

A genome-wide association study on photic sneeze reflex in the Chinese population

Mengqiao Wang^{#*}, Xinghan Sun[#], Yang Shi, Xiaojun Song, Hao Mi

[#] these authors contributed equally

^{*} corresponding author

Supplementary Figures

Figure S1

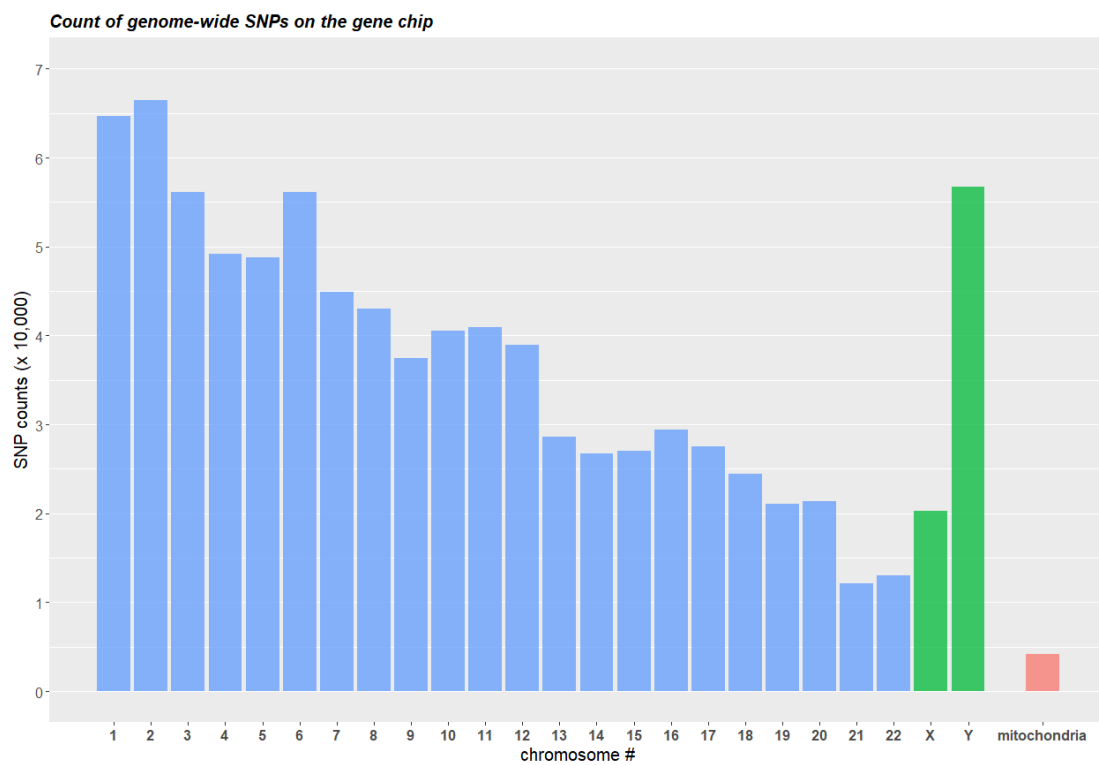
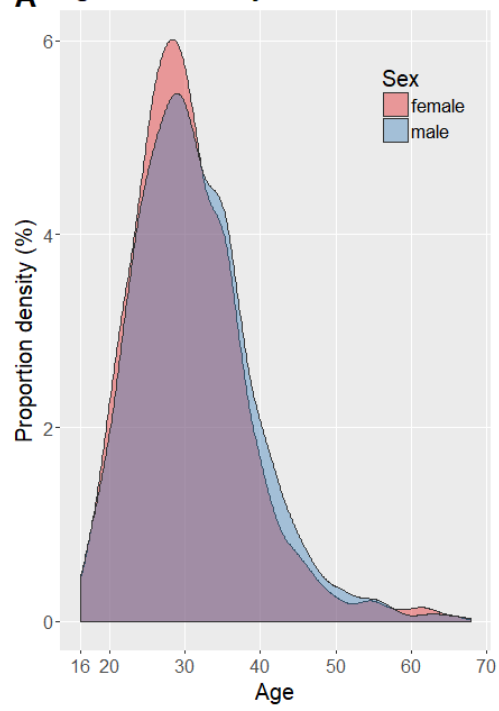


Figure S2

A *Age distribution by sex*



B *Age distribution by phenotype*

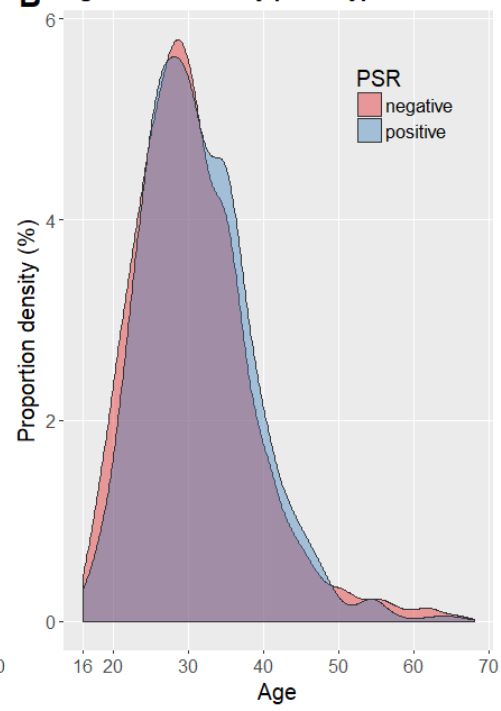


Figure S3

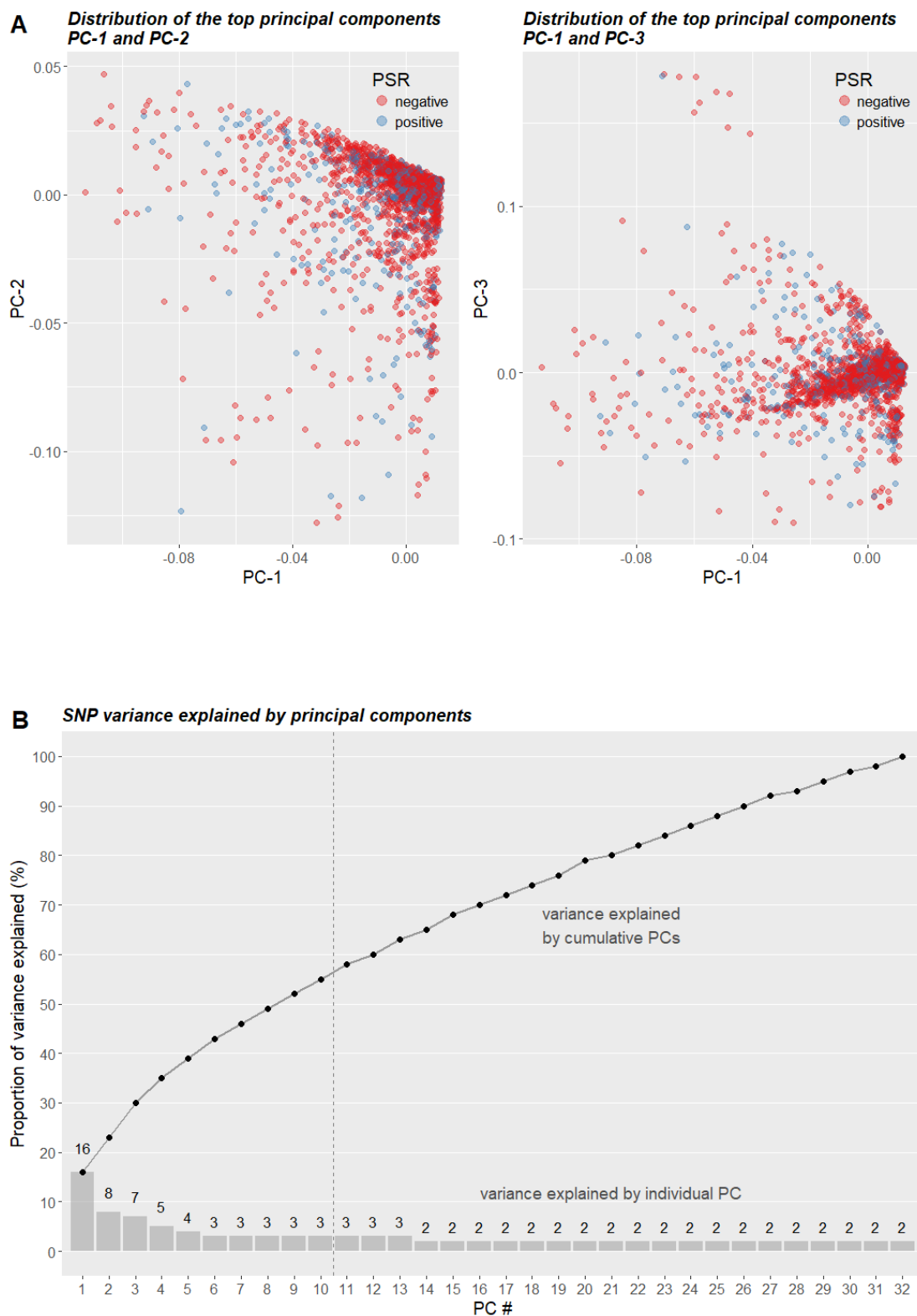
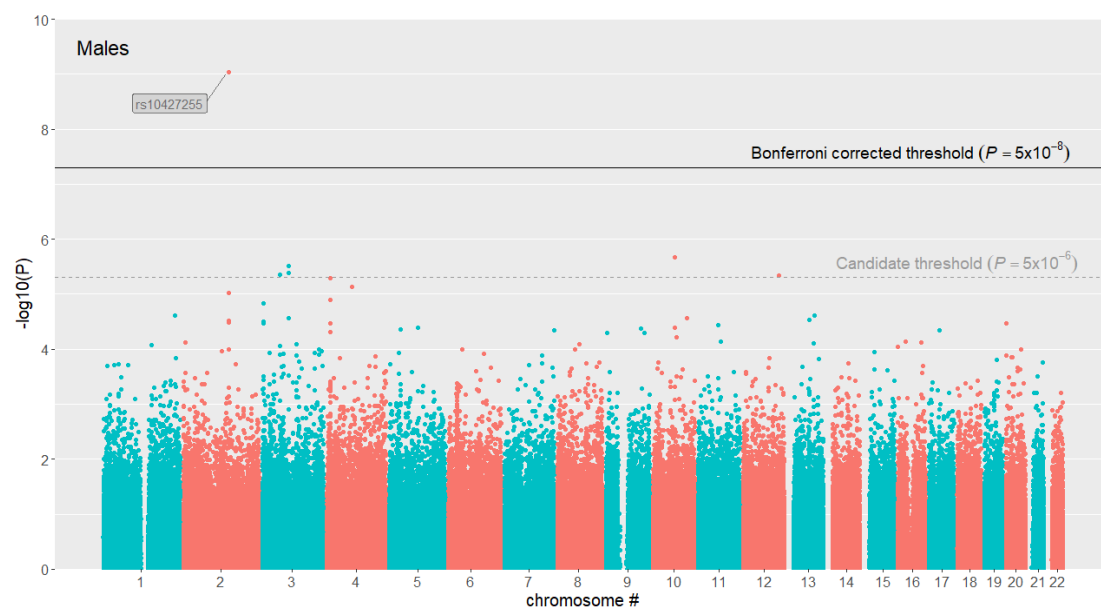


Figure S4

A



B

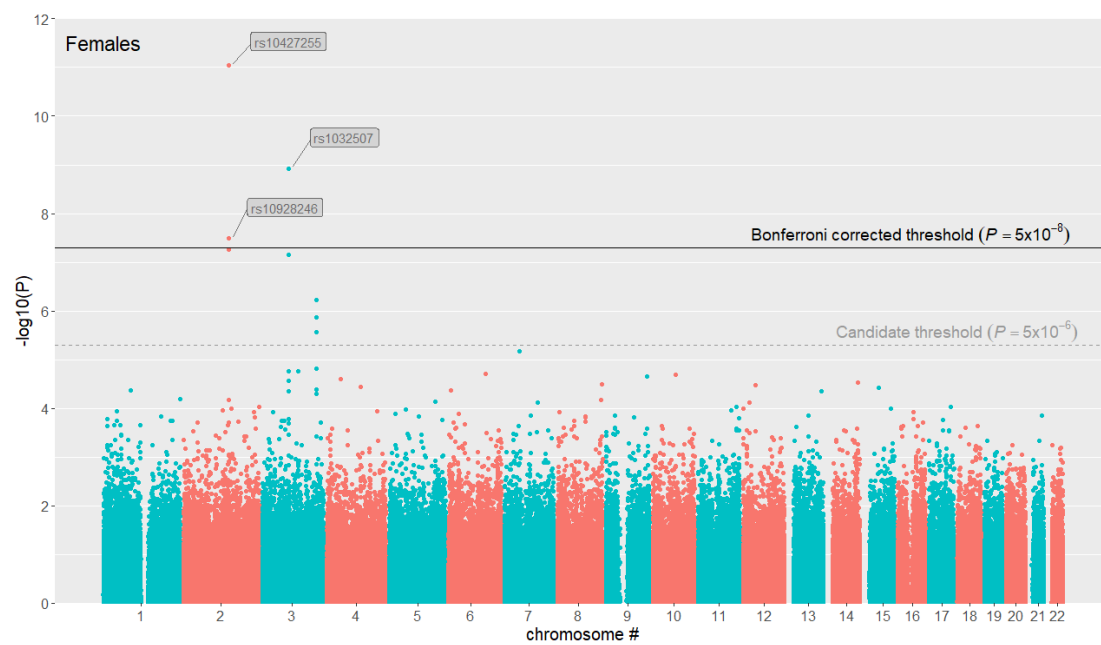
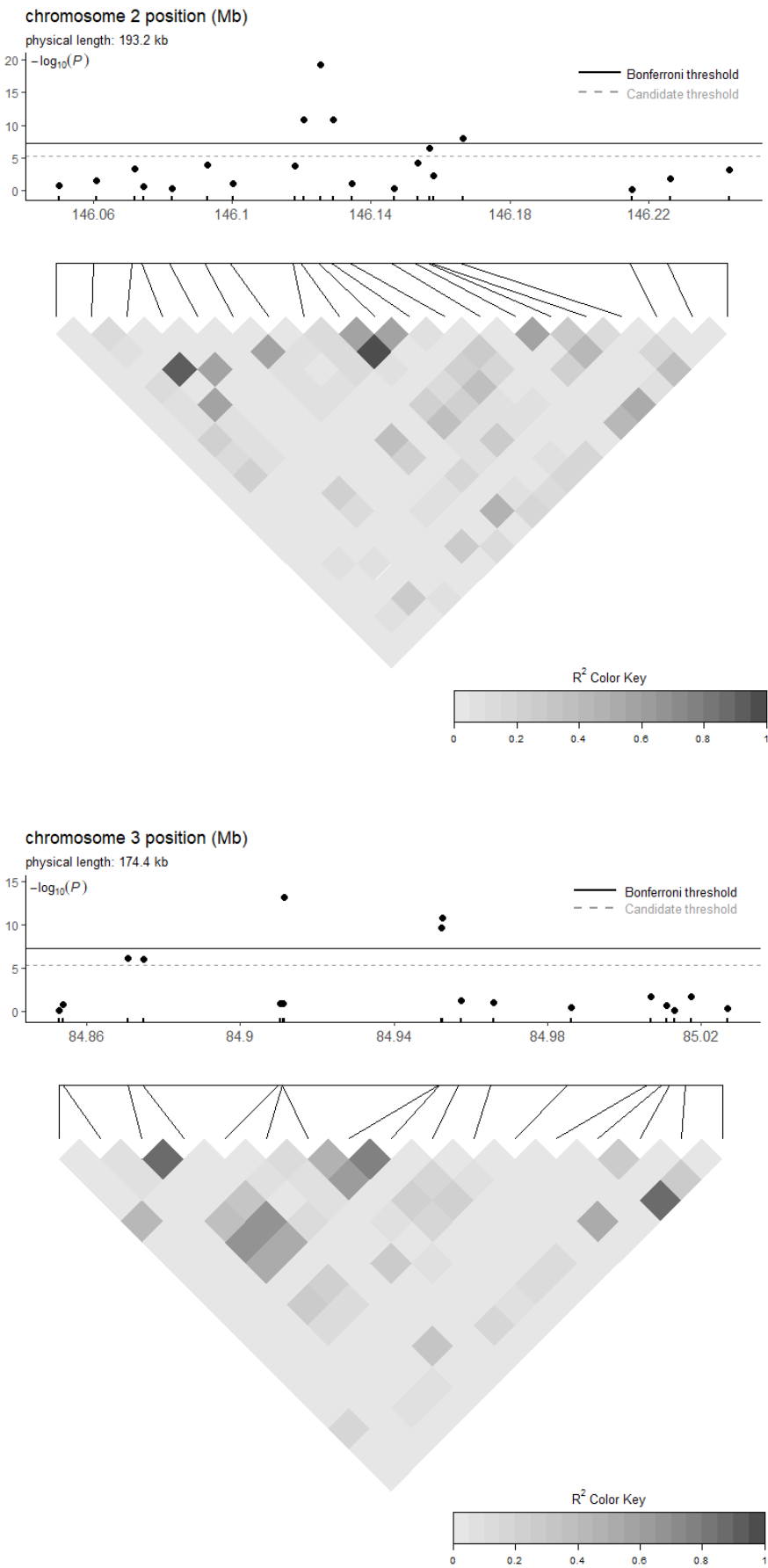


Figure S5



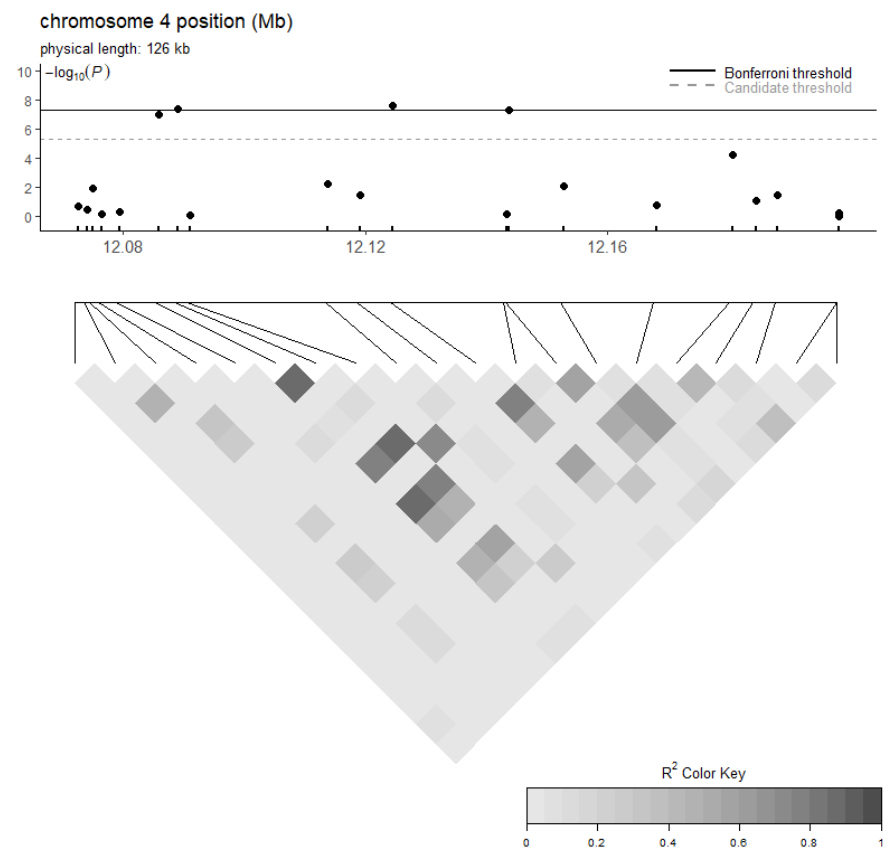


Figure S6

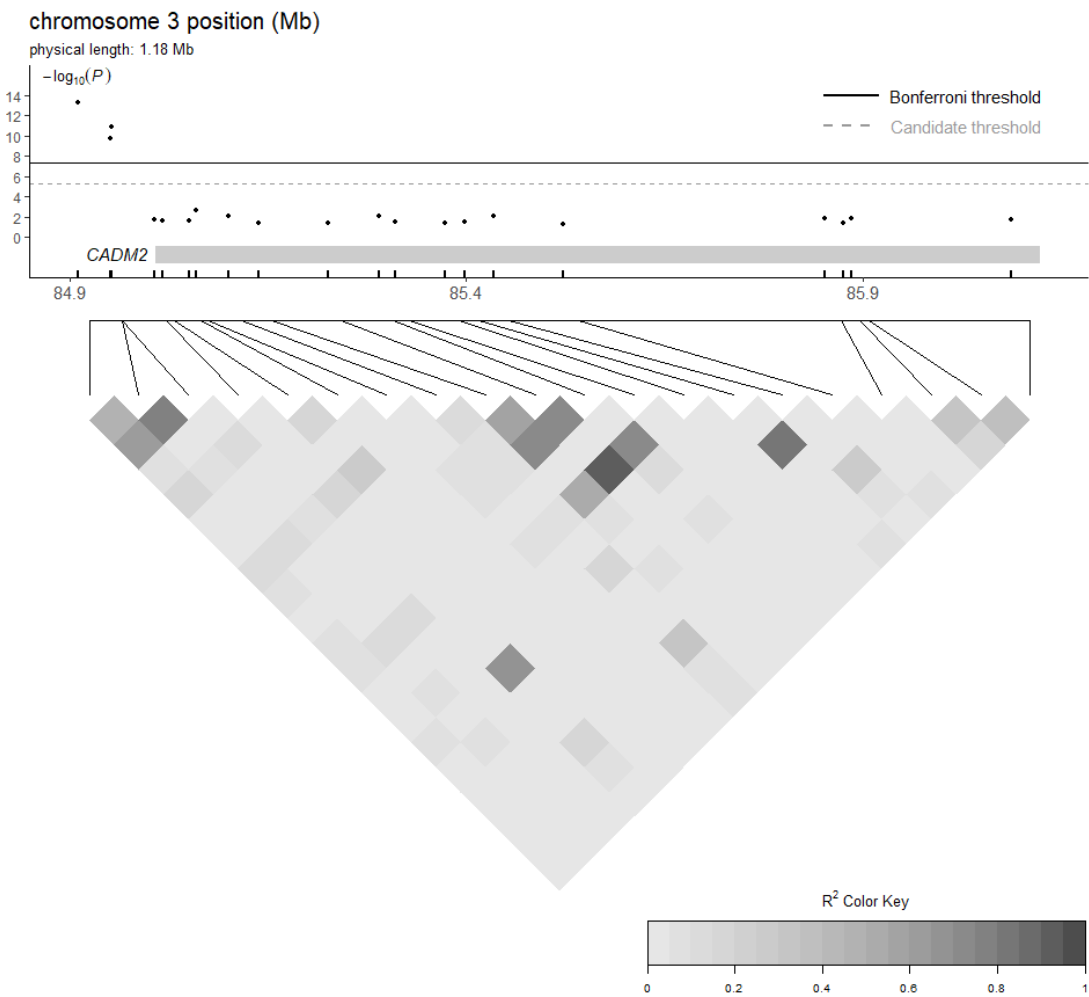


Figure S7

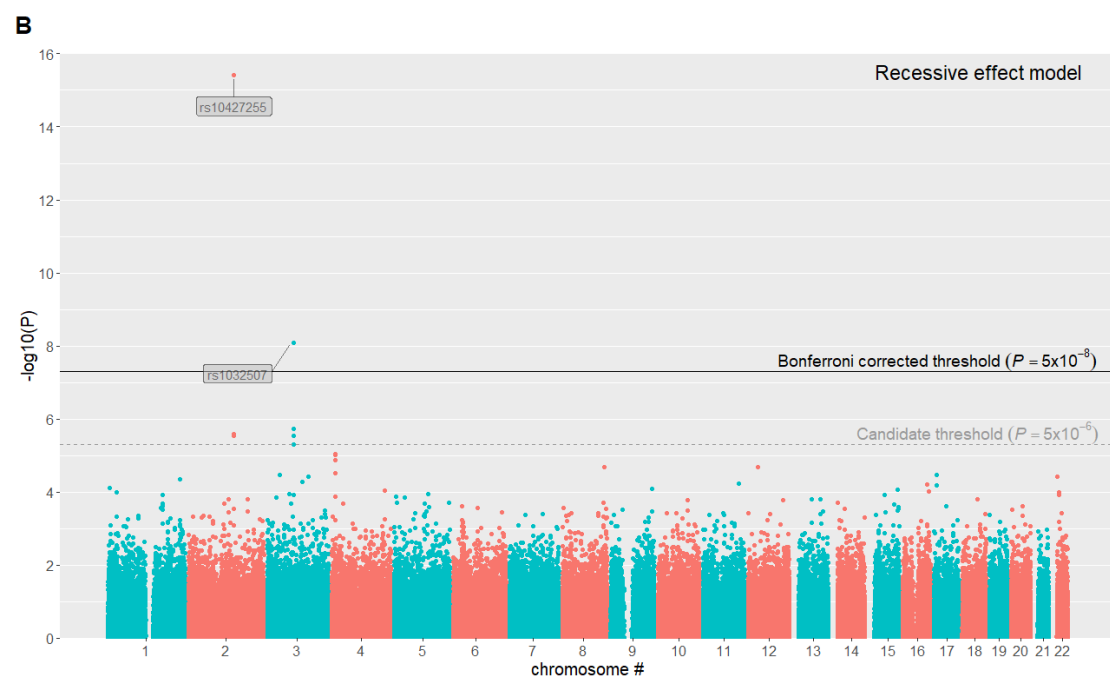
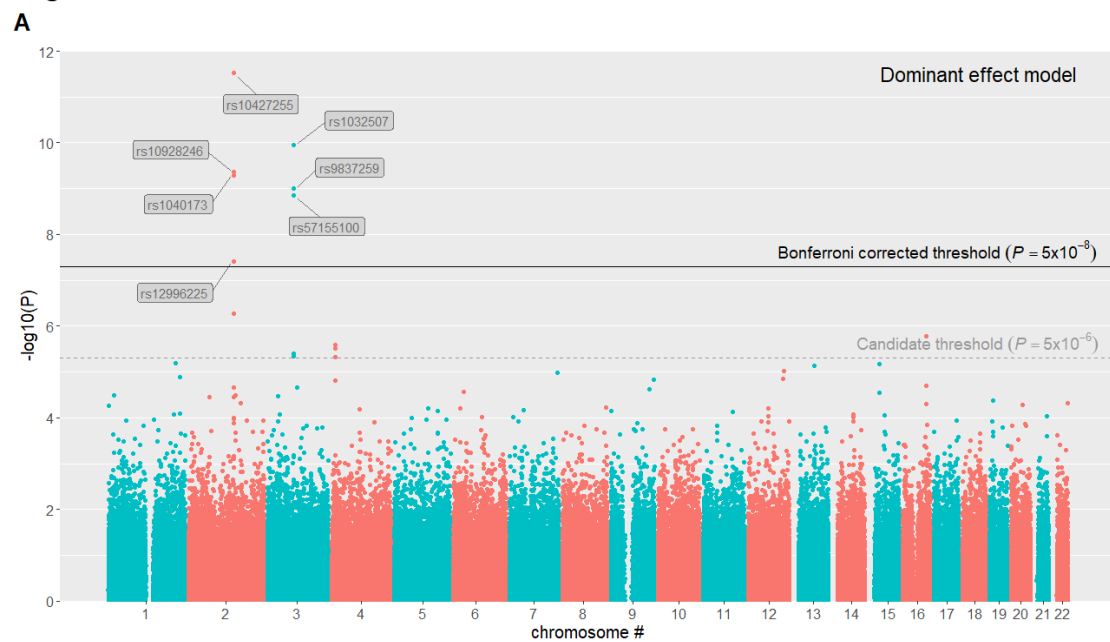
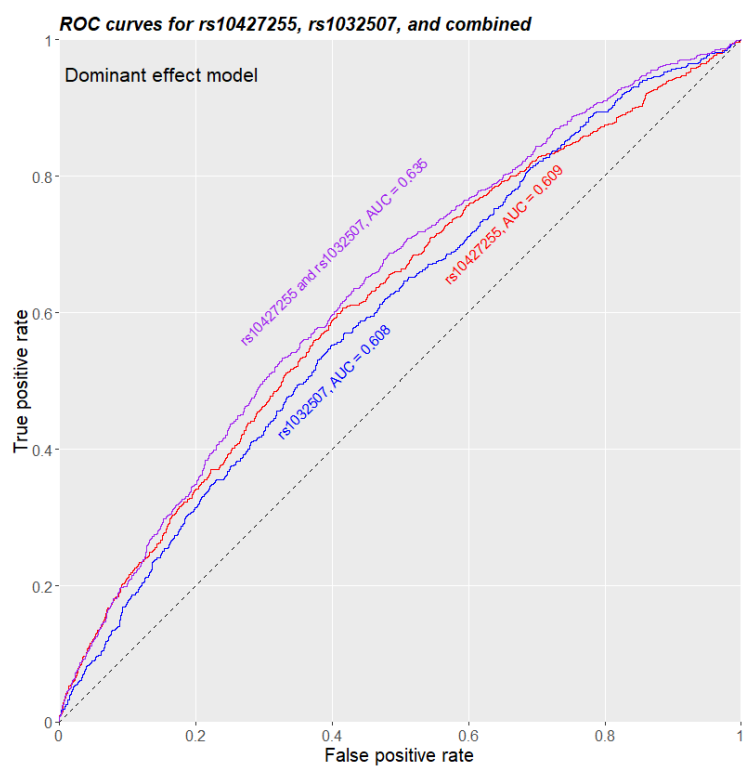
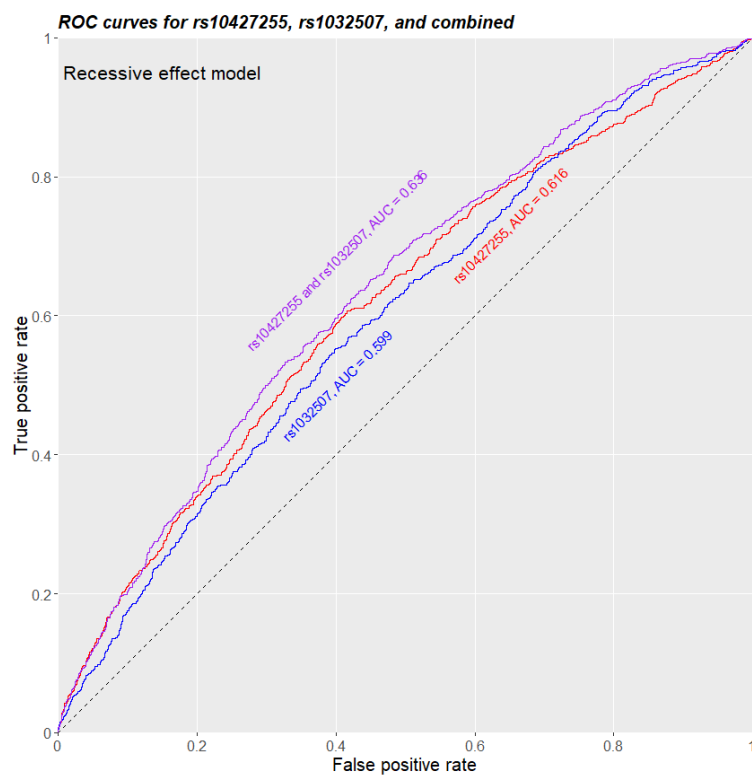


Figure S8

A



B



Supplementary Figure Legends

Figure S1. Distribution for the count of SNPs on the gene chip used for genotyping.

Figure S2. Age distribution by sex (A) or by phenotype (B) of sample individuals.

Figure S3. Principal component analysis of SNPs. PC-1, PC-2, and PC-3 represented the top 3 principal components (PCs) explaining the highest portion of genetic variation. A. Scatter plots of PC-1 vs. PC-2, and PC-1 vs. PC-3, displayed the clustering of the surveyed individuals as a homogenous sample. B. Individual and cumulative proportion of variance explained by the corresponding PCs. The top 10 PCs cumulatively explained over 55% of total variance.

Figure S4. Sex-separated analysis: Manhattan plot of GWAS based on an additive effect model for males (A) or for females (B). Bonferroni corrected threshold and candidate threshold respectively correspond to 7.30 and 5.30 with regard to $-\log_{10}(P)$.

Figure S5. Zoomed-in regional association plot of 3 hit SNP regions.

Figure S6. Association plot of hit SNPs on chromosome 3 to SNPs within their nearest downstream gene of *CADM2*.

Figure S7. Manhattan plot of GWAS based on the dominant effect model (A) or the recessive effect model (B).

Figure S8. ROC curves for predicting PSR phenotype using rs10427255 alone (red), rs1032507 alone (blue), or rs10427255 and rs1032507 combined (purple), based on the dominant effect model (A) or the recessive effect model (B).

Supplementary Tables

Table S1.

Pre-processing steps to assemble the final dataset for GWAS

Step	Pre-processing	Sample	SNP	Note
1	initial	3519	819427	
2	sex check	3519	819427	check the match between self-reported sex and gene-chip-revealed sex
3	age check	3468	819427	exclusion of 51 individuals with self-reported age outside the [16, 70] range
4	SNPs on 22 autosomes	3468	738243	exclusion of 81184 non-autosome SNPs
5	SNP-level filtering	3468	430822	exclusion of 307421 SNPs due to low call rate (37711, threshold = 0.95) or low minor allele frequency (233609, threshold = 0.01) or both (36101)
6	sample-level filtering	3432	430822	exclusion of 36 individuals due to low call rate (14, threshold = 0.95) or high inbreeding coefficient (20, threshold = 0.1) or both (2)
7	kinship filtering	3417	430822	exclusion of 15 individuals due to cryptic relatedness using identity-by-descent analysis (threshold = 0.1)
8	ancestry filtering	3417	430822	no individual excluded given the largely homogeneous Chinese ancestry in PCA
9	HWE filtering	3417	419093	exclusion of 11729 SNPs due to the rejection of Hardy-Weinberg Equilibrium at adjusted α of 1×10^{-6}