The American Journal of Human Genetics, Volume *99*

Supplemental Data

Ancestral Origins and Genetic History of Tibetan Highlanders

Dongsheng Lu, Haiyi Lou, Kai Yuan, Xiaoji Wang, Yuchen Wang, Chao Zhang, Yan Lu, Xiong Yang, Lian Deng, Ying Zhou, Qidi Feng, Ya Hu, Qiliang Ding, Yajun Yang, Shilin Li, Li Jin, Yaqun Guan, Bing Su, Longli Kang, and Shuhua Xu

Sample ID	Ethnicity	Sex	Location	Sample ID	Ethnicity	Sex	Location
AAGC022151	Han	male	Henan	AAGC022108	Tibetan	female	Chamdo
AAGC022152	Han	female	Gansu	AAGC022109	Tibetan	female	Nagqu
AAGC022153	Han	female	Xinjiang	AAGC022110	Tibetan	female	Nagqu
AAGC022154	Han	female	Henan	AAGC022111	Tibetan	female	Nagqu
AAGC022155	Han	female	Xinjiang	AAGC022112	Tibetan	female	Lhasa
AAGC022156	Han	female	Shandong	AAGC022113	Tibetan	female	Shigatse
AAGC022157	Han	female	Gansu	AAGC022114	Tibetan	female	Shigatse
AAGC022158	Han	male	Sichuan	AAGC022107	Tibetan	female	Shannan
AAGC022159	Han	female	Xinjiang	AAGC022051	Tibetan	female	Shannan
AAGC022160	Han	female	Henan	AAGC022052	Tibetan	female	Shigatse
AAGC022161	Han	male	Shandong	AAGC022053	Tibetan	female	Shannan
AAGC022162	Han	male	Jiangsu	AAGC022054	Tibetan	female	Shannan
AAGC022163	Han	male	Jiangsu	AAGC022055	Tibetan	female	Shannan
AAGC022164	Han	female	Sichuan	AAGC022056	Tibetan	male	Chamdo
AAGC022165	Han	female	Gansu	AAGC022057	Tibetan	male	Chamdo
AAGC022166	Han	female	Sichuan	AAGC022058	Tibetan	male	Chamdo
AAGC022167	Han	female	Henan	AAGC022059	Tibetan	male	Chamdo
AAGC022168	Han	female	Jiangsu	AAGC022060	Tibetan	male	Chamdo
AAGC022169	Han	male	Gansu	AAGC022061	Tibetan	male	Lhasa
AAGC022170	Han	male	Shandong	AAGC022062	Tibetan	male	Lhasa
AAGC022171	Han	male	Anhui	AAGC022063	Tibetan	male	Shannan
AAGC022172	Han	male	Henan	AAGC022064	Tibetan	male	Shannan
AAGC022173	Han	female	Henan	AAGC022065	Tibetan	male	Shigatse
AAGC022174	Han	male	Sichuan	AAGC022066	Tibetan	male	Shigatse
AAGC022175	Han	female	Henan	AAGC022067	Tibetan	male	Shigatse
AAGC022176	Han	female	Shandong	AAGC022068	Tibetan	male	Shigatse
AAGC022177	Han	female	Henan	AAGC022069	Tibetan	male	Shigatse
AAGC022178	Han	male	Sichuan	AAGC022070	Tibetan	male	Nyingchi
AAGC022179	Han	male	Hubei	AAGC022071	Tibetan	male	Nyingchi
AAGC022180	Han	female	Guangdong	AAGC022072	Sherpa	male	Shigatse
AAGC022181	Han	female	Sichuan	AAGC022073	Sherpa	male	Shigatse
AAGC022182	Han	male	Xinjiang	AAGC022074	Sherpa	female	Shigatse
AAGC022183	Han	male	Jiangsu	AAGC022075	Sherpa	female	Shigatse
AAGC022184	Han	male	Gansu	AAGC022076	Sherpa	male	Shigatse
AAGC022185	Han	male	Chongqing	AAGC022077	Tibetan	male	Shigatse
AAGC022186	Han	male	Henan	AAGC022078	Tibetan	male	Shigatse
AAGC022187	Han	male	Henan	AAGC022079	Tibetan	female	Shigatse
AAGC022188	Han	male	Chongqing	AAGC022080	Tibetan	female	Shigatse
AAGC022189	Han	male	Xinjiang				

Table S1 **Information of the samples for sequencing in this study**

Sample ID	Ethnicity	Mapped reads	Read depth	Sample ID	Ethnicity	Mapped reads	Read depth
AAGC022151	Han	1,591,144,384	79.6×	AAGC022051	Tibetan	749,183,856	37.5×
AAGC022152	Han	1,673,465,613	83.7×	AAGC022052	Tibetan	797,679,786	39.9×
AAGC022153	Han	1,695,829,400	84.8×	AAGC022053	Tibetan	812,473,983	40.6×
AAGC022154	Han	1,691,085,840	84.6×	AAGC022054	Tibetan	806,315,598	40.3×
AAGC022155	Han	1,616,239,134	80.8×	AAGC022055	Tibetan	617,510,254	30.9×
AAGC022156	Han	1,632,671,866	81.6×	AAGC022056	Tibetan	720,647,639	36.0×
AAGC022157	Han	1,692,498,209	84.6×	AAGC022057	Tibetan	731,673,851	36.6×
AAGC022158	Han	1,555,333,449	77.8×	AAGC022058	Tibetan	727,579,311	36.4×
AAGC022159	Han	812,464,199	40.6×	AAGC022059	Tibetan	651,295,687	32.6×
AAGC022160	Han	830,888,051	41.5×	AAGC022060	Tibetan	627,802,914	31.4×
AAGC022161	Han	837,055,727	41.9×	AAGC022061	Tibetan	748,037,548	37.4×
AAGC022162	Han	831,193,590	41.6×	AAGC022062	Tibetan	682,690,826	34.1×
AAGC022163	Han	842,546,144	42.1×	AAGC022063	Tibetan	740,232,029	37.0×
AAGC022164	Han	839,036,197	42.0×	AAGC022064	Tibetan	833,814,794	41.7×
AAGC022165	Han	833,031,290	41.7×	AAGC022065	Tibetan	827,002,773	41.4×
AAGC022166	Han	823,596,584	41.2×	AAGC022066	Tibetan	782,327,372	39.1×
AAGC022167	Han	822,539,963	41.1×	AAGC022067	Tibetan	761,167,865	38.1×
AAGC022168	Han	841,427,099	42.1×	AAGC022068	Tibetan	793,520,776	39.7×
AAGC022169	Han	819,660,225	41.0×	AAGC022069	Tibetan	796,307,201	39.8×
AAGC022170	Han	841,146,434	42.1×	AAGC022070	Tibetan	801,217,703	40.1×
AAGC022171	Han	843,735,336	42.2×	AAGC022071	Tibetan	806,962,026	40.3×
AAGC022172	Han	842,541,191	42.1×	AAGC022072	Sherpa	782,905,194	39.1×
AAGC022173	Han	832,455,997	41.6×	AAGC022073	Sherpa	761,382,444	38.1×
AAGC022174	Han	791,968,738	39.6×	AAGC022074	Sherpa	752,908,895	37.6×
AAGC022175	Han	757,869,466	37.9×	AAGC022075	Sherpa	738,558,824	36.9×
AAGC022176	Han	790,897,953	39.5×	AAGC022076	Sherpa	750,982,567	37.5×
AAGC022177	Han	801,776,959	40.1×	AAGC022077	Tibetan	712,250,441	35.6×
AAGC022178	Han	779,127,405	39.0×	AAGC022078	Tibetan	787,354,852	39.4×
AAGC022179	Han	797,138,859	39.9×	AAGC022079	Tibetan	768,522,496	38.4×
AAGC022180	Han	779,907,616	39.0×	AAGC022080	Tibetan	773,418,725	38.7×
AAGC022181	Han	806,697,742	40.3×	AAGC022107	Tibetan	695,413,410	34.8×
AAGC022182	Han	770,819,826	38.5×	AAGC022108	Tibetan	1,485,808,123	73.8×
AAGC022183	Han	733,161,684	36.7×	AAGC022109	Tibetan	1,552,598,646	76.6×
AAGC022184	Han	754,995,394	37.7×	AAGC022110	Tibetan	1,560,413,927	76.8×
AAGC022185	Han	685,607,075	34.3×	AAGC022111	Tibetan	1,471,230,647	72.3×
AAGC022186	Han	693,816,750	34.7×	AAGC022112	Tibetan	1,528,021,903	74.8×
AAGC022187	Han	610,984,692	30.5×	AAGC022113	Tibetan	1,524,930,381	74.8×
AAGC022188	Han	691,654,456	34.6×	AAGC022114	Tibetan	1,533,664,538	75.4×
AAGC022189	Han	694,617,260	34.7×				

Table S2 **Summary statistics of NGS data of TIB and HAN samples**

Continent region ID	# of Populations (simple ID)	Sample size
Africa	49	535
WestEurasia	60	781
SouthAsia	22	280
CentralAsiaSiberia	23	266
EastAsia	22	243
Oceania	3	27
America	24	213
Total	203	2,345

Table S3 Number of HuOrigin populations included in this study

Note: HuOrigin, Affymetrix human Origins genotyping dataset for modern human. We used the 'Simple population ID' instead of the 'Verbose population ID' to assign the population ID for each individual.

Y Chromosome Haplogroup	HAN (%) 20	TBN (%) 18	SHP (%) 3
C3f	1 (5.00)	0 (0.00)	0 (0.00)
C3f1	2 (10.00)	0 (0.00)	0 (0.00)
D1a	0 (0.00)	8 (44.44)	0 (0.00)
D3a	0 (0.00)	4 (22.22)	1 (33.33)
N	2 (10.00)	1 (5.56)	0 (0.00)
N1b	1 (5.00)	0 (0.00)	0 (0.00)
N1c1	0 (0.00)	1 (5.56)	0 (0.00)
O1a	1 (5.00)	0 (0.00)	0 (0.00)
O1a1	1 (5.00)	0 (0.00)	0 (0.00)
O2a1	2 (10.00)	0 (0.00)	0 (0.00)
O2a1a	1 (5.00)	0 (0.00)	0 (0.00)
O3a1c1	2 (10.00)	0 (0.00)	0 (0.00)
O3a2b1	1 (5.00)	0 (0.00)	0 (0.00)
O3a2c	4 (20.00)	0 (0.00)	0 (0.00)
O3a2c1a	0 (0.00)	4 (22.22)	1 (33.33)
O3a2c2	1 (5.00)	0 (0.00)	0 (0.00)
Q1a1a1	1 (5.00)	0 (0.00)	0 (0.00)
R1a1a1b2a1a	0 (0.00)	0 (0.00)	1 (33.33)

Table S5 Distribution of Y chromosome haplogroups in the three populations

MtDNA Haplogroup	HAN (%) 39	TBN (%) 33	SHP (%) 5
A	3 (7.69)	3 (9.09)	2 (40.00)
B4	2 (5.13)	1 (3.03)	0 (0.00)
B5	4 (10.26)	0 (0.00)	0 (0.00)
C*	1 (2.56)	0 (0.00)	0 (0.00)
C4	1 (2.56)	4 (12.12)	1 (20.00)
C5	0 (0.00)	1 (3.03)	0 (0.00)
C7	3 (7.69)	0 (0.00)	0 (0.00)
D4	7 (17.95)	3 (9.09)	0 (0.00)
D5	0 (0.00)	1 (3.03)	0 (0.00)
F1	3 (7.69)	3 (9.09)	0 (0.00)
F2	1 (2.56)	1 (3.03)	0 (0.00)
F3	2 (5.13)	0 (0.00)	0 (0.00)
G2	1 (2.56)	2 (6.06)	0 (0.00)
G3	1 (2.56)	1 (3.03)	0 (0.00)
M10	3 (7.69)	0 (0.00)	0 (0.00)
M13	0 (0.00)	4 (12.12)	0 (0.00)
M5	0 (0.00)	0 (0.00)	1 (20.00)
M62	0 (0.00)	1 (3.03)	0 (0.00)
M8	1 (2.56)	0 (0.00)	0 (0.00)
M9	0 (0.00)	6 (18.18)	1 (20.00)
N9	2 (5.13)	0 (0.00)	0 (0.00)
R	3 (7.69)	0 (0.00)	0 (0.00)
Y	3 (7.69)	0 (0.00)	0 (0.00)
Z	0 (0.00)	2 (6.06)	0 (0.00)

Table S6 **Distribution of mtDNA haplogroups in the three populations**

Population ID	F_{ST}	Population ID	F_{ST}	Population ID	F_{ST}	Population ID	F_{ST}	Population ID	F_{ST}
AA	0.151	Dinka	0.177	Kalash	0.103	Nogais	0.064	Turkish	0.086
Abkhasians	0.102	Druze	0.105	Kalmyks	0.018	NorthOssetians	0.091	Turkish_Jew	0.104
Adyghe	0.095	Egyptians	0.098	Kaqchikel	0.099	Norwegian	0.105	Turkmens	0.056
Albanian	0.105	English	0.106	Karitiana	0.162	Ojibwa	0.073	Tuscany	0.106
Aleutian	0.071	Esan	0.184	Kgalagadi	0.203	Onge	0.142	Tuvinian	0.024
Algerian	0.116	Eskimo	0.074	Kharia	0.056	Orcadians	0.107	Ukranians	0.103
Algonquin	0.107	Estonians	0.103	Khomani	0.203	Oroqens	0.018	Ulchi	0.031
Altaians	0.025	Ethiopian_Jews	0.117	Khwe	0.187	Palestinians	0.099	Uygurs	0.030
Ami	0.047	Evens	0.033	Kinh	0.022	Papuan	0.167	Uzbeks	0.044
Armenians	0.102	Finnish	0.098	Korean	0.012	Pathans	0.076	Vishwabrahmin	0.066
Ashkenazy_Jews	0.105	French	0.103	Koryaks	0.070	Pima	0.128	Wambo	0.186
Atayal	0.069	French_South	0.110	Kumyks	0.092	Punjabi	0.070	Xibo	0.012
Aymara	0.107	Gambian	0.181	Kusunda	0.037	Quechua	0.099	Xuun	0.217
Balkars	0.090	Gana	0.211	Kyrgyzians	0.023	Russians_North	0.091	Yakuts	0.032
Balochi	0.084	Georgian_Jews	0.107	Lahu	0.032	Saharawi	0.118	Yemenese	0.100
BantuKenya	0.177	Georgians	0.105	Lebanese	0.101	Sandawe	0.144	Yemenite_Jews	0.113
BantuSouthAfrica	0.182	Greek	0.102	Lezgins	0.098	Sardinian	0.116	**Yizu**	**0.007**
Basque	0.112	Guarani	0.108	Libyan_Jew	0.110	Saudis	0.111	Yoruba	0.170
BedouinA	0.098	Gui	0.241	Lithuanians	0.108	Scottish	0.109	Yukaghirs	0.036
BedouinB	0.120	GujaratiA	0.077	Lodhi	0.068	She	0.023	Zapotec	0.100
Belarusians	0.103	GujaratiB	0.074	Luhya	0.172	Shua	0.185		
Bengali	0.058	GujaratiC	0.074	Luo	0.173	Sicilian	0.103		
BiakaPygmy	0.201	GujaratiD	0.075	Makrani	0.088	Sindhi	0.075		
Bolivian	0.097	**HAN**	**0.011**	Mala	0.066	Somali	0.125		
Bougainville	0.131	Hadza	0.193	Maltese	0.105	Spanish	0.102		
Brahui	0.086	Haiom	0.193	Mandenka	0.177	Spanish_North	0.112		
Bulgarians	0.102	Han	0.015	Mansi	0.058	Surui	0.172		
Burusho	0.065	**Han_North**	**0.010**	Masai	0.140	Syrians	0.102		
Cabecar	0.162	Hazara	0.033	Maya	0.098	TIB	0.000		
Cambodian	0.026	Hezhens	0.017	MbutiPygmy	0.230	Taa_East	0.238		
Chechens	0.098	Hoan	0.235	Mende	0.182	Taa_North	0.242		
Chipewyan	0.091	Hungarians	0.101	Miaozu	0.018	Taa_West	0.241		
Chukchis	0.065	Icelandic	0.107	Mixe	0.118	Tajiks_Pamir	0.083		
Chuvashes	0.073	Iranian_Jews	0.106	Mixtec	0.105	Thai	0.028		
Cochin_Jews	0.080	Iranians	0.092	**Mongolas**	**0.010**	Tiwari	0.070		
Cree	0.074	Iraqi_Jews	0.108	Mordovians	0.094	Tshwa	0.215		
Croatian	0.104	Italian_North	0.107	Moroccan_Jews	0.107	Tswana	0.190		
Cypriots	0.106	Itelmen	0.073	Mozabites	0.111	**Tu**	**0.006**		
Czech	0.104	Japanese	0.018	Nama	0.186	Tubalar	0.044		
Dai	0.026	Jordanians	0.100	Naro	0.238	Tujia	0.013		
Damara	0.188	Ju_hoan_North	0.245	**Naxi**	**0.009**	Tunisian	0.103		
Daurs	0.015	Ju_hoan_South	0.248	Nganasans	0.066	Tunisian_Jew	0.111		

Table S7 **Estimation of F_{ST} between TIB and worldwide populations**

chr	start	end	Length(kb)***	Genes	Ancestry	Tibetan frequency*	Supporting methods**
2	46566351	46600894	34.5	*EPAS1*	Denisovan	0.855	A,E,S,CP
2	46600895	46673393	72.5	*EPAS1,LOC101805491*	Ust'-Ishim	0.842	A,CP
2	46673394	46719121	45.7	*TMEM247*	Denisovan	0.763	A,E,S,CP
2	46719122	46728427	9.3	.	Neanderthal	0.763	CP
2	46728428	46742298	13.9	*ATP6V1E2*	Ust'-Ishim	0.737	CP
2	46742299	46749337	7.0	*ATP6V1E2*	Denisovan	0.737	CP
2	46749338	46759493	10.2	.	Neanderthal	0.737	CP
2	46759494	46772998	13.5	*RHOQ*	Denisovan	0.724	E,CP
2	46772999	46818801	45.8	*RHOQ,LOC100506142*	Denisovan/Han	0.368/0.605	A,E,S,CP
2	46818802	46873937	55.1	*PIGF,CRIPT*	Denisovan	0.605	A,E,S,CP

Table S16 Ancestry inference of *EPAS1* and its downstream region in Tibetan population.

* The Tibetan frequency here refers to the maximum Tibetan frequency with the corresponding ancestry in the region.

** A, ArchaicSeeker; S, *S**; E, *E-allele*; CP, ChromoPainter.

*** The boundary of each region was determined manually: some of the small segments (<1 kb) with different ancestry from flanking region was removed and also the 'unknown' region was filled by its most likely neighouring ancestry.

Haplotype	**Start**	**End**	**size(kb)**	**Sample size**	**TMRCA**	$T_{\mathrm{D(Den)}}$	$T_{\mathrm{D(Nean)}}$
new	46577796	46870806	293.0	23	38.4 ± 13.0	944 ± 14	1,178 ± 88
old	46577796	46870806	293.0	17	62.2 ± 17.6	868 ± 16	1,222 ±92

Table S17 TMRCA and divergence time of major Tibetan haplotypes

The time estimation in the table are in thousand year (kya). A slow mutation rate of 0.5×10^{-9} per site per year was assumed. $T_{\mathrm{D(Den)}}$: Divergence time between TIB's haplotypes and Denisovan's sequences; $T_{\mathrm{D(Nean)}}$: Divergence time between TIB's haplotypes and Neanderthal's sequences; Sample size: number of haplotypes with corresponding haplotypes (new and old see definition in maintext Figure 7) at the 293kb region..

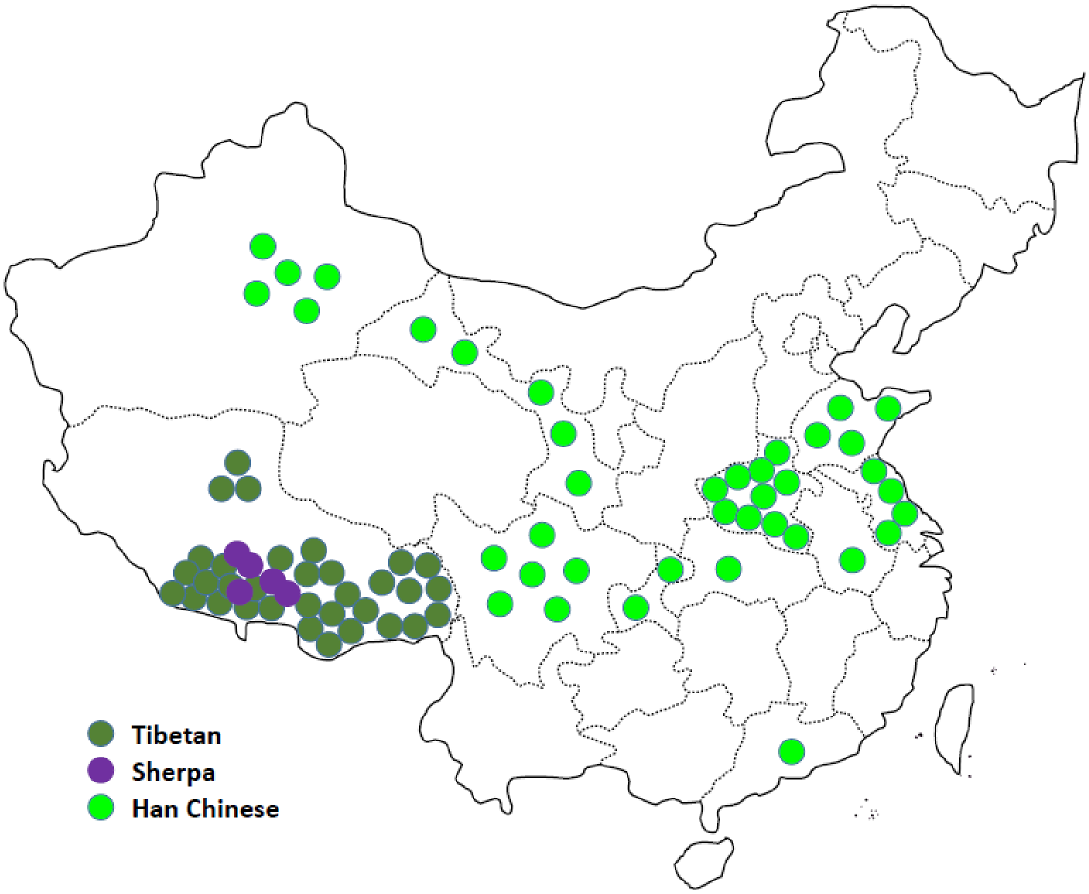

Figure S1 Geographical distribution of Tibetan, Sherpa and Han Chinese samples.

Showing here are the rough locations of the individuals with genomes sequenced in this study. The 33 Tibetan and 5 Sherpa samples were collected from six prefectures (Lhasa, Chamdo, Nagqu, Nyingchi, Shannan, and Shigatse) in the Tibet Autonomous Region. The 39 Han Chinese samples were collected from 10 regionsin China (see Table S1 for more details).

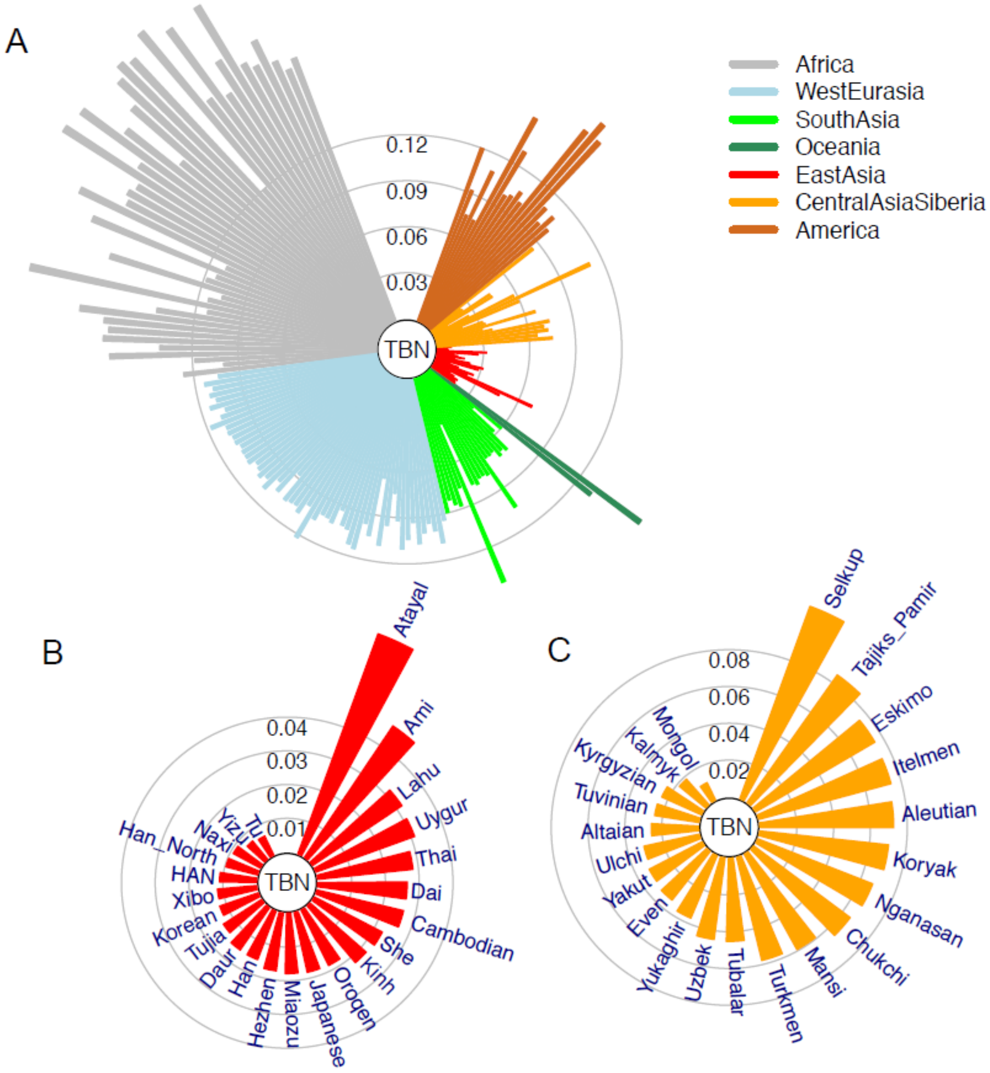

Figure S2 **Genetic affinities of TBN in the context of worldwide populations.**

a. a fan-like chart showing genetic difference (F_{ST}) between TBN and worldwide populations. Each branch represent a comparison between TBN and one of the 256 populations and the length is proportional to the F_{ST} value as indicated in grey circles. The populations are classified by geographical regions and indicated with colors as shown in the legend. The populations in each region are presented in a clockwise order according to great-circle distance to the Tibet. **b.** a fan-like chart showing F_{ST} between TBN and East Asian populations. **c.** a fan-like chart showing F_{ST} between TBN and Central Asian/Siberian populations.

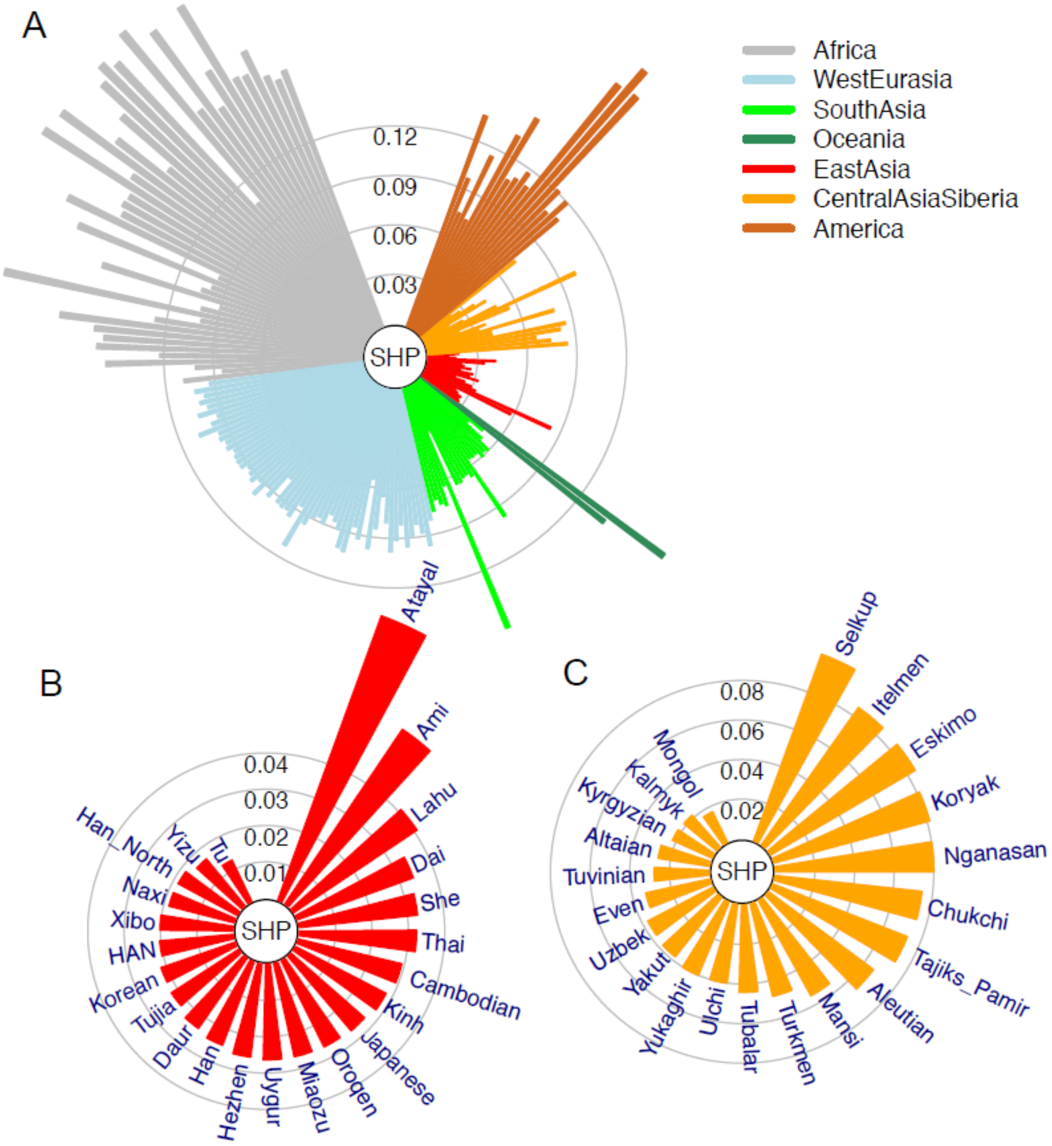

Figure S3 Genetic affinities of SHP in the context of worldwide populations.

a. a fan-like chart showing genetic difference (F_{ST}) between SHP and worldwide populations. Each branch represent a comparison between SHP and one of the 256 populations and the length is proportional to the F_{ST} value as indicated in grey circles. The populations are classified by geographical regions and indicated with colors as shown in the legend. The populations in each region are presented in a clockwise order according to great-circle distance to the Tibet. **b.** a fan-like chart showing F_{ST} between SHP and East Asian populations. **c.** a fan-like chart showing F_{ST} between SHP and Central Asian/Siberian populations.

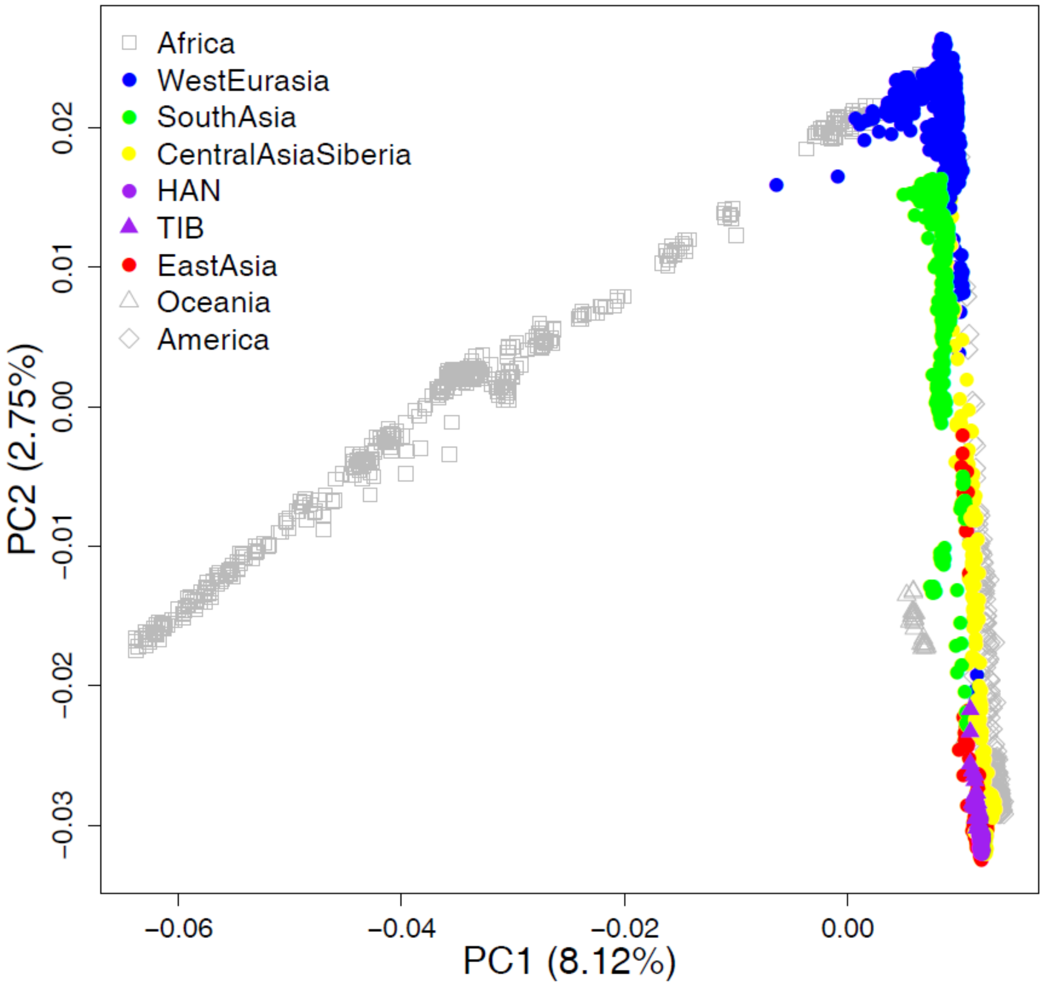

Figure S4 Principal component analysis (PCA) of TIB and HAN Individuals and world-wide population samples.

Analysis of the first two principal components of TIB individuals and world-wide population samples which are consist of 2,035 Human Origin samples; Geographical regions where the individuals are located are indicated with colors as shown in the legend. Numbers in brackets denote variance explained by each principal component (PC).

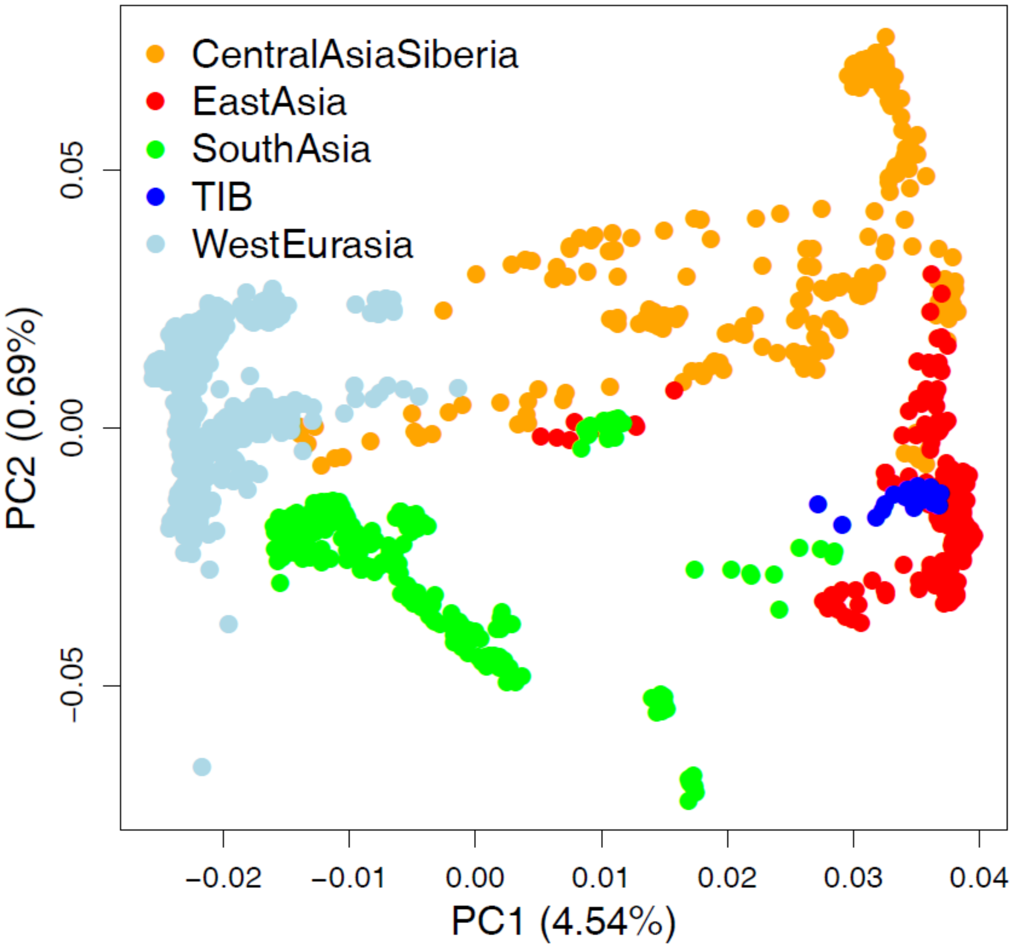

Figure S5 PCA of TIB Individuals and non-African samples (with Oceanian samples excluded).

Analysis of the first two principal components of TIB individuals and non-African samples (with Oceanian samples excluded); Geographical regions where the individuals are located are indicated with colors as shown in the legend. Numbers in brackets denote variance explained by each PC.

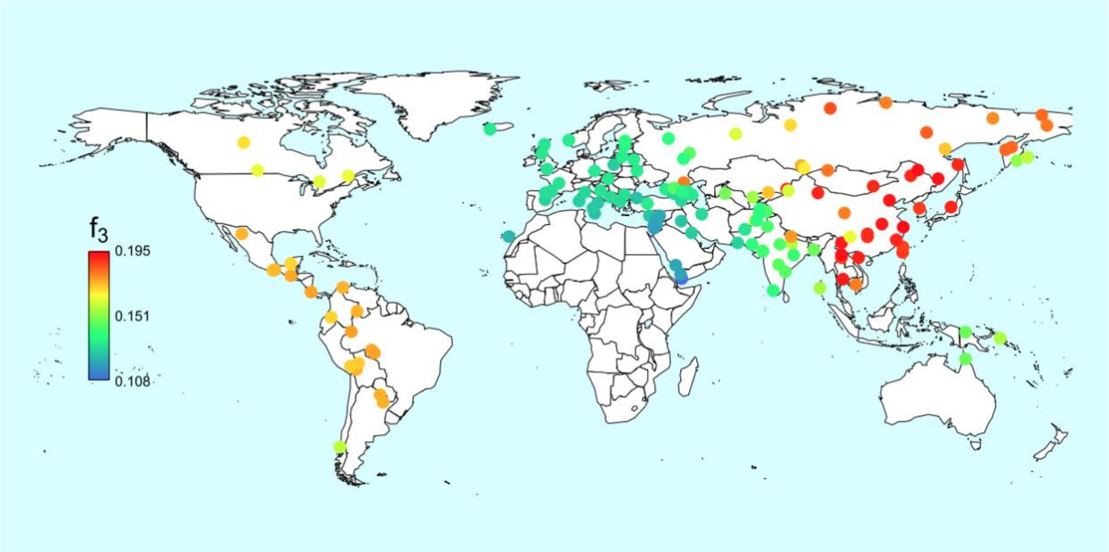

Figure S6 Heat map of outgroup *f*3 statistics testing for relatedness of non-African populations to TIB.

The values shown here are based on computing f_3(African; TIB, X), where X is a non-African population. A warmer color denotes a higher f_3 value, which indicates population X shared greater drift with Tibetan derived from African and suggests a closer relationship to TIB. For list of population labels, see Table S4.

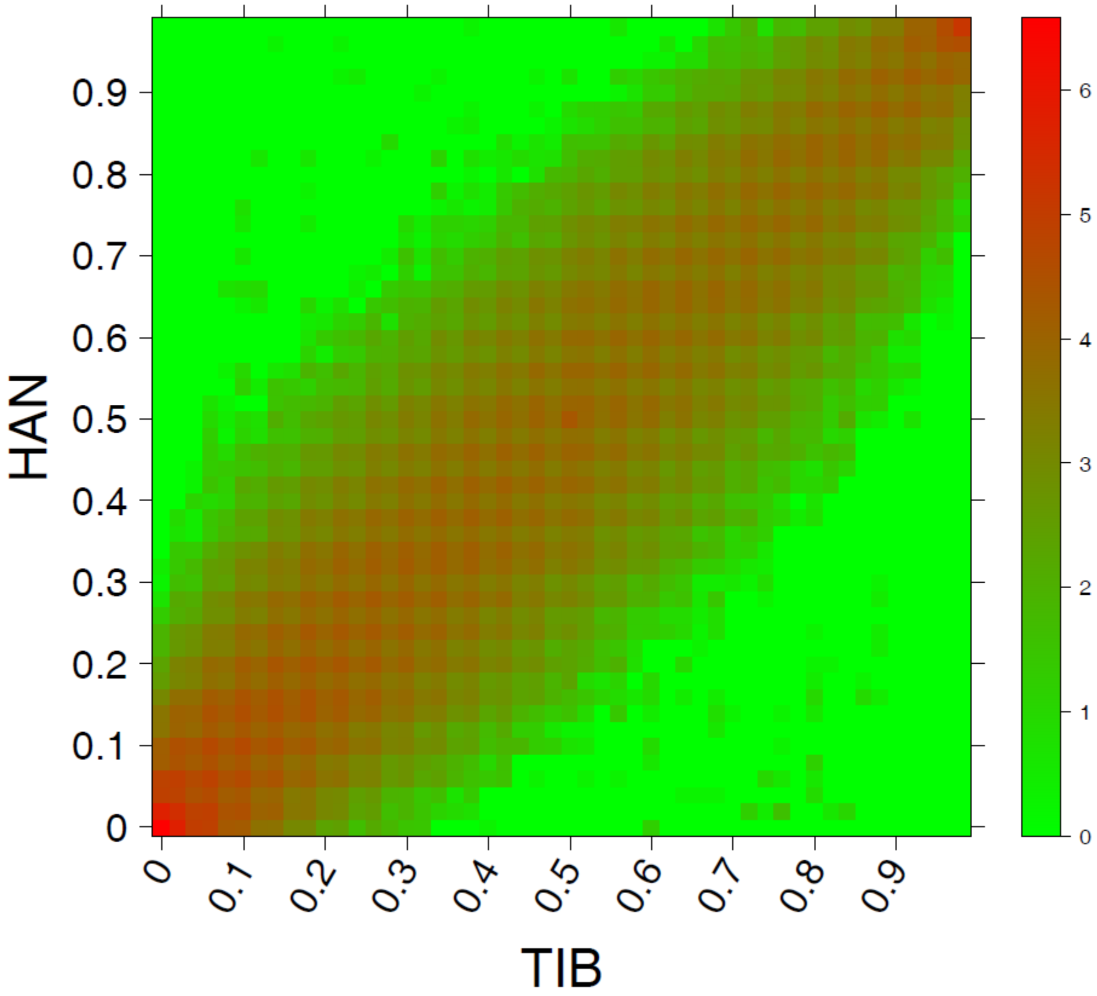

Figure S7 Correlation of derived allele frequency between TIB and HAN.

For each polymorphic SNV we calculated the allele frequency of the derived allele in both TIB and HAN. The color in each bin represents the number of SNPs that display each given set of derived allele frequencies. The light green regions show that very few SNVs with derived allele in high frequency in TIB but in low frequency in HAN, or verse versa. The red regions show that there are many SNVs that have similar low derived allele frequencies in both TIB and HAN. The right legend with color coding for the number of SNVs in log scale with warmer color indicating larger number of SNVs.

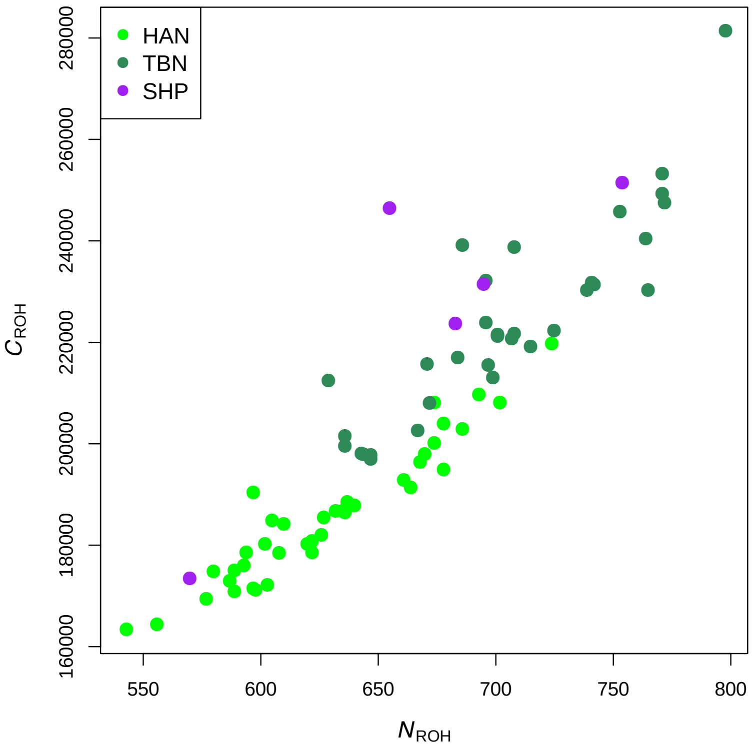

Figure S8 **Distribution of number of ROH (*N*$_{\mathrm{ROH}}$) per individual and cumulative length of ROH (*C*$_{\mathrm{ROH}}$) in TBN, SHP and HAN.**

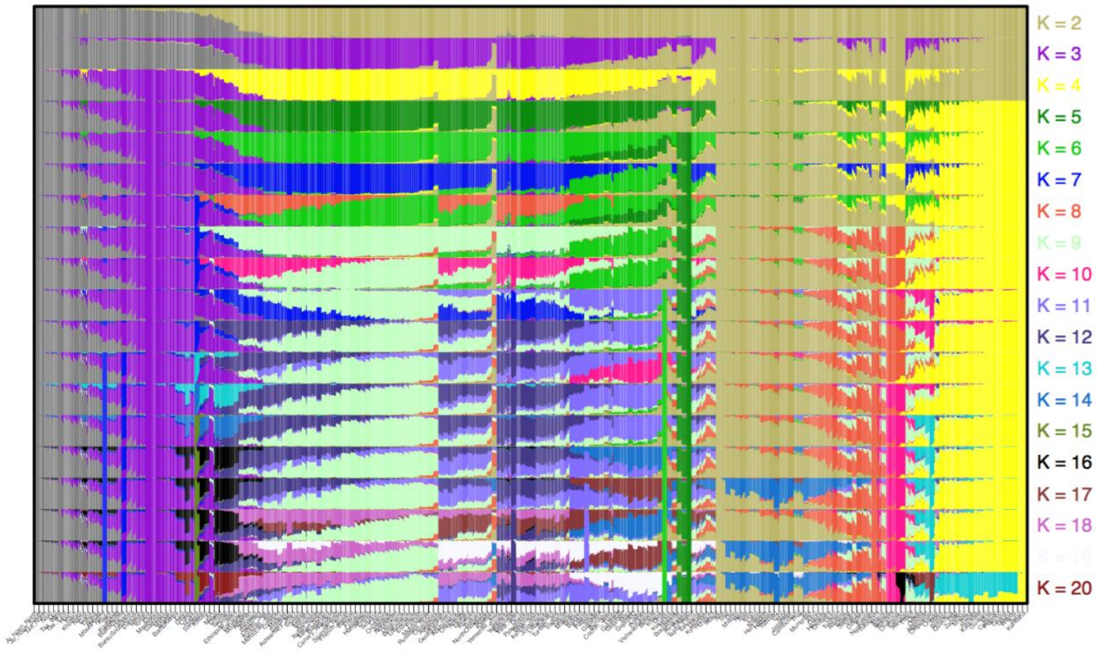

Figure S9 Summary plot of genetic admixture of worldwide populations.

The results of individual admixture proportions estimated from 288,979 SNPs with genotype data available for 38 TIB, 39 HAN and 2,345 HuOrigin samples. Each individual is represented by a single vertical line broken into K colored segments, with lengths proportional to each of the K inferred clusters, where the K values are shown in the legend. The predefined populations IDs are presented at the bottom of the plot, the geographic region of each population is shown on the top of the plot.

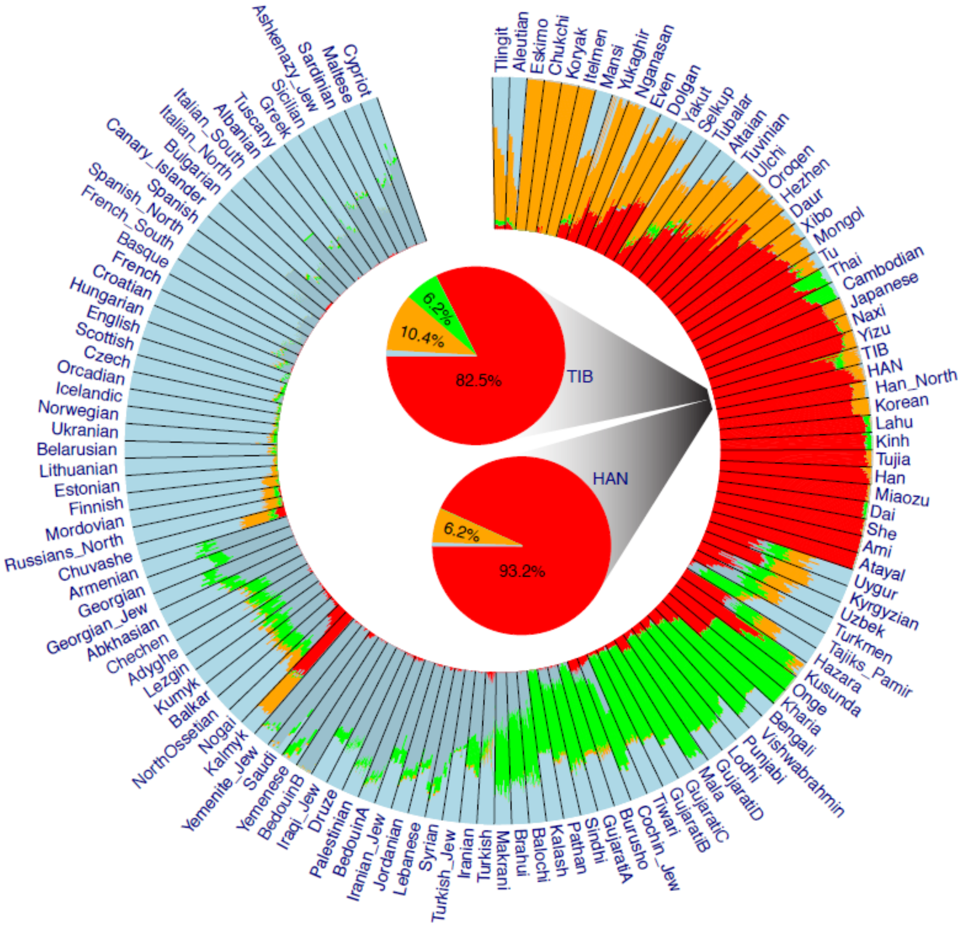

Figure S10 **Summary plot of genetic admixture (with Oceanian and American samples excluded).**

The results of individual admixture proportions estimated from 592,799 autosomal SNPs with genotype data available for 38 TIB, 39 HAN and 2,345 HuOrigin samples (African, American and Oceanian samples were not included). Each individual is represented by a single line broken into K = 5 colored segments, with lengths proportional to the K = 5 inferred clusters. The population IDs are presented outside of the circle of the plot. The results of population level admixture of TIB and HAN are further summarized and displayed in the two pie-charts in the center of the circle plot with admixture proportion denoted in percentage.

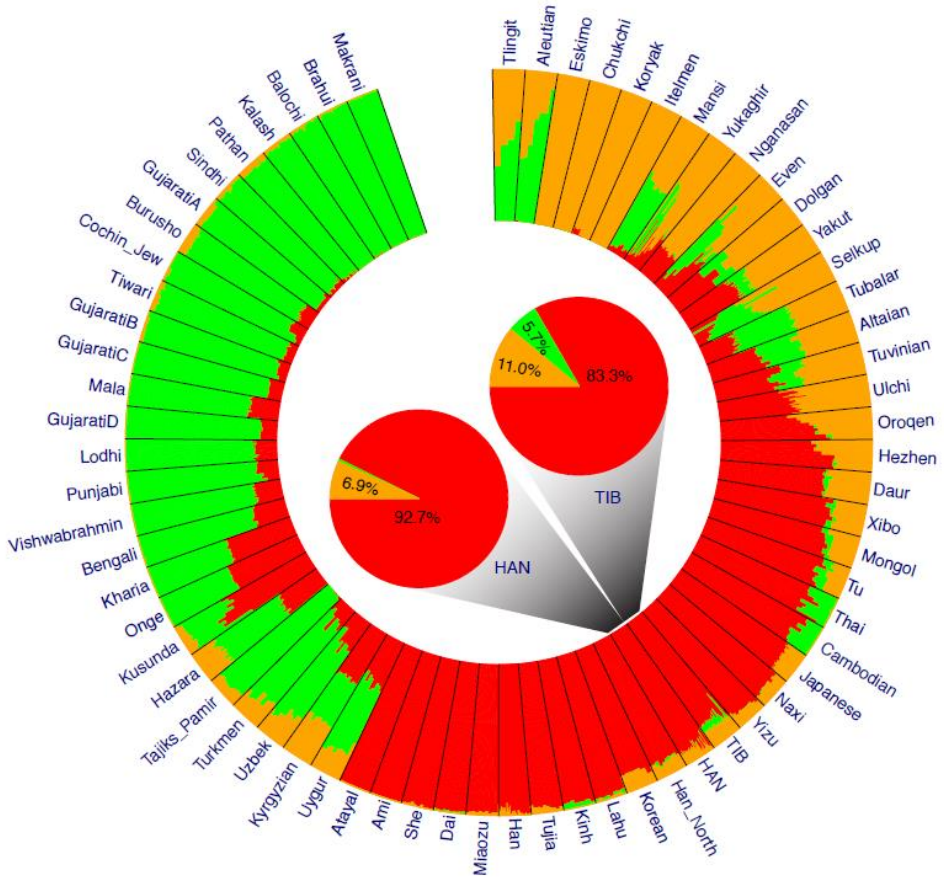

Figure S11 Summary plot of genetic admixture (with western Eurasian samples excluded).

The results of individual admixture proportions estimated from 592,799 SNPs with genotype data available for 38 TIB, 39 HAN and 2,345 HuOrigin samples (African, American, Oceanian and western Eurasian samples were not included). Each individual is represented by a single line broken into K = 3 colored segments, with lengths proportional to the K = 3 inferred clusters. The population IDs are presented outside of the circle of the plot. The results of population level admixture of TIB and HAN are further summarized and displayed in the two pie-charts in the center of the circle plot with admixture proportion denoted in percentage.

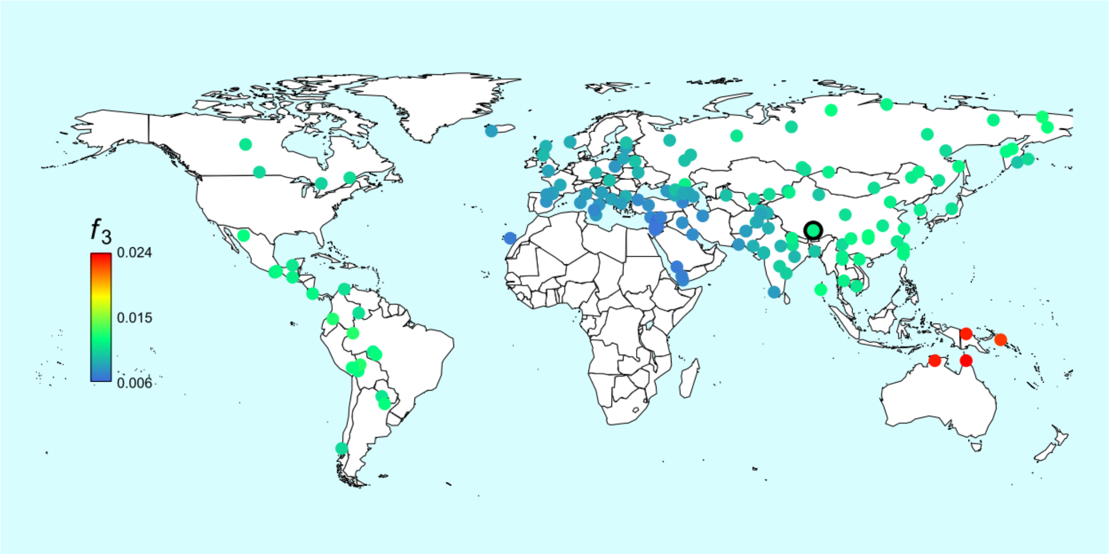

Figure S12 Heat map of outgroup f_3 statistics testing for relatedness of non-African populations to Neanderthal.

The values showing here are based on computing f_3(*African*; Neanderthal, *X*), where *X* is a non-African population. A warmer color indicates a higher f_3 value, which suggests greater derived allele sharing between population *X* and Neanderthal genome. TIB is highlighted with a black circle. For list of population labels, see Table S4.

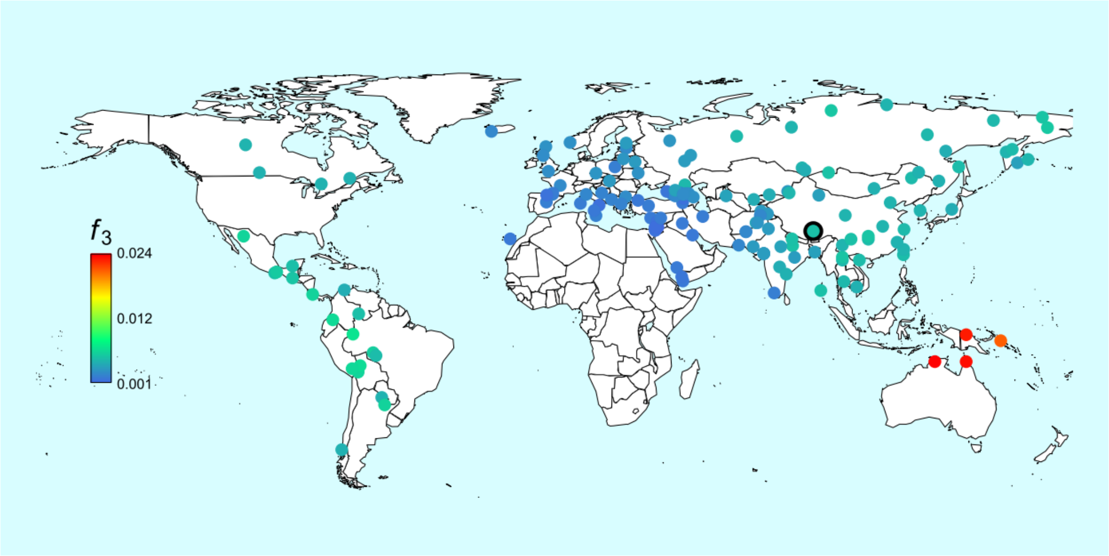

Figure S13 Heat map of outgroup f_3 statistics testing for relatedness of non-African populations to Denisovan.

The values showing here are based on computing f_3(*African*; Denisovan, *X*), where *X* is a non-African population. A warmer color indicates a higher f_3 value, which suggests greater derived allele sharing between population *X* and Denisovan genome. TIB is highlighted with a black circle. For list of population labels, see Table S4..

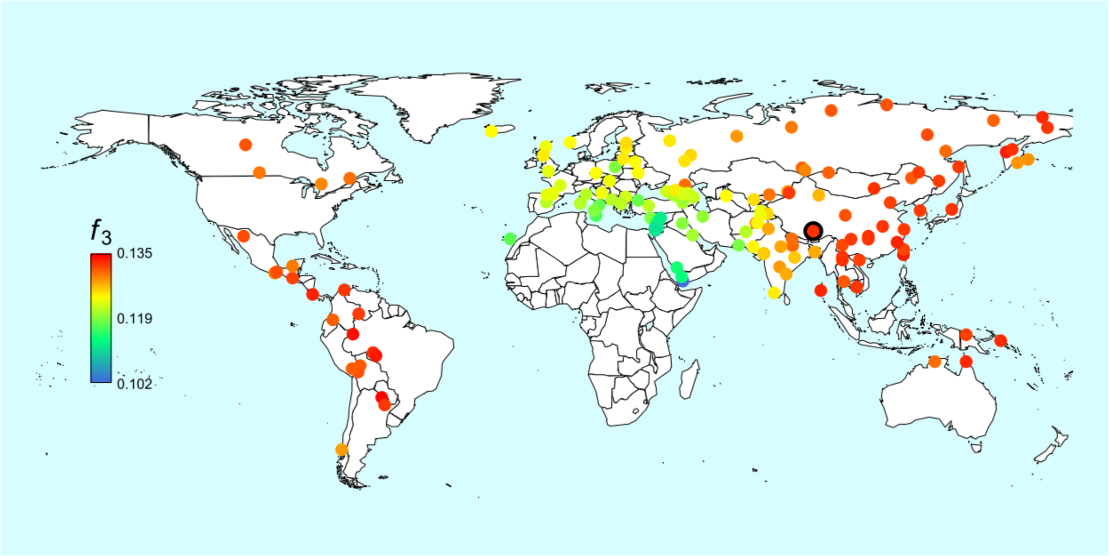

Figure S14 **Heat map of outgroup f_3 statistics testing for relatedness of non-African populations to Ust'-Ishim.**

The values showing here are based on computing f_3(*African*; Ust'-Ishim, *X*), where *X* is a non-African population. A warmer color indicates a higher f_3 value, which suggests greater derived allele sharing between population *X* and Ust'-Ishim genome. TIB is highlighted with a black circle. For list of population labels, see Table S4.

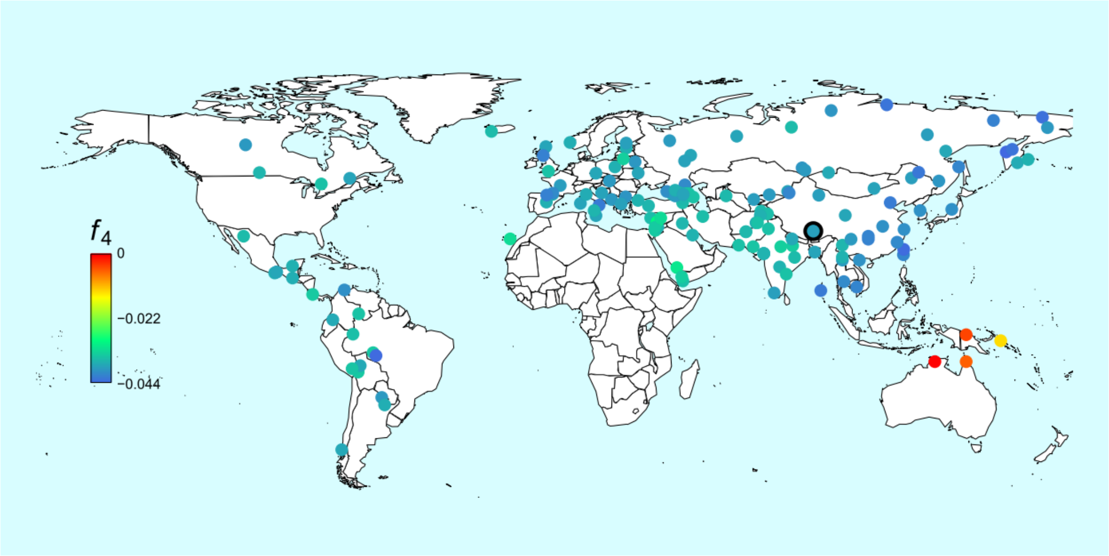

Figure S15 Heat map of outgroup f_4 statistics (*D*-test) testing for relatedness of non-African populations to Neanderthal and Denisovan.

The values showing here are based on computing f_4(*African*, *X*; Neanderthal, Denisovan), where *X* is a non-African population. A colder color indicates a more negative f_4 value, which suggests greater derived allele sharing between population *X* and Neanderthal genome than that between *X* and Denisovan genome. TIB is highlighted with a black circle. For list of population labels, see Table S4.

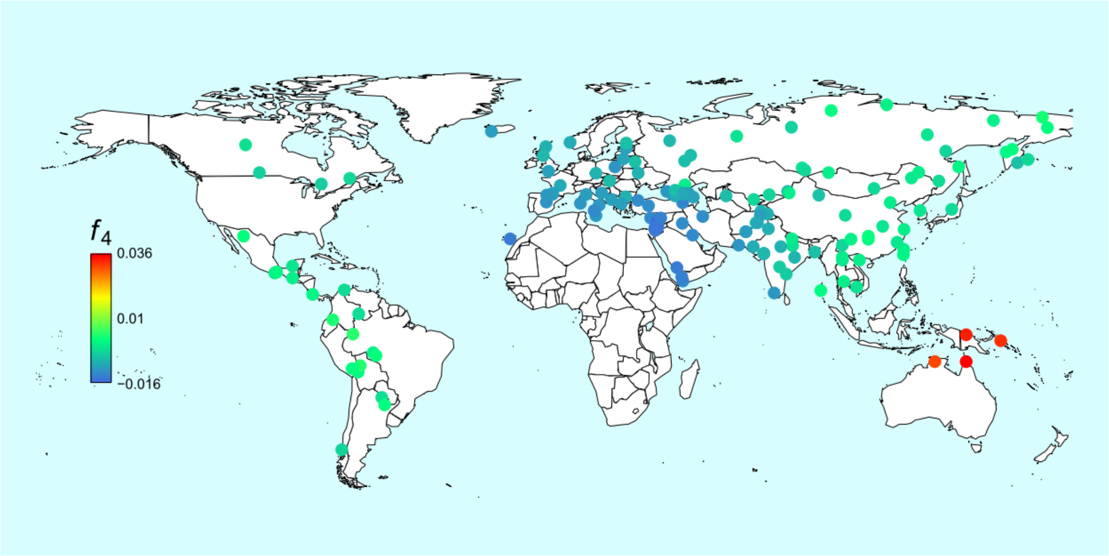

Figure S16 Heat map of f_4 statistics (*D*-test) testing for relatedness of Neanderthal to TIB and the other non-African populations.

The values showing here are based on computing f_4(*African*, Neanderthal; TIB, *X*), where *X* is a non-African population. A warmer color indicates a more positive f_4 value, which suggests greater derived allele sharing between population *X* and Neanderthal genome than that between TIB and Neanderthal. For list of population labels, see Table S4.

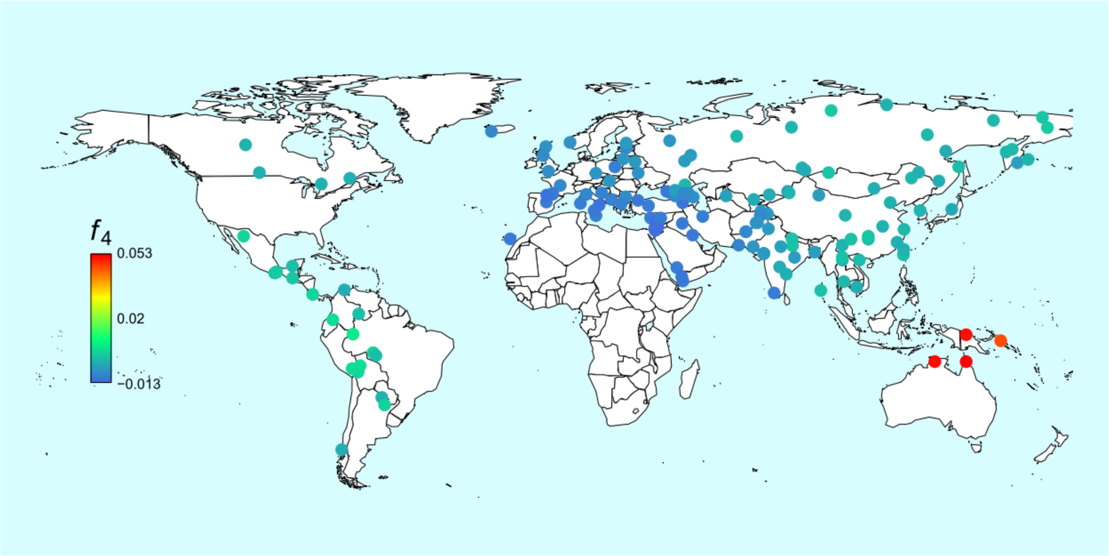

Figure S17 Heat map of f_4 statistics (*D*-test) testing for relatedness of Denisovan to TIB and the other non-African populations.

The values showing here are based on computing f_4(*African*, Denisovan; TIB, *X*), where *X* is a non-African population. A warmer color indicates a more positive f_4 value, which suggests greater derived allele sharing between population *X* and Denisovan genome than that between TIB and Denisovan. For list of population labels, see Table S4.

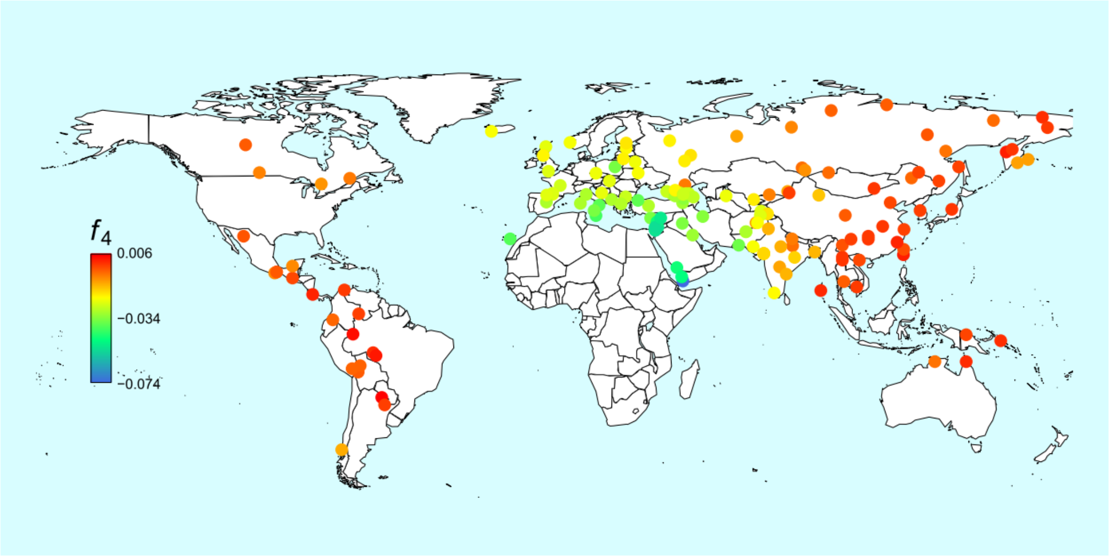

Figure S18 Heat map of f_4 statistics (*D*-test) testing for relatedness of Ust'-Ishim to TIB and the other non-African populations.

The values showing here are based on computing f_4(*African*, Ust'-Ishim; TIB, *X*), where *X* is a non-African population. A warmer color indicates a more positive f_4 value, which suggests greater derived allele sharing between population *X* and Ust'-Ishim genome than that between TIB and Ust'-Ishim. For list of population labels, see Table S4.

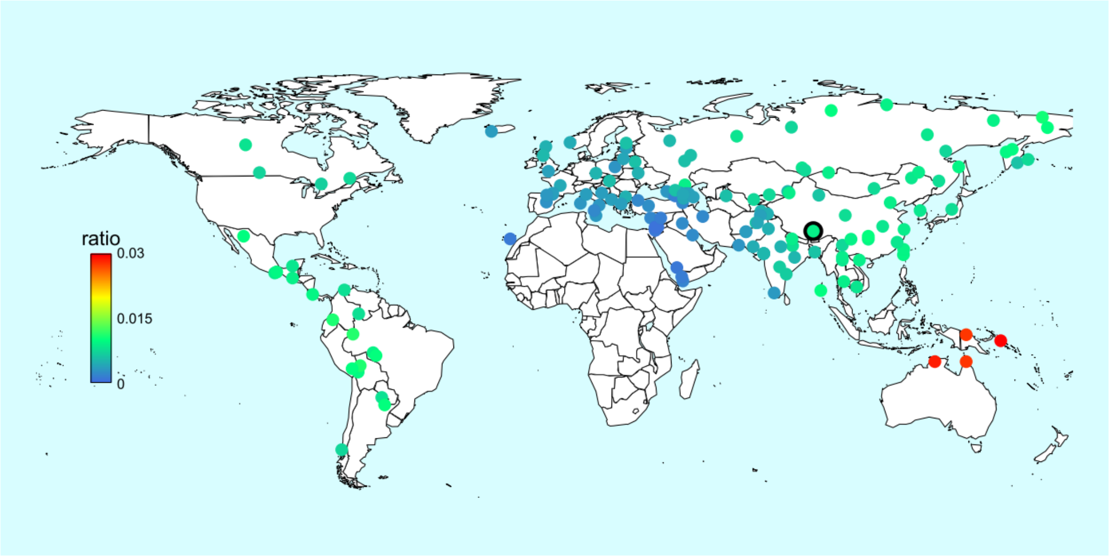

Figure S19 **Heat map of Neanderthal gene introgression ratio (based on Formula 1, see Material and Methods).**

The graded heat key represents magnitude of Neanderthal gene introgression ratio, which is estimated by Formula 1. A warmer color indicates a higher ancient introgression level from Neanderthal. TIB is highlighted with a black circle. For list of population labels, see Table S4.

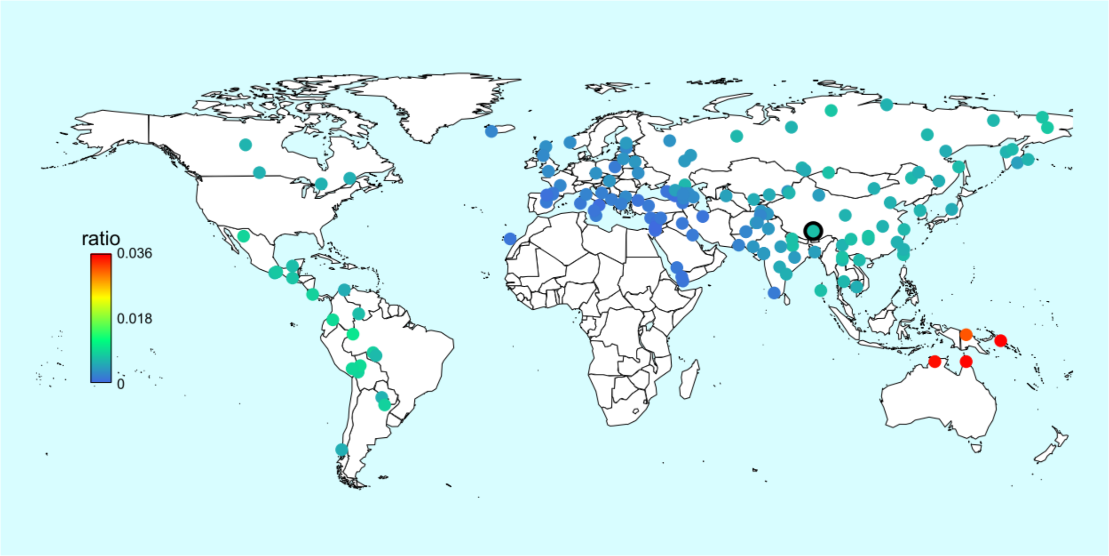

Figure S20 **Heat map of Denisovan gene introgression ratio (based on Formula 1, see Material and Methods).**

The graded heat key represents magnitude of Denisovan gene introgression ratio, which is estimated by Formula 1. A warmer color means a higher ancient introgression level from Denisovan. TIB is highlighted with a black circle. For list of population labels, see Table S4.

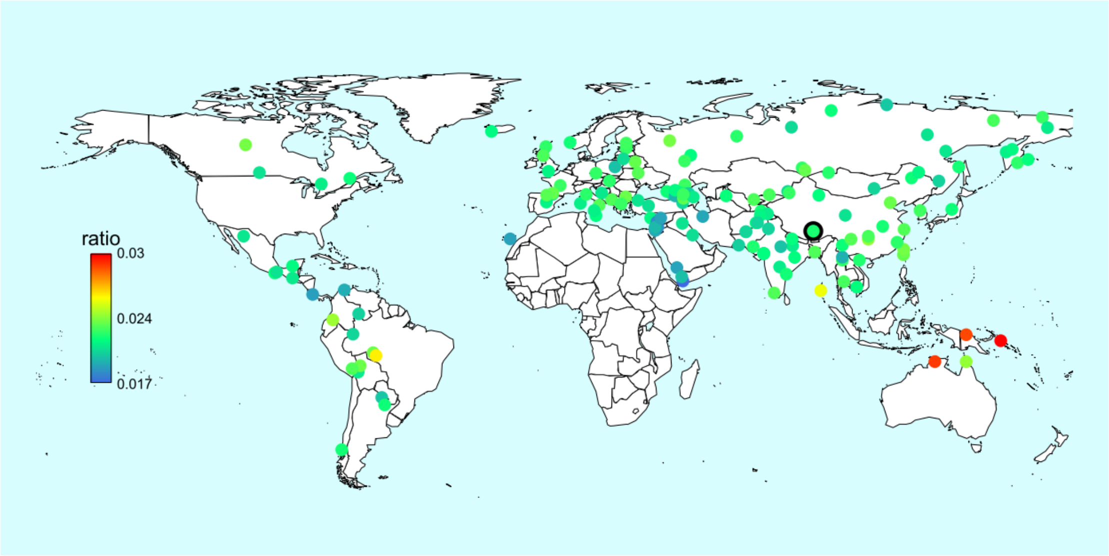

Figure S21 Heat map of Neanderthal gene introgression ratio (based on Formula 2, see Material and Methods).

The graded heat key represents magnitude of Neanderthal gene introgression ratio, which is estimated by Formula 2. A warmer color means a higher ancient introgression level from Neandertal. TIB is highlighted with a black circle. For list of population labels, see Table S4.

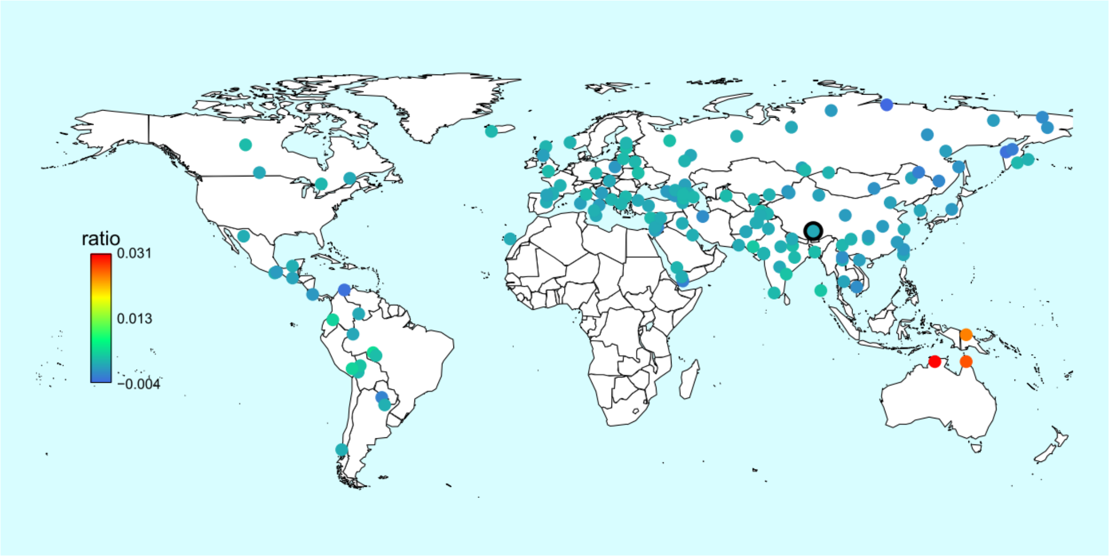

Figure S22 **Heat map of Denisovan gene introgression ratio (based on Formula 2, see Material and Methods).**

The graded heat key represents magnitude of Denisovan gene introgression ratio, which is estimated by Formula 2. A warmer color means a higher ancient introgression level from Denisovan. TIB is highlighted with a black circle. For list of population labels, see Table S4.

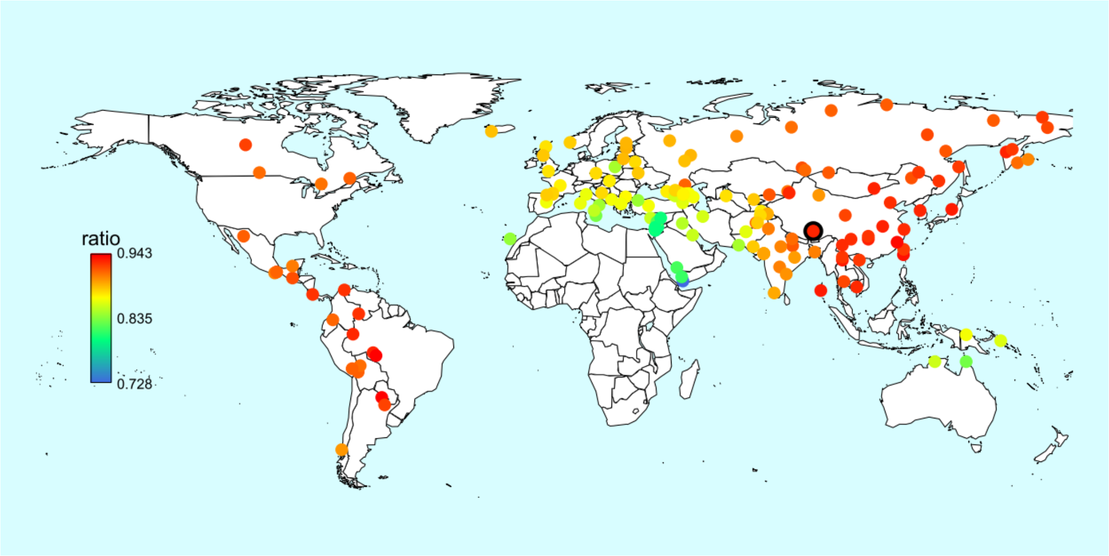

Figure S23 Heat map of Ust'-Ishim gene introgression ratio (based on Formula 2, see Material and Methods).

The graded heat key represents magnitude of Ust'-Ishim gene introgression ratio, which is estimated by Formula 2. A warmer color means a higher ancient introgression level from Ust'-Ishim. TIB is highlighted with a black circle. For list of population labels, see Table S4.

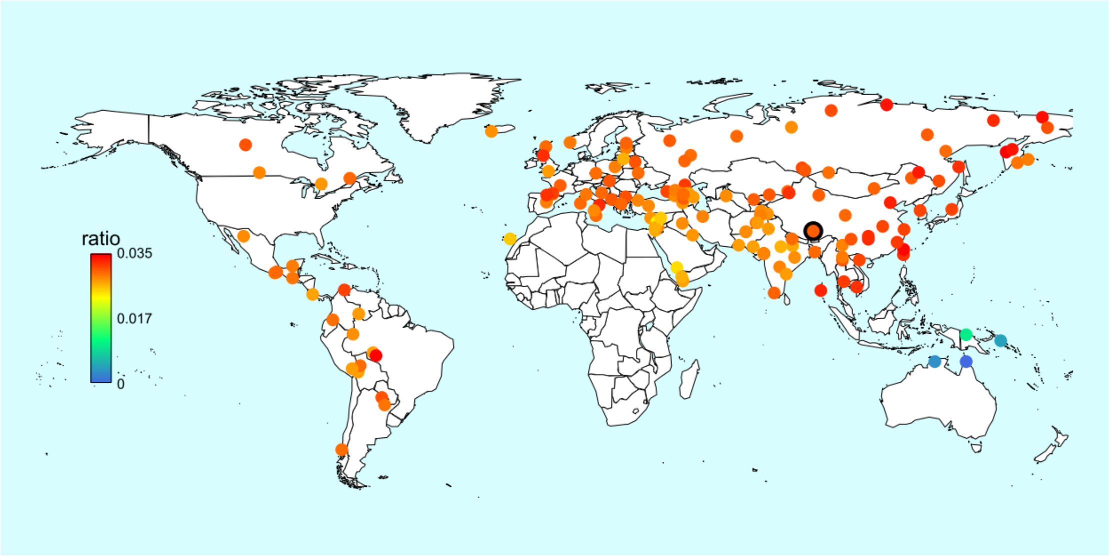

Figure S24 Heat map of Neanderthal gene introgression ratio (based on Formula 3, see Material and Methods).

The graded heat key represents magnitude of Neanderthal gene introgression ratio, which is estimated by Formula 3. A warmer color means a higher ancient introgression level from Neanderthal. TIB is highlighted with a black circle. For list of population labels, see Table S4.

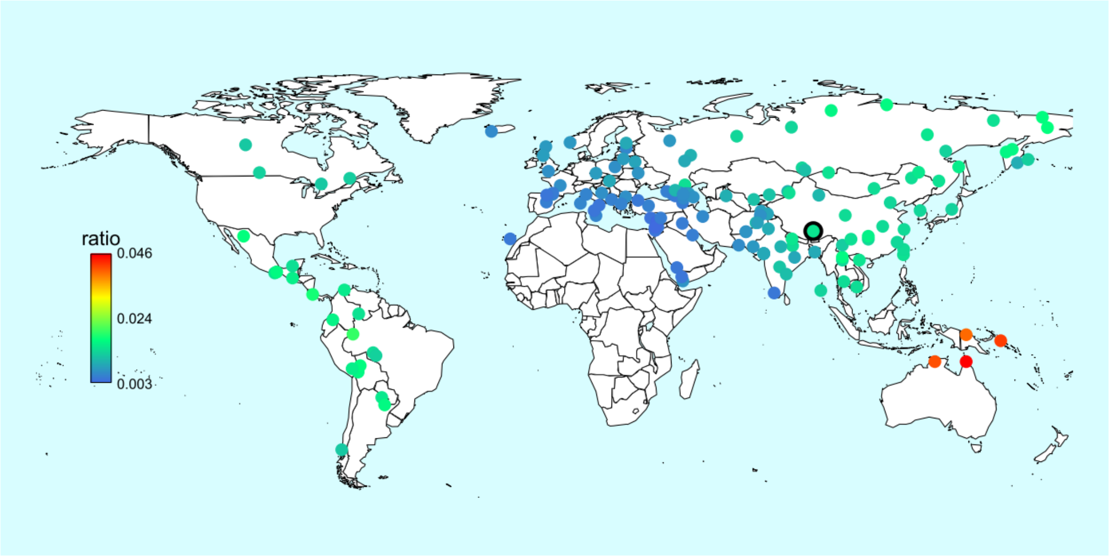

Figure S25 Heat map of Neanderthal gene introgression ratio (based on Formula 4, see Material and Methods).

The graded heat key represents magnitude of Neanderthal gene introgression ratio, which is estimated by Formula 4. A warmer color means a higher ancient introgression level from Neanderthal. TIB is highlighted with a black circle. For list of population labels, see Table S4.

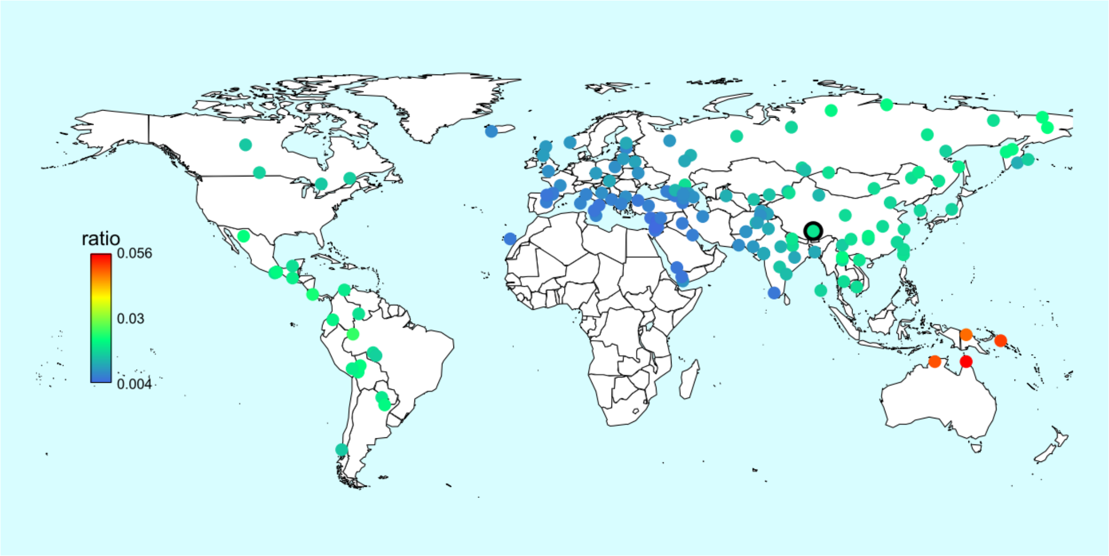

Figure S26 **Heat map of Denisovan gene introgression ratio (based on Formula 4, see Material and Methods).**

The graded heat key represents magnitude of Denisovan gene introgression ratio, which is estimated by Formula 4. A warmer color means a higher ancient introgression level from Denisovan. TIB is highlighted with a black circle. For list of population labels, see Table S4.

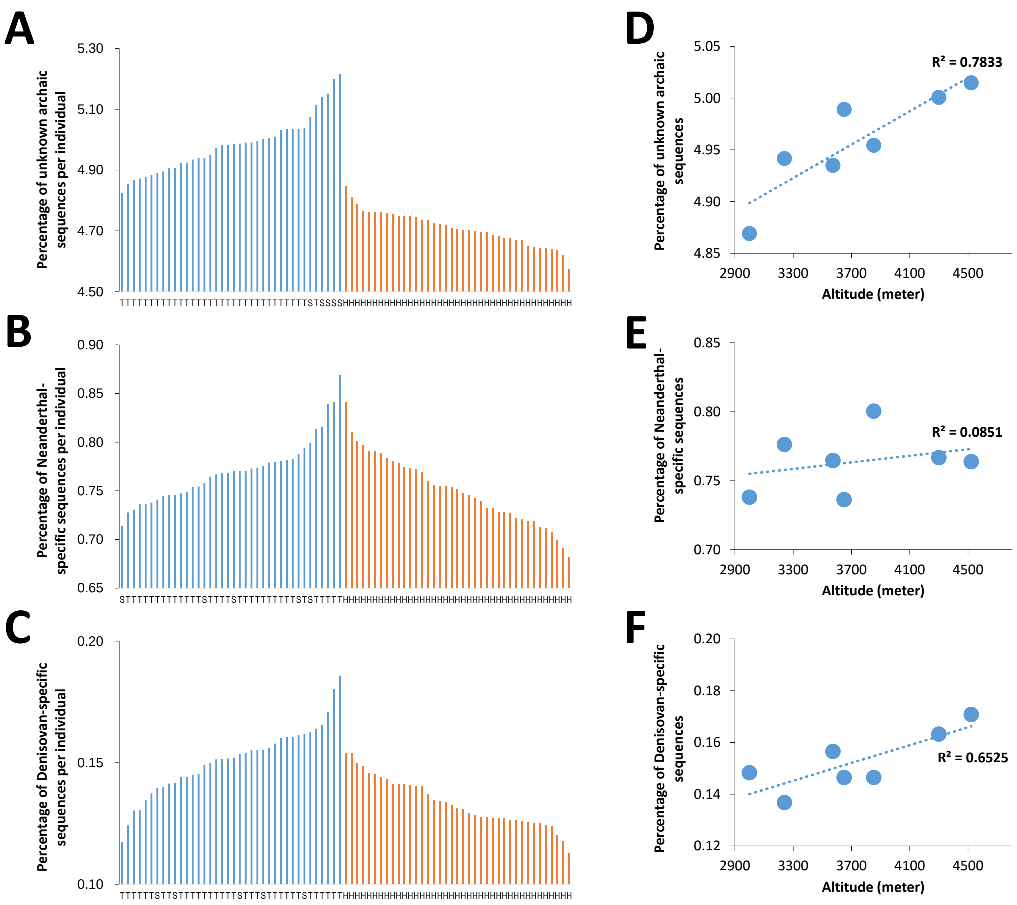

Figure S27 Distribution of non-modern-human-specific sequences among individuals and correlation with altitude. a. A comparison of the percentage of unknown non-modern human sequences per individual between TIB (4.98 ±0.09) and HAN (4.71 ± 0.06) had significant differences ($p < 10^{-5}$, permutation test); **b**. A comparison of percentage of Neanderthal-specific sequences per individual between TIB (0.77 ± 0.03) and HAN (0.75 ± 0.04) had no significant differences ($p = 0.01$, permutation test); **c**. A comparison of the percentage of Denisovan-specific sequences per individual between TIB (0.15 ±0.01) and HAN (0.13 ±0.01) had slightly significant differences ($p < 10^{-5}$, permutation test); the statistical significance (*p* value) was obtained by permutation tests that were repeated 100,000 times; on the horizontal axis of **a**, **b**, and **c**, T denotes Tibetan, S for Sherpa, and H for Han Chinese; **d**. Correlation between the average proportion of unknown non-modern human sequences in regional Tibetan populations and altitude (Pearson $R^2 = 0.783$; $p = 0.019$); **e**. Correlation between the average proportion of Neanderthal-specific sequences per individual and altitude (Pearson $R^2 = 0.085$; $p = 0.575$); **f**. Correlation between the average proportion of Denisovan-specific sequences per individual and altitude (Pearson $R^2 = 0.653$; $p = 0.052$).

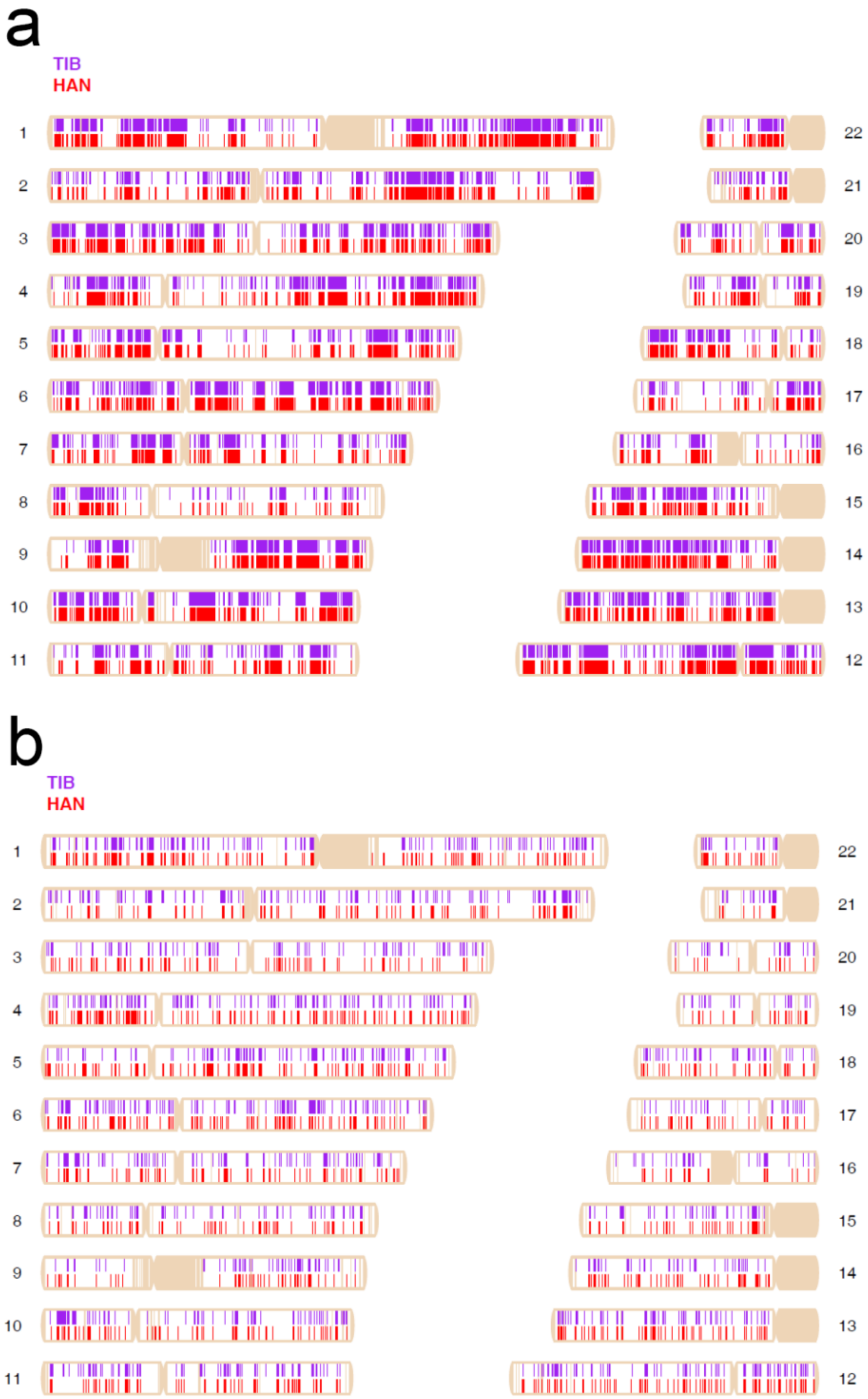

Figure S28 Identified archaic lineages in the TIB and HAN genomes. a. Spatial distributions of Neanderthal sequences in TIB and HAN. **b**. Spatial distributions of Denisovan sequences in TIB and HAN. Please note both panels (a and b) show results on population level, thus the positions indicated by purple and red colors represent an union of those observed in 38 TIB individuals and 39 HAN individuals, respectively.

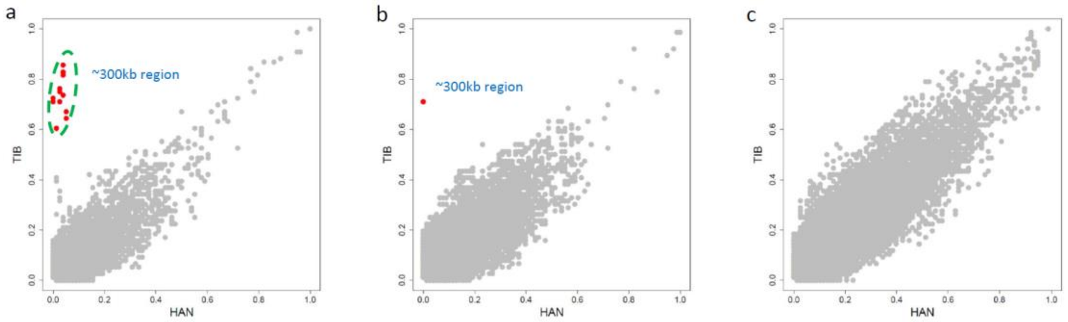

Figure S29 Frequency distribution of ancient sequences in TIB and HAN. The ancient segments in TIB and HAN were detected by ArchaicSeeker and confirmed by other two methods (see Material and Methods). Each dot represents an ancient DNA segment with the frequency in TIB and HAN indicated by vertical and horizontal axis, respectively. Dots in red represent highly-differentiated segment between TIB and HAN (frequency difference larger than 0.4). a) Comparison of the frequency distribution of Denisovan-like sequences between TIB and HAN; the highly-differentiated segments enriched in the ~300kb *EPAS1* and its downstream region which encompass the 5-SNP-motif, TED, *TMEM247*, and other genes. b) Comparison of the frequency distribution of Neanderthal-like sequences between TIB and HAN; the highly-differentiated segment is also in the ~300kb region and its downstream. Neanderthal shared alleles with Denisovan in this segment. c) Comparison of the frequency distribution of Ust'-Ishim-like sequences between TIB and HAN; No highly-differentiated Ust'-Ishim-like segments were found between Tibetan and Han Chinese.