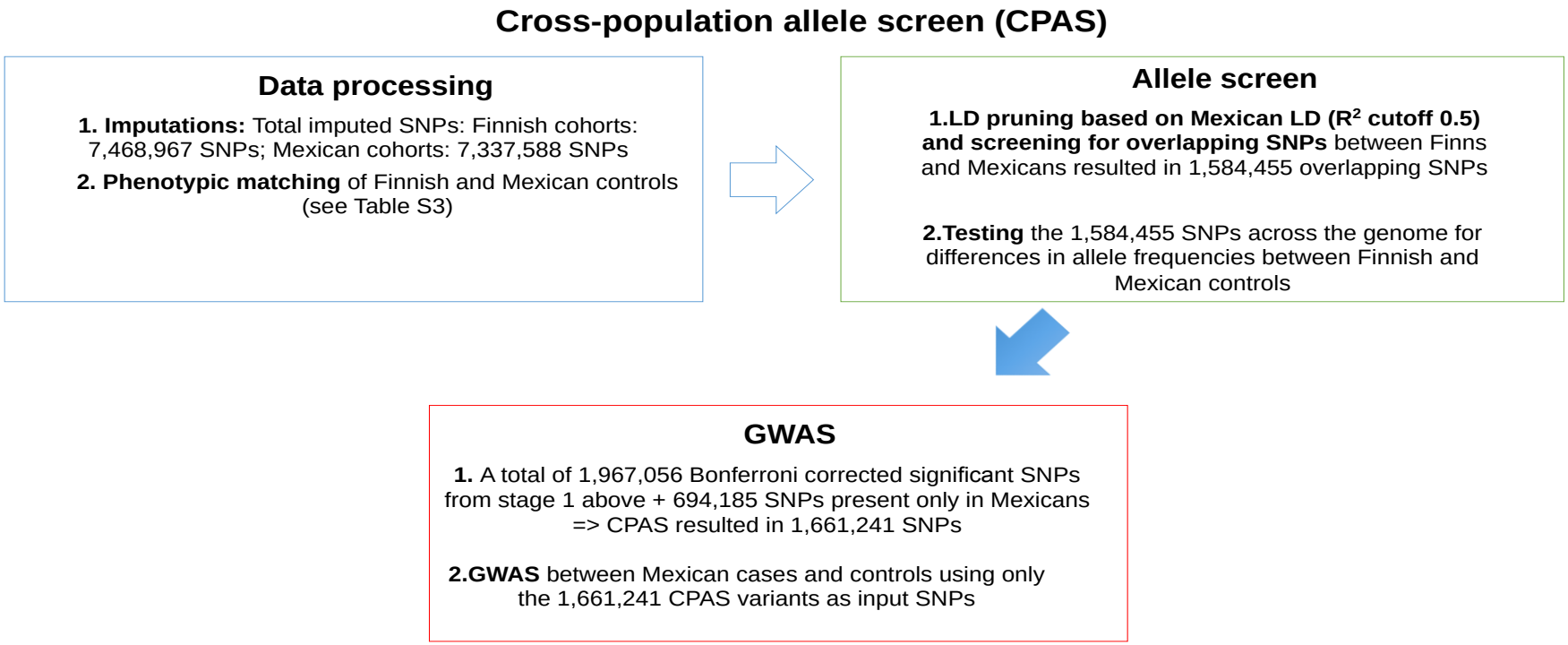
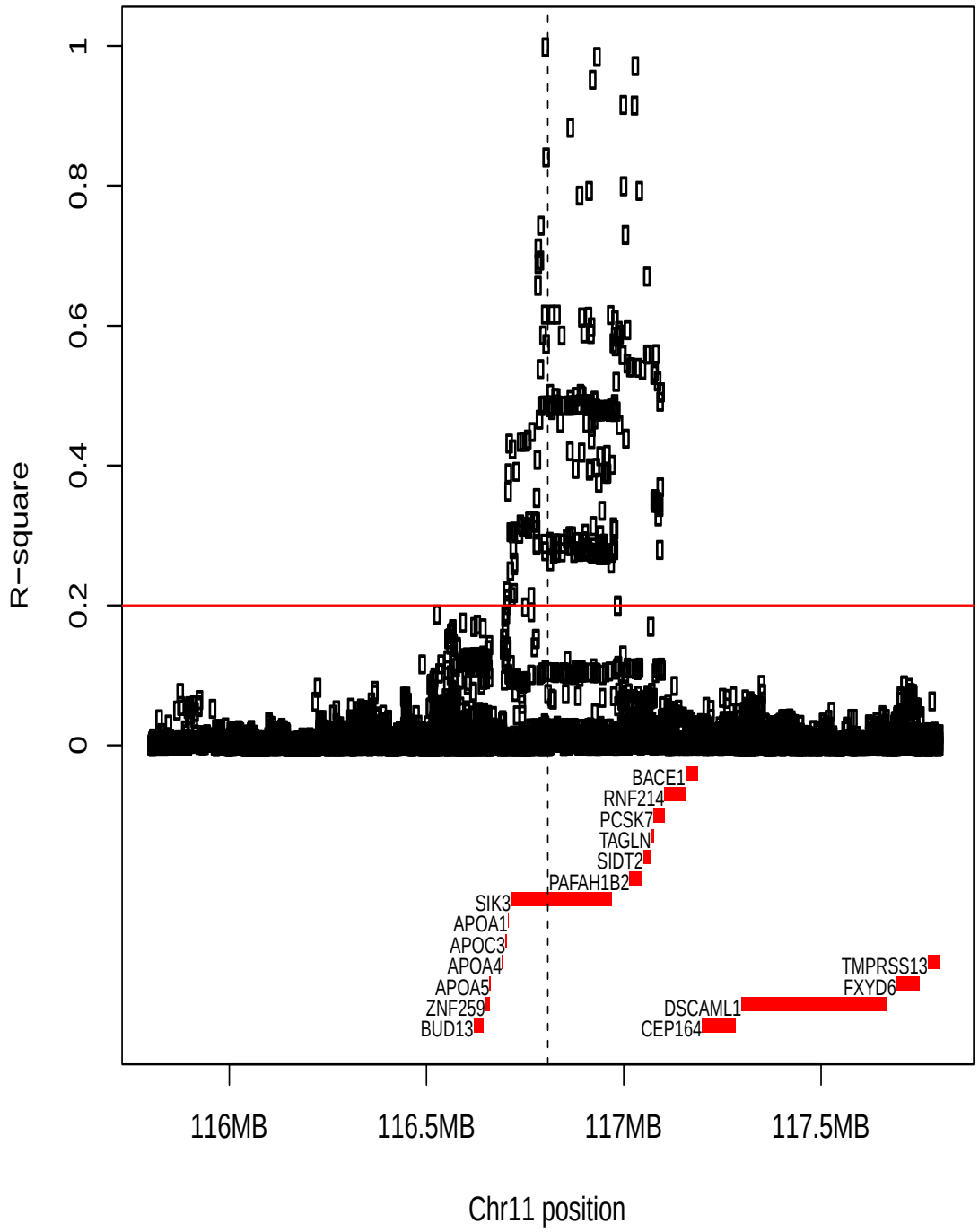
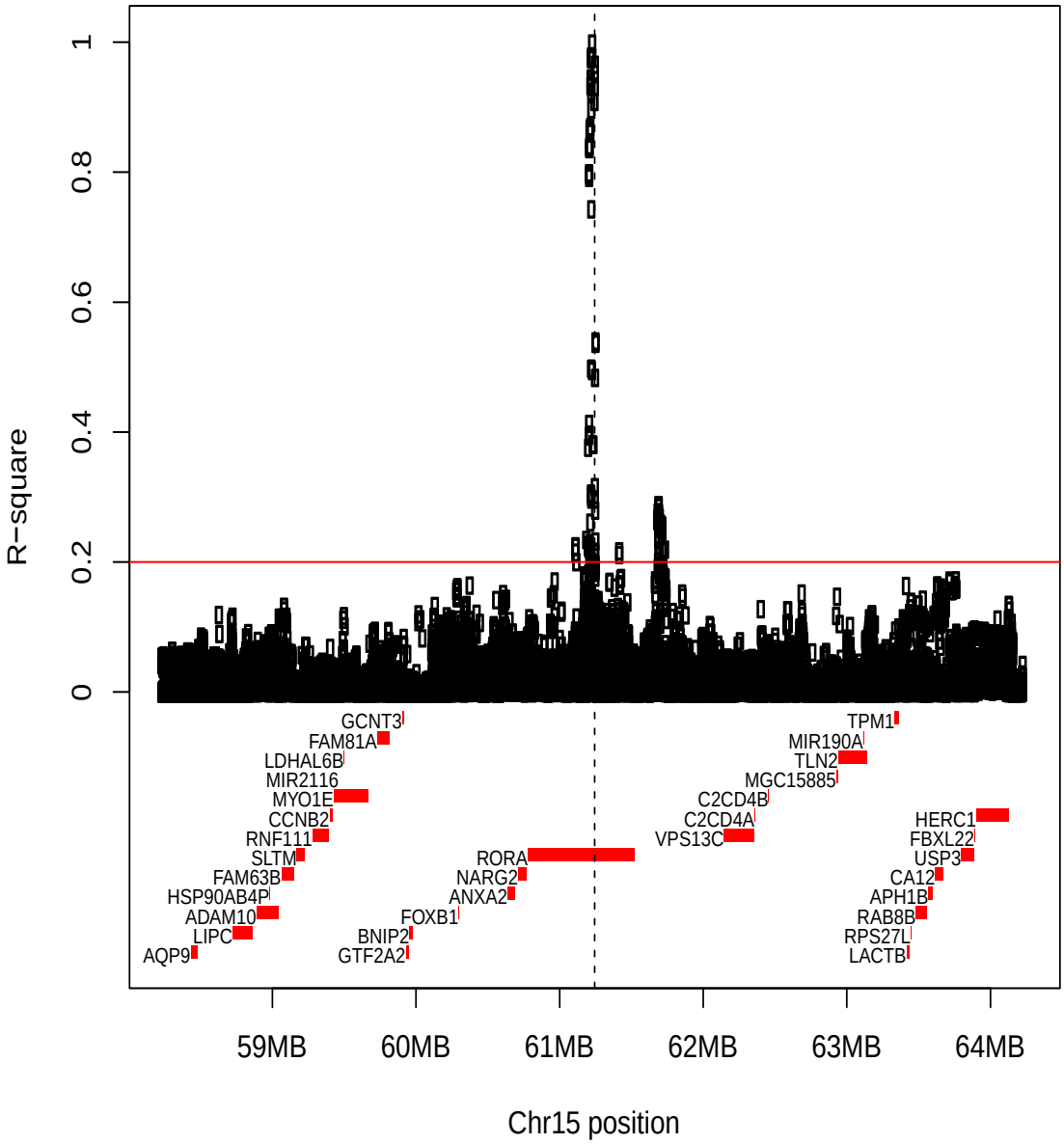


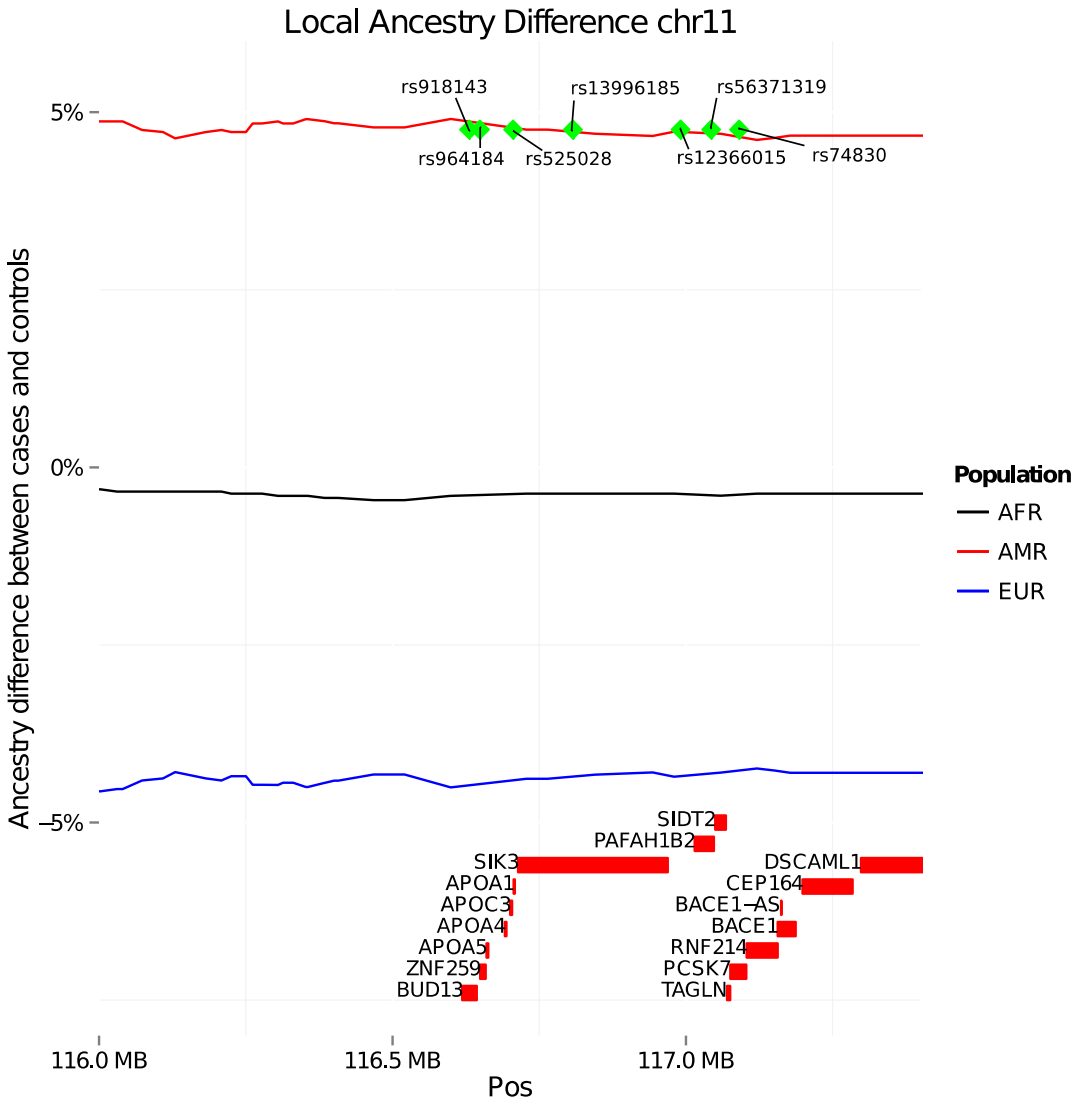
Supplementary Figure 1. Overall design and workflow of Cross-population allele screen (CPAS). To search for Amerindian-specific genetic variants contributing to the high risk of dyslipidemia and obesity in Mexicans, we developed a CPAS-GWAS approach that first screens across the genome for variants that differ in frequency between the two ancestry populations, Europeans and Amerindians, and subsequently includes only these variants (CPAS-variants) in the actual Mexican GWAS. In the data processing step, we carried out imputations in the Finnish and Mexican cohorts. Then, the two cohorts were phenotypically matched to ensure that the differences in allele frequencies are strictly due to a population structure instead of being confounded by other phenotypes. In the allele screen step, we first pruned the SNPs to exclude redundant SNPs, and subsequently we tested the SNPs present in both cohorts for allele frequency differences. All SNPs the allele frequencies of which differed significantly between the Mexicans and Finns as well as the SNPs present only in the Mexicans were labeled as CPAS-variants. Finally, in the GWAS step, we tested the CPAS-variants for association with the trait between the Mexican cases and controls.



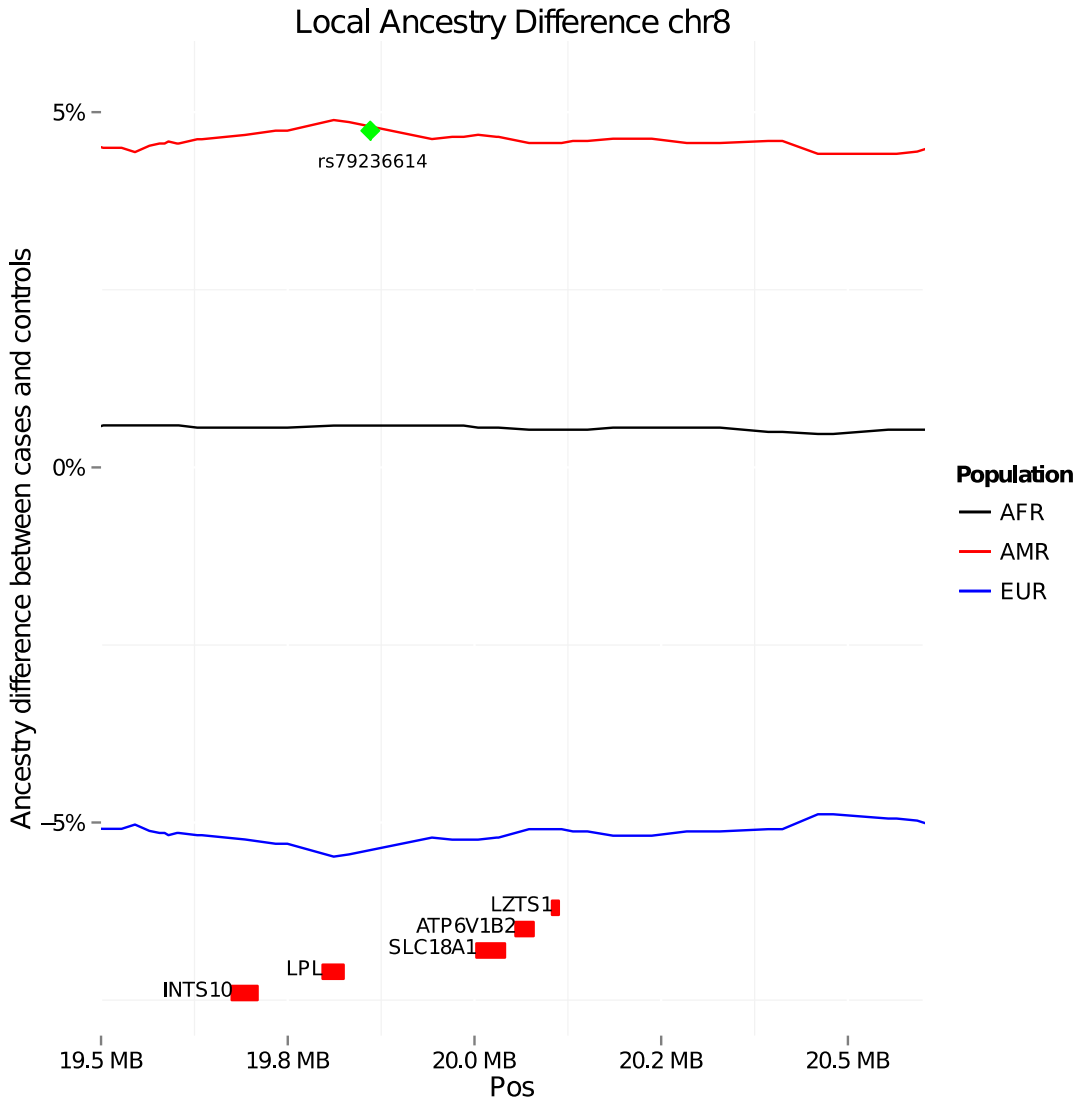
Supplementary Figure 2. Regional LD analysis of rs139961185. The novel TG-associated variant is not in LD with variants in *APOA5* or *APOC3* (R^2<0.2), indicating that there is a separate TG-associated signal in *SIK3* from these known TG-associated genes. The red line indicates R^2=0.2.



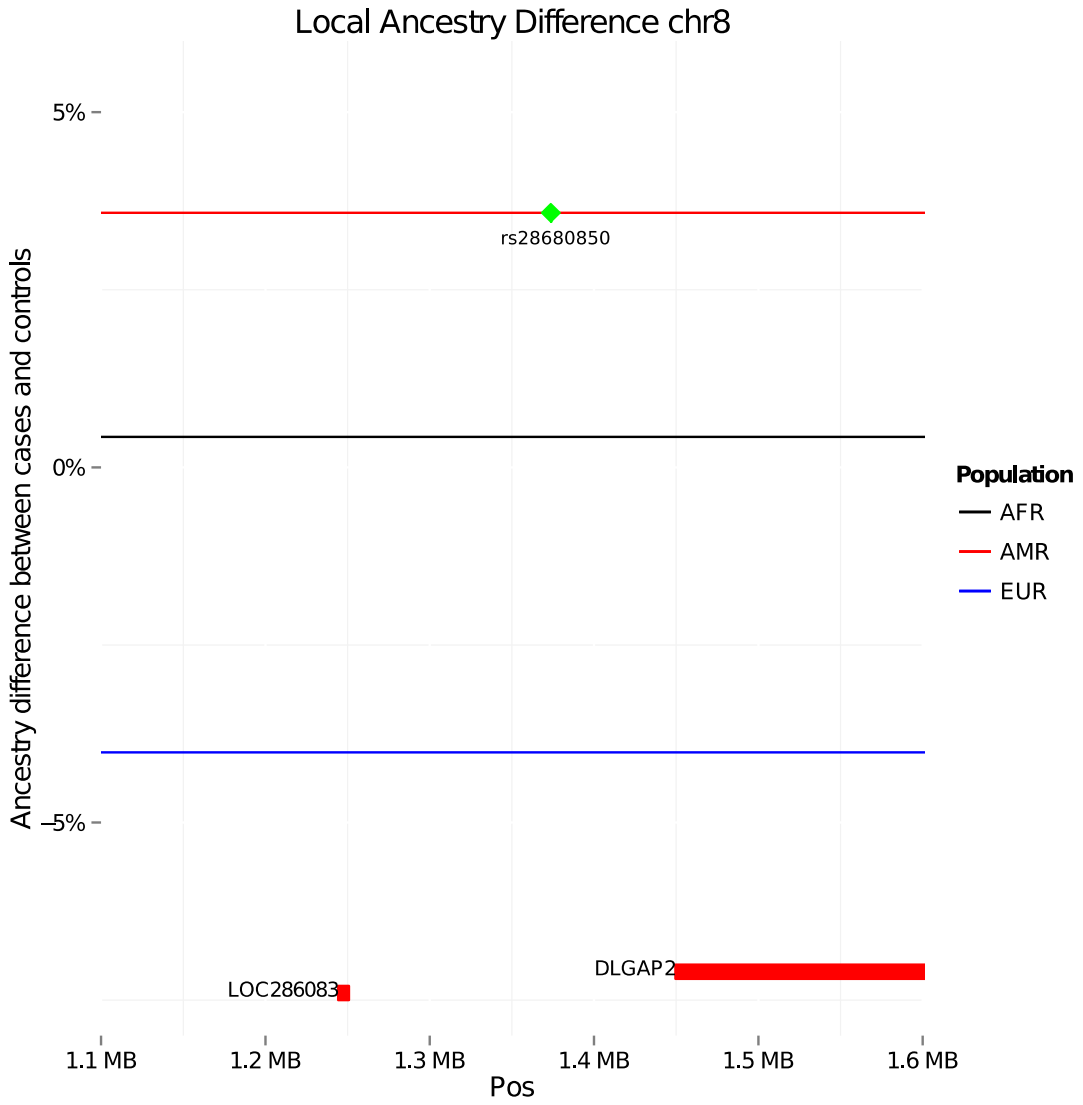
Supplementary Figure 3. Regional LD analysis of rs148533712. The new HDL-C-associated variant does not show strong LD with any variants within *LIPC* and it is 2.3 Mb away from the transcription start site of *LIPC*, indicating that rs148533712 in *RORA* is an independent HDL-C signal from *LIPC*. The red line indicates R^2=0.2.



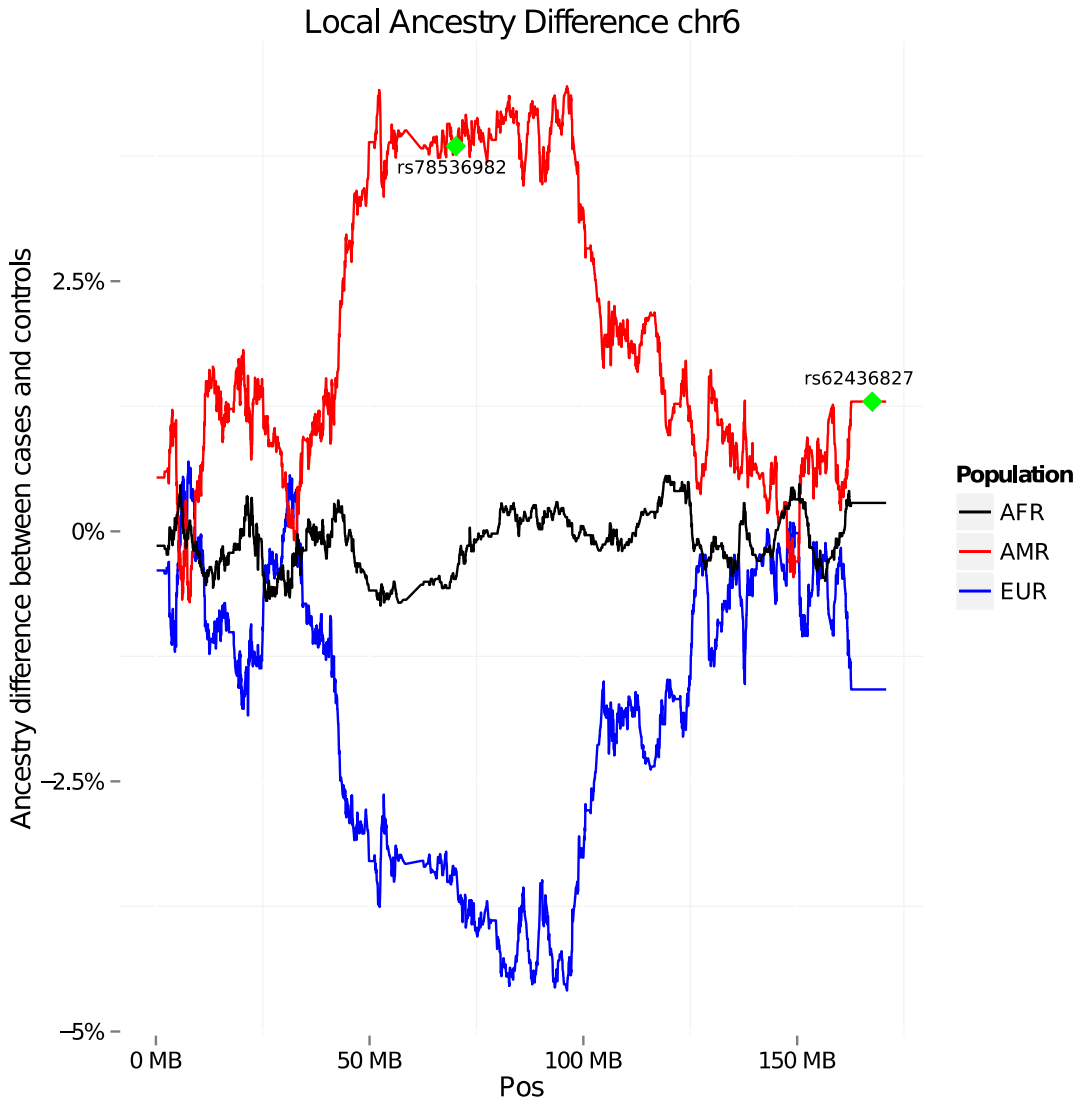
Supplementary Fig. 4. Closed-up view of the ancestry difference between the Mexican TG cases and controls in the chr11 TG risk haplotype region. Variants defining the TG risk haplotype are shown as green diamonds. Genes in the region are shown at the bottom.



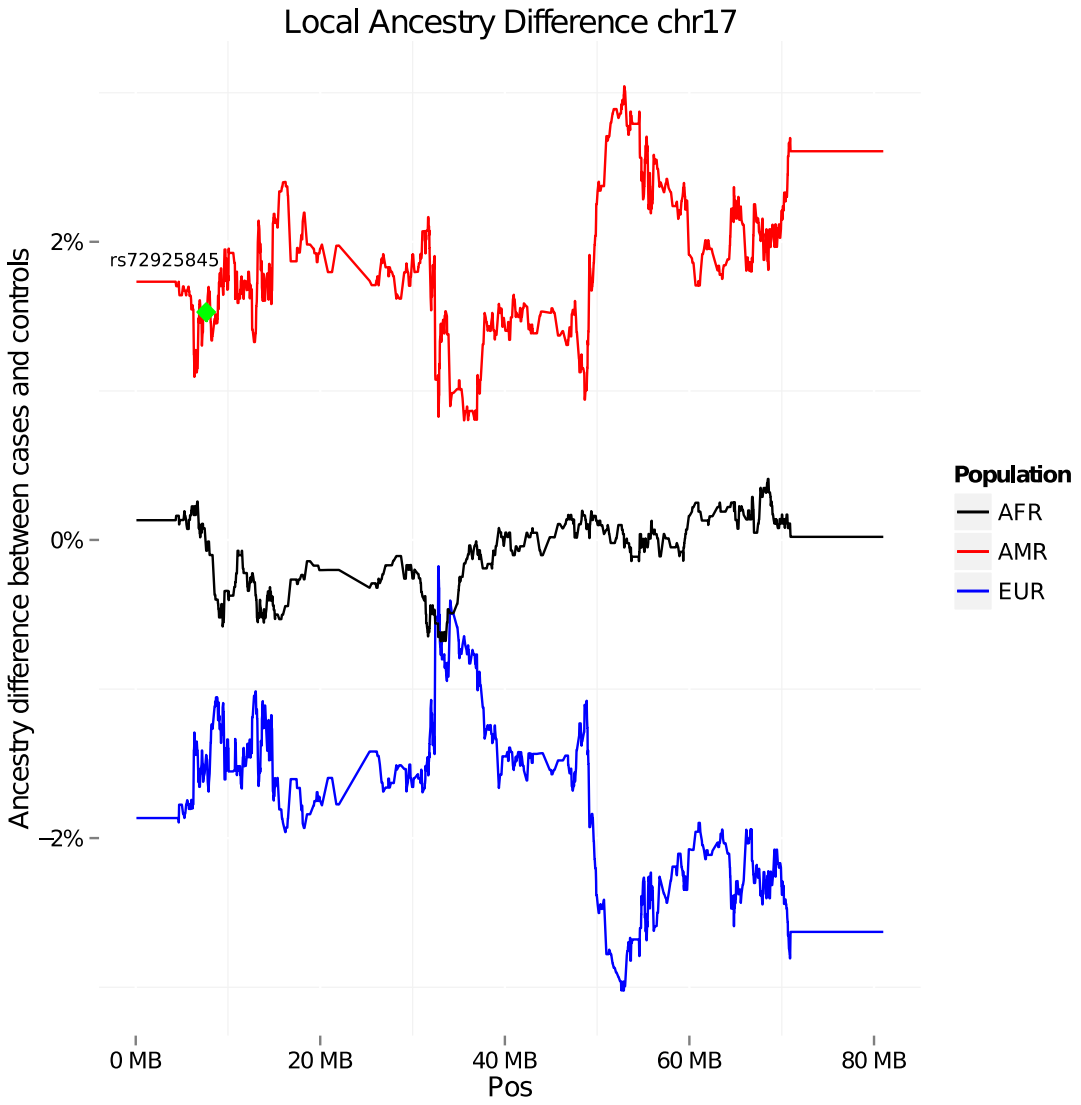
Supplementary Fig. 5. Closed-up view of the ancestry difference between the Mexican TG cases and controls on chr8p23.3. The variant rs79236614 shows a TG association in an Amerindian ancestry enriched region of the Mexican TG cases. Genes in the region are displayed at the bottom.



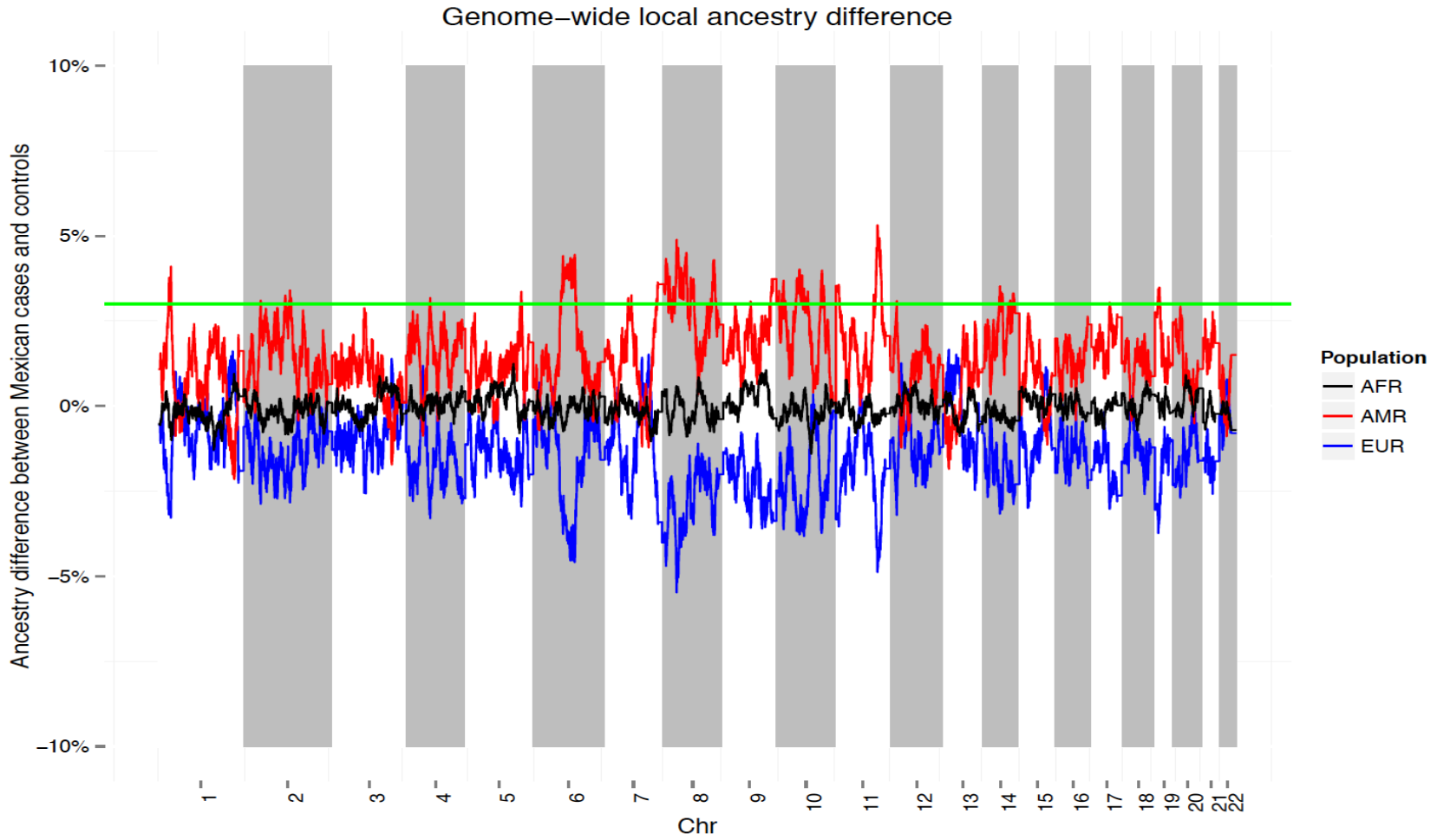
Supplementary Fig. 6. Closed-up view of the ancestry difference between the Mexican TG cases and controls on chr8p21. The variant rs28680850 shows a TG association in an Amerindian ancestry enriched region of the Mexican TG cases. Genes in the region are displayed at the bottom.



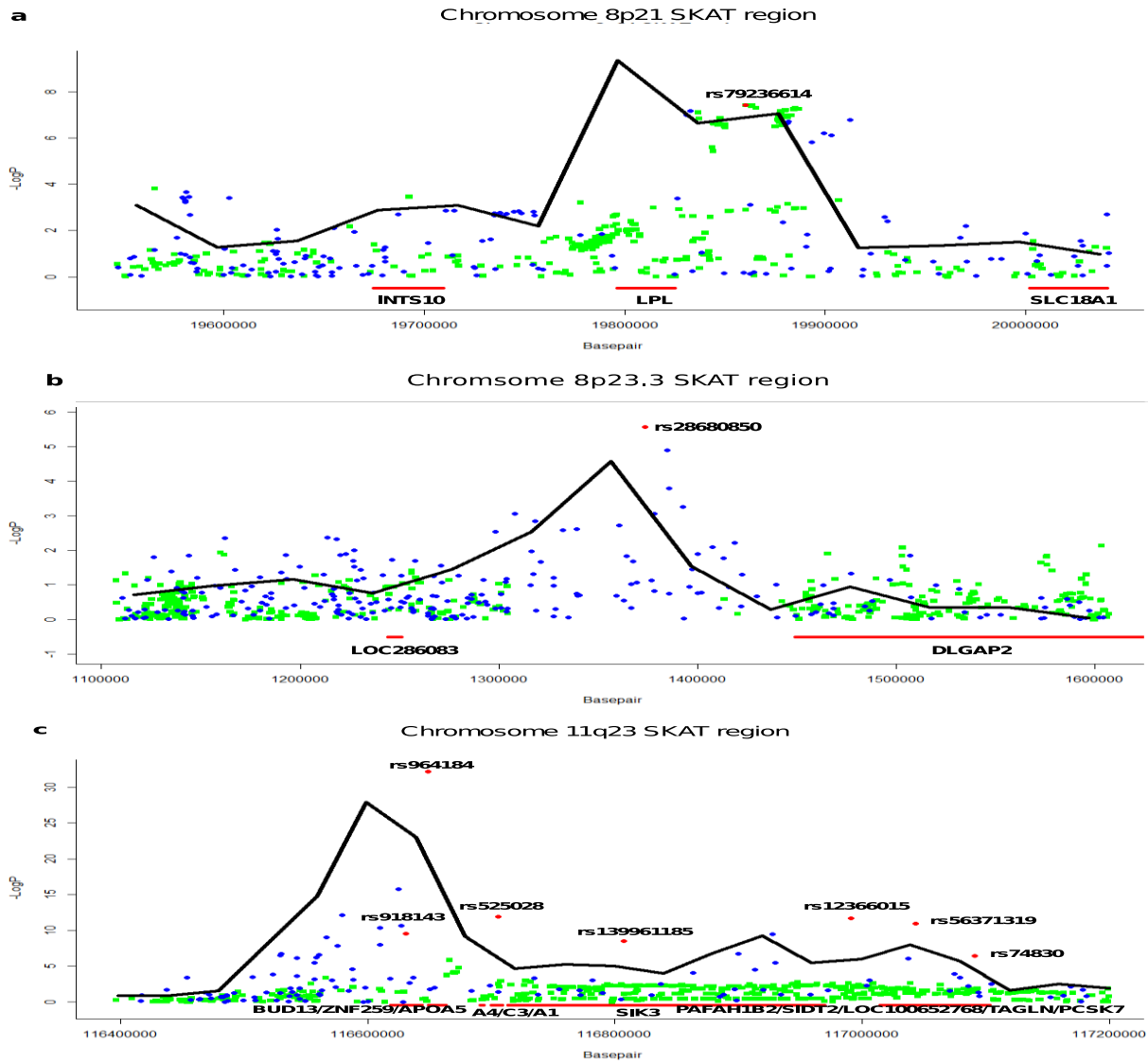
Supplementary Fig. 7. Local ancestry difference between the Mexican low TG controls and high TG cases on chr6 with suggestive TG CPAS-variants. Both variants indicated by green diamonds show suggestive TG signals at the replication stage. Rs78536982 lands in a region that is highly enriched for the Amerindian ancestry in the Mexican TG cases (>3%).



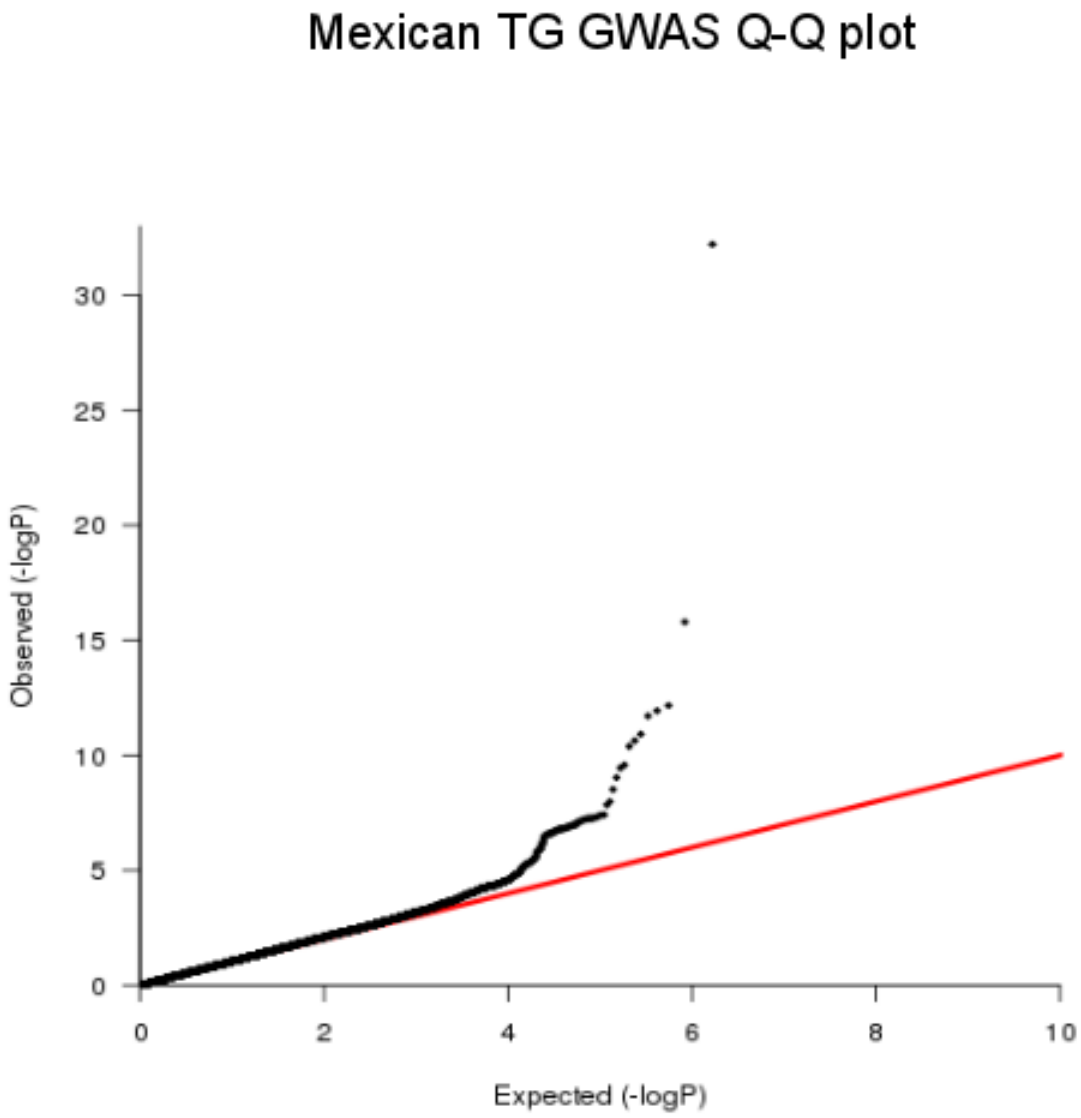
Supplementary Fig. 8. Local ancestry difference between the Mexican low TG controls and high TG cases on chr17 with suggestive TG CPAS-variants. The variant rs72925845 displayed as a green diamond has a suggestive TG-association signal at the replication stage.



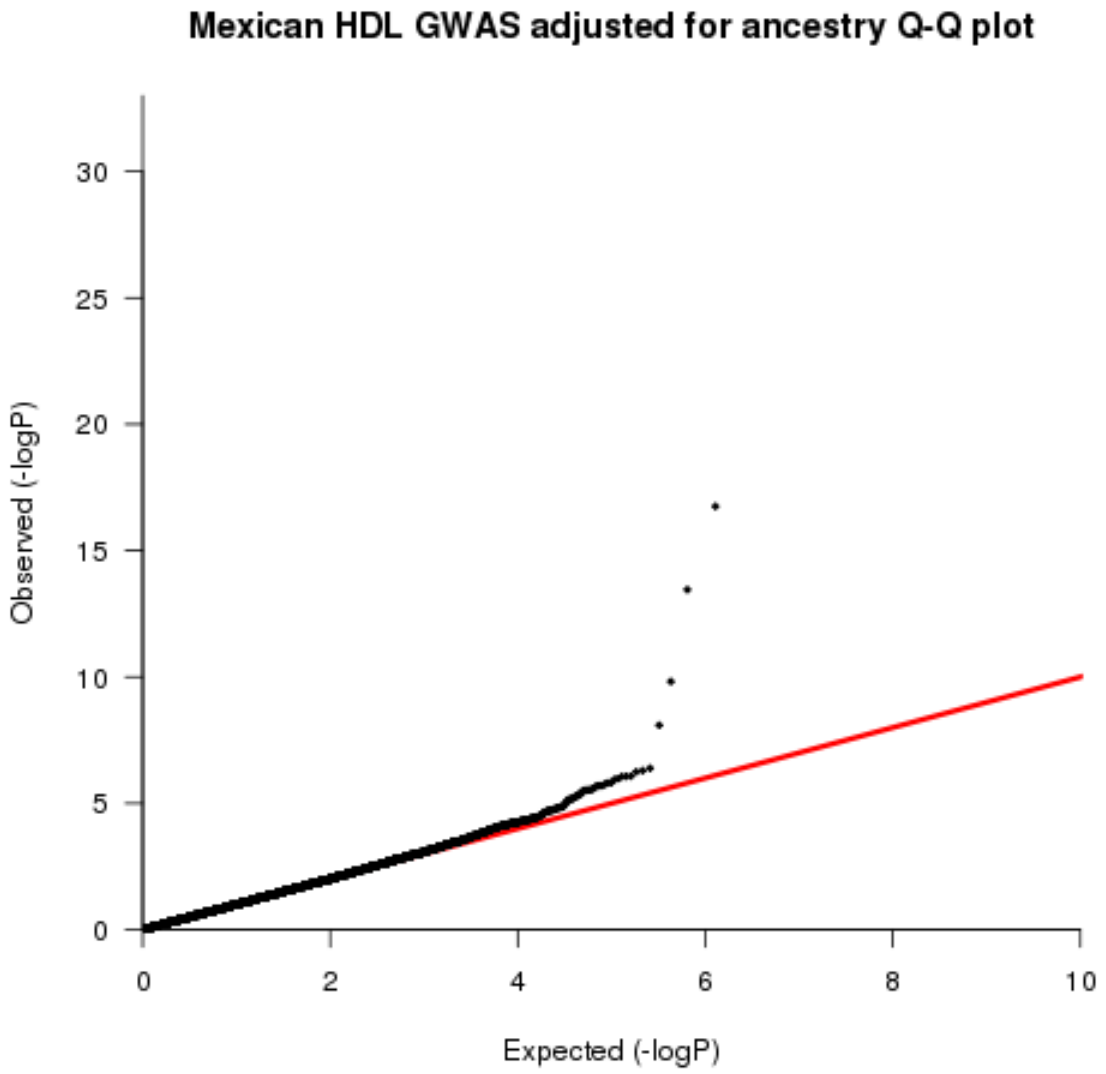
Supplementary Figure 9. Genome-wide local ancestry difference between the Mexican low TG controls and high TG cases. Local ancestry results are shown for the whole genome. We observed the highest Amerindian ancestry enrichment on chr6, chr8, and chr11, corresponding to our top CPAS-GWAS TG loci. The green line indicates a 3% ancestry difference.



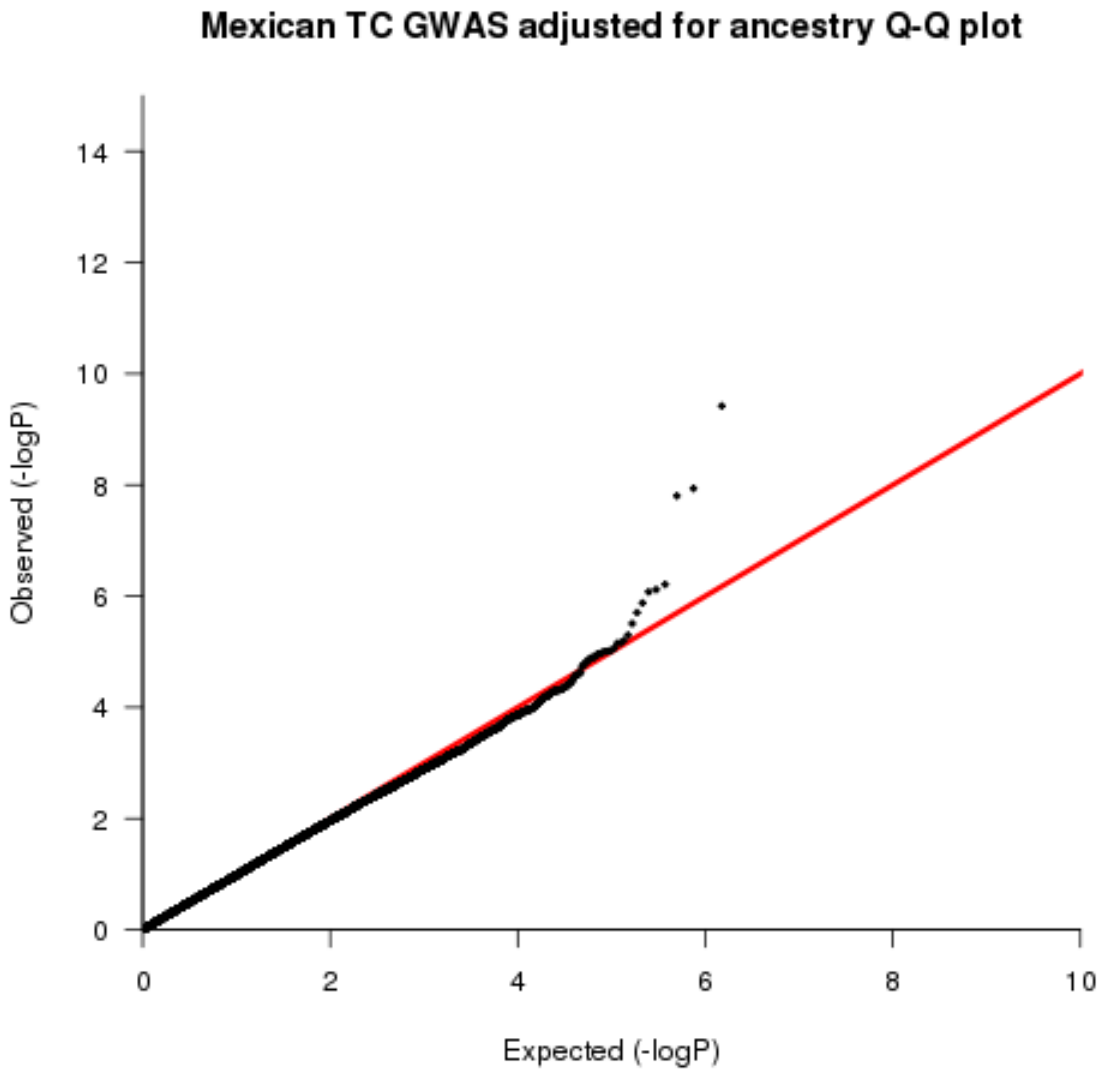
Supplementary Figure 10. Regional SKAT-C and GWAS association results with CPAS-SNPs. SKAT-C computes region-based test statistics and weights test statistics of variants based on their frequencies. Because variants with MAF<1% were disregarded as part of our imputation QC, we set the threshold between common and rare variants to 5%. Accordingly, we considered variants with a MAF 1-5% as rare to increase the number of rare variants in SKAT. As an alternative approach, we also utilized a rare variant frequency cutoff of 1.2% based on our sample size (see online Methods). The black line indicates the p-values from each SKAT window. Rare variants are labeled as green squares and common ones as blue dots. (A) The SKAT-C results on chr8p21. A peak in front of the lipoprotein lipase (LPL) gene is the second most significant p-value across the whole genome, possibly due to a large amount of TG-associated Amerindian rare variants in the region. (B) The SKAT-C results on chr8p23.3. The SKAT peak next to rs28680850 provides additional evidence for the regional association with TGs. (C) The SKAT-C results on chr11q23. The chr11q23 region displays the strongest SKAT signal in the whole genome, indicating a strong combined effect of rare and common Amerindian risk variants on the TG status.



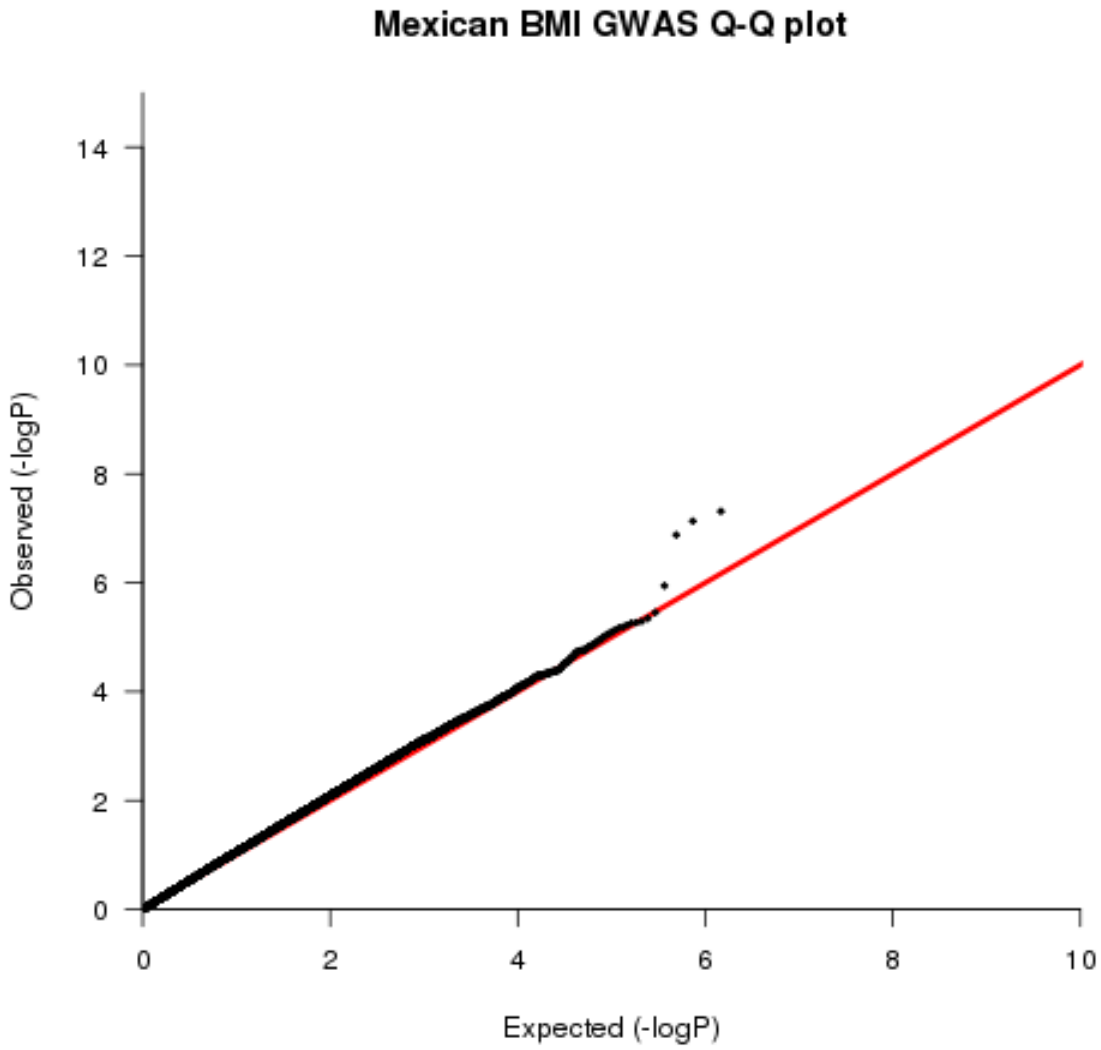
Supplementary Figure 11. The quantile-quantile (Q-Q) plot of the GWAS results with the CPAS SNPs between the Mexican TG cases and controls. Most of the distribution behaves as the expected null, ruling out major confounders. P-values were calculated using logistic regression (n = 1,678 TG cases and 1,645 controls).



Supplementary Figure 12. The quantile-quantile (Q-Q) plot of the HDL GWAS results with the CPAS SNPs adjusting for age, sex, BMI, global ancestry, and high TG status. Most of the distribution behaves as the expected null, ruling out major confounders. P-values were calculated using linear regression (n = 3,701 Mexicans).



Supplementary Figure 13. The quantile-quantile (Q-Q) plot of the TC GWAS results with the CPAS SNPs adjusting for age, sex, BMI, global ancestry, and high TG status. Most of the distribution behaves as the expected null, ruling out major confounders. P-values were calculated using linear regression (n = 3,701 Mexicans).



Supplementary Figure 14. The quantile-quantile (Q-Q) plot of the BMI GWAS results with the CPAS SNPs adjusting for age and sex. Most of the distribution behaves as the expected null, ruling out major confounders. P-values were calculated using linear regression (n= 3,701 Mexicans).

Supplementary Table 1. CPAS-GWAS and replication results of other SNPs for TGs.

SNP	Chr	Position	MAF (risk allele)	GWAS P	GWAS OR(95% CI)	Replication: qualitative** P	Replication: qualitative** Z	Replication: quantitative** P	Replication: quantitative** Z	Type	Nearby genes
rs72880046	2	171026420	16/3/6(C)	$5.25x10^{-6}$	1.73(1.36-2.20)	ns	0.40	ns	0.59	intergenic	UBR3,MYO3B
rs80020137	4	147964061	0/2/4(G)	$8.86x10^{-6}$	1.96(1.45-2.66)	ns	-0.23	ns	-0.96	intergenic	TTC29,EDNRA
rs80324749	8	16363770	0/4/2(C)	$8.78x10^{-6}$	0.52(0.39-0.70)	ns	0.14	ns	0.19	intergenic	MSR1,FGF20
rs62511975	8	58712381	16/4/2(A)	$1.38x10^{-6}$	0.51(0.38-0.67)	ns	0.67	ns	0.24	intergenic	LINC00588,FAM110B
rs62536962	9	7061167	16/36/42(T)	$6.59x10^{-6}$	1.25(1.13-1.38)	ns	1.35	ns	1.51	intronic	KDM4C
rs7854003	9	130310834	13/48/53(G)	$8.30x10^{-6}$	1.25(1.14-1.38)	ns	-0.54	ns	-1.56	intronic	FAM129B
rs2845791	21	38035224	16*/25/30(T)	$8.64x10^{-6}$	1.26(1.13-1.40)	ns	-0.13	ns	1.21	intergenic	CLDN14,SIM2

MAFs are listed in the following order: Finnish low TG controls/Mexican low TG controls/Mexican high TG cases. MAF indicates the minor allele frequency; P, p-value calculated using linear regression and logistic regression for quantitative and qualitative TG traits, respectively; OR, odds ratio; Chr, chromosome; CI, confidence interval; ns, $P \geq 0.05$; and Z, the standard score from meta-analysis. *Finnish MAFs of these SNPs were obtained from the Finnish population in the1000 Genomes Project as they were missing in our Finnish cohort. **Meta-analysis of the family and unrelated cohorts in the replication stage.

Supplementary Table 2. Clinical characteristics of the Finnish low TG controls, Mexican low TG controls, Mexican high TG cases, and Mexican replication cohorts.

		N (% female)	Age	BMI	TG (mmol/l)
Finnish	NFBC66	4,427(55%)	31*	24.16±3.91	0.95±0.32
	YFS	1,428(39%)	41.73±5.01	25.77±4.72	0.93±0.31
	HBCS	991(61%)	61.44±2.94	26.63±4.36	1.11±0.30
	GenMets	1,301(58%)	50.83±10.96	26.35±4.29	1.11±0.28
	Twins	421(60%)	22.85±1.43	22.82±3.37	0.92±0.31
	FINRISK	1,223(53%)	53.63±12.60	26.12±4.35	1.05±0.32
	Finnish low TG controls	9,791(54%)	40.76±13.23	25.12±4.32	1.00±0.32
Mexican	low TG controls	1,645(57%)	47.79±13.32	27.15±4.03	1.18±0.31
	high TG cases	1,678(50%)	46.98±11.62	28.94±3.95	3.49±1.38
	Total	3,323(54%)	47.38±12.49	28.06±4.09	2.34±1.53
Mexican replication	Unrelated cohort	5,256	41.51±30.61	27.92±4.92	2.15±1.41
	Family cohort	903	36.97±18.04	25.67±5.21	2.44±2.61

*NFBC is a Finnish birth cohort in which all clinical information was collected at the age of 31.
TG indicates triglycerides; and BMI, body mass index. For age, BMI, and TGs, the mean +/- standard deviation is shown.

Supplementary Table 3. Chromosome 11 haplotype results between the Mexican TG cases and controls.

Haplotype	Frequency in cases	Frequency in controls	CHISQ	DF	P-value
Omnibus	na	na	158.10	18	1.93×10^{-24}
2221211	0.12	0.17	24.83	1	6.25×10^{-7}
1211112	0.11	0.13	6.95	1	0.008
1221112	0.01	0.02	3.11	1	ns
2211112	0.01	0.02	0.95	1	ns
1221211	0.17	0.21	10.90	1	0.001
1112122 (=HT1)	0.18	0.11	50.34	1	1.29×10^{-12}
1121211	0.03	0.02	2.88	1	ns
1111211	0.02	0.01	2.84	1	ns
1111112 (=HT2)	0.12	0.08	32.85	1	9.96×10^{-9}
1121212	0.01	0.01	0.56	1	ns
1221212	0.02	0.03	8.84	1	0.003
1211212	0.02	0.02	0.85	1	ns
1111122 (=HT3)	0.07	0.05	18.88	1	1.40×10^{-5}
2212122	0.01	0.01	0.60	1	ns
1212122	0.02	0.02	0.79	1	ns
1211122	0.03	0.03	1.21	1	ns
2221212	0.01	0.01	2.62	1	ns
2211212	0.01	0.02	6.96	1	0.008
2221222	0.01	0.01	0.22	1	ns

The order of the SNPs on the haplotype is rs918143 (1/C), rs964184 (1/G), rs525028 (1/G), rs139961185 (2/A), rs12366015 (1/A), rs56371319 (2/A), and rs74830 (2/T) with the TG-increasing allele given in parenthesis. CHISQ indicates the chi-square value; na, not applicable; ns, not significant (P>0.05); and DF, degrees of freedom. P-values were calculated using the haplotype case/control test in PLINK.

Supplementary Table 4. Results of the linear regression of the chr11 SNPs by the HT1 and HT2 haplotype carrier-status.

SNP ID	Position on chr11	Risk allele	MAF in haplotype carriers*	MAF in non-haplotype carriers**	OR	P-value	LD (D')***
HT1 haplotype							
rs11820589	116633862	A	0.37	0.08	6.62	$1.67x10^{-202}$	0.53
rs10488698	116633947	A	0.09	0.16	0.52	$2.71x10^{-15}$	1.00
rs675	116691675	A	0.04	0.11	0.37	$3.72x10^{-19}$	1.00
rs4520	116701535	C	0.27	0.56	0.29	$9.24x10^{-126}$	1.00
rs12225230	116728630	C	0.47	0.15	5.01	$2.79x10^{-183}$	0.65
chr11:117076875	117076875	A	0.01	0.05	0.16	$3.68x10^{-18}$	1.00
rs2277287	117096652	A	0.64	0.21	6.72	$9.46x10^{-276}$	1.00
rs2306473	117097952	T	0.63	0.16	9.17	0	1.00
HT2 haplotype							
rs662799	116663707	G	0.51	0.12	7.37	$4.31x10^{-206}$	0.95

*Indicates the MAF in the carriers of the chr11 risk haplotype (1/X, 1/1), and **indicates the MAF in the non-carriers on the risk haplotype (X/X), 1 being the risk haplotype and X any other haplotype of the 7 SNPs (rs918143, rs964184, rs525028, rs139961185, rs12366015, rs56371319, and rs74830) for HT1 or HT2 haplotype (see Table 2). OR indicates odds ratio for the SNP between the haplotype carriers (1/1, 1/X) and non-carriers (X/X); and LD (D') linkage disequilibrium (in D') between the SNP and haplotype carrier status coded as a genotype (1/1, 1/X, X/X) for HT1 or HT2 haplotype. ***LD in R^2 was <0.50 for all SNPs between the SNP and haplotype carrier status coded as a genotype (1/1, 1/X, X/X) for HT1 or HT2 haplotype.