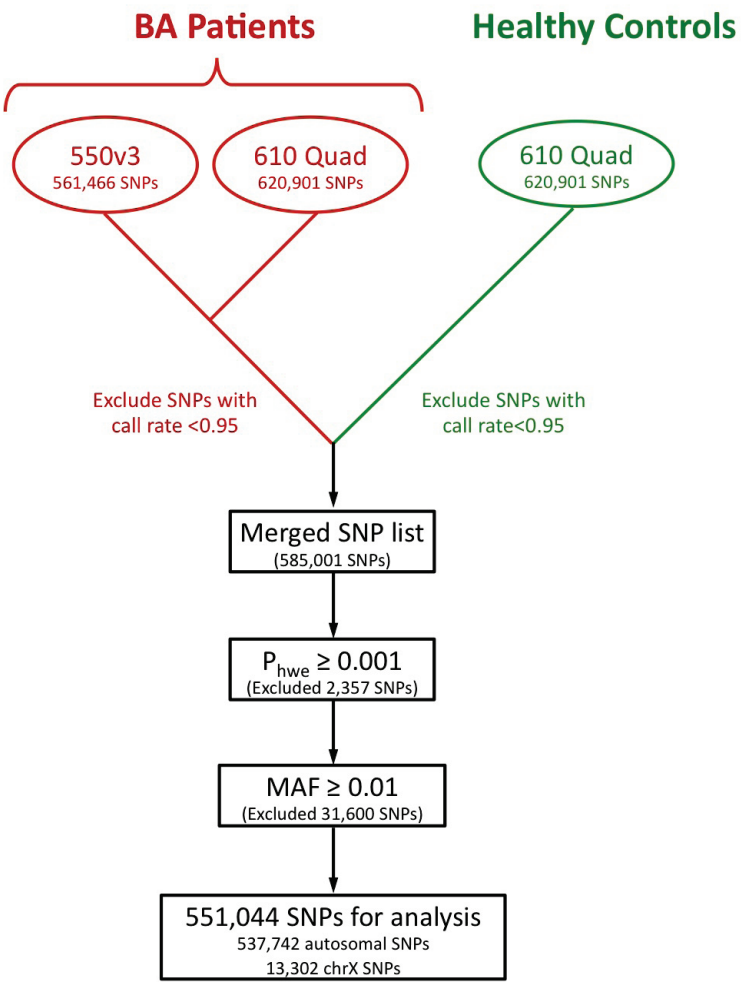
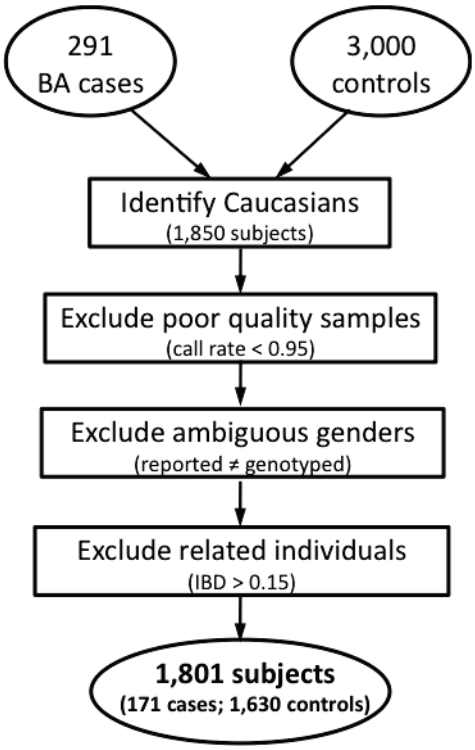
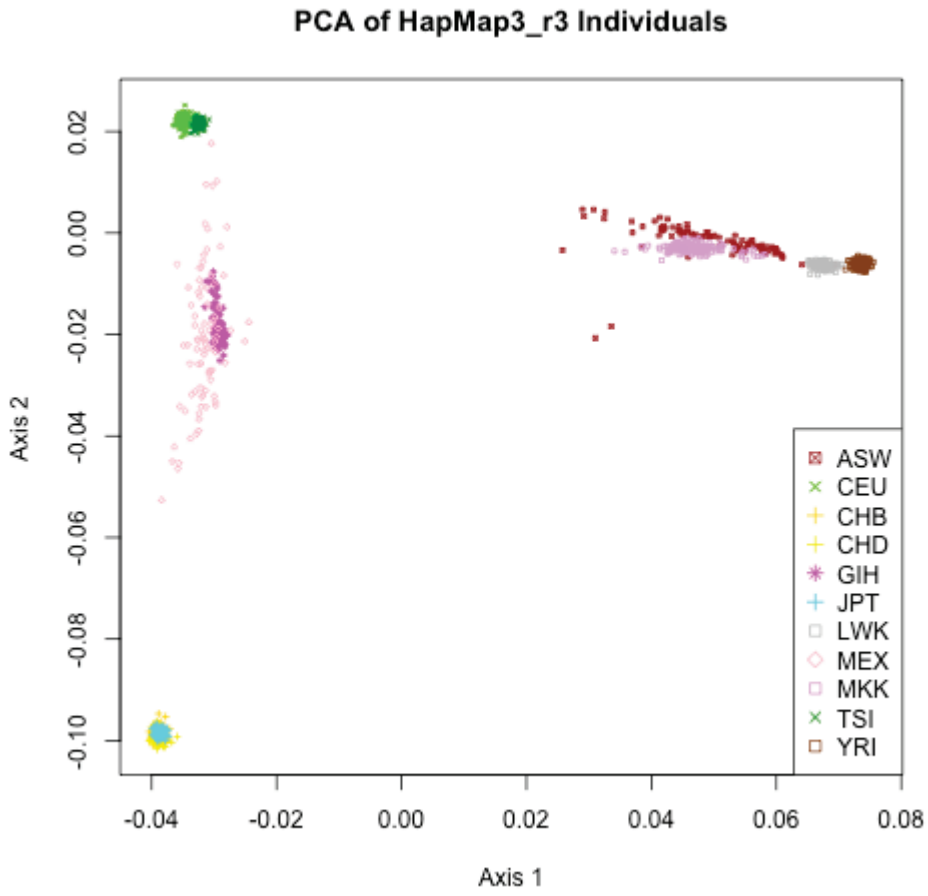


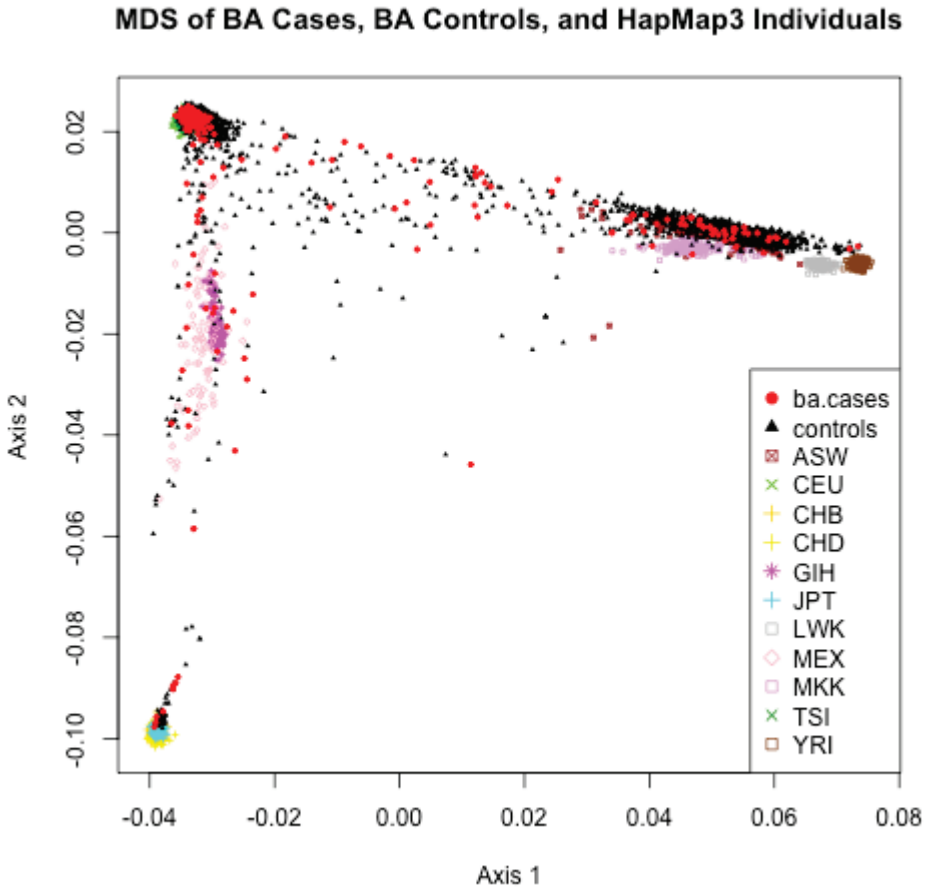
Supplementary Figure 1. SNP Selection Flowchart
SNPs that were only present on the 550v3 were removed. Standard thresholds were set and SNPs that did not meet these thresholds were removed from the association test.



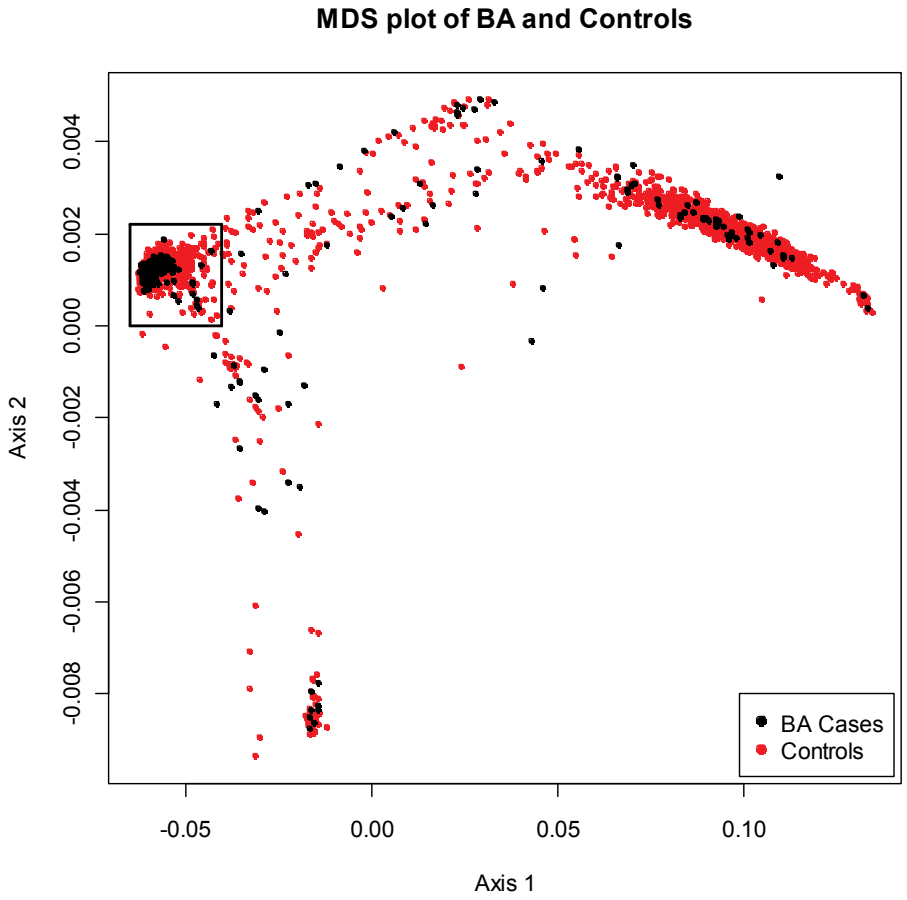
Supplementary Figure 2. Sample Selection Flowchart
The entire cohort of available BA DNA samples was genotyped, but only the samples identified as Caucasian from the genotype data were used for the association test. The 171 BA cases and 1,630 healthy controls passed the sample quality control and were used in our association study.



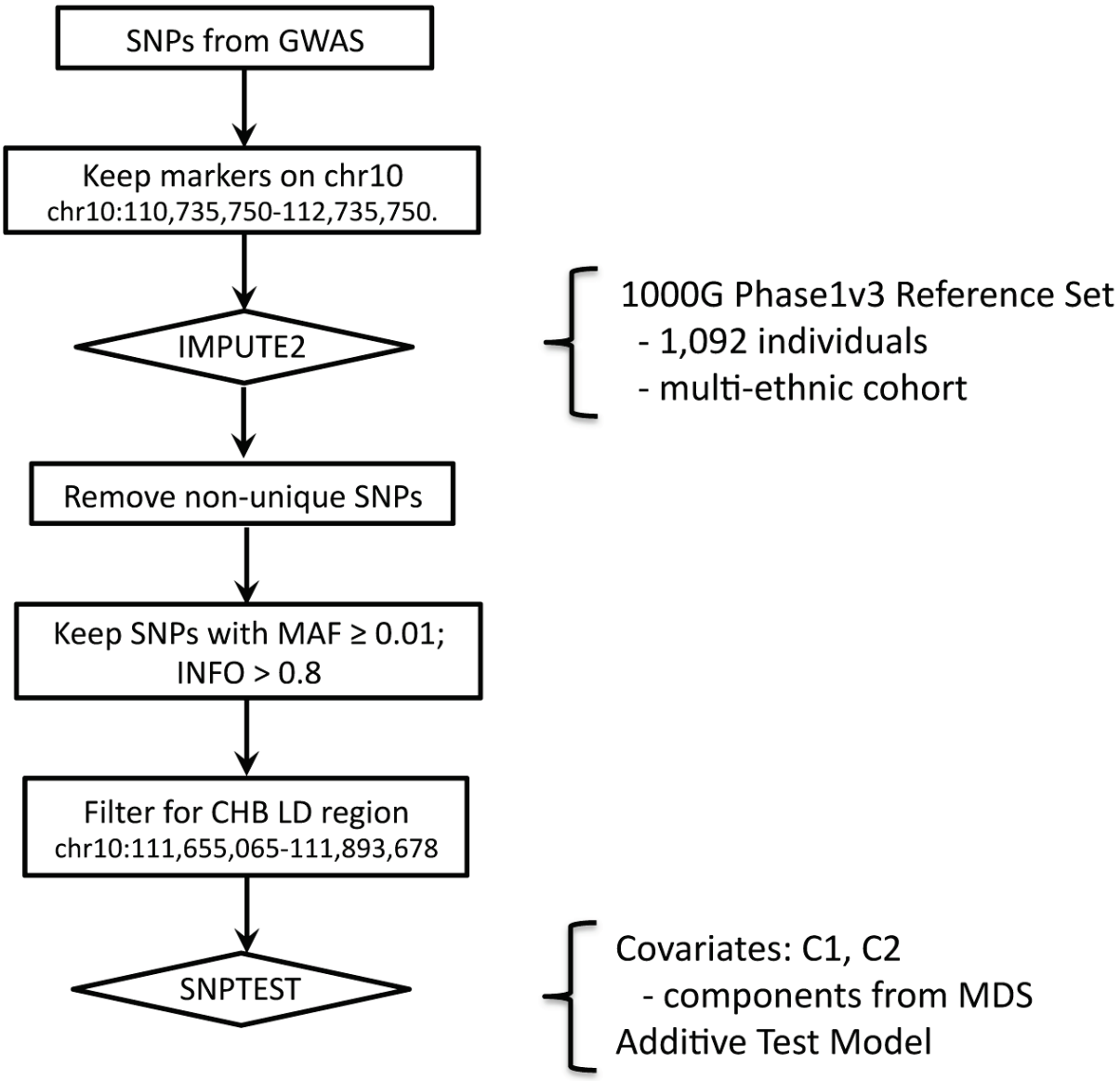
Supplementary Figure 3A. MDS plot of HapMap 3 Samples.
The distribution of the 13 HapMap 3 populations are first plotted here for comparison with the BA case and control samples in the next figure.



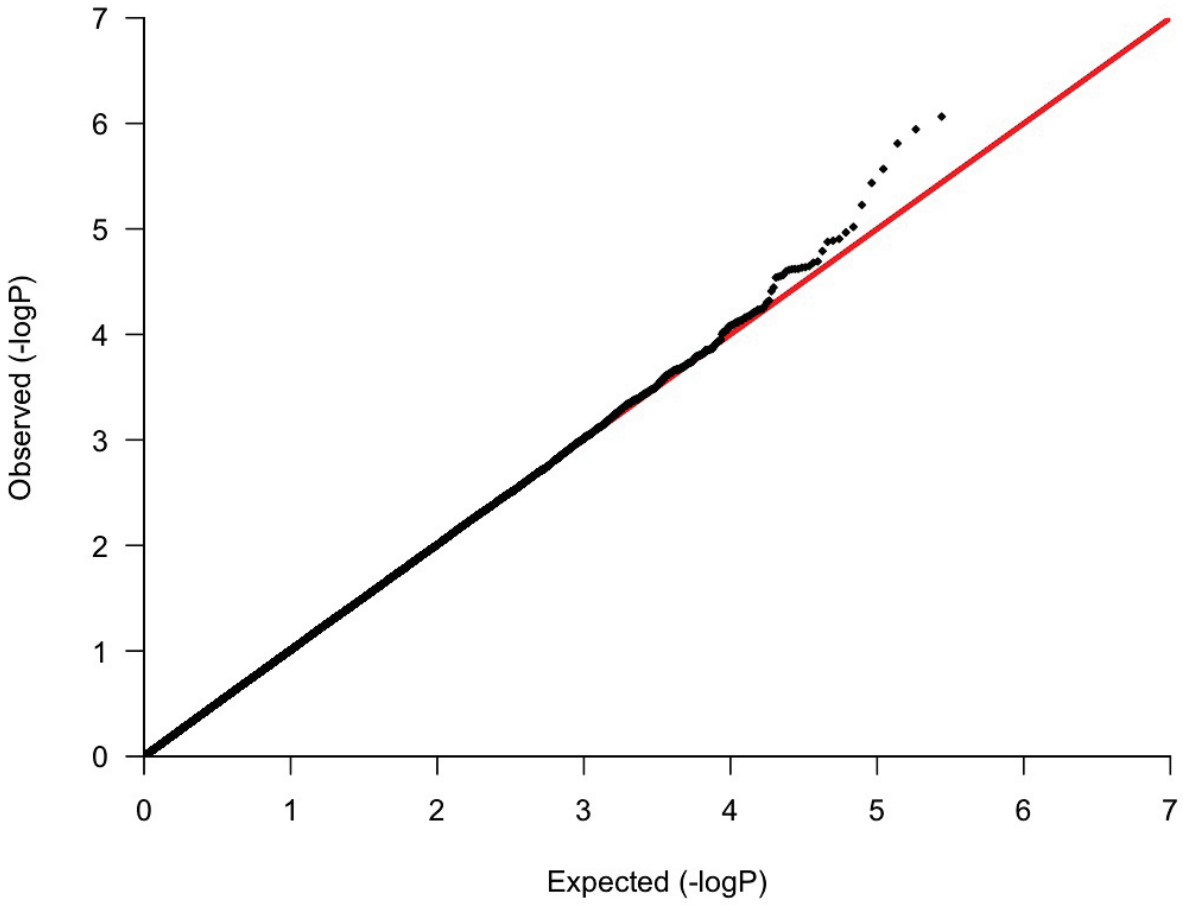
Supplementary Figure 3B. MDS plot of BA cases and controls overlaid over HapMap 3 samples. This figure plots the first two MDS components of the BA case and control samples over the 11 HapMap 3 populations. As expected from the self-reported ethnicities, most of the BA cases overlie the CEU and TSI Caucasian populations.



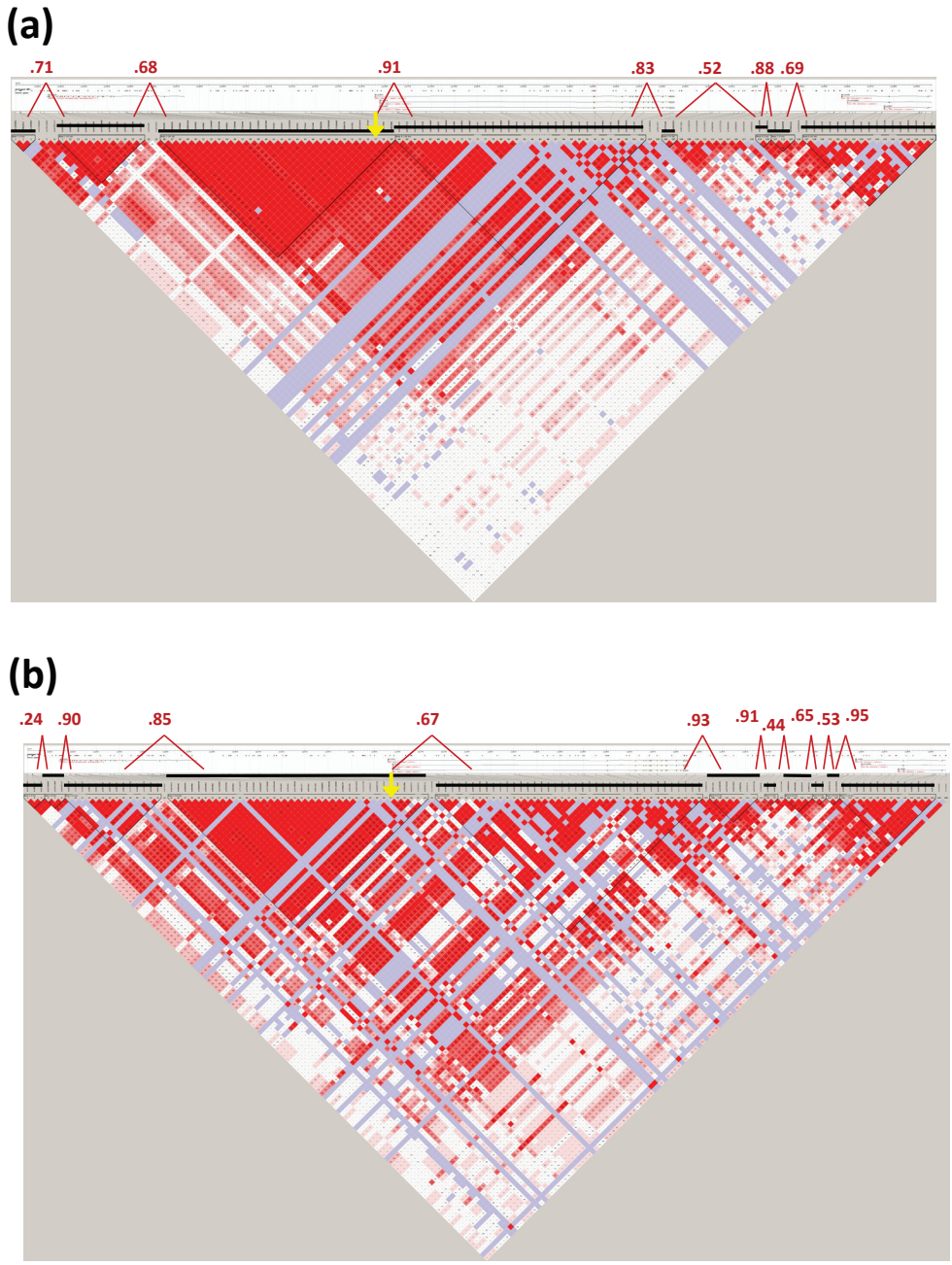
Supplementary Figure 3C. Plot of the First Two MDS Components of BA cases and controls. The MDS components 1 and 2 for all the 291 BA samples and 3,000 healthy controls are plotted above. The samples inside the box are within two standard deviations from the means of the self-reported Caucasian samples cohort and are used in the association study.



Supplementary Figure 4. Imputation and Association Test Flowchart.
Regional imputation was performed in the 2 Mb interval surrounding the reported SNP association signal. Only the imputed SNPs with a high information score and a minor allele frequency greater than 1% were used for our association test. Furthermore, SNPs outside the Chinese haplotype block were excluded.

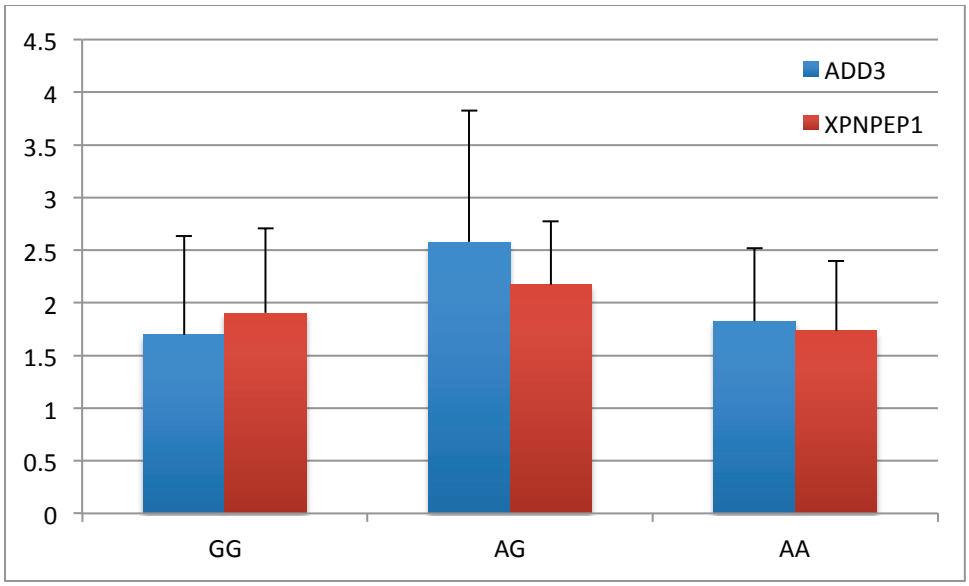


Supplementary Figure 5. Q-Q Plot of the Genotyped SNP Association Test.
In this Q-Q plot, the observed p-values are from the logistic regression association test adjusted for MDS components 1 and 2. There is no evidence of genomic inflation; most of the markers do not deviate from their expected values and the genomic inflation factor is close to 1 (λ=1.01). The SNPs with p < 1e-5 in order of increasing p-value are rs2345662 (7p22.1), rs12683683 (9q33.3), rs1867896 (2q21.2), rs9359428 (6q14.1), rs1612815 (18q23), rs3779312 (7q21.11), rs481777 (6q14.1), rs2287595 (7p22.1), rs946442 (1p32.3), rs11069817 (13q33.3), rs4702142 (5p15.1), rs7452665 (6q12) and rs11057354 (12q24.31).



Supplementary Figure 6. Linkage Disequilibrium Plots of the Region Surrounding *XPNPEP1* and *ADD3*.

LD plots of the region chr10:111,600,000-112,000,000 (hg18) of HapMap 3 Release 2 populations were generated in Haploview.[1] The linkage disequilibrium block in the Chinese of Han descent in Beijing (CHB) is pictured in (a). The linkage disequilibrium block in the Caucasians of European ancestry in Utah (CEU) and Tuscans in Italy (TSI) is pictured in (b). The black bars above the graph represent the haplotype blocks as calculated by the default algorithm in Haploview. The values above these bars in red represent the correlation between neighboring blocks. The yellow arrow points to rs17095355.



Supplementary Figure 7. Expression levels of candidate genes by the genotype of rs7099604. Data from 30 BA patients with available genotypes (4 GG, 7 AG, 19 AA), were plotted with their corresponding expression levels. There is no correlation of *XPNPEP1* (p=0.34) and *ADD3* (p=0.61) expression with the underlying patient genotype.

Sample Ascertainment			Genotype Detail					Alleles	
Population	Individual Group	Chrom. Sample Cnt.	C/C	C/T	T	T/T	HWP	C	T
HapMap-HCB	Asian	86	0.349	0.419		0.233	0.343	0.558	0.442
HAPMAP-GIH	Asian	176	0.352	0.432		0.216	0.273	0.568	0.432
HAPMAP-CHD	Asian	168	0.369	0.440		0.190	0.439	0.589	0.411
HAPMAP-CHB	Asian	82	0.390	0.439		0.171	0.655	0.610	0.390
HAPMAP-ASW	African	98	0.367	0.551		0.082	0.200	0.643	0.357
HAPMAP-LWK	African	180	0.422	0.467		0.111	0.752	0.656	0.344
HapMap-YRI	Sub-Saharan African	226	0.460	0.407		0.133	0.371	0.664	0.336
HapMap-JPT	Asian	170	0.412	0.518		0.071	0.150	0.671	0.329
HAPMAP-TSI	Southern European	176	0.545	0.386		0.068	1.000	0.739	0.261
HAPMAP-MKK	African	276	0.601	0.355		0.043	0.752	0.779	0.221
HapMap-CEU	Western & Northern European	226	0.805	0.195			0.655	0.903	0.097
HAPMAP-MEX		100	0.860	0.120		0.020	0.200	0.920	0.080

Supplementary Figure 8. Population Diversity of rs17095355 in HapMap 3.
This table of allele frequencies for rs17095355 in the HapMap populations was retrieved from dbSNP.[2] The T allele of this C/T SNP varies from 8-44% in different populations. The populations are sorted in descending order of minor allele frequencies. Interestingly, most of the Asian populations have a higher minor allele frequency than the Caucasian populations. The CEU and MEX populations have reported minor allele frequency less than 10%, while other populations have a reported value more than 20%.

Reference:

1. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005 Jan 15.

2. Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine. dbSNP accession:rs17095355, (dbSNP Build ID: 137). Available from: http://www.ncbi.nlm.nih.gov/SNP/