Figure legends

Figure 1: Associations of SNPs with gene expression of SPRED2 on chromosome 2. Panels contrast the results obtained using phase I HapMap SNPs and phase II HapMap SNPs. Coordinates are in NCBI Build 35. Blue arrows represent the location (not to scale) and direction of transcription of the associated gene.

Figure 2: Comparison of detected *cis* associations between single and multi-population analysis. **(A)** Numbers of genes with significant *cis*- associations as uncovered by single and multi-population analysis and proportion overlap of associations across the two methodologies. **(B)** Associations of SNPs of the phase II HapMap with gene expression of SGPP2 on chromosome 2. Coordinates are in NCBI Build 35. Panels show results of 4-population multi-population analysis, and individual population analysis for CEU, CHB, JPT, and YRI. Blue arrows represent the location (not to scale) and direction of transcription of the associated gene. In this case the SNP was not rare in any of the populations (MAF was between 0.08 and 0.44) but the effect was small ($R^2 = 0.25$ and slope = 0.25) so it could only be detected when we pooled the populations increasing the sample size.
(C) Comparison of the adjusted R^2 values (proportion of the variance in expression explained by the linear relationship between genotype and phenotype) of *cis* significant associations obtained from single and multi-population analysis (0.001 permutation threshold).

Figure 3: Comparison of the direction of shared SNP-gene allelic effects across all pairs of populations (0.001 permutation threshold). White panels indicate effects in the same direction.

Figure 4: Statistical significance, adjusted R^2 of the association (proportion of the expression variance explained by the linear relationship between genotype and phenotype), and absolute value of the slope of the linear regression, as a function of distance from the transcription start site, of the most significantly associated SNP per gene in each of the 4 populations (order CEU, CHB, JPT and YRI) and the pooled sample of all 4 populations.

Supplementary Figure legends:

Figure S1: The triangle indicates the number of genes that have significantly different median value for each population comparison. ASN indicates the pooled CHB+JPT population. The external lines and numbers indicate genes that are differentiated only for that given pair. The internal lines and numbers indicate genes differentiated in multiple comparisons. For example, there 1320 genes that are only different between CEU and YRI, 799 genes are different between YRI-CEU and YRI-ASN but not CEU-ASN and 342 genes are different in all 3 populations. Please note that the CEU population is more differentiated than the other two possibly due to the age of the cell lines

Figure S2: Correlation of gene expression variance for identical probes on the whole-genome and custom arrays, plotted against the p-value of the correlation.

Figure S3: Overlap of genes with significant *cis*- association as detected by the Phase I vs. Phase II HapMap.

Figure S4: The plot shows the maximum –logp-value (Y-axis) for the cis analysis (distance < 1Mb) for each of the 14072 genes relative to estimated heritability (slope of the mid-parent-offspring regression; X-axis). Note that genes with high –logp-values tend to have high heritability estimates. Negative values of estimated heritability are expected when there is no genetic effect on the expression phenotype.

Figure S5: Difference in expected heterozygosity (2pq) of SNPs in shared and unshared associations. A. Same SNP and gene significantly associated in 2 populations: $2pq_{population1} - 2pq_{population2}$ B. Same gene, but not same SNP significantly associated in 2 populations: $2pq_{associated\ pop} - 2pq_{not\ associated\ pop}$ C. Different genes and different SNPs significantly associated in 2 populations: $2pq_{associated\ pop} - 2pq_{not\ associated\ pop}$.

Figure S6: This figure shows a specific example where the outlier high value in the CC genotype category drives a very significant nominal p-value for the linear regression (-logpval = 8.01), while the Spearman rank correlation p-value is much higher ($P > 0.01$).

Figure S1

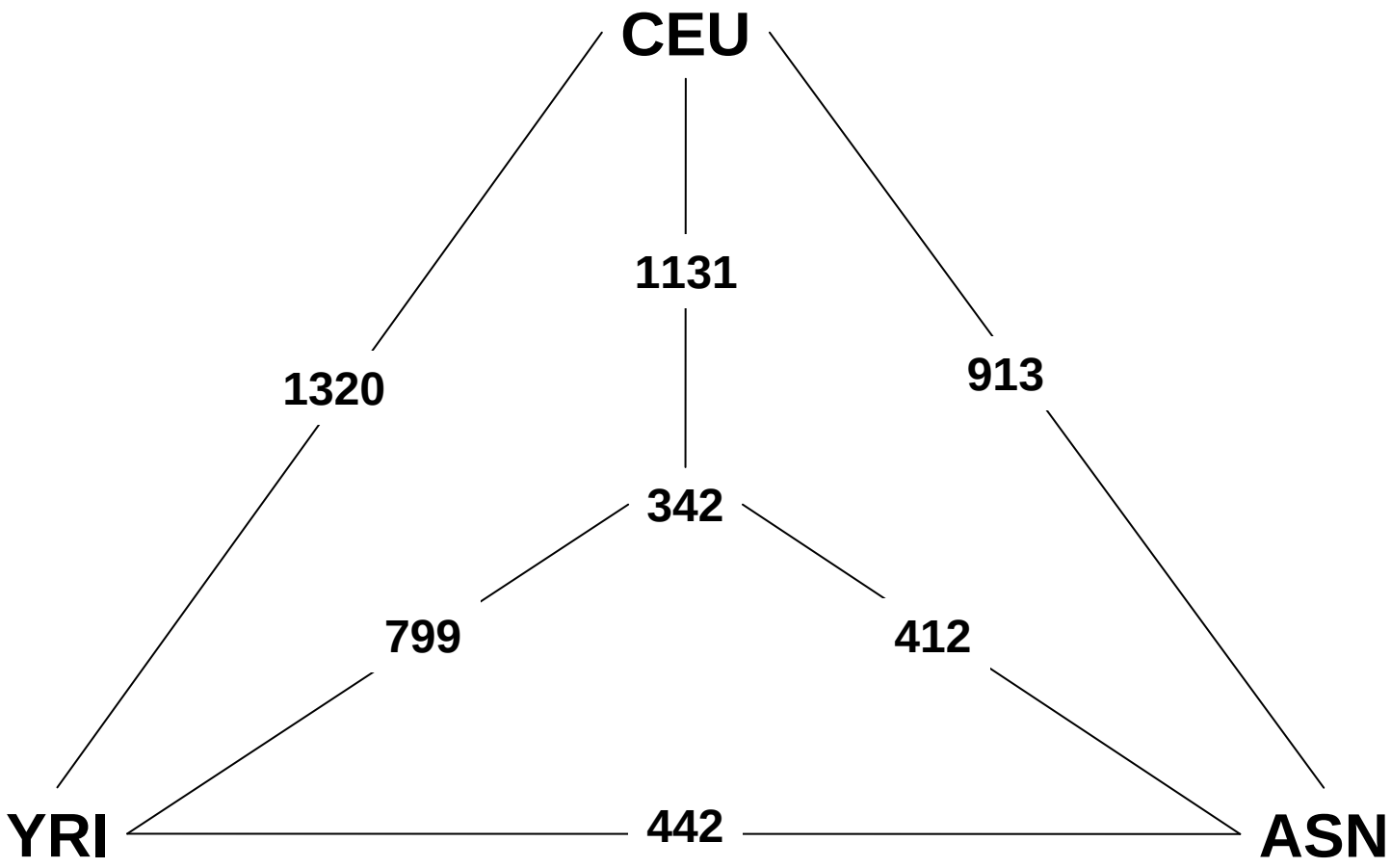


Figure S2

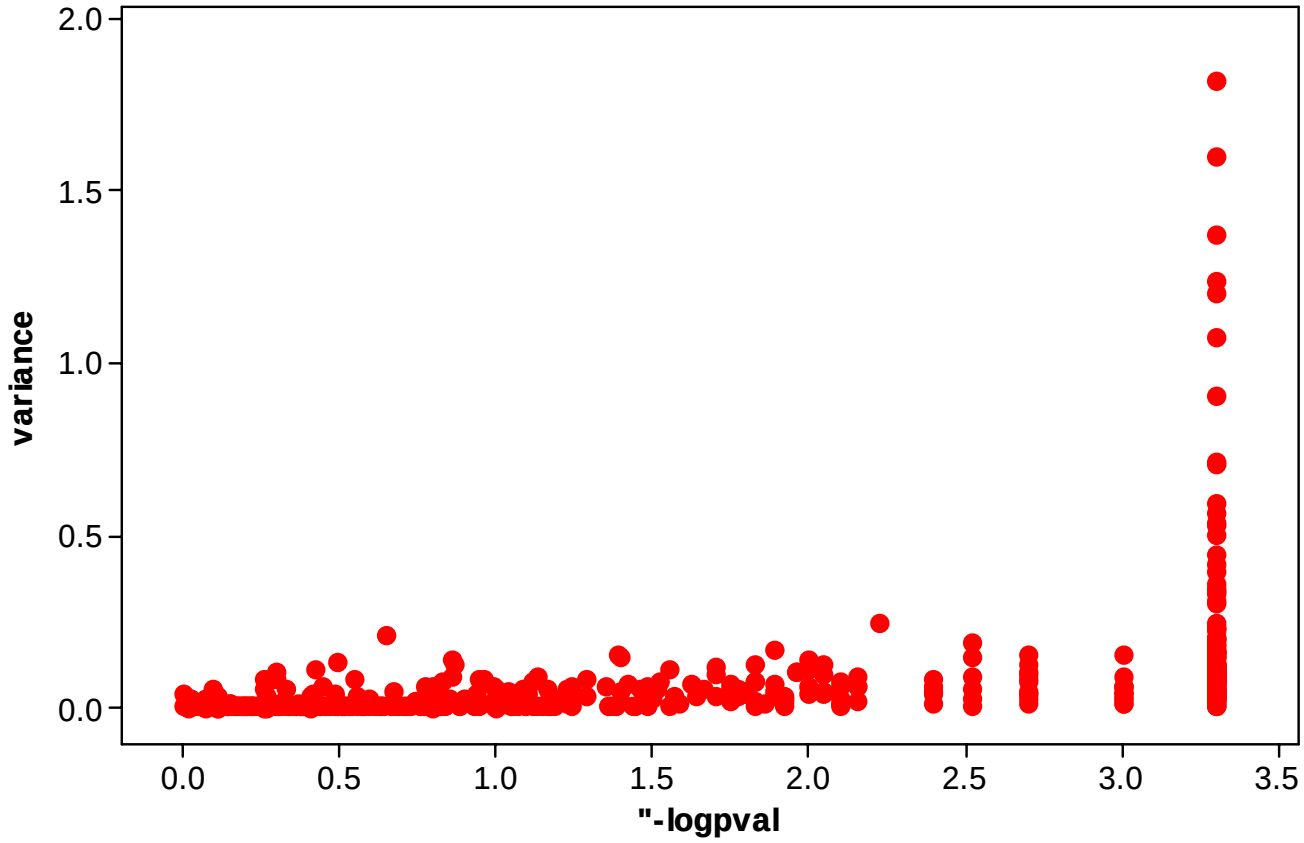


Figure S3

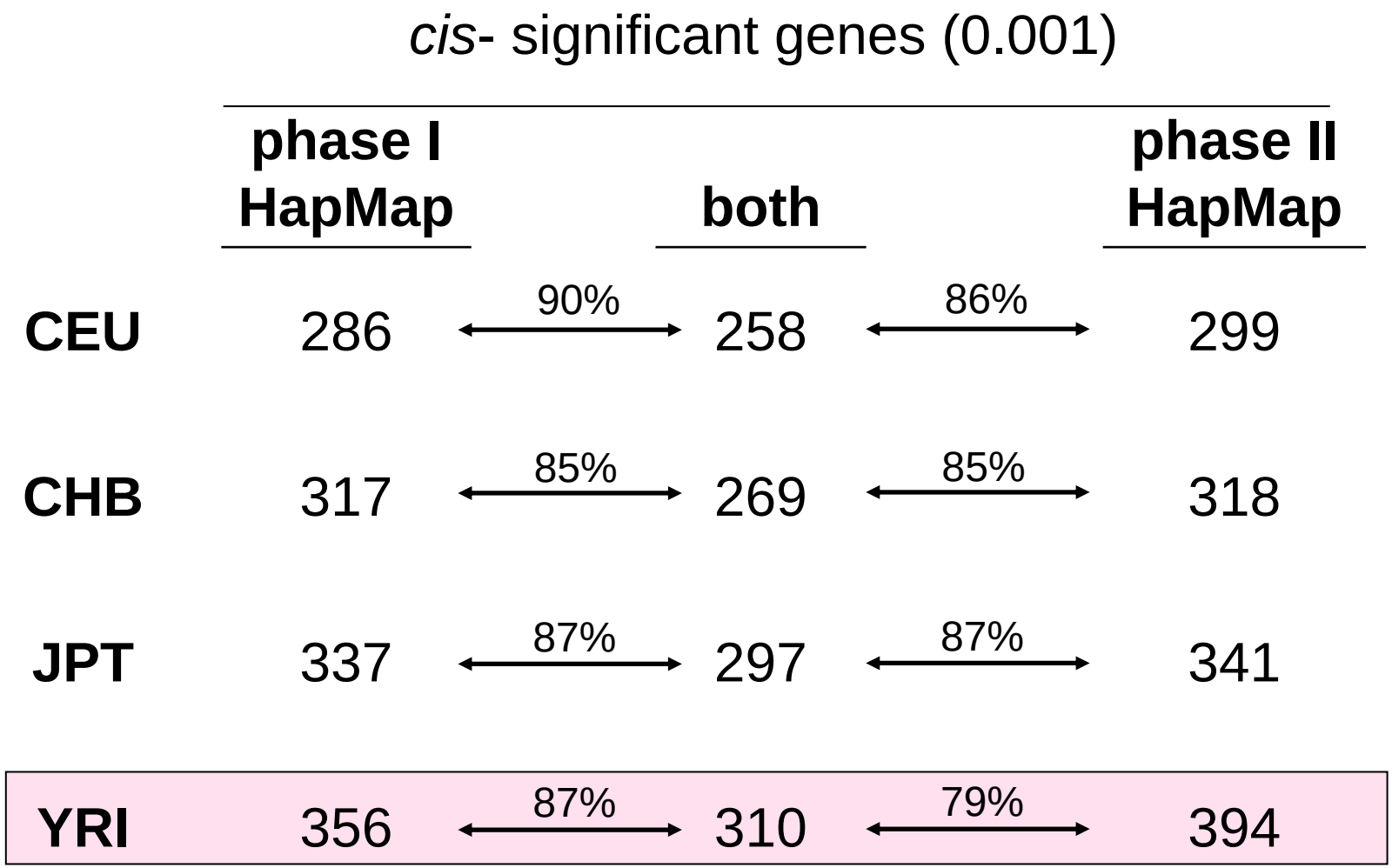
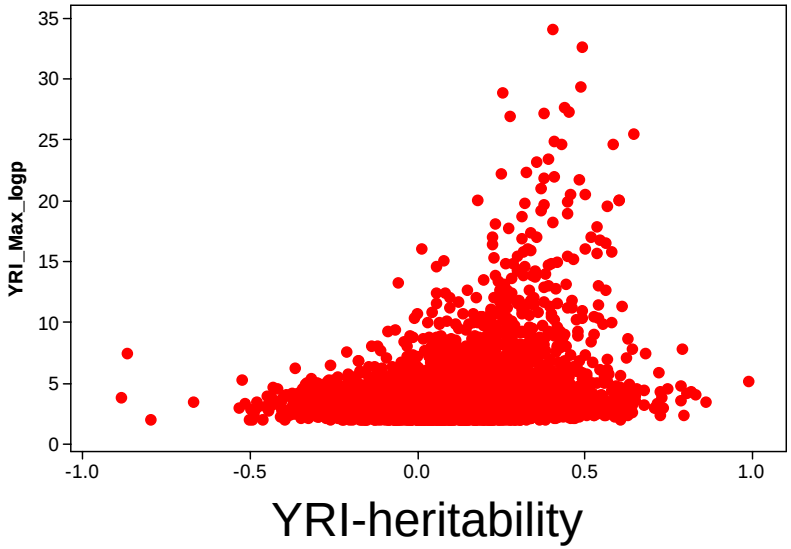
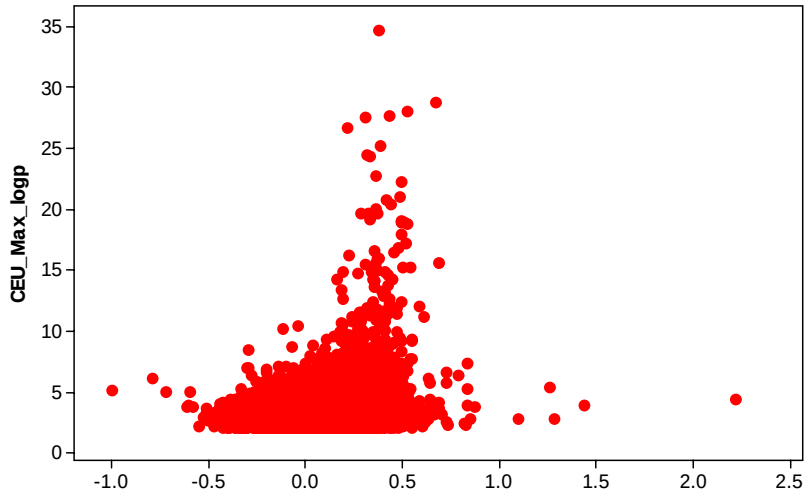

cis- significant genes (0.001)
phase I HapMap
both
phase II HapMap
CEU
286
90%
258
86%
299
CHB
317
85%
269
85%
318
JPT
337
87%
297
87%
341
YRI
356
87%
310
79%
394

Figure S4



YRI-heritability



CEU-heritability

Figure S5

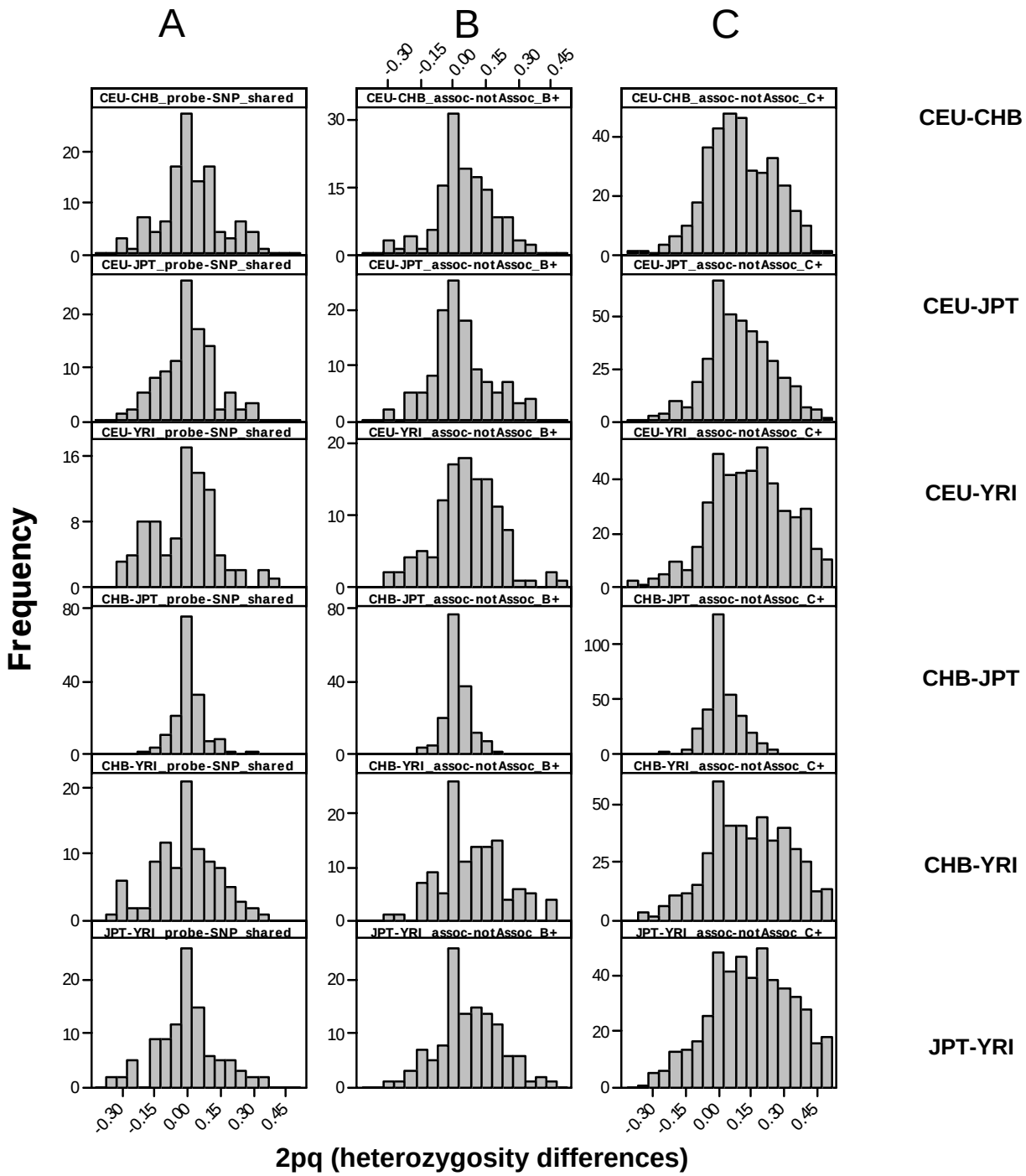
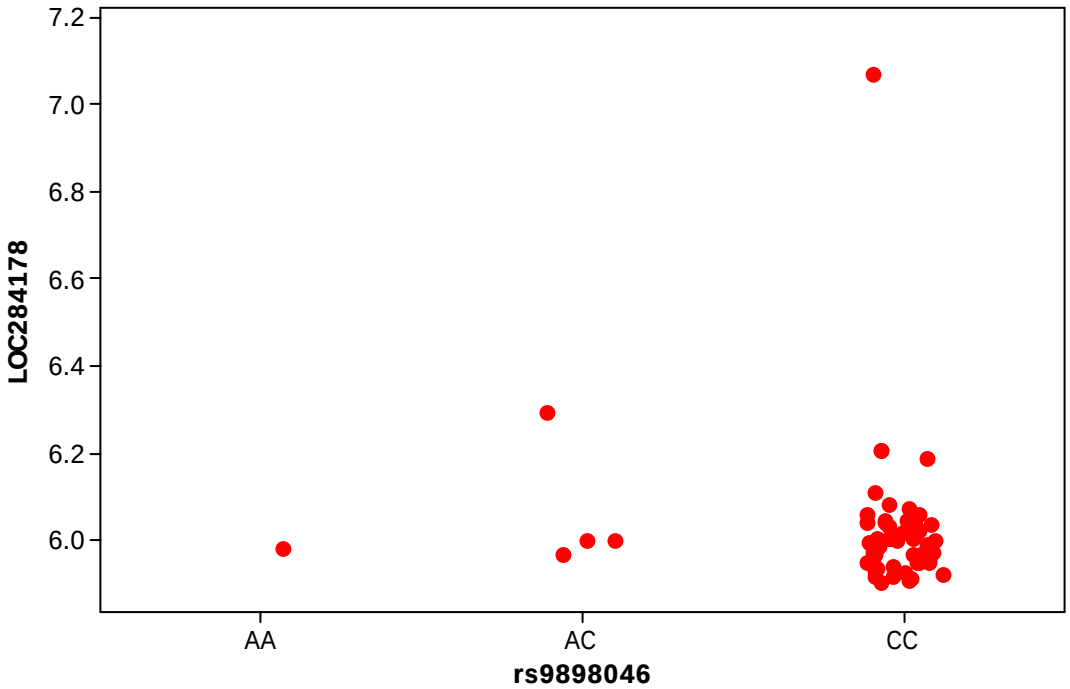

A
B
C
CEU-CHB_probe-SNP_shared
CEU-CHB_assoc-notAssoc_B+
CEU-CHB_assoc-notAssoc_C+
CEU-JPT_probe-SNP_shared
CEU-JPT_assoc-notAssoc_B+
CEU-JPT_assoc-notAssoc_C+
CEU-YRI_probe-SNP_shared
CEU-YRI_assoc-notAssoc_B+
CEU-YRI_assoc-notAssoc_C+
CHB-JPT_probe-SNP_shared
CHB-JPT_assoc-notAssoc_B+
CHB-JPT_assoc-notAssoc_C+
CHB-YRI_probe-SNP_shared
CHB-YRI_assoc-notAssoc_B+
CHB-YRI_assoc-notAssoc_C+
JPT-YRI_probe-SNP_shared
JPT-YRI_assoc-notAssoc_B+
JPT-YRI_assoc-notAssoc_C+
CEU-CHB
CEU-JPT
CEU-YRI
CHB-JPT
CHB-YRI
JPT-YRI
Frequency
-0.30
-0.15
0.00
0.15
0.30
0.45
2pq (heterozygosity differences)

Figure S6



LR –logp-value = 8.0121
SRC –logp-value > 0.01

Table S1

Number of genes with sigificant cis- associations shared by population pairs

	CEU	**CHB**	**JPT**	**YRI**	
CEU		133	122	119	299
CHB	230		172	123	318
JPT	240	291		125	341
YRI	219	203	228		394
	606	634	679	742	

Above the diagonal, 0.001 significant.
Below the diagonal, 0.01 significant.

Extra column and row: 0.001 and 0.01 within-population significant gene counts.

Table S2

Number and source category of SNPs used in trans analysis

	CEU	CHB	JPT	YRI
cis- associated (rSNPs)	13221	13133	13191	13375
Non-synonymous	9904	9383	9378	10727
Splicing	1756	1585	1594	1950
miRNA	34	34	32	37
Non-redundant[a]	24635	23854	23907	25797

[a]Note: all SNPs are > 0.05 frequency in the respective population.