altitude

ABIGAIL W. BIGHAM,^{1*} MEGAN J. WILSON,^{2,3} COLLEEN G. JULIAN,^{2,3} MELISA KIYAMU,⁴ ENRIQUE VARGAS,⁵ FABIOLA LEON-VELARDE,⁶ MARIA RIVERA-CHIRA,⁶ CARMELO RODRIQUEZ,⁷ VAUGHN A. BROWNE,³ ESTEBAN PARRA,⁸ TOM D. BRUTSAERT,⁹ LORNA G. MOORE,^{2,10} AND MARK D. SHRIVER¹¹

¹Department of Anthropology, University of Michigan, Ann Arbor, Michigan

²Altitude Research Center, University of Colorado Denver, Aurora, Colorado

³Department of Health and Behavioral Science, University of Colorado Denver, Denver, Colorado

⁴Department of Anthropology, The University at Albany, SUNY, Albany, New York

⁵Instituto Boliviano de Biología de Altura, La Paz, Bolivia

⁶Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Peru

⁷Clinica Sirani, Santa Cruz, Bolivia

⁸Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada

⁹Department of Exercise Science, Syracuse University, Syracuse, New York

¹⁰Department of Obstetrics and Gynecology, University of Colorado Denver, Aurora, Colorado

¹¹Department of Anthropology, Pennsylvania State University, State College, Pennsylvania

Objectives: High-altitude hypoxia, or decreased oxygen levels caused by low barometric pressure, challenges the ability of humans to live and reproduce. Despite these challenges, human populations have lived on the Andean Altiplano and the Tibetan Plateau for millennia and exhibit unique circulatory, respiratory, and hematological adaptations to life at high altitude. We and others have identified natural selection candidate genes and gene regions for these adaptations using dense genome scan data. One gene previously known to be important in cellular oxygen sensing, egl nine homolog 1 (*EGLN1*), shows evidence of positive selection in both Tibetans and Andeans. Interestingly, the pattern of variation for this gene differs between the two populations. Continued research among Tibetan populations has identified statistical associations between hemoglobin concentration and single nucleotide polymorphism (SNP) genotype at *EGLN1* and a second gene, endothelial PAS domain protein 1 (*EPAS1*).

Methods: To measure for the effects of *EGLN1* and *EPAS1* altitude genotypes on hemoglobin concentration among Andean highlanders, we performed a multiple linear regression analysis of 10 candidate SNPs in or near these two genes.

Results: Our analysis did not identify significant associations between *EPAS1* or *EGLN1* SNP genotypes and hemoglobin concentration in Andeans.

Conclusions: These results contribute to our understanding of the unique set of adaptations developed in different highland groups to the hypoxia of high altitude. Overall, the results provide key insights into the patterns of genetic adaptation to high altitude in Andean and Tibetan populations. *Am. J. Hum. Biol.* 25:190–197, 2013. © 2013 Wiley Periodicals, Inc.

High-altitude hypoxia is caused by a decrease in available oxygen levels brought on by lowered barometric pressure at high elevations. It presents severe physiological challenges to the human body and may result in several forms of altitude illness, such as acute or chronic mountain sickness. However, long-term resident populations of high altitude (defined as regions lying above 2,500 m, as this is the altitude where most people display a fall in arterial O₂ saturation) present a unique suite of physiological adaptations to this environmental pressure, which allows them to survive and reproduce in this extreme niche (Baker, 1976; Moore, 2001; Schull, 1990). In the field of biological anthropology, high-altitude adaptation has been a classic area of research. Using a natural, experimental design presented by having largely independent human populations present on three continents, biological anthropologists, physiologists, and others have studied the three geographic regions where humans have adapted to life on high: the Qinghai-Tibetan Plateau, the Andean Altiplano, and the Semien Plateau of Ethiopia. Beginning with the work of Francois-Gilbert Vault in the 1890s, early studies on hypoxia focused on the Andean pattern of physiological adaptation (Cueto, 1986). It was not until the 1970s that research began on understanding the physiological adaptations present in populations of the Qinghai-Tibetan Plateau (see Moore and Zamudio, 1998 for review). Even more recently, attention has turned to

understand the Ethiopian pattern of adaptation (Beall et al., 2002). Each of these three highland populations exhibit unique circulatory, respiratory, and hematological adaptations to life at high altitude. Moreover, they display physiological differences with respect to one another. Overall, these efforts have led to a body of the literature documenting the suite of human physiological responses to high-altitude habitation. This article focuses on patterns of adaptation in two highland groups; namely those from the Andes and the Tibetan plateau.

Andeans and Tibetans exhibit distinct physiological adaptations to high altitude. Tibetan populations have

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: National Institutes of Health; contract grant numbers: HLBI-079647 (L.G.M.), TW001188 (L.G.M.), and HG-002154 (M.D.S.); Contract grant sponsor: National Science Foundation; contract grant numbers: BNS-8919645 (L.G.M.), 0622337 (A.W.B. and M.D.S.), and BCS-0129377 (T.D.B.); Contract grant sponsor: Wenner-Gren Foundation; contract grant number: 7538 (A.W.B.); Contract grant sponsor: National Science Foundation Graduate Research Fellowships to A.W.B. and M.J.W., and an American Heart Association Predoctoral Fellowship to C.G.J.

*Correspondence to: Abigail W. Bigham. E-mail: awbigham@umich.edu

Received 12 November 2012; Accepted 19 November 2012

DOI: 10.1002/ajhb.22358

Published online 24 January 2013 in Wiley Online Library (wileyonlinelibrary.com).

relatively low-hemoglobin concentrations (Adams and Strang, 1975; Beall and Goldstein, 1987; Beall and Reichsman, 1984). This is in contrast to the “classic” Andean physiological adaptation (also seen in high-altitude sojourners), where hemoglobin concentrations are elevated compared to low-altitude groups (Beall et al., 1990, 1998). In fact, Tibetans exhibit as much as a 3.5 g/dl difference when compared with Andeans (Beall et al., 1998). Other differences between these two populations’ physiological responses to altitude exist as well. Tibetans have higher resting ventilation compared to other Asian and European populations measured at the same altitude (Beall et al., 1997; Ge et al., 1994; Zhuang et al., 1993). This increase in ventilation is not observed among Andeans, who exhibit a blunted (low) hypoxic ventilatory response (Beall et al., 1997; Chiodi, 1957). In fact, Tibetan resting ventilation is 1.5 times higher than that observed among the Aymara (Beall et al., 1997). In contrast, Andeans display a high resting arterial oxygen saturation compared to Tibetans (Beall, 2007). Despite the respiratory and hematological differences observed between Andeans and Tibetans, both groups show phenotypic adaptations with respect to pregnancy phenotypes. In particular, Andean and Tibetan women have a greater rise in uteroplacental blood flow and give birth to infants who are relatively protected from altitude-associated fetal growth restriction when compared to infants born at high altitude to European and Han (“Chinese”) women, respectively, who reside at the same altitude (for review see Moore et al., 2011).

Although the physiological differences between low- and high-altitude populations have been well documented, their genetic bases are only beginning to be unraveled. Some work has explored the heritability of specific altitude phenotypes. For example, arterial oxygen saturation and hemoglobin concentration show significant heritability in Tibetans (Beall et al., 1994, 1998). Further, among Tibetan women, a major autosomal dominant allele for high oxygen saturation has been identified wherein women carrying the high oxygen saturation allele exhibit a greater offspring survival rate than women possessing the low oxygen saturation allele (Beall et al., 1994). In addition to heritability studies, other studies have looked for associations between altitude phenotypes and Native American ancestry proportion (NAAP) measured using ancestry informative markers (AIMs). AIMs are genetic markers that partition an individual’s genetic ancestry into major continental groups, such as East Asian, Native American, European, and West African. Genotyping a panel of AIMs at the individual level allows for the genetic ancestry of an individual to be estimated. Among studies at high altitude, a trihybrid model of ancestry is often assumed, and AIMs that distinguish between Native American, European, and West African ancestry are genotyped. Research has shown that Andeans’ blunted ventilatory response is associated with Quechua ancestry (Brutsaert et al., 2005). Other research has shown an association between NAAP (as a proxy for Andean ancestry) and birth weight as well as uterine artery diameter (Julian et al., 2009; Wilson et al., 2007). Although research of this nature documents the potential for natural selection to act on phenotypic traits, it does not identify the gene(s) controlling the phenotype. Therefore, the genetic factors undergirding these populations’ phenotypic responses to high-altitude habitation may

have involved different genes, functionally different changes in the same genes, or similar changes in the same genes.

Genomic scans for high-altitude adaptation

In previous work, we used population genomic techniques to identify gene regions that showed strong signatures of natural selection among Andeans and Tibetans using dense genome scan data (Bigham et al., 2010, 2009). We genotyped high- and low-altitude populations using the Affymetrix (Santa Clara, CA) 6.0 genotyping array. The high-altitude populations included 49 Tibetans and 49 Andeans (Quechua and Aymara). The low-altitude comparison populations consisted of 49 Mesoamericans (Maya, Nahua, Mixtec, and Tlapanec), and two populations from the Haplotype Mapping project (HapMap), 60 Centre d’Étude Polymorphisme Humain Europeans and 90 East Asians comprised of 45 Han Chinese from Beijing and the 45 Japanese from Tokyo. We applied four statistical tests that detect departures from neutrality to identify candidate loci demonstrating evidence of natural selection in high-altitude Andeans and Tibetans. These tests included the locus-specific branch length (LSBL), the log of the ratios of heterozygosity ($\ln RH$), Tajima’s D , and a whole genome long-range haplotype (WGLRH) test (Shriver et al., 2004; Storz et al., 2004; Tajima, 1989; Zhang et al., 2006). By comparing Andeans and Tibetans to groups with varying degrees of shared genetic ancestry living at low altitude, we generated a list of natural selection candidate genes for each test statistic and for each high-altitude population. Statistical significance was assessed using the empirical distribution generated by the data for LSBL, Tajima’s D , and $\ln RH$, with genes falling in the top 5% of the distribution considered to be statistically significant, and the gamma distribution with a false discovery rate correction for multiple tests for the WGLRH test. The significant results were then compared across the test statistics to generate a list of selection-nominated candidate genes separately for Andeans and Tibetans.

We focused on genes that were part of the hypoxia-inducible transcription factor (HIF) pathway. This pathway, important in embryogenesis, development, and oxygen homeostasis, regulates many of the physiological responses to high-altitude habitation. Therefore, genes involved in this pathway are a priori candidates for natural selection to hypoxia (Moore et al., 2004). We found evidence of positive selection in HIF genes for both Andeans and Tibetans with most of the specific HIF-targeted or -regulatory genes differing between the two highland populations. However, a single HIF-regulatory candidate gene, egl nine homolog 1 (*EGLN1* or *PHD2*), was statistically significant in both Andeans and Tibetans (Fig. 1A, B) with both populations exhibiting extended regions of significant statistics within this gene as well as 500 kb upstream and downstream (Fig. 2A, B). For the Tibetan population, many of the single nucleotide polymorphism (SNP) comparisons showed high LSBLs coupled with extended regions of significant $\ln RH$ and standardized difference of D values. Andeans also exhibited a concentration of high LSBL values, and extended regions of negative $\ln RH$ and standardized difference of D , but the pattern differed from that observed in the Tibetans. Each population displayed a distinctive haplotype structure; both Andeans and Tibetans showed a single dominant

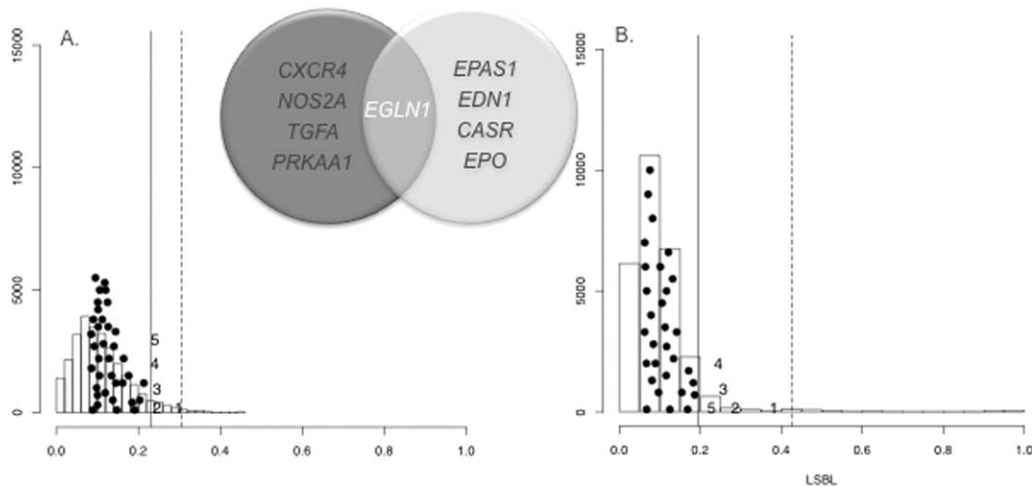


Fig. 1. *EGLN1* shows evidence of recent positive selection in Andeans and Tibetans. The plots depict the most significant SNP for each empirically significant ($p_E < 0.05$) HIF pathway candidate gene for the LSBL statistic in (A) Andeans and (B) Tibetans. Each significant SNP is plotted against the null distribution generated by plotting the highest ranked (i.e., most significant) LSBL test statistic for each gene from the dataset. The dashed line indicates $p_E < 0.01$ and the solid line indicates $p_E < 0.05$. SNPs falling in the top 5% or top 1% of the distribution are plotted as numbers, where 1 = *EGLN1*, 2 = *NOS2*, 3 = *TGFA*, 4 = *CXCR4*, and 5 = *PRKAA1* in Andeans and 1 = *EPAS1*, 2 = *EPO*, 3 = *CASR*, 4 = *EDN1*, and 5 = *EGLN1* in Tibetans. HIF pathway SNPs falling outside of the zone of statistical significance are depicted as solid circles.

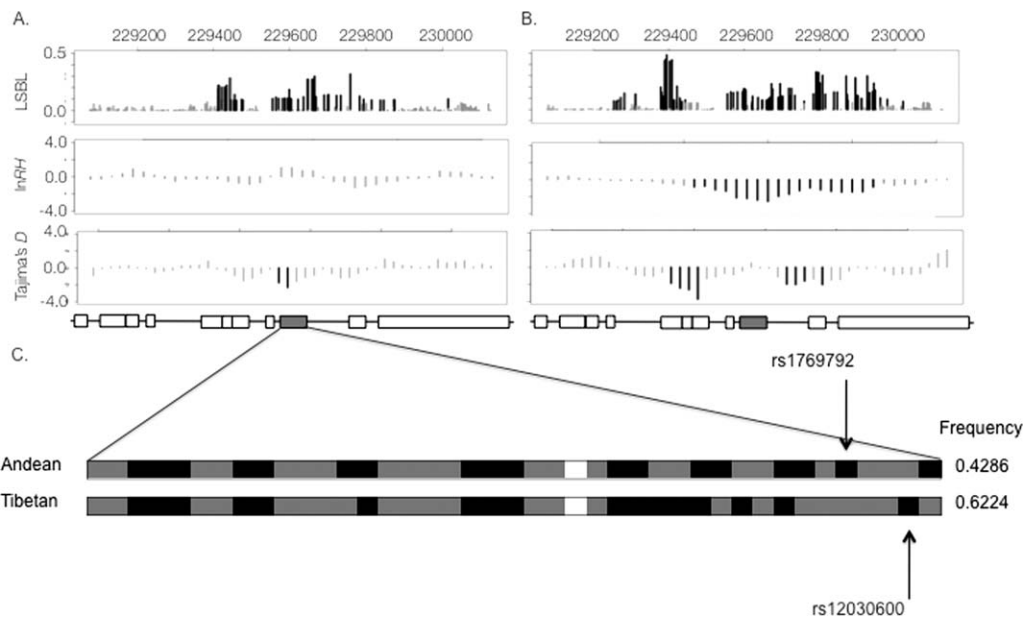


Fig. 2. The pattern of variation for *EGLN1* in Andeans and Tibetans. LSBL, $\ln RH$, and Tajima's D are plotted for *EGLN1* including 500 kb upstream and downstream of the gene in Andeans (A) and Tibetans (B). Significant test statistics ($p_E < 0.05$) are shown in black and nonsignificant test statistics are shown in gray. A gene map of the region is depicted below the plots, with *EGLN1* depicted as a gray box and nearby genes shown as unfilled boxes. The most frequent haplotype in the 100kb region surrounding *EGLN1* is depicted for Andeans and Tibetans (C), with its frequency in each population indicated to the right of each haplotype. Derived alleles are shown in gray and ancestral alleles are shown in black. The ancestral/derived state is unknown for a single SNP shown in white. The two SNPs exhibiting the highest ranked LSBL are rs1769792 in Andeans and rs12030600 in Tibetans. The location of each is indicated in the figure by the arrows. Both of these SNPs are located outside of the *EGLN1* coding regions 3' to the gene.

haplotype, but the haplotype was unique (Fig. 2C). This result was especially interesting in light of parallel human adaptations and suggested that Andeans and Tibetans may have adapted to high-altitude hypoxia via distinct modifications in the same gene, *EGLN1*.

In addition to studying the HIF-pathway candidate genes, we also scanned across each chromosome to identify genes or gene regions with previously unknown function with respect to altitude adaptation but that showed evidence of positive directional selection. To reduce the

false positive rate, we identified megabase regions that exhibited continuously significant LSBL and $\ln RH$ statistics or $\ln RH$ and Tajima's D statistics for Andeans and Tibetans (for review see Bigham et al., 2010). The hypergeometric distribution, calculated for one megabase non-overlapping windows along each chromosome, was used to identify extended regions of statistical significance for each test statistic. The P -value for each window was corrected for multiple tests using the Bonferroni correction and significant P -values ($P < 0.004$) were defined such that one false positive would be expected for all observed windows. The results of this analysis revealed 14 candidate regions for high-altitude adaptation in Tibetans and 37 regions in Andeans. All of these gene regions were unique to each of the two highland populations. Among Andeans, the largest of the regions was a four megabase region located on chromosome 12, ranging from 109,000,000 to 113,000,000, that contained 47 genes, some of which were involved in cellular housekeeping and immunity. Among Tibetans, a single region located on chromosome 2, ranging from 46,000,000 to 47,000,000, was identified. This genomic region contains a HIF pathway candidate gene, endothelial PAS domain protein 1 (*EPAS1* also known as [aka] *HIF2A*).

Several other studies have performed genomic scans for positive selection among Tibetan highlanders (Beall et al., 2010; Peng et al., 2011; Simonson et al., 2010; Wang et al., 2011; Xu et al., 2011; Yi et al., 2010). Importantly, all of these studies have shown evidence of natural selection for the HIF pathway genes. One of the same gene regions that we identified, *EPAS1*, shows evidence of positive directional selection in all of these genome-wide analyses (Beall et al., 2010; Bigham et al., 2010; Peng et al., 2011; Simonson et al., 2010; Wang et al., 2011; Xu et al., 2011; Yi et al., 2010) and the majority of these analyses have shown evidence of natural selection for *EGLN1* (Peng et al., 2011; Simonson et al., 2010; Wang et al., 2011; Xu et al., 2011; Yi et al., 2010). This convergence and replication of results using distinct cohorts strongly supports the roles of both of these genes, *EPAS1* and *EGLN1*, in the Tibetan pattern of high-altitude adaptation.

Although genome scans are powerful tools to identify and rank candidate genes, they have the potential to identify false positives (Thornton and Jensen, 2007). Moreover, without determining if an association exists between a purported candidate gene and a relevant phenotype, evolutionary explanations are merely suggestive. Therefore, it is imperative to follow-up with independent lines of evidence that further support the evidence for positive selection at a particular gene. One method for doing so is to conduct association studies that link genotype to phenotype. A number of such studies have used hemoglobin as the phenotype of interest, given its role in transporting oxygen from vertebrate lungs to other tissues of the body. Thus, it has been hypothesized that hemoglobin could be a potential target of natural selection to hypoxia tolerance. Previous work has shown significant associations for *EPAS1* and *EGLN1* SNP genotypes and haplotypes with hemoglobin concentration in Tibetans (Beall et al., 2010; Simonson et al., 2010; Yi et al., 2010). However, similar associations have not been demonstrated among Andeans. This is especially important as *EGLN1* exhibits signatures of natural selection among Andeans as well as Tibetans. To explore the genetic basis of hematological adaptations to high altitude among Andeans, we set out

to test for *EPAS1* and *EGLN1* SNP associations with hemoglobin concentration in a cohort of South American highlanders.

MATERIALS AND METHODS

Study participants, phenotype collection, and DNA isolation

The Andean sample consisted of 71 Bolivian and 35 Peruvian individuals of largely Aymara and Quechua ancestry, respectively. Bolivian samples were collected in La Paz, Bolivia (3,600 m) or the neighboring city of El Alto (4,200 m). Peruvian samples consisted of multigenerational high-altitude residents of Cerro de Pasco, Peru (4,338 m). All samples included in this analysis were female.

For the Bolivian cohort, sample collection took place at the Instituto Boliviano de Biología de Altura (Bolivian High Altitude Biology Institute) in La Paz, Bolivia. At screening, blood was drawn from the antecubital vein. Hemoglobin concentration (g/dl) was measured in triplicate using the cyanmethemoglobin technique. For the Peruvian cohort, hemoglobin concentration (g/dl) was measured at high altitude in the city of Cerro de Pasco, Peru (see Brutsaert et al. 2005 for details of the study design) using a Hemocue blood hemoglobin analyzer (Angelholm, Sweden). Subjects with hemoglobin less than sex-specific cutoff values for anemia were excluded from the study. In addition, highland resident subjects with excessive polycythemia, defined as hemoglobin >21.5 gm/dl, were also excluded. Details of the procedures used to collect the phenotypic data have been previously reported (Brutsaert et al., 2005, 2003; Moore, 2003).

DNA was isolated from archived venous blood using the Puregene DNA purification system (Qiagen, Valencia, CA) according to the manufacturer's instructions. All subjects provided informed, written consent, and the study was approved by the Colorado Multiple Institutions Review Board, Tibet Institute of Medical Sciences, Colegio Médico in Bolivia, University at Albany, SUNY, the Universidad Peruana Cayetano Heredia, Penn State University, Syracuse University, and the University of Michigan.

Selection of single nucleotide polymorphisms and genotyping

The results of genome-wide screens for natural selection in Andeans and Tibetans showed strong signatures of natural selection in two genes, *EPAS1* and *EGLN1*, that are part of the HIF pathway (Bigham et al., 2010, 2009). *EGLN1* exhibited statistical evidence of past natural selection among both Andeans and Tibetans, whereas *EPAS1* showed such evidence only among Tibetans. For each gene, we selected (1) SNPs that distinguished high-altitude Andean (i.e., Quechua and Aymara) haplotypes exhibiting high frequency and long-range linkage disequilibrium (LD) compared to low-altitude Mesoamericans (i.e., Nahua, Mixtec, Tlapanec, and Maya) and (2) SNPs that distinguished high-altitude Tibetan haplotypes with high frequency and long range LD compared to low-altitude East Asians (i.e., Han Chinese from Beijing and Japanese from Tokyo). To do so, genotype data spanning 500 kb pairs upstream and downstream of the start and end coordinates of each gene were phased using Fastphase (v1.1) (Scheet and Stephens, 2006) and the resulting haplotypes were inferred for each population (Andean, Mesoamerican, and European American) using Haploview

TABLE 1. SNPs genotyped in or near EPAS1 and EGLN1

Gene symbol or chromosomal region	Gene name	Chr	SNP(s)	SNP count
<i>EGLN1</i>	egl nine homolog 1 (<i>Caenorhabditis elegans</i>)	1	rs1339896, rs2812385, rs16854779, rs2572259, rs11122273	5
<i>EPAS1</i>	Endothelial PAS domain protein 1	2	rs12619696, rs10187368, rs13420857, rs7587138, rs2881504	5

Chr, chromosome.

TABLE 2. Subject characteristics

Variable	Combined	Bolivian (3,600 m)	Peruvian (4,300 m)	<i>P</i>
Hb (g/dl)	15.2 ± 1.74 (106)	14.5 ± 1.3 (71)	16.6 ± 1.7 (35)	<0.001
Age (year)	26.5 ± 5.7 (106)	27.3 ± 6.4 (71)	24.9 ± 3.0 (35)	<0.05
Height (cm)	150.3 ± 5.3 (106)	151.1 ± 5.2 (71)	148.8 ± 5.2 (35)	<0.05
Weight (g)	56.2 ± 8.2(106)	58.1 ± 8.7 (71)	52.4 ± 5.2 (35)	<0.001
Female (%)	100 ± 0 (106)	100 ± 0 (71)	100 ± 0 (35)	NS
NAAP	86.2 ± 10.0 (106)	83.2 ± 10.8 (69)	92.1 ± 3.4 (35)	<0.001

All characteristics are reported as mean ± standard deviation or 95% confidence intervals for proportions. Samples sizes for each variable are provided in parentheses.

NS, not significant. Comparisons between groups were conducted using Student's *t*-tests for unequal or equal variances, depending on the characteristic.

(v 4.1; Barrett et al., 2005). In this way, we included SNPs located within each candidate gene as well as SNPs upstream and downstream as the latter have been shown to regulate adjacent genes (Tishkoff et al., 2007). See Table 1 and Supporting Information Table S1 for a listing of each SNP selected.

Genotyping of 78 AIMs that distinguish between West African, European and Native American parental populations was performed at Prevention Genetics (Marshfield, WI) for the Peruvian and Bolivian samples. Details, including allele frequencies in the parental populations, DNA sequences, and the genomic positions of SNPs are available from the dbSNP database (www.ncbi.nlm.nih.gov/SNP) under the submitter handle PSU-ANTH (Bonilla et al., 2004; Brutsaert et al., 2005; Shriver et al., 2003) and are provided in Supporting Information Table S2. Individual ancestry estimates were calculated for each individual in the population sample using a dihybrid model implemented in the program maximum likelihood (Hanis et al., 1986).

Association analysis

Candidate SNPs were selected in or near the genes *EPAS1* and *EGLN1* (Table 1). Candidate SNP genotyping was performed using the ABI SNplex genotyping technology (Applied Biosystems, Foster City, CA). Genemapper v3.5 (Applied Biosystems, Foster City, CA) software was used for genotype calling. In total, 10 SNPs produced reliable genotype calls and were subject to quality control (QC). SNPs were filtered from our dataset if they exhibited ≥20% missing data ($n = 1$ SNPs, rs2881504). No SNPs were filtered from our dataset because they violated Hardy–Weinberg equilibrium ($P < 0.001$). In total, nine out of 10 genotyped SNPs were tested for an association with hemoglobin. In addition to filtering SNPs from the dataset, we also excluded individuals from our analysis if they exhibited >20% missing data. We removed 17 individuals (nine Bolivians and eight Peruvians) from the original 106 study subjects based on this criterion.

Quantitative association analyses were performed on each SNP to identify SNP genotype associations with

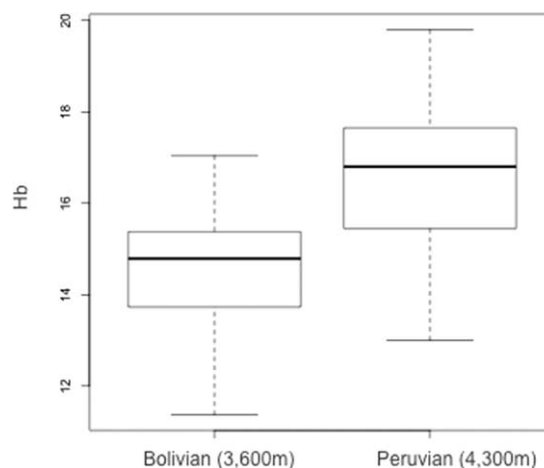


Fig. 3. Boxplot of hemoglobin levels estimated in Bolivian and Peruvian study participants. Hemoglobin concentration (g/dl) is shown along the y-axis. The populations and their altitude of residence are shown along the x-axis.

hemoglobin in the Andean population using R version 2.11.1 (<http://www.R-project.org>; Team, 2008). We performed standard linear regression on hemoglobin g/dl to calculate regression coefficients and 95% confidence intervals using dominant, recessive, and additive models of inheritance. Models were adjusted for age, altitude of residence, and NAAP to control for population substructure. We applied a Bonferroni correction for multiple testing with a *P*-value cutoff of $P < 0.006$ ($\alpha = 0.05$, nine tests).

RESULTS

Study participants

Subject characteristics including hemoglobin, height, weight, age, sex, and NAAP are presented in Table 2. Peruvian and Bolivian study participants exhibited largely Native American ancestry. The average individual NAAP for Peruvians (92.1 ± 3.4) was higher than that observed among Bolivians (83.2 ± 10.8). In addition, Peruvians had higher hemoglobin concentration than their Bolivian counterparts. As hemoglobin concentration is influenced by altitude of residence, other physiological characteristics of the individual, and method of analysis, this increase in hemoglobin concentration observed in Peruvians compared to Bolivians is likely due to several factors. Bolivians were taller, weighed more, and were older than Peruvians (all $P < 0.05$). Of the total cohort, 89 individuals passed QC including 62 high-altitude Andeans from Bolivia and 27 high-altitude Andeans from Peru.

Hemoglobin

Hemoglobin concentration differed between Bolivian (mean = 14.5 g/dl) and Peruvian (mean = 16.6 g/dl)

TABLE 3. Multiple linear regression results for SNP genotype on Hb (g/dl)

SNP	n	MA	Additive model			Recessive model			Dominant model		
			Beta	P	Bonferroni	Beta	P	Bonferroni	Beta	P	Bonferroni
rs2572259	78	G	-0.22	0.08	0.68	-1.07	0.08	0.68	0.05	0.91	8.15
rs11122273	88	G	-0.37	0.04	0.38	-1.21	0.04	0.38	-0.18	0.61	5.49
rs1339896	89	C	-0.35	0.17	1.49	-0.81	0.07	0.63	-0.19	0.60	5.36
rs16854779	86	C	0.16	0.48	4.31	0.56	0.27	2.42	0.08	0.79	7.07
rs2812385	85	T	0.00	0.64	5.79	0.22	0.64	5.79	-0.11	0.75	6.71
rs10187368	87	G	0.30	0.50	4.50	0.29	0.50	4.50	0.43	0.24	2.20
rs7587138	86	G	0.72	0.26	2.36	1.21	0.26	2.36	0.78	0.03	0.29
rs12619696	88	G	0.46	0.66	5.90	0.21	0.66	5.90	0.81	0.02	0.16
rs13420857	89	C	0.23	0.29	2.60	0.39	0.30	2.70	0.25	0.47	4.25

Beta, uncorrected *P*-values, and Bonferroni corrected *P*-values are reported for the minor allele for additive, recessive and dominant models of inheritance. MA, minor allele; *n*, number of individuals.

women ($P < 0.001$; Fig. 3). On average, Peruvian hemoglobin concentration was 2.09 g/dl higher than the Bolivian's. Such a difference was likely due to the higher altitude at which the Peruvian sample was collected in addition to the other factors mentioned above. We also examined the association between NAAP, height, weight, age, and altitude of residence and hemoglobin concentration. We found that hemoglobin concentration was associated with NAAP, height, and altitude of residence when each of these variables was tested independently, but only with altitude of residence when these three independent variables were included in the multiple linear regression model.

Candidate gene associations with hemoglobin

We tested nine SNPs from two genes (*EPAS1* and *EGLN1*) previously shown to be associated with hemoglobin concentration in Tibetans. Standard linear regression was performed for each SNP after controlling for NAAP, height, weight, age, and altitude of residence. Using additive, dominant, and recessive models of inheritance, no SNPs were significantly associated with hemoglobin concentration in our combined Bolivian and Peruvian cohort after correcting for multiple tests (Table 3).

In addition to analyzing the Bolivian and Peruvian cohorts together to increase our power of detecting an association with hemoglobin, we also interrogated the Bolivian cohort independently for such associations to determine if the slight differences in hemoglobin concentration contributed to the lack of genotype-phenotype associations. We chose not to interrogate the Peruvian cohort for SNP-hemoglobin associations as the sample size for this population lacked statistical power. Again, standard linear regression was performed for each SNP after controlling for NAAP. We did not detect any significant association with SNP genotype and hemoglobin concentration (data not shown).

DISCUSSION

Multiple linear regression analysis was performed to measure for the effects of altitude genotypes on hemoglobin concentration among Andean highlanders. Using additive, dominant, and recessive models of inheritance, we did not identify any associations between SNP genotype at two genes previously shown to be associated with hemoglobin concentration in Tibetans, *EPAS1* and *EGLN1*. These results contribute to our understanding of the unique set of adaptations developed in different high-

land groups to the hypoxia of high altitude. Physiological studies have previously shown that Tibetans exhibit lower hemoglobin concentrations than observed in other high-altitude groups but values that are still elevated compared to what is observed at sea level, but that Andeans show a greater hemopoietic response to altitude. The lower hemoglobin concentrations observed among Tibetans than Andeans might be adaptive based on the hemodynamic advantages of reduced blood viscosity. Andeans may compensate for the disadvantages of high blood viscosity (higher hemoglobin concentrations make the blood more viscous) by increased levels of SaO₂ (Beall et al., 1999), which together with hemoglobin determines the oxygen content of blood.

At the genomic level, Andeans and Tibetans show evidence of positive directional selection in the same HIF pathway candidate gene, *EGLN1*. The protein product of this gene, along with those of *EGLN2* and *EGLN3*, is a molecular oxygen sensor that regulates the HIF transcriptional pathway (Hirsila et al., 2003). In normoxia, EGLN1 hydroxylates HIF-1 α 's oxygen-dependent degradation domain, which targets this protein for breakdown by the E3 ubiquitin ligase complex (Ivan et al., 2001; Jaakkola et al., 2001; Maxwell et al., 1999). Under hypoxic stress, decreases in oxygen tension lead to a reduction in prolyl hydroxylation of HIF-1 α by EGLN1, thus increasing HIF levels (Epstein et al., 2001). This permits HIF1 to continually target downstream genes to maintain cellular oxygen homeostasis. Given EGLN1's basal role as a molecular oxygen sensor, evidence that the pattern of genetic variation is distinct between Andean and Tibetan populations, and the associations of *EGLN1* genotypes with hemoglobin concentration in Tibetans but not Andeans, it is possible that different positively selected mutations in this gene may be responsible for affecting different phenotypes in Andeans versus Tibetans. Further research elucidating the genotype-phenotype relationship between this gene and Andean phenotypes will be an important step in understanding the functional significance of *EGLN1* variation in this population.

EPAS1 is a HIF regulatory gene encoding the transcription factor HIF2A. It induces downstream genes when cellular oxygen levels decrease. In addition, HIF2A is partly responsible for erythropoiesis and therefore can affect the concentration of hemoglobin in the blood stream. As noted above, variants in this gene have been shown to associate with hemoglobin concentration in Tibetans (Beall et al.,

2010; Simonson et al., 2010; Yi et al., 2010). Given the lack of genomic signatures of natural selection in or near this gene in Andean populations, it is unsurprising that associations between hemoglobin concentration and SNP genotype were not identified among Andeans. Further research elucidating the molecular mechanisms underlying this adaptation in Tibetans will be an important step in understanding the functional significance of *EPAS1* variation on hemoglobin concentration.

It is important to emphasize that the results of this study do not suggest that the higher hemoglobin concentration observed among Andeans is not genetic nor do they suggest that genetic variation does not control this phenotype. Rather, they show that the two loci associated with variation in hemoglobin concentration among Tibetans do not associate with this phenotypic variable in Andeans. However, we did not interrogate Andeans for the same SNPs that showed associations with hemoglobin concentrations in Tibetans. It remains possible that if the same SNPs were assayed in a cohort of Andeans, the findings could change.

In future work, it will be interesting to follow-up with continued genotypic and phenotypic comparisons between Andean and Tibetan high-altitude adaptation, as these two populations are distinct in geographic locale as well as in duration of high-altitude habitation. Archaeological data indicate that Himalayan populations first inhabited the Tibetan Plateau as early as 25,000 years ago whereas populations first moved onto the Andean Altiplano 11,000 years ago (Aldenderfer, 2003; Moseley, 2001). Such data suggest that both populations have likely experienced natural selection in response to hypoxia and recent genome scan studies support the existence of naturally selected genes in both groups. Although this has been broadly acknowledged for Tibetans in recent years (Beall et al., 2010; Peng et al., 2011; Simonson et al., 2010; Wang et al., 2011; Xu et al., 2011; Yi et al., 2010), it is less often recognized that considerable evidence for natural selection in the Andean populations exists. In particular, populations with greater Andean ancestry are protected from intrauterine growth restriction at high altitude, a trait that clearly exerts a major influence over reproductive success and therefore possibly fitness. Moreover, there exists considerable evidence to suggest that such Andean genes influence the maternal vascular response to pregnancy, thus enabling a normal sea-level rise in uteroplacental blood flow (Wilson et al., 2007; Julian et al., 2009; Moore et al., 2011). However, many broad questions remain unanswered, particularly those that involve Tibetan–Andean comparisons. For example, (1) Are there genotype–phenotype associations in Andeans and, in particular, do these differ from those seen in Tibetans? (2) If so, what are the convergent and/or divergent evolutionary pathways that underlie the Tibetan–Andean contrast? and (3) What role do developmental versus genetic factors play, and are there differences in this regard between Tibetans and Andeans?

The questions posed above are answerable, but they will require an integration of data at the genetic and physiological levels, as well as at the demographic, ecological, and historical levels. Continued research looking at genotype–phenotype associations with alternative phenotypes in Andeans, such as resting SaO_2 or phenotypes related to pregnancy, will be necessary to further elucidate the pattern of adaptation observed among Andeans and to compare their pattern of genetic adaptation to that

of Tibetans. In the high-altitude native literature, genotype–phenotype integration is largely absent. Hence, moving forward with additional genotype–phenotype studies at this time is critical not only for answering evolutionary questions, but also because identifying the genes involved in high-altitude adaptive phenotypes has implications for the over 140 million people who currently reside at high altitude (Moore and Zamudio, 1998) or the even larger numbers of persons suffering from hypoxia-related disorders. By understanding how similar environmental pressures with varying evolutionary time frames can result in either the same or different genetic adaptations, we will be better situated to understand the molecular basis for convergent human adaptations. Furthermore, we hope this research will establish a basis for comparative studies looking at the process of genetic adaptation to altitude or other environments in human populations.

ACKNOWLEDGMENTS

The authors would like to thank Michael Little and Ralph Garruto for putting together the special edition on high-altitude adaptation and for providing instructive comments on earlier versions of this manuscript. We would also like to thank all of the study participants who participated in this research.

LITERATURE CITED

- Adams WH, Strang LJ. 1975. Hemoglobin levels in persons of tibetan ancestry living at high altitude. *Proc Soc Exp Biol Med* 149:1036–1039.
- Aldenderfer M. 2003. Moving up in the world. *American Scientist* 91:542–549.
- Baker P, Little MA, editors. 1976. *Man in the Andes: A Multidisciplinary Study of High-Altitude Quechua*. Stroudsburg: Dowden, Hutchinson, and Ross, INC.
- Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265.
- Beall CM. 2007. Two routes to functional adaptation: Tibetan and Andean high-altitude natives. *Proc Natl Acad Sci U S A* 104 Suppl 1:8655–8660.
- Beall CM, Almasy LA, Blangero J, Williams-Blangero S, Brittenham GM, Strohl KP, Decker MJ, Vargas E, Villena M, Soria R, Alarcon AM, Gonzales C. 1999. Percent of oxygen saturation of arterial hemoglobin among Bolivian Aymara at 3,900–4,000 m. *Am J Phys Anthropol* 108:41–51.
- Beall CM, Blangero J, Williams-Blangero S, Goldstein MC. 1994. Major gene for percent of oxygen saturation of arterial hemoglobin in Tibetan highlanders. *Am J Phys Anthropol* 95:271–276.
- Beall CM, Brittenham GM, Macuaga F, Barragan M. 1990. Variation in Hemoglobin Concentration among Samples of High-Altitude Natives in the Andes and the Himalayas. *Am J Hum Biol* 2:639–651.
- Beall CM, Brittenham GM, Strohl KP, Blangero J, Williams-Blangero S, Goldstein MC, Decker MJ, Vargas E, Villena M, Soria R, Alarcon AM, Gonzales C. 1998. Hemoglobin concentration of high-altitude Tibetans and Bolivian Aymara. *Am J Phys Anthropol* 106:385–400.
- Beall CM, Cavalleri GL, Deng L, Elston RC, Gao Y, Knight J, Li C, Li JC, Liang Y, McCormack M, Montgomery HE, Pan H, Robbins PA, Shianna KV, Tam SC, Tsering N, Veeramah KR, Wang W, Wangdui P, Weale ME, Xu Y, Xu Z, Yang L, Zaman MJ, Zeng C, Zhang L, Zhang X, Zhaxi P, Zheng YT. 2010. Natural selection on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan highlanders. *Proc Natl Acad Sci U S A* 107:11459–11464.
- Beall CM, Decker MJ, Brittenham GM, Kushner I, Gebremedhin A, Strohl KP. 2002. An Ethiopian pattern of human adaptation to high-altitude hypoxia. *Proc Natl Acad Sci U S A* 99:17215–17218.
- Beall CM, Goldstein MC. 1987. Hemoglobin concentration of pastoral nomads permanently resident at 4,850–5,450 meters in Tibet. *Am J Phys Anthropol* 73:433–438.
- Beall CM, Reichsman AB. 1984. Hemoglobin levels in a Himalayan high altitude population. *Am J Phys Anthropol* 63:301–306.
- Beall CM, Strohl KP, Blangero J, Williams-Blangero S, Almasy LA, Decker MJ, Worthman CM, Goldstein MC, Vargas E, Villena M, Soria R, Alarcon AM, Gonzales C. 1997. Ventilation and hypoxic ventilatory response

- of Tibetan and Aymara high altitude natives. *Am J Phys Anthropol* 104:427–447.
- Bigham A, Bauchet M, Pinto D, Mao X, Akey JM, Mei R, Scherer SW, Julian CG, Wilson MJ, Lopez Herraes D, Brutsaert T, Parra EJ, Moore LG, Shriver MD. 2010. Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. *PLoS Genet* 6.
- Bigham AW, Mao X, Mei R, Brutsaert T, Wilson MJ, Julian CG, Parra EJ, Akey JM, Moore LG, Shriver MD. 2009. Identifying positive selection candidate loci for high-altitude adaptation in Andean populations. *Hum Genomics* 4:79–90.
- Bonilla C, Shriver MD, Parra EJ, Jones A, Fernandez JR. 2004. Ancestral proportions and their association with skin pigmentation and bone mineral density in Puerto Rican women from New York city. *Hum Genet* 115:57–68.
- Brutsaert TD, Parra EJ, Shriver MD, Gamboa A, Leon-Velarde F. 2005. Ancestry explains the blunted ventilatory response to sustained hypoxia and lower exercise ventilation of Quechua altitude natives. *Am J Physiol-Reg I* 289:R225–R234.
- Brutsaert TD, Parra EJ, Shriver MD, Gamboa A, Palacios JA, Rivera M, Rodriguez I, Leon-Velarde F. 2003. Spanish genetic admixture is associated with larger V(O₂) max decrement from sea level to 4338 m in Peruvian Quechua. *J Appl Physiol* 95:519–528.
- Chiodi H. 1957. Respiratory adaptations to chronic high altitude hypoxia. *J Appl Physiol* 10:81–87.
- Cueto M. 1986. *Andean Biology in Peru: Scientific Styles on the Periphery*. Isis 80:640–658.
- Epstein AC, Gleadle JM, McNeill LA, Hewitson KS, O'Rourke J, Mole DR, Mukherji M, Metzén E, Wilson MI, Dhanda A, Tian YM, Masson N, Hamilton DL, Jaakkola P, Barstead R, Hodgkin J, Maxwell PH, Pugh CW, Schofield CJ, Ratcliffe PJ. 2001. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* 107:43–54.
- Ge RL, Chen QH, He LG. 1994. [Characteristics of hypoxic ventilatory response in Tibetan living at moderate and high altitudes]. *Chinese Journal of Tuberculosis and Respiratory Diseases* 17:364–366, 384.
- Hanis CL, Chakraborty R, Ferrell RE, Schull WJ. 1986. Individual Admixture Estimates - Disease Associations and Individual Risk of Diabetes and Gallbladder-Disease among Mexican-Americans in Starr County, Texas. *Am J Phys Anthropol* 70:433–441.
- Hirsila M, Koivunen P, Gunzler V, Kivirikko KI, Myllyharju J. 2003. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *The Journal of biological chemistry* 278:30772–30780.
- Ivan M, Kondo K, Yang H, Kim W, Valiando J, Ohh M, Salic A, Asara JM, Lane WS, Kaelin WG, Jr. 2001. HIF α targeted for VHL-mediated destruction by proline hydroxylation: implications for O₂ sensing. *Science* 292:464–468.
- Jaakkola P, Mole DR, Tian YM, Wilson MI, Gielbert J, Gaskell SJ, Kriegshelm A, Hebestreit HF, Mukherji M, Schofield CJ, Maxwell PH, Pugh CW, Ratcliffe PJ. 2001. Targeting of HIF- α to the von Hippel-Lindau ubiquitylation complex by O₂-regulated prolyl hydroxylation. *Science* 292:468–472.
- Julian CG, Wilson MJ, Lopez M, Yamashiro H, Tellez W, Rodriguez A, Bigham AW, Shriver MD, Rodriguez C, Vargas E, Moore LG. 2009. Augmented uterine artery blood flow and oxygen delivery protect Andeans from altitude-associated reductions in fetal growth. *Am J Physiol-Reg I* 296:R1564–R1575.
- Maxwell PH, Wiesener MS, Chang GW, Clifford SC, Vaux EC, Cockman ME, Wykoff CC, Pugh CW, Maher ER, Ratcliffe PJ. 1999. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* 399:271–275.
- Moore LG. 2001. Human genetic adaptation to high altitude. *High Alt Med Biol* 2:257–279.
- Moore LG. 2003. Fetal growth restriction and maternal oxygen transport during high altitude pregnancy. *High Alt Med Biol* 4:141–156.
- Moore LG, Charles SM, Julian CG. 2011. Humans at high altitude: hypoxia and fetal growth. *Resp Physiol Neurobi* 178:181–190.
- Moore LG, Shriver M, Bemis L, Hickler B, Wilson M, Brutsaert T, Parra E, Vargas E. 2004. Maternal adaptation to high-altitude pregnancy: an experiment of nature—a review. *Placenta* 25 Suppl A:S60–71.
- Moore LGN, Zamudio S. 1998. Human adaptation to high altitude: Regional and life cycle perspectives. *American Journal of Physical Anthropology Yearbook* 41:25–64.
- Moseley MM. 2001. *The Incas and Their Ancestors*. London: Thames and Hudson.
- Peng Y, Yang Z, Zhang H, Cui C, Qi X, Luo X, Tao X, Wu T, Ouzhuluobu, Basang, Ciwangangbu, Danzengduojie, Chen H, Shi H, Su B. 2011. Genetic variations in Tibetan populations and high-altitude adaptation at the Himalayas. *Mol Biol Evol* 28:1075–1081.
- Scheet P, Stephens M. 2006. A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet* 78:629–644.
- Schull WJaR, F, editor. 1990. *The Aymara: Strategies in Human Adaptation to a Rigorous Environment*. Boston: Kluwer Academic Publishers.
- Shriver MD, Kennedy GC, Parra EJ, Lawson HA, Sonpar V, Huang J, Akey JM, Jones KW. 2004. The genomic distribution of population substructure in four populations using 8,525 autosomal SNPs. *Hum Genomics* 1:274–286.
- Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart CJ, McKeigue PM, Kittles RA. 2003. Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum Genet* 112:387–399.
- Simonson TS, Yang YZ, Huff CD, Yun HX, Qin G, Witherspoon DJ, Bai ZZ, Lorenzo FR, Xing JC, Jorde LB, Prchal JT, Ge RL. 2010. Genetic Evidence for High-Altitude Adaptation in Tibet. *Science* 329:72–75.
- Storz JF, Payseur BA, Nachman MW. 2004. Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. *Mol Biol Evol* 21:1800–1811.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- R Development Core Team. 2008. *R: A language and environment for statistical computing*. Version 2.11.1. Vienna, Austria.
- Thornton KR, Jensen JD. 2007. Controlling the false-positive rate in multilocus genome scans for selection. *Genetics* 175:737–750.
- Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, Nyambo TB, Ghorji J, Bumpstead S, Pritchard JK, Wray GA, Deloukas P. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet* 39:31–40.
- Wang B, Zhang YB, Zhang F, Lin H, Wang X, Wan N, Ye Z, Weng H, Zhang L, Li X, Yan J, Wang P, Wu T, Cheng L, Wang J, Wang DM, Ma X, Yu J. 2011. On the origin of Tibetans and their genetic basis in adapting high-altitude environments. *PLoS One* 6:e17002.
- Wilson MJ, Lopez M, Vargas M, Julian C, Tellez W, Rodriguez A, Bigham A, Armaza JF, Niermeyer S, Shriver M, Vargas E, Moore LG. 2007. Greater uterine artery blood flow during pregnancy in multigenerational (Andean) than shorter-term (European) high-altitude residents. *Am J Physiol-Reg I* 293:R1313–R1324.
- Xu S, Li S, Yang Y, Tan J, Lou H, Jin W, Yang L, Pan X, Wang J, Shen Y, Wu B, Wang H, Jin L. 2011. A genome-wide search for signals of high-altitude adaptation in Tibetans. *Mol Biol Evol* 28:1003–1011.
- Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, Zheng H, Liu T, He W, Li K, Luo R, Nie X, Wu H, Zhao M, Cao H, Zou J, Shan Y, Li S, Yang Q, Asan, Ni P, Tian G, Xu J, Liu X, Jiang T, Wu R, Zhou G, Tang M, Qin J, Wang T, Feng S, Li G, Huasang, Luosang J, Wang W, Chen F, Wang Y, Zheng X, Li Z, Bianba Z, Yang G, Wang X, Tang S, Gao G, Chen Y, Luo Z, Gusang L, Cao Z, Zhang Q, Ouyang W, Ren X, Liang H, Huang Y, Li J, Bolund L, Kristiansen K, Li Y, Zhang Y, Zhang X, Li R, Yang H, Nielsen R, Wang J. 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329:75–78.
- Zhang C, Bailey DK, Awad T, Liu G, Xing G, Cao M, Valmeekam V, Retief J, Matsuzaki H, Taub M, Seielstad M, Kennedy GC. 2006. A whole genome long-range haplotype (WGLRH) test for detecting imprints of positive selection in human populations. *Bioinformatics* 22:2122–2128.
- Zhuang J, Droma T, Sun S, Janes C, McCullough RE, McCullough RG, Cuymerman A, Huang SY, Reeves JT, Moore LG. 1993. Hypoxic ventilatory responsiveness in Tibetan compared with Han residents of 3,658 m. *J Appl Physiol* 74:303–311.