Supplemental Information

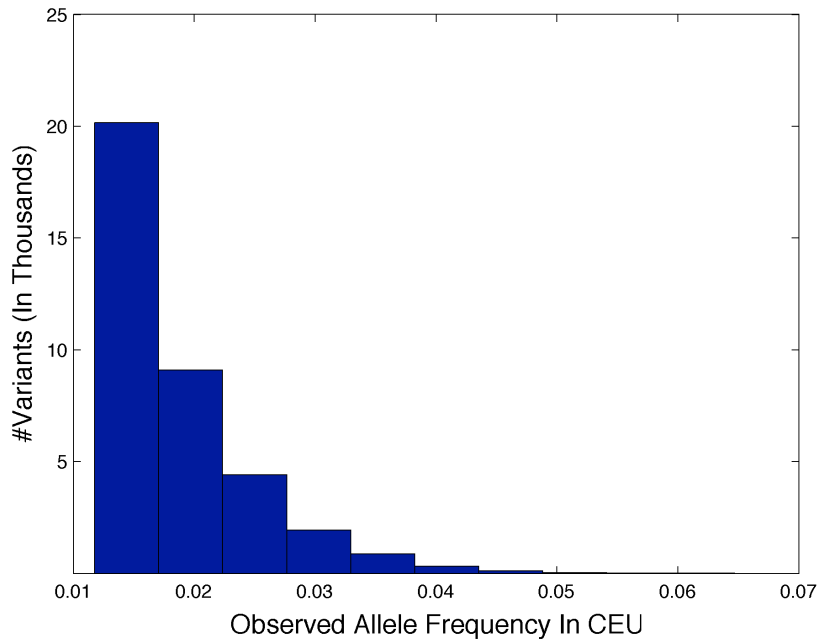Figure S1: Histogram of variants with allele frequencies <1% in 1000 Genomes but >1% in the CEU. It shows that allele frequencies can be highly structured for rare variants and that by averaging across too many genetically dissimilar populations can have a downwardly bias affect on frequency estimates of alleles present in a given population.
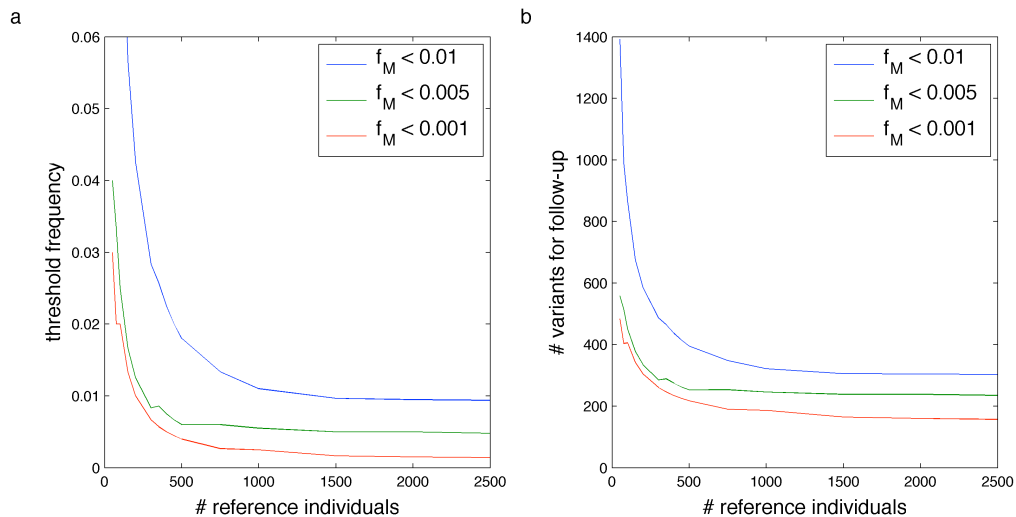

Figure S2: Reference panel size impacts the efficacy of filtering in exome sequencing in African-American simulations from the EVS data.
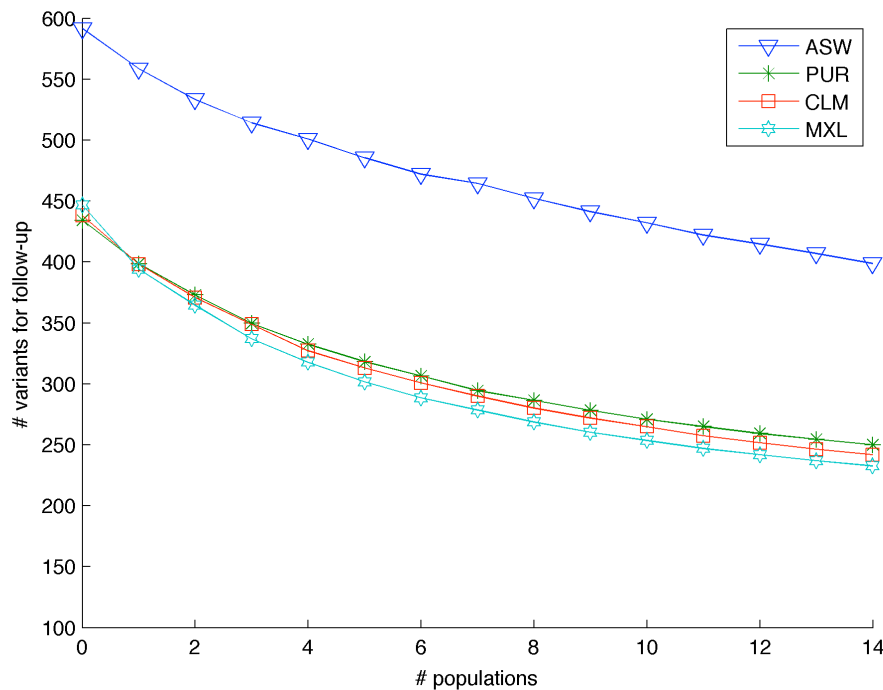
Figure S3. Average number of variants for follow-up in the ASW, CLM, PUR and MXL when leveraging cross-population information. The number of variants for follow-up after this filtering for each of the legend populations is shown along the Y-axis of the graph. Further filtering removed alleles if they were observed in any of the additional populations above the 5% FNR threshold determined for those populations. The number of randomly selected populations on which this filtering is based is given in the x-axis.

Individual	Country	Inheritance	Causal Variant zygosity	Variants Remaining Pre-Frequency Filtering	No Ancestry, f>1%	PopMatched, FNR <5%	AllPop, FNR <5%
1	Jordan	Recessive	Homozygous	5016	65	43	30
2	Turkey	Recessive	Homozygous	5049	50	20	19
3	Turkey	Recessive	Homozygous	4752	25	23	19
4	Iran	Recessive	Homozygous	4674	39	15	11
5	Syria	Recessive	Homozygous	4751	26	18	16
6	Tunisia	None Given	Homozygous	4834	37	19	11
7	Jordan	Recessive	Homozygous	4722	22	11	4
8	Jordan	Recessive	Homozygous	4973	104	102	52
9	Iran	None Given	Homozygous	5653	112	67	48
10	Jordan	None Given	Homozygous	5938	97	83	63

11	Jordan	None Given	Heterozygous	15694	872	750	498
12	Turkey	Recessive	Heterozygous	12171	619	445	342
13	Tunisia	Dominant	Heterozygous	14111	813	747	461
14	Iran	Dominant	Heterozygous	13727	726	535	421
15	Iran	Dominant	Heterozygous	13417	661	512	379
16	Jordan	None Given	Heterozygous	14713	871	690	511
17	Jordan	None Given	Heterozygous	13472	668	543	384
18	Jordan	None Given	Heterozygous	13789	748	623	450
19	Iran	None Given	Heterozygous	15128	769	598	465
20	Turkey	Recessive	Compound Heterozygous	12919	604	426	352

Table S1. Analysis of real data by individual exome. Variants remaining pre-frequency filtering reflects the number of variants remaining when all variants without damaging annotations are filtered out and when variants inconsistent with the disorder architecture are removed but before frequency-based filtering has been preformed.