iScience, Volume *24*

Supplemental information

Cell type-specific and cross-population polygenic risk score analyses of *MIR137* gene pathway in schizophrenia

Yin Yao, Wei Guo, Siwei Zhang, Hao Yu, Hao Yan, Hanwen Zhang, Alan R. Sanders, Weihua Yue, and Jubao Duan

Supplementary Figures and Legends

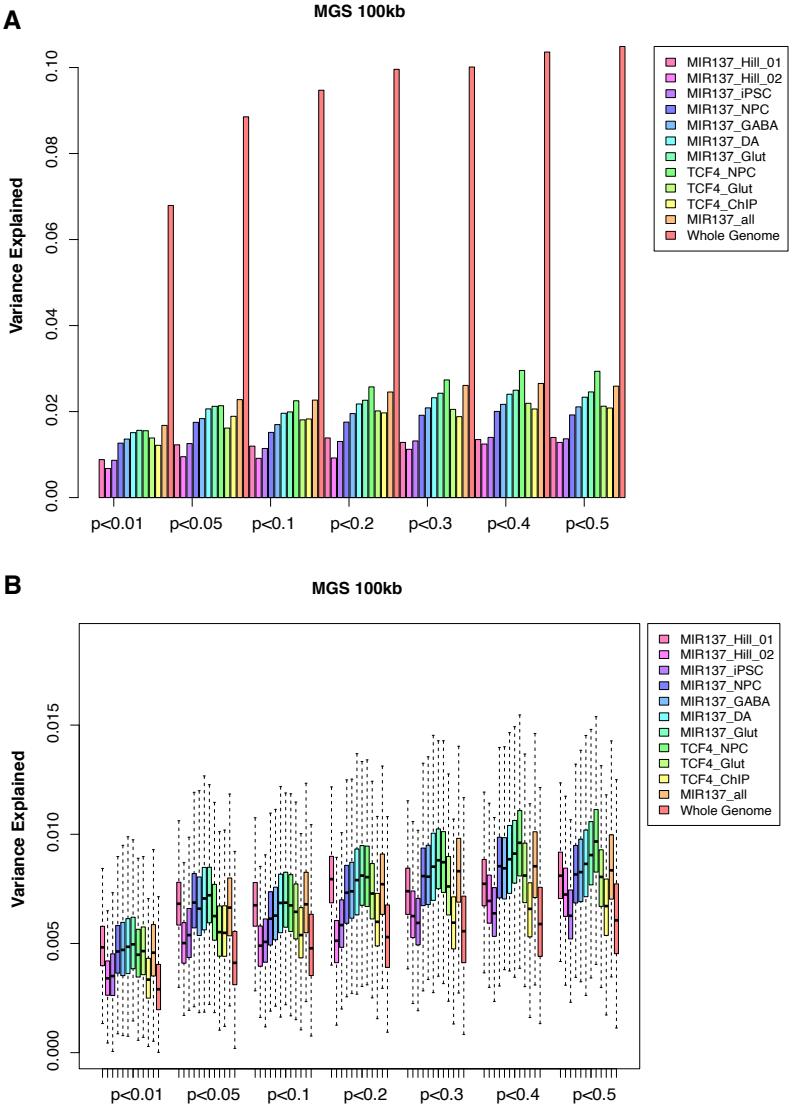

Figure S1. PRS analysis of different SNP lists of *MIR137* and *TCF4* target genes annotated in a 100kb window for the MGS sample. Related to Figure 2. (A) The PRS results from using the LD-pruned GWAS SNPs for each gene list. The variance explained in the target sample is based on risk scores derived from an aggregated sum of weighted SNP risk allele effect sizes estimated from the discovery samples at seven significance thresholds ($P < 0.01, 0.05, 0.1, 0.2,$

0.3, 0.4, and 0.5). The y-axis indicates the percentage of phenotypic variance explained by the PRS (Nagelkerke's pseudo R^2). (B) The permutation (N=1,000) PRS results based on randomly selected 3,000 SNPs for each gene set. Gene-sets in (A) and (B) were described in Table 1.

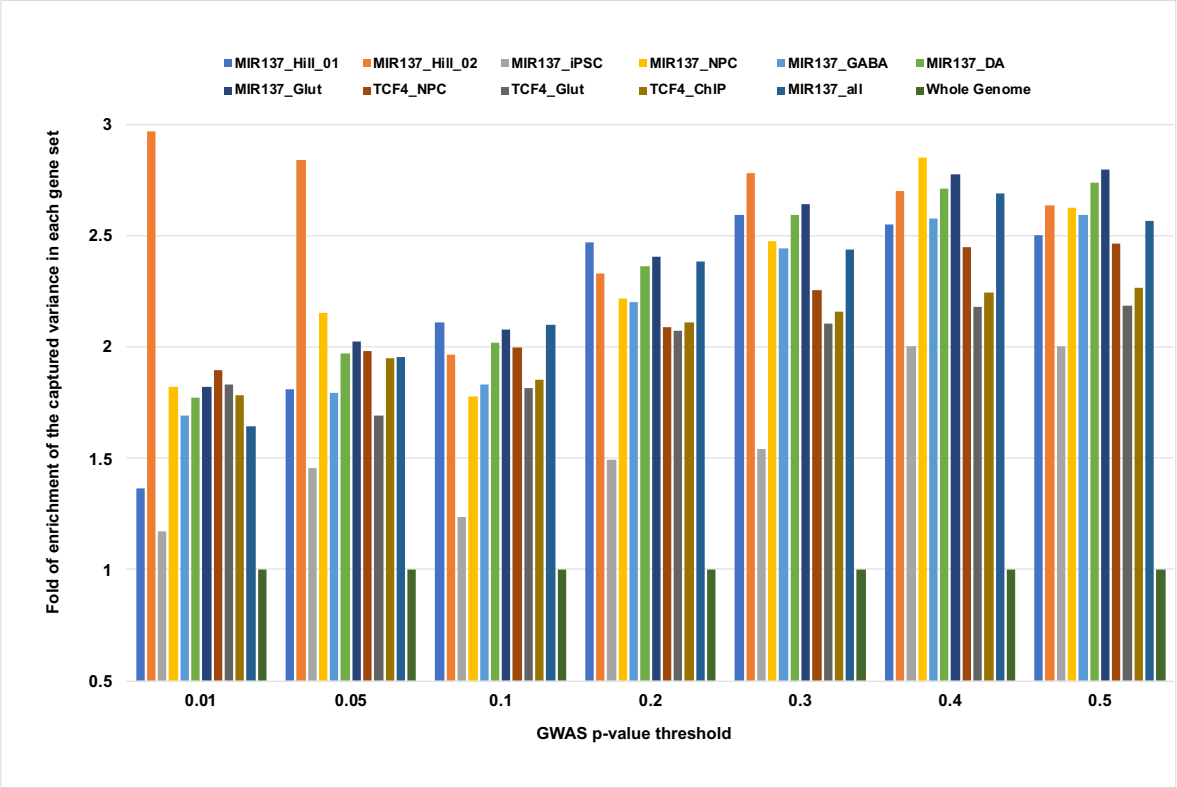

Figure S2. Relative SZ PRS enrichment analysis for different gene sets annotated in a 20 kb window for the MGS sample at different GWAS *P*-value thresholds. Related to Figure 2. The analysis used the PRS data in Figures 2 and Table S3. The fold of enrichment of the captured variance in each gene set (y-axis) was derived from correcting the gene-set PRS for the SNP number and by normalizing to the variance captured by whole genome SNP set.

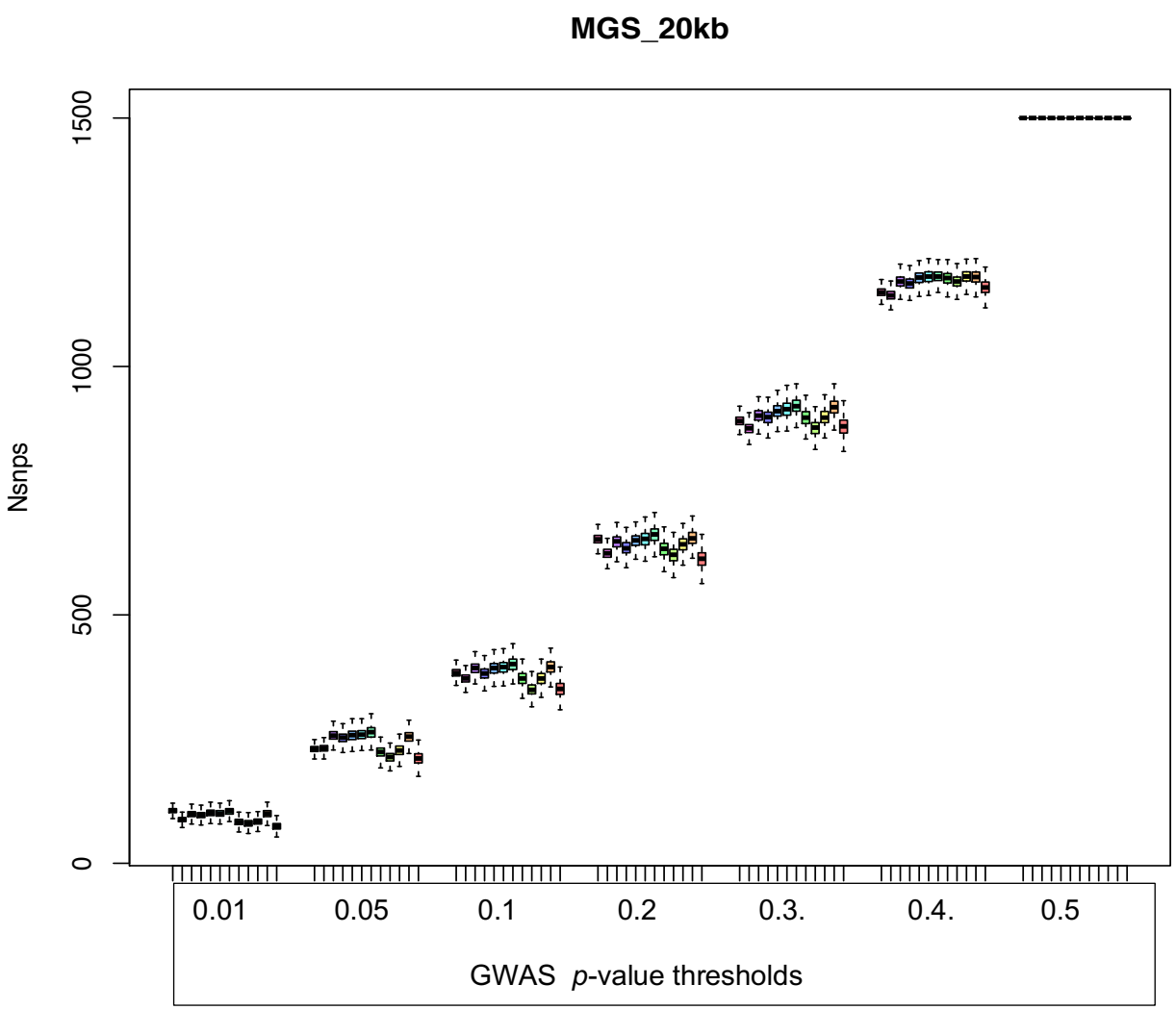

Figure S3. SNP number used in each permutation (n = 1000) for the PRS analysis results based on the randomly selected 1,500 SNPs for each gene set at different SZ GWAS *p*-value cut-offs. Related to Figure 2. Note the same number of SNPs (n=1,500) used in the permutation test for *p*-value < 0.5. The gene sets at each *p*-value cut-off are in the same order as in Figure 2B, i.e., from left to right: *MIR137*_Hill_01, *MIR137*_Hill_02, *MIR137*_iPSC, *MIR137*_NPC, *MIR137*_GABA, *MIR137*_DA, *MIR137*_Glut, *TCF4*_NPC, *TCF4*_Glut, *TCF4*_ChIP, *MIR137*_all, and Whole Genome.

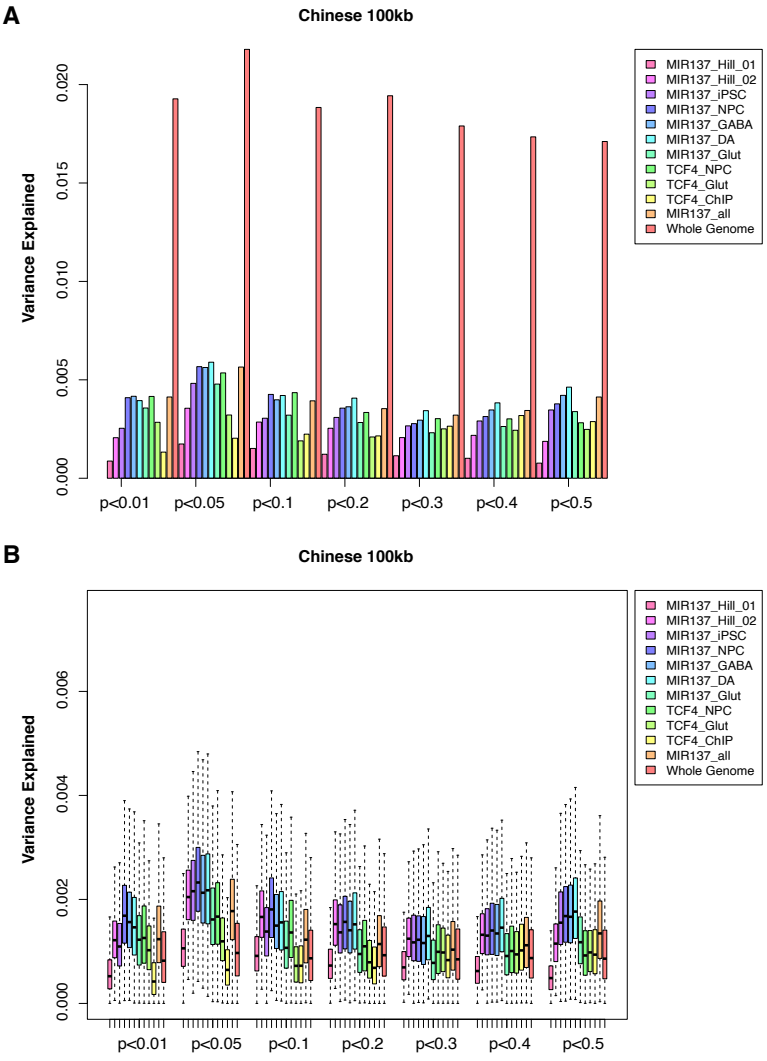

Figure S4. PRS analysis of different SNP lists of *MIR137* and *TCF4* target genes annotated in a 100kb window for the Han Chinese sample. Related to Figure 5. (A) The PRS results from using all the LD-pruned GWAS SNPs for each gene list. The variance explained in the target sample is based on risk scores derived from an aggregated sum of weighted SNP risk allele effect sizes estimated from the discovery samples at seven significance thresholds (P < 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5). The y-axis indicates the percentage of phenotypic variance explained by the PRS (Nagelkerke's pseudo R^2). (B) The permutation (N = 1,000) PRS results based on

randomly selected 3,000 SNPs for each gene set. Gene-sets in (A) and (B) were described in Table 1.