Supplementary Table 1 Description of the 1000 Genomes Project Phase 3 representing 2504 individuals from 26 different global populations that are assigned to five super-populations

Number of individuals	Population code	Population description	Super-population	Super-population code
103	CHB	Han Chinese in Beijing, China	East Asian	EAS
104	JPT	Japanese in Tokyo, Japan	East Asian	EAS
105	CHS	Southern Han Chinese	East Asian	EAS
93	CDX	Dai in Xishuangbanna, China	East Asian	EAS
99	KHV	Kinh in Ho Chi Minh City, Vietnam	East Asian	EAS
99	CEU	Utah Residents (CEPH)	European	EUR
107	TSI	Toscani in Italia	European	EUR
99	FIN	Finnish in Finland	European	EUR
91	GBR	British in England and Scotland	European	EUR
107	IBS	Iberian population in Spain	European	EUR
108	YRI	Yoruba in Ibadan, Nigera	African	AFR
99	LWK	Luhya in Webuye, Kenya	African	AFR
113	GWD	Gambians from The Gambia	African	AFR
85	MSL	Mende in Sierra Leone	African	AFR
99	ESN	Esan in Nigeria	African	AFR
61	ASW	African Americans in SW USA	African	AFR
96	ACB	African Caribbeans in Barbados	African	AFR
64	MXL	Mexican Ancestry from Los Angeles	Admixed American	AMR
104	PUR	Puerto Ricans from Puerto Rico	Admixed American	AMR
94	CLM	Colombians from Medellin, Colombia	Admixed American	AMR
85	PEL	Peruvians from Lima, Peru	Admixed American	AMR
103	GIH	Gujarati Indian from Texas	South Asian	SAS
96	PJL	Punjabi from Lahore, Pakistan	South Asian	SAS
86	BEB	Bengali from Bangladesh	South Asian	SAS
102	STU	Sri Lankan Tamil from the UK	South Asian	SAS
102	ITU	Indian Telugu from the UK	South Asian	SAS

Data generated from: ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/release/20130502/
http://www.1000genomes.org/category/frequently-asked-questions/population

Supplementary Table 2: *please see separate supplementary attachment for these data and related key that provides column heading descriptions.*

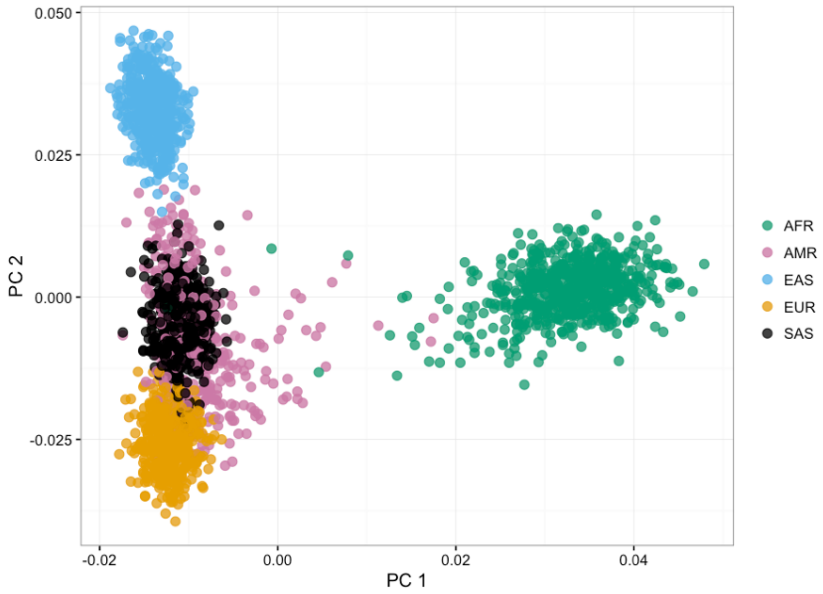
Supplementary Table 3: *please see separate supplementary attachment for these data and related key that provides column heading descriptions.*

Supplementary Table 4. Functional annotations of highly differentiated pharmacogenomic variants in the 17 pharmacogenes that contained variants were common in one population (>5%), but rare (0.05%) in the global cohort.

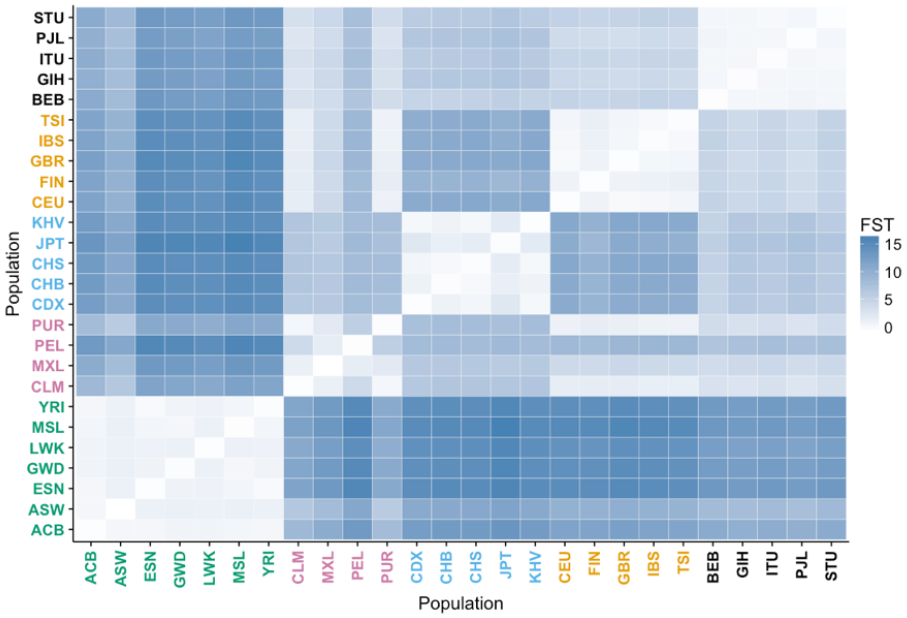
ID	Population	Gene	Annotation	CADD
rs35865660	GWD	*ACE*	Missense	0.14
rs185351296	PEL	*ACE*	Intron	8.93
rs75067113	**ESN**	***ACE***	**Synonymous**	**16.31**
rs4979	LWK	*ACE*	Synonymous	14.05
rs139471271	LWK	*ACYP2*	Intron	1.05
rs376175333	GIH	*ALK*	Missense	21.50
rs6003592	ESN	*BCR*	Intron	6.64
rs4986854	JPT	*BRCA1*	Missense	0.023
rs113857788	**KHV**	***CFTR***	**Missense**	**26.50**
rs116593005	LWK	*CRHR1*	Intron	14.21
rs28399501	FIN	*CYP2B6*	3'UTR	3.47
rs181297724	**FIN**	***CYP2C19***	**Missense and splice region**	**22.90**
rs11572079	FIN	*CYP2C8*	Splice region	0.13
rs141759372	**ESN**	***GSTT1***	**Missense**	**27.00**
rs142766358	LWK	*HAS3*	Missense	14.10
rs17181024	ESN	*OPRM1*	Intron	6.20
rs200451188	ESN	*RYR1*	Synonymous	11.70
rs200023171	GIH	*RYR1*	Splice region	2.45
rs17215493	GWD	*SCN5A*	Synonymous	2.90
rs140504750	ASW	*SULT1A1*	Intron	3.97
rs186462665	FIN	*TXNRD2*	Synonymous	10.98

Potentially deleterious variants (CADD scores ≥15) are displayed in bold.
Independent signals (r2 < 0.4) with the largest allele frequency differences between the remaining populations for each gene region are reported.

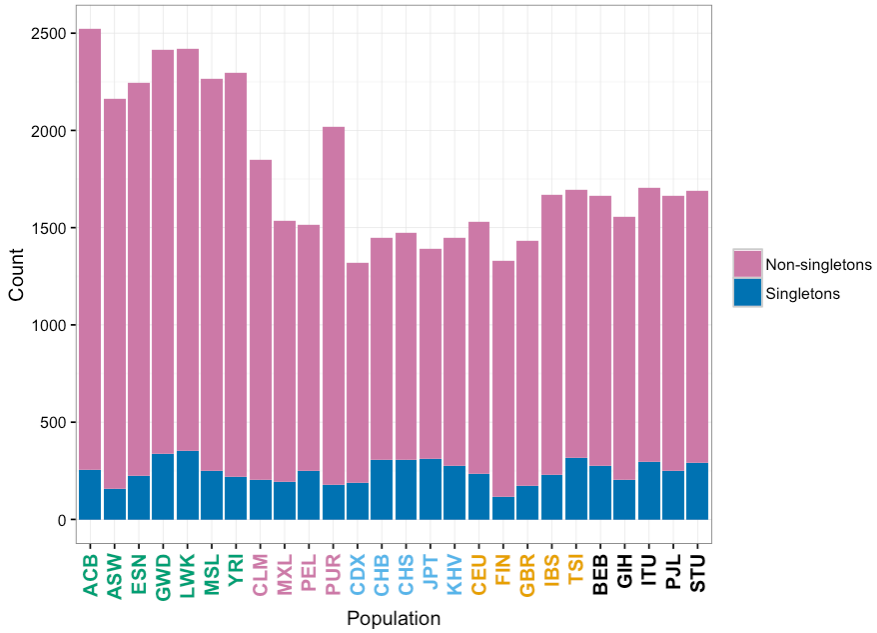
Supplementary Table 5: *please see separate supplementary attachment for these data and related key that provides column heading descriptions.*



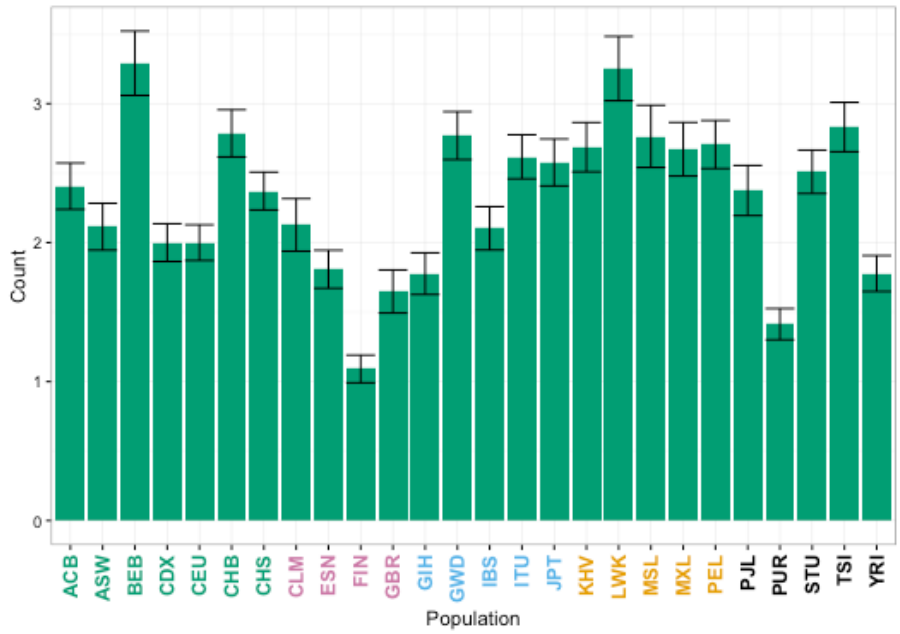
Supplementary Figure 1 Principal component analysis (PCA) of pharmacogenomic variants stratified individuals into continental super populations, with the admixed American individuals separating along clines between the main ancestral clusters. Principal component 1 separates the African individuals from the rest of the continental populations, while principal component 2 distinguishes European, South and East Asian population groups.



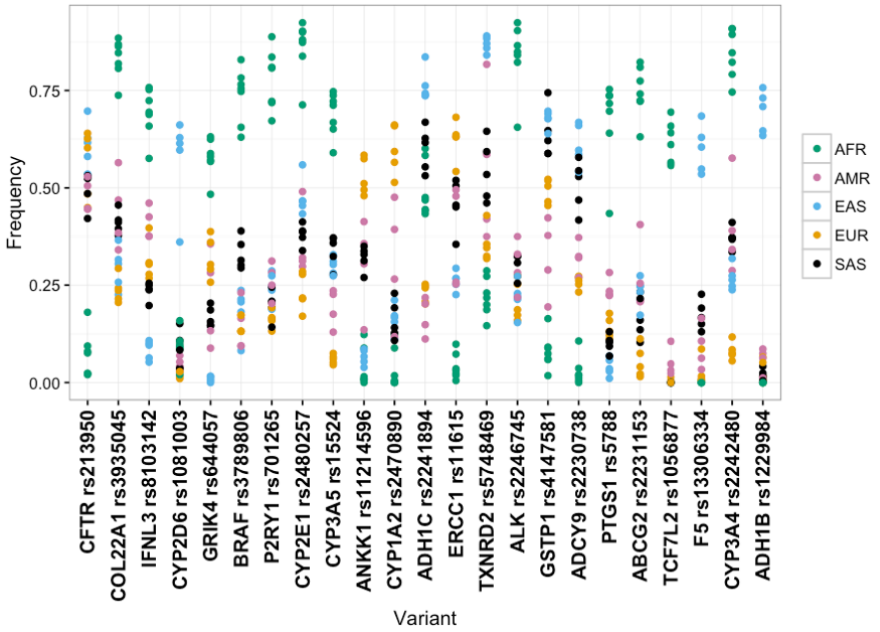
Supplementary Figure 2 Weir and Cockerham's fixation index (F_{ST}) statistics x 100 for synonymous pharmacogenomic variants in the 1000 Genomes Project individuals. Populations tended to be the most similar on the continental level (indicated by lighter shades of blue) and the admixed American populations showing differing degrees of differentiation from the European populations.



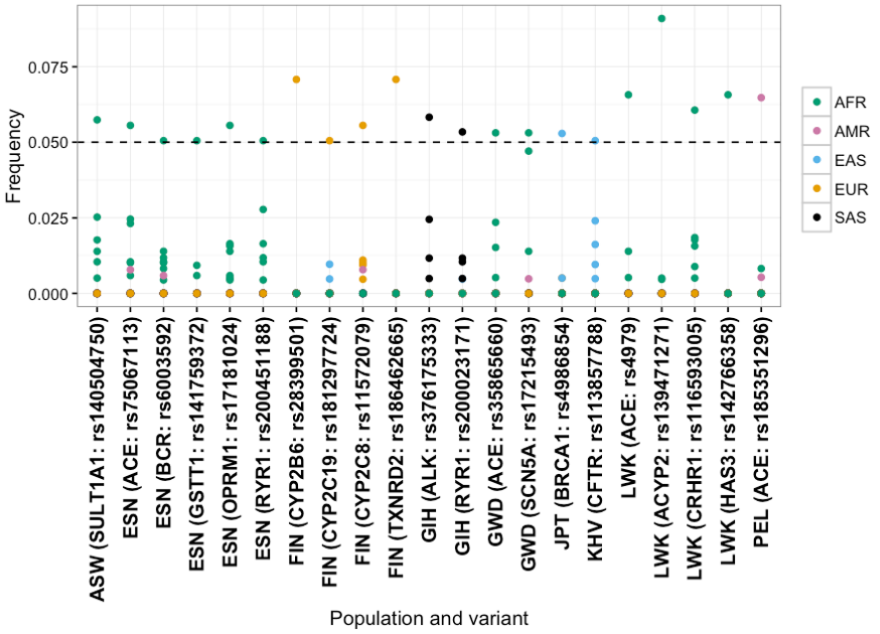
Supplementary Figure 3 The number of polymorphic sites in each of the individual 1000 Genomes Project individual populations stratified by singleton sites. African populations had the highest number of polymorphic sites in their pharmacogenes.



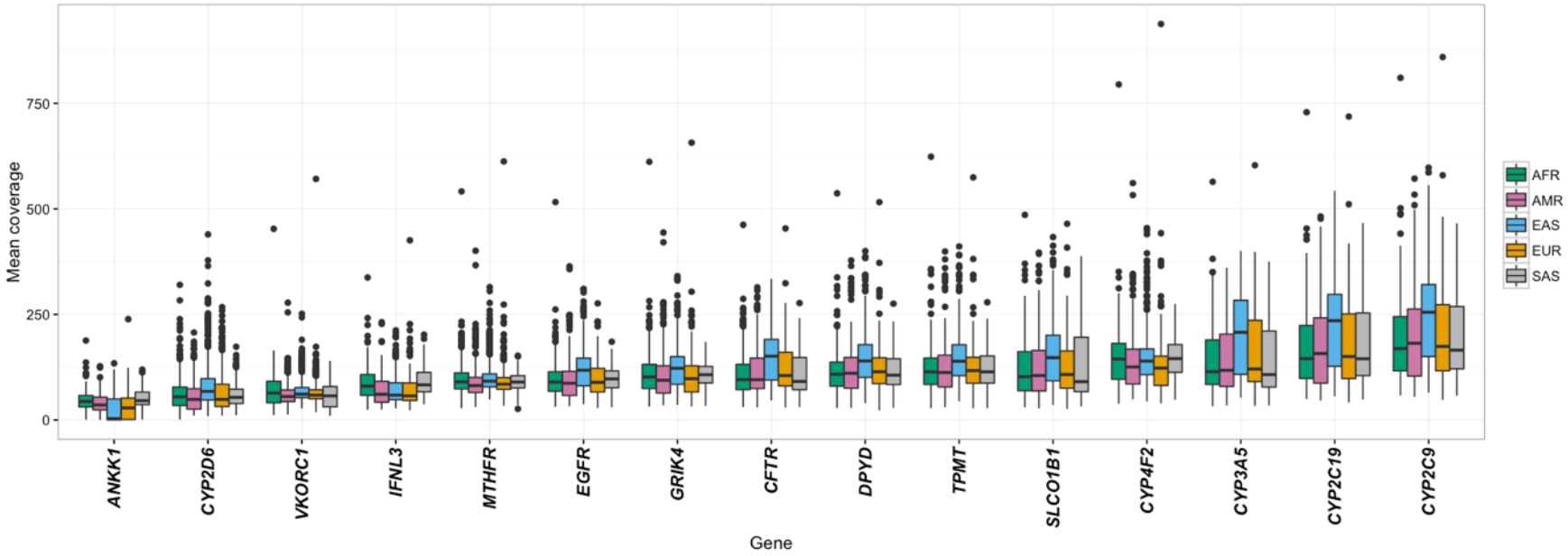
Supplementary Figure 4 The mean number of pharmacogenomic singletons per individual in each of the 1000 Genomes Populations. Individuals carried between 1 and 3 unique pharmacogenomic variants on average and this class of variation made up 52.9% of the dataset.



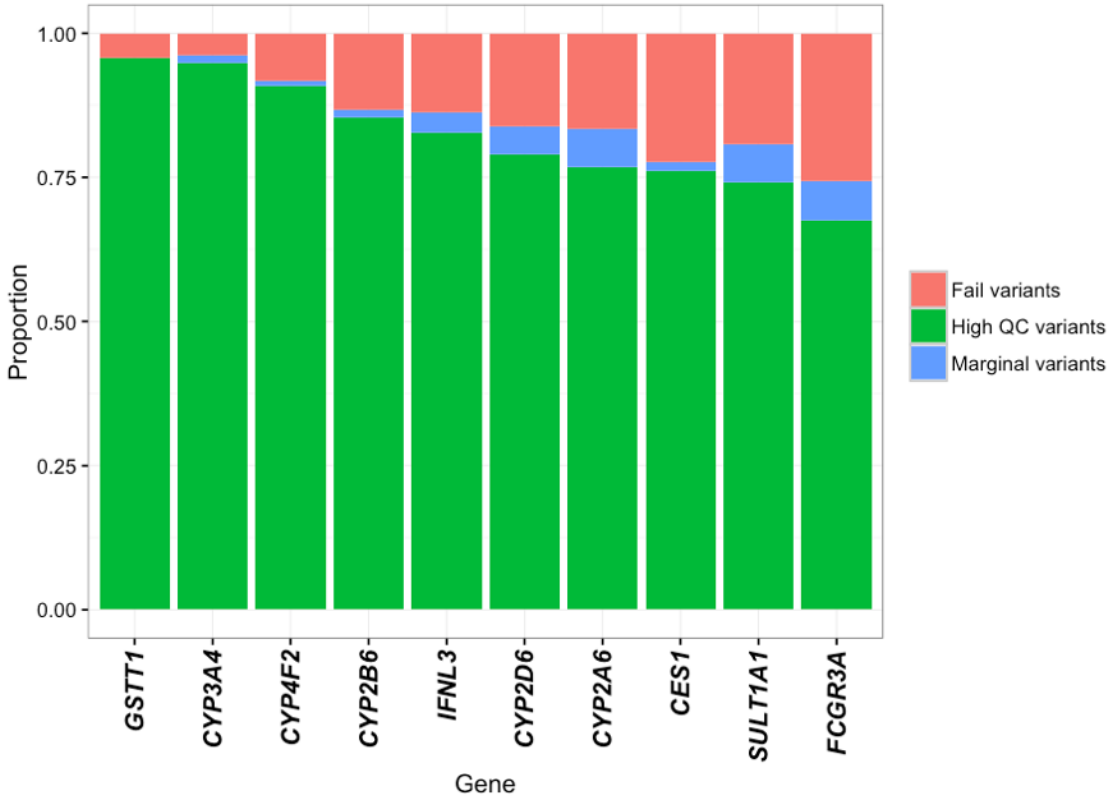
Supplementary Figure 5 Allele frequencies of pharmacogenomic variants in the 23 pharmacogenes that displayed at least one highly differentiated signal (pairwise $F_{ST} > 0.5$ for one or more super-population comparison). The highest mean F_{ST} variant for independent signal ($r^2 < 0.4$) each gene region is reported.



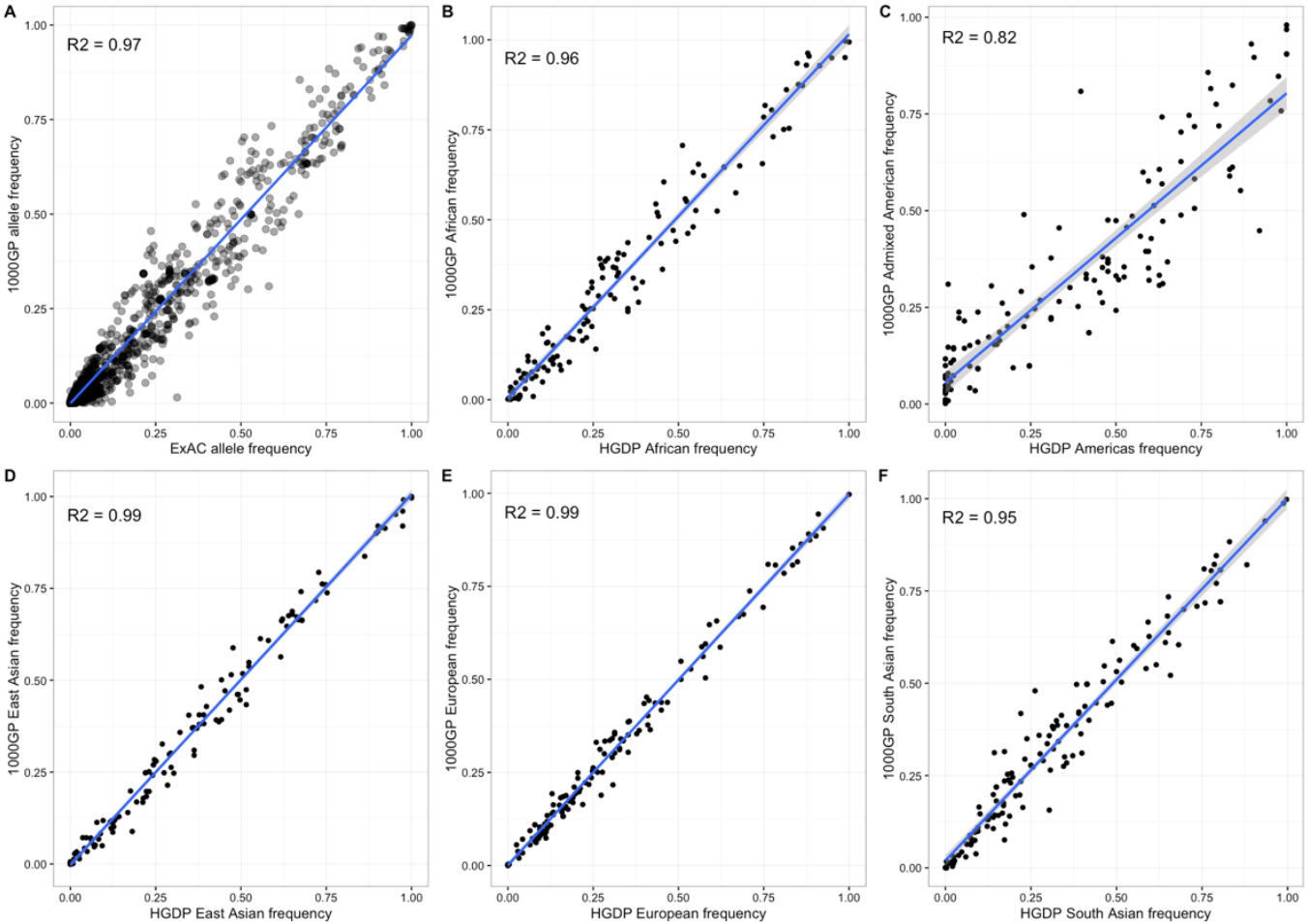
Supplementary Figure 6 Minor allele frequencies of highly differentiated pharmacogenomic variants in the 17 pharmacogenes that contained variants were common in one population (>5%), but rare (0.05%) in the global cohort. The dashed line represents the 5% minor allele frequency. Independent signals ($r^2 < 0.4$) with the largest allele frequency differences between the remaining populations for each gene region are reported.



Supplementary Figure 7 Mean sequencing coverage for each of the 1000 Genomes Project individuals for pharmacogenes that are known to carry high clinical evidence variants (i.e. PharmGKB 1A/B) separated by super-population.



Supplementary Figure 8 Proportion of different quality variants in the ten pharmacogenes that were flagged as potentially problematic for short read sequencing data. High QC variants had SVM scores >0.3, while marginal variants had scores between 0 and 0.3. Fail variants were not included in the final release of the 1000 Genomes Project and had SVM scores <0.



Supplementary Figure 9 A strong correlation in allele frequencies were observed between overlapping pharmacogenomic variants in the 1000 Genomes Project and those generated by external projects: **A** Comparison between global allele frequency in the 1000 Genomes Project and sequencing data generated through the Exome Aggregation Consortium (10871 variants) **B-F** Comparison between super-population allele frequencies of the 1000 Genomes Project and the Human Genome Diversity Project (136 array-genotyped variants).