

Supplementary Tables

Table S1. Sample information of the populations genotyped in the present study.

Populations	N	Sampling Location	Language phylum	Grouping
Han Chinese from Guangdong (Han-GD)	169	Guangdong province	Sino-Tibetan	Han Chinese populations
Han Chinese from Shanghai (Han-SH)	264	Shanghai city	Sino-Tibetan	
Han Chinese from Shandong (Han-SD)	179	Shandong province	Sino-Tibetan	
Lic	32	Hainan province	Austronesian	Ethnic minorities from South China
Zhuang	68	Guangxi Zhuang autonomous region	Austronesian	
Miao	92	Guizhou province	Hmong–Mien	
Tibetan	91	Tibet Autonomous Region	Tibeto-Burman	Ethnic minorities from Northwest China
Mongolian	65	Inner Mongolia autonomous region	Mongolic	
Uygur	76	Xinjiang Uygur autonomous region	Turkic	

Table S2. The AMOVA result. Population structure was dissected through comparison of the genetic variance among different grouping hypotheses. A grouping hypothesis, which allocates a higher proportion of the variance among groups and a lower percentage among populations within groups, can reflect a more reliable genetic divergence. The within-population genetic variance contributes the largest percentage in all grouping-based analyses. A high variance percentage among groups was produced when separating the three northern minority populations from the other. The highest among-group variance was obtained when grouping in agreement with the language phylum. The variance percentage among groups turned larger when grouping Han-GD with Lic and Zhuang than with Miao, indicating a closer relation between Han-GD with Lic and Zhuang.

Grouping		Among groups	Among populations	Within populations
Lic, Zhuang, Miao, Han-GD, Han-SH, Han-SD / Tibetan, Mongolian, Uygur	Percentage of variance (%)	1.61	1.13	97.26
	Fixation Index	0.01605	0.01149	0.02735
	<i>P</i> Value	0.01238	$< 10^{-5}$	$< 10^{-5}$
Lic, Zhuang, Miao / Han-GD, Han-SH, Han-SD / Tibetan, Mongolian, Uygur	Percentage of variance (%)	1.19	1	97.81
	Fixation Index	0.01193	0.01013	0.02194
	<i>P</i> Value	0.00396	$< 10^{-5}$	$< 10^{-5}$
Lic, Zhuang / Miao / Han-GD, Han-SH, Han-SD / Tibetan, Mongolian, Uygur	Percentage of variance (%)	1.36	0.86	97.78
	Fixation Index	0.01362	0.00872	0.02222
	<i>P</i> Value	0.00178	$< 10^{-5}$	$< 10^{-5}$
Lic, Zhuang, Miao, Han-GD / Han-SH, Han-SD / Tibetan, Mongolian, Uygur	Percentage of variance (%)	0.95	1.08	97.97
	Fixation Index	0.00949	0.01089	0.02028
	<i>P</i> Value	0.00069	$< 10^{-5}$	$< 10^{-5}$
Lic, Zhuang, Han-GD / Miao / Han-SH, Han-SD / Tibetan, Mongolian, Uygur	Percentage of variance (%)	1.08	0.92	98
	Fixation Index	0.0108	0.00931	0.02001
	<i>P</i> Value	0.01079	$< 10^{-5}$	$< 10^{-5}$
Lic, Zhuang / Han-GD, Miao / Han-SH, Han-SD / Tibetan, Mongolian, Uygur	Percentage of variance (%)	0.85	1.11	98.04
	Fixation Index	0.00847	0.01119	0.01956
	<i>P</i> Value	0.01446	$< 10^{-5}$	$< 10^{-5}$

Lic, Zhuang, Han-GD / Miao / Han-SH, Han-SD / Tibetan / Mongolian / Uyгур	Percentage of variance (%) Fixation Index <i>P</i> Value	1.84 0.01839 0.00386	0.23 0.00233 < 10 ⁻⁵	97.93 0.02068 < 10 ⁻⁵
Lic, Zhuang / Miao / Han-GD, Han-SH, Han-SD / Tibetan / Mongolian / Uyгур	Percentage of variance (%) Fixation Index <i>P</i> Value	2.08 0.02076 0.00059	0.29 0.00292 < 10 ⁻⁵	97.64 0.02361 < 10 ⁻⁵

Note- The bold slash symbols denote grouping among the populations. The fixation index indicates *F_{st}*, *F_{ic}* and *F_{cs}*, respectively. The *P* values in each grouping analysis were produced by significance tests of 10,000 permutations.

Table S3. The top percentiles (99.9th, 99th, 98th and 95th) were calculated from the F_{ST} value distributions. This table is related to Fig. S4.

Percentile	99.9 th	99 th	98 th	95 th
All Fst Values	0.308238	0.140464	0.096288	0.049691
All Populations	0.162259	0.102287	0.073021	0.042191
Han_GD-Han_SD	0.069555	0.031864	0.021576	0.013843
Han_GD-Han_SH	0.049156	0.020654	0.014987	0.008958
Han_GD-Lic	0.117307	0.081194	0.055104	0.026384
Han_GD-Miao	0.148564	0.07607	0.052627	0.033012
Han_GD-Mongolian	0.244702	0.093479	0.079458	0.037427
Han_GD-Tibetan	0.148564	0.07607	0.052627	0.033012
Han_GD-Uygur	0.385947	0.281161	0.208081	0.121702
Han_GD-Zhuang	0.071436	0.035902	0.027765	0.0125
Han_SD-Han_SH	0.025394	0.011495	0.008974	0.00465
Han_SD-Lic	0.19749	0.104962	0.075868	0.045843
Han_SD-Miao	0.179096	0.094993	0.082268	0.038749
Han_SD-Mongolian	0.106235	0.052365	0.041844	0.02911
Han_SD-Tibetan	0.190449	0.06999	0.045714	0.027795
Han_SD-Uygur	0.34922	0.235829	0.173103	0.104513
Han_SD-Zhuang	0.123178	0.067388	0.053439	0.02996
Han_SH-Lic	0.187085	0.109817	0.082945	0.040557
Han_SH-Miao	0.204314	0.083817	0.062683	0.034364
Han_SH-Mongolian	0.120748	0.074468	0.052494	0.033529
Han_SH-Tibetan	0.232904	0.076166	0.054571	0.031349
Han_SH-Uygur	0.403863	0.295512	0.193242	0.119184
Han_SH-Zhuang	0.105872	0.05998	0.04427	0.026697
Lic-Miao	0.228173	0.118559	0.088162	0.053682
Lic-Mongolian	0.223872	0.142676	0.096594	0.058907
Lic-Tibetan	0.35312	0.185933	0.137305	0.072593
Lic-Uygur	0.42051	0.228524	0.198025	0.118071
Lic-Zhuang	0.124061	0.064665	0.049906	0.024539
Miao-Mongolian	0.232542	0.131982	0.102022	0.067011
Miao-Tibetan	0.239562	0.156779	0.131205	0.071044
Miao-Uygur	0.35816	0.269331	0.205502	0.127766
Miao-Zhuang	0.132178	0.101336	0.069802	0.038911
Mongolian-Tibetan	0.160495	0.057955	0.042667	0.024886
Mongolian-Uygur	0.234891	0.144261	0.088226	0.045656
Mongolian-Zhuang	0.208006	0.117777	0.078811	0.047161
Tibetan-Uygur	0.314594	0.189605	0.148858	0.081797
Tibetan-Zhuang	0.262158	0.150264	0.112434	0.066818
Uygur-Zhuang	0.441832	0.243288	0.172667	0.109453

Table S6. Summary statistics of selection tests for the SNPs with significant F_{ST} values between the different clustered Chinese subpopulations. iHS scores, Fay and Wu's H, and DAF for retrieved from a resource of genome-wide scan for recent positive selection, Haplotter (<http://haplotter.uchicago.edu/>), based on Hapmap phase II data.

rsID	CEU			YRI			ASN		
	iHS	Fay Wu's H	DAF	iHS	Fay Wu's H	DAF	iHS	Fay Wu's H	DAF
rs1883322	-2.389	-45.567	0.725	-0.137	16.672	0.292	0.589	-41.03	0.778
rs7751481	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs2267667	-2.469	-44.308	0.725	-0.868	12.36	0.267	0.626	-39.956	0.772
rs2076167	-1.82	-42.016	0.750	-0.77	5.901	0.3	0.944	-42.268	0.789
rs1003973	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs2267665	2.242	-46.811	0.192	0.435	16.695	0.108	-0.841	-40.849	0.172
rs2038067	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs2267668	2.346	-44.922	0.192	0.509	10.124	0.108	-0.981	-39.723	0.161
rs2267669	2.311	-43.866	0.192	0.509	9.705	0.108	-0.981	-40.958	0.161
rs2277624	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs4148416	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs3748933	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs1884725	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs274548	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs2283458	NA	NA	NA	NA	NA	NA	NA	NA	NA
rs4646227	NA	NA	NA	NA	NA	NA	NA	NA	NA

Abbreviations: iHS, integrated haplotype score; DAF, derived allele frequency; NA, not available.

ASN includes the Japanese of Tokyo and Han Chinese in Beijing from the Hapmap data.

Table S7. The neutrality tests for the upstream and downstream parts of the recombination point at *PPARD*. Three regions were surveyed including a long upstream interval, a 2.2kb fragment around exon 4 harboring the translation start site (TSS) and a 4.5kb stretch spanning from exon 5 to exon 8. The latter two are from the downstream. The statistics for Fu and Li's *D* and *F*, Fay and Wu's *H* and normalized *H* were estimated with an outgroup of ancestral information. The significance was generated from 10,000 coalescent simulations assuming a standard neutral model with no recombination.

Population	Test	Upstream		Downstream			
		Up SNPs (55.3kb)		TSS SNPs (2.2kb)		Exon SNPs (4.5kb)	
		Statistic	<i>P</i> Value	Statistic	<i>P</i> Value	Statistic	<i>P</i> Value
YRI	Taj-D	-0.3576	0.3432	-1.1278	0.1236	-0.24	0.4738
	FL-D	-0.4491	0.3222	-1.5272	0.095	-2.0285	0.0381
	FL-F	-0.4839	0.356	-1.5982	0.0833	-1.6355	0.0679
	FW-H	20.372	0.1574	0.097	0.3481	0.2791	0.42
	FW-Hn	0.6168		0.0622		0.1674	
CEU	Taj-D	-2.5203	0	-0.9032	0.1895	-0.6575	0.2935
	FL-D	-3.1921	0.0051	-1.1425	0.1559	-0.051	0.4535
	FL-F	-3.4144	0.002	-1.2634	0.1231	-0.3281	0.3627
	FW-H	-48.537	0.0001	0.0204	0.315	-2.3308	0.0511
	FW-Hn	-2.1144		0.0115		-1.4904	
CHB1	Taj-D	-1.8113	0.0035	-1.0842	0.0194	-1.8663	0.0057
	FL-D	-3.3325	0.0041	-1.0991	0.0979	-1.777	0.0887
	FL-F	-3.1676	0.0021	-1.2825	0.0764	-2.1551	0.0248
	FW-H	-42.041	0.0052	-1.7236	0.0048	-5.0789	0.0005
	FW-Hn	-2.3145		-3.2756		-3.5707	
CHB2	Taj-D	-1.83	0.0023	1.6496	0.044	2.3234	0.0124
	FL-D	-3.0479	0.0067	-0.7667	0.2342	1.06154	0.1447
	FL-F	-3.0205	0.0065	0.04173	0.4895	1.73967	0.0278
	FW-H	-35.137	0.0146	0.35112	0.4424	0.5261	0.5376
	FW-Hn	-1.5971		0.32178		0.5469	

Table S11. An independent, LD-defined candidate set for the colocalization. A shared variant is assumed under positive selection across major Eurasian populations. A candidate set of variants in LD with rs9470007 were acquired after intersecting the variants in LD of $r^2 > 0.75$ in CHB+CHS and CEU, respectively.

SNP	Chr	Position	r^2 in CHB+CHS	r^2 in CEU
rs11571504	6	35310749	1	1
rs111244160	6	35318338	1	1
rs6901410	6	35330030	1	1
rs6902123	6	35330421	1	1
rs9470001	6	35333741	1	1
rs9296148	6	35337467	0.9395	1
rs7739752	6	35339035	1	1
rs6919334	6	35341875	1	1
rs7754871	6	35347118	1	1
rs9470007	6	35349739	1	1
rs373497464	6	35351397	0.873209	0.781784
rs6922548	6	35353523	1	1
rs9658100	6	35356640	1	1
rs73413718	6	35360177	0.9395	0.929487
rs6457816	6	35362848	0.9395	0.929487