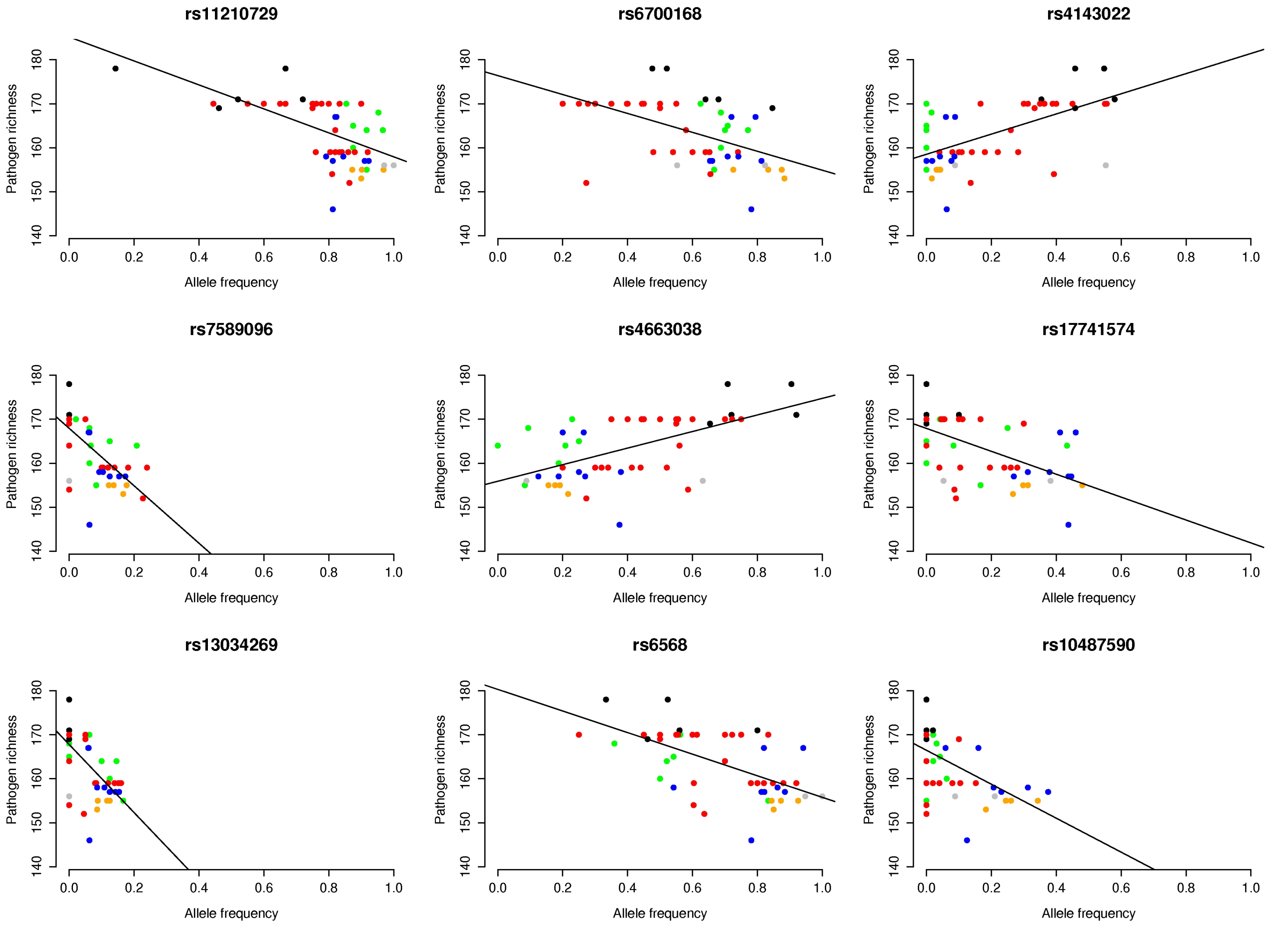
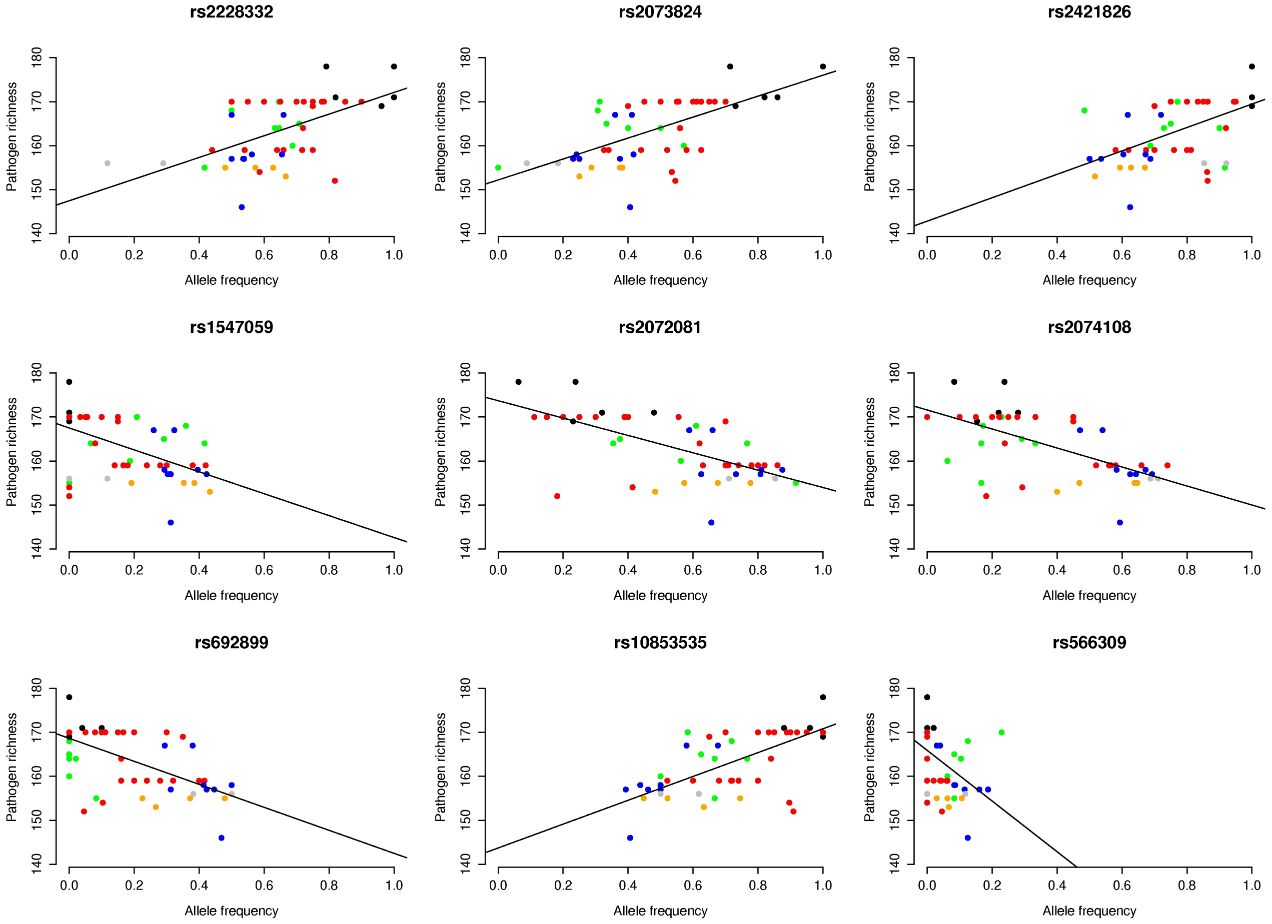
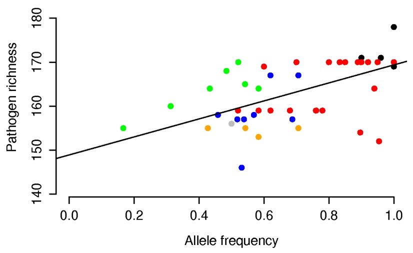
Supplementary Figure 1

Correlation between pathogen richness and allele frequency for all significantly associated BGA SNPs. Populations from different broad geographic areas are coded by different colors: Sub-Saharan Africa (green), America (black), Asia (red), Europe (blue), Middle East (orange), Oceania (gray).

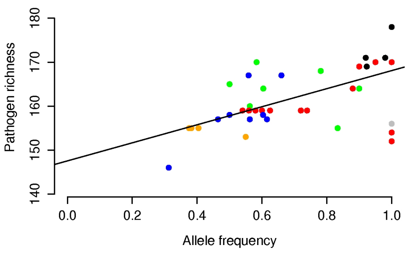



rs2228332
rs2073824
rs2421826
rs1547059
rs2072081
rs2074108
rs692899
rs10853535
rs566309
Pathogen richness
140
150
160
170
180
Allele frequency
0.0
0.2
0.4
0.6
0.8
1.0

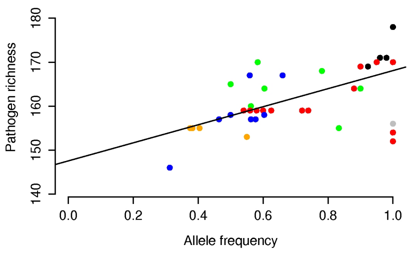
rs6507641



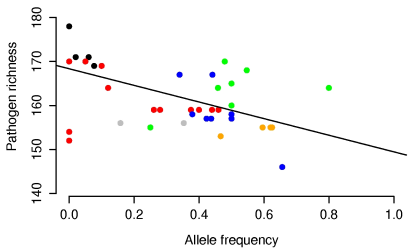
rs602662



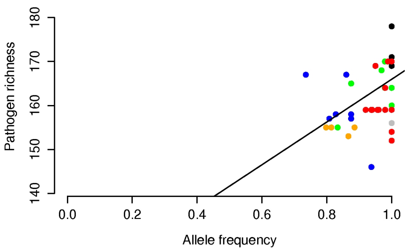
rs485186



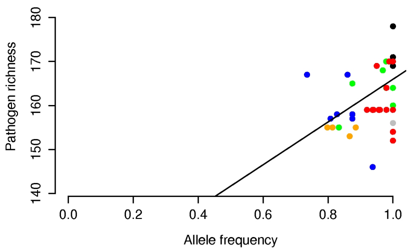
rs504963



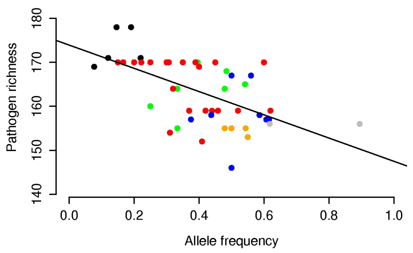
rs9466910



rs9466912

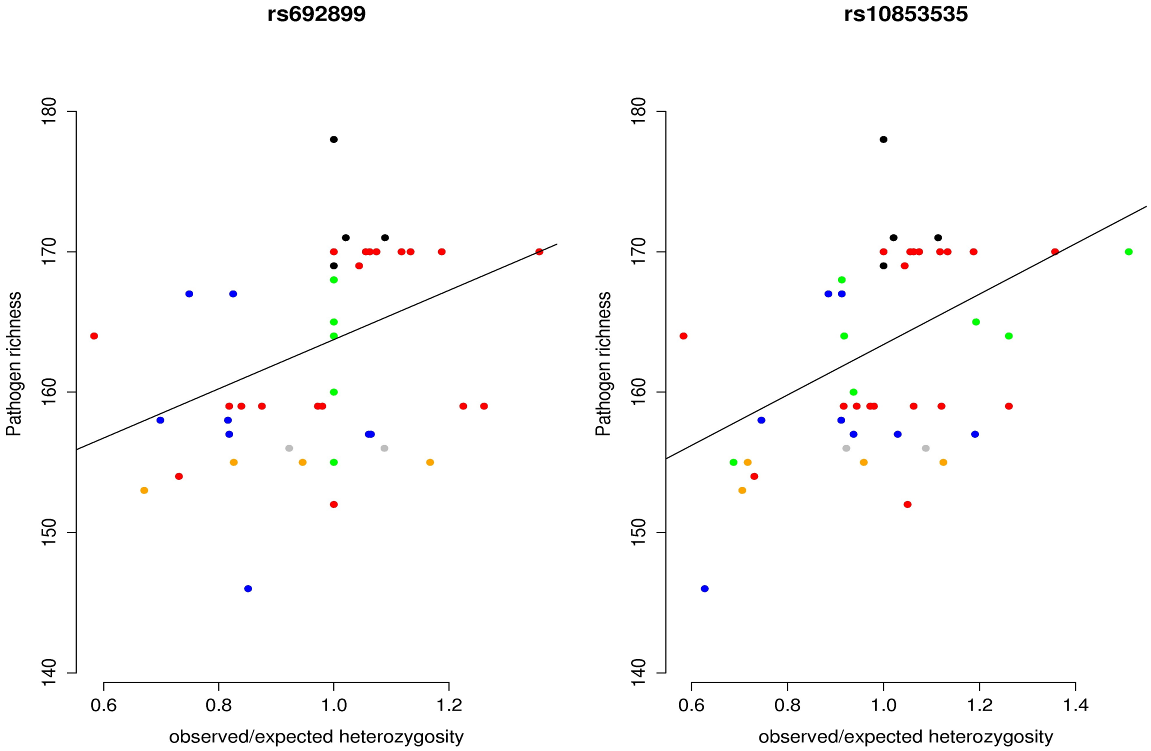


rs17576994



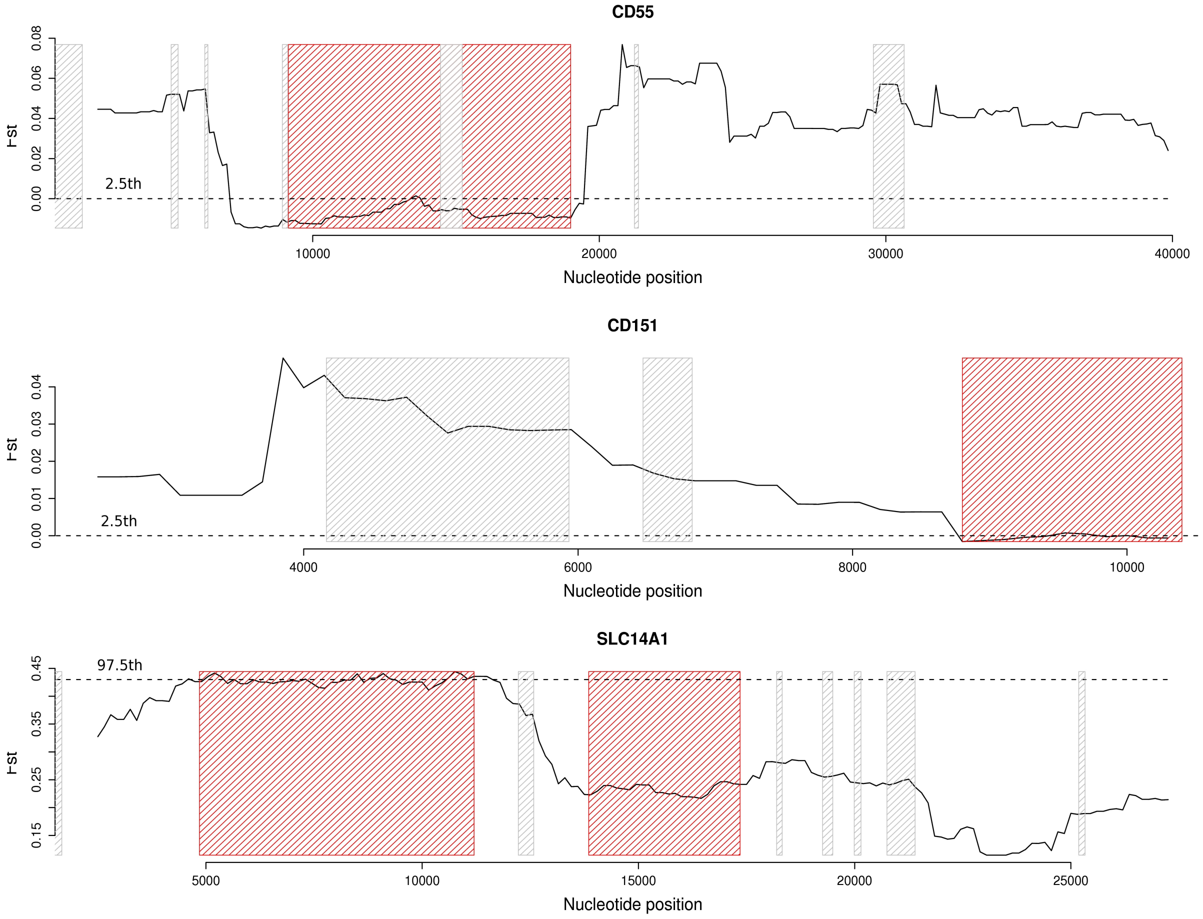
Supplementary Figure 2

Correlation between observed/expected heterozygosity and pathogen richness for two SNPs in *SLC14A1* (see text). Populations from different broad geographic areas are coded by different colors: Sub-Saharan Africa (green), America (black), Asia (red), Europe (blue), Middle East (orange), Oceania (gray).



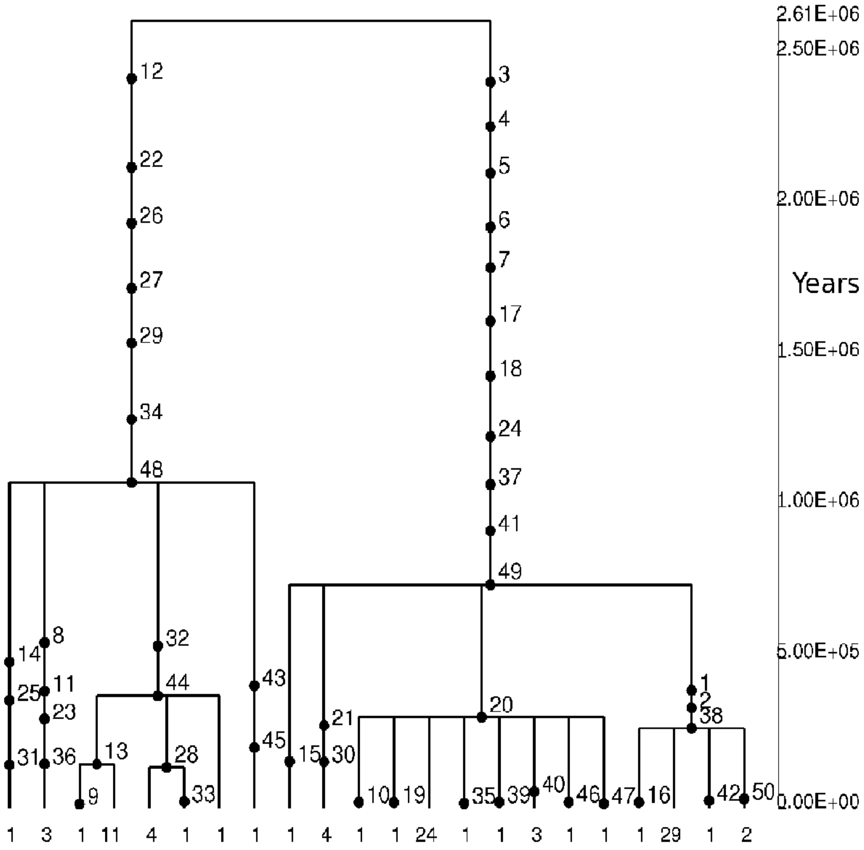
Supplementary Figure 3

Sliding window analysis of F_{ST}. Red shading identifies the regions we analyzed. Gray shading indicates resequencing gaps. We used windows of 5 kb sliding with a step of 150 bp. Hatched lines represent the 2.5th and 97.5th percentiles in the distribution of F_{ST} values calculated for SeattleSNPs genes. Negative F_{ST} values should be interpreted as equal to 0.



Supplementary Figure 4

Estimated tree for the *CD55* gene region we analyzed. Mutations are represented as black dots and named for their physical position along the region. The absolute frequency of each haplotype is also reported.

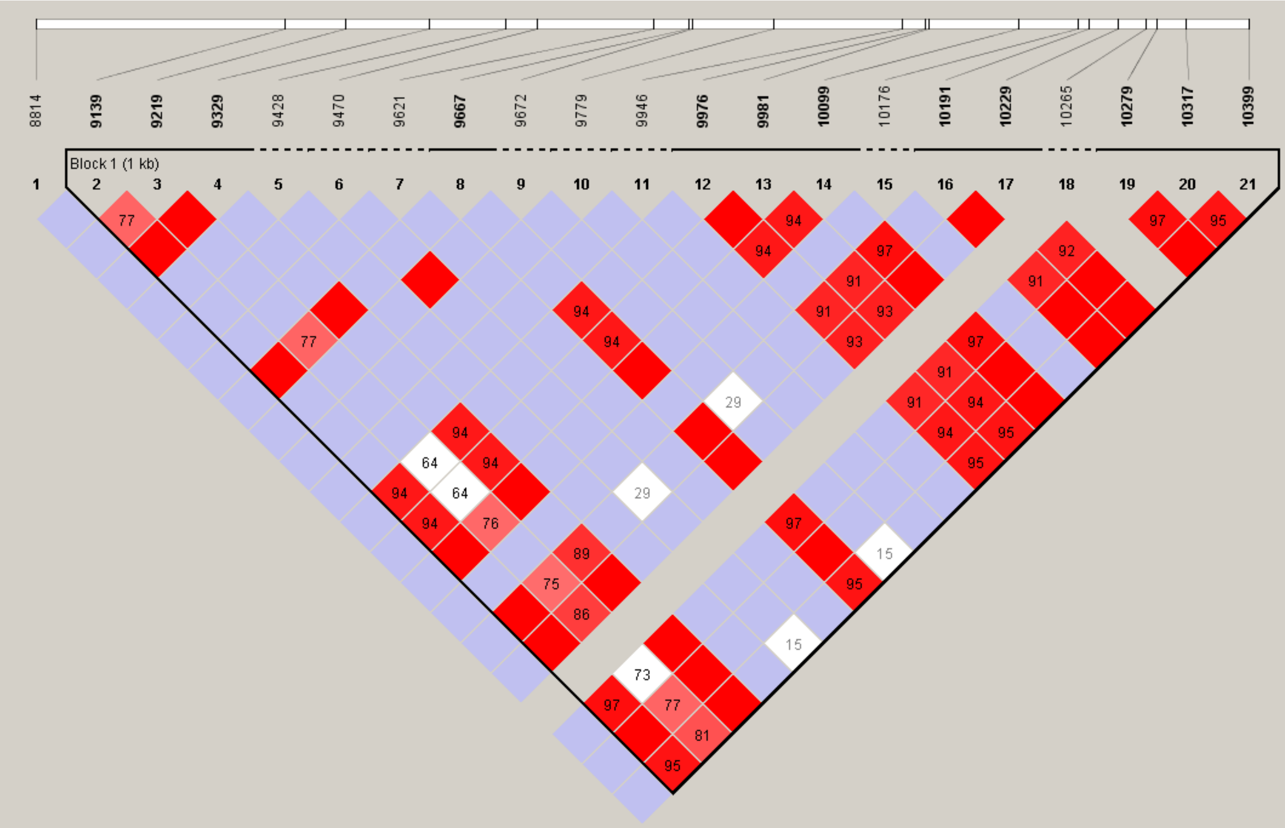


Supplementary Figure 5

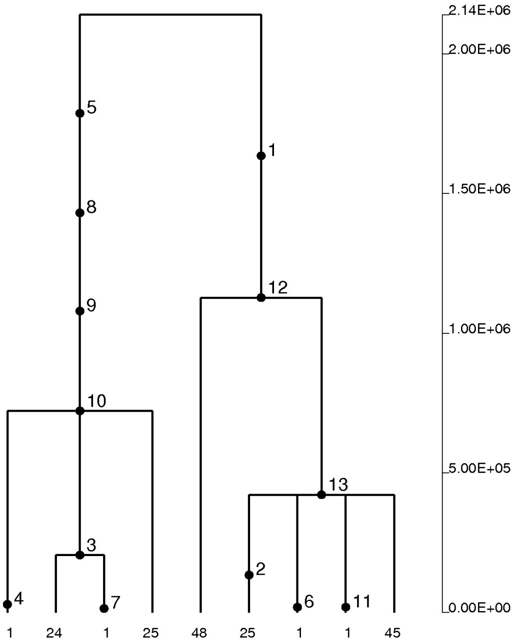
(A) Linkage disequilibrium (LD) blocks for the *CD151* gene region we analyzed.

(B) Estimated tree for the *CD151* gene region corresponding to the large LD block. Mutations are represented as black dots and named for their physical position along the region. The absolute frequency of each haplotype is also reported.

A.



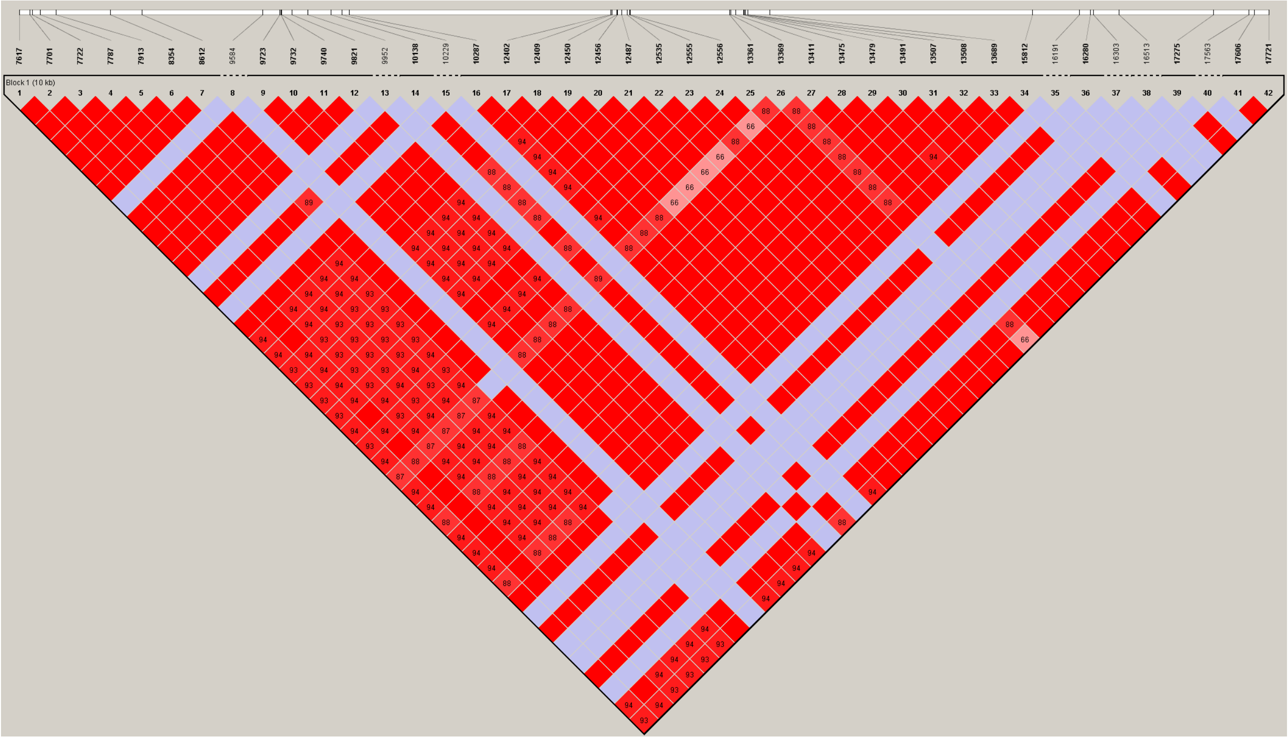
B.



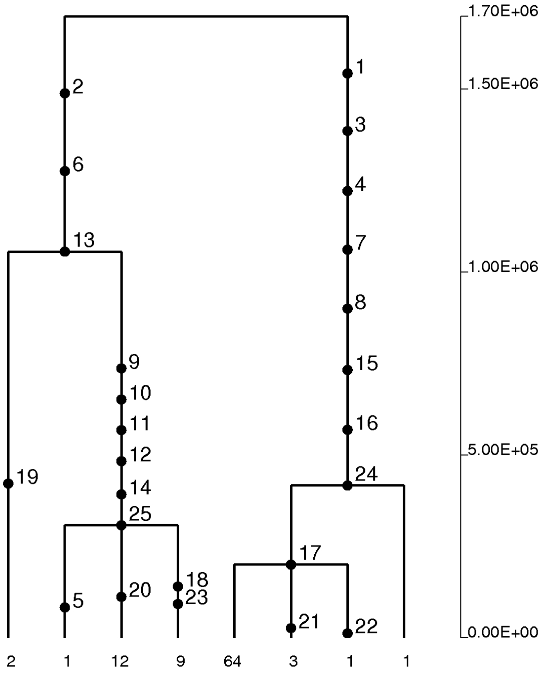
Supplementary Figure 6

LD blocks (A) and estimated tree (B) for the *FUT2* promoter region.

A.



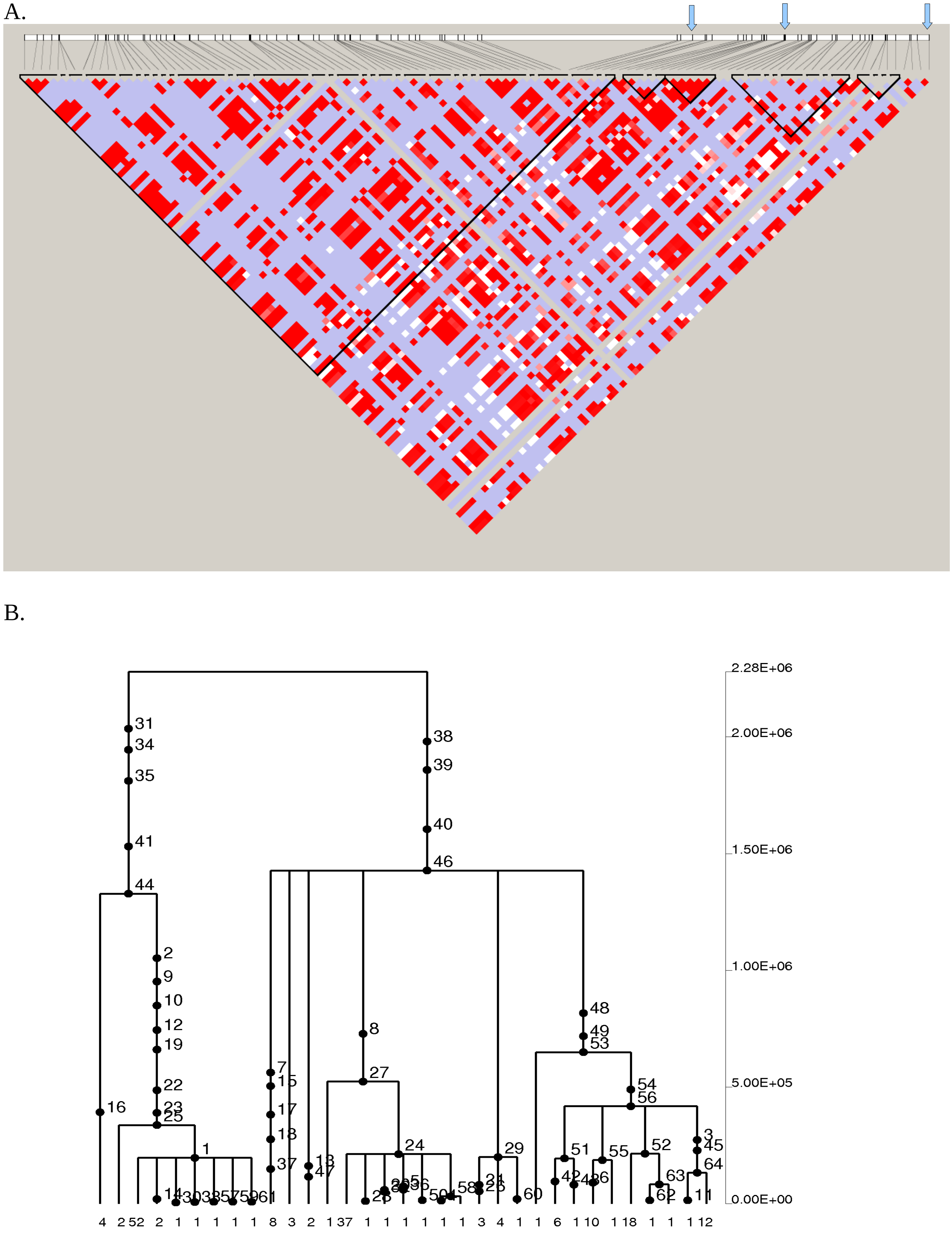
B.



Supplementary Figure 7

(A) Linkage disequilibrium blocks for the *SLC14A1* gene region we analyzed. Arrows indicate the following SNPs (from left to right): rs692899 (correlates with pathogens), rs10853535 (correlates with pathogens) and rs1058396 (Ans280Asp).

(B) Estimated tree for the *SLC14A1* gene region corresponding to the larger LD block.

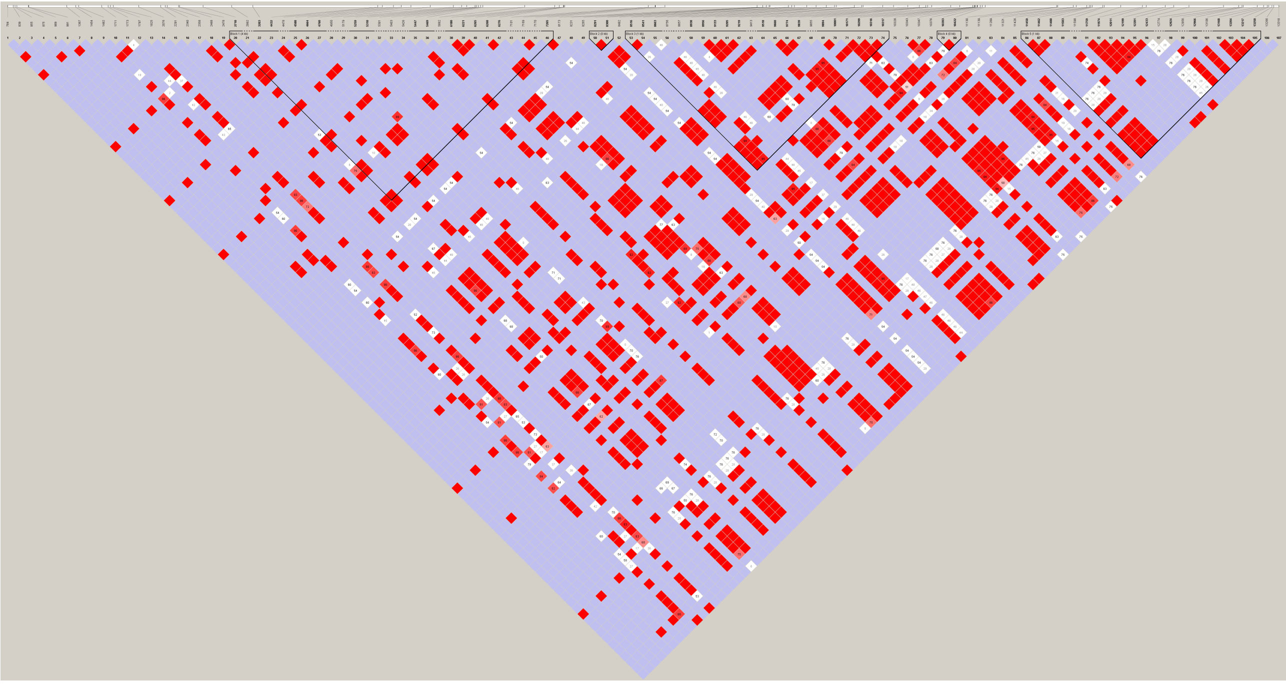


Supplementary Figure 8

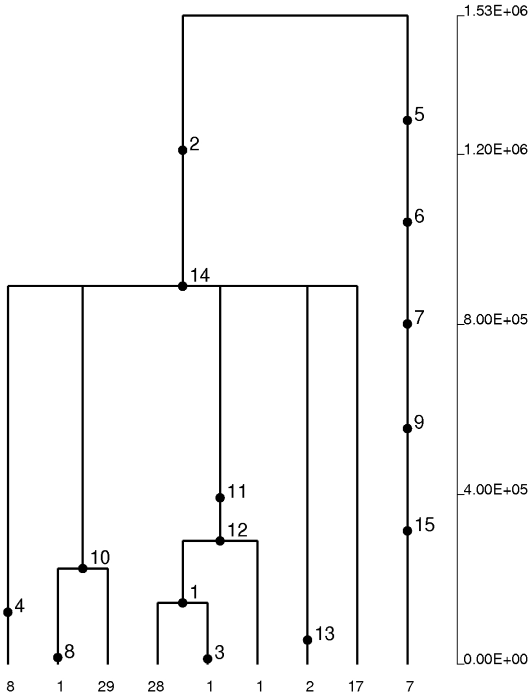
(A) LD blocks for the *BSG* gene.

(B) Estimated tree for the *BSG* gene region corresponding to the second LD block.

A.



B.



Supplementary table 1.

θ_W and π values corresponding to the 2.5th and 97.5th percentile in the distribution of 5 kb windows deriving from NIEHS genes.

	θ ($\times 10^{-4}$)		**π ($\times 10^{-4}$)**	
	2.5th	*97.5th*	*2.5th*	*97.5th*
YRI	2.14	18.61	1.17	19.89
AA	3.53	18.17	2.04	19.06
EA	1.38	14.25	0.65	18.36
AS	1.35	13.63	0.35	17.21

Supplementary Table 2.

Percentile ranks for selected BGA gene regions.

Gene	**Pop.**	**Tajima's D**	**Fu & Li's D***	**Fu & Li's F***
		rank	rank	rank
CD55	AA	.88	.42	.58
	EA	.80	.71	.78
CD151	YRI	.97	.81	.92
	AA	>.99	.86	.98
	EA	.85	.66	.77
FUT2 pm	YRI	>.99	.95	>.99
	EA	.10	.20	.16
FUT2 cds	YRI	.98	.46	.71
	EA	.95	.69	.87
SLC14A1	YRI	.61	.81	.75
	AA	.91	.99	.98
	EA	.94	.78	.91
BSG	YRI	.81	.83	.81
	EA	.13	.059	.059

Note: percentile ranks were calculated in 5 kb windows deriving from SeattleSNPs genes resequenced across panel 1 or panel 2 (see methods in main text for details).

Supplementary table 3.

Percentile rank of Tajima's D statistics for MLHKA control genes.

Gene	Tajima's D *rank*	Gene	Tajima's D *rank*
VNN3		TUBA1	
YRI	0.1722689	YRI	0.7857143
AA	0.7647059	AA	0.7352941
EA	0.5210084	EA	0.5672269
AS	0.8277311	AS	0.3235294
PLA2G2D		MT3	
YRI	0.7605042	YRI	0.684874
AA	0.6302521	AA	0.407563
EA	0.2899160	EA	0.4453782
AS	0.4327731	AS	0.8445378
MB		NUDT1	
YRI	0.2605042	YRI	0.6386555
AA	0.512605	AA	0.6890756
EA	0.8865546	EA	0.512605
AS	0.07142857	AS	0.7268908
MAD2L2		PRDX5	
YRI	0.2478992	YRI	0.7268908
AA	0.3823529	AA	0.302521
EA	0.8277311	EA	0.6554622
AS	0.789916	AS	0.3445378
HRAS		RETN	
YRI	0.4243697	YRI	0.7983193
AA	0.8529412	AA	0.7815126
EA	0.3865546	EA	0.8991597
AS	0.7857143	AS	0.8067227
CYP17A1		JUND	
YRI	0.09243697	YRI	0.9033613
AA	0.4705882	AA	0.802521
EA	0.9705882	EA	0.9117647
AS	0.8991597	AS	0.920168
ATOX1			
YRI	0.197479		
AA	0.4411765		
EA	0.3697479		
AS	0.05882353		
BNIP3			
YRI	0.1008403		
AA	0.2647059		
EA	0.4117647		
AS	0.802521		
CDC20			
YRI	0.5420168		
AA	0.4327731		
EA	0.2815126		
AS	0.1008403		
NGB			
YRI	0.3361345		
AA	0.1596639		
EA	0.3109244		
AS	0.4831933		