iScience, Volume *27*

Supplemental information

Genomic imputation of ancient Asian populations contrasts local adaptation in pre- and post-agricultural Japan

Niall P. Cooke, Madeleine Murray, Lara M. Cassidy, Valeria Mattiangeli, Kenji Okazaki, Kenji Kasai, Takashi Gakuhari, Daniel G. Bradley, and Shigeki Nakagome

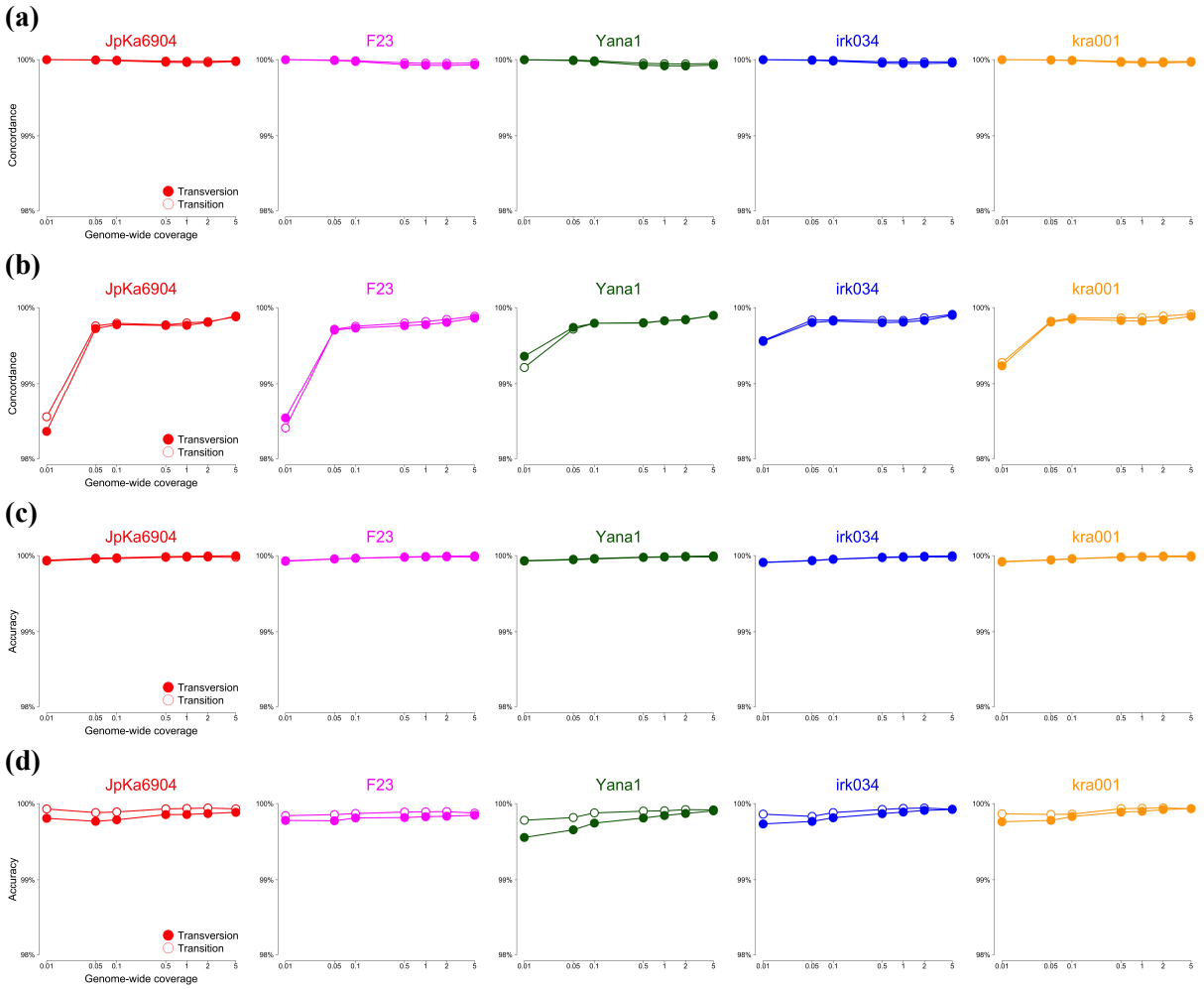

Figure S1. Performance of imputing homozygotes in five ancient individuals, related to Figure 1. The ability to impute ancient Asian individuals is assessed by two metrics: **(a-b)** the concordance measures how many of homozygous sites carrying **(a)** reference or **(b)** alternative alleles in the original diploid genomes are successfully recovered from imputation, while **(c-d)** the accuracy quantifies how many of the sites that are imputed as homozygotes with **(c)** reference or **(d)** alternative alleles are validated with the diploid data. These metrics are calculated for seven different depths of lower coverage data downsampled from the original data (*i.e.*, 0.01, 0.05, 0.1, 0.5, 1, 2, and 5×). The lines with open circles represent transition mutations, while those with closed circles represent transversions. The ancient genomes included are as follows: two Jomon individuals (JpKa6904 and F23, represented by red and pink respectively), a 31,600-year-old Upper Palaeolithic North Siberian (Yana1 represented by green), and two Siberian individuals from the Baikal Region and Krasnoyarsk Krai (irk034 and kra001, represented by blue and orange respectively).

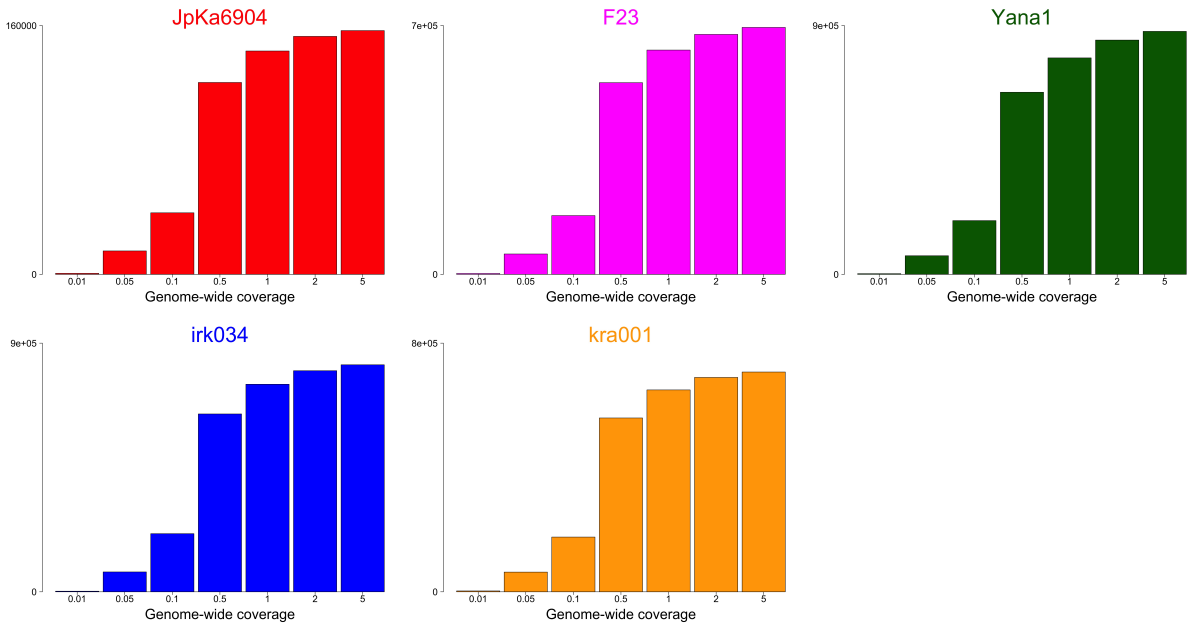

Figure S2. Overlapping heterozygous sites produced by imputation, related to Figure 1. Bar plots show the total number of heterozygous sites that reach a genotype probability of above 99% while also having a corresponding call in the diploid data. Results are presented for each downsampled coverage, with different colours representing each of the five tested individuals.

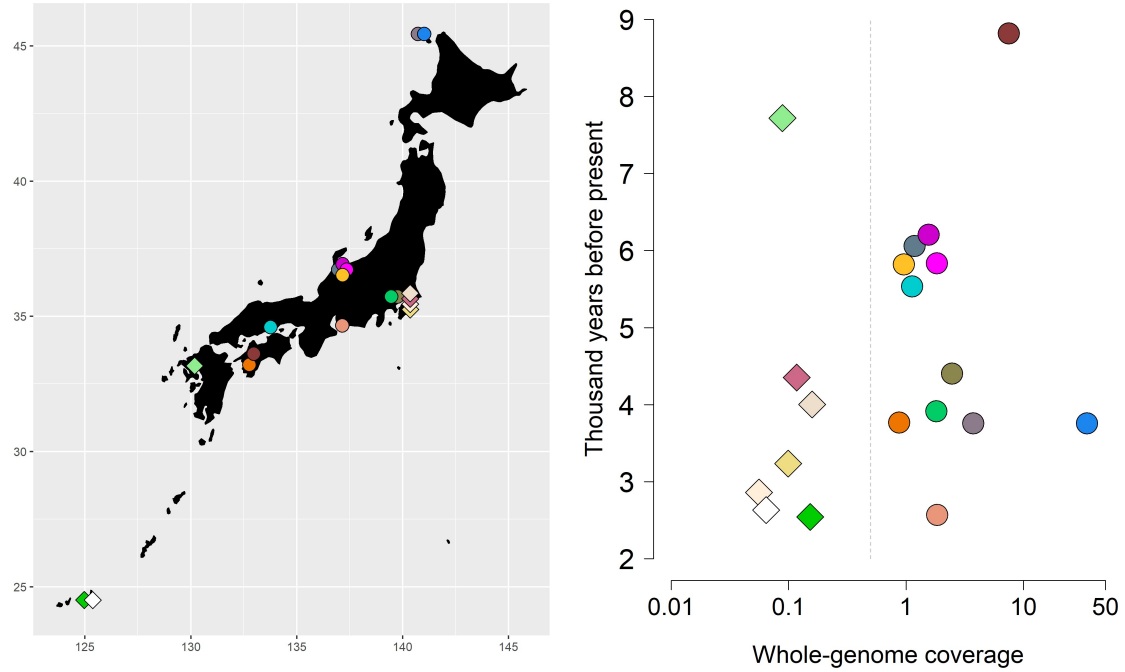

Figure S3. Sampling locations, dates, and genome coverages of Jomon individuals analysed in this study, related to Figure 2. Archaeological sites are marked with circles if the coverage is above 0.5× or diamonds otherwise. Each individual is plotted with whole-genome coverage on the *x*-axis and median age (years before present) on the *y*-axis.

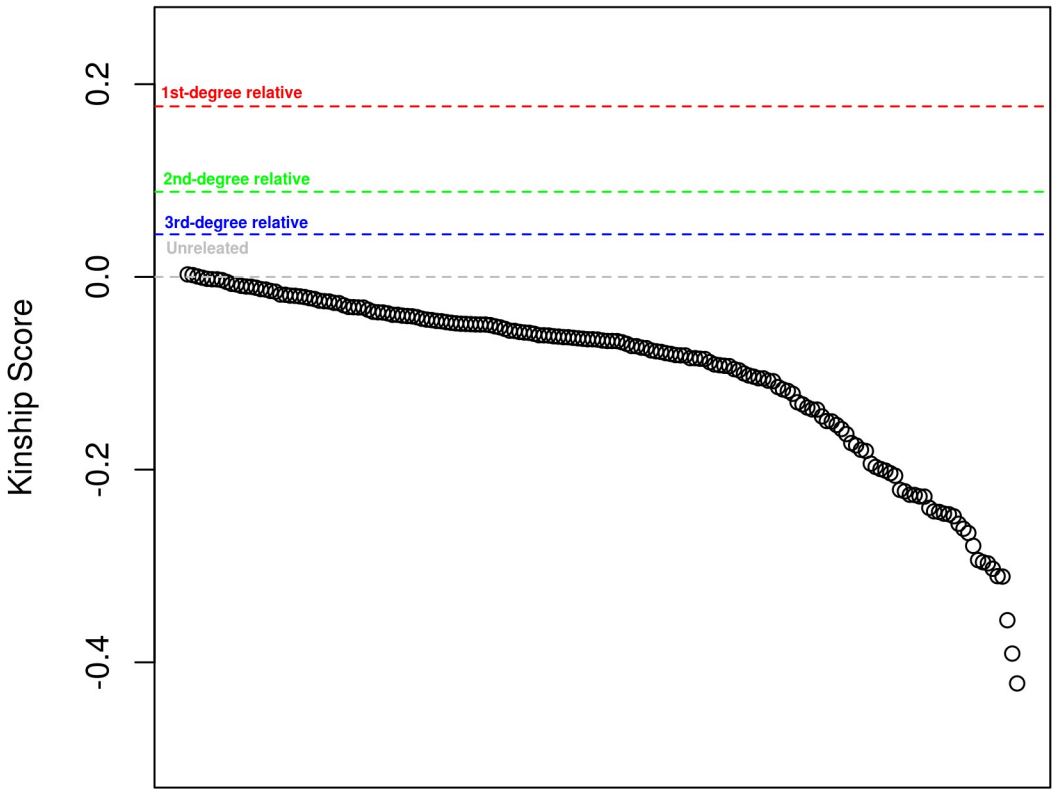

Figure S4. KING results for kinship analysis of imputed Jomon, related to Figure 2. None of the 172 combinations of the 19 imputed Jomon genomes tested reaches a high enough kinship score of above 0.0442, which is defined as a threshold for third-degree relatives. The overwhelming number of tested combinations produces negative values, supporting no genetic relatedness among these individuals. These results are consistent with the kinship analysis based on pseudo-haploid data with a subset of this population in Cooke *et al*. (2021)[1].

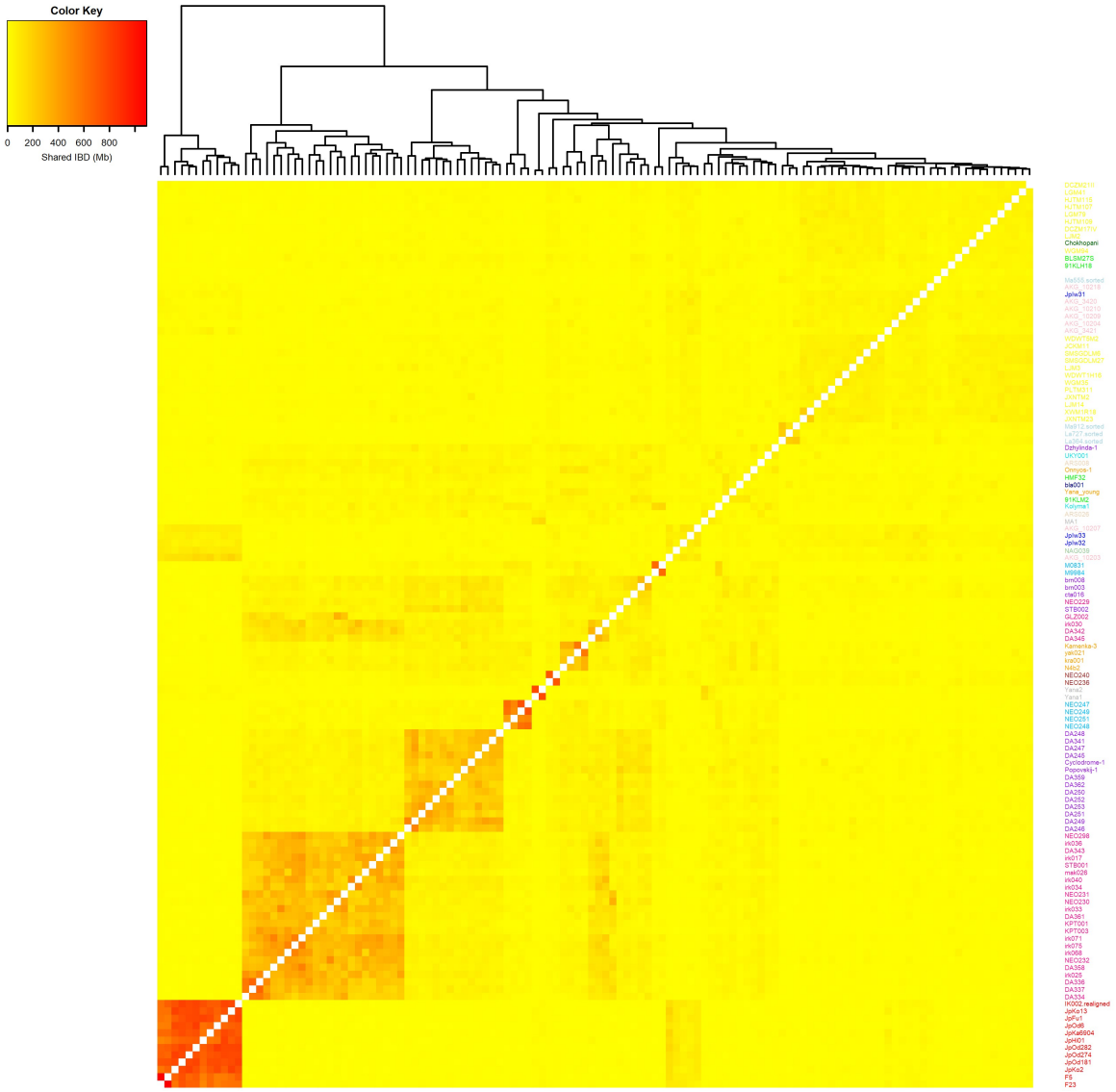

Figure S5. Heatmap of raw IBD sharing across 126 imputed ancient Asian individuals, related to Figure 2. This plot is based on the raw lengths of identity-by-descent (IBD) sharing between each combination of imputed ancient individuals with missingness of <15%. The total value for shared IBD is reflected in a colour spectrum, spanning the lowest value in yellow to the highest in red. The labels used reflect the same colours used for visualisations of principal component analysis shown in Fig. 2a (from bottom to top): Jomon – red, Baikal (Bronze Age) – deeppink, Baikal (Neolithic) – purple, Far East Siberians – deepskyblue, Upper Paleolithic – gray, Devil's Cave – brown, Yakutia – orange, West Liao River – green, Ancient Paleosiberian – cyan, Mongolia – lightbrown, Nagabaka_H – seagreen, Kofun – darkblue, Southeast Asia – lightblue, Baikal (historic) – black, Yellow River basin – yellow, Chokhopani – darkgreen, and Historic Korean - pink.

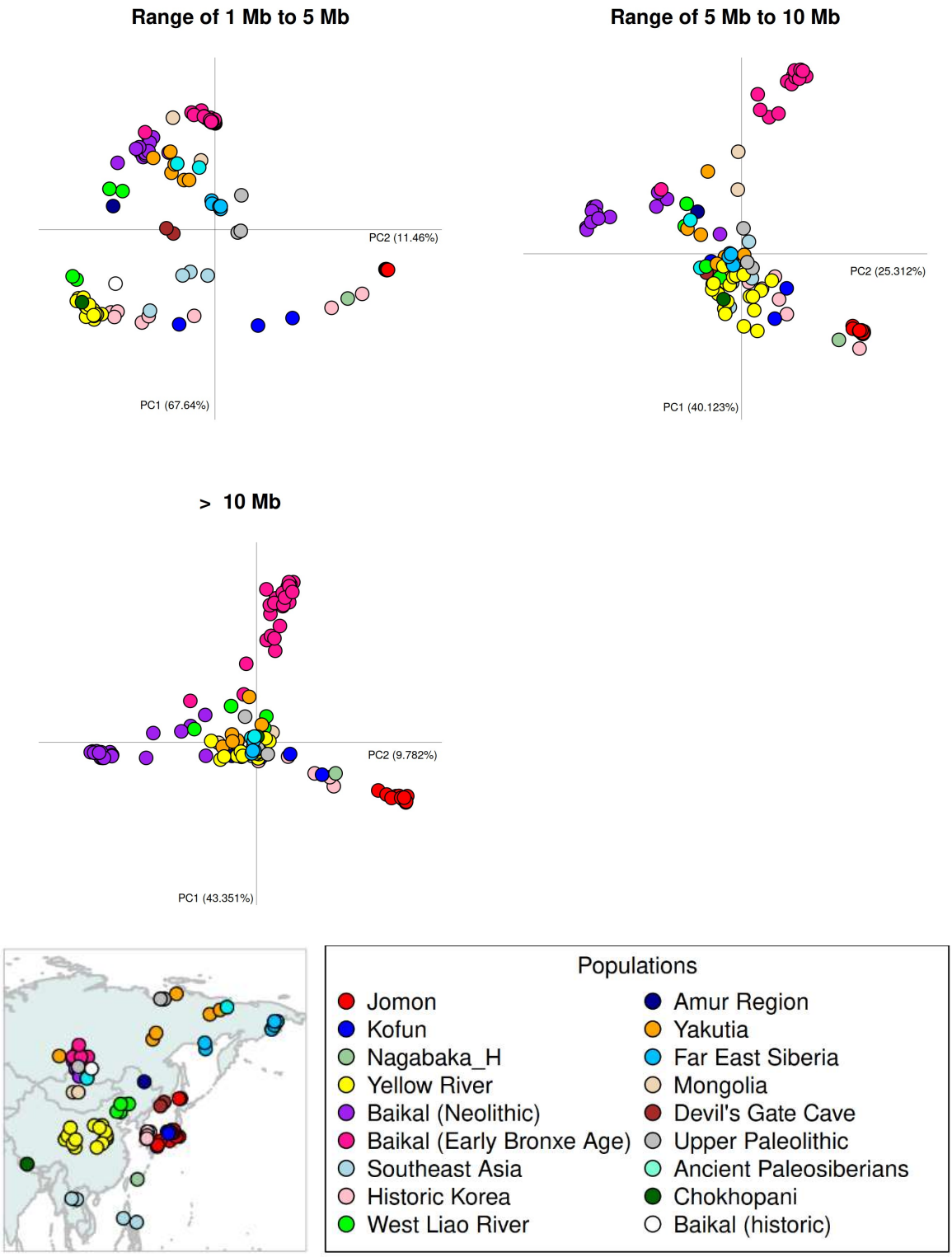

Figure S6. Principal component analysis (PCA) of ancient Asian individuals based on IBD sharing with different lengths, related to Figure 2. The PCA presented in Fig. 2a is repeated with restrictions to shared segments of different lengths (1-5 Mb, 5-10 Mb, and >10 Mb). Only individuals with <15% of missingness are included in this analysis. Different colours show different geographic or cultural backgrounds.

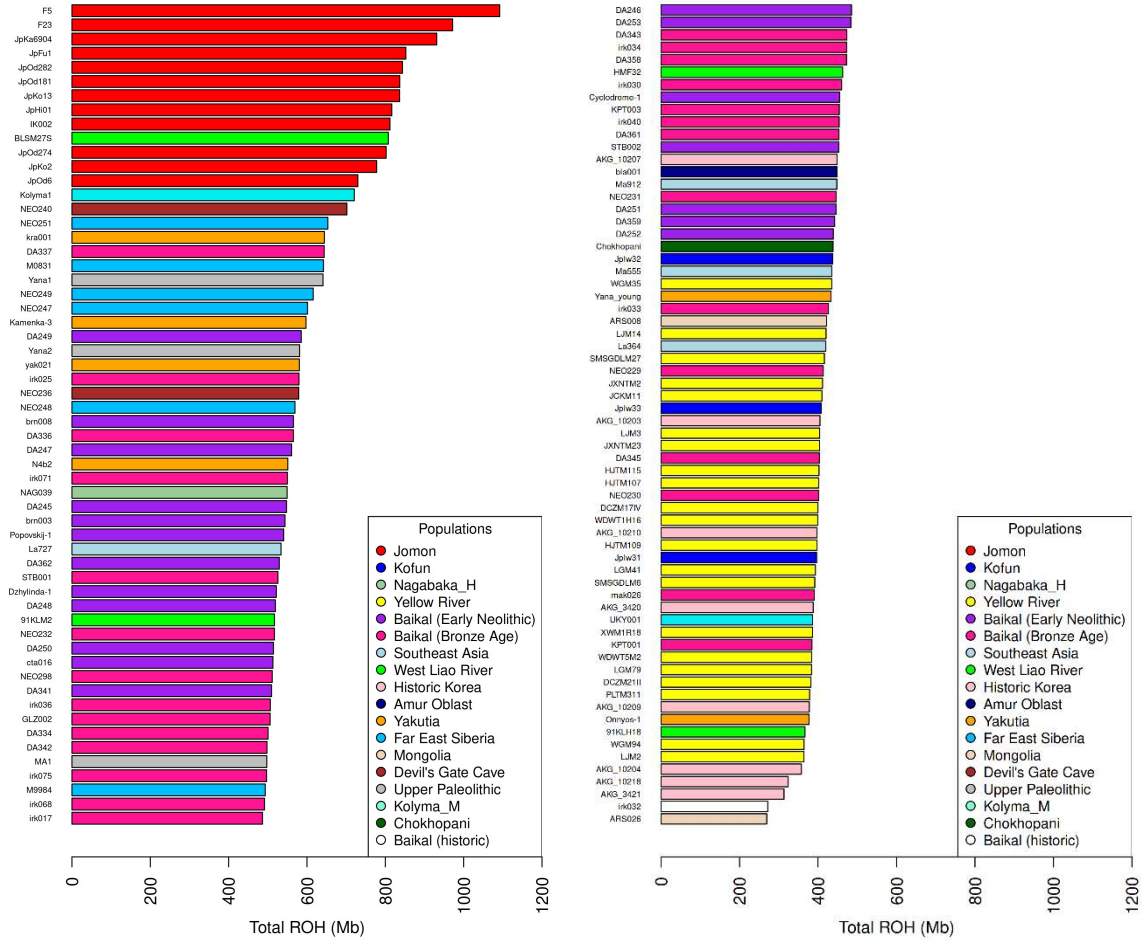

Figure S7. Total ROH lengths for 126 imputed Asian individuals, related to Figure 2. A bar plot shows the total lengths (in Mb) of stretches of chromosomes found to be homozygous. Only individuals with <15% of missingness are included in this analysis. Colours represent their geographic or cultural designation as outlined in the legend. Populations are ordered from highest amount of total ROH to lowest.

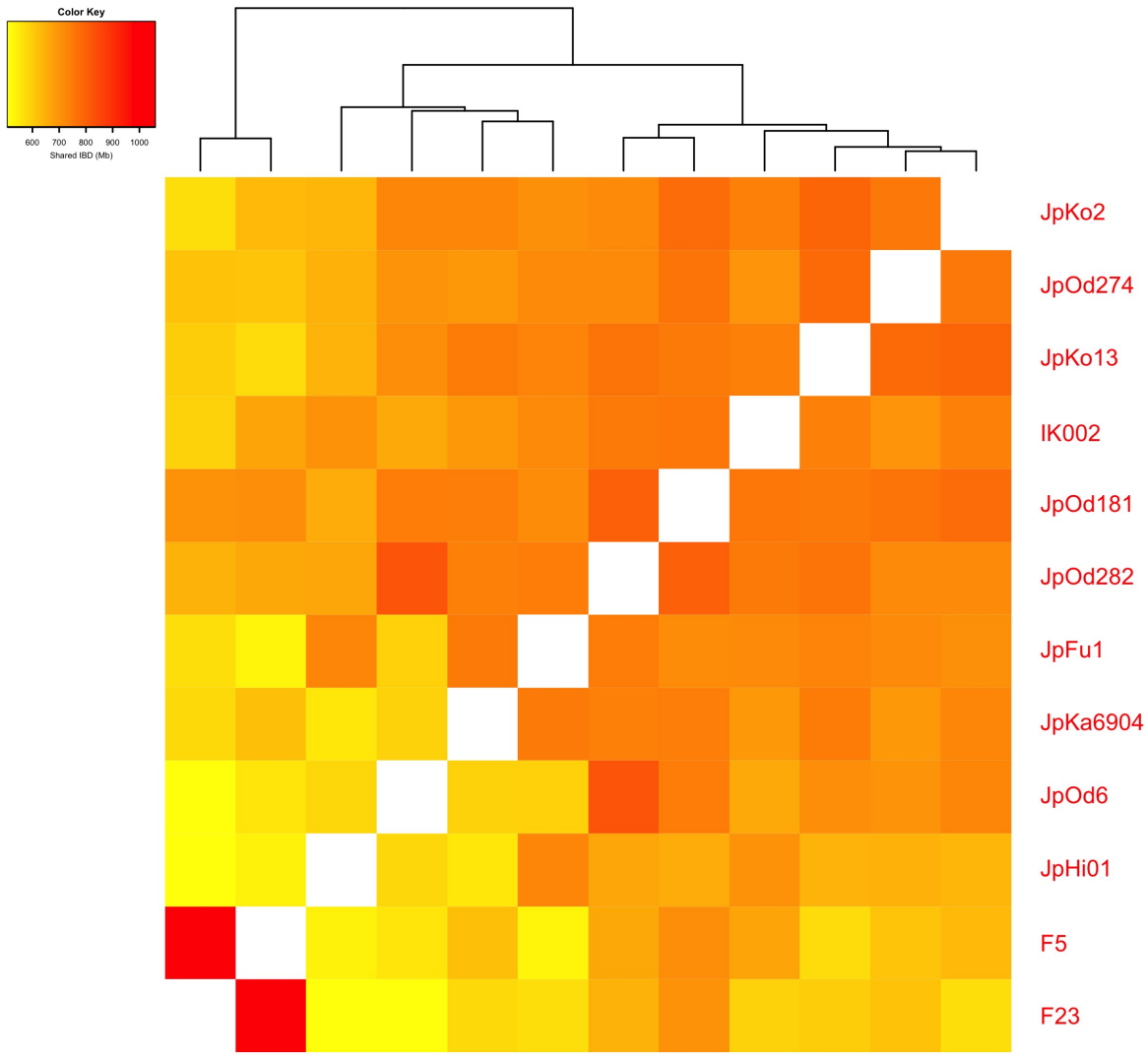

Figure S8. Heatmap of raw IBD sharing across 12 imputed Jomon individuals, related to Figure 2. This plot is based on the raw lengths of identity-by-descent (IBD) sharing between each combination of imputed Jomon individuals. Only individuals with <15% of missingness are included in this analysis. The total value for shared IBD is reflected in a colour spectrum, spanning the lowest value in yellow to the highest in red.

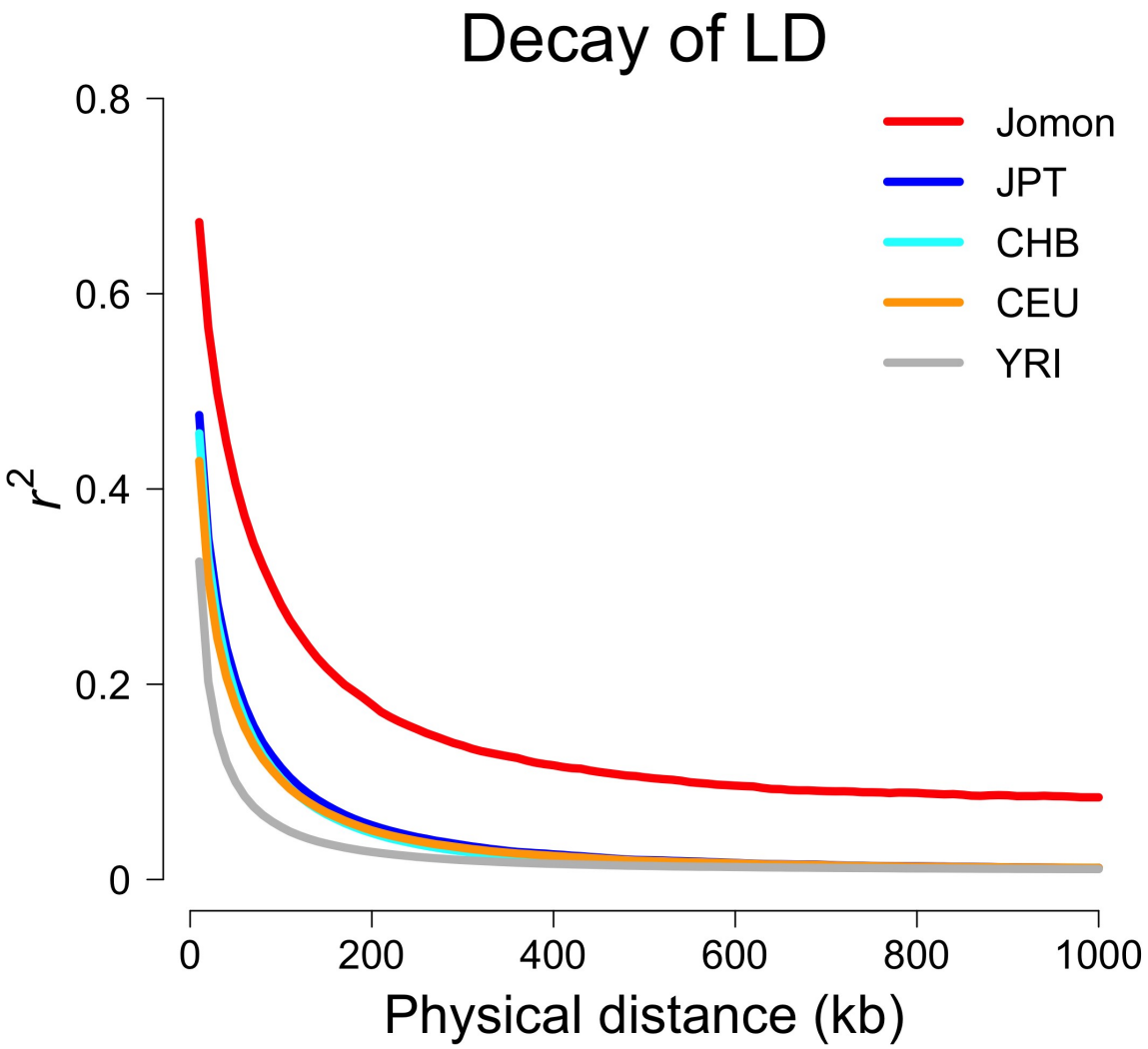

Figure S9. Decay of linkage disequilibrium across ancient and modern individuals, related to Figure 3. The linkage disequilibrium (r^2) values for Jomon (red) and modern populations (JPT, CHB, CEU, and YRI; blue, cyan, orange, and grey) are plotted against distance between pairs of SNPs, while the *y*-axis represents r^2 values.

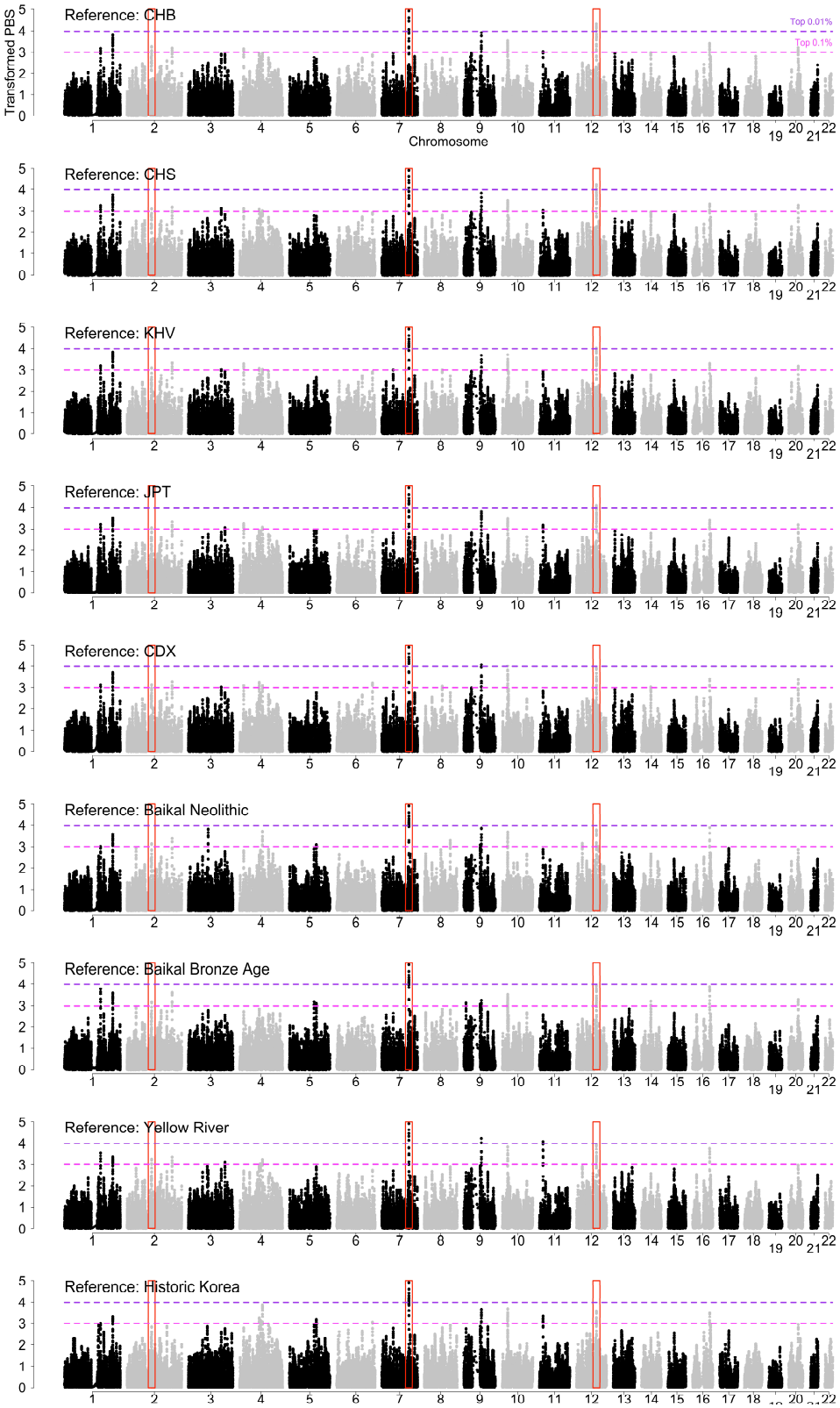

Figure S10. Genome-wide scans for positive selection in the Jomon with different reference populations, related to Figure 3. Transformed PBS on the *y*-axis represents rankings of windows (size: 80 SNPs; step between windows: 10 SNPs) by their PBS values. The dashed horizontal lines show the 99.9th (magenta) and 99.99th (purple) percentiles of the empirical distribution. Three regions are highlighted by red rectangles: chromosome 2: 109251048-109657715, chromosome 7: 120246643-120777717, and chromosome 12: 88290182-89157081.

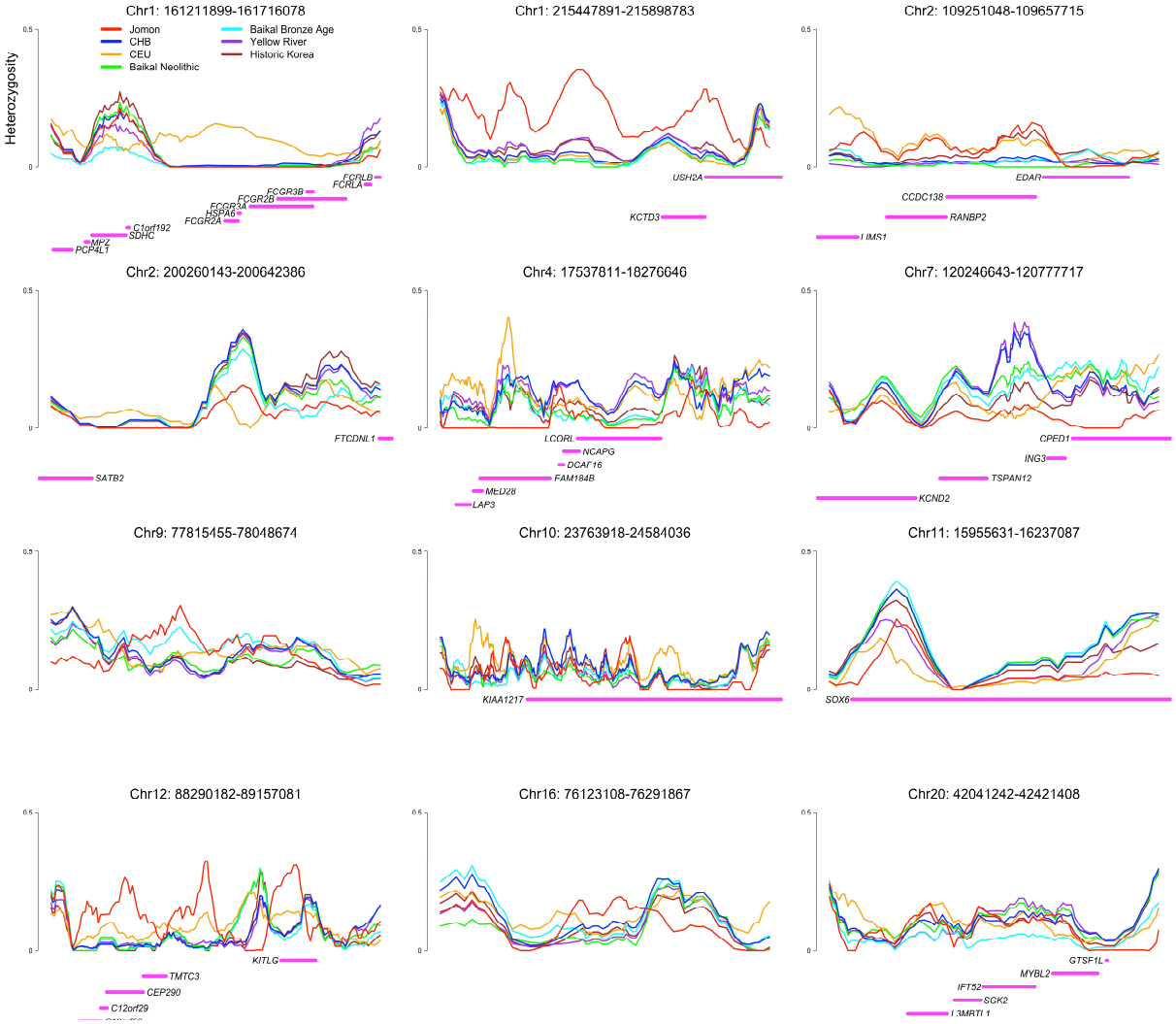

Figure S11. Heterozygosity profiles and genomic contexts of 12 regions identified from selection scans, related to Figure 3. These regions include successive windows where mean PBS values are ranked in the top 0.1% of the empirical distribution. Different colours in the plots represent different populations: Jomon (red), CHB (blue), CEU (orange), Baikal Neolithic (green), Baikal Bronze Age (cyan), Yellow River (purple), and Historic Korea (brown). The protein-coding genes present in each region are shown by solid lines with magenta.

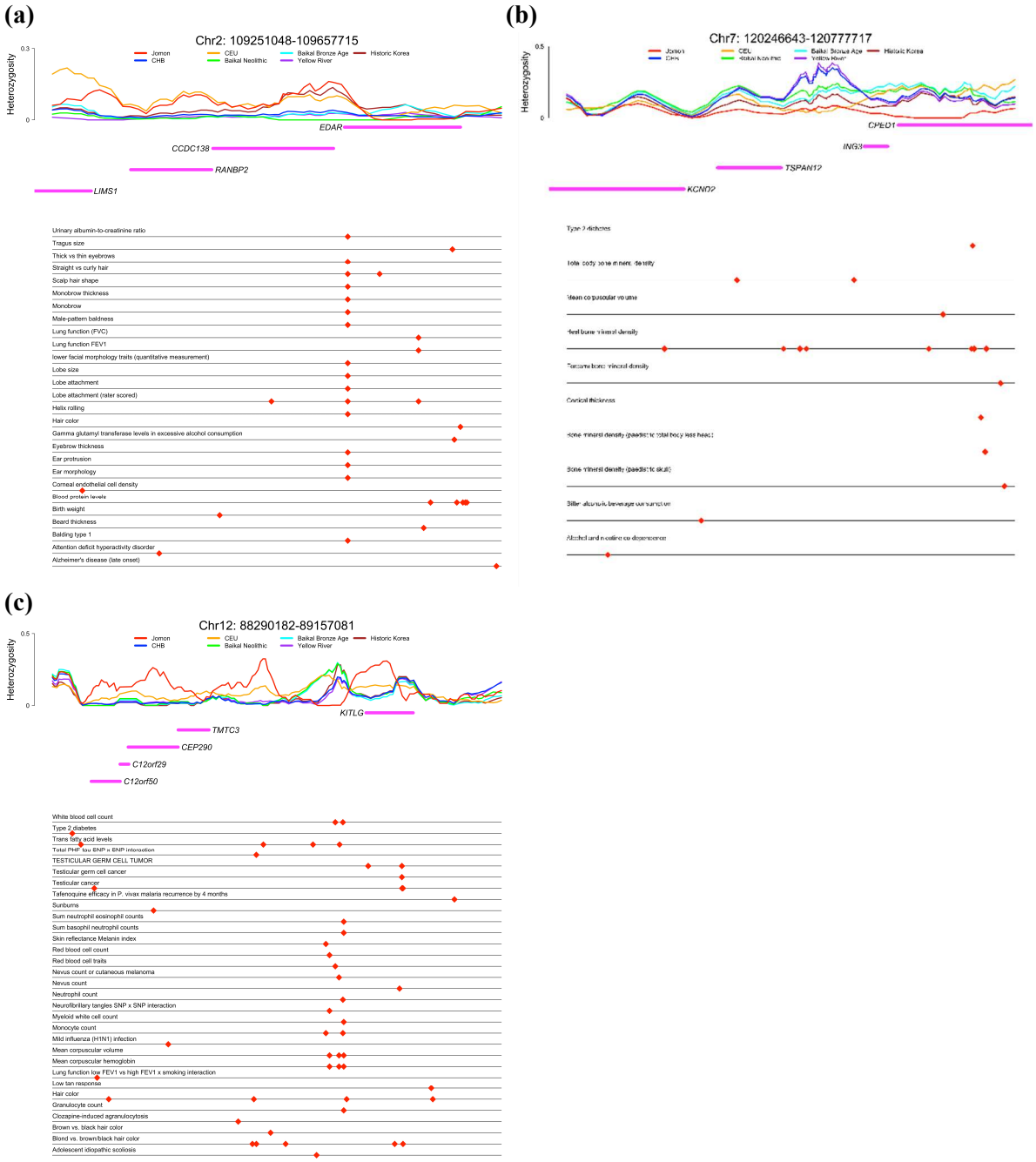

Figure S12. Phenotypic associations of the three different regions identified from selection scans: (a) chromosome 2: 109251048-109657715, (b) chromosome 7: 120246643-120777717, and (c) chromosome 12: 88290182-89157081, related to Figure 3. In each region, the top and middle plots show the heterozygosity and genomic contexts. The bottom plots represent SNP sites (red diamonds) and their phenotypic associations retrieved from BioMart in Ensembl (Ensembl Variation 104) with the filters for a human reference genome (GRCh37.p13) and "Study type: GWAS".

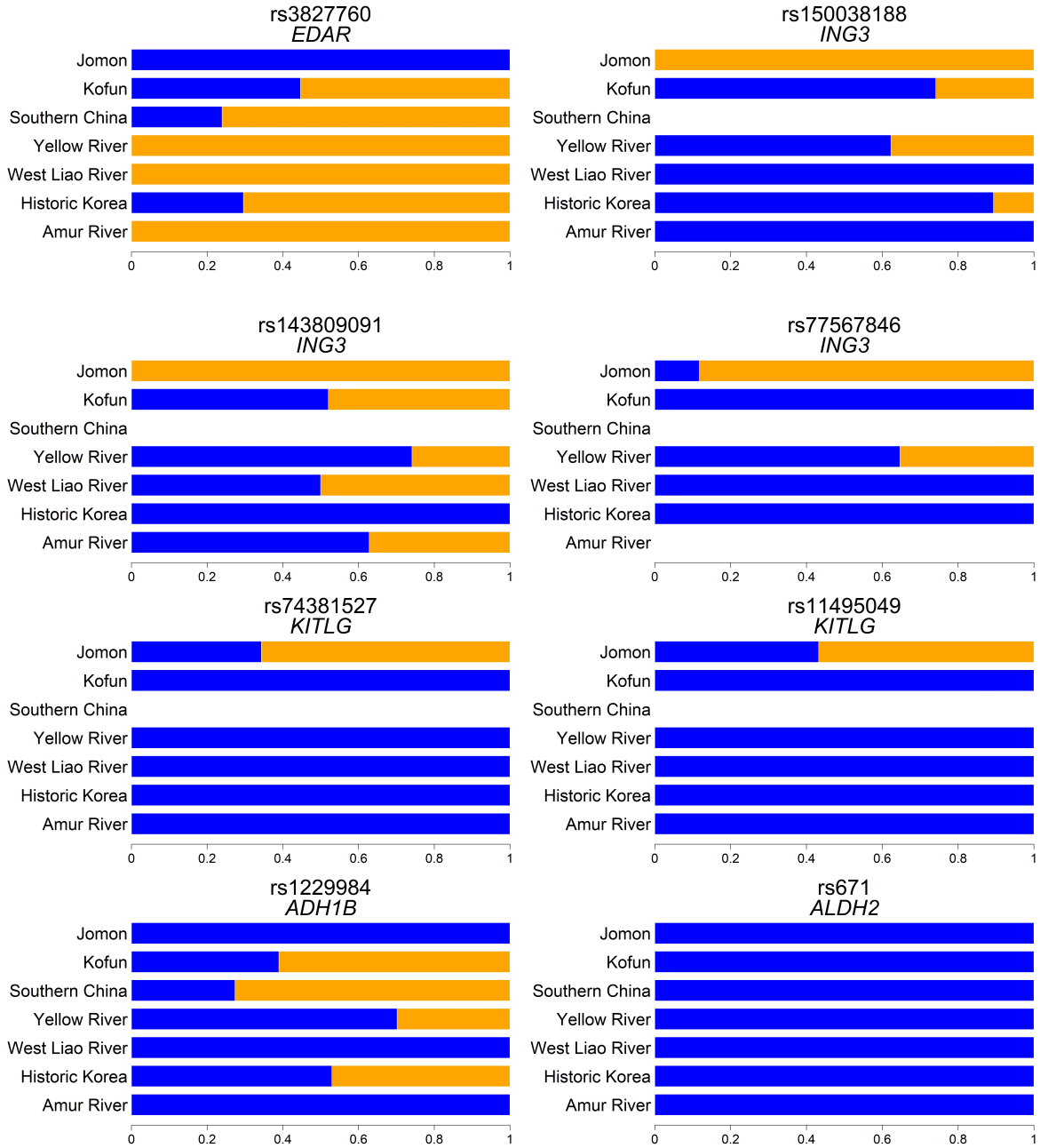

Figure S13. Allele frequencies in ancient populations estimated from raw read counts carrying ancestral or derived alleles, related to Figures 3-4. The bar plots represent the frequencies of ancestral (blue) and derived (orange) alleles in ancient populations (*i.e.*, Jomon, Kofun, Southern China, Yellow River, West Liao River, Historic Korea, and Amur River; see Table S1 for groupings). The absence of bar plots in a certain population represents that there is no aligned read covering a given site and the frequencies are not measurable. It should be noted that all individuals in Southern China are captured genomes; their allele frequencies tend to be more frequently missing compared with the other shotgun sequenced populations. Selection signals identified using imputed data are verified in the pre-imputed data, which provide evidence that the observed frequency differences in imputed genotypes are not due to imputation errors. Raw allelic counts and imputed genotypes for a subset of the populations (*i.e.*, Jomon, Kofun, and Historic Korea) are shown in Table S2.

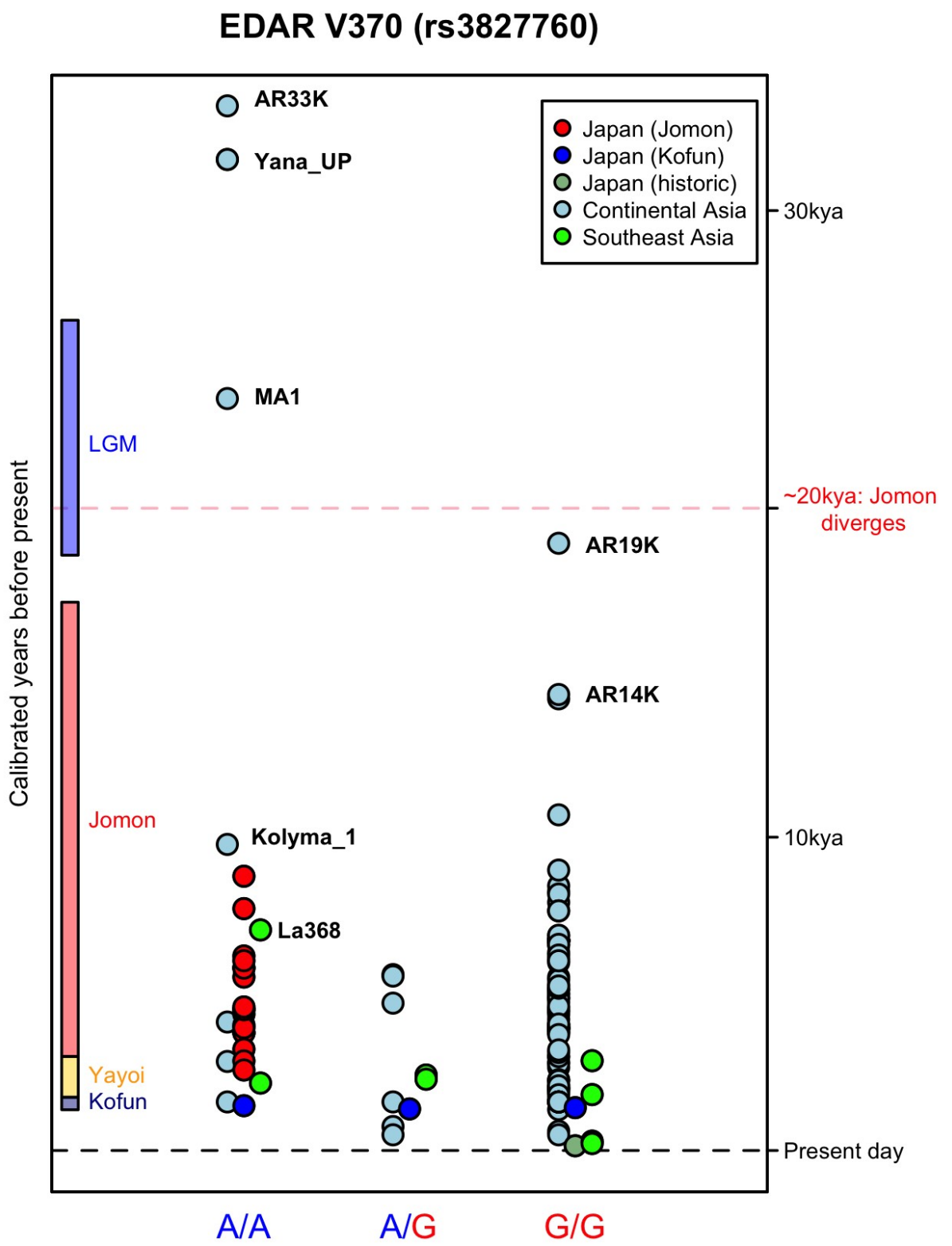

Figure S14. **Regional and temporal differences in *EDAR* genotypes (rs38227760) across imputed Asian data, related to Figure 3.** "A" or "G" represent the ancestral or derived allele at this SNP site. The plots show genotypes in ancient individuals from continental Asia (lightblue), the Japanese archipelago (red or blue), and Southeast Asia (green). The last glacial maximum (LGM) and cultural periods of Japanese prehistory are marked on the *y*-axis.

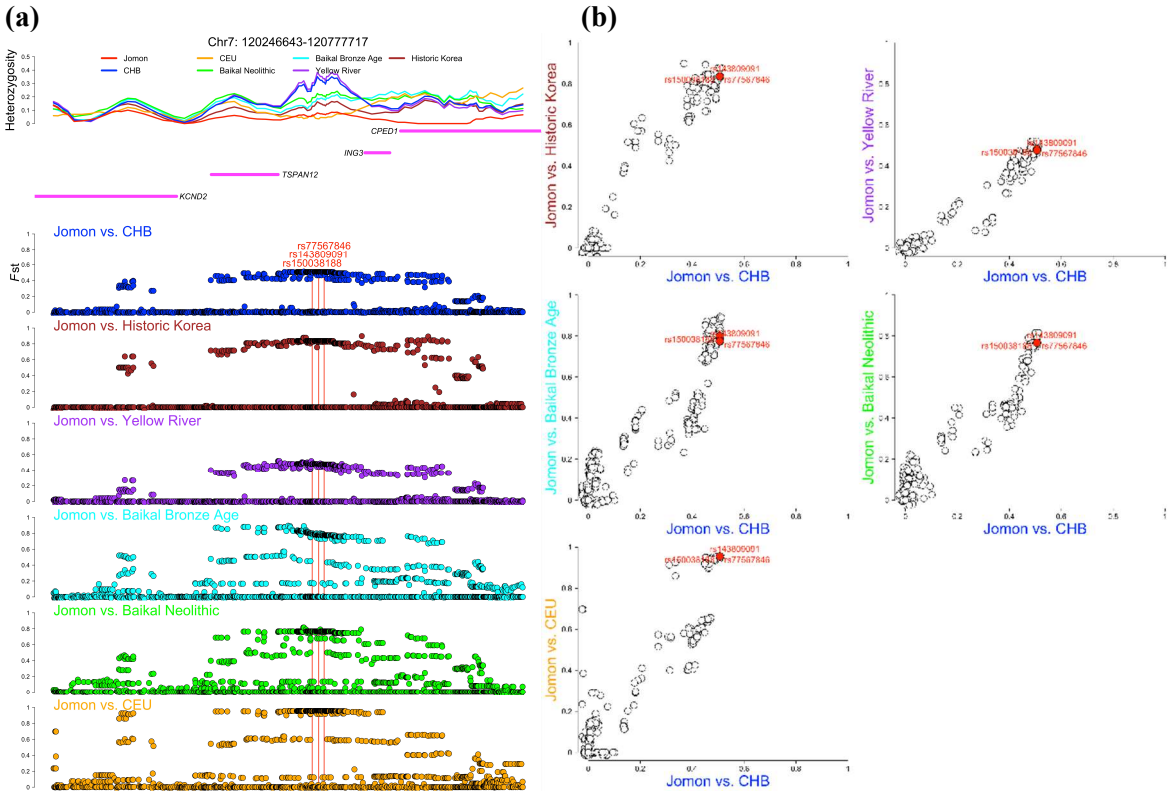

Figure S15. Cross-population comparisons of allele frequencies in the region on chromosome 7, related to Figure 3. (a) Genetic differentiation is measured for all imputed SNPs in the region between Jomon and each of ancient (Historic Korea, Yellow River, Baikal Bronze Age, and Baikal Neolithic; see Table S1 for groupings) or modern continental populations (CHB and CEU). Three SNPs (rs143809091, rs150038188, and rs77567846) that are highly differentiated across all pairwise comparisons are highlighted by red solid lines with their rs numbers. **(b)** *F*st values between Jomon and CHB are compared with those in any pairs of Jomon and the other populations. The three SNPs are amongst the top variants in any comparisons.

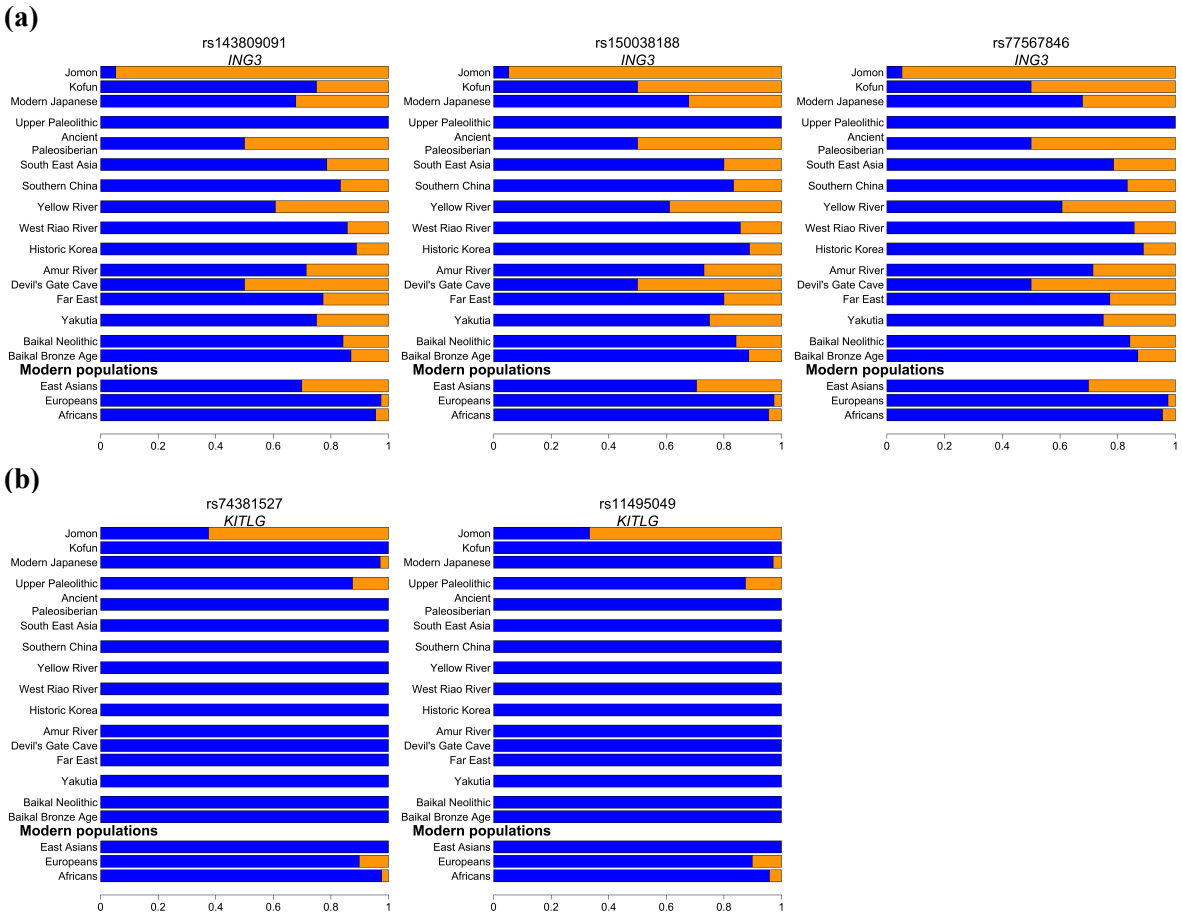

Figure S16. Frequencies of focal alleles in imputed ancient populations and modern-day references, related to Figure 3. The bar plots represent the frequencies of ancestral (blue) and derived (orange) alleles for SNPs of the region on **(a)** chromosome 7 and **(b)** chromosome 12 respectively. Grouping of ancient populations/individuals are listed in Table S1.

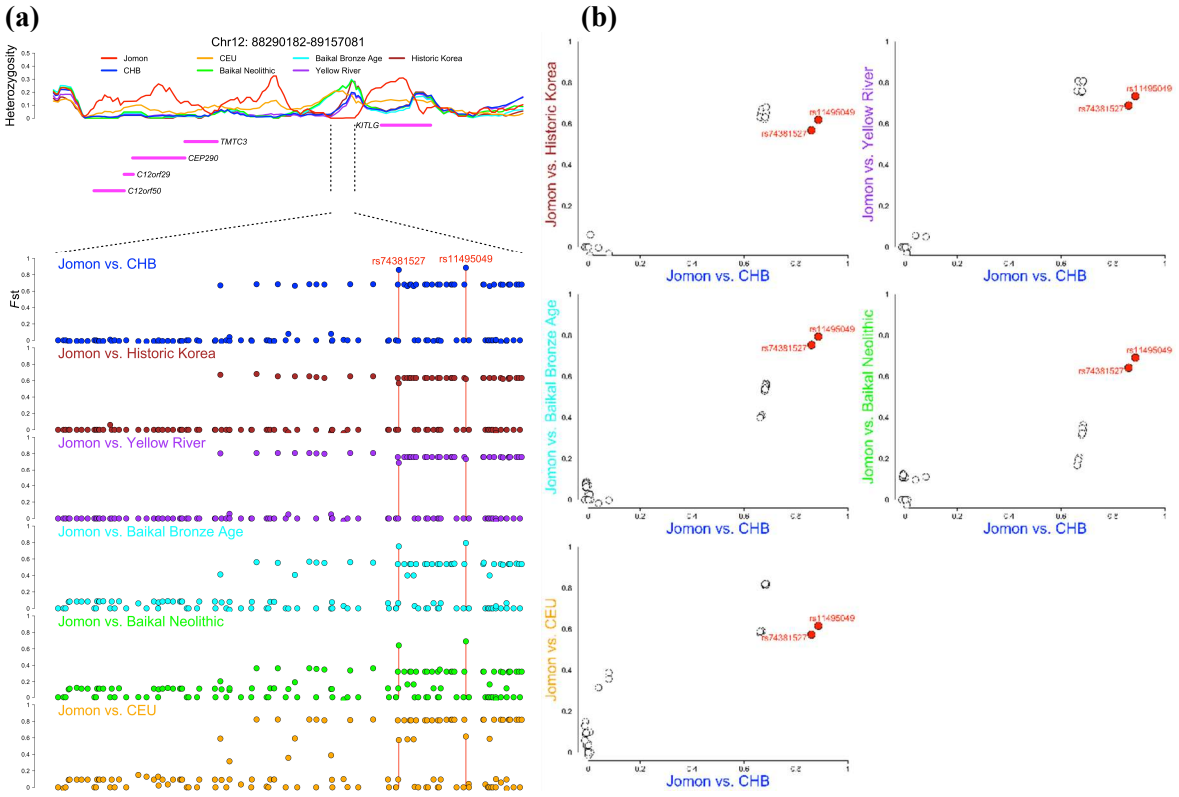

Figure S17. Cross-population comparisons of allele frequencies in a region in which heterozygosity is extremely low in Jomon on chromosome 12, related to Figure 3. (a) Genetic differentiation is measured for all imputed SNPs in the region between Jomon and each of ancient (Historic Korea, Yellow River, Baikal Bronze Age, and Baikal Neolithic; see Table S1 for groupings) and modern continental populations (CHB and CEU). Two SNPs (rs74381527 and rs77567846) that are highly differentiated across all pairwise comparisons are highlighted by red solid lines with their rs numbers. **(b)** *F*st values between Jomon and CHB are compared with those in any pairs of Jomon and the other populations. The two SNPs are amongst the top variants in any comparisons.

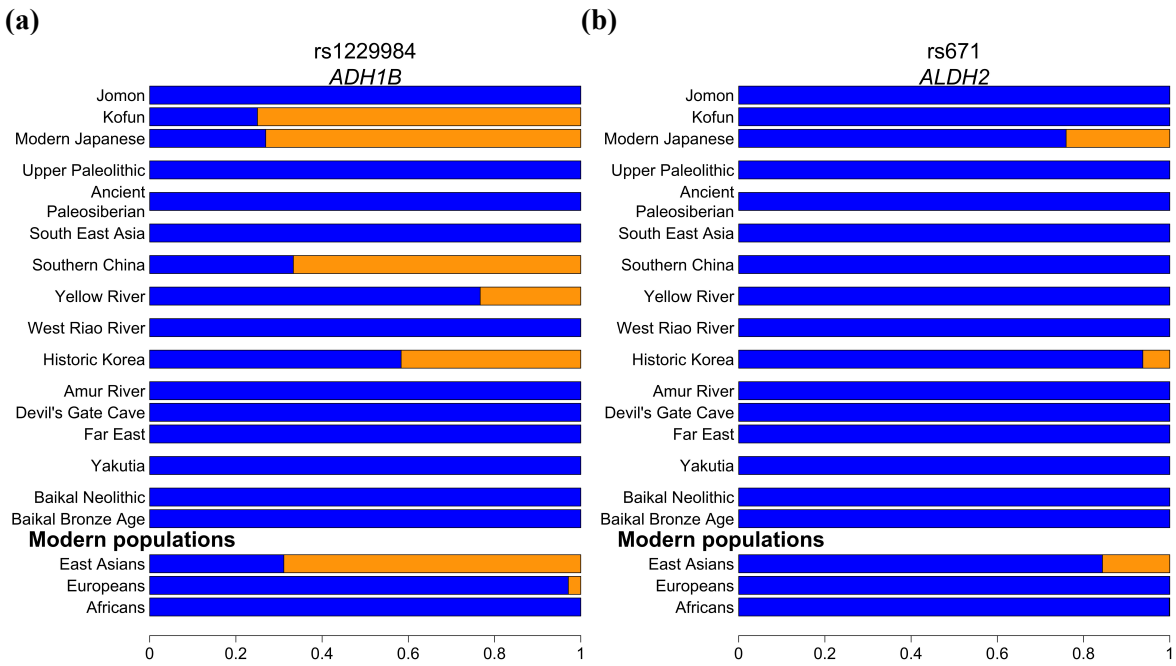

Figure S18. Frequencies of focal alleles for the metabolism of alcohol in imputed ancient populations and modern-day references, related to Figure 4. The bar plots represent the frequencies of ancestral (blue) and derived (orange) alleles for SNPs at **(a)** *ADH1B* and **(b)** *ALDH2* respectively. Grouping of ancient populations/individuals are listed in Table S1.

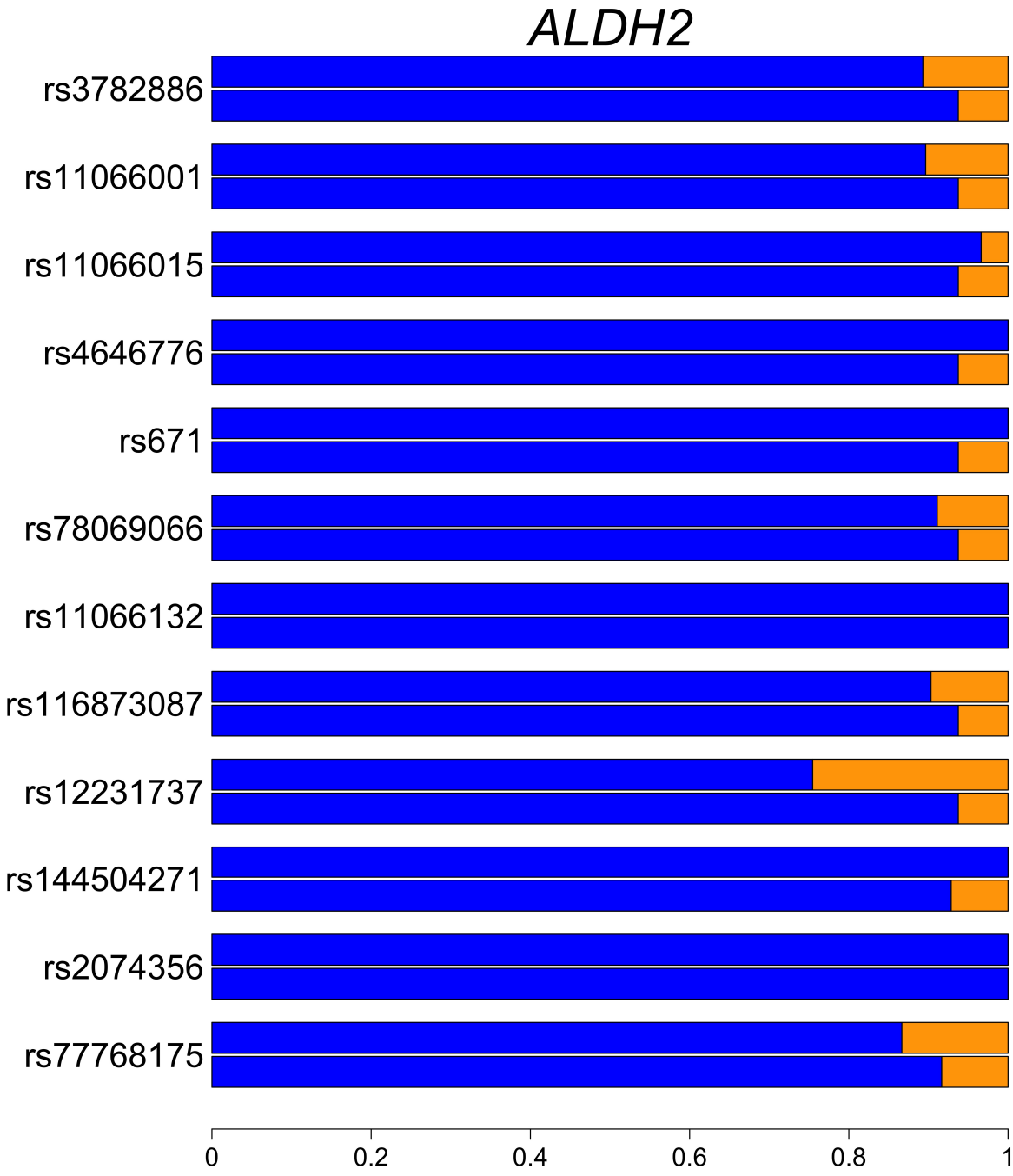

Figure S19. Allele frequencies of rs671 and SNP sites that are in strong LD with rs671 (r^2 > 0.8), related to Figure 4. The bar plots represent the frequencies of ancestral (blue) and derived (orange) alleles in Historic Korea. In each site, the allele frequencies are estimated from raw read counts carrying ancestral or derived alleles (top) or measured from imputed genotypes (bottom).

Table S1. A list of 260 ancient Asian individuals imputed by *GLIMPSE*, related to Figures 2-4. Individuals highlighted in red are deemed too low coverage for analysis based on identifying segments between individuals that are identical-by descent (IBD) or stretches runs of homozygosity (ROH) within an individual; after their exclusion, 126 individuals remained. The following 9 individuals are not listed as they were excluded in analysis due to being in close kinship with other higher coverage individuals: irk022 (CisBaikal_6100_3700_BP), NEO238 (DevilsCave_N), DA340 (Lokomotiv_EN), GUC001, GUC004, GUC006, GUC007 (MedievalKorea), BLSM45 (WLR_MN), and PLTM310 (YR_LN).

Population ID	#inds	Individual ID	Grouping (IBDSeq)	Grouping (Selection scan)	Reference
Blagoveshensk_1345_1270_BP	1	bla001	Amur Region	Amur River	2
AR_EN	2	WQM4, ZLNR2	-		3
AR_IA	1	ZLNR1	-		3
AR_Xianbei_IA	2	MGSM6, MGSM7R	-		3
AR13_10K	2	NE-3, NE-4	-		4
AR14K	2	NE-5, NE34	-		4
ARpost9K	5	NE-9, NE19, NE29, NE35, NE39	-		4
UKY	1	UKY	Ancient Paleosiberian	Ancient Paleosiberian	5
Kolyma1	1	Kolyma1	Ancient Paleosiberian		6
AR19K	1	AR19K	-	-	4
AR33K	1	AR33K	-	-	4
Baikal_BA_outlier	2	GLZ001, GLZ002	Baikal (Bronze Age)	Baikal (Bronze Age)	5
Baikal_LNBA	11	GLZ003, KAG002, KPT001, KPT002, KPT003, KPT004, KPT005, KPT006, STB001, ZPL001, ZPL002	Baikal (Bronze Age)		5
CisBaikal_6100_3700_BP	15	Anosovo-1, irk008, irk017, irk025, irk030, irk033, irk034, irk036, irk040, irk057, irk068, irk071, irk075, irk076, mak026	Baikal (Bronze Age)		2
Kurma_EBA	3	DA354, DA358, DA360	Baikal (Bronze Age)		7
Shamanka_EBA	4	DA334, DA336, DA337, DA339	Baikal (Bronze Age)		8
Ust_Ida_EBA	4	DA343, DA353, DA356, DA361	Baikal (Bronze Age)		8
Ust_Ida_LN	4	DA342, DA344, DA345, DA355	Baikal (Bronze Age)		8
UstBelaya_EBA	4	NEO230, NEO231, NEO232, NEO298	Baikal (Bronze Age)		6
CisBaikal_7980_7855_BP	1	Cyclodrome-1	Baikal (Early Neolithic)	Baikal (Early Neolithic)	2
CisBaikal_8980_8640_BP	1	Popovskij-1	Baikal (Early Neolithic)		2
Baikal_EN	4	ANG001, IUO001, KAG001, STB002	Baikal (Early Neolithic)		5
Lokomotiv_EN	3	DA341, DA357, DA359	Baikal (Early Neolithic)		8
Shamanka_EN	10	DA245, DA246, DA247, DA248, DA249, DA250, DA251, DA252, DA253, DA362	Baikal (Early Neolithic)		8
TransBaikal_8345_3000_BP	6	brn001, brn003, brn008, brn012, cta016, irk078	Baikal (Early Neolithic)		2
TransBaikal_8515_8380_BP	1	Dzhylinda-1	Baikal (Early Neolithic)		2
UstBelaya_N	1	NEO229	Baikal (Early Neolithic)		6
CisBaikal_673_557_BP	1	irk032	Baikal (Historic)	-	2
BZK002	1	BZK002	-	-	5
Chokhopani	1	Chokhopani	Chokhopani	-	9

DevilsCave_N	3	NEO236, NEO239, NEO240	Devil's Gate Cave	Devil's Gate Cave	6
Ekven_IA	10	NEO241, NEO242, NEO243, NEO246, NEO247, NEO248, NEO249, NEO250, NEO251, NEO253	Far East Siberia	Far East	6
Magadan_BA	2	M0831, M9984	Far East Siberia		6
AncientKorean	8	AKG_3420, AKG_3421, AKG_10203, AKG_10204, AKG_10207, AKG_10209, AKG_10210, AKG_10218	Historic Korea	Historic Korea	10,11
MedievalKorea	3	GUC002, GUC003, GUC005	-		10,11
Funadomari_Jomon	2	F23, F5	Jomon	Jomon	12
Higashimyo	1	Higashimyo	-		13
IK002	1	IK002	Jomon		14
Jomon	12	JpFu1, JpHi01, JpKa6904, JpKo13, JpKo2, JpOd181, JpOd274, JpOd282, JpOd6	Jomon		1
Jomon_Capture	4	I13882, I13883, I13884, I13886	-		15
Nagabaka_Jomon	2	NAG019, NAG038	-		16
Khaiyrgas-1	1	Khaiyrgas-1	-	-	2
Kofun	3	JpIw31, JpIw32, JpIw33	Kofun	Kofun	1
LeangPanninge	1	LeangPanninge	-	-	17
Mongolia	2	ARS008, ARS026	Mongolia	-	18
Nagabaka_H	1	NAG039	Nagabaka_H	-	16
Salkhit	1	Salkhit	-	-	19
S_China	24	BalongKD07, BalongKD10, BandaKD11, BandaKD15, Baojiansan5_M1, CenxunKP05, CenxunKP07, CenxunKP13, Dushan4_1, GaofengNL23, HuaqiaoNL26, HuatuyanNL0, HuatuyanNL11, HuatuyanNL17, HuatuyanNL19, HuatuyanNL21, LadaKH01, LayiKD01, Longlin_1, Qihe3, QinchangKD13, QinchangKD14, ShenxianKP09, YiyangKP17	-	Southern China (Baojiansan5_M1, Dushan4_1, Longlin_1, and Qihe3 are excluded from “Southern China” to define “Historic Southern China”)	20
McColl_Group1	2	La368, Ma911	-	Southeast Asia	14
McColl_Group2	6	La364, La727, La898, Ma912, Th531, Vt880	Southeast Asia		14
McColl_Group3	5	Vt778, Vt781, Vt796, Vt808, Vt833	-		14
McColl_Group3_1	1	Th703	-		14
McColl_Group4	4	Th519, Th521, Th530, Vt719	-		14
McColl_Group4_1	2	Vt777, Vt779	-		14
McColl_Group5	2	In661, In662	-		14
McColl_Group6	3	Ma554, Ma555, Phl534	Southeast Asia		14
TY	1	TY	-	-	21
MA1	1	MA1	Upper Paleolithic	Upper Paleolithic	22
Yana_UP	2	Yana1, Yana2	Upper Paleolithic		6
AG2	1	AG2	-		22
HMMH_MN	1	HMF32	West Liao River	West Liao River	3

WLR_BA_o	1	91KLM2	West Liao River		[3]
WLR_LN	3	EDM124, EDM139, EDM176	-		[3]
WLR_BA	2	91KLH11, 91KLH18	West Liao River		[3]
WLR_MN	1	BLSM27S	West Liao River		[3]
Yakutia_Lena_6850_6190_BP	2	Matta-1, Onnyos-1	Yakutia	Yakutia	[2]
Yakutia_Lena_Kolyma_4780_2490_BP	6	Kamenka-3, kra001, N4a1, N4b2, yak021, yak030	Yakutia		[2]
Yana_MED	1	Yana_young	Yakutia		[6]
Yayoi	1	Yayoi_2	-	-	[23]
Miaozigou_MN	3	MZGM101, MZGM16, MZGM252	-	-	[3]
Shimao_LN	3	SMSGDLM27, SMSGDLM6, SMSGDLM7X	Yellow River	-	[3]
Upper_YR_IA	4	DCZM17IV, DCZM21II, DCZM22IV, DCZM6	Yellow River	Yellow River	[3]
Upper_YR_LN	6	JCKM11, LJM14, LJM2, LJM25, LJM3, LJM4	Yellow River		[3]
YR_LBIA	6	HJTM115, HJTW13, JXNTM2, JXNTM23, LGM41, LGM79	Yellow River		[3]
YR_LN	5	HJTM107, HJTM109, PLTM311, WDWT1H16, WDWT5M2	Yellow River		[3]
YR_MN	8	WGH351, WGM20, WGM35, WGM43, WGM70, WGM76S, WGM94, XWM1R18	Yellow River		[3]

Table S2. Allelic counts ("Count") and imputed genotypes ("Geno") for individuals from Jomon, Kofun, and Historic Korea, related to Figures 3-4. In the "Geno" column, "0/0", "0/1", or "1/1" indicates an imputed genotype for a homozygote of the ancestral allele, heterozygote, or homozygote of the derived allele, respectively. A hyphen ("-") denotes a missing imputed genotype for a particular individual, thus allelic counts from this individual are not used for allele frequency estimation.

Population	Individual	rs38227760 (*EDAR*)			rs77567846 (*ING3*)			rs74381527 (*KITLG*)			rs1229984 (*ADH1B*)			rs671 (*ALDH2*)		
		Count		Geno	Count		Geno	Count		Geno	Count		Geno	Count		Geno
		Anc.	Der.		Anc.	Der.		Anc.	Der.		Anc.	Der.		Anc.	Der.	
Jomon	F23	45	2	0/0	0	40	1/1	44	2	0/0	41	0	0/0	33	0	0/0
	F5	6	0	0/0	0	3	1/1	4	0	0/0	-	-	-	6	0	0/0
	Higashimyo	0	0	0/0	0	0	1/1	-	-	-	-	-	-	0	0	0/0
	I13882	5	0	0/0	0	0	1/1	-	-	-	-	-	-	1	0	0/0
	I13883	0	0	0/0	0	0	1/1	0	0	0/0	-	-	-	0	0	0/0
	I13884	6	0	0/0	0	0	1/1	0	0	0/0	-	-	-	1	0	0/0
	I13886	10	0	0/0	0	0	1/1	0	0	1/1	-	-	-	1	0	0/0
	IK002	4	0	0/0	0	1	1/1	0	2	1/1	1	0	0/0	1	0	0/0
	JpFu1	1	0	0/0	0	0	1/1	0	1	0/1	-	-	-	0	0	0/0
	JpHi01	-	-	-	0	2	1/1	0	3	1/1	-	-	-	1	0	0/0
	JpKa6904	13	0	0/0	0	2	1/1	0	7	1/1	-	-	-	5	0	0/0
	JpKo2	2	0	0/0	0	4	1/1	0	0	1/1	-	-	-	1	0	0/0
	JpKo13	2	0	0/0	0	2	1/1	0	1	1/1	3	0	0/0	1	0	0/0
	JpOd6	0	0	0/0	1	1	0/1	0	1	0/1	-	-	-	0	0	0/0
	JpOd181	1	0	0/0	2	1	0/1	3	0	0/0	-	-	-	2	0	0/0
	JpOd274	0	0	0/0	0	1	1/1	0	1	1/1	-	-	-	2	0	0/0
	JpOd282	2	0	0/0	0	0	1/1	0	1	1/1	-	-	-	0	0	0/0
	NAG019	7	0	0/0	0	0	1/1	0	0	1/1	4	0	0/0	0	0	0/0
	NAG038	2	0	0/0	0	0	1/1	-	-	-	-	-	-	0	0	0/0
Kofun	JpIw31	1	2	0/1	0	0	1/1	0	0	0/0	-	-	-	3	0	0/0
	JpIw32	0	2	1/1	4	0	0/1	9	1	0/0	0	9	1/1	3	0	0/0
	JpIw33	1	0	0/0	0	0	0/0	0	0	0/0	2	0	0/1	-	-	-
Historic Korea	GUC002	0	5	1/1	-	-	-	0	0	0/0	-	-	-	-	-	-
	GUC003	-	-	-	0	0	0/0	0	0	0/0	-	-	-	-	-	-
	GUC005	2	3	0/1	-	-	-	0	0	0/0	-	-	-	0	0	0/0
	AKG_3420	3	3	0/1	3	0	0/0	5	0	0/0	-	-	-	1	0	0/1
	AKG_3421	0	1	1/1	2	0	0/0	0	0	0/0	0	0	0/1	0	0	0/0
	AKG_10203	0	1	1/1	2	0	0/0	1	0	0/0	0	0	0/1	-	-	-
	AKG_10204	0	0	1/1	0	0	0/0	0	0	0/0	2	0	0/0	0	0	0/0
	AKG_10207	6	0	0/0	0	0	0/0	5	0	0/0	6	0	0/0	5	0	0/0
	AKG_10209	0	2	1/1	0	0	0/1	3	0	0/0	-	-	-	2	0	0/0
	AKG_10210	0	7	1/1	3	0	0/0	3	0	0/0	0	6	1/1	4	0	0/0

	AKG_10218	0	0	1/1	0	0	0/1	1	0	0/0	0	1	0/1	1	0	0/0

Table S3. Associations of SNPs (rs150038188, rs143809091, and rs77567846) with traits, related to Figure 3. *P*-values and effect sizes are obtained from Open Targets Genetics (https://genetics.opentargets.org/; accessed March 13th, 2024).

Phenotype (data field ID in UK Biobank)		rs150038188	rs143809091	rs77567846	Phenotypic effect of derived alleles
Heel bone mineral density	*P*-value	7.0e-36	5.9e-36	6.1e-36	↑ Heel bone mineral density
	Effect size	0.0753	0.0753	0.0753	
Heel bone mineral density (bmd) (3148)	*P*-value	9.5e-14	8.1e-14	8.1e-14	↑ Heel bone mineral density
	Effect size	0.00901	0.00904	0.00904	
Heel bone mineral density	*P*-value	1.2e-13	1.3e-13	1.4e-13	↑ Heel bone mineral density
	Effect size	0.0678	0.0677	0.0676	
Heel broadband ultrasound attenuation, direct entry (3144)	*P*-value	1.6e-13	1.4e-13	1.4e-13	↑ Heel broadband ultrasound attenuation, direct entry
	Effect size	1.24	1.24	1.24	
Heel bone mineral density (bmd) t-score, automated (78)	*P*-value	2.0e-13	1.7e-13	1.7e-13	↑ Heel bone mineral density (bmd) t-score, automated
	Effect size	0.0797	0.0799	0.0799	
Heel quantitative ultrasound index (qui), direct entry (3147)	*P*-value	2.0e-13	1.7e-13	1.7e-13	↑ Heel quantitative ultrasound index (qui), direct entry
	Effect size	1.41	1.42	1.42	
Heel bone mineral density (bmd) (right) (4124)	*P*-value	1.3e-8	1.5e-8	1.4e-8	↑ Heel bone mineral density (right)
	Effect size	0.00935	0.00933	0.00933	
Heel bone mineral density (bmd) t-score, automated (right) (4125)	*P*-value	1.7e-8	1.9e-8	1.8e-8	↑ Heel bone mineral density t-score, automated (right)
	Effect size	0.0833	0.0830	0.0831	
Heel quantitative ultrasound index (qui), direct entry (right) (4123)	*P*-value	1.7e-8	1.9e-8	1.8e-8	↑ Heel quantitative ultrasound index (qui), direct entry (right)
	Effect size	1.48	1.47	1.47	
Heel broadband ultrasound attenuation (right) (4120)	*P*-value	3.4e-8	3.8-e8	3.7e-8	↑ Heel broadband ultrasound attenuation (right)
	Effect size	1.27	1.26	1.26	

Table S4. Associations of SNPs (rs74381527 and rs11495049) with traits and gene expression, related to Figure 3. *P*-values and effect sizes are obtained from Open Targets Genetics (https://genetics.opentargets.org/; accessed October 4th, 2023) or GTEx Portal (https://gtexportal.org/home/; accessed October 4th, 2023).

Phenotype (data field ID) and tissue of eQTLs		**rs74381527**	**rs11495049**	**Phenotypic effect of derived alleles**
Blonde | hair colour (natural, before greying) (1747)	*P*-value	3.5E-104	4.7E-104	Darker hair colour
	Effect size	-0.254	-0.254	
Dark brown | hair colour (natural, before greying) (1747)	*P*-value	1.2E-103	9.7E-103	
	Effect size	0.167	0.166	
Light brown | hair colour (natural, before greying) (1747)	*P*-value	2.9E-45	7.8E-45	
	Effect size	-0.107	-0.107	
Black | hair colour (natural, before greying) (1747)	*P*-value	4.1E-38	4.0E-38	
	Effect size	0.232	0.232	
Skin colour (1717)	*P*-value	1.8E-21	1.8E-21	Darker skin pigmentation
	Effect size	0.0207	0.0207	
Cultured fibroblasts	*P*-value	2.30E-23	5.70E-22	↑ *KITLG*
	Effect size	0.77	0.74	

Table S5. Frequencies of SNP sites in linkage disequilibrium with SNPs related to alcohol metabolism in Japanese and Chinese populations, related to Figure 4. In addition to calculating the frequencies of derived alleles at rs1229984 and rs671 (see also Figure S18), which are both marked in bold and linked to alcohol dehydrogenase and aldehyde dehydrogenase respectively, frequencies for all sites in linkage disequilibrium ($r^2 > 0.8$) within 500 kb from the sites are calculated from the imputed data.

Chr	SNP ID	r^2	Mutation type	Jomon	Kofun	Modern Japanese	Historic Korea	YR MN	UYR LN	YR LN	YR LBIA	UYR IA	Historic Southern China	Modern Han
4	**rs1229984**	**-**	**Transition**	**0.00**	**0.75**	**0.73**	**0.42**	**0.00**	**0.00**	**0.33**	**0.50**	**1.00**	**1.00**	**0.71**
	rs78456448	0.83	Transversion	0.00	0.67	0.73	0.31	0.00	0.00	0.10	0.33	1.00	1.00	0.66
	rs75721934	0.95	Transition	0.00	0.75	0.73	0.29	0.00	0.00	0.00	0.25	1.00	1.00	0.67
	rs3811801	0.80	Transition	0.00	0.50	0.70	0.40	0.00	0.00	0.00	0.50	1.00	0.50	0.59
12	**rs671**	**-**	**Transition**	**0.00**	**0.00**	**0.24**	**0.06**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.16**
	rs3782886	0.95	Transition	0.00	0.00	0.25	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.16
	rs11066001	1.00	Transition	0.00	0.00	0.24	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.16
	rs11066015	1.00	Transition	0.00	0.00	0.24	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.16
	rs4646776	1.00	Transversion	0.00	0.00	0.24	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.16
	rs78069066	0.95	Transition	0.00	0.00	0.25	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.16
	rs11066132	1.00	Transition	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16
	rs116873087	0.97	Transversion	0.00	0.00	0.25	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.16
	rs12231737	0.92	Transition	0.00	0.00	0.25	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.17
	rs144504271	0.87	Transition	0.00	0.00	0.24	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.16
	rs2074356	0.85	Transition	0.00	0.00	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10
	rs77768175	0.92	Transition	0.00	0.00	0.23	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.15

References

1. Cooke, N.P., Mattiangeli, V., Cassidy, L.M., Okazaki, K., Stokes, C.A., Onbe, S., Hatakeyama, S., Machida, K., Kasai, K., Tomioka, N., et al. (2021). Ancient genomics reveals tripartite origins of Japanese populations. Sci Adv *7*, eabh2419.

2. Kılınç, G.M., Kashuba, N., Koptekin, D., Bergfeldt, N., Dönertaş, H.M., Rodríguez-Varela, R., Shergin, D., Ivanov, G., Kichigin, D., Pestereva, K., et al. (2021). Human population dynamics and Yersinia pestis in ancient northeast Asia. Sci Adv *7*. 10.1126/sciadv.abc4587.

3. Ning, C., Li, T., Wang, K., Zhang, F., Li, T., Wu, X., Gao, S., Zhang, Q., Zhang, H., Hudson, M.J., et al. (2020). Ancient genomes from northern China suggest links between subsistence changes and human migration. Nat. Commun. *11*, 2700.

4. Mao, X., Zhang, H., Qiao, S., Liu, Y., Chang, F., Xie, P., Zhang, M., Wang, T., Li, M., Cao, P., et al. (2021). The deep population history of northern East Asia from the Late Pleistocene to the Holocene. Cell *184*, 3256–3266.e13.

5. Yu, H., Spyrou, M.A., Karapetian, M., Shnaider, S., Radzevičiūtė, R., Nägele, K., Neumann, G.U., Penske, S., Zech, J., Lucas, M., et al. (2020). Paleolithic to Bronze Age Siberians Reveal Connections with First Americans and across Eurasia. Cell *181*, 1232–1245.e20.

6. Sikora, M., Pitulko, V.V., Sousa, V.C., Allentoft, M.E., Vinner, L., Rasmussen, S., Margaryan, A., de Barros Damgaard, P., de la Fuente, C., Renaud, G., et al. (2019). The population history of northeastern Siberia since the Pleistocene. Nature *570*, 182–188.

7. de Barros Damgaard, P., Martiniano, R., Kamm, J., Moreno-Mayar, J.V., Kroonen, G., Peyrot, M., Barjamovic, G., Rasmussen, S., Zacho, C., Baimukhanov, N., et al. (2018). The first horse herders and the impact of early Bronze Age steppe expansions into Asia. Science *360*. 10.1126/science.aar7711.

8. Damgaard, P. de B., Marchi, N., Rasmussen, S., Peyrot, M., Renaud, G., Korneliussen, T., Moreno-Mayar, J.V., Pedersen, M.W., Goldberg, A., Usmanova, E., et al. (2018). 137 ancient human genomes from across the Eurasian steppes. Nature *557*, 369–374.

9. Jeong, C., Ozga, A.T., Witonsky, D.B., Malmström, H., Edlund, H., Hofman, C.A., Hagan, R.W., Jakobsson, M., Lewis, C.M., Aldenderfer, M.S., et al. (2016). Long-term genetic stability and a high-altitude East Asian origin for the peoples of the high valleys of the Himalayan arc. Proc. Natl. Acad. Sci. U. S. A. *113*, 7485–7490.

10. Gelabert, P., Blazyte, A., Chang, Y., Fernandes, D.M., Jeon, S., Hong, J.G., Yoon, J., Ko, Y., Oberreiter, V., Cheronet, O., et al. (2022). Northeastern Asian and Jomon-related genetic structure in the Three Kingdoms period of Gimhae, Korea. Curr. Biol. *32*, 3232–3244.e6.

11. Lee, D.-N., Jeon, C.L., Kang, J., Burri, M., Krause, J., Woo, E.J., and Jeong, C. (2022). Genomic detection of a secondary family burial in a single jar coffin in early Medieval Korea. Am. J. Biol. Anthropol. *179*, 585–597.

12. Kanzawa-Kiriyama, H., Jinam, T.A., Kawai, Y., Sato, T., Hosomichi, K., Tajima, A., Adachi, N., Matsumura, H., Kryukov, K., Saitou, N., et al. (2019). Late Jomon male and female genome sequences from the Funadomari site in Hokkaido, Japan. Anthropol. Sci. *advpub*. 10.1537/ase.190415.

13. Adachi, N., Kanzawa-Kiriyama, H., Nara, T., Kakuda, T., Nishida, I., and Shinoda, K.-I. (2021). Ancient genomes from the initial Jomon period: new insights into the genetic history of the Japanese archipelago. Anthropol. Sci. *129*, 13–22.

14. McColl, H., Racimo, F., Vinner, L., Demeter, F., Gakuhari, T., Moreno-Mayar, J.V., van Driem,

G., Gram Wilken, U., Seguin-Orlando, A., de la Fuente Castro, C., et al. (2018). The prehistoric peopling of Southeast Asia. Science *361*, 88–92.

15. Wang, C.-C., Yeh, H.-Y., Popov, A.N., Zhang, H.-Q., Matsumura, H., Sirak, K., Cheronet, O., Kovalev, A., Rohland, N., Kim, A.M., et al. (2021). Genomic insights into the formation of human populations in East Asia. Nature *591*, 413–419.

16. Robbeets, M., Bouckaert, R., Conte, M., Savelyev, A., Li, T., An, D.-I., Shinoda, K.-I., Cui, Y., Kawashima, T., Kim, G., et al. (2021). Triangulation supports agricultural spread of the Transeurasian languages. Nature *599*, 616–621.

17. Carlhoff, S., Duli, A., Nägele, K., Nur, M., Skov, L., Sumantri, I., Oktaviana, A.A., Hakim, B., Burhan, B., Syahdar, F.A., et al. (2021). Genome of a middle Holocene hunter-gatherer from Wallacea. Nature *596*, 543–547.

18. Jeong, C., Wilkin, S., Amgalantugs, T., Bouwman, A.S., Taylor, W.T.T., Hagan, R.W., Bromage, S., Tsolmon, S., Trachsel, C., Grossmann, J., et al. (2018). Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. Proc. Natl. Acad. Sci. U. S. A. *115*, E11248–E11255.

19. Massilani, D., Skov, L., Hajdinjak, M., Gunchinsuren, B., Tseveendorj, D., Yi, S., Lee, J., Nagel, S., Nickel, B., Devièse, T., et al. (2020). Denisovan ancestry and population history of early East Asians. Science *370*, 579–583.

20. Wang, T., Wang, W., Xie, G., Li, Z., Fan, X., Yang, Q., Wu, X., Cao, P., Liu, Y., Yang, R., et al. (2021). Human population history at the crossroads of East and Southeast Asia since 11,000 years ago. Cell *184*, 3829–3841.e21.

21. Yang, M.A., Gao, X., Theunert, C., Tong, H., Aximu-Petri, A., Nickel, B., Slatkin, M., Meyer, M., Pääbo, S., Kelso, J., et al. (2017). 40,000-Year-Old Individual from Asia Provides Insight into Early Population Structure in Eurasia. Curr. Biol. *27*, 3202–3208.e9.

22. Raghavan, M., Skoglund, P., Graf, K.E., Metspalu, M., Albrechtsen, A., Moltke, I., Rasmussen, S., Stafford, T.W., Jr, Orlando, L., Metspalu, E., et al. (2014). Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature *505*, 87–91.

23. Shinoda, K.-I., Kanzawa-Kiriyama, H., Kakuda, T., and Adachi, N. (2019). Genetic characteristics of Yayoi people in Northwestern Kyushu. Preprint, 10.1537/asj.1904231 10.1537/asj.1904231.