www.sciencemag.org/cgi/content/full/science.1189406/DC1

Supporting Online Material for

# Genetic Evidence for High-Altitude Adaptation in Tibet

Tatum S. Simonson, Yingzhong Yang, Chad D. Huff, Haixia Yun, Ga Qin, David J. Witherspoon, Zhenzhong Bai, Felipe R. Lorenzo, Jinchuan Xing, Lynn B. Jorde,* Josef T. Prchal,* RiLi Ge*

*To whom correspondence should be addressed. E-mail: lbj@genetics.utah.edu (L.B.J); josef.prchal@hsc.utah.edu (J.T.P.); geriligao@hotmail.com (R.G.)

Published 13 May 2010 on *Science* Express
DOI: 10.1126/science.1189406

**This PDF file includes:**

**Supporting Online Material:**

**Methods, Figure and Table Legends**

## Materials and Methods

### DNA sample collection

DNA was extracted from whole blood samples for 49 individuals (non-smokers, no chronic diseases) residing in Madou County in Qinghai province (~4,350 m). Informed consent was obtained for all participants according to guidelines approved by the Institutional Review Boards at the High Altitude Medical Research Institute (Xining, Qinghai, People's Republic of China).

### SNP genotyping

Forty-nine individual DNA samples were genotyped using Affymetrix 6.0 SNP Array technology (>900,000 SNPs) at Capital Bio Corporation (Beijing, China). We used default parameters for the Birdseed algorithm (version 2) to determine genotypes for all samples (Affymetrix, Santa Clara, CA, USA). Genotypic data were analyzed using the Affymetrix Genotyping Console 3.1 (Affymetrix) and included all autosomes but excluded the X and Y chromosomes and mitochondrial genome.

### Principal components analysis

We performed principal components analysis (fig. S1) based on genetic distances as previously described (*S1*). This analysis indicates that Tibetans form a distinct group but are genetically similar to the combined CHB-JPT population (CHB = Chinese in Beijing, China; JPT = Japanese in Tokyo, Japan), which serves as an appropriate lowland East Asian comparison population (panel A of fig. S1). The CEU (U.S. Utah residents with ancestry from northern and

western Europe) and YRI (YRI = Yoruba in Ibadan, Nigeria) HapMap populations provide context for the patterns of variation observed among these populations (*S2*).

**Estimates of relatedness**

We collected samples at a small town health center where groups of semi-nomadic clans visit. Prior to genotyping, we excluded first-degree relatives who visited the clinic. It is possible that related individuals were examined at different times throughout the collection process; therefore, relatedness could only be determined after genotype analysis. We used pair-wise genetic distances and the proportion of shared genomic segments to determine relatedness between subjects (*S1, S3*). When pairs of individuals exhibited genetic distances less than 4.95 x $10^{-2}$ or had genome-wide identity-by-descent of greater than 400cM (minimum segment size 2.5cM), one member of the pair was excluded from the analyses. Based on these criteria, a total of 31 unrelated individuals were included in the analyses.

***A priori* functional candidate list**

We generated a list of genes likely related to high-altitude adaptation based on categories provided in table S1. We coupled genes associated with Gene Ontology (GO) categories (*S4*) that may be involved in the observed high-altitude Tibetan phenotypes (481 genes), with genes listed in the "Hypoxia response via HIF activation" defined by Panther Pathways (33 genes) (*S5*). Potential candidate genes identified in the mitochondrial genome and on the X chromosome were not considered for this study.

Although the intersection of functional and selection candidate lists is enriched for hypoxia-related signals of selection, the ten genes we identified probably do not account for all high-altitude adaptive traits in this population. For example, the genomic region containing *HIF1AN*, an inhibitor of HIF in normoxic conditions, was significant in both the XP-EHH and

iHS tests, although it was not included in our *a priori* list of functional candidates (Fig. 2; tables S2, S11, and S12). Other genes in our functional candidate list are found within regions identified in the top 2% of the selection scans, such as the human ß-globin gene cluster (table S2, S11, and S12). These post hoc findings reflect the conservative approach used to define our list of genes for high-altitude adaptation in Tibetans.

**Admixture analysis**

A model-based algorithm implemented in *ADMIXTURE* (*S6*) was used to determine the genetic ancestries of each individual in a given number of populations without using information about population designation. To eliminate the effects of SNPs that are in linkage disequilibrium (LD), we first filtered out SNPs that had $r^2 > 0.2$ within 100kb using PLINK (*S7*), as recommended by the authors of *ADMIXTURE*. The pruned data set contains 142,888 SNPs.

While the demographic history of the Tibetan Plateau is unclear (i.e., whether modern Tibetans descended from populations who occupied this region during the mid-Holocene, the Late Pleistocene, or if they are an admixed population (*S8*), this analysis indicates a distinct relationship between Tibetans and East Asian (CHB-JPT) populations. We see no strong evidence of admixture in our samples (fig. S2), although signals of selection should be detectable even if Tibetans were admixed or descended from populations who occupied this region prior to or during the mid-Holocene.

**Selection analyses**

We used the Beagle software package to estimate phase in the 31 unrelated Tibetan individuals (*S9*) and calculated all selection statistics from the phased data. To calculate iHH for each allele at each site, we integrated the expected EHH in both directions from the core SNP until expected EHH was less than or equal to 0.10 (*S10, S11*). To calculate iHS, we calculated

the log of the ratio of iHH scores at each site for the derived and ancestral alleles, standardizing within each population by the derived allele frequency. We computed iHS scores in this manner for all SNPs on the Affymetrix 6.0 microarray with at least 10 copies of the derived allele and the ancestral allele in a given population. For the iHS selection scan, our test statistic for each 200kb genomic region was the fraction of SNPs in each region where |iHS|>2.0, excluding regions with fewer than 5 SNPs (*S10, S11*). We calculated XP-EHH at each site using the default settings of the XP-EHH software (*S12*). For the XP-EHH selection scan, our test statistic was the maximum XP-EHH score in each 200kb region (*S13*). We determined statistical significance for each 200kb region from the empirical distribution of each test statistic. Our selection candidates are those genes contained in any of the 200kb regions significant at the 0.01 level in either test, excluding regions where the iHS test was significant at the 0.01 level in neighboring populations from Mongolia (n = 25 ), India (n = 25), Nepal (n = 25), China and Japan (CHB-JPT: n = 90), Kyrgyzstan (n = 25), and Thailand (n = 25) (unpublished data for samples provided by Scott Woodward and the Sorenson Molecular Genealogy Foundation, Salt Lake City, Utah, USA). The goal of this exclusion step was to enrich for signals of local adaptation in the Tibetan population by filtering out signals of selection present in other Asian populations. Our exclusion criteria only included significant iHS results because the comparison population (CHB-JPT) in the XP-EHH test directly controls for genomic variation in a neighboring population. Tables S11 and S12 contain all 200kb genomic regions identified in the top 2% of each selection scan.

**Analyses for localization of selection signal**

Although the composite of multiple tests (CMS) statistic is not applicable for localization of our selection signals (due to a lack of detailed information about Tibetan demographic history), we have conducted analyses for the three additional statistics reported by Grossman et

al.: $F_{ST}$, ΔDAF, and ΔiHH (*S14*). If demographic information were available, these statistics could be combined into a single test, but we present the separate results of each test for the ten gene regions described (fig. S4).

**Phenotype collection**

Hemoglobin concentration, hematocrit, and percent oxygen saturation were determined from venous blood samples using the Mindray Hematology Analyzer (BC-2300, Shenzhen, People's Republic of China) and the Pulse Oximeter (Ohmeda 3700 Pulse Oximeter, Datex-Ohmeda, Boulder, Colorado, USA), respectively. Hematocrit values are highly correlated with [Hb] ($r = 0.861$, $p < 10^{-9}$). See table S6 for phenotype measurements.

**Genotype-phenotype association**

For the five iHS selection candidates that intersect our functional candidate list (Fig. 1; Table 2), we identified the putatively advantageous haplotypes as those carrying the SNP alleles responsible for the most extreme iHS scores within the corresponding 200 kb genomic region. Ideally, we would test for a direct correlation between the advantageous genetic variants and Hb concentration. However, our selection scan results provide only indirect inferences about SNPs that are linked to the putatively advantageous variant. Because the sign of an iHS score indicates an excess of homozygosity around the ancestral (+) or derived (-) allele, the allele designated by an extreme iHS score is frequently linked closely to the advantageous allele during a selective sweep (table S6). Therefore, we selected the three alleles exhibiting the most extreme iHS scores within each 200kb genomic region to construct haplotypes that partially tag the putatively advantageous variants. We are able to test for an association between the putatively advantageous haplotypes at these loci and a phenotype. Stepwise linear regression (MATLAB R2009b) was used to detect significant relationships between these genotypes and hemoglobin

concentration in 30 Tibetan individuals (after excluding one tobacco smoker) (Fig. 3; fig. S5; tables S7-S9) and oxygen saturation in 29 individuals (after exclusion of one missing data point) (fig. S7; table S10).

**Verification of performance of linear multivariate regression between hemoglobin concentration and SNP genotypes in the Tibetan data set**

To check the reliability of the stepwise linear multivariate regression method that we used to test for an association of genotypes with [Hb], we used the genome-wide SNP data to generate an empirical p-value distribution to compare with the theoretically expected distribution. We repeated the regression analysis for the same 30 unrelated Tibetan individuals for each SNP, retaining the original values for the sex, male age, female age, and [Hb], but changing the value of the SNP genotype variable. This value was replaced with the genotype (in the corresponding individuals) at each SNP selected from the autosomes, subject to the constraints that (1) the SNP contained no missing data (no 'No Call' genotypes); (2) the minor allele frequency of each SNP was ≥ 15% (i.e., the minor allele was observed at least nine times in the 60 chromosomes). 398,020 SNPs meet these criteria. The regression method tests the variable against the null hypothesis that the effect size coefficient for that variable would be zero if it were included in the regression model and reports the resulting p-values.

Fig. S5 shows a quantile-quantile (QQ) plot of those p-values compared with a uniform distribution (the expected distribution for p-values). There is a near-perfect match between the theoretical and observed p-value distributions, as expected. Moreover, the mean correlation (Pearson's $r$) between the individual genotypes at these 398,020 SNPs and [Hb] is essentially zero (mean $1.08\times10^{-4}$, median $1.67\times10^{-4}$, standard deviation 0.185). There is no evidence of

population stratification in the data set that could produce excess associations between genotypes and [Hb] (see also the PCA and Admixture results, figs. S1 and S2).

Supplemental References


S1. J. Xing *et al.*, *Genome Research* **19**, 815 (2009).
S2. C. The International HapMap, *Nature* **437**, 1299 (2005).
S3. A. Gusev *et al.*, *Genome Research* **19**, 318 (2009).
S4. C. Gene Ontology, *Nucl. Acids Res.* **34**, D322 (2006).
S5. H. Mi, N. Guo, A. Kejariwal, P. D. Thomas, *Nucl. Acids Res.* **35**, D247 (2007).
S6. D. H. Alexander, J. Novembre, K. Lange, *Genome Research* **19**, 1655 (2009).
S7. S. Purcell *et al.*, *The American Journal of Human Genetics* **81**, 559 (2007).
S8. M. Zhao *et al.*, *Proceedings of the National Academy of Sciences* **106**, 21230 (2009).
S9. B. L. Browning, S. R. Browning, *The American Journal of Human Genetics* **84**, 210 (2009).
S10. C. Huff, H. Harpending, A. Rogers, *BMC Genomics* **11**, 8 (2010).
S11. B. F. Voight, S. Kudaravalli, X. Wen, J. K. Pritchard, *PLoS Biology* **4**, e72 (2006).
S12. http://hgdp.uchicago.edu/Software/).
S13. P. C. Sabeti *et al.*, *Nature* **449**, 913 (2007).
S14. S. R. Grossman *et al.*, *Science* **327**, 883 (2010).
S15. C. M. Beall, J. Blangero, S. Williams-Blangero, M. C. Goldstein, *American Journal of Physical Anthropology* **95**, 271 (1994).
S16. H. Cai, D. Liu, J. G. N. Garcia, *Cardiovascular Research* **77**, 30 (2008).
S17. S. E. J. Williams *et al.*, *Science* **306**, 2093 (2004).
S18. J. Z. He *et al.*, *Journal of Biological Chemistry* **285**, 10 (2010).
S19. W. Zundel *et al.*, *Genes & Development* **14**, 391 (2000).

**Supplemental Figures**

Fig. S1. Principal Components Analysis (PCA) of Tibetans and the HapMap populations based on genome-wide SNP data. CHB = Chinese in Beijing, China; JPT = Japanese in Tokyo, Japan; YRI = Yoruba in Ibadan, Nigeria; CEU (U.S. Utah residents with ancestry from northern and western Europe) (10). (**A**) PCA of Tibetans and East Asian populations (CHB and JPT). (**B**) PCA of Tibetans and HapMap populations. The Tibetans are genetically most similar to (but still distinct from) the CHB-JPT populations; therefore, the CHB-JPT population comparison for XP-EHH analysis is appropriate for identifying regions of the genome subject to positive selection in Tibetans.

Fig. S2. Individual grouping inferred by *ADMIXTURE* at K = 4. Each individual's genome is represented by a vertical bar composed of colored sections, where each section represents the proportion of an individual's ancestry derived from one of the K ancestral populations. Individuals are arrayed horizontally and grouped by population as indicated.

Fig. S3. Haplotype structure for selection candidates.

The defined selection candidate haplotype for regions with (A) significant XP-EHH scores and (B) significant iHS scores that intersect our list of *a priori* functional candidate genes. The top and bottom half of each figure represents chromosome regions in the Tibetan (n = 62 chromosomes) and CHB-JPT populations (n = 62 chromosome regions randomly sampled from 90 unrelated individuals), respectively. An * indicates the SNP used to define the haplotype for each region. A reference haplotype was assigned as the longest haplotype which contains this core set of SNPs. All haplotypes were sorted based on length of uninterrupted matches to the

reference sequence. The longer haplotypes are sorted to the midline for both halves of each panel (to the bottom of the top half for the Tibetans and to the top of bottom half for CHB-JPT). Genetic lengths are provided in table S5.

Fig. S4. Selection signal in ten genomic regions for iHS, XP-EHH, $F_{ST}$, ΔDAF, and ΔiHH statistics. The bottom axis represents the one cM region which contains the selected 200kb region identified and gene of interest as indicated. The names and values for each statistic are provided on the left axis.

Fig. S5. Quantile-Quantile plot for the regression analysis comparing the distribution of observed genome-wide SNP p values vs. a uniform random distribution on (0,1).

Fig. S6. Genotype-phenotype association of the inferred adaptive iHS haplotypes with oxygen saturation. Because previous studies have suggested that offspring survival is positively correlated with a high-oxygen saturation allele (*S15*), we also tested for a relationship between oxygen saturation and haplotype variation. Although the *HMOX2* haplotype shows a positive relationship with oxygen saturation, the relationship is only marginally significant ($p = 0.07$) (see table S10).

**Supplemental Tables**

Supplemental tables

**Table S1. Description of putatively advantageous genes identified by the intersection of functional and selection candidate gene lists**

| Gene (alias) identified in 200kb region | Description | Category |
|---|---|---|
| ***EPAS1*(*HIF2A*)** | HIF-family transcription factor; up-regulated under hypoxic conditions, regulates vascular endothelial growth factor expression, erythropoietin expression in the brain and liver | GO: Response to hypoxia, Response to hypoxia levels |
| ***CYP2E1*(*CPE1*)** | Catalyzes many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids | GO: Oxygen binding |
| ***EDNRA (ET-A)*** | Encodes a cell surface receptor for endothelin-1 | GO: Response to hypoxia, Response to hypoxia levels |
| ***ANGPTL4(FIAF)*** | Directly involved in regulating glucose homeostasis, lipid metabolism, and insulin sensitivity and also acts as an apoptosis survival factor for vascular endothelial cells | GO: Response to hypoxia, Response to hypoxia levels |
| ***CAMK2D*(*CAMKD*)** | Serine/threonine protein kinase family and to the Ca(2+)/calmodulin-dependent protein kinase subfamily; mediates nitric oxide production in response to changes in intracellular calcium (*S16*) | GO: Response to hypoxia, Response to oxygen levels |
| ***EGLN1(PHD2)*** | Catalyzes post-translational hydroxylation of the two HIF alpha proteins (HIF1a and HIF2a), targeting them for proteasomal degradation in normoxic conditions | Panther: Hypoxia response via HIF activation; GO: Response to hypoxia, Response to oxygen levels |
| ***HMOX2* (*HO-2*)** | Involved in oxygen sensing independent of the HIF pathway (*S17*); enhanced expression preserves endothelial cell viability during hypoxia (*S18*); the NmrA-like family domain containing 1 (*NMRAL1*) gene, which encodes a protein involved in NO synthesis, is located within the region containing *HMOX2* | GO: Response to hypoxia, Response to oxygen levels |
| ***CYP17A1*(*CPT7*)** | Mono-oxygenases which catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids | GO: Oxygen binding |
| ***PPARA* (*NR1C1*)** | Nuclear hormone-binding protein transcriptional regulator that controls the peroxisomal beta-oxidation pathway of fatty acids | GO: Response to hypoxia, Response to oxygen levels |
| ***PTEN* (*TEP1*)** | Lipid phosphatase, tumor suppressor that antagonizes the PI3K-AKT/PKB signaling pathway, thereby modulating cell cycle progression and cell survival. Loss of *PTEN* increases HIF activity (*S19*) | Panther: Hypoxia response via HIF activation |

**Table S2. List of *a priori* functional candidate genes**

[*] Genes within regions identified in the functional candidate list at $p < 0.01$ level
[†] Genes within regions identified in the functional candidate list at $p < 0.02$ level
[††] Genes identified in selected genomic regions of neighboring Asian populations at $p < 0.01$ level

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *ABAT* | *BCL2* | *CYP1B1* | *EGLN2* | *HRH1* | *MT3* | *PIK3CD* | *SLC8A1*[†] | *UCP3* |
| *ACE* | *BCL2L1* | *CYP26A1*[††] | *EGLN3* | *HSD11B2* | *NARFL* | *PIK3CG* | *SMAD3* | *USF1* |
| *ACTN4* | *BIRC2* | *CYP2A7* | *ENG* | *HSP90AA1* | *NF1* | *PIK3R1* | *SMAD4* | *VEGFA* |
| *ADA* | *BNIP3* | *CYP2B6* | *EP300* | *HSP90AB1* | *NGB* | *PIK3R2* | *SMAD9* | *VHL* |
| *ADAM17*[††] | *CA9* | *CYP2C18* | *EPAS1*[*] | *HSP90B1* | *NOS1* | *PIK3R3* | *SOCS3* | *VHLL* |
| *ADIPOQ* | *CABC1* | *CYP2C19* | *EPHX2* | *HYOU1* | *NOS2* | *PIP3-E* | *SOD1* | *VLDLR* |
| *ADM* | *CALCA* | *CYP2C8* | *EPO* | *ICAM1* | *NOS3* | *PLAT* | *SOD2* | *XRCC1* |
| *ADORA1* | *CALCB* | *CYP2C9* | *ERCC3* | *IFNG* | *NOX4*[†] | *PLAU* | *SOD3* | |
| *ADORA2A* | *CAMK2D*[*] | *CYP2E1*[*] | *FLT1* | *IL10* | *NPPB* | *PLOD1* | *SPR*[††] | |
| *ADORA2B* | *CAPN2* | *CYP2F1* | *FRAP1* | *IL18* | *NPPC* | *PLOD2* | *STAT5B*[††] | |
| *ADRB1* | *CASP1* | *CYP2U1* | *GCH1* | *IL1B* | *NPR1* | *PML* | *TDO2* | |
| *ADRB2* | *CAV1* | *CYP3A4* | *GCHFR* | *INDO* | *NQO1* | *PPARA*[*] | *TFRC* | |
| *AGTR1* | *CCL2* | *CYP3A5* | *GIMAP1* | *INS* | *NR4A2* | *PRKAA1* | *TGFB1* | |
| *AKT1* | *CD38* | *CYP3A7* | *GIMAP5* | *INSR* | *OXTR* | *PRKCQ* | *TGFB2* | |
| *AKT2* | *CDKN1A* | *CYP4A11* | *GPR182* | *ITGA1* | *P2RX3*[††] | *PSEN2* | *TGFB3* | |
| *AKT3* | *CFTR* | *CYP4B1* | *GPX1*[††] | *ITGA2* | *P2RX4* | *PTEN*[*] | *TGFBR1* | |
| *ALB* | *CHRNA4* | *CYP4F12* | *GUCY1A3* | *ITPR1* | *PDE5A* | *PTK2B* | *TH* | |
| *ALDH2* | *CHRNA7* | *CYP4F2* | *HAAO* | *ITPR2* | *PDGFA* | *PTX3* | *THBS1*[†] | |
| *ALDOC* | *CHRNB2* | *CYP4F3* | *HBA1* | *JAG2* | *PDGFB* | *PYGM* | *TICAM1* | |
| *ANG* | *CITED2* | *CYP8B1* | *HBB*[†] | *JAK2* | *PDGFRA* | *RORA* | *TNF* | |
| *ANGPT1* | *CLDN3* | *DDAH1* | *HBD*[†] | *KCNA5* | *PDIA2* | *RORB* | *TPTE* | |
| *ANGPTL4*[*] | *CPS1* | *DDAH2* | *HBE1*[†] | *KCNJ8* | *PDLIM1* | *RORC* | *TPTE2* | |
| *APOE* | *CREBBP* | *DDIT4* | *HBG2*[†] | *KCNMA1* | *PDPN* | *RYR1* | *TRH* | |
| *APOLD1* | *CXCR4* | *DPP4* | *HBM* | *KLRK1* | *PGF* | *RYR2* | *TXN* | |
| *ARG2* | *CYB5R4* | *ECE1* | *HBQ1* | *KNG1* | *PIK3C2A* | *SCNN1B* | *TXN2* | |
| *ARNT* | *CYGB* | *EDN1* | *HBZ* | *LCT* | *PIK3C2B* | *SCNN1G* | *TXNDC2* | |
| *ARNT2* | *CYP17A1*[*] | *EDNRA* | *HIF1A* | *LONP1* | *PIK3C2G*[††] | *SERPINA1* | *UBE2B* | |
| *ASCL2* | *CYP19A1* | *EDNRB* | *HIF3A* | *MB* | *PIK3C3* | *SHH* | *UBQLN1* | |
| *ATG5* | *CYP1A1* | *EGFR* | *HMOX1* | *MMP14* | *PIK3CA* | *SLC11A2* | *UCN* | |
| *ATP1B1* | *CYP1A2* | *EGLN1*[*] | *HMOX2*[*] | *MMP2* | *PIK3CB* | *SLC2A8* | *UCN3* | |

**Table S3. Information used for XP-EHH randomization test**

| | XP-EHH | |
|---|---|---|
| | *A priori* functional candidate genes | Non-*a priori* genes |
| Number of genes within 200kb regions identified by XP-EHH | 6 | 207 |
| Number of genes not identified by XP-EHH | 236 | 18762 |

**Table S4. Information used for iHS randomization test**

| | iHS | |
|---|---|---|
| | *A priori* functional candidate genes | Non-*a priori* genes |
| Number of genes within 200kb regions identified by iHS | 5 | 109 |
| Number of genes not identified by iHS | 237 | 18762 |

**Table S5. Haplotype region information for Fig. 1 and Fig. S3**

| Gene | HG 18 Position | Number of SNPs | Length (bp) | Genetic Length (cM) |
|---|---|---|---|---|
| *EGLN1* | Chr1: 229296131-230232917 | 202 | 936786 | 1.26 |
| *EPAS1* | Chr2: 46304028-46851921 | 207 | 547893 | 0.94 |
| *CAMK2D* | Chr4: 114384034-114723610 | 81 | 339576 | 0.68 |
| *EDNRA* | Chr4: 148218788-149122838 | 159 | 904050 | 0.78 |
| *PTEN* | Chr8: 89408710-89951510 | 99 | 542800 | 0.88 |
| *CYP17A1* | Chr10: 103834925-105315160 | 194 | 1480235 | 0.44 |
| *CYP2E1* | Chr10: 134765822-135284541 | 87 | 518719 | 0.65 |
| *HMOX2* | Chr16: 4146014-4950914 | 121 | 804900 | 1.18 |
| *ANGPTL4* | Chr19: 8039905-8615907 | 62 | 576002 | 1.56 |
| *PPARA* | Chr22: 44338204-45322351 | 139 | 984147 | 2.32 |

**Table S6. SNPs used to define the selection candidate haplotypes for iHS genotype-phenotype analysis**

| Gene identified in 200kb region | First Allele of Core Selection Haplotype | | | | Second Allele of Core Selection Haplotype | | | | Third Allele of Core Selected Haplotype | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | HG 18 Position* | iHS Score | Selected Allele | Alternate Allele | HG 18 Position* | iHS Score | Selected Allele | Alternate Allele | HG 18 Position* | iHS Score | Selected Allele | Alternate Allele |
| *EGLN1* | Chr1: 229793717 | 2.68 | A | T | Chr1: 229667980 | 2.45 | T | C | Chr1: 229665156 | 2.34 | T | C |
| *CYP17A1* | Chr10: 104568521 | 4.00 | G | C | Chr10: 104594906 | -3.54 | G | T | Chr10: 104517420 | 3.49 | C | G |
| *PTEN* | Chr10: 89770364 | 3.90 | G | C | Chr10: 89790851 | 2.72 | C | T | Chr10: 89778618 | 2.68 | C | T |
| *HMOX2* | Chr16: 4456093 | 3.00 | C | T | Chr16: 4465266 | 2.87 | T | C | Chr16: 4442515 | 2.84 | T | C |
| *PPARA* | Chr22: 44827140 | 3.58 | A | G | Chr22: 44832376 | -2.72 | C | A | Chr22: 44842095 | -2.55 | T | C |

*Based on UCSC Genome Browser Human Reference Build 18 (*S4*)

**Table S7. Phenotype and core haplotype data for 30 Tibetan individuals**

| | | | | | Number of selection candidate haplotypes per individual | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Age | Gender | Hemoglobin (g/dL) | Hematocrit (%) | Oxygen Saturation (%) | *EGLN1* | *PPARA* | *HMOX2* | *CYP17A1* | *PTEN* |
| 22 | F | 14.4 | 40.1 | 91 | 0 | 2 | 2 | 0 | 1 |
| 62 | F | 19.4 | 57.5 | 86 | 1 | 1 | 1 | 1 | 2 |
| 31 | M | 10.1 | 32.7 | 88 | 2 | 2 | 2 | 0 | 2 |
| 56 | F | 13.2 | 38.3 | 89 | 2 | 1 | 2 | 1 | 1 |
| 56 | M | 16.0 | 44.3 | 87 | 1 | 2 | 2 | 1 | 0 |
| 36 | F | 13.8 | 38.4 | 92 | 2 | 2 | 2 | 0 | 1 |
| 66 | F | 16.1 | 45.5 | 85 | 2 | 2 | 2 | 0 | 1 |
| 45 | M | 17.8 | 69.3 | 86 | 0 | 2 | 2 | 1 | 0 |
| 56 | F | 14.9 | 44.4 | 83 | 2 | 2 | 2 | 1 | 1 |
| 29 | M | 18.7 | 53.2 | 91 | 0 | 1 | 2 | 0 | 0 |
| 34 | M | 19.4 | 53.5 | 85 | 1 | 1 | 1 | 0 | 1 |
| 32 | M | 22.3 | 64.3 | 90 | 1 | 1 | 2 | 0 | 1 |
| 36 | F | 14.9 | 43.4 | 91 | 0 | 2 | 2 | 2 | 2 |
| 44 | M | 15.9 | 44.9 | 84 | 2 | 2 | 2 | 1 | 1 |
| 23 | F | 12.7 | 37.3 | 88 | 2 | 1 | 1 | 0 | 0 |
| 60 | F | 14.6 | 41.9 | 84 | 2 | 1 | 1 | 1 | 1 |
| 33 | F | 15.5 | 44.7 | 86 | 1 | 2 | 2 | 2 | 1 |
| 23 | F | 15.6 | 43.4 | 86 | 1 | 2 | 1 | 1 | 2 |
| 17 | F | 14.7 | 42.2 | 90 | 1 | 2 | 2 | 0 | 0 |
| 62 | M | 15.5 | 42.8 | 87 | 1 | 2 | 1 | 0 | 0 |
| 38 | M | 17.7 | 50.5 | 81 | 2 | 0 | 1 | 2 | 1 |
| 29 | M | 16.4 | 49.4 | 84 | 2 | 2 | 1 | 0 | 1 |
| 46 | F | 17.9 | 51.4 | 83 | 0 | 2 | 1 | 1 | 2 |
| 44 | M | 16.0 | 45.2 | 86 | 1 | 2 | 2 | 0 | 1 |
| 40 | F | 16.9 | 47 | 87 | 2 | 1 | 2 | 1 | 1 |
| 36 | F | 16.4 | 47 | 85 | 1 | 2 | 1 | 0 | 0 |
| 40 | F | 17.0 | 49 | - | 1 | 2 | 2 | 1 | 2 |
| 53 | F | 19.5 | 56.6 | 79 | 2 | 0 | 2 | 2 | 0 |
| 30 | F | 13.4 | 38.4 | 89 | 2 | 2 | 2 | 0 | 1 |
| 27 | F | 12.9 | 38.4 | 85 | 2 | 2 | 2 | 1 | 2 |
| 41 | F | 16.9 | 43.1 | 85 | 2 | 1 | 2 | 0 | 2 |

**Table S8. Regression analysis between selection candidate haplotypes and Hb phenotype**

| Predictor variables | p* | Effect size | Standard deviation |
|---|---|---|---|
| Sex (0 = Male, 1 = Female) | 0.44 | -5.27 | 6.78 |
| Male Age (0 for Females) | 0.41 | 0.13 | 0.16 |
| Female Age (0 for Males) | 0.74 | 0.05 | 0.15 |
| *EGLN1* | 0.002 | -20.06 | 5.39 |
| *PPARA* | 0.0009 | -14.88 | 4.42 |
| *HMOX2* | 0.49 | -4.85 | 6.97 |
| *CYP17A1* | 0.74 | 1.66 | 4.87 |
| *PTEN* | 0.87 | -0.71 | 4.47 |

* For the regression model: $F_{(2, 27)} = 9.78$, $p < 0.0006$. Only variables with $p < 0.01$ individually were included in the final model.

**Table S9. Regression analysis between selection candidate haplotypes and Hb phenotype with *EGLN1* and *PPARA* haplotypes combined**

| Predictor variable | p* | Effect size | Standard deviation |
|---|---|---|---|
| Sex | 0.48 | -5.27 | 6.78 |
| Male Age | 0.47 | 0.11 | 0.16 |
| Female Age | 0.59 | 0.08 | 0.14 |
| *EGLN1+PPARA*[†] | 0.0002 | -16.76 | 3.85 |
| *HMOX2* | 0.41 | -5.67 | 6.83 |
| *CYP17A1* | 0.61 | 2.47 | 4.72 |
| *PTEN* | 0.82 | -1.01 | 4.43 |

* For the regression model: $F_{(1, 28)} = 18.93$, $p < 0.0002$. Only variables with $p < 0.01$ individually were included in the final model.
[†] Total number of putatively selected *EGLN1* and *PPARA* alleles in an individual (0 – 4).

**Table S10. Regression analysis between selection candidate haplotypes and percent $SaO_2$ phenotype**

| Predictor variables | p* | Effect size | Standard deviation |
|---|---|---|---|
| Sex (0 = Male, 1 = Female) | 0.39 | 0.78 | 0.88 |
| Male Age (0 for Females) | 0.36 | -0.02 | 0.020 |
| Female Age (0 for Males) | 0.60 | 0.010 | 0.02 |
| Hb | 0.0007 | -0.08 | 0.020 |
| *EGLN1* | 0.0008 | -2.15 | 0.56 |
| *PPARA* | 0.25 | -1.01 | 0.86 |
| *HMOX2* | 0.07 | 1.61 | 0.83 |
| *CYP17A1* | 0.04 | -1.28 | 0.58 |
| *PTEN* | 0.76 | -0.18 | 0.59 |

* For the regression model: $F_{(3,25)}$ 10.70, $p < 0.0001$. Only variables with $p < 0.01$ individually were included in the final model.

**Table S11. 200kb genomic regions identified in the top two percent of the XP-EHH selection scan, not excluding regions identified in other populations**

| Genes in XP-EHH regions | Chromosome | 200kb Region | XP-EHH value | P value |
|---|---|---|---|---|
| *PRDM2* | Chr1 | 70 | 0.60 | 0.0196 |
| *PAX7* | Chr1 | 94 | 0.65 | 0.0132 |
| *CLIC4,RUNX3* | Chr1 | 125 | 0.65 | 0.0127 |
| No genes in this region | Chr1 | 153 | 0.63 | 0.0163 |
| *DAB1* | Chr1 | 288 | 0.69 | 0.0082 |
| No genes in this region | Chr1 | 403 | 0.66 | 0.0114 |
| *OR10J5,OR10J1* | Chr1 | 788 | 0.69 | 0.0082 |
| No genes in this region | Chr1 | 1060 | 0.80 | 0.0023 |
| *USH2A* | Chr1 | 1070 | 0.65 | 0.0131 |
| *CNIH3* | Chr1 | 1114 | 0.64 | 0.0146 |
| No genes in this region | Chr1 | 1117 | 0.69 | 0.0084 |
| *ITPKB,C1orf95* | Chr1 | 1124 | 0.72 | 0.0064 |
| *TRIM67,C1orf124,GNPAT,C1orf131,* ***EGLN1****,* | | | | |
| *EXOC8* | Chr1 | 1147 | 0.86 | 0.0012 |
| *TSNAX,**EGLN1*** | Chr1 | 1148 | 0.96 | 0.0002 |
| *DISC1* | Chr1 | 1149 | 0.92 | 0.0002 |
| *AKT3* | Chr1 | 1210 | 0.65 | 0.0134 |
| *KLF6* | Chr10 | 19 | 0.79 | 0.0028 |
| *DHTKD1,NUDT5,SEC61A2,CDC123* | Chr10 | 61 | 0.61 | 0.0185 |
| *SEPHS1,BEND7* | Chr10 | 67 | 0.74 | 0.0054 |
| *RSU1,CUBN* | Chr10 | 84 | 0.65 | 0.0125 |
| No genes in this region | Chr10 | 147 | 0.66 | 0.0110 |
| *MBL2* | Chr10 | 271 | 0.60 | 0.0198 |
| No genes in this region | Chr10 | 272 | 0.76 | 0.0039 |
| No genes in this region | Chr10 | 273 | 0.69 | 0.0086 |
| *ANK3* | Chr10 | 308 | 0.73 | 0.0055 |
| *LRRC20,NPFFR1,PPA1,SAR1A* | Chr10 | 358 | 0.64 | 0.0149 |
| *CDH23,C10orf54* | Chr10 | 365 | 0.61 | 0.0183 |
| *RPS24,POLR3A* | Chr10 | 397 | 0.64 | 0.0148 |
| *CHUK,ERLIN1,CWF19L1,SNORA12,CPN1* | Chr10 | 509 | 0.86 | 0.0011 |
| *BLOC1S2,SCD,PKD2L1,CWF19L1* | Chr10 | 510 | 0.89 | 0.0007 |
| *WNT8B,**HIF1AN**,SEC31B,NDUFB8* | Chr10 | 511 | 0.81 | 0.0017 |
| *PAX2* | Chr10 | 512 | 0.60 | 0.0199 |
| *NCRNA00081,PDCD4,SHOC2* | Chr10 | 563 | 0.78 | 0.0035 |
| No genes in this region | Chr10 | 609 | 0.62 | 0.0170 |
| *ADAM12* | Chr10 | 639 | 0.62 | 0.0169 |
| *FRG2B,SYCE1,DUX4,**CYP2E1**,LOC728410,* *LOC653544,LOC653545,LOC653543* | Chr10 | 676 | 0.71 | 0.0074 |
| *HBD,OR51B2,OR51B5,HBG2,HBBP1,* *OR51M1,HBE1,HBG1,HBB,OR51B6,* *OR51B4* | Chr11 | 26 | 0.63 | 0.0158 |

| | | | | |
|---|---|---|---|---|
| *TRIM34,UBQLNL,TRIM6-TRIM34,OR52H1,OR52B6,TRIM6,OR52D1, UBQLN3,OR51I1,OR51Q1,OR51I2* | Chr11 | 27 | 0.64 | 0.0150 |
| *TEAD1,RASSF10* | Chr11 | 64 | 0.62 | 0.0169 |
| No genes in this region | Chr11 | 65 | 0.66 | 0.0109 |
| *TMEM86A,SPTY2D1,PTPN5,IGSF22* | Chr11 | 93 | 0.66 | 0.0109 |
| *PRMT3,SLC6A5* | Chr11 | 102 | 0.84 | 0.0013 |
| *NELL1,SLC6A5* | Chr11 | 103 | 0.67 | 0.0100 |
| *LUZP2* | Chr11 | 124 | 0.71 | 0.0075 |
| *ALX4,EXT2* | Chr11 | 221 | 0.62 | 0.0178 |
| *APLNR,LRRC55* | Chr11 | 283 | 0.66 | 0.0112 |
| *KCNE3,POLD3* | Chr11 | 369 | 0.74 | 0.0052 |
| *RNF169,CHRDL2,POLD3* | Chr11 | 370 | 0.61 | 0.0189 |
| *C11orf87* | Chr11 | 544 | 0.67 | 0.0102 |
| *SC5DL* | Chr11 | 603 | 0.79 | 0.0032 |
| *BLID,LOC399959* | Chr11 | 607 | 0.76 | 0.0042 |
| *OPCML* | Chr11 | 659 | 0.62 | 0.0174 |
| *CCDC77,SLC6A13,JARID1A* | Chr12 | 1 | 0.74 | 0.0052 |
| *CCDC77,NINJ2,B4GALNT3* | Chr12 | 2 | 0.64 | 0.0140 |
| *CACNA1C* | Chr12 | 10 | 0.73 | 0.0060 |
| *CACNA1C* | Chr12 | 11 | 0.66 | 0.0119 |
| *BCL2L14,LOH12CR2,LRP6,MANSC1* | Chr12 | 61 | 0.66 | 0.0118 |
| No genes in this region | Chr12 | 143 | 0.67 | 0.0101 |
| *SYT10* | Chr12 | 167 | 0.82 | 0.0015 |
| *OR10AD1,PFKM,C12orf68,ASB8* | Chr12 | 234 | 0.73 | 0.0058 |
| *DYRK2* | Chr12 | 331 | 0.80 | 0.0025 |
| *FGD6,VEZT* | Chr12 | 470 | 0.69 | 0.0079 |
| *NUAK1* | Chr12 | 524 | 0.62 | 0.0179 |
| No genes in this region | Chr12 | 562 | 0.65 | 0.0136 |
| No genes in this region | Chr12 | 587 | 0.67 | 0.0105 |
| *CCDC64* | Chr12 | 594 | 0.75 | 0.0048 |
| *TMEM120B,LOC338799,HPD,SETD1B, RHOF* | Chr12 | 603 | 0.63 | 0.0156 |
| *PSMD9,WDR66,BCL7A* | Chr12 | 604 | 0.72 | 0.0063 |
| No genes in this region | Chr12 | 631 | 0.62 | 0.0179 |
| No genes in this region | Chr13 | 429 | 0.61 | 0.0181 |
| *NALCN,ITGBL1* | Chr13 | 504 | 0.61 | 0.0195 |
| *ING1,CARKD,RAB20,CARS2* | Chr13 | 550 | 0.78 | 0.0034 |
| *ANKRD10* | Chr13 | 551 | 0.63 | 0.0165 |
| *MAX* | Chr14 | 323 | 0.61 | 0.0194 |
| *RAD51L1* | Chr14 | 340 | 0.68 | 0.0089 |
| *WDR21A,DPF3,RBM25,ZFYVE1* | Chr14 | 362 | 0.62 | 0.0180 |
| *TMEM63C,KIAA1737,ZDHHC22* | Chr14 | 383 | 0.66 | 0.0120 |
| *TTC8* | Chr14 | 442 | 0.63 | 0.0157 |
| *FSIP1,THBS1* | Chr15 | 188 | 0.64 | 0.0138 |
| No genes in this region | Chr15 | 219 | 0.75 | 0.0044 |
| *KIAA0256,SHC4* | Chr15 | 235 | 0.63 | 0.0159 |
| *WDR72* | Chr15 | 257 | 0.62 | 0.0171 |

| | | | | |
|---|---|---|---|---|
| *RFX7* | Chr15 | 271 | 0.77 | 0.0036 |
| *MCTP2* | Chr15 | 464 | 0.61 | 0.0183 |
| No genes in this region | Chr15 | 465 | 0.69 | 0.0086 |
| *TELO2,C16orf38,UNKL,LOC283951, TMEM204,IFT140,CLCN7* | Chr16 | 7 | 0.90 | 0.0006 |
| *HN1L,IGFALS,SPSB3,EME2,NUBP2, MRPS34, NME3,IFT140,HAGH,MAPK8IP3, CRAMP1L* | Chr16 | 8 | 1.04 | 0.0001 |
| *NDE1,MYH11,KIAA0430* | Chr16 | 78 | 0.77 | 0.0035 |
| *GSG1L* | Chr16 | 139 | 0.74 | 0.0055 |
| No genes in this region | Chr16 | 266 | 0.74 | 0.0049 |
| No genes in this region | Chr16 | 311 | 0.76 | 0.0039 |
| *WWOX* | Chr16 | 385 | 0.69 | 0.0080 |
| *CDH13* | Chr16 | 408 | 0.62 | 0.0173 |
| *MLYCD,HSBP1,OSGIN1,NECAB2* | Chr16 | 412 | 0.65 | 0.0126 |
| *MAP1LC3B,ZCCHC14,FBXO31* | Chr16 | 429 | 0.65 | 0.0124 |
| *DHX40P,HEATR6,CA4* | Chr17 | 277 | 0.74 | 0.0051 |
| *WIPI1,ARSG* | Chr17 | 319 | 0.66 | 0.0115 |
| *ACTG1,BAHCC1,FSCN2,C17orf70,NPLOC4* | Chr17 | 385 | 0.66 | 0.0112 |
| No genes in this region | Chr18 | 7 | 0.63 | 0.0153 |
| *LAMA1,ARHGAP28* | Chr18 | 34 | 0.67 | 0.0097 |
| No genes in this region | Chr18 | 54 | 0.62 | 0.0173 |
| No genes in this region | Chr18 | 56 | 0.87 | 0.0009 |
| *TCF4* | Chr18 | 255 | 0.74 | 0.0050 |
| *DSEL* | Chr18 | 316 | 0.70 | 0.0076 |
| No genes in this region | Chr18 | 348 | 0.71 | 0.0071 |
| *BRUNOL5,NFIC* | Chr19 | 16 | 0.75 | 0.0043 |
| *C19orf29,FZR1,TBXA2R,C19orf71,GIPC3, DOHH,HMG20B,NFIC,LOC284422, C19orf28, PIP5K1C* | Chr19 | 17 | 0.72 | 0.0065 |
| *FBN3,LASS4,CCL25* | Chr19 | 40 | 0.67 | 0.0104 |
| ***ANGPTL4**,KANK3,RPS28,MARCH2, NDUFA7, LASS4,CD320,RAB11B* | Chr19 | 41 | 0.69 | 0.0083 |
| *RHOB* | Chr2 | 102 | 0.64 | 0.0139 |
| *ALK* | Chr2 | 149 | 0.64 | 0.0147 |
| *XDH* | Chr2 | 157 | 0.80 | 0.0024 |
| *SRD5A2* | Chr2 | 158 | 0.74 | 0.0049 |
| *BIRC6* | Chr2 | 162 | 0.65 | 0.0122 |
| *BIRC6,TTC27* | Chr2 | 163 | 0.72 | 0.0062 |
| *SLC8A1* | Chr2 | 202 | 0.64 | 0.0146 |
| No genes in this region | Chr2 | 215 | 0.79 | 0.0033 |
| *PRKCE* | Chr2 | 229 | 0.67 | 0.0096 |
| *PRKCE,**EPAS1*** | Chr2 | 231 | 0.82 | 0.0015 |
| *ATP6V1E2,EPAS1* | Chr2 | 232 | 0.67 | 0.0103 |
| *C2orf61,CALM2* | Chr2 | 236 | 0.70 | 0.0079 |
| *FSHR* | Chr2 | 246 | 0.64 | 0.0144 |

| | | | | |
|---|---|---|---|---|
| *AFTPH,SERTAD2* | Chr2 | 323 | 0.63 | 0.0162 |
| *MCEE,PAIP2B,MPHOSPH10* | Chr2 | 356 | 0.81 | 0.0018 |
| *ZNF638,DYSF* | Chr2 | 357 | 0.81 | 0.0018 |
| *LOC150568* | Chr2 | 522 | 0.72 | 0.0066 |
| No genes in this region | Chr2 | 646 | 0.63 | 0.0154 |
| No genes in this region | Chr2 | 686 | 0.61 | 0.0184 |
| *GTDC1,ZEB2* | Chr2 | 724 | 0.65 | 0.0133 |
| *ARL5A,NEB* | Chr2 | 761 | 0.68 | 0.0096 |
| *CACNB4* | Chr2 | 762 | 0.72 | 0.0067 |
| *XIRP2* | Chr2 | 837 | 0.67 | 0.0098 |
| *XIRP2* | Chr2 | 838 | 0.68 | 0.0087 |
| *DNAJC10* | Chr2 | 916 | 0.80 | 0.0022 |
| *CYP20A1,ABI2* | Chr2 | 1019 | 0.65 | 0.0130 |
| *PARD3B* | Chr2 | 1028 | 0.61 | 0.0191 |
| *PARD3B* | Chr2 | 1029 | 0.80 | 0.0022 |
| *LOC643905,MYEOV2,NDUFA10,OR6B3, OR6B2,OTOS* | Chr2 | 1203 | 0.66 | 0.0117 |
| No genes in this region | Chr20 | 21 | 0.63 | 0.0163 |
| *PLCB4* | Chr20 | 46 | 0.73 | 0.0061 |
| *B4GALT5,PTGIS* | Chr20 | 238 | 0.79 | 0.0029 |
| *SPATA2,SLC9A8,RNF114* | Chr20 | 239 | 0.72 | 0.0068 |
| *TMEM189,TMEM189-UBE2V1,SNAI1,UBE2V1,RNF114* | Chr20 | 240 | 0.67 | 0.0099 |
| *ERG* | Chr21 | 193 | 0.79 | 0.0029 |
| *DGCR11,TSSK2,CLTCL1,DGCR2,SLC25A1, DGCR14,GSC2* | Chr22 | 87 | 0.63 | 0.0159 |
| *CLTCL1,HIRA* | Chr22 | 88 | 0.79 | 0.0031 |
| *KIAA1644* | Chr22 | 215 | 0.73 | 0.0056 |
| *CELSR1* | Chr22 | 226 | 0.92 | 0.0003 |
| *EDEM1* | Chr3 | 26 | 0.67 | 0.0108 |
| No genes in this region | Chr3 | 29 | 0.70 | 0.0077 |
| No genes in this region | Chr3 | 31 | 0.61 | 0.0187 |
| *SRGAP3* | Chr3 | 45 | 0.67 | 0.0106 |
| *THUMPD3,SRGAP3* | Chr3 | 46 | 0.68 | 0.0095 |
| *SETD5,THUMPD3,LHFPL4* | Chr3 | 47 | 0.64 | 0.0152 |
| *VGLL4* | Chr3 | 58 | 0.81 | 0.0019 |
| *SATB1* | Chr3 | 92 | 0.88 | 0.0008 |
| No genes in this region | Chr3 | 93 | 0.92 | 0.0004 |
| *CADPS* | Chr3 | 313 | 0.64 | 0.0141 |
| *ADAMTS9* | Chr3 | 323 | 0.68 | 0.0093 |
| *RYBP* | Chr3 | 362 | 0.63 | 0.0156 |
| *KIAA1407,QTRTD1,DRD3* | Chr3 | 576 | 0.68 | 0.0088 |
| *ADPRH,TMEM39A,KTELC1,C3orf1,CD80, PLA1A,CDGAP* | Chr3 | 603 | 0.76 | 0.0038 |
| *H1FOO,IFT122,RHO,RPL32P3,PLXND1, C3orf25,MBD4* | Chr3 | 653 | 0.62 | 0.0166 |
| *TRIM42* | Chr3 | 709 | 0.64 | 0.0149 |
| No genes in this region | Chr3 | 726 | 0.64 | 0.0143 |
| *VEPH1,PTX3,C3orf55* | Chr3 | 793 | 0.67 | 0.0107 |
| *DGKG,ETV5* | Chr3 | 936 | 0.80 | 0.0021 |

| | | | | |
|---|---|---|---|---|
| *RTP1,RPL39L,ST6GAL1* | Chr3 | 941 | 0.65 | 0.0126 |
| No genes in this region | Chr4 | 150 | 0.88 | 0.0008 |
| *PCDH7* | Chr4 | 152 | 0.64 | 0.0137 |
| No genes in this region | Chr4 | 158 | 0.61 | 0.0182 |
| *TBC1D1* | Chr4 | 189 | 0.67 | 0.0099 |
| *SLC4A4* | Chr4 | 362 | 0.64 | 0.0145 |
| *DMP1,MEPE,IBSP* | Chr4 | 444 | 0.71 | 0.0069 |
| *UNC5C,BMPR1B* | Chr4 | 481 | 0.66 | 0.0111 |
| *CENPE,BDH2,NHEDC2* | Chr4 | 521 | 0.79 | 0.0030 |
| No genes in this region | Chr4 | 522 | 0.62 | 0.0166 |
| No genes in this region | Chr4 | 559 | 0.64 | 0.0153 |
| ***CAMK2D**,ANK2* | Chr4 | 572 | 0.68 | 0.0092 |
| *TMEM155,CCNA2,EXOSC9,ANXA5, LOC100192379,BBS7* | Chr4 | 614 | 0.76 | 0.0040 |
| *TRPC3,BBS7* | Chr4 | 615 | 0.79 | 0.0027 |
| No genes in this region | Chr4 | 637 | 0.68 | 0.0091 |
| No genes in this region | Chr4 | 742 | 0.68 | 0.0089 |
| ***EDNRA**,TMEM184C,LOC90826* | Chr4 | 743 | 0.70 | 0.0078 |
| No genes in this region | Chr4 | 817 | 0.62 | 0.0175 |
| *C4orf43,MARCH1,TKTL2* | Chr4 | 823 | 0.62 | 0.0172 |
| No genes in this region | Chr4 | 825 | 0.62 | 0.0177 |
| No genes in this region | Chr5 | 16 | 0.83 | 0.0014 |
| No genes in this region | Chr5 | 23 | 0.61 | 0.0193 |
| No genes in this region | Chr5 | 95 | 0.64 | 0.0136 |
| *RNASEN,C5orf22* | Chr5 | 157 | 0.65 | 0.0129 |
| *THBS4,SERINC5* | Chr5 | 397 | 0.71 | 0.0072 |
| *MCC* | Chr5 | 562 | 0.63 | 0.0164 |
| *IL12B,UBLCP1* | Chr5 | 793 | 0.66 | 0.0116 |
| No genes in this region | Chr5 | 823 | 0.71 | 0.0070 |
| *JARID2* | Chr6 | 77 | 0.64 | 0.0139 |
| *HLA-DRB6,HLA-DQA1,HLA-DRB1, HLA-DRB5,HLA-DQB1* | Chr6 | 163 | 0.65 | 0.0123 |
| *IL17A,PKHD1* | Chr6 | 260 | 0.64 | 0.0142 |
| *PRIM2* | Chr6 | 286 | 0.61 | 0.0189 |
| *PRIM2* | Chr6 | 287 | 0.81 | 0.0020 |
| *B3GAT2,SMAP1* | Chr6 | 358 | 0.75 | 0.0045 |
| No genes in this region | Chr6 | 434 | 0.63 | 0.0161 |
| No genes in this region | Chr6 | 493 | 0.68 | 0.0092 |
| No genes in this region | Chr6 | 494 | 0.69 | 0.0081 |
| *SOBP,SCML4* | Chr6 | 540 | 0.61 | 0.0186 |
| *NKAIN2* | Chr6 | 624 | 0.86 | 0.0012 |
| *AKAP7,EPB41L2* | Chr6 | 657 | 0.75 | 0.0045 |
| No genes in this region | Chr6 | 707 | 0.69 | 0.0085 |
| No genes in this region | Chr6 | 708 | 0.63 | 0.0160 |
| *UTRN* | Chr6 | 724 | 0.65 | 0.0132 |
| *UTRN* | Chr6 | 725 | 0.90 | 0.0005 |
| *SAMD5* | Chr6 | 739 | 0.62 | 0.0168 |
| *ULBP3,PPP1R14C* | Chr6 | 752 | 0.65 | 0.0135 |
| *NOX3* | Chr6 | 779 | 0.65 | 0.0129 |
| *MAD1L1* | Chr7 | 9 | 0.64 | 0.0143 |

| | | | | |
|---|---|---|---|---|
| *CYTH3,EIF2AK1,USP42,PMS2,JTV1* | Chr7 | 30 | 0.61 | 0.0188 |
| *MACC1,ITGB8* | Chr7 | 101 | 0.76 | 0.0042 |
| *ITGB8* | Chr7 | 102 | 0.67 | 0.0106 |
| *BBS9* | Chr7 | 167 | 0.60 | 0.0197 |
| No genes in this region | Chr7 | 172 | 0.65 | 0.0128 |
| *HERPUD2* | Chr7 | 178 | 0.66 | 0.0116 |
| *ELMO1* | Chr7 | 186 | 0.71 | 0.0073 |
| *EPDR1,TXNDC3,SFRP4* | Chr7 | 189 | 0.72 | 0.0069 |
| *ADCY1* | Chr7 | 227 | 0.60 | 0.0200 |
| *PCLO* | Chr7 | 411 | 0.82 | 0.0016 |
| *PCLO* | Chr7 | 412 | 0.79 | 0.0032 |
| *SSPO,KRBA1,ZNF862,ZNF467* | Chr7 | 745 | 0.63 | 0.0155 |
| *C7orf29,REPIN1,RARRES2,LRRC61, GIMAP8, ZNF775* | Chr7 | 748 | 0.62 | 0.0167 |
| *CSMD1* | Chr8 | 18 | 0.60 | 0.0196 |
| No genes in this region | Chr8 | 79 | 0.66 | 0.0113 |
| No genes in this region | Chr8 | 104 | 0.62 | 0.0176 |
| *INTS9,EXTL3* | Chr8 | 143 | 0.71 | 0.0075 |
| *INTS9,KIF13B,HMBOX1* | Chr8 | 144 | 0.68 | 0.0090 |
| *LOC286135* | Chr8 | 149 | 0.76 | 0.0041 |
| No genes in this region | Chr8 | 246 | 0.90 | 0.0005 |
| No genes in this region | Chr8 | 247 | 0.86 | 0.0010 |
| *EFCAB1* | Chr8 | 248 | 0.76 | 0.0037 |
| No genes in this region | Chr8 | 327 | 0.66 | 0.0121 |
| *MYBL1,VCPIP1,C8orf44,SGK3* | Chr8 | 338 | 0.61 | 0.0190 |
| *ARFGEF1,CSPP1* | Chr8 | 341 | 0.75 | 0.0046 |
| *ARFGEF1,CPA6* | Chr8 | 342 | 0.79 | 0.0026 |
| *CPA6* | Chr8 | 344 | 0.73 | 0.0057 |
| No genes in this region | Chr8 | 384 | 0.67 | 0.0102 |
| No genes in this region | Chr8 | 639 | 0.73 | 0.0059 |
| No genes in this region | Chr8 | 640 | 0.66 | 0.0119 |
| No genes in this region | Chr8 | 646 | 0.61 | 0.0192 |
| No genes in this region | Chr9 | 124 | 0.73 | 0.0059 |
| *TMEM215,TAF1L* | Chr9 | 163 | 0.75 | 0.0047 |
| *MAMDC2* | Chr9 | 359 | 0.80 | 0.0025 |
| *KLF9,SMC5,MAMDC2* | Chr9 | 360 | 0.73 | 0.0062 |
| *SLC28A3* | Chr9 | 430 | 0.62 | 0.0176 |
| *IARS,OGN,SNORA84,NOL8,CENPP* | Chr9 | 470 | 0.64 | 0.0151 |
| *BARX1* | Chr9 | 478 | 0.71 | 0.0072 |
| *PTPDC1* | Chr9 | 479 | 0.74 | 0.0053 |
| No genes in this region | Chr9 | 542 | 0.61 | 0.0193 |
| *HSDL2,KIAA1958* | Chr9 | 571 | 0.66 | 0.0122 |
| *ZNF618* | Chr9 | 578 | 0.68 | 0.0094 |
| *ASTN2,SNORA70C* | Chr9 | 594 | 0.72 | 0.0065 |
| No genes in this region | Chr9 | 598 | 0.61 | 0.0186 |

**Table S12. 200kb genomic regions identified in the top two percent of the iHS selection scan, not excluding regions identified in other populations**

| Genes in iHS regions | Chromosome | 200kb Region | P value |
|---|---|---|---|
| *XPR1* | Chr1 | 894 | 0.0135 |
| *ATXN7L2,CYB561D1,AMIGO1,GPR61,AMPD2,GNAI3,SYPL2,GNAT2* | Chr1 | 549 | 0.0106 |
| No genes in this region | Chr1 | 964 | 0.0159 |
| *TCEA3,ZNF436,HNRNPR,C1orf213* | Chr1 | 117 | 0.0141 |
| No genes in this region | Chr1 | 959 | 0.0122 |
| *CACNA1E* | Chr1 | 898 | 0.0095 |
| No genes in this region | Chr1 | 1117 | 0.0135 |
| *PTPRF,ST3GAL3,JMJD2A* | Chr1 | 219 | 0.0176 |
| *DISC1* | Chr1 | 1149 | 0.0142 |
| *SGIP1,PDE4B* | Chr1 | 333 | 0.0097 |
| *PFKFB2,C4BPB,C4BPA,FCAMR,YOD1,C1orf116* | Chr1 | 1026 | 0.0115 |
| No genes in this region | Chr1 | 151 | 0.0094 |
| *FAM163A,TOR1AIP2,IFRG15,TOR1AIP1,CEP350* | Chr1 | 890 | 0.0035 |
| *TSNAX,**EGLN1*** | Chr1 | 1148 | 0.0098 |
| No genes in this region | Chr1 | 963 | 0.0063 |
| *AGBL4,ELAVL4* | Chr1 | 251 | 0.0083 |
| No genes in this region | Chr1 | 344 | 0.0093 |
| *FCER1A,DARC,CADM3,OR10J3* | Chr1 | 787 | 0.0091 |
| *GBAP,PKLR,C1orf104,SCAMP3,MTX1,KRTCAP2,TRIM46,GBA,HCN3,MUC1,FDPS,C1orf2,CLK2,THBS3,RUSC1,ASH1L* | Chr1 | 767 | 0.0200 |
| No genes in this region | Chr10 | 295 | 0.0197 |
| No genes in this region | Chr10 | 286 | 0.0187 |
| *ATRNL1* | Chr10 | 584 | 0.0009 |
| *WNT8B,**HIF1AN**,SEC31B,NDUFB8* | Chr10 | 511 | 0.0020 |
| *ENTPD1,CCNJ,LOC100127889,CC2D2B* | Chr10 | 488 | 0.0042 |

| | | | |
|---|---|---|---|
| *TRIM8,**CYP17A1**,ARL3,SFXN2,C10orf26* | Chr10 | 522 | 0.0071 |
| *MYOF* | Chr10 | 475 | 0.0143 |
| No genes in this region | Chr10 | 294 | 0.0090 |
| No genes in this region | Chr10 | 550 | 0.0008 |
| *FRMPD2* | Chr10 | 245 | 0.0026 |
| ***PTEN**,KILLIN* | Chr10 | 448 | 0.0072 |
| *CUEDC2,C10orf95,GBF1,NFKB2,PSD, FBXL15* | Chr10 | 520 | 0.0159 |
| *C10orf32,AS3MT,CNNM2* | Chr10 | 523 | 0.0164 |
| *PAX2* | Chr10 | 512 | 0.0001 |
| *ZNF37A,LOC100129055* | Chr10 | 192 | 0.0070 |
| *ATRNL1* | Chr10 | 585 | 0.0051 |
| *C10orf125,PRAP1,CALY,ECHS1,CYP2E1, SPRN,PAOX,MTG1,LOC619207* | Chr10 | 675 | 0.0148 |
| *PCDH15* | Chr10 | 278 | 0.0002 |
| *KIAA1217* | Chr10 | 121 | 0.0039 |
| *BLOC1S2,SCD,PKD2L1,CWF19L1* | Chr10 | 510 | 0.0075 |
| *ATRNL1* | Chr10 | 586 | 0.0185 |
| No genes in this region | Chr11 | 486 | 0.0166 |
| *APLNR,LRRC55* | Chr11 | 283 | 0.0183 |
| *TTC17,API5* | Chr11 | 216 | 0.0149 |
| No genes in this region | Chr11 | 126 | 0.0016 |
| *QSER1,DEPDC7,PRRG4* | Chr11 | 164 | 0.0089 |
| *P2RX3,PRG3,SLC43A3,RTN4RL2,SSRP1, TNKS1BP1,PRG2* | Chr11 | 284 | 0.0101 |
| *NOX4* | Chr11 | 444 | 0.0170 |
| *GYS2,LDHB* | Chr12 | 108 | 0.0081 |
| No genes in this region | Chr12 | 364 | 0.0104 |
| *KITLG* | Chr12 | 437 | 0.0006 |
| *SOX5* | Chr12 | 122 | 0.0108 |
| *BCL2L14,LOH12CR2,LRP6,MANSC1* | Chr12 | 61 | 0.0028 |
| *SOX5,C12orf67* | Chr12 | 123 | 0.0198 |
| *ETNK1* | Chr12 | 113 | 0.0039 |
| *ANKS1B* | Chr12 | 489 | 0.0173 |
| *PZP,LOC642846* | Chr12 | 46 | 0.0107 |
| No genes in this region | Chr12 | 362 | 0.0085 |
| No genes in this region | Chr13 | 336 | 0.0120 |
| No genes in this region | Chr13 | 371 | 0.0080 |
| No genes in this region | Chr13 | 325 | 0.0048 |

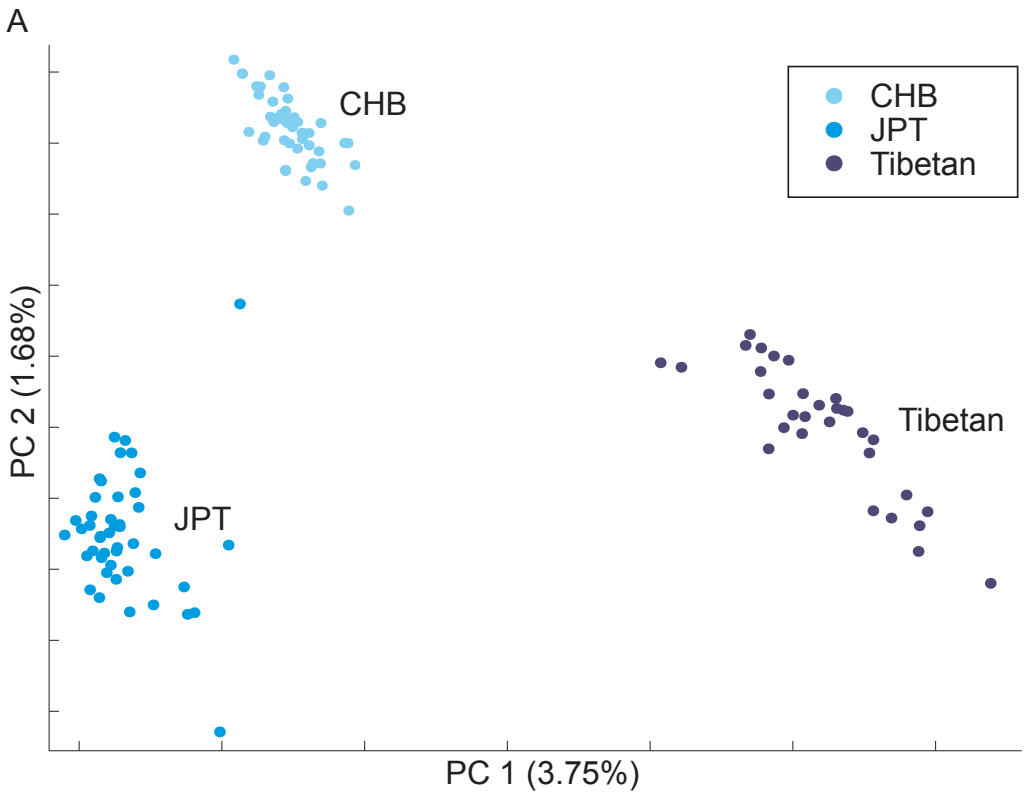
| | | | |
|---|---|---|---|
| *SLITRK6* | Chr13 | 426 | 0.0147 |
| No genes in this region | Chr13 | 357 | 0.0184 |
| *KLF12* | Chr13 | 367 | 0.0019 |
| No genes in this region | Chr14 | 434 | 0.0043 |
| *SYT16,FLJ43390* | Chr14 | 308 | 0.0160 |
| No genes in this region | Chr14 | 221 | 0.0110 |
| *ERH,SLC39A9,GALNTL1* | Chr14 | 344 | 0.0151 |
| *RTN1,GPR135,C14orf149,C14orf100* | Chr14 | 295 | 0.0089 |
| *SV2B* | Chr15 | 447 | 0.0161 |
| *LBXCOR1,MAP2K5* | Chr15 | 329 | 0.0018 |
| *SH2D7,CIB2,TBC1D2B* | Chr15 | 380 | 0.0139 |
| *RGMA,CHD2* | Chr15 | 456 | 0.0036 |
| *PRDM7,GAS8,DBNDD1,C16orf3* | Chr16 | 443 | 0.0060 |
| No genes in this region | Chr16 | 321 | 0.0038 |
| *NMRAL1,CORO7,DNAJA3,C16orf5,* ***HMOX2****,FAM100A* | Chr16 | 22 | 0.0013 |
| No genes in this region | Chr16 | 127 | 0.0176 |
| *BLMH,TMIGD1,CPD* | Chr17 | 128 | 0.0052 |
| *DBF4B,CCDC43,ADAM11* | Chr17 | 200 | 0.0111 |
| *GHDC,STAT5B,STAT5A,STAT3* | Chr17 | 188 | 0.0137 |
| No genes in this region | Chr17 | 336 | 0.0194 |
| *BLMH,EFCAB5,CCDC55,SLC6A4* | Chr17 | 127 | 0.0186 |
| *TAF4B,PSMA8* | Chr18 | 110 | 0.0034 |
| *FAM59A,MEP1B* | Chr18 | 140 | 0.0191 |
| No genes in this region | Chr18 | 73 | 0.0084 |
| No genes in this region | Chr18 | 332 | 0.0133 |
| No genes in this region | Chr18 | 89 | 0.0150 |
| *TCF4* | Chr18 | 256 | 0.0060 |
| *GALNT13* | Chr2 | 772 | 0.0073 |
| *CYP20A1,ABI2* | Chr2 | 1019 | 0.0056 |
| No genes in this region | Chr2 | 86 | 0.0012 |
| *GPBAR1,C2orf62,SLC11A1,TMBIM1,* *PNKD,CTDSP1,ARPC2,AAMP,VIL1* | Chr2 | 1094 | 0.0056 |
| *CCDC148* | Chr2 | 794 | 0.0195 |
| No genes in this region | Chr2 | 627 | 0.0124 |
| *MKI67IP,TSN* | Chr2 | 611 | 0.0145 |
| No genes in this region | Chr2 | 112 | 0.0061 |
| *ANKRD44* | Chr2 | 988 | 0.0085 |

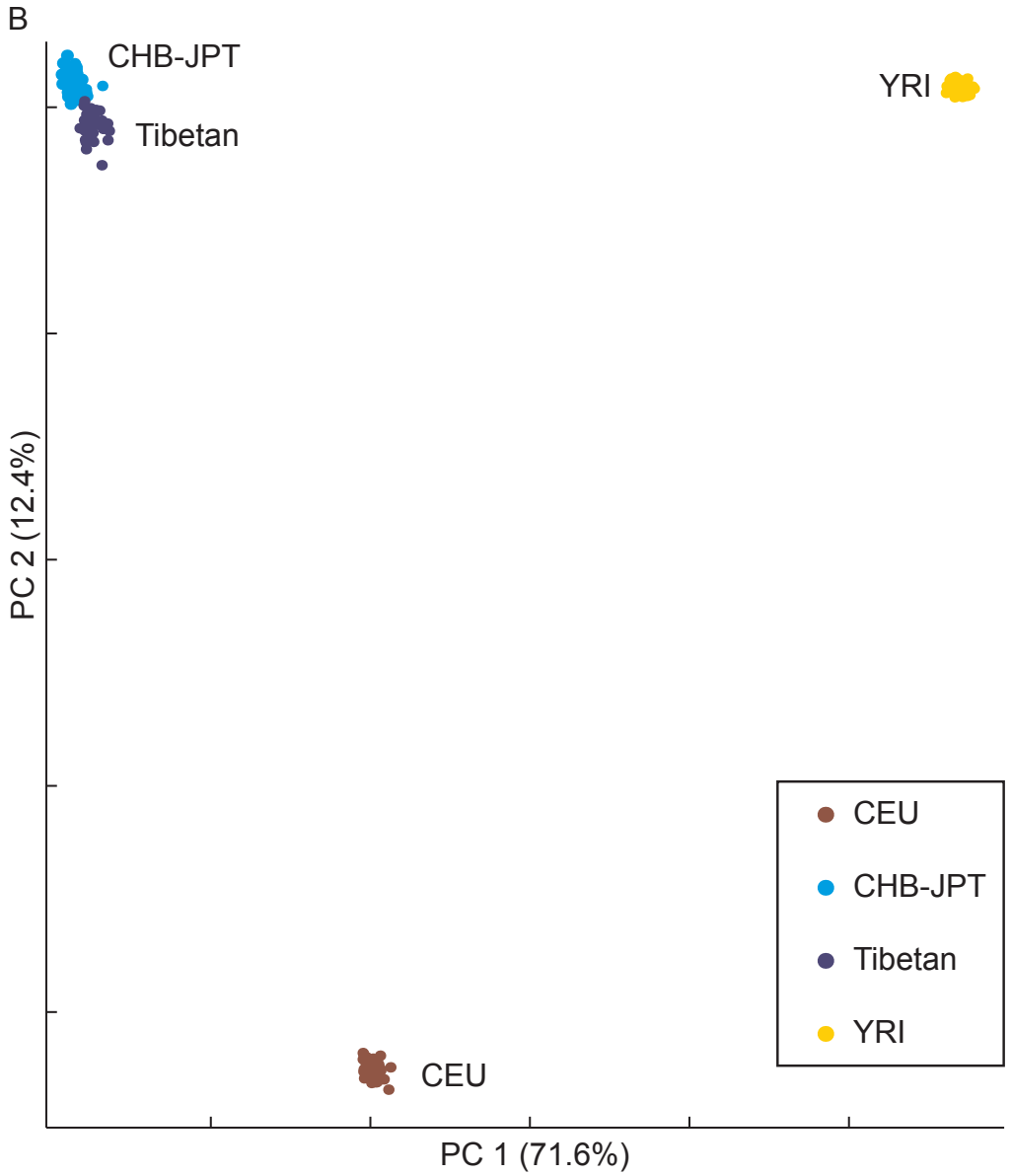
| | | | |
|---|---|---|---|
| No genes in this region | Chr2 | 109 | 0.0079 |
| *ORMDL1,PMS1,ANKAR,OSGEPL1, ASNSD1* | Chr2 | 951 | 0.0118 |
| *LYPD6* | Chr2 | 749 | 0.0169 |
| No genes in this region | Chr2 | 977 | 0.0177 |
| *MBOAT2* | Chr2 | 45 | 0.0010 |
| *EFEMP1* | Chr2 | 280 | 0.0154 |
| *XIRP2* | Chr2 | 837 | 0.0077 |
| No genes in this region | Chr2 | 632 | 0.0119 |
| *VSNL1,SMC6,GEN1* | Chr2 | 88 | 0.0023 |
| *RHOB* | Chr2 | 102 | 0.0181 |
| No genes in this region | Chr2 | 286 | 0.0096 |
| *SULT1C3,SULT1C2,SULT1C4* | Chr2 | 541 | 0.0105 |
| *ARL6IP6,FMNL2,PRPF40A* | Chr2 | 766 | 0.0130 |
| *EPC2,KIF5C* | Chr2 | 746 | 0.0174 |
| *ASAP2* | Chr2 | 46 | 0.0078 |
| No genes in this region | Chr2 | 84 | 0.0180 |
| *RAPH1,ABI2* | Chr2 | 1020 | 0.0117 |
| *FOXA2* | Chr20 | 112 | 0.0099 |
| *PARD6B,ADNP,BCAS4,DPM1* | Chr20 | 244 | 0.0168 |
| *LOC100130264,SLC24A3* | Chr20 | 95 | 0.0068 |
| *FAM83C,EIF6,UQCC,PROCR,MMP24* | Chr20 | 166 | 0.0019 |
| *SCAND1,PHF20,EPB41L1,C20orf152* | Chr20 | 170 | 0.0024 |
| *DHX35,FAM83D* | Chr20 | 185 | 0.0143 |
| *UQCC,GDF5,CEP250,ERGIC3* | Chr20 | 167 | 0.0022 |
| *C21orf33,AGPAT3,TRAPPC10,PWP2* | Chr21 | 221 | 0.0172 |
| *C21orf7,BACH1* | Chr21 | 147 | 0.0128 |
| *C21orf34* | Chr21 | 81 | 0.0041 |
| *APOL4,APOL2,APOL3,APOL1* | Chr22 | 174 | 0.0172 |
| ***PPARA**,C22orf26* | Chr22 | 224 | 0.0092 |
| *PLXNB1,CCDC51,ATRIP,TREX1,PFKFB4, FBXW12,UCN2,CCDC72,SHISA5,COL7A1* | Chr3 | 242 | 0.0031 |
| *OTOL1* | Chr3 | 813 | 0.0126 |
| *NAALADL2* | Chr3 | 883 | 0.0101 |
| *MORC1* | Chr3 | 551 | 0.0121 |
| *PA2G4P4,LEKR1* | Chr3 | 790 | 0.0146 |
| *ABHD5,ANO10* | Chr3 | 218 | 0.0153 |
| *SLC15A2,EAF2,ILDR1,IQCB1* | Chr3 | 615 | 0.0068 |
| No genes in this region | Chr3 | 374 | 0.0064 |

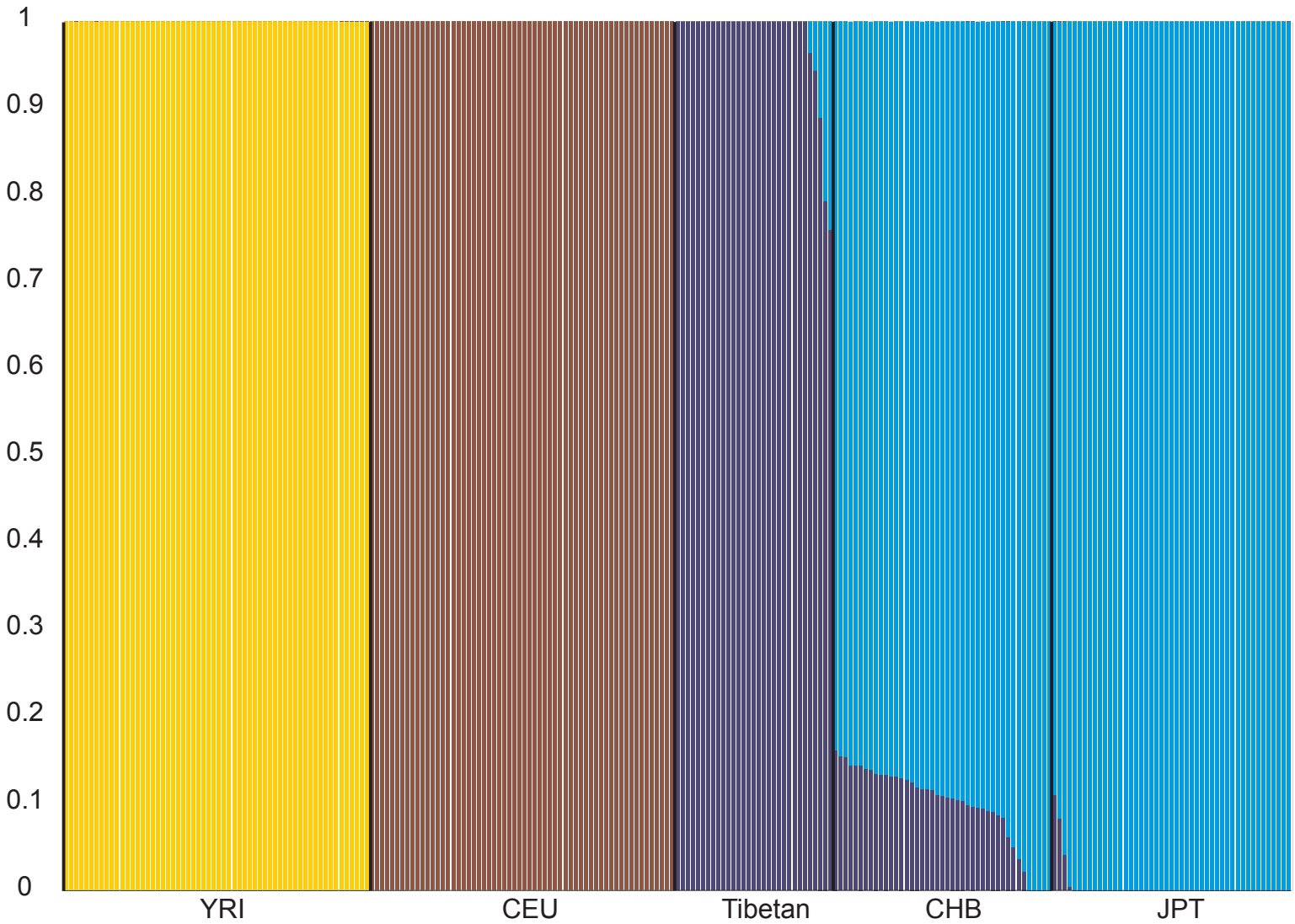
| | | | |
|---|---|---|---|
| *FLJ46210,LOC389151,LOC729627* | Chr3 | 701 | 0.0125 |
| *SLC25A26,LRIG1* | Chr3 | 332 | 0.0157 |
| *ZBTB20* | Chr3 | 578 | 0.0155 |
| *TIPARP* | Chr3 | 789 | 0.0130 |
| No genes in this region | Chr3 | 220 | 0.0055 |
| *PPARG,TSEN2,MKRN2* | Chr3 | 62 | 0.0178 |
| *WDR49,SERPINI1,PDCD10* | Chr3 | 844 | 0.0196 |
| *FSTL1,NDUFB4* | Chr3 | 608 | 0.0162 |
| *EPHA6* | Chr3 | 490 | 0.0168 |
| No genes in this region | Chr3 | 839 | 0.0015 |
| No genes in this region | Chr3 | 695 | 0.0062 |
| *CADPS* | Chr3 | 313 | 0.0047 |
| *ZBTB20* | Chr3 | 579 | 0.0037 |
| *PAK2,SENP5,PIGZ,LOC152217,NCBP2* | Chr3 | 990 | 0.0131 |
| No genes in this region | Chr3 | 219 | 0.0014 |
| *FOXP1* | Chr3 | 357 | 0.0193 |
| *MAGI1* | Chr3 | 327 | 0.0114 |
| *DPPA2,DPPA4* | Chr3 | 552 | 0.0082 |
| *INPP4B* | Chr4 | 718 | 0.0175 |
| *TRPC3,BBS7* | Chr4 | 615 | 0.0029 |
| *GRSF1,MOBKL1A,RUFY3* | Chr4 | 359 | 0.0193 |
| *GAB1* | Chr4 | 722 | 0.0192 |
| No genes in this region | Chr4 | 492 | 0.0064 |
| *TMEM155,EXOSC9,CCNA2,ANXA5, LOC100192379,BBS7* | Chr4 | 614 | 0.0044 |
| No genes in this region | Chr4 | 742 | 0.0114 |
| *KCNIP4* | Chr4 | 106 | 0.0058 |
| No genes in this region | Chr4 | 795 | 0.0134 |
| *GRID2* | Chr4 | 468 | 0.0003 |
| No genes in this region | Chr4 | 171 | 0.0158 |
| *RAPGEF2* | Chr4 | 802 | 0.0005 |
| No genes in this region | Chr4 | 172 | 0.0190 |
| No genes in this region | Chr4 | 140 | 0.0136 |
| *SCRG1,SAP30,GALNT7,HMGB2* | Chr4 | 872 | 0.0087 |
| *USP38* | Chr4 | 721 | 0.0025 |
| No genes in this region | Chr4 | 803 | 0.0004 |
| No genes in this region | Chr4 | 526 | 0.0011 |
| *GRID2* | Chr4 | 469 | 0.0032 |
| *CPEB2* | Chr4 | 73 | 0.0072 |
| *INPP4B* | Chr4 | 719 | 0.0053 |

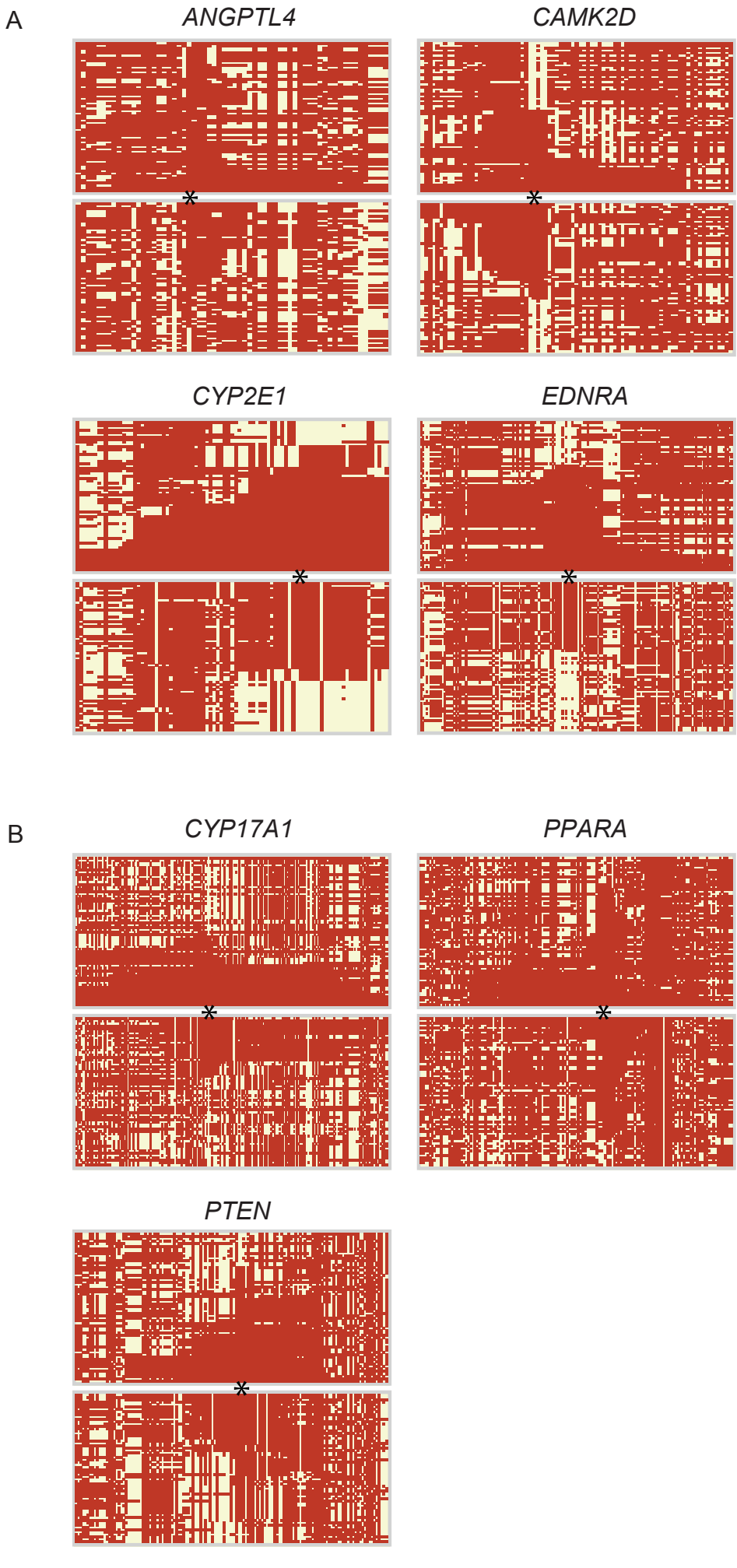
| | | | |
|---|---|---|---|
| *SLIT2* | Chr4 | 99 | 0.0086 |
| No genes in this region | Chr4 | 593 | 0.0103 |
| | | | |
| *CXCL5,CXCL3,PPBP,CXCL2,PF4* | Chr4 | 375 | 0.0189 |
| *NCRNA00099,KCNIP4* | Chr4 | 107 | 0.0007 |
| *EDNRA,TMEM184C,LOC90826* | Chr4 | 743 | 0.0118 |
| No genes in this region | Chr4 | 170 | 0.0067 |
| *MARCH11* | Chr5 | 81 | 0.0093 |
| *PDE4D* | Chr5 | 294 | 0.0122 |
| *RNF180* | Chr5 | 318 | 0.0188 |
| *RASA1* | Chr5 | 432 | 0.0033 |
| *HAVCR1,TIMD4,PPP1R2P3* | Chr5 | 781 | 0.0043 |
| *MCC* | Chr5 | 562 | 0.0027 |
| No genes in this region | Chr5 | 823 | 0.0021 |
| *MCC,YTHDC2* | Chr5 | 564 | 0.0050 |
| No genes in this region | Chr5 | 485 | 0.0123 |
| *CCT5,FAM173B,CMBL* | Chr5 | 51 | 0.0065 |
| No genes in this region | Chr5 | 586 | 0.0100 |
| *APC,SRP19,MCC,REEP5,DCP2* | Chr5 | 561 | 0.0035 |
| No genes in this region | Chr5 | 588 | 0.0002 |
| *ADAMTS6* | Chr5 | 323 | 0.0155 |
| *CLINT1* | Chr5 | 786 | 0.0045 |
| No genes in this region | Chr5 | 680 | 0.0110 |
| *CDH9* | Chr5 | 135 | 0.0197 |
| No genes in this region | Chr5 | 585 | 0.0147 |
| *ZNF131,HMGCS1,MGC42105* | Chr5 | 216 | 0.0156 |
| | | | |
| *HCP5,MICA,HLA-B,MICB,HCG26* | Chr6 | 157 | 0.0074 |
| *COL21A1* | Chr6 | 281 | 0.0126 |
| No genes in this region | Chr6 | 629 | 0.0113 |
| No genes in this region | Chr6 | 417 | 0.0027 |
| *HMGN3* | Chr6 | 400 | 0.0199 |
| *BMP5* | Chr6 | 278 | 0.0144 |
| *ASF1A,FAM184A,MCM9* | Chr6 | 596 | 0.0030 |
| No genes in this region | Chr6 | 416 | 0.0127 |
| *CDYL,RPP40* | Chr6 | 24 | 0.0066 |
| | | | |
| *GABRR1,PNRC1,SRrp35,PM20D2* | Chr6 | 449 | 0.0069 |
| *SOBP,SCML4* | Chr6 | 540 | 0.0180 |
| *PHACTR1* | Chr6 | 65 | 0.0052 |
| *OPN5,C6orf138* | Chr6 | 239 | 0.0152 |
| *CD2AP,GPR111,GPR115* | Chr6 | 238 | 0.0010 |
| *PRIM2* | Chr6 | 286 | 0.0017 |

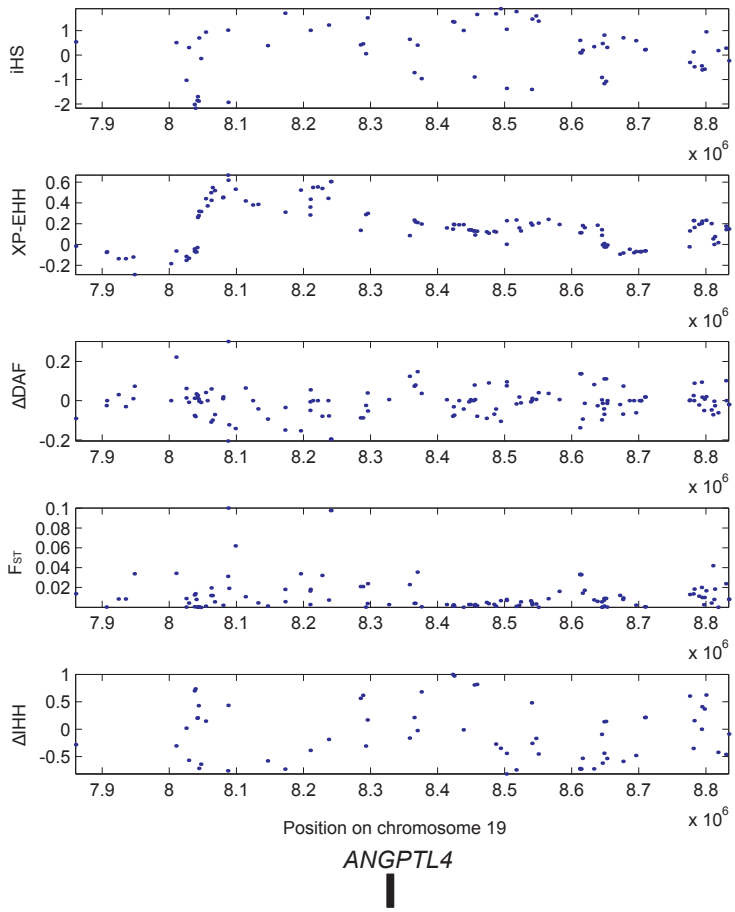
| | | | |
|---|---|---|---|
| No genes in this region | Chr6 | 521 | 0.0163 |
| *GCNT2* | Chr6 | 53 | 0.0088 |
| *PRIM2* | Chr6 | 287 | 0.0006 |
| *PKHD1* | Chr6 | 258 | 0.0182 |
| No genes in this region | Chr7 | 677 | 0.0138 |
| *NRF1* | Chr7 | 645 | 0.0164 |
| *TRPV6,C7orf34,EPHB6,KEL,TRPV5* | Chr7 | 711 | 0.0054 |
| *ZNF800* | Chr7 | 634 | 0.0023 |
| *GUSB,VKORC1L1,ASL* | Chr7 | 325 | 0.0057 |
| *ZPBP* | Chr7 | 250 | 0.0014 |
| *PRSS1,TRY6,PRSS2* | Chr7 | 710 | 0.0040 |
| No genes in this region | Chr7 | 99 | 0.0102 |
| No genes in this region | Chr7 | 425 | 0.0165 |
| *SDK1* | Chr7 | 20 | 0.0109 |
| *GCC1,FSCN3,ARF5,PAX4,SND1* | Chr7 | 635 | 0.0081 |
| *LOC401397,GPR85* | Chr7 | 562 | 0.0059 |
| *EXOC4* | Chr7 | 664 | 0.0139 |
| No genes in this region | Chr7 | 274 | 0.0129 |
| *IFRD1,C7orf53* | Chr7 | 559 | 0.0077 |
| *C7orf58,FAM3C,WNT16* | Chr7 | 603 | 0.0106 |
| *MAD1L1* | Chr7 | 10 | 0.0151 |
| No genes in this region | Chr8 | 247 | 0.0046 |
| No genes in this region | Chr8 | 103 | 0.0184 |
| No genes in this region | Chr8 | 385 | 0.0048 |
| No genes in this region | Chr8 | 634 | 0.0097 |
| No genes in this region | Chr8 | 384 | 0.0031 |
| No genes in this region | Chr8 | 246 | 0.0049 |
| No genes in this region | Chr8 | 102 | 0.0140 |
| *HNF4G* | Chr8 | 383 | 0.0112 |
| *CHRNA6,CHRNB3* | Chr8 | 213 | 0.0132 |
| *MAMDC2* | Chr9 | 359 | 0.0116 |
| *LINGO2* | Chr9 | 139 | 0.0171 |
| *SDCCAG3,C9orf163,SNAPC4,SEC16A, PMPCA,NOTCH1,INPP5E* | Chr9 | 692 | 0.0076 |
| *NXNL2,SPIN1* | Chr9 | 451 | 0.0167 |
| *WNK2,C9orf129* | Chr9 | 475 | 0.0179 |
| *HSPA5,GAPVD1,RABEPK* | Chr9 | 635 | 0.0188 |

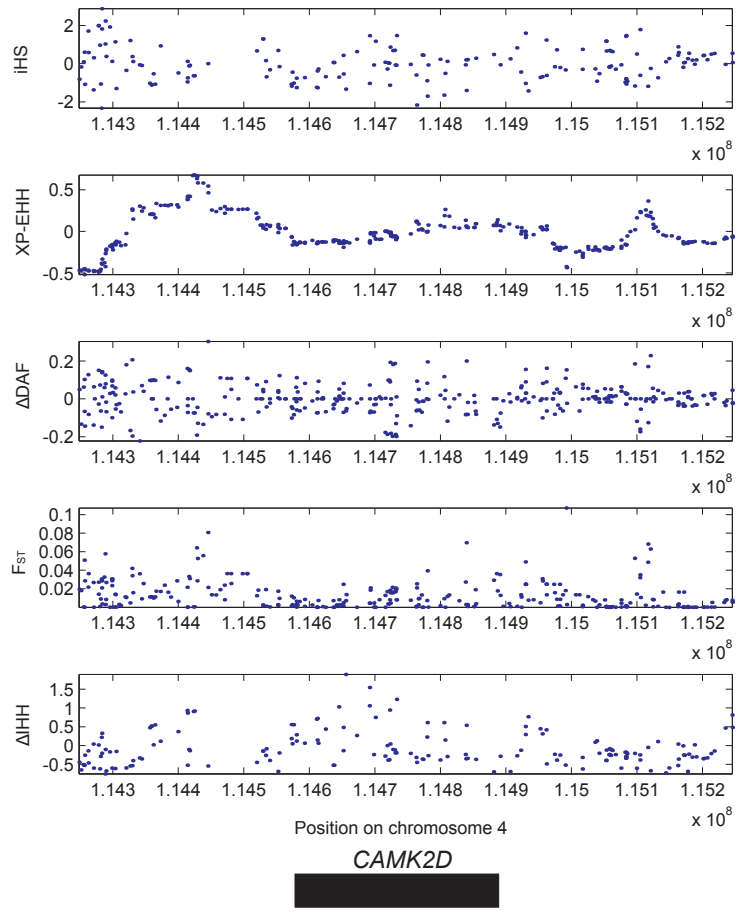

A
CHB
JPT
Tibetan
CHB
JPT
Tibetan
PC 2 (1.68%)
PC 1 (3.75%)

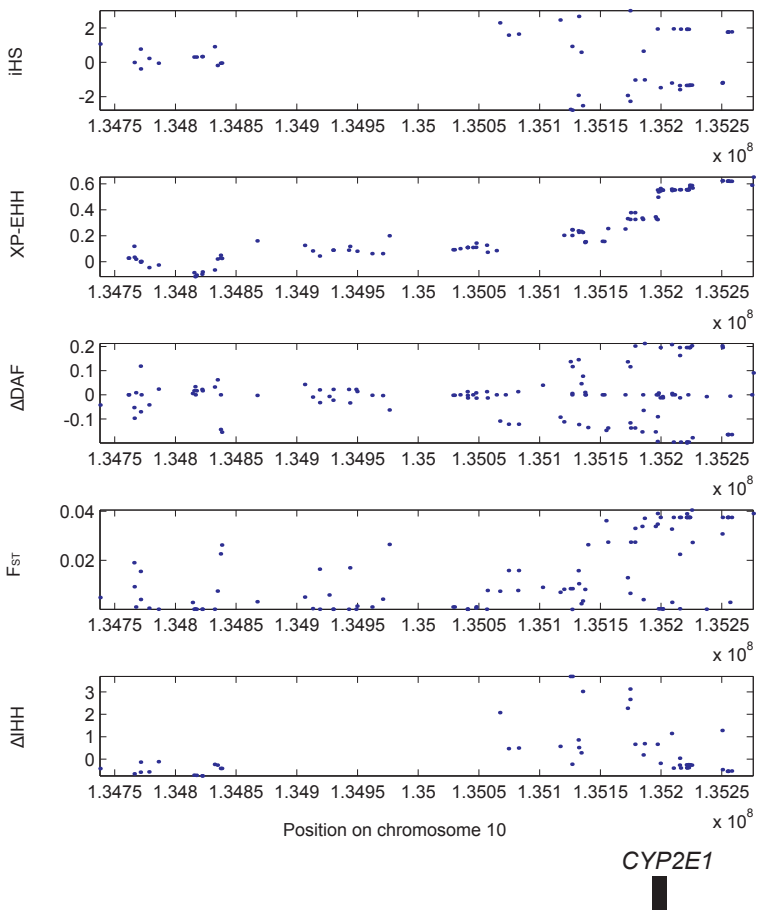

B
CHB-JPT
Tibetan
YRI
CEU
CHB-JPT
Tibetan
YRI
PC 2 (12.4%)
PC 1 (71.6%)

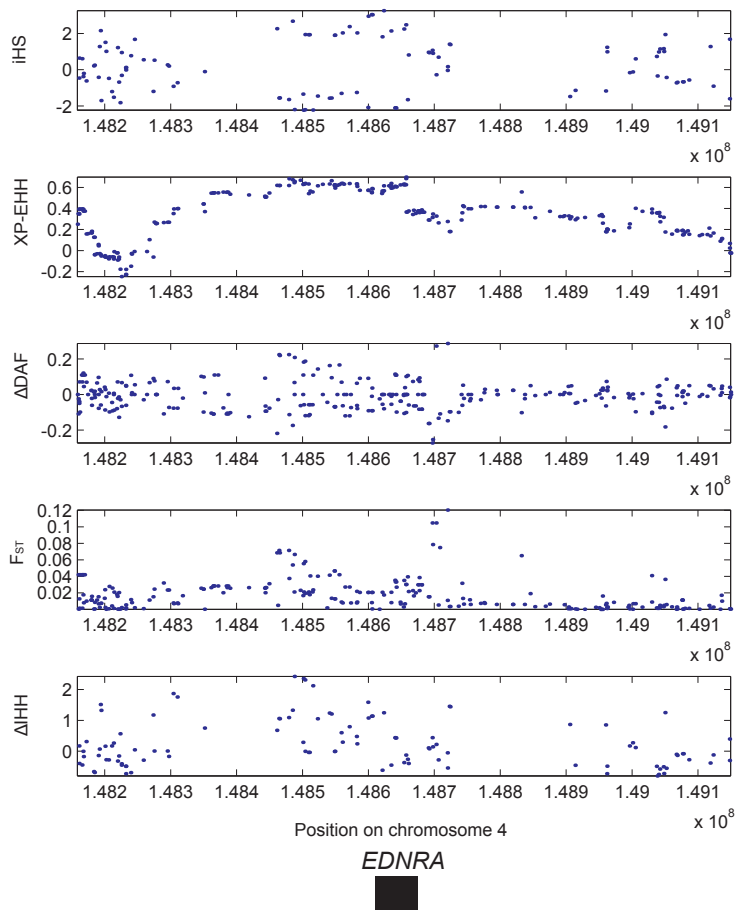

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
YRI
CEU
Tibetan
CHB
JPT

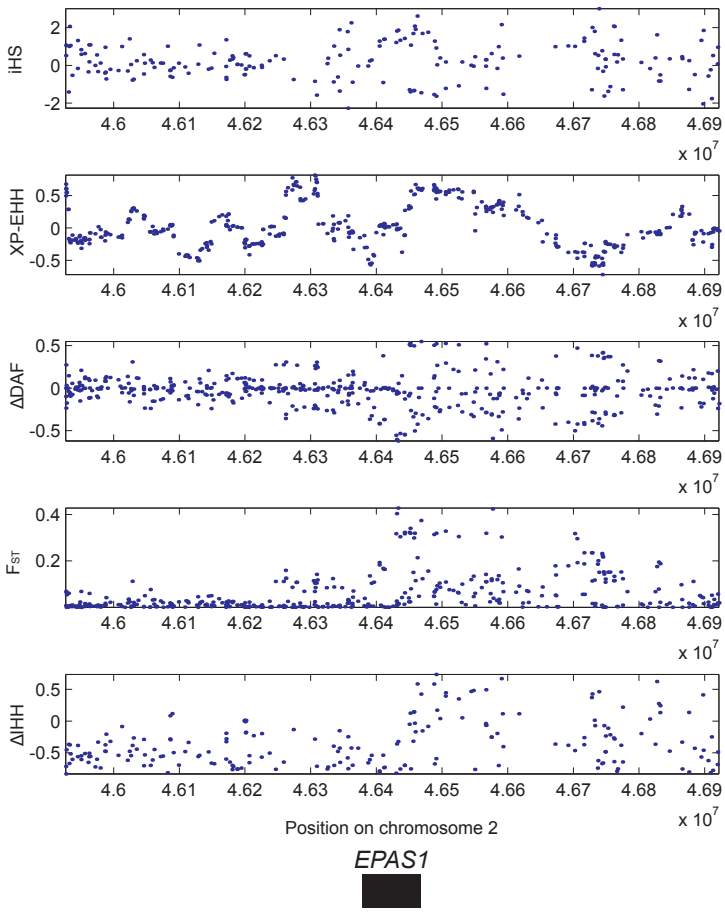

A
ANGPTL4
CAMK2D
CYP2E1
EDNRA
B
CYP17A1
PPARA
PTEN

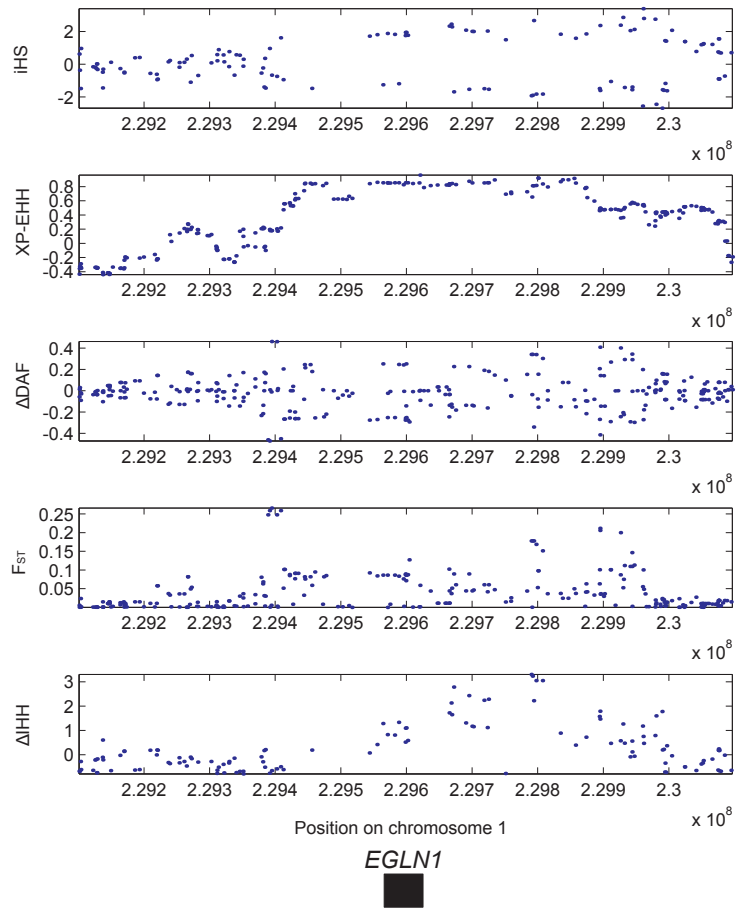

iHS
XP-EHH
ΔDAF
FST
ΔIHH
x 10^6
Position on chromosome 19
ANGPTL4

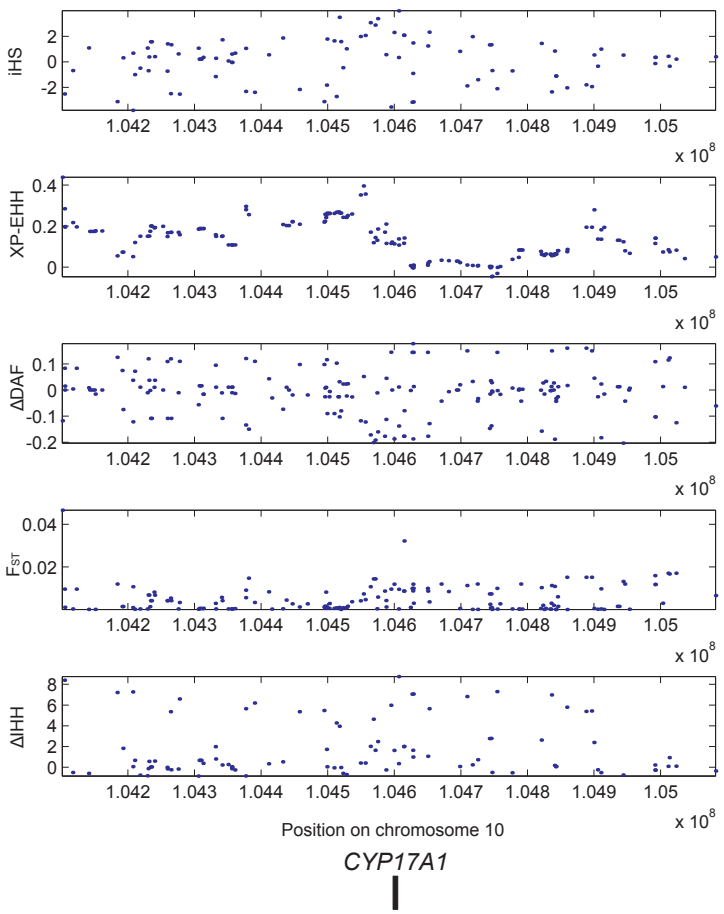

iHS
XP-EHH
ΔDAF
FST
ΔIHH
x 10^8
Position on chromosome 4
CAMK2D

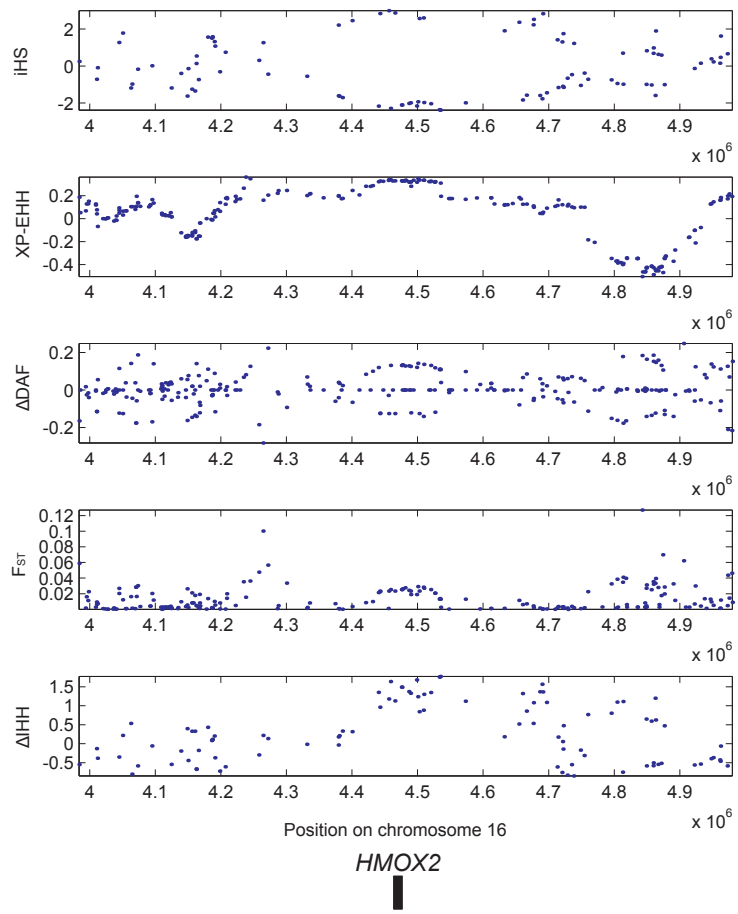

iHS
XP-EHH
ΔDAF
FST
ΔIHH
x 10^8
Position on chromosome 10
CYP2E1

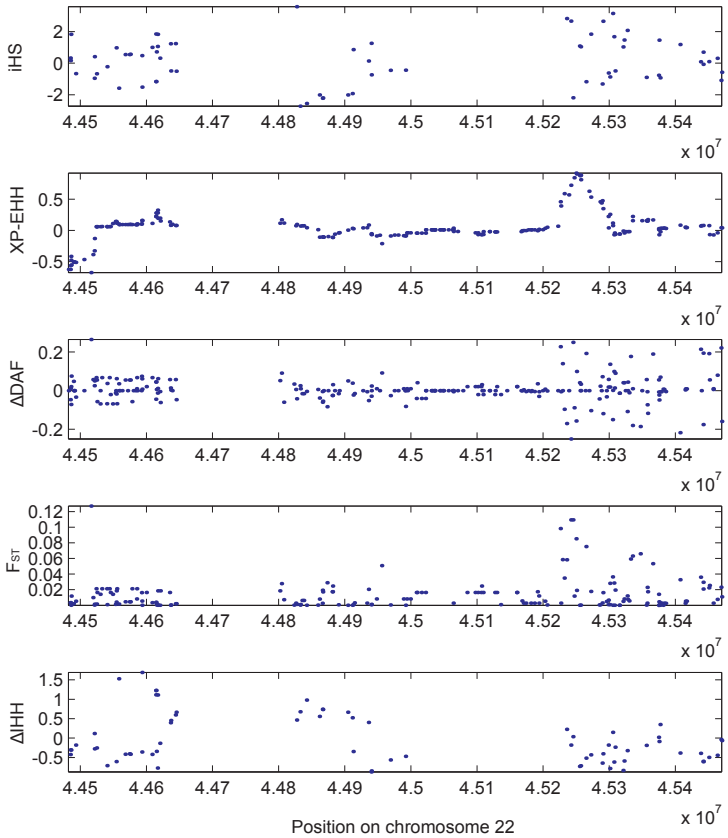

iHS
XP-EHH
ΔDAF
FST
ΔIHH
x 10^8
Position on chromosome 4
EDNRA

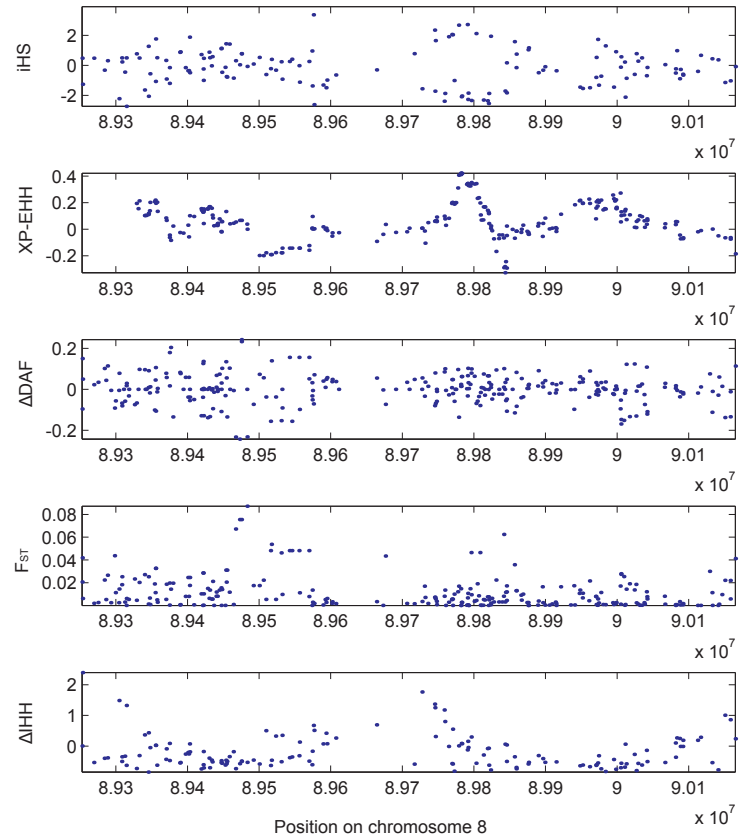

iHS
XP-EHH
ΔDAF
FST
ΔIHH
x 10^7
Position on chromosome 2
EPAS1

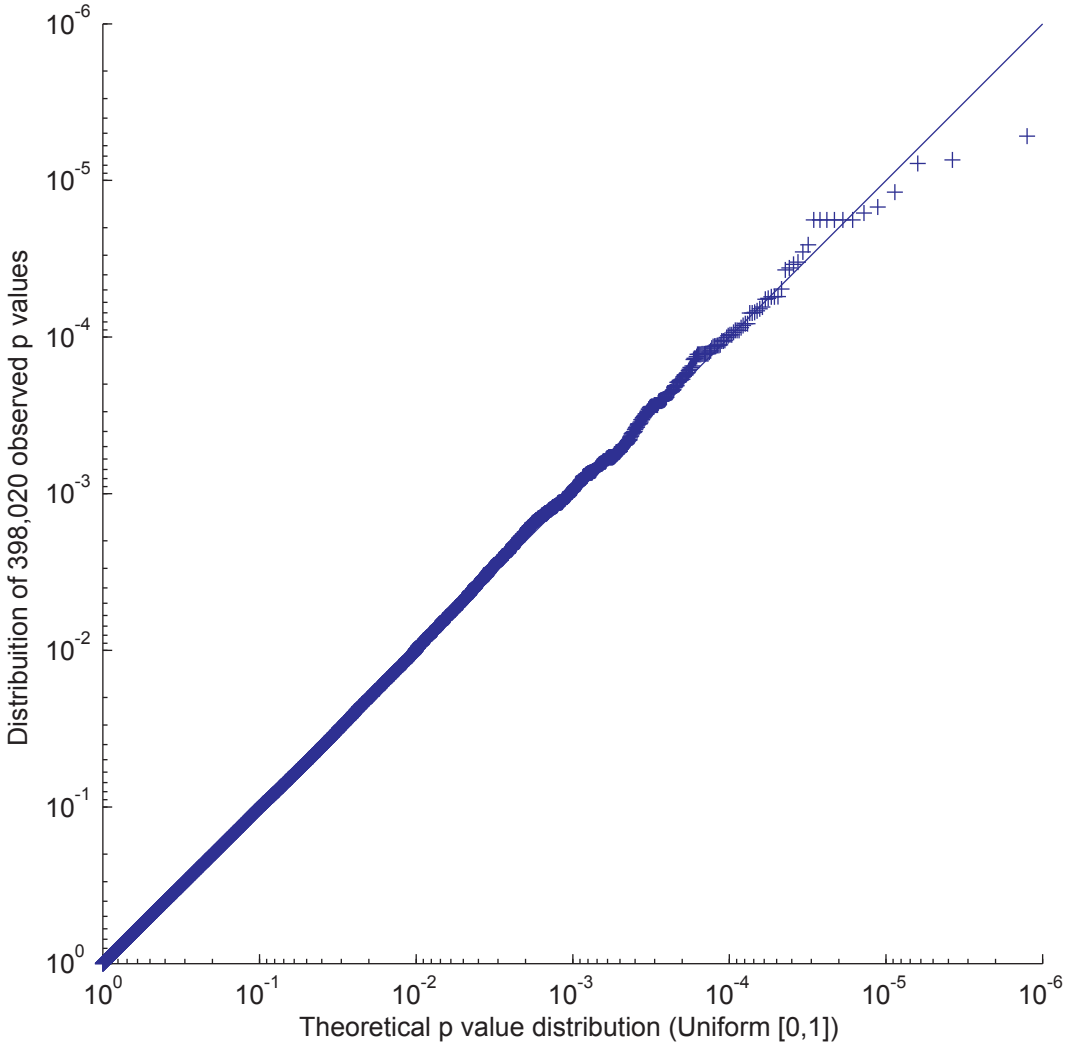

iHS
XP-EHH
ΔDAF
FST
ΔIHH
x 10^8
Position on chromosome 1
EGLN1

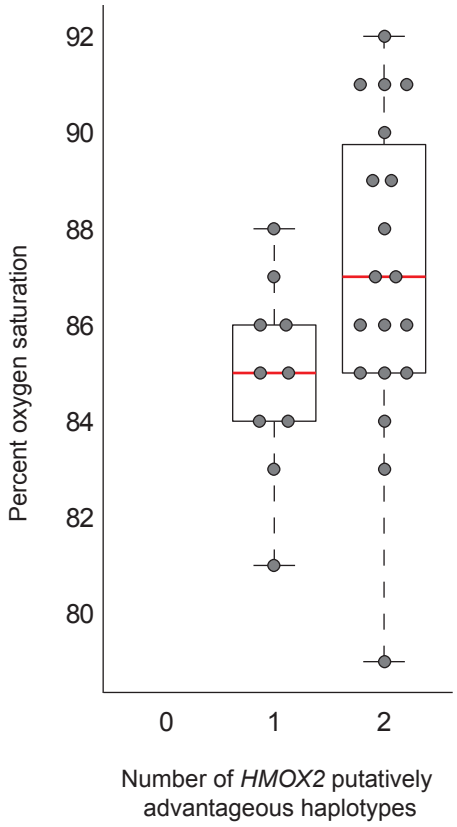

iHS
XP-EHH
ΔDAF
FST
ΔIHH
x 10^8
Position on chromosome 10
CYP17A1


iHS
XP-EHH
ΔDAF
FST
ΔIHH
x 10^6
Position on chromosome 16
HMOX2


iHS
XP-EHH
ΔDAF
FST
ΔIHH
4.45 4.46 4.47 4.48 4.49 4.5 4.51 4.52 4.53 4.54
x 10^7
Position on chromosome 22
PPARA


iHS
XP-EHH
ΔDAF
FST
ΔIHH
8.93 8.94 8.95 8.96 8.97 8.98 8.99 9 9.01
x 10^7
Position on chromosome 8
PTEN


Distribuition of 398,020 observed p values
Theoretical p value distribution (Uniform [0,1])
10^-6
10^-5
10^-4
10^-3
10^-2
10^-1
10^0


92
90
88
86
84
82
80
Percent oxygen saturation
0
1
2
Number of HMOX2 putatively advantageous haplotypes