Supplemental Information for

"Trans-ethnic Evaluation Identifies Novel Low Frequency Loci Associated with 25-Hydroxyvitamin D Concentrations"

Hong and Hatchell, et al.

Supplemental Table 1. TRANSCEND cohorts' specific quality control.

Ethnicity	Cohort	Genotyping Platform and SNP Panel	n	SNP quality control				Sample quality control	
				Call rate	Minor Allele Frequency	Hardy-Weinberg Equilibrium	# Passed SNPs	Call rate	Exlcusions
African American	African American Diabetes Heart Study (AADHS)	Illumina Omni5 array	531	95%	$\geq$0.01	$p>10^{-6}$	2,298,937	99%	Ethnic outliers
	Atherosclerosis Risk in Communities Study (ARIC)	Affymetrix Genome-Wide Human SNP Array 6.0	2658	95%	$\geq$0.01	$p>10^{-5}$	806,416	95%	Missing covariates, Ethnic outliers, Non-consent, Duplicates, Relatedness, Mixed/contaminated samples, Samples from a problem plate, Sex mismatches, Genotype mismatches, Genetic outliers
	Bone Mineral Density in Childhood Study (BMDCS)	Illumina Infinium™ II OMNI Express plus Exome BeadChip	161	95%	$\geq$0.01	$p>10^{-6}$	739,284	95%	N/A
	Children's Hospital of Philadelphia (CHOP)	Illumina 550v1, 550v3, 610 Quad	379	95%	$\geq$0.01	$p>10^{-6}$	529,957	95%	Ethnic outliers
	Cardiovascular Health Study (CHS)	HumanOmni1-Quad_v1 BeadChip	303	97%	$\geq$0.01	$p>10^{-5}$	963,248	95%	Serum vitamin D levels >100ng/ml
	Health, Aging and Body Composition Study (HABC)	Illumina 1M	981	95%	$\geq$0.01	$p>10^{-5}$	14,391,447	95%	Missing/impossible covariates, Ethnic outliers, No GWAS data
	Mount Sinai BioMe BioBank	Illumina HumanOmniExpress-12v1_H	361	95%	$\geq$0.01	$p>10^{-5}$	608,364	90%	Ethnic outliers
	Jackson Heart Study (JHS)	Affymetrix Genome-Wide Human SNP Array 6.0 (906K)	2132	95%	$\geq$0.01	$p>10^{-6}$	15,000,000	95%	Duplicates from ARIC cohort, Missing covariates
	Multi-Ethnic Study of Atherosclerosis (MESA)	Affymetrix Genome-Wide Human SNP Array 6.0	1035	95%	$\geq$0.01	$p>10^{-6}$	637,631	95%	Relatedness, Serum vitamin D levels >100ng/ml
Hispanic	Boston Puerto Rican Health Study Projects (BPRHS)	Affymetrix Axiom World LAT 1 Array Plate	1360	95%	$\geq$0.01	$p>10^{-6}$	717,275	95%	Missing covariates, No GWAS data
	Insulin Resistance Atherosclerosis Study Family Study (IRASFS)	Illumina HumanOmniExpress BeadChip	738	95%	$\geq$0.05	$p>10^{-4}$	672,436	99%	Ethnic outliers, Sex mismatches, Fingerprinting non-concordance, Failed pedcheck, Duplicates
	Multi-Ethnic Study of Atherosclerosis (MESA)	Affymetrix Genome-Wide Human SNP Array 6.0	1387	95%	$\geq$0.01	$p>10^{-6}$	398,246	95%	Relatedness Serum vitamin D levels >100ng/ml

Supplemental Table 2. Genome-wide Significant Vitamin D SNPs in a Trans-ethnic Evaluation without removal for Linkage Disequilibrium

SNP INFO						Trans-ethnic Zscore Metal			
SNP	**Chr**	**Position (build37)**	**Ref.**	**Alt.**	**Locus**	**RAF**	**Sample Size**	**Zscore**	**P.value**
rs2282679	4	72608383	a	c	*GC*	0.7742	24497	16.628	4.40E-62
rs3755967	4	72609398	a	g	*GC*	0.2228	24982	-15.469	5.59E-54
rs1155563	4	72643488	t	c	*GC*	0.7584	22851	15.031	4.57E-51
rs7041	4	72618334	a	c	*GC*	0.5617	24729	-14.796	1.56E-49
rs2298850	4	72614267	c	g	*GC*	0.2126	23446	-14.435	3.10E-47
rs222040	4	72616932	a	g	*GC*	0.5125	21286	14.402	5.05E-47
rs17467825	4	72605517	a	g	*GC*	0.7799	21193	14.372	7.76E-47
rs222047	4	72610208	a	c	*GC*	0.559	23607	-12.91	3.96E-38
rs222035	4	72621674	t	g	*GC*	0.5811	21202	-12.826	1.17E-37
rs705120	4	72614140	a	c	*GC*	0.4378	23505	-12.205	2.92E-34
rs4694105	4	72592214	a	g	*GC*	0.2806	24113	-11.858	1.95E-32
rs842999	4	72611690	c	g	*GC*	0.4055	19092	11.802	3.80E-32
rs2201124	4	72597009	t	c	*GC*	0.2764	21502	-11.711	1.11E-31
rs16846876	4	72592491	a	t	*GC*	0.7251	24441	11.424	3.19E-30
rs1526692	4	72578724	a	g	*GC*	0.5851	24420	11.306	1.23E-29
rs7679037	4	72573949	c	g	*GC*	0.5682	21960	10.507	8.03E-26
rs7682810	4	72574208	a	g	*GC*	0.4307	22034	-10.22	1.62E-24
rs12512631	4	72601331	a	g	*GC*	0.6555	21962	-10.144	3.52E-24
rs1491711	4	72602252	c	g	*GC*	0.6547	22274	-9.277	1.75E-20
rs17383291	4	72605410	t	g	*GC*	0.9377	18222	8.817	1.17E-18
rs377687	4	72569065	a	g	*GC*	0.345	23402	8.389	4.92E-17
rs395371	4	72572374	a	g	*GC*	0.3408	26411	8.072	6.91E-16
rs962227	4	72707517	a	g	*GC*	0.3667	24120	-7.747	9.37E-15
rs394295	4	72575635	c	g	*GC*	0.3402	21963	7.742	9.82E-15
rs842873	4	72696465	c	g	*GC*	0.3794	19054	7.696	1.41E-14
rs1402155	4	72705908	a	g	*GC*	0.3706	24431	-7.642	2.14E-14
rs1476272	4	72706235	a	g	*GC*	0.3586	24117	-7.572	3.67E-14
rs396047	4	72572134	a	g	*GC*	0.3412	24412	7.557	4.14E-14
rs842872	4	72695059	a	g	*GC*	0.3691	20415	7.483	7.25E-14
rs393417	4	72570636	a	t	*GC*	0.6537	22041	-7.446	9.61E-14
rs367399	4	72573000	t	c	*GC*	0.6535	22043	-7.391	1.46E-13
rs6837549	4	72596821	a	c	*GC*	0.4601	21193	7.368	1.73E-13
rs17765181	4	72569468	a	t	*GC*	0.277	22035	7.34	2.13E-13
rs17767445	4	72745267	a	g	*GC*	0.1541	20417	7.198	6.10E-13

Supplemental Table 2 (cont'd)

SNP	Chr	Position (build37)	Ref.	Alt.	Locus	RAF	Sample Size	Zscore	P.value
rs962225	4	72707356	a	g	*GC*	0.3655	22042	-6.96	3.39E-12
rs406416	4	72570885	a	g	*GC*	0.3442	19834	6.514	7.31E-11
rs10013907	4	72720231	t	c	*GC*	0.3425	22003	-6.384	1.72E-10
rs13113067	4	72739098	a	c	*GC*	0.4051	21976	-5.942	2.81E-09
rs1356876	4	72728303	a	g	*GC*	0.5958	21972	5.92	3.21E-09
rs6858790	4	72738679	a	g	*GC*	0.4035	23011	-5.789	7.10E-09
rs10021949	4	72764039	a	g	*GC*	0.2583	19813	5.74	9.45E-09
rs1352844	4	72647749	a	g	*GC*	0.1251	24119	5.667	1.45E-08
rs6814839	4	72751722	a	g	*GC*	0.8849	26505	5.61	2.02E-08
rs10518092	4	72696825	t	c	*GC*	0.1249	23283	5.5	3.79E-08
rs12643936	4	72783474	t	c	*GC*	0.8288	22025	5.472	4.45E-08
rs10937986	4	72792521	a	g	*NPFFR2*	0.5617	24388	-7.022	2.18E-12
rs1520497	4	72800493	a	c	*NPFFR2*	0.4437	24115	6.597	4.18E-11
rs13107347	4	72974748	a	g	*NPFFR2*	0.6745	26517	-6.145	7.98E-10
rs10035034	4	72792667	a	g	*NPFFR2*	0.7705	23062	-5.91	3.42E-09
rs17774118	4	72796639	t	c	*NPFFR2*	0.7667	23035	-5.766	8.12E-09
rs1520499	4	72791849	a	g	*NPFFR2*	0.8366	24104	5.696	1.23E-08
rs4694110	4	72801249	t	c	*NPFFR2*	0.7955	23540	-5.545	2.93E-08
rs6838307	4	73052613	c	g	*NPFFR2*	0.9471	17154	-5.483	4.19E-08
rs7675387	4	72921718	a	c	*NPFFR2*	0.2779	23298	5.463	4.68E-08
rs11606033	11	71154204	a	g	*DHCR7*	0.3639	21969	-7.674	1.66E-14
rs1044482	11	71155171	a	g	*DHCR7*	0.6363	21976	7.673	1.67E-14
rs7928249	11	71161063	a	g	*DHCR7*	0.3595	24210	-7.478	7.57E-14
rs3750997	11	71158841	a	c	*DHCR7*	0.3536	26449	-7.129	1.01E-12
rs1630498	11	71150785	a	c	*DHCR7*	0.7637	24018	6.347	2.20E-10
rs11603330	11	71153459	a	c	*DHCR7*	0.5797	21972	6.337	2.34E-10
rs12422045	11	71154820	a	g	*DHCR7*	0.5798	21975	6.333	2.40E-10
rs12797951	11	71143266	a	c	*DHCR7*	0.4001	24326	-6.065	1.32E-09
rs1790345	11	71145941	a	g	*DHCR7*	0.2473	21938	-5.58	2.40E-08
rs12419334	11	71139472	a	g	*DHCR7*	0.3985	21233	-5.478	4.31E-08
rs7118246	11	71231124	c	g	*KRTAP5-7*	0.6564	22037	6.36	2.02E-10
rs12800438	11	71171003	a	g	*NADSYN1*	0.6463	24114	8.284	1.20E-16
rs7938885	11	71170043	t	c	*NADSYN1*	0.366	22039	-8.125	4.47E-16
rs4423214	11	71173254	t	c	*NADSYN1*	0.6335	22034	8.063	7.45E-16
rs2276362	11	71174452	a	g	*NADSYN1*	0.369	22039	-7.967	1.62E-15

Supplemental Table 2 (cont'd)

SNP	Chr	Position (build37)	Ref.	Alt.	Locus	RAF	Sample Size	Zscore	P.value
rs2040323	11	71172783	a	c	*NADSYN1*	0.3392	22038	-7.702	1.35E-14
rs4944062	11	71187294	t	g	*NADSYN1*	0.634	22044	7.653	1.97E-14
rs4944957	11	71168035	a	g	*NADSYN1*	0.3278	23071	-7.604	2.88E-14
rs12785878	11	71167449	t	g	*NADSYN1*	0.6151	21105	7.468	8.13E-14
rs7944926	11	71165625	a	g	*NADSYN1*	0.3963	24109	-7.387	1.50E-13
rs4945008	11	71221248	a	g	*NADSYN1*	0.4302	23398	-7.36	1.84E-13
rs4944958	11	71168073	a	g	*NADSYN1*	0.4149	22036	-7.338	2.17E-13
rs3794060	11	71187679	a	g	*NADSYN1*	0.5972	24118	7.291	3.08E-13
rs11600569	11	71173747	c	g	*NADSYN1*	0.4192	22036	-7.153	8.47E-13
rs4944959	11	71168234	t	c	*NADSYN1*	0.4118	22039	-7.098	1.26E-12
rs7935125	11	71184190	a	c	*NADSYN1*	0.5808	22044	7.08	1.44E-12
rs12790558	11	71170905	t	c	*NADSYN1*	0.4122	22039	-7.074	1.50E-12
rs12808368	11	71170817	t	c	*NADSYN1*	0.5913	22034	7.001	2.55E-12
rs2282620	11	71183477	t	c	*NADSYN1*	0.6132	22041	6.744	1.55E-11
rs2282621	11	71183474	a	g	*NADSYN1*	0.6132	22041	6.743	1.56E-11
rs1792226	11	71176574	a	g	*NADSYN1*	0.2349	24113	-6.443	1.18E-10
rs7940244	11	71221574	a	g	*NADSYN1*	0.2944	26513	-6.379	1.79E-10
rs4245442	11	71220909	t	c	*NADSYN1*	0.2954	22045	-6.369	1.90E-10
rs2186777	11	71219803	a	c	*NADSYN1*	0.6496	22713	6.359	2.03E-10
rs2002064	11	71163420	a	g	*NADSYN1*	0.7635	22039	6.316	2.69E-10
rs1540127	11	71179038	a	g	*NADSYN1*	0.2397	22035	-6.177	6.51E-10
rs1792234	11	71182018	t	c	*NADSYN1*	0.7603	22037	6.164	7.10E-10
rs1629220	11	71174553	t	c	*NADSYN1*	0.2396	22039	-6.12	9.38E-10
rs1790343	11	71177207	t	c	*NADSYN1*	0.7548	23400	5.883	4.02E-09
rs3829251	11	71194559	a	g	*NADSYN1*	0.1759	23958	-5.674	1.40E-08
rs2276360	11	71169547	c	g	*NADSYN1*	0.6566	23896	5.659	1.52E-08
rs1792229	11	71179395	a	g	*NADSYN1*	0.7589	24432	5.626	1.85E-08
rs10898223	11	71219264	a	g	*NADSYN1*	0.1647	22020	-5.624	1.87E-08
rs10898203	11	71203436	t	c	*NADSYN1*	0.1625	22017	-5.563	2.65E-08
rs719700	12	46029209	t	c	*ANOA6/ARID2*	0.9634	12087	5.555	2.78E-08
rs1410656	13	47542521	t	c	*HTR2A*	0.0592	12841	-5.896	3.72E-09

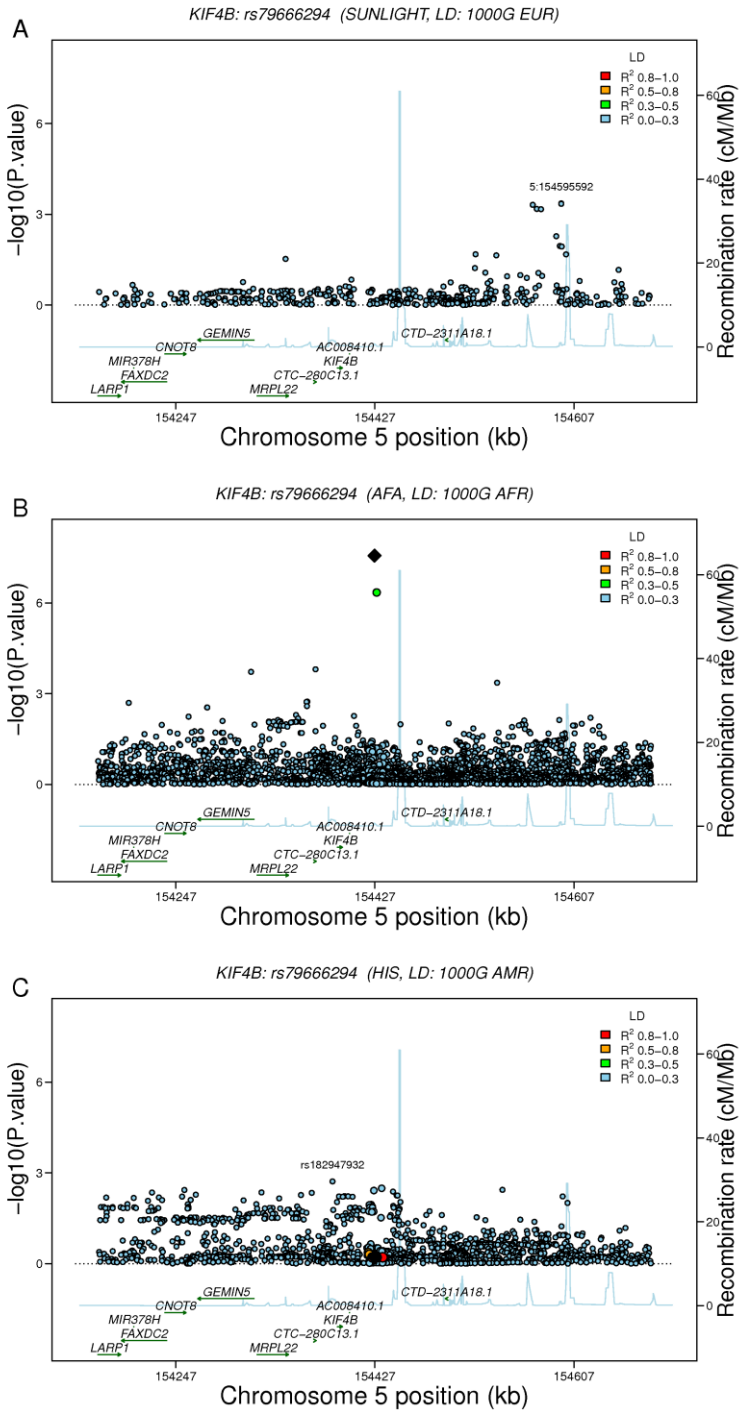
Genome-wide significance is defined as p <5E-08. Abbreviations used: Chr, chromosome; and RAF, reference allele frequency. The reference allele is the vitamin D raising allele. Position corresponds to build 37. The model was adjusted for age, sex, BMI, UV Index and prinicple components.

Supplemental Table 3. Tranferability of SNPs in European population to African, Hispanic and transethnic ancestry populations.

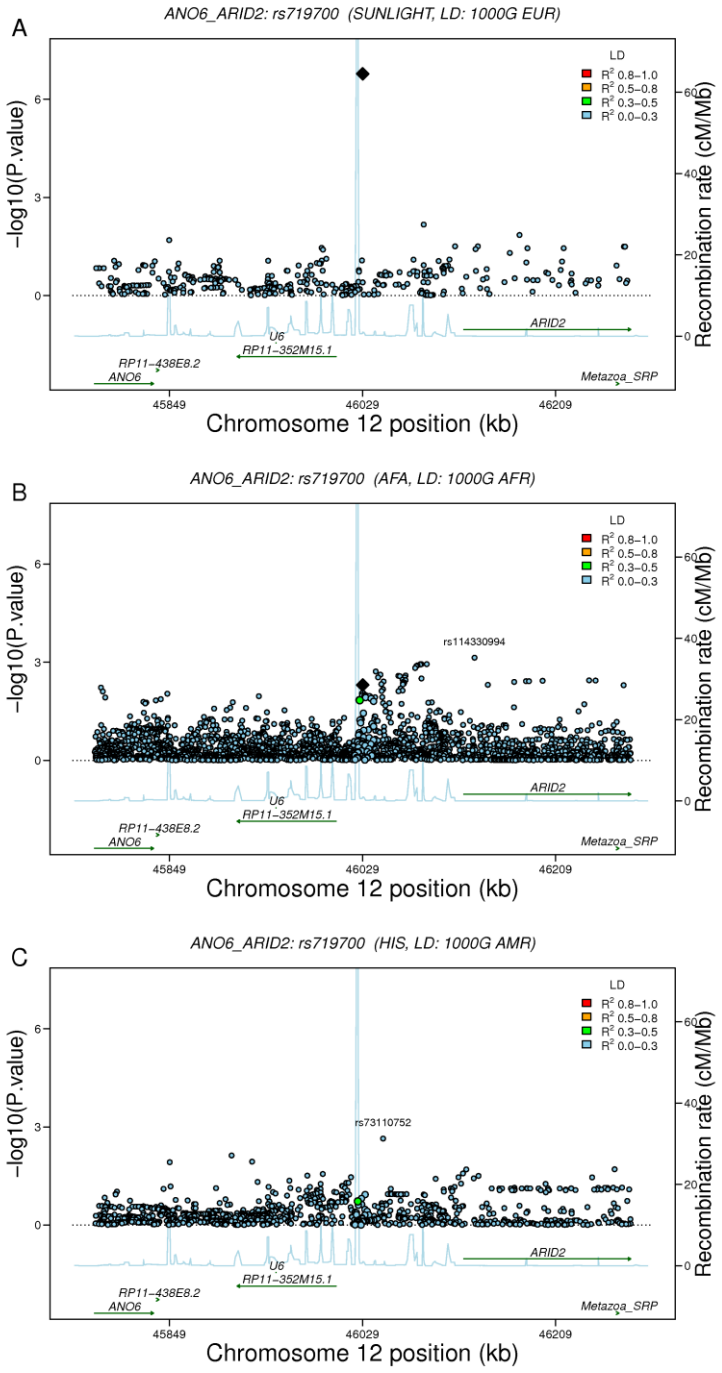
					SUNLIGHT		TRANSCEN-D trans-ethnic		TRANSCEN-D African		TRANSCEN-D Hispanic	
SNP	**Chr**	**Locus**	**Trans-ethnic Allele1**	**Trans-ethnic Allele2**	**direction of effect (N=33,996)**	**p-value**	**direction of effect (N=12,025)**	**p-value**	**direction of effect (N=8,540)**	**p-value**	**direction of effect (N=3,485)**	**p-value**
rs2282679	4	*GC*	a	c	+	$1.9X10^{-109}$	+	$4.398X10^{-62}$	+	$2.70X10^{-4}$	+	$1.00X10^{-6}$
rs2298850	4	*GC*	c	g	-	$2.03X10^{-71}$	-	$3.099X10^{-47}$	-	$5.98X10^{-5}$	-	$5.13X10^{-4}$
rs1155563	4	*GC*	t	c	+	$2.37X10^{-73}$	+	$4.568X10^{-51}$	+	$3.22X10^{-4}$	NA	NA
rs3755967	4	*GC*	a	g	-	$2.42X10^{-75}$	-	$5.585X10^{-54}$	-	$1.58X10^{-4}$	-	$8.95X10^{-9}$
rs7041	4	*GC*	a	c	-	$6.31X10^{-59}$	-	$1.561X10^{-49}$	-	$1.07X10^{-7}$	-	$2.60X10^{-5}$
rs17467825	4	*GC*	a	g	+	$6.75X10^{-74}$	+	$7.763X10^{-47}$	+	$3.39X10^{-4}$	NA	NA
rs12785878	11	*DHCR7/NADSYN1*	t	g	+	$2.12X10^{-27}$	+	$8.13X10^{-14}$	+	$1.56X10^{-2}$	+	$9.26X10^{-2}$
rs12800438	11	*DHCR7/NADSYN1*	a	g	+	$2.54X10^{-15}$	+	$1.20X10^{-16}$	+	$2.87X10^{-5}$	+	$4.40X10^{-1}$
rs3794060	11	*DHCR7/NADSYN1*	a	g	+	$3.38X10^{-15}$	+	$3.03X10^{-13}$	+	$1.70X10^{-2}$	+	$3.81X10^{-1}$
rs4945008	11	*DHCR7/NADSYN1*	a	g	-	$4.55X10^{-15}$	-	$1.84X10^{-13}$	-	$1.48X10^{-2}$	-	$1.00X10^{-1}$
rs4944957	11	*DHCR7/NADSYN1*	a	g	-	$8.70X10^{-15}$	-	$2.88X10^{-14}$	-	$8.62X10^{-4}$	-	$5.31X10^{-1}$
rs7944926	11	*DHCR7/NADSYN1*	a	g	-	$8.96X10^{-16}$	-	$1.50X10^{-13}$	-	$2.71X10^{-2}$	-	$3.46X10^{-1}$
rs2060793	11	*CYP2R1*	a	g	+	$1.73X10^{-11}$	+	$1.71X10^{-4}$	-	$9.12X10^{-1}$	-	$9.93X10^{-1}$
rs12794714	11	*CYP2R1*	a	g	-	$1.84X10^{-9}$	-	$1.52X10^{-5}$	-	$7.79X10^{-2}$	-	$5.38X10^{-1}$
rs10500804	11	*CYP2R1*	a	c	+	$2.67X10^{-9}$	+	$1.04X10^{-5}$	+	$1.14X10^{-1}$	+	$1.03X10^{-1}$
rs10741657	11	*CYP2R1*	a	g	+	$3.27X10^{-20}$	+	$2.28X10^{-6}$	+	$8.90X10^{-1}$	+	$1.66X10^{-1}$
rs7116978	11	*CYP2R1*	t	c	+	$4.99X10^{-9}$	+	$1.19X10^{-3}$	-	$7.45X10^{-1}$	-	$8.25X10^{-1}$
rs1993116	11	*CYP2R1*	a	g	+	$6.25X10^{-11}$	+	$2.83X10^{-5}$	+	$9.01X10^{-1}$	+	$1.34X10^{-1}$

European GWAS by Ahn, et al (N=4,501) findings: rs2282679 (direction: +, p-value: 2.0×10^{-30}), rs1155563 (direction: +, p-value: 3.8×10^{-25}), rs7041 (direction: -, p-value: 4.1×10^{-22}), rs1993116 (direction: +, p-value: 1.6×10^{-17})

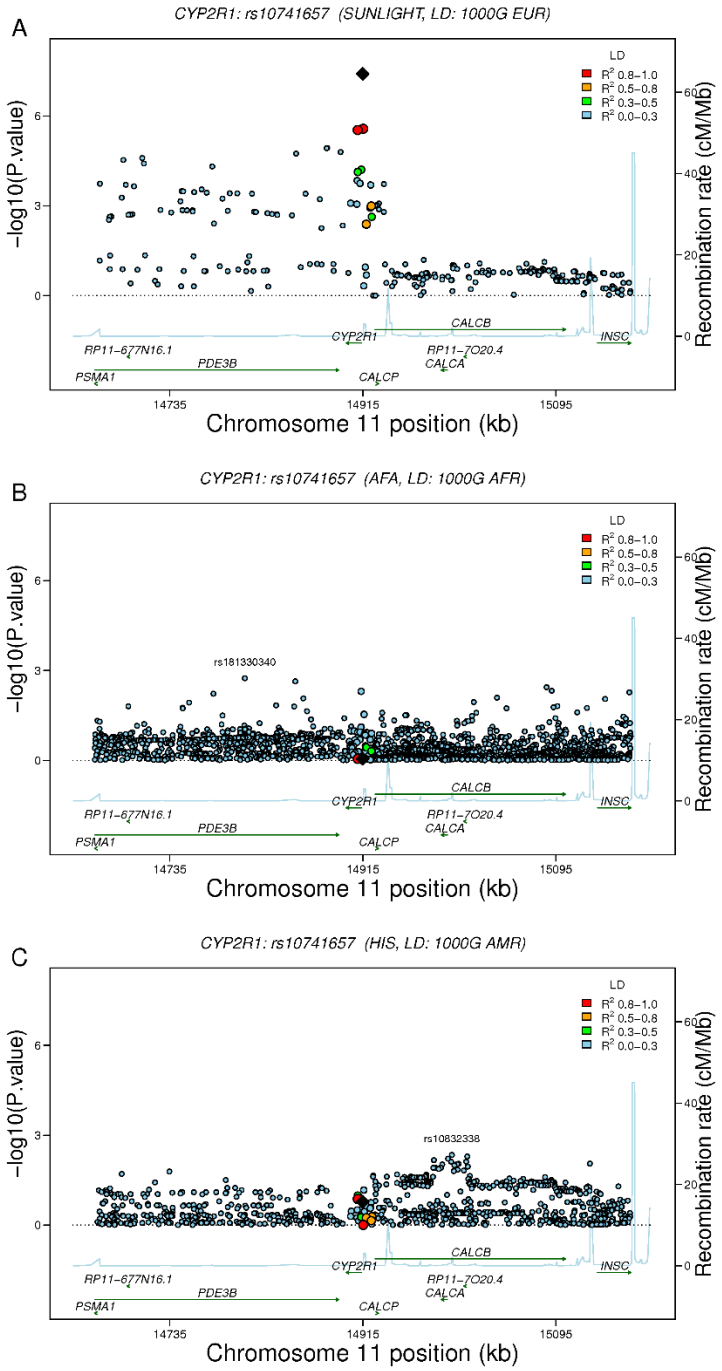
Supplemental Figure 1. Regional association plots for *KIF4B* SNPs on chromosome 4 for European (EUR), African (AFR) and Hispanic (AMR) ancestries. Log 10 p-values of genotyped SNPs are plotted against their position in the genome (build 37). The top SNP from the trans-ethnic analysis is shown with the black diamond, and shading of the circles corresponds to the LD (measured with r^2 using relevant populations from 1000G Phase I data) with this SNP. Of note, the top SNP identified by the trans-ethnic GWAS was not genotyped or imputed in SUNLIGHT (EUR).



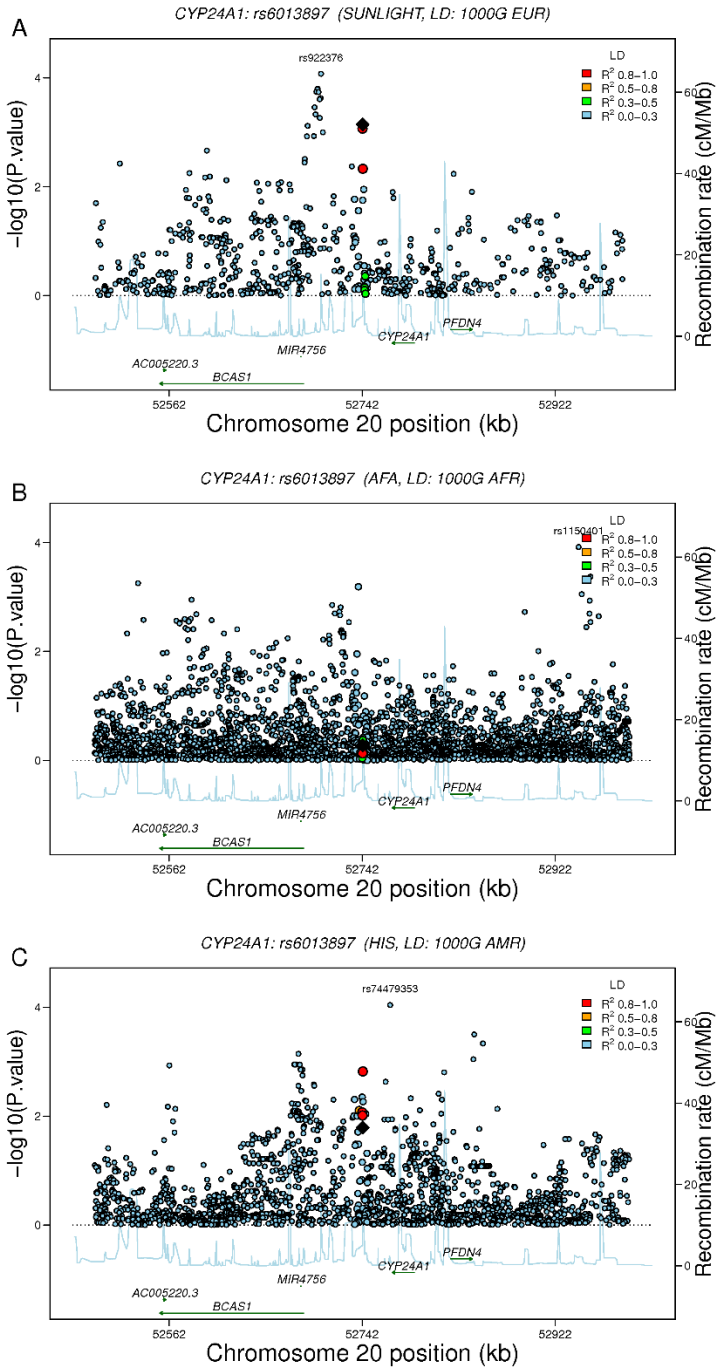
Supplemental Figure 2. Regional association plots for *ANO6/ARID2* SNPs on chromosome 4 for European (EUR), African (AFR) and Hispanic (AMR) ancestries. Log 10 p-values of genotyped SNPs are plotted against their position in the genome (build 37). The top SNP from the trans-ethnic analysis is shown with the black diamond (this SNP was not genotyped for the Hispanic ancestry so no black diamond is displayed), and shading of the circles corresponds to the LD with the top trans-ethnic SNP (measured with r^2 using relevant populations from 1000G Phase I data). The LD information is sparse in this regional association plot due to the low minor allele frequency of *ANO6/ARID2*.



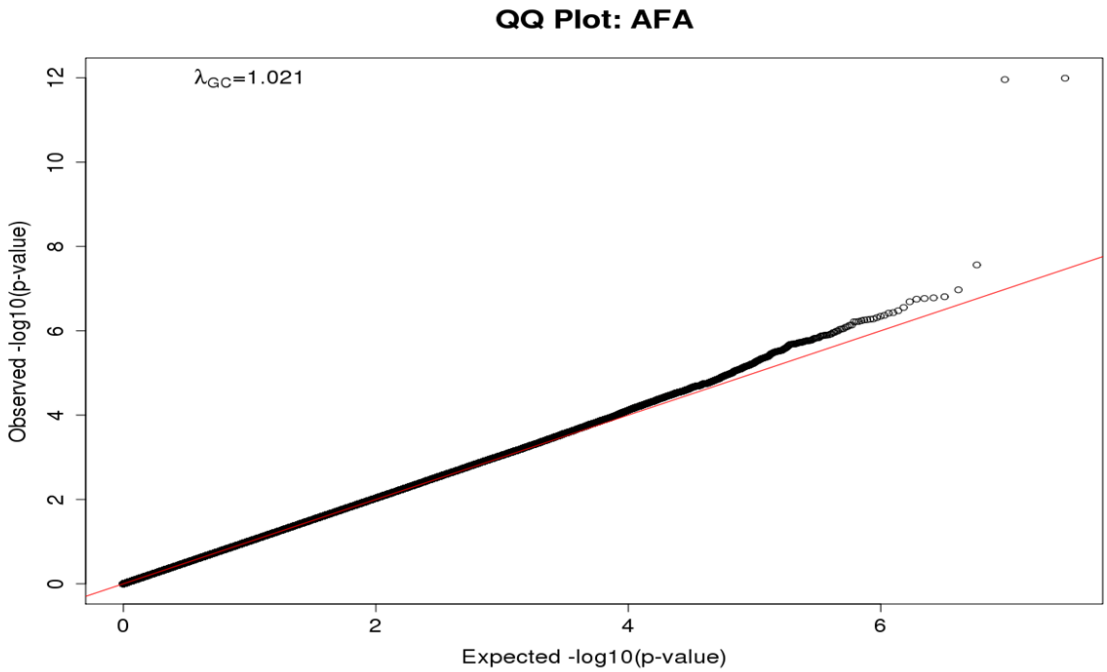
Supplemental Figure 3 . Regional association plots for *CYP2R1* SNPs on chromosome 4 for European (EUR), African (AFR) and Hispanic (AMR) ancestries. Log 10 p-values of genotyped SNPs are plotted against their position in the genome (build 37). The top SNP from the trans-ethnic analysis is shown with the black diamond, and shading of the circles corresponds to the LD (measured with r^2 using relevant populations from 1000G Phase I data) with this SNP.



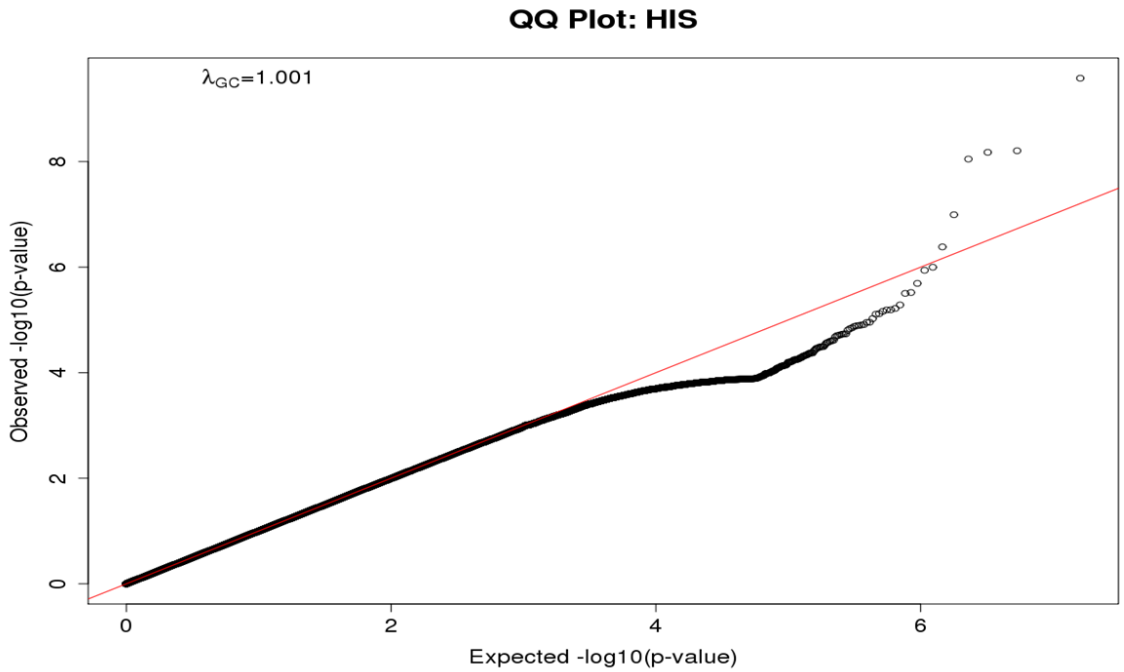
Supplemental Figure 4. Regional association plots for *CYP24A1* SNPs on chromosome 4 for European (EUR), African (AFR) and Hispanic (AMR) ancestries. Log 10 p-values of genotyped SNPs are plotted against their position in the genome (build 37). The top SNP from the trans-ethnic analysis is shown with the black diamond, and shading of the circles corresponds to the LD (measured with r^2 using relevant populations from 1000G Phase I data) with this SNP.



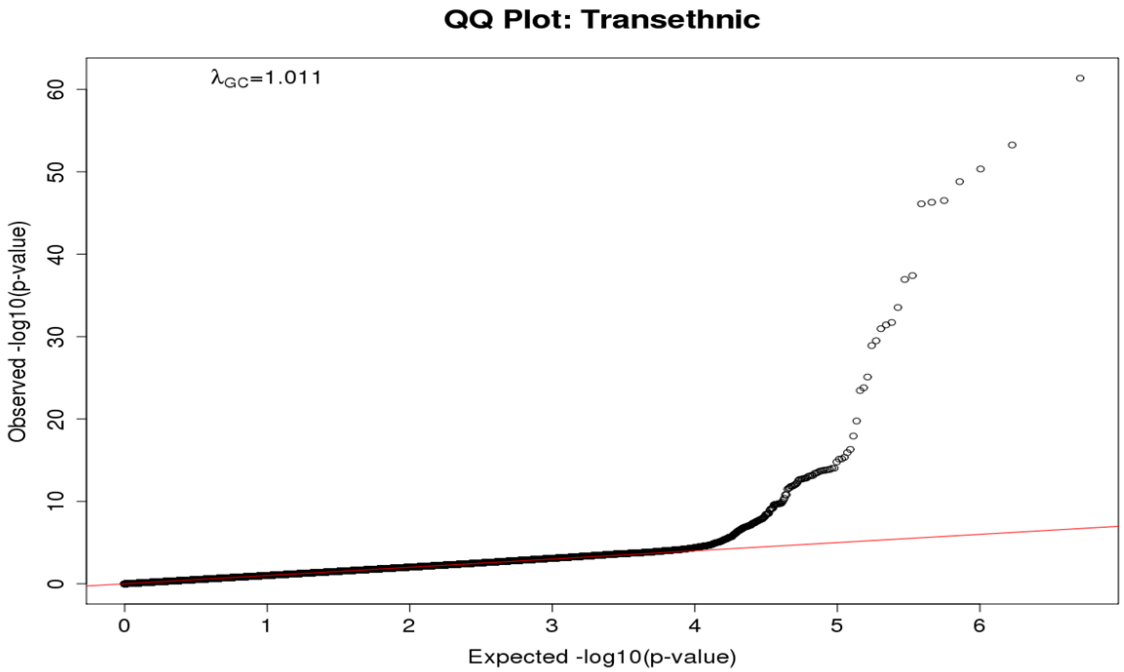
Supplemental Figure 5. Quantile-Quantile plot for the African-ancestry cohorts' analyses is shown. Each black dot represents an observed statistic [$\log_{10}(P)$] (y-axis) versus the corresponding expected statistic (x-axis). The red line denotes the null distribution, which assumes no association. Deviation is only observed in the tail for each of the plots (lamba<1.021 in all analyses), suggesting that population stratification was adequately controlled.



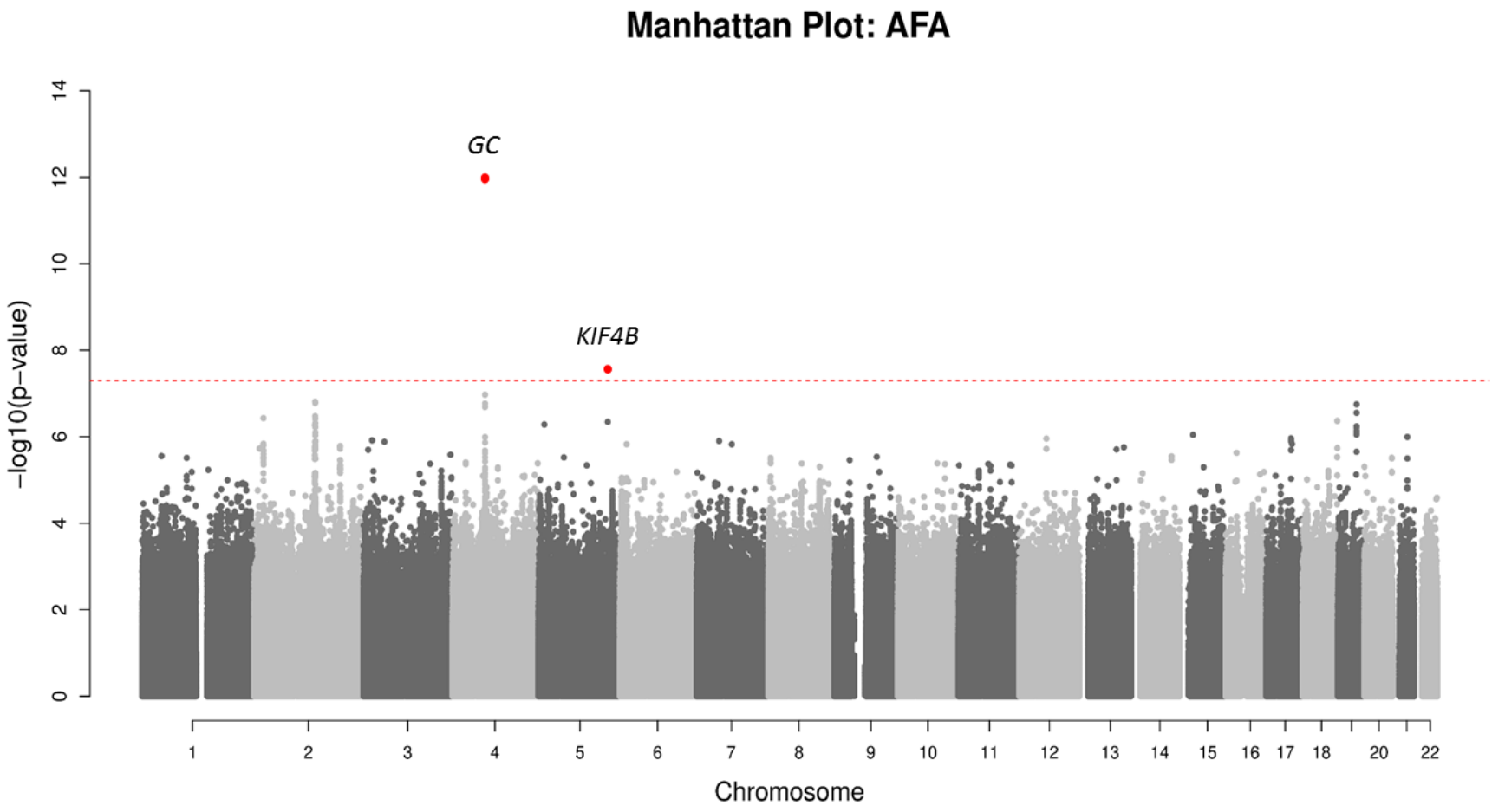
Supplemental Figure 6. Quantile-Quantile plot for the Hispanic-ancestry cohorts is shown. Each black dot represents an observed statistic [$\log_{10}(P)$] (y-axis) versus the corresponding expected statistic (x-axis). The red line denotes the null distribution, which assumes no association. Deviation is only observed in the tail for each of the plots (lamba<1.021 in all analyses), suggesting that population stratification was adequately controlled.



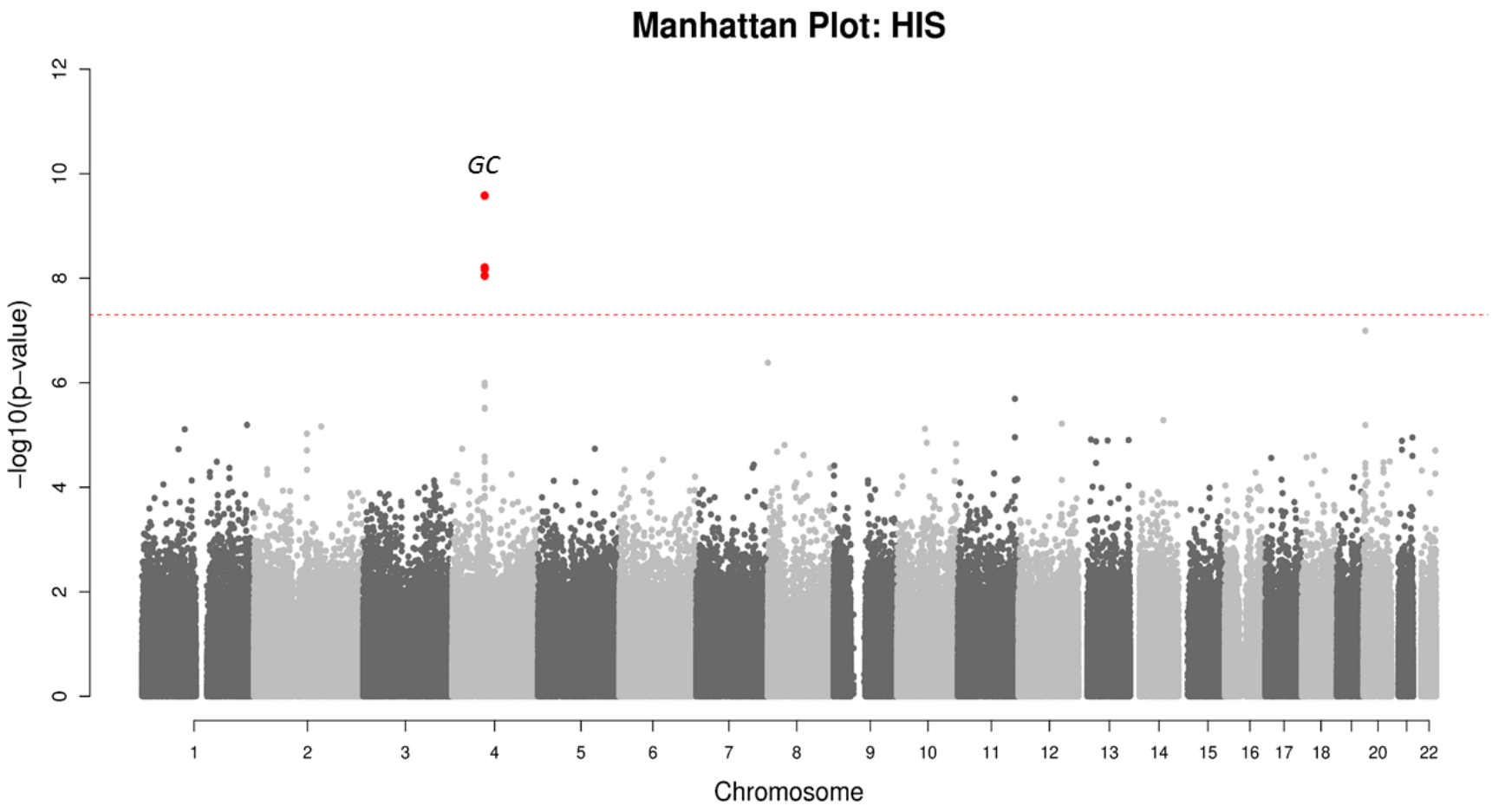
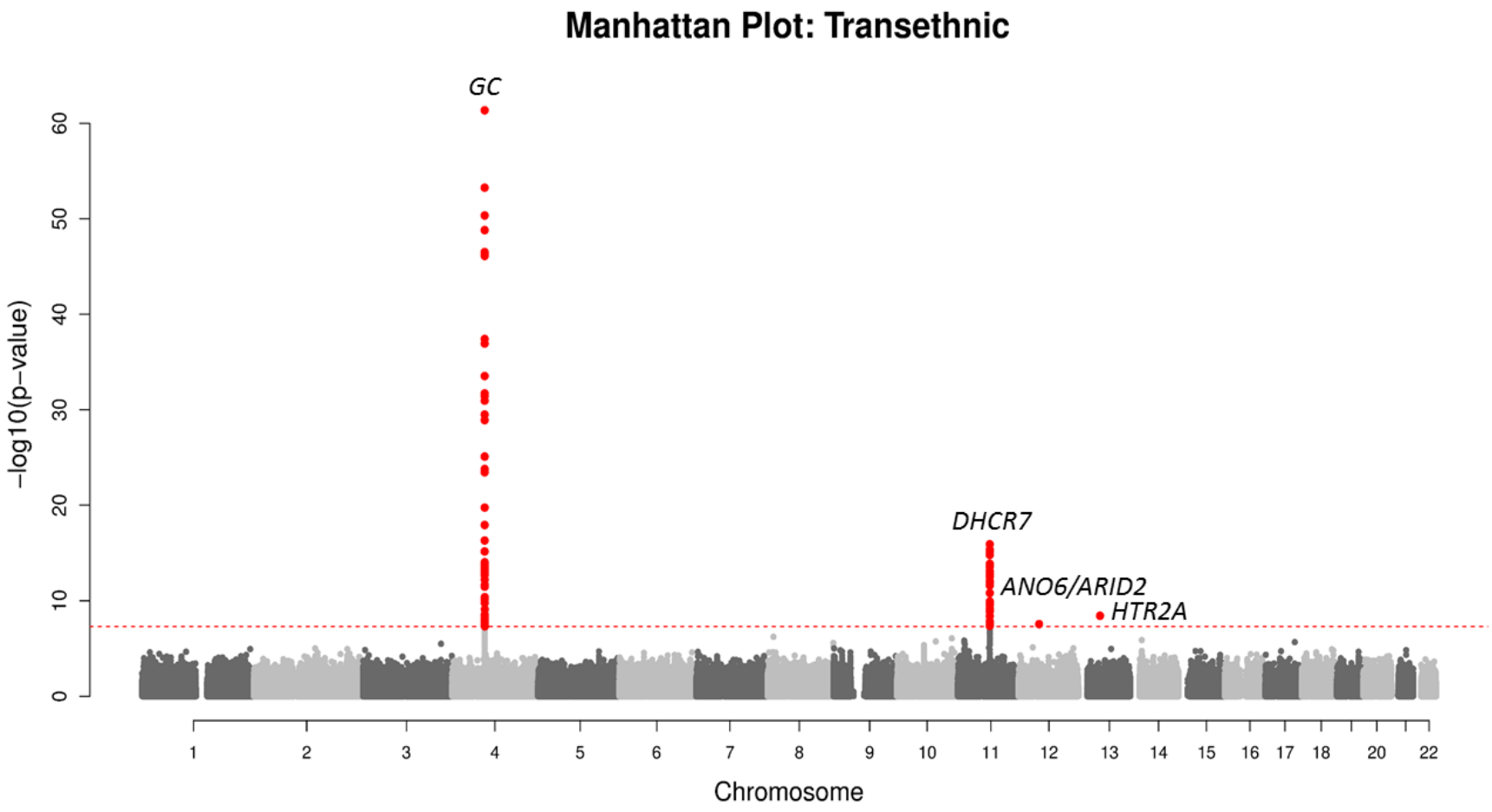
Supplemental Figure 7. Quantile-Quantile plot for the trans-ethnic analysis is shown. Each black dot represents an observed statistic [$\log_{10}(P)$] (y-axis) versus the corresponding expected statistic (x-axis). The red line denotes the null distribution, which assumes no association. Deviation is only observed in the tail for each of the plots (lamba<1.021 in all analyses), suggesting that population stratification was adequately controlled.



Supplemental Figure 8. Genome-wide associations of circulating 25(OH)D are graphed by chromosome position and −log10 *P*-value for the African analysis. Each grey or black dot represents a SNP; SNPS above the red dashed line achieved genome-wide significance.



Supplemental Figure 9. Genome-wide associations of circulating 25(OH)D are graphed by chromosome position and −log10 *P*-value for the Hispanic analysis. Each grey or black dot represents a SNP; SNPS above the red dashed line achieved genome-wide significance.



Supplemental Figure 10. Genome-wide associations of circulating 25(OH)D are graphed by chromosome position and −log10 *P*-value for the trans-ethnic analysis. Each grey or black dot represents a SNP; SNPS above the red dashed line achieved genome-wide significance.