NIH Public Access
Author Manuscript
Science. Author manuscript; available in PMC 2013 July 15.

Published in final edited form as:
Science. 2010 July 2; 329(5987): 75–78. doi:10.1126/science.1190371.

Sequencing of Fifty Human Exomes Reveals Adaptation to High Altitude

Xin Yi[1,2,*], **Yu Liang**[1,2,*], **Emilia Huerta-Sanchez**[3,*], **Xin Jin**[1,4,*], **Zha Xi Ping Cuo**[2,5,*], **John E. Pool**[3,6,*], **Xun Xu**[1], **Hui Jiang**[1], **Nicolas Vinckenbosch**[3], **Thorfinn Sand Korneliussen**[7], **Hancheng Zheng**[1,4], **Tao Liu**[1], **Weiming He**[1,8], **Kui Li**[2,5], **Ruibang Luo**[1,4], **Xifang Nie**[1], **Honglong Wu**[1,9], **Meiru Zhao**[1], **Hongzhi Cao**[1,9], **Jing Zou**[1], **Ying Shan**[1,4], **Shuzheng Li**[1], **Qi Yang**[1], **Asan**[1,2], **Peixiang Ni**[1], **Geng Tian**[1,2], **Junming Xu**[1], **Xiao Liu**[1], **Tao Jiang**[1,9], **Renhua Wu**[1], **Guangyu Zhou**[1], **Meifang Tang**[1], **Junjie Qin**[1], **Tong Wang**[1], **Shuijian Feng**[1], **Guohong Li**[1], **Huasang**[1], **Jiangbai Luosang**[1], **Wei Wang**[1], **Fang Chen**[1], **Yading Wang**[1], **Xiaoguang Zheng**[1,2], **Zhuo Li**[1], **Zhuoma Bianba**[10], **Ge Yang**[10], **Xiznping Wang**[11], **Shuhui Tang**[11], **Guoyi Gao**[12], **Yong Chen**[5], **Zhen Luo**[5], **Lamu Gusang**[5], **Zheng Cao**[1], **Qinghui Zhang**[1], **Weihan Ouyang**[1], **Xiaoli Ren**[1], **Huiqing Liang**[1], **Huisong Zheng**[1], **Yebo Huang**[1], **Jingxiang Li**[1], **Lars Bolund**[1], **Karsten Kristiansen**[1,7], **Yingrui Li**[1], **Yong Zhang**[1], **Xiuqing Zhang**[1], **Ruiqiang Li**[1,7], **Songgang Li**[1], **Huanming Yang**[1], **Rasmus Nielsen**[1,3,7,#], **Jun Wang**[1,7,#], and **Jian Wang**[1,#]

[1]BGI-Shenzhen, Shenzhen 518083, China

[2]The Graduate University of Chinese Academy of Sciences, Beijing 100062, China

[3]Departments of Integrative Biology and Statistics, UC Berkeley, Berkeley CA 94820, USA

[4]Innovative program for undergraduate students, School of Bioscience and Biotechnology, South China University of Technology, Guangzhou, 510641, China

[5]The People's Hospital of the Tibet Autonomous Region, Lhasa, 850000, China

[6]Department of Evolution and Ecology, UC Davis, Davis, CA 95616, USA

[7]Department of Biology, University of Copenhagen, Copenhagen, Denmark

[8]Innovative program for undergraduate students, School of Science, South China University of Technology, Guangzhou 510641, China

[9]Genome Research Institute, Shenzhen University Medical School, Shenzhen, China

[10]The People's Hospital of Lhasa, Lhasa, 850000, China

[11]The military general hospital of Tibet, Lhasa, 850007, China

[12]The hospital of XiShuangBanNa Dai Nationalities Autonomous Jinghong 666100, China

Abstract

Residents of the Tibetan Plateau show heritable adaptations to extreme altitude. We sequenced 50 exomes of ethnic Tibetans, encompassing coding sequences of 92% of human genes, with an average coverage of 18X per individual. Genes showing population-specific allele frequency changes, which represent strong candidates for altitude adaptation, were identified. The strongest signal of natural selection came from *EPAS1*, a transcription factor involved in response to hypoxia. One SNP at *EPAS1* shows a 78% frequency difference between Tibetan and Han samples, representing the fastest allele frequency change observed at any human gene to date. This

[#]To whom correspondence should be addressed. rasmus_nielsen@berkeley.edu (R.N); wangj@genomics.org.cn (Ju.W.); wangjian@genomics.org.cn (Ji.W.).
[*]These authors contributed equally to this work.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

SNP's association with erythrocyte abundance supports the role of *EPAS1* in adaptation to hypoxia. Thus, a population genomic survey has revealed a functionally important locus in genetic adaptation to high altitude.

The expansion of humans into a vast range of environments may have involved both cultural and genetic adaptation. Among the most severe environmental challenges to confront human populations is the low oxygen availability of high altitude regions such as the Tibetan Plateau. Many residents of this region live at elevations exceeding 4000 meters, experiencing oxygen concentrations about 40% lower than at sea level. Ethnic Tibetans possess heritable adaptations to their hypoxic environment, as indicated by birth weight (1), hemoglobin levels (2) and oxygen saturation of blood in infants (3) and adults after exercise (4). These results imply a history of natural selection for altitude adaptation, which may be detectable from a scan of genetic diversity across the genome.

We sequenced the exomes of 50 unrelated individuals from two villages in the Tibet Autonomous Region of China, both at least 4300 m in altitude (5). Exonic sequences were enriched with the NimbleGen 2.1M exon capture array (6), targeting 34Mb of sequence from exons and flanking regions in nearly 20,000 genes. Sequencing was performed with the Illumina Genome Analyzer II platform and reads were aligned using SOAP (7) to the reference human genome (NCBI 36.3).

Exomes were sequenced to a mean depth of 18X (Table S1), which does not guarantee confident inference of individual genotypes. Therefore, we statistically estimated the probability of each possible genotype with a Bayesian algorithm (5) that also estimated single nucleotide polymorphism (SNP) probabilities and population allele frequencies for each site. 151,825 SNPs were inferred to have >50% probability of being variable within the Tibetan sample, and 101,668 had >99% SNP probability (Table S2). Sanger sequencing validated 53 of 56 SNPs that had at least 95% SNP probability and minor allele frequencies between 3% and 50%. Allele frequency estimates showed an excess of low frequency variants (Fig. S1), particularly for nonsynonymous SNPs.

The exome data was compared to 40 genomes from ethnic Han individuals from Beijing [(the HapMap CHB sample), part of the 1000 genomes project (http://1000genomes.org)], sequenced to about 4-fold coverage per individual. Beijing's altitude is less than 50 m above sea level, and nearly all Han come from altitudes below 2000 m. The Han sample represents an appropriate comparison for the Tibetan sample on the basis of low genetic differentiation between these samples (F_{ST} = 0.026). The two Tibetan villages show minimal evidence of genetic structure (F_{ST} = 0.014), and we therefore treated them as one population for most analyses. We observed a strong covariance between Han and Tibetan allele frequencies (Fig. 1), but with an excess of SNPs at low frequency in the Han and moderate frequency in the Tibetans.

Population historical models were estimated (8) from the two-dimensional frequency spectrum of synonymous sites in the two populations. The best-fitting model suggested that the Tibetan and Han populations diverged 2,750 years ago, with the Han population growing from a small initial size, and the Tibetan population contracting from a large initial size (Fig. S2). Migration was inferred from the Tibetan to the Han sample, with recent admixture in the opposite direction.

Genes with strong frequency differences between populations are potential targets of natural selection. However, a simple ranking of F_{ST} values would not reveal which population was affected by selection. Therefore, we estimated population-specific allele frequency change by including a third, more distantly related population. We thus examined exome sequences

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

from 200 Danish individuals, collected and analyzed as described for the Tibetan sample. By comparing the three pairwise F_{ST} values between these three samples, we can estimate the frequency change that occurred in the Tibetan population since its divergence from the Han population (5, 9). We found that this population branch statistic (*PBS*) has strong power to detect recent natural selection (Fig. S3).

Genes showing extreme Tibetan *PBS* values represent strong candidates for the genetic basis of altitude adaptation. The strongest such signals include several genes with known roles in oxygen transport and regulation (Table 1; Table S3). Overall, the 34 genes in our data set that fell under the gene ontology category "response to hypoxia" had significantly greater *PBS* values than the genome-wide average ($P = 0.00796$).

The strongest signal of selection came from the endothelial PAS domain protein 1 (*EPAS1*) gene. On the basis of frequency differences among the Danes, Han, and Tibetans, *EPAS1* was inferred to have a very long Tibetan branch relative to other genes in the genome (Fig. 2). In order to confirm the action of natural selection, *PBS* values were compared against neutral simulations under our estimated demographic model. None of one million simulations surpassed the *PBS* value observed for *EPAS1*, and this result remained statistically significant after accounting for the number of genes tested ($P < 0.02$ after Bonferroni correction). Many other genes had uncorrected *P* values below 0.005 (Table 1), and while none of these were statistically significant after correcting for multiple tests, the functional enrichment suggests that some of these genes may also contribute to altitude adaptation.

EPAS1 is also known as hypoxia-inducible factor 2α (*HIF-2α*). The *HIF* family of transcription factors consist of two subunits, with three alternate α subunits (*HIF-1α*, *HIF-2α*/*EPAS1*, *HIF-3α*) that dimerize with a β subunit encoded by *ARNT* or *ARNT2. HIF-1α* and *EPAS1* each act on a unique set of regulatory targets (10), and the narrower expression profile of *EPAS1* includes adult and fetal lung, placenta, and vascular endothelial cells (11). A protein-stabilizing mutation in *EPAS1* is associated with erythrocytosis (12), suggesting a link between *EPAS1* and the regulation of red blood cell production.

Although our sequencing primarily targeted exons, some flanking intronic and UTR sequence was included. The *EPAS1* SNP with the greatest Tibetan-Han frequency difference was intronic (with a derived allele at 9% frequency in the Han and 87% in the Tibetan sample; Table S4), whereas no amino acid-changing variant had a population frequency difference of greater than 6%. Selection may have acted directly on this variant, or another linked non-coding variant, to influence the regulation of *EPAS1*. Detailed molecular studies will be needed to investigate the direction and magnitude of gene expression changes associated with this SNP, the tissues and developmental time points affected, and the downstream target genes that show altered regulation.

Associations between SNPs at *EPAS1* and athletic performance have been demonstrated (13). Our data set contains a different set of SNPs, and we conducted association testing on the SNP with the most extreme frequency difference, located just upstream of the sixth exon. Alleles at this SNP tested for association with blood-related phenotypes showed no relationship with oxygen saturation. However, significant associations were discovered for erythrocyte count (F-test $P = 0.00141$) and for hemoglobin concentration (F-test $P = 0.00131$), with significant or marginally significant *P* values for both traits when each village was tested separately (Table S5). Comparison of the *EPAS1* SNP to genotype data from 48 unlinked SNPs confirmed that its *P* value is a strong outlier (5; Fig. S4).

The allele at high frequency in the Tibetan sample was associated with lower erythrocyte quantities, and correspondingly lower hemoglobin levels (Table S4). Since elevated

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

erythrocyte production is a common response to hypoxic stress, it may be that carriers of the "Tibetan allele" of *EPAS1* are able to maintain sufficient oxygenation of tissues at high altitude without the need for increased erythrocyte levels. Thus, the hematological differences observed here may not represent the phenotypic target of selection, and could instead reflect a side effect of *EPAS1*-mediated adaptation to hypoxic conditions. While the precise physiological mechanism remains to be discovered, our results suggest that the allele targeted by selection is likely to confer a functionally relevant adaptation to the hypoxic environment of high altitude.

We also identified components of adult and fetal hemoglobin (*HBB* and *HBG2*, respectively) as putatively under selection. These genes are located only ~20 kb apart (Fig. S5), so their *PBS* values could reflect a single adaptive event. For both genes, the SNP with the strongest Tibetan-Han frequency difference is intronic. Although altered globin proteins have been found in some altitude-adapted species (14), in this case regulatory changes appear more likely. A parallel result was reported in Andean highlanders, with promoter variants at *HBG2* varying with altitude and associated with a delayed transition from fetal to adult hemoglobin (15).

Aside from *HBB*, two other anemia-associated genes were identified: *FANCA* and *PKLR*, associated with erythrocyte production and maintenance, respectively (16, 17). We also identified genes associated with diseases linked to low oxygen during pregnancy or birth: schizophrenia (*DISC1* and *FXYD6*) (18, 19) and epilepsy (*OTX1*) (20). However, the strong signal of selection affecting *DISC1*, along with *C1orf124*, might instead trace to a regulatory region of *EGLN1*, which lies between these loci (Fig. S5) and functions in the hypoxia response pathway (21).

Other genes identified here are also located near candidate genes. *OR10X1* and *OR6Y1* are within ~60 kb of the *SPTA1* gene (Fig. S5), associated with erythrocyte shape (22). Additionally, the three histones implicated here (Table 1) are clustered around *HFE* (Fig. S5), a gene involved in iron storage (23). The influence of population genetic signals on neighboring genes is consistent with recent and strong selection imposed by the hypoxic environment. Stronger frequency changes at flanking genes might be expected if adaptive mutations have targeted candidate gene regulatory regions that are not near common exonic polymorphisms.

Of the genes identified here, only *EGLN1* was mentioned in a recent SNP variation study in Andean highlanders (24). This result is consistent with the physiological differences observed between Tibetan and Andean populations (25), suggesting that these populations have taken largely distinct evolutionary paths in altitude adaptation.

Several loci previously studied in Himalayan populations showed no signs of selection in our data set (Table S6), whereas *EPAS1* has not been a focus of previous altitude research. Although *EPAS1* may play an important role in the oxygen regulation pathway, this gene was identified based on a non-candidate population genomic survey for natural selection, illustrating the utility of evolutionary inference in revealing functionally important loci.

Given our estimate that Han and Tibetans diverged 2,750 years ago and experienced subsequent migration, it appears that our focal SNP at *EPAS1* may have experienced a faster rate of frequency change than even the lactase persistence allele in northern Europe, which rose in frequency over the course of about 7,500 years (26). *EPAS1* may therefore represent the strongest instance of natural selection documented in a human population, and variation at this gene appears to have had important consequences for human survival and/or reproduction in the Tibetan region.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are indebted to many additional faculty and staff of BGI-Shenzhen who contributed to this team work, and to X. Wang (South China University of Technology). This research was funded by the Danish Natural Science Research Council, the Solexa project (272-07-0196), the Shenzhen Municipal Government of China, the National Natural Science Foundation of China (30725008), the International Science and Technology Cooperation Project (0806), China (CXB200903110066A; ZYC200903240076A), the Danish Strategic Research Council grant (2106-07-0021), the Lundbeck Foundation, the Swiss National Science Foundation (PBLAP3-124318), the US National Institute of Health (R01MHG084695, R01HG003229), the U.S. National Science Foundation (DBI-0906065), the Ministry of Science and Technology of China (863 program: 2006AA02A302, 2009AA022707; 973 program: 2006CB504103), the Chinese Academy of Sciences (KSCX2-YW-R-76), Science and Technology Plan of the Tibet Autonomous Region (#2007-2-18), and the Basic Science Foundation of ShenZhen (JC200903190772A). The data have NCBI Short Read Archive accession number SRA012603.

References and Notes

1. Moore LG. Human genetic adaptation to high altitude. High Alt Med Biol. 2001; 2:257–279. [PubMed: 11443005]
2. Wu T, et al. Hemoglobin levels in Qinghai-Tibet: different effects of gender for Tibetans vs. Han. J Appl Physiol. 2005; 98:598–604. [PubMed: 15258131]
3. Niermeyer S, et al. Arterial oxygen saturation in Tibetan and Han infants born in Lhasa, Tibet. NEJM. 1995; 333:1248–1252. [PubMed: 7566001]
4. Zhuang J, et al. Smaller alveolar-arterial O2 gradients in Tibetan than Han residents of Lhasa (3658 m). Respiration Physiology. 1996; 103:75–82. [PubMed: 8822225]
5. Materials and methods are available as supporting material on *Science* Online.
6. Albert TJ, et al. Direct selection of human genomic loci by microarray hybridization. Nat Methods. 2007; 4:903–905. [PubMed: 17934467]
7. Li R, et al. SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics. 2009; 25:1966–1967. [PubMed: 19497933]
8. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genetics. 2008; 5:e1000695. [PubMed: 19851460]
9. Shriver MD, et al. The genomic distribution of population substructure in four populations using 8,525 autosomal SNPs. Human Genomics. 2004; 1:274–286. [PubMed: 15588487]
10. Hu C-J, Wang L-Y, Chodosh LA, Keith B, Simon MC. Differential roles of hypoxia-indicible factor 1α (HIF-1α) and HIF-2α in hypoxic gene regulation. Mol Cell Biol. 2003; 23:9361–9374. [PubMed: 14645546]
11. Jain S, Maltepe E, Lu MM, Simon C, Bradfield CA. Expression of ARNT, ARNT2, HIF1α, HIF2α and Ah receptor mRNAs in the developing mouse. Mech Dev. 1998; 73:117–123. [PubMed: 9545558]
12. Percy MJ, et al. A gain-of-function mutation in the *HIF2A* gene in familial erythrocytosis. NJEM. 2008; 358:162–168.
13. Henderson J, et al. The *EPAS1* gene in uences the aerobic–anaerobic contribution in elite endurance athletes. Hum Genet. 2005; 118:416–423. [PubMed: 16208515]
14. Storz JF, et al. Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. Proc Natl Acad Sci USA. 2009; 106:14450–14455. [PubMed: 19667207]
15. Rottgardt I, Rothhammer F, Dittmar M. Native highland and lowland populations differ in γ-globin gene promoter polymorphisms related to altered fetal hemoglobin levels and delayed fetal to adult globin switch after birth. Anthropol Sci. 2010; 118:41–48.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

16. Kanno H, Fujii H, Hirono A, Omime M, Miwa S. Identical point mutations of the R-type pyruvate kinase (PK) cDNA found in unrelated PK variants associated with hereditary hemolytic anemia. Blood. 1992; 79:1347–1350. [PubMed: 1536957]
17. Zhang X, Li J, Sejas DP, Pang Q. Hypoxia-reoxygenation induces premature senescence in FA bone marrow hematopoietic cells. Blood. 2005; 106:75–85. [PubMed: 15769896]
18. Hodgkinson CA, et al. Disrupted in Schizophrenia 1 (DISC1): association with schizophrenia, schizoaffective disorder, and bipolar disorder. Am J Hum Genet. 2004; 75:862–872. [PubMed: 15386212]
19. Choudhury K, et al. A genetic association study of chromosome 11q22–24 in two different samples implicates the FXYD6 gene, encoding phosphohipplolin, in susceptibility to schizophrenia. Am J Hum Genet. 2007; 80:664–672. [PubMed: 17357072]
20. Acampora D, et al. Epilepsy and brain abnormalities in mice lacking the Otx1 gene. Nat Genet. 1996; 14:218–222. [PubMed: 8841200]
21. To KKW, Huang LE. Suppression of hypoxia-inducible factor 1α (HIF1α) transcriptional activity by the HIF prolyl hydroxylase EGLN1. J Biol Chem. 2005; 280:38102–38107. [PubMed: 16157596]
22. Gaetani M, Mootien S, Harper S, Gallagher PG, Speicher DW. Structural and functional effects of hereditary hemolytic anemia-associated point mutations in the alpha spectrin tetramer site. Blood. 2008; 111:5712–5720. [PubMed: 18218854]
23. Hentze MW, Muckenthaler MU, Andrews NC. Balancing acts: Molecular control of mammalian iron metabolism. Cell. 2004; 117:285–297. [PubMed: 15109490]
24. Bigham AW, et al. Identifying positive selection candidate loci for high-altitude adaptation in Andean populations. Human Genomics. 2009; 4:79–90. [PubMed: 20038496]
25. Beall CM. Two routes to functional adaptation: Tibetan and Andean high-altitude natives. Proc Natl Acad Sci USA. 2007; 104:8655–8660. [PubMed: 17494744]
26. Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. The origins of lactase persistence in Europe. PLoS Comput Biol. 2009; 5:e1000491. [PubMed: 19714206]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Summary

Sequencing and analysis of 50 exomes from ethnic Tibetans has led to the discovery of genes involved in adaptation to extreme altitude, including the *EPAS1* gene which shows evidence of the strongest natural selection observed at any human gene.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

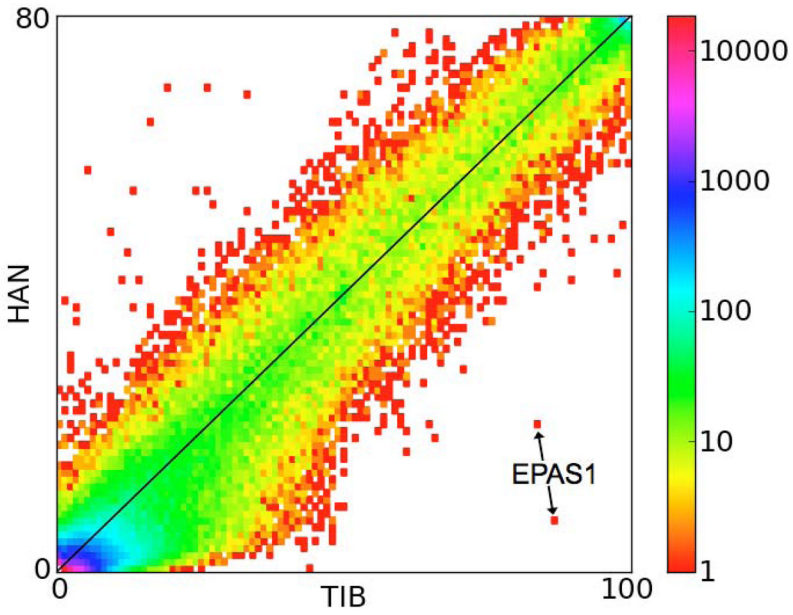

Figure 1.

Two-dimensional unfolded site frequency spectrum for SNPs in Tibetan (x-axis) and Han (y-axis) population samples. The number of SNPs detected is colored coded according to the logarithmic scale plotted on the right. Arrows indicate a pair of intronic SNPs from the *EPAS1* gene that show strikingly elevated derived allele frequencies in the Tibetan sample compared to the Han.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

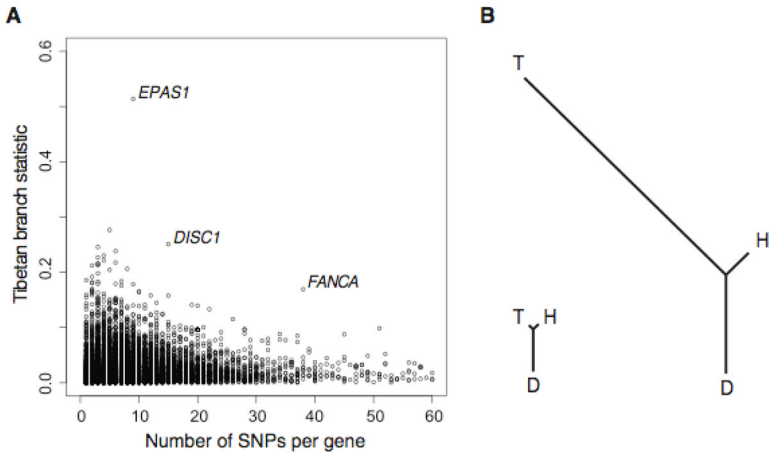

Figure 2.

Population-specific allele frequency change. **(A)** The distribution of F_{ST}-based *PBS* statistics for the Tibetan branches, according to the number of variable sites in each gene. Outlier genes are indicated in red. **(B)** The signal of selection on *EPAS1*: Genomic average F_{ST}-based branch lengths for Tibetan (T), Han (H), and Danish (D) branches (left), and branch lengths for *EPAS1*, indicating substantial differentiation along the Tibetan lineage (right)

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1
Genes with strongest frequency changes in the Tibetan population

The top 30 *PBS* values for the Tibetan branch are listed. Oxygen-related candidate genes within 100kb of these loci are noted.

Gene	Description	Nearby candidate	*PBS*	*P* value
EPAS1	endothelial PAS domain protein 1 (HIF-2α)	(self)	0.514	<0.000001
C1orf124	hypothetical protein LOC83932	*EGLN1*	0.277	0.000203
DISC1	disrupted in schizophrenia 1	*EGLN1*	0.251	0.000219
ATP6V1E2	ATPase, H+ transporting, lysosomal 31kDa, V1	*EPAS1*	0.246	0.000705
SPP1	secreted phosphoprotein 1		0.238	0.000562
PKLR	pyruvate kinase, liver and RBC	(self)	0.230	0.000896
C4orf7	chromosome 4 open reading frame 7		0.227	0.001098
PSME2	proteasome activator subunit 2		0.222	0.001103
OR10X1	olfactory receptor, family 10, subfamily X	*SPTA1*	0.218	0.000950
FAM9C	family with sequence similarity 9, member C	*TMSB4X*	0.216	0.001389
LRRC3B	leucine rich repeat containing 3B		0.215	0.001405
KRTAP21–2	keratin associated protein 21–2		0.213	0.001470
HIST1H2BE	histone cluster 1, H2be	*HFE*	0.212	0.001568
TTLL3	tubulin tyrosine ligase-like family, member 3		0.206	0.001146
HIST1H4B	histone cluster 1, H4b	*HFE*	0.204	0.001404
ACVR1B	activin A type IB receptor isoform a precursor	*ACVRL1*	0.198	0.002041
FXYD6	FXYD domain-containing ion transport regulator		0.192	0.002459
NAGLU	alpha-N-acetylglucosaminidase precursor		0.186	0.002834
MDH1B	malate dehydrogenase 1B, NAD (soluble)		0.184	0.002113
OR6Y1	olfactory receptor, family 6, subfamily Y	*SPTA1*	0.183	0.002835
HBB	beta globin	(self), *HBG2*	0.182	0.003128
OTX1	orthodenticle homeobox 1		0.181	0.003235
MBNL1	muscleblind-like 1		0.179	0.002410
IFI27L1	interferon, alpha-inducible protein 27-like 1		0.179	0.003064
C18orf55	hypothetical protein LOC29090		0.178	0.002271
RFX3	regulatory factor X3		0.176	0.002632
HBG2	G-gamma globin	(self), *HBB*	0.170	0.004147
FANCA	Fanconi anemia, complementation group A	(self)	0.169	0.000995
HIST1H3C	histone cluster 1, H3c	*HFE*	0.168	0.004287
TMEM206	transmembrane protein 206		0.166	0.004537