Report 11-12: Addressing Questions with Analysis

The questions guiding this overall research are:

Q1: Are there genes involved in human phenotypes/diseases harboring SNPs which are population-specific?

Q2: Do the handful of known human examples (e.g., sickle-cell in African Americans, EPAS1 in Tibetans) validate our results? (i.e., Are our results consistent with current scientific literature on monogenic / single SNP driven phenotypes?

Q3: Do diseases associated with population-specific SNPs show significant disparities between populations?

At this point the data is sufficiently structured to allow an attempt to create insight into these questions. Several perspectives by visualization have been used in order to help build intuition about the data, and I will hereby present the analyses and give my interpretation of the meaning behind the patterns we see. Q1 is the first major question to approach generally, and for that we have a visualization of all variants classified as diseases variants (according to the Experimental Factor Ontology [EFO] imposed by the GWAS catalog) organized by subpopulation and colored by super population:

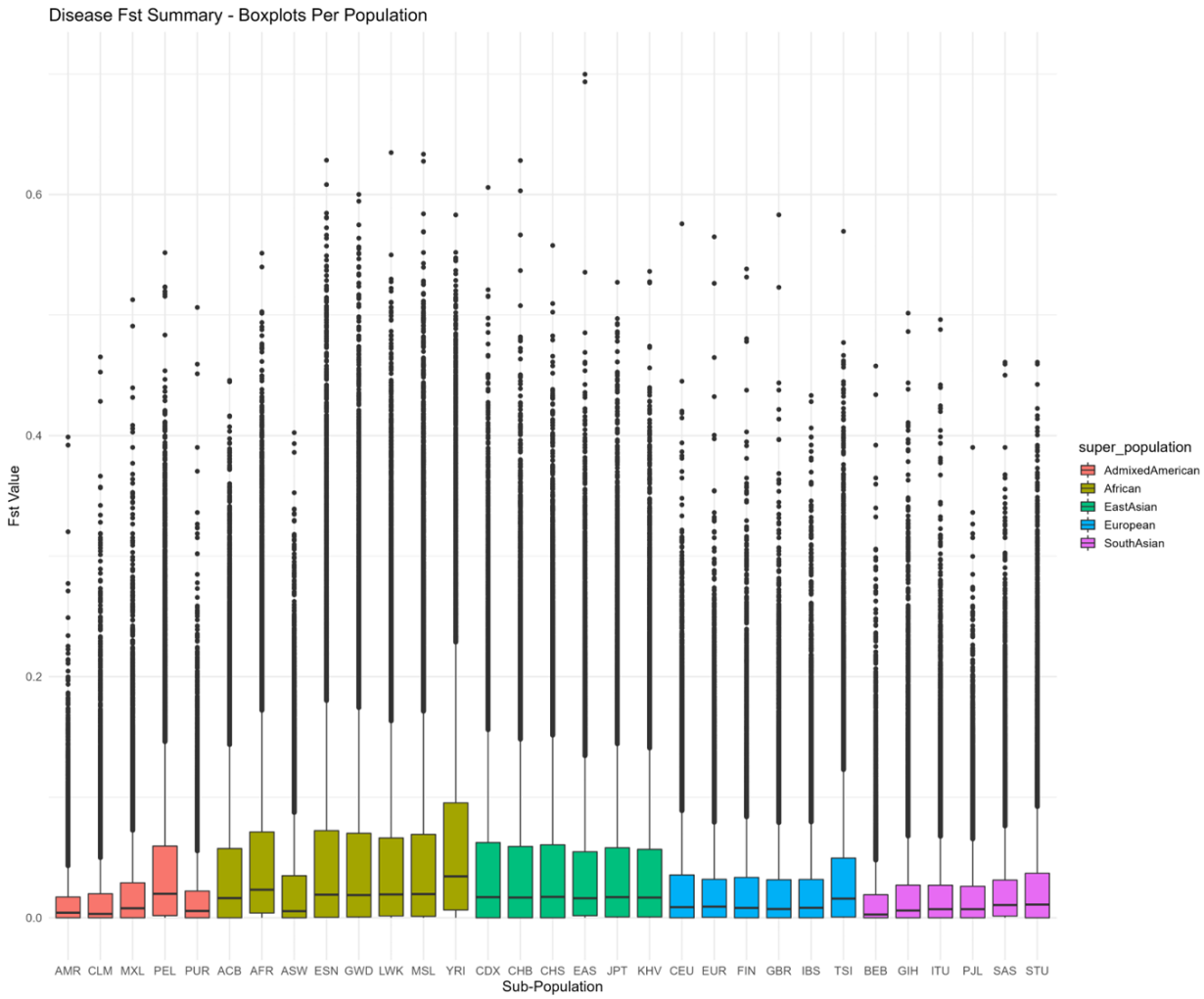


Fig - 1: Each population has some number of variants peaking above 0.4 Fst against the metapopulation. Those of East Asian and African ancestry evidently contain more high Fst variants than the other super-populations. The data otherwise looks as expected, with means quite close to 0 regardless of the population.

Looking more specifically into the data, we can see just how many variants are above 0.4 Fst for each subpopulation:

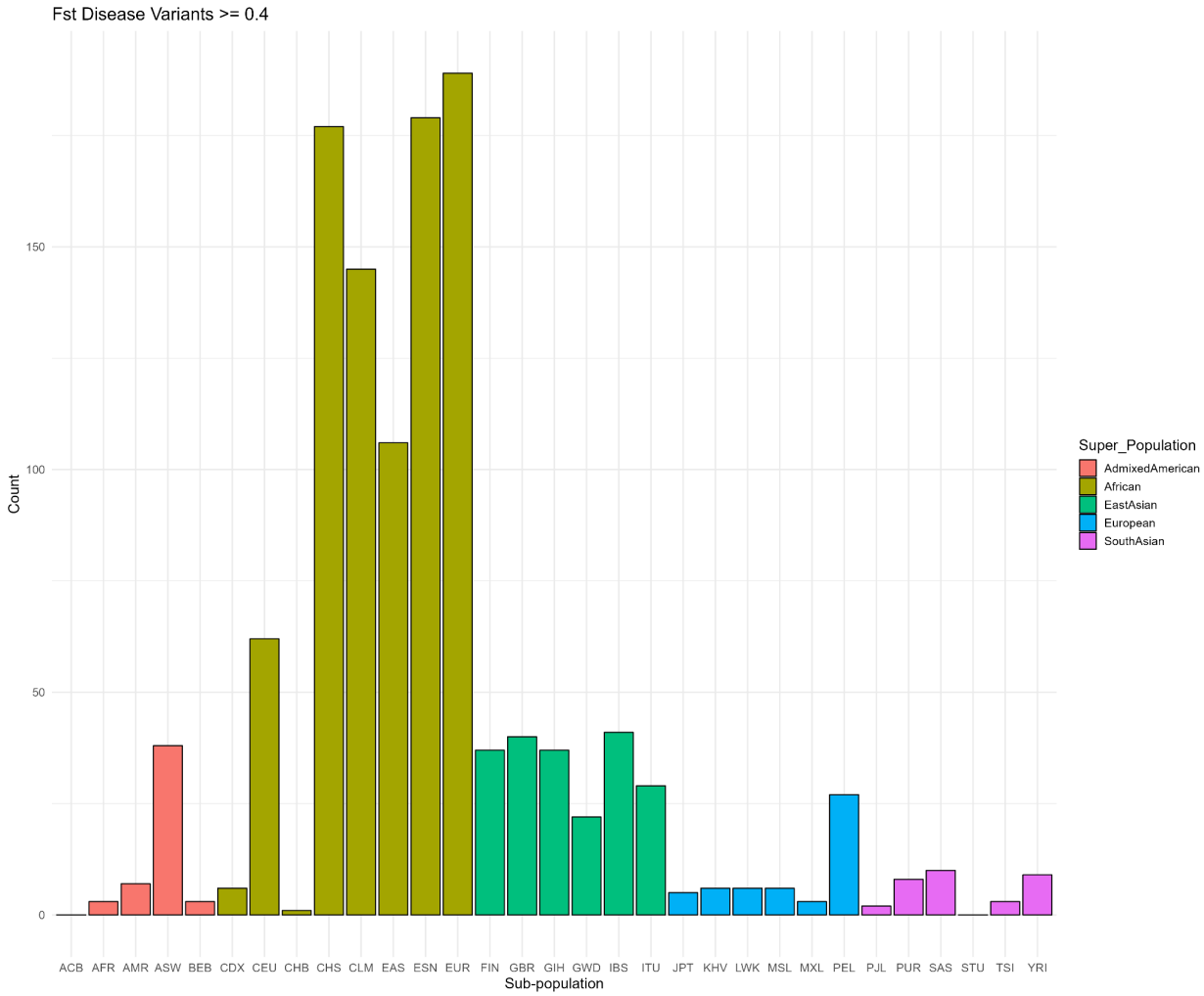


Fig 2 - African sub-populations posses the majority of high Fst variants beyond this threshold. This trend of differentiation according to super-population is one which appears consistently throughout our data. For more graphs of this form with different threshold Fst values (0.2-0.6) see the accompanying power point.

It is clear that each subpopulation posses a handful of high Fst variants, and it is thus interesting to ask: What are those variants, and what diseases and genes are they associated with? In order to answer these questions I assembled a table which captures the top 20 variants for each sub population, as well as the genes they're mapped to, their disease/trait as reported by the studies from which these variants were identified, and further what kinds of mutations they are. The table is 428 rows long, which makes for poor viewing, though a small example is shown below:

	VariantID	MAPPED_GENE	EnsVar_most_severe_consequence	DISEASE/TRAIT	MAPPED_TRAIT	Parent term	population_labels
1	rs10032909	LINC01088	intron_variant	Systemic lupus erythematosus	obsolete_systemic lupus erythematosus	Immune system disorder	EUR
2	rs10035291	SSBP2	intron_variant	Bipolar disorder	obsolete_bipolar disorder	Neurological disorder	LWK
4	rs1012621	TGM3	intron_variant	Inflammatory skin disease	atopic eczema, psoriasis	Other disease	ASW
5	rs1012621	TGM3	intron_variant	Inflammatory skin disease	atopic eczema, psoriasis	Immune system disorder	ASW
6	rs1016189	MAGI2	intron_variant	Acute graft-versus-host disease (gut) (recipient effect)	acute graft vs. host disease	Immune system disorder	FIN
7	rs10245867	JAZF1	intron_variant	Multiple sclerosis	multiple sclerosis	Immune system disorder	CHS
8	rs10245867	JAZF1	intron_variant	Type 1 diabetes	type 1 diabetes mellitus	Digestive system disorder	CHS
9	rs10245867	JAZF1	intron_variant	Systemic lupus erythematosus	obsolete_systemic lupus erythematosus	Immune system disorder	CHS
10	rs10245867	JAZF1	intron_variant	Type 1 diabetes	type 1 diabetes mellitus	Immune system disorder	CHS
13	rs103294	MIR4752 - LILRA5	3_prime_UTR_variant	Prostate cancer	prostate carcinoma	Cancer	CHB
14	rs103294	MIR4752 - LILRA5	3_prime_UTR_variant	Takayasu arteritis	Takayasu arteritis	Cardiovascular disease	CHB
17	rs10421769	GPATCH1	stop_gained	Thrombosis in response to liver transplant	coronary thrombosis, GM11992	Cardiovascular disease	CEU|EUR

Fig 3 - the complete table can be found as an attached csv file. (top_20_variants_per_population.csv)

If you bring your attention to the final row, and final column the population label "CEU|EUR" can be seen, which comes from the fact that there is overlap between two populations with respect to this variant. That is to say within the top 20 variants of the CEU (Utah residents) sub-population there is a shared variant with the EUR sub-population (European ancestry). Many variants within this table are unique to a single population, but some are found within several populations. To quantify this degree of overlap, with 31 subpopulations and 20 variants per population there were a total of 620 variants brought up. Of that 620, there are only 300 unique variant IDs. Of those 300 unique variant IDs 176 belong uniquely to a sub-population. These one-to-one variants of high Fst are all found within the complete table, and are also listed separately in their own csv (top_20_variants_per_population.csv). Further there are 196 unique disease/trait terms contained within the table as well, which doesn't point to 196 unique disease conditions necessarily, as many of these terms can be categorized as sub-traits associated with some major disease term.

As a complement to this detailed table capturing the most interesting details of each top 20 Fst variant I have generated 5 associated figures displaying the variants associated with each sub-population and their Fst, each figure is organized by super-population. The remaining 4 figures can be found at the bottom of this document as well as within the associated powerpoint.

ACB		AFR		ASW		ESN		GWD		LWK		MSL		YRI	
rs269863	0.4458	rs269863	0.5513	rs269863	0.4024	rs60276348	0.6284	rs11673591	0.6001	rs269863	0.6347	rs269863	0.6334	rs2226738	0.583
rs2226738	0.4446	rs2226738	0.5398	rs5757922	0.3933	rs2226738	0.6082	rs269863	0.5944	rs11673591	0.5499	rs2226738	0.6276	rs60276348	0.5519
rs4245229	0.4163	rs8068952	0.5027	rs2226738	0.3861	rs6447129	0.5845	rs11814448	0.5748	rs8068952	0.5298	rs8068952	0.5839	rs269863	0.5474
rs11680058	0.4158	rs57676627	0.5023	rs6014499	0.3526	rs1129038	0.581	rs8068952	0.5638	rs2226738	0.5276	rs7185636	0.5692	rs6447129	0.5465
rs12425451	0.4073	rs60276348	0.5014	rs6547598	0.3391	rs12913832	0.5805	rs7951870	0.5565	rs4643526	0.5222	rs4643526	0.5688	rs8068952	0.5453
rs6447129	0.4035	rs6447129	0.4938	rs6427419	0.3387	rs8068952	0.5724	rs4643526	0.5551	rs8038734	0.5198	rs4245229	0.5518	rs12425451	0.5449
rs4403550	0.3974	rs4643526	0.4901	rs56401710	0.3354	rs1716183	0.5663	rs2226738	0.5514	rs2965030	0.5107	rs2854437	0.5428	rs2965030	0.5369
rs12074934	0.394	rs11673591	0.4878	rs60276348	0.3349	rs4403550	0.5637	rs7185636	0.5507	rs9623117	0.5105	rs8072449	0.5396	rs116945310	0.5342
rs60276348	0.3936	rs7951870	0.4829	rs57676627	0.3346	rs8072449	0.5572	rs4245229	0.5466	rs9607685	0.5103	rs28498223	0.5283	rs8072449	0.5287
rs57676627	0.3934	rs11038927	0.4761	rs927485	0.3297	rs269863	0.5492	rs8064088	0.5401	rs6447129	0.5065	rs9908820	0.5271	rs7185636	0.5243
rs4643526	0.3891	rs61882743	0.4748	rs11038927	0.329	rs2965030	0.5455	rs35324223	0.5394	rs35085068	0.506	rs8046545	0.5256	rs7125179	0.5202
rs11673591	0.3874	rs35324223	0.4732	rs61882743	0.318	rs7951870	0.5407	rs6447129	0.5342	rs10035291	0.5033	rs12425451	0.5252	rs1716183	0.5172
rs12075	0.3859	rs8072449	0.4725	rs2965030	0.3172	rs562638	0.5401	rs11038927	0.5332	rs7951870	0.5018	rs60276348	0.518	rs802571	0.5146
rs8038734	0.3856	rs7185636	0.4721	rs10948071	0.3167	rs11038927	0.5371	rs61882743	0.533	rs2413631	0.4993	rs873994	0.5179	rs2525776	0.5141
rs3827760	0.3835	rs2965030	0.4719	rs6631122	0.3162	rs61882743	0.5369	rs60276348	0.5265	rs13003464	0.4962	rs2020854	0.5171	rs4643526	0.514
rs7951870	0.383	rs12425451	0.4693	rs1012621	0.3138	rs12425451	0.5327	rs57676627	0.5263	rs4565870	0.4956	rs1681087	0.5116	rs2647995	0.5132
rs58629129	0.3809	rs11814448	0.4615	rs7951870	0.3137	rs35324223	0.5287	rs11680058	0.5238	rs60276348	0.4865	rs2257883	0.5112	rs2279162	0.5128
rs8068952	0.3803	rs8038734	0.4543	rs75748221	0.313	rs2257883	0.5241	rs2257883	0.5168	rs11038927	0.4833	rs2066808	0.5097	rs11038927	0.5127
rs4908343	0.3785	rs13003464	0.4539	rs6447129	0.3097	rs9894206	0.5222	rs8072449	0.5167	rs61882743	0.4831	rs13003464	0.509	rs61882743	0.5125
rs1716183	0.3778	rs35085068	0.4496	rs2854437	0.3081	rs343093	0.5221	rs9908820	0.5161	rs35324223	0.4826	rs62956461	0.5071	rs35324223	0.512

Fig 4 - African super-population grouping of top 20 variants. Some variants are shared across several groups in this subset, such as the first two within ACB, rs269863 and rs2226738. Overlapping top variants by Fst are common amongst sub-populations that share a super-population.

Clustering With PCA:

To address Q1 broadly I think it is helpful to look at the data from a different perspective, as we can see that there are clearly many variants unique to single sub-populations with high Fst, but that knowledge isn't enough to understand the broad relationship between sub-populations. Towards this end I have generated a set of clustering plots using Principle Component Analysis (PCA). Each plot used Fst values against the meta-population for different sets and subsets of variants with respect to all sub-populations as input for PCA.

First I think it would be interesting to look at a plot created using only disease associated variants which are above 0.4 Fst:

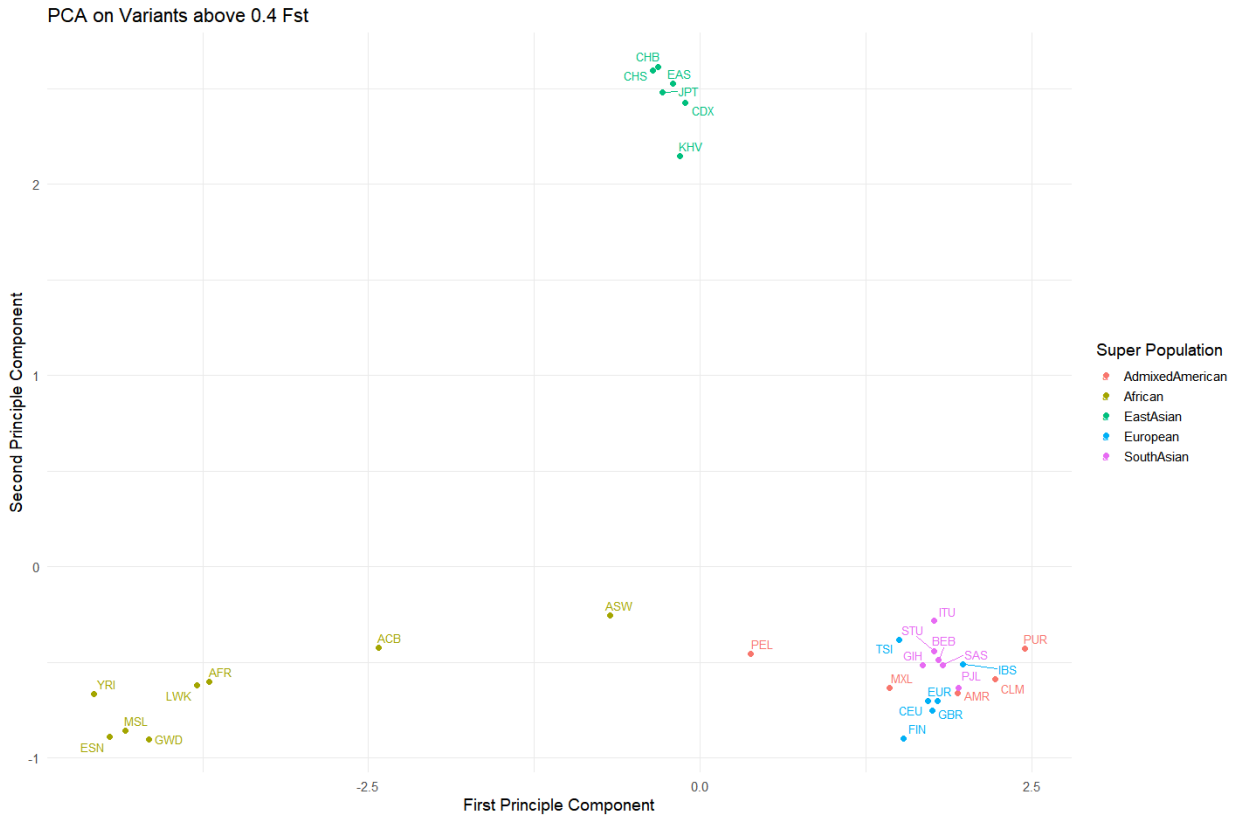
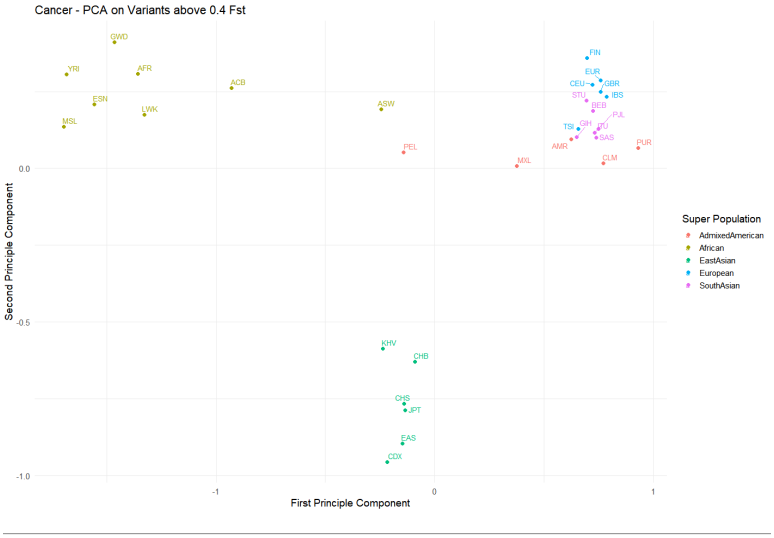
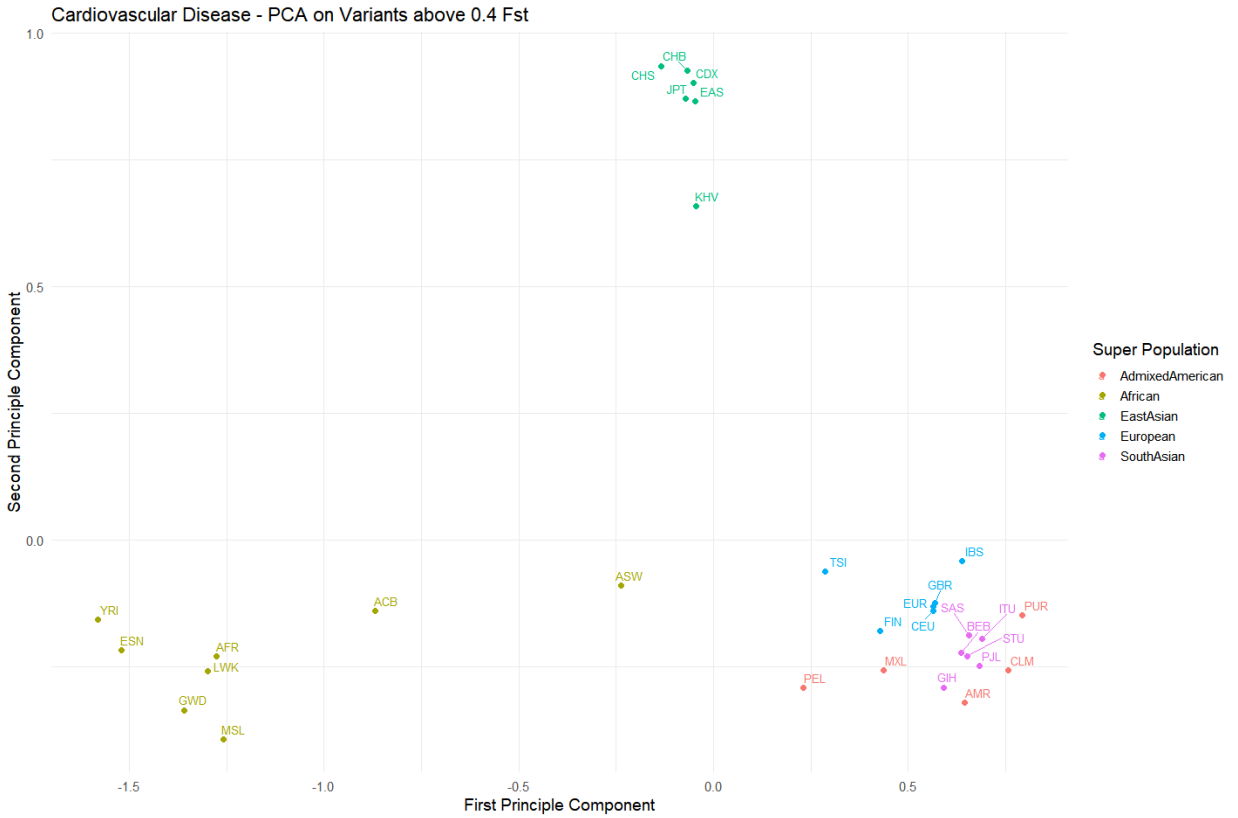


Fig 5 - PCA clustering using only disease Fst above 0.4, this includes all variants within the top 20 for each subpopulation.

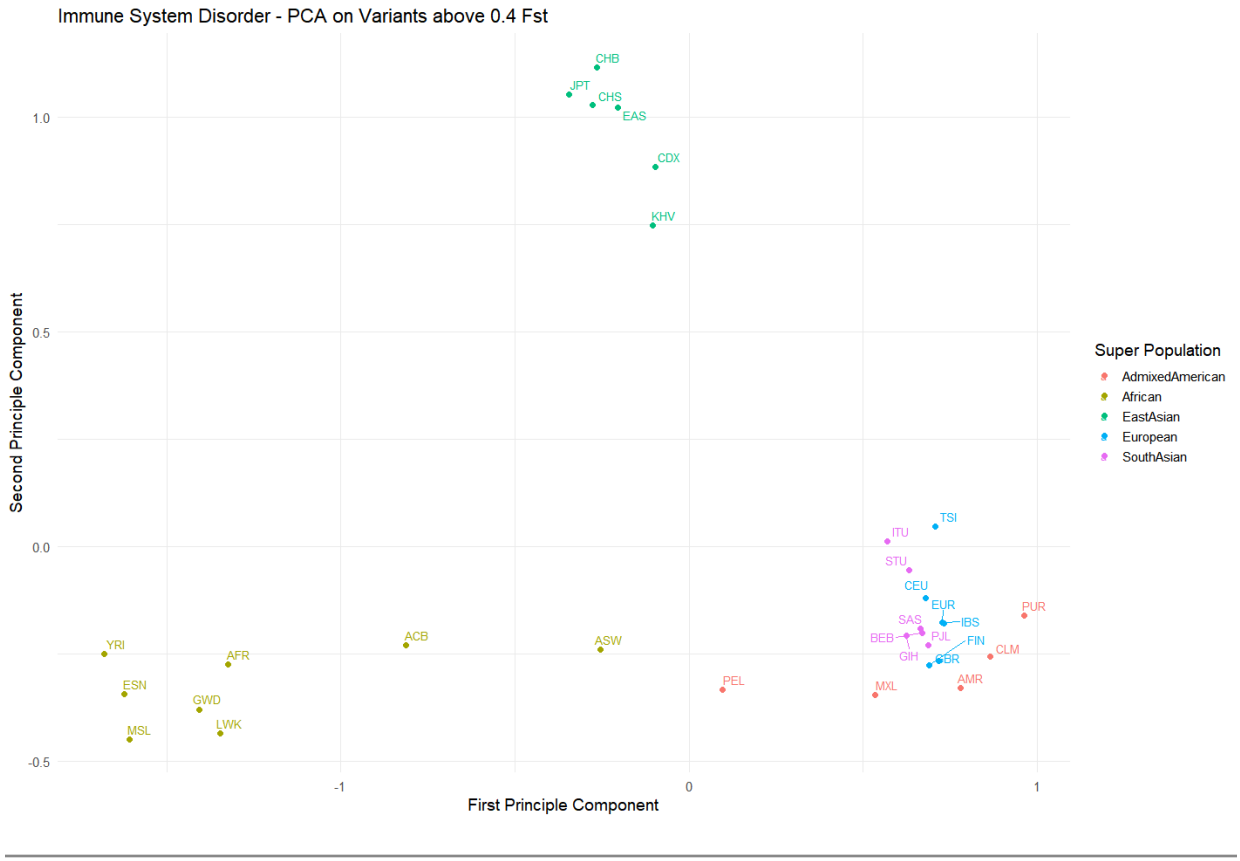
This first PCA demonstrates a patter which will hold across all PCA and is indicative of the underlying factor which seems most prominent in genetic differentiation, which is that of broad ancestry. Super-populations seem to be a primary driver behind measureable differentiation. Most interesting is to observe how this pattern holds regardless of the subset of data used to generate the PCA plot. I have used parent terms from the EFO in order to carve out subsets from the total set of variants above 0.4 Fst, 7 sub terms in total, each of which has their own PCA plot. The below table states the number of variants ≥ 0.4 Fst for each parent term.

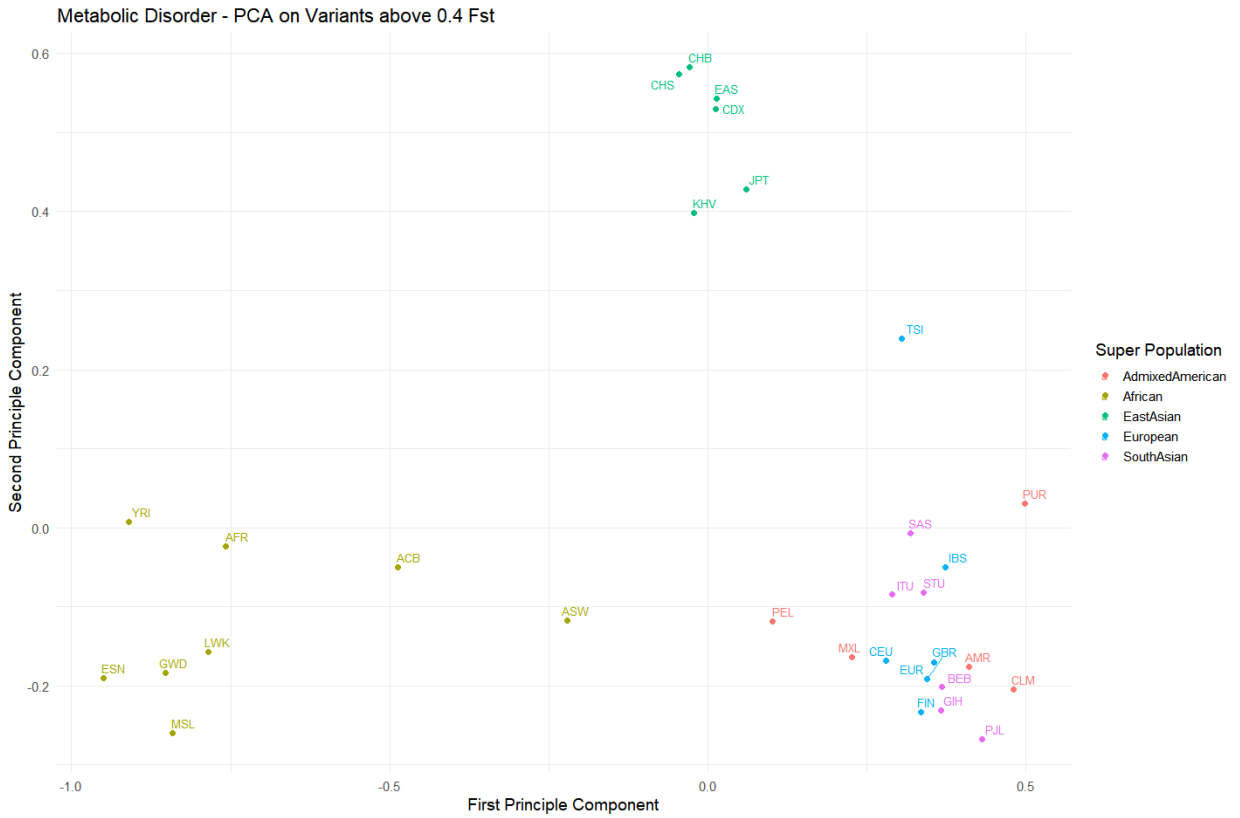
Parent Term	Cancer	Cardiovascular Disease	Digestive System Disorder	Immune System Disorder	Metabolic Disorder	Neurological Disorder	Other Disease	Total
Number Variants ≥ 0.4 Fst	66	51	75	74	18	162	105	477



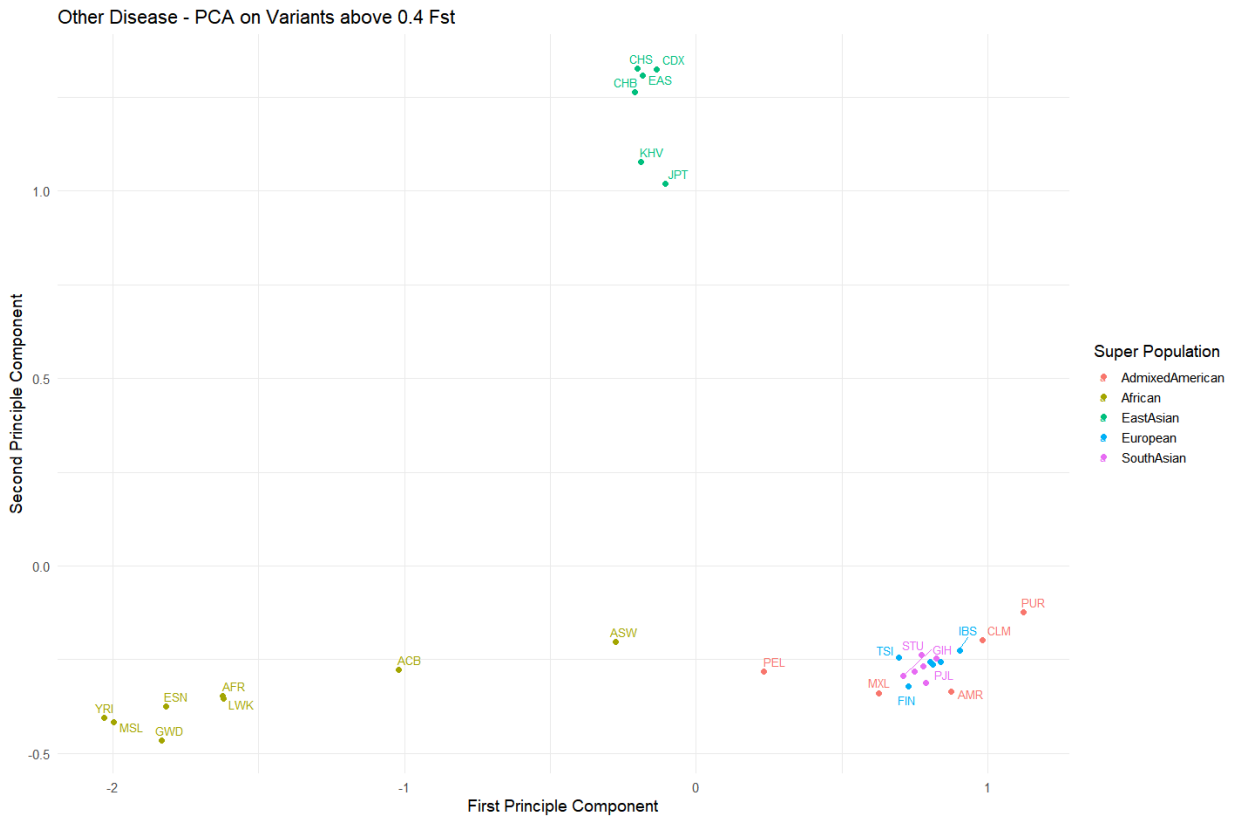

Cardiovascular Disease - PCA on Variants above 0.4 Fst
Second Principle Component
First Principle Component
1.0
0.5
0.0
-1.5
-1.0
-0.5
0.0
0.5
CHB
CHS
CDX
JPT
EAS
KHV
ASW
ACB
YRI
ESN
AFR
LWK
GWD
MSL
TSI
IBS
GBR
EUR
SAS
FIN
CEU
ITU
PUR
BEB
STU
PJL
CLM
MXL
PEL
GIH
AMR
Super Population
AdmixedAmerican
African
EastAsian
European
SouthAsian


Digestive System Disorder - PCA on Variants above 0.4 Fst
Second Principle Component
First Principle Component
1.0
0.5
0.0
-0.5
-1
0
1
CHS
CHB
JPT
EAS
CDX
KHV
ITU
ASW
GIH
TSI
STU
BEB
SAS
PJL
IBS
ACB
YRI
AFR
LWK
ESN
GWD
MSL
CEU
EUR
FIN
GBR
PUR
CLM
MXL
AMR
PEL
Super Population
AdmixedAmerican
African
EastAsian
European
SouthAsian


Immune System Disorder - PCA on Variants above 0.4 Fst
Second Principle Component
First Principle Component
1.0
0.5
0.0
-0.5
-1
0
1
CHB
JPT
CHS
EAS
CDX
KHV
TSI
ITU
STU
CEU
EUR
IBS
SAS
BEB
PJL
FIN
GIH
GBR
PUR
CLM
AMR
MXL
PEL
YRI
AFR
ESN
GWD
MSL
LWK
ACB
ASW
Super Population
AdmixedAmerican
African
EastAsian
European
SouthAsian


Metabolic Disorder - PCA on Variants above 0.4 Fst
Second Principle Component
First Principle Component
0.6
0.4
0.2
0.0
-0.2
-1.0
-0.5
0.0
0.5
CHB
CHS
EAS
CDX
JPT
KHV
TSI
PUR
SAS
IBS
ITU
STU
PEL
MXL
CEU
GBR
AMR
EUR
BEB
CLM
FIN
GIH
PJL
YRI
AFR
ACB
ASW
LWK
ESN
GWD
MSL
Super Population
AdmixedAmerican
African
EastAsian
European
SouthAsian


Neurological Disorder - PCA on Variants above 0.4 Fst
Second Principle Component
First Principle Component
1.5
1.0
0.5
0.0
-0.5
-2
-1
0
1
CHB
CHS
EAS
JPT
CDX
KHV
PEL
ASW
ACB
LWK
AFR
MSL
GWD
YRI
ESN
ITU
BEB
AMR
STU
SAS
CLM
PUR
MXL
PJL
GIH
IBS
TSI
EUR
CEU
GBR
FIN
Super Population
AdmixedAmerican
African
EastAsian
European
SouthAsian



Figs 6-12 - Each PCA plot has a similar shape, the Cancer plot alone seems to be flipped, but still displaying the same broad groupings by super-population

The PCA plot set is indicative of a broad pattern in the data, as one would expect to see some greater intra super-population differentiation at the granularity of broad disease categories if it were extant in the data. Obviously this approach is not as fine as taking variants associated with a single disease and seeing how populations differ with respect to such conditions, which may be a valid next step in digesting this data. With regards to Q1 - *"Are there genes involved in human phenotypes/diseases harboring SNPs which are population-specific?",* I think there is a clear pattern of differentiation across populations, however I think it is evidently the super-population level where differentiation is most evident, which may suggest something about the nature of disease associated variants more generally.

As for Q2 and Q3, data must be gathered on disease incidence and effect on populations for monogenic and multigenic diseases with representation in our data respectively to provide more insight. This next step of analysis will be more granular in nature, as targeting specific diseases rather than broad categories would require subsets even smaller than those used to generate the above PCA plots and thus may yield intra super-population differentiation yet, though I would expect not.

Side Note on Included Materials:

Within the power point are some plots not included in the report itself, such as the number of monogenic associated disease SNPs above 0.4 Fst (and other threshold values). A .txt file containing unique gene names which can be used in *pantherdb.org* for pathway analysis exists as well; I decided not to report on the findings of this task in exploration as they are highly variant and not evidently meaningful without more specific, targeted questions.

Top 20 Variants Per Super-Population:

	AMR		CLM		MXL		PEL		PUR
rs11665674	0.3988	rs1446585	0.4652	rs6088466	0.5127	rs57676627	0.5517	rs63406760	0.5063
rs1047781	0.3921	rs1790099	0.4527	rs13054099	0.4908	rs66783663	0.5234	rs12075	0.4592
rs3827760	0.3202	rs6730157	0.4284	rs9611460	0.4396	rs59998884	0.523	rs12068879	0.4513
rs76317718	0.2773	rs3827760	0.3663	rs10490031	0.4316	rs7140	0.5196	rs3827760	0.3902
rs75748221	0.2711	rs12913832	0.3573	rs10828735	0.4085	rs8127691	0.5194	rs7254272	0.3704
rs144895060	0.2489	rs1129038	0.3567	rs7327620	0.4058	rs762421	0.5173	rs4764080	0.3362
rs11119434	0.2341	rs16824395	0.3421	rs174533	0.4028	rs4456788	0.5155	rs75748221	0.3266
rs77237189	0.2253	rs7254272	0.3341	rs3827760	0.3902	rs151181	0.4834	rs2238151	0.3238
rs80533	0.2228	rs5951698	0.328	rs76509900	0.377	rs2971760	0.4538	rs10744777	0.3233
rs655484	0.2193	rs28538685	0.3188	rs10432638	0.368	rs12484074	0.4466	rs144895060	0.3186
rs17866443	0.2142	rs8112559	0.3184	rs2092563	0.3638	rs77125470	0.4401	rs45471499	0.3152
rs117483894	0.2135	rs45471499	0.3152	rs2093210	0.3604	rs8056890	0.4365	rs3800917	0.302
rs13054099	0.2134	rs293566	0.3134	rs17123757	0.3528	rs174549	0.4323	rs604723	0.3018
rs11201999	0.211	rs4833103	0.3111	rs2237892	0.3517	rs11612312	0.432	rs57676627	0.2849
rs2413631	0.2046	rs1567084	0.3091	rs9607782	0.3494	rs11610143	0.432	rs12450323	0.2779
rs76442143	0.1998	rs17623373	0.3067	rs12485034	0.3493	rs4822020	0.4289	rs11119434	0.2731
rs1385374	0.1989	rs7305242	0.3042	rs4821942	0.3485	rs2785988	0.4208	rs1865680	0.2657
rs11059927	0.1974	rs4766898	0.3012	rs10483727	0.345	rs2820446	0.4205	rs655484	0.2587
rs11059928	0.1974	rs10919928	0.2959	rs34935520	0.3443	rs4846567	0.4205	rs143886976	0.2572
rs11673591	0.1968	rs7521492	0.2893	rs2411306	0.3399	rs2820443	0.4197	rs1737894	0.2551

Fig - Admixed American top 20 variants

	BEB		GIH		ITU		PJL		SAS		STU
rs16824395	0.4577	rs12074934	0.5015	rs7254272	0.4962	rs3827760	0.3902	rs2643826	0.4608	rs2643826	0.4608
rs3827760	0.4339	rs57676627	0.4863	rs12171500	0.4879	rs4764080	0.3362	rs12075	0.4592	rs12075	0.4592
rs2526678	0.392	rs943451	0.4437	rs12627970	0.4421	rs75748221	0.3266	rs1790099	0.4501	rs12297948	0.4425
rs12918327	0.3648	rs4796791	0.4384	rs3026433	0.4402	rs144895060	0.3186	rs3827760	0.3902	rs12479436	0.4225
rs73885319	0.3597	rs143384	0.4107	rs28538685	0.4398	rs45471499	0.3152	rs112238765	0.3676	rs11692588	0.4165
rs5757922	0.3399	rs58629129	0.4091	rs63406760	0.4247	rs62153901	0.2998	rs9532984	0.3674	rs11114149	0.4139
rs9785971	0.3325	rs1517352	0.4075	rs9358912	0.4227	rs57676627	0.2849	rs11715126	0.3648	rs3741353	0.4065
rs1047781	0.3058	rs11114149	0.4043	rs58629129	0.4224	rs11798231	0.2742	rs35257692	0.3555	rs4842266	0.4035
rs117949785	0.3051	rs618205	0.3966	rs2358973	0.4199	rs11119434	0.2731	rs63406760	0.3488	rs12918327	0.4004
rs11065828	0.2988	rs12297948	0.3966	rs12450323	0.4043	rs12918327	0.2723	rs820430	0.3436	rs1790099	0.3919
rs62388754	0.2967	rs12479436	0.3942	rs2602813	0.3989	rs1865680	0.2657	rs72375069	0.3398	rs3827760	0.3902
rs62392216	0.291	rs4842266	0.3938	rs12068879	0.3936	rs655484	0.2587	rs4764080	0.3362	rs11603634	0.3879
rs35629860	0.2907	rs11603634	0.3883	rs12918327	0.3814	rs143886976	0.2572	rs35955841	0.332	rs56116847	0.3869
rs17623373	0.2904	rs11692588	0.388	rs3827760	0.3793	rs11192283	0.2539	rs12918327	0.3293	rs6930773	0.3833
rs34480360	0.2885	rs3741353	0.3872	rs12075	0.3774	rs1156389	0.2529	rs75748221	0.3266	rs34480360	0.3794
rs75748221	0.2809	rs6736175	0.3784	rs11150602	0.3673	rs830620	0.2521	rs3805236	0.3252	rs13233308	0.373
rs77871618	0.2753	rs1687657	0.3674	rs11203032	0.3447	rs12931235	0.2507	rs293566	0.3216	rs35629860	0.3693
rs147967693	0.2749	rs13233308	0.363	rs9651899	0.3431	rs1932040	0.2477	rs6021247	0.3196	rs874168	0.3676
rs11066359	0.2739	rs56116847	0.3484	rs12625546	0.3326	rs11905172	0.2475	rs144895060	0.3186	rs112238765	0.3676
rs57676627	0.2707	rs2602813	0.3473	rs1790099	0.3305	rs17749211	0.2471	rs13206608	0.3154	rs9532984	0.3674

Fig - South Asian top 20 variants

	CDX
rs260643	0.6058
rs12074934	0.521
rs1129038	0.5158
rs12913832	0.5153
rs3827760	0.4975
rs5743618	0.4922
rs1446585	0.4855
rs67969609	0.4759
rs6730157	0.4666
rs2306125	0.4658
rs12621647	0.4507
rs12075	0.4503
rs174565	0.4492
rs174570	0.4492
rs10772040	0.4478
rs2153219	0.4438
rs13010313	0.4431
rs2441727	0.4386
rs9277348	0.4374
rs815609	0.4369

	CHB
rs57676627	0.6282
rs260643	0.603
rs14235	0.5665
rs11150602	0.5369
rs1047781	0.5077
rs815609	0.4819
rs12074934	0.4802
rs897984	0.4785
rs2306125	0.4723
rs760500	0.4699
rs590616	0.4635
rs13708	0.4504
rs10782001	0.4498
rs6801781	0.4495
rs739496	0.4486
rs103294	0.4484
rs67250450	0.4452
rs11250135	0.4427
rs9568402	0.439
rs2009262	0.4329

	CHS
rs260643	0.5577
rs63406760	0.5095
rs760500	0.5024
rs2306125	0.4827
rs13168358	0.4793
rs815609	0.4658
rs14235	0.461
rs4805962	0.4578
rs13010313	0.4517
rs2384000	0.4383
rs9385400	0.4368
rs4833103	0.4349
rs67969609	0.4333
rs10245867	0.4332
rs9388490	0.432
rs6801781	0.4312
rs11150602	0.4307
rs7904519	0.4265
rs2012820	0.4258
rs897984	0.4205

	EAS
rs12913832	0.6997
rs1129038	0.6934
rs260643	0.5355
rs9785971	0.4853
rs12074934	0.469
rs2306125	0.4609
rs815609	0.4597
rs14235	0.4538
rs760500	0.4424
rs590616	0.4358
rs13168358	0.4336
rs10772040	0.4316
rs4805962	0.4223
rs6801781	0.422
rs2009262	0.4198
rs9651899	0.4158
rs11249906	0.413
rs13010313	0.4123
rs67250450	0.4101
rs11250135	0.4091

	JPT
rs815609	0.5272
rs2306125	0.4971
rs4805962	0.493
rs75412658	0.492
rs760500	0.4862
rs968451	0.4845
rs14235	0.4818
rs2157453	0.4782
rs13168358	0.4782
rs9286879	0.4696
rs174565	0.4611
rs174570	0.4611
rs2267407	0.4527
rs11250135	0.4522
rs5757628	0.4513
rs6801781	0.4498
rs9385400	0.4469
rs67250450	0.4455
rs9388490	0.4421
rs4836913	0.4412

	KHV
rs2904880	0.5362
rs7184597	0.5276
rs260643	0.5276
rs11814448	0.5267
rs3790553	0.4742
rs2009262	0.473
rs2153219	0.4561
rs10772040	0.4398
rs4787458	0.438
rs4796791	0.4358
rs590616	0.4334
rs174546	0.4285
rs174551	0.4285
rs174547	0.428
rs17023134	0.4245
rs174537	0.422
rs174541	0.4205
rs739496	0.4187
rs58629129	0.4128
rs11249906	0.4103

Fig - East Asian top 20 variants

	EUR
rs5743618	0.5649
rs185146	0.5263
rs57676627	0.5263
rs9785971	0.4648
rs959071	0.4323
rs10421769	0.4003
rs11590283	0.3973
rs35390	0.3543
rs2275247	0.3542
rs10032909	0.3539
rs10828735	0.3361
rs6430538	0.3323
rs618205	0.3304
rs28538685	0.329
rs8042680	0.3204
rs1419138	0.3197
rs45471499	0.3152
rs2441727	0.3067
rs75748221	0.3041
rs7170666	0.2977

	FIN
rs7403279	0.5382
rs185146	0.5315
rs5743618	0.4803
rs63406760	0.478
rs12068879	0.4376
rs6847640	0.403
rs2032624	0.3948
rs28538685	0.3915
rs959071	0.3811
rs1865680	0.3648
rs1016189	0.3603
rs9651899	0.3581
rs11665674	0.3568
rs117949785	0.3539
rs35390	0.3526
rs4908343	0.3519
rs6457709	0.3462
rs11590283	0.3448
rs9277332	0.3294
rs2275247	0.3244

	GBR
rs185146	0.5831
rs7254272	0.523
rs959071	0.4437
rs12756986	0.4377
rs1790099	0.4217
rs35390	0.4136
rs7403279	0.3968
rs10919928	0.3965
rs11590283	0.3685
rs1419138	0.3634
rs4665630	0.3627
rs7170666	0.3604
rs2275247	0.3537
rs117949785	0.3447
rs2177223	0.3412
rs6867265	0.3347
rs4921542	0.3317
rs2255280	0.3278
rs6488140	0.3266
rs314580	0.3214

	IBS
rs959071	0.4333
rs12068879	0.4281
rs11590283	0.4063
rs11203032	0.3989
rs2526678	0.3986
rs185146	0.3921
rs3805236	0.3869
rs10828735	0.3814
rs2441727	0.3796
rs1790099	0.3697
rs3827760	0.3613
rs11202345	0.3604
rs63406760	0.3595
rs12075	0.3458
rs12878003	0.3423
rs2275247	0.3417
rs7254272	0.3403
rs12931235	0.3381
rs16824395	0.3379
rs1459513	0.3322

	TSI
rs185146	0.5694
rs959071	0.4771
rs11814448	0.4665
rs943451	0.462
rs2643826	0.4608
rs12075	0.4592
rs993471	0.459
rs56335113	0.4567
rs58629129	0.4447
rs12297948	0.4425
rs143384	0.4387
rs10873298	0.4381
rs4796791	0.4376
rs10873299	0.4375
rs8181996	0.4353
rs11590283	0.4325
rs2058619	0.4273
rs12479436	0.4225
rs897984	0.4191
rs11692588	0.4165

Fig - European top 20 variants