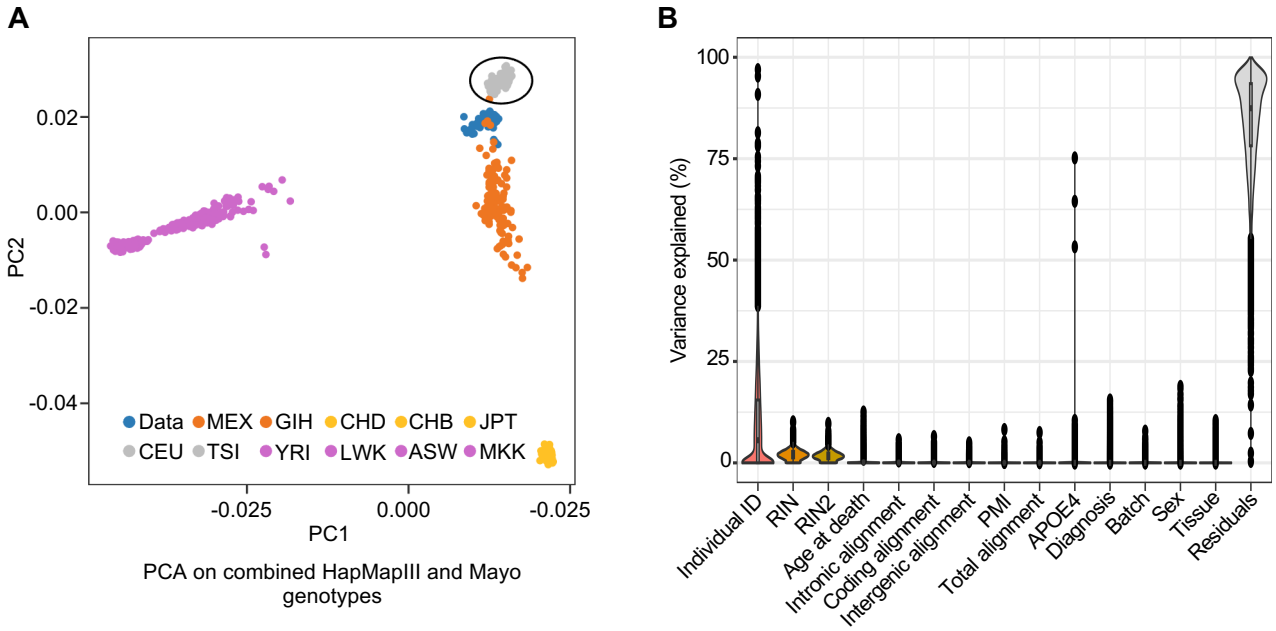

Figure S1. Data quality control of the three brain biobanks. (**A**) Ancestry estimation of the subjects from the three brain biobanks based on reference samples of HapMap III. (**B**) VariancePartition computes variance explained per gene for each variable independently. Boxplots show the median, first and third quartile of the distribution. Jittered points are genes outside 1.5 times the interquartile range. More variance is explained on average by technical factors than clinical factors.

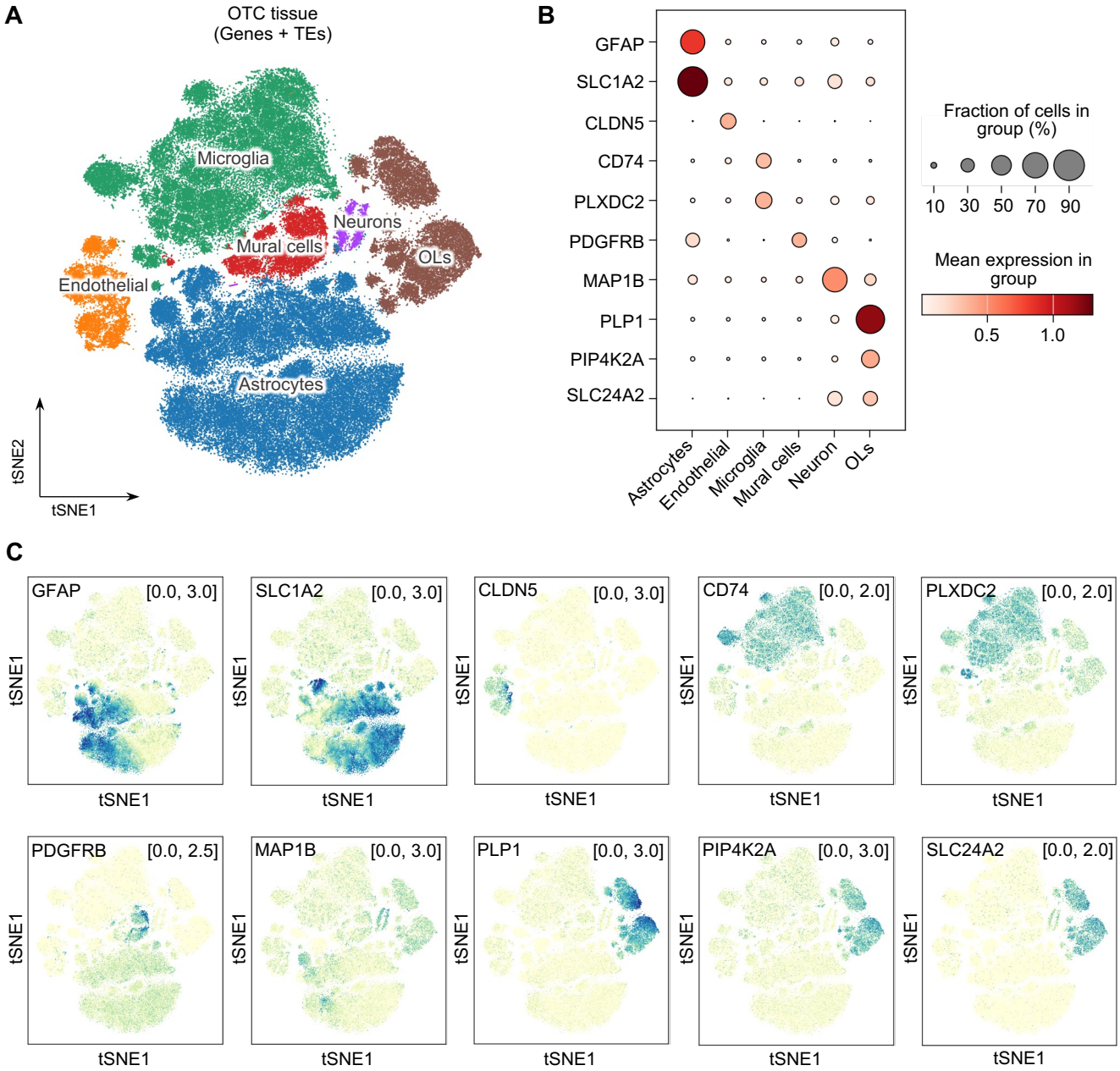

Figure S2. Cell type-specific analysis of subfamily-based differentially expressed TEs in human AD brain, Related to Figure 2. (**A**) t-distributed stochastic neighbor embedding (t-SNE) plot of clustering of 130,888 nuclei in the occipitotemporal cortex (OTC) regions. (**B**) Heatmap depicting expression changes of cell type-specific marker genes in the OTC regions. (**C**) tSNE plot depicting expression of marker gene per clusters in the OTC regions.

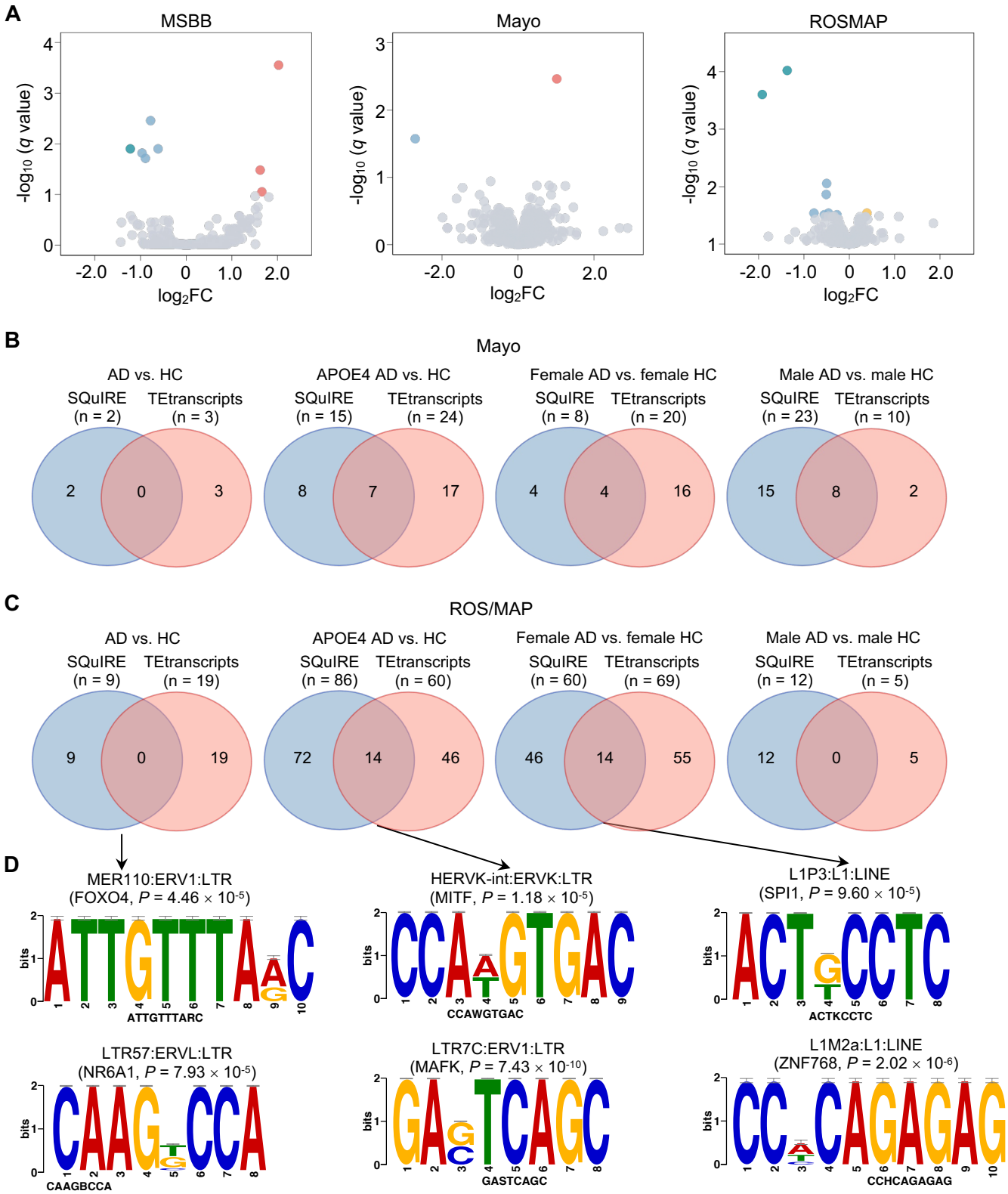

Figure S3. Family-based differential expression of transposable elements (TE) across the three brain biobanks, Related to Figure 2. (**A**) Volcano plots of subfamily-based differentially expressed TEs when comparing AD patients to cognitive healthy controls in MSBB, Mayo, and ROS/MAP cohorts. We highlighted several significant subfamilies in the three brain biobanks. TE families are colored by whether not differentially expressed (q < 0.05; gray), differentially expressed but with modest effects (|$\log_2$FC| < 1.0; orange and light blue) or with stronger effects ((|$\log_2$FC| > 1.0; red and dark blue). (**B**, **C**) Venn diagram showing number of subfamily-based differentially expressed TEs between SQuIRE and TEtranscripts in Mayo (**B**) and ROS/MAP (**C**) brain biobanks. (**D**) Motif enrichment analysis of family-based differential expressed TEs using MEME. TEs highlighted here represent the top two differentially expressed TEs identified in ROS/MAP biobank. TFs in parentheses refer to the enriched TFs of subfamily-based differentially expressed TEs under consideration.

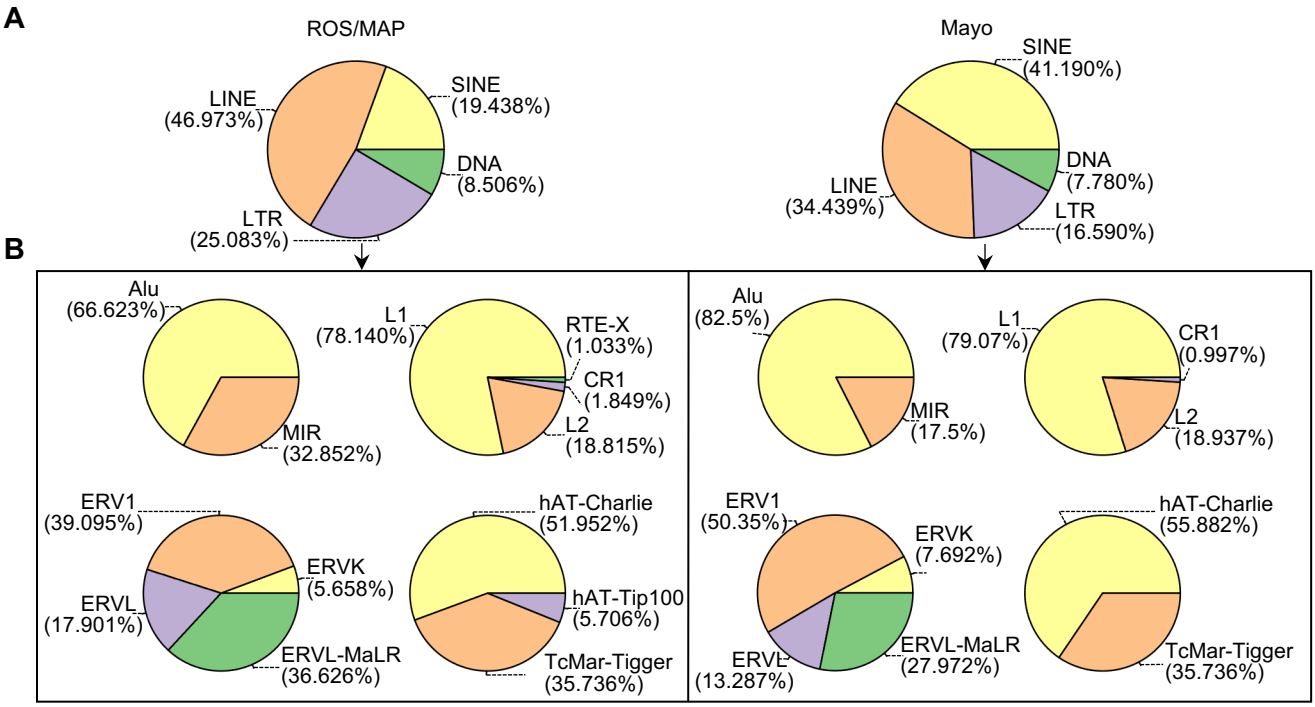

Figure S4. Analysis of locus-based differentially expressed TEs, Related to Figure 2. (**A**) Pie charts representing the proportion of locus-based differential expressed TE in each TE class (SINE, LINE, LTR, and DNA) across Mayo and ROS/MAP brain biobanks. (**B**) Pie charts showing the proportion of locus-based differentially expressed TEs among TE subfamilies across ROS/MAP and Mayo brain biobanks.

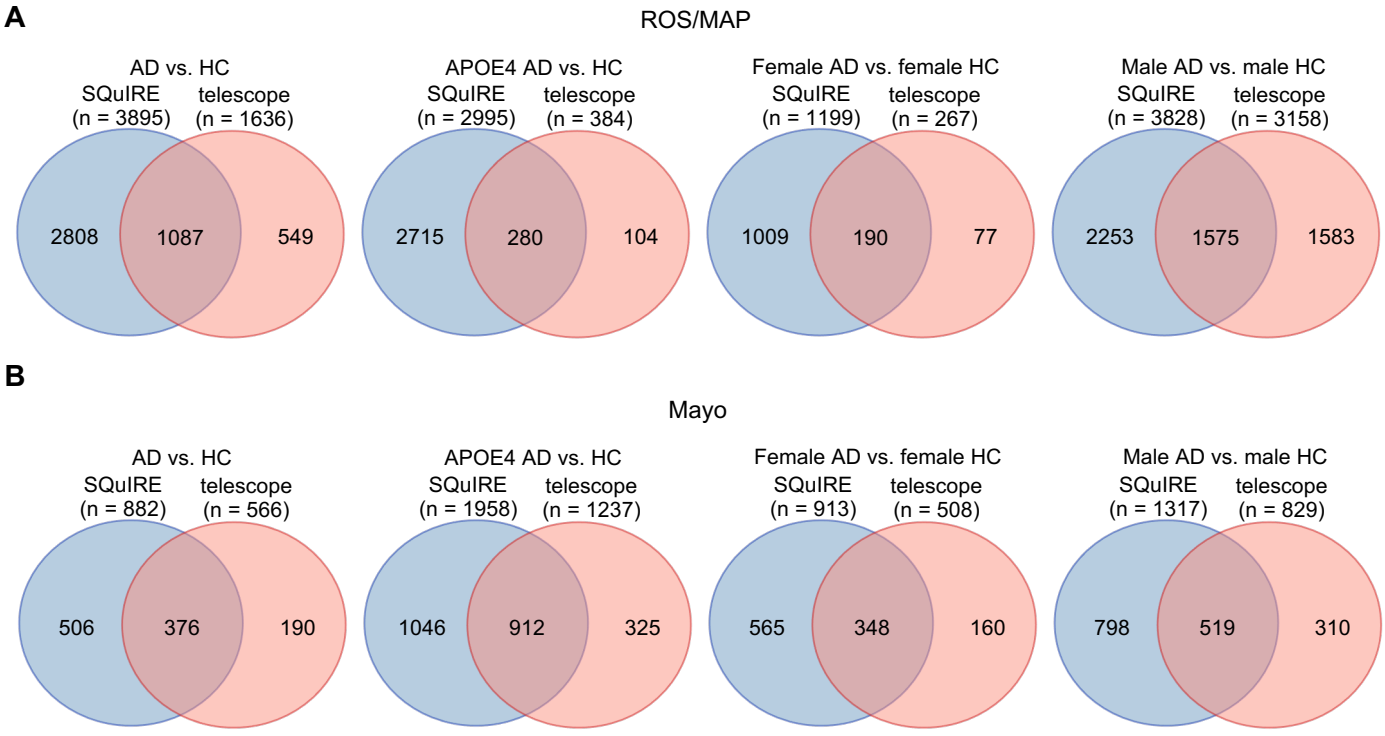

Figure S5. Venn diagrams showing number of locus-based differentially expressed TEs between SQuIRE and Telescope in different biological comparison groups in ROS/MAP (**A**) and Mayo (**B**) brain biobanks.

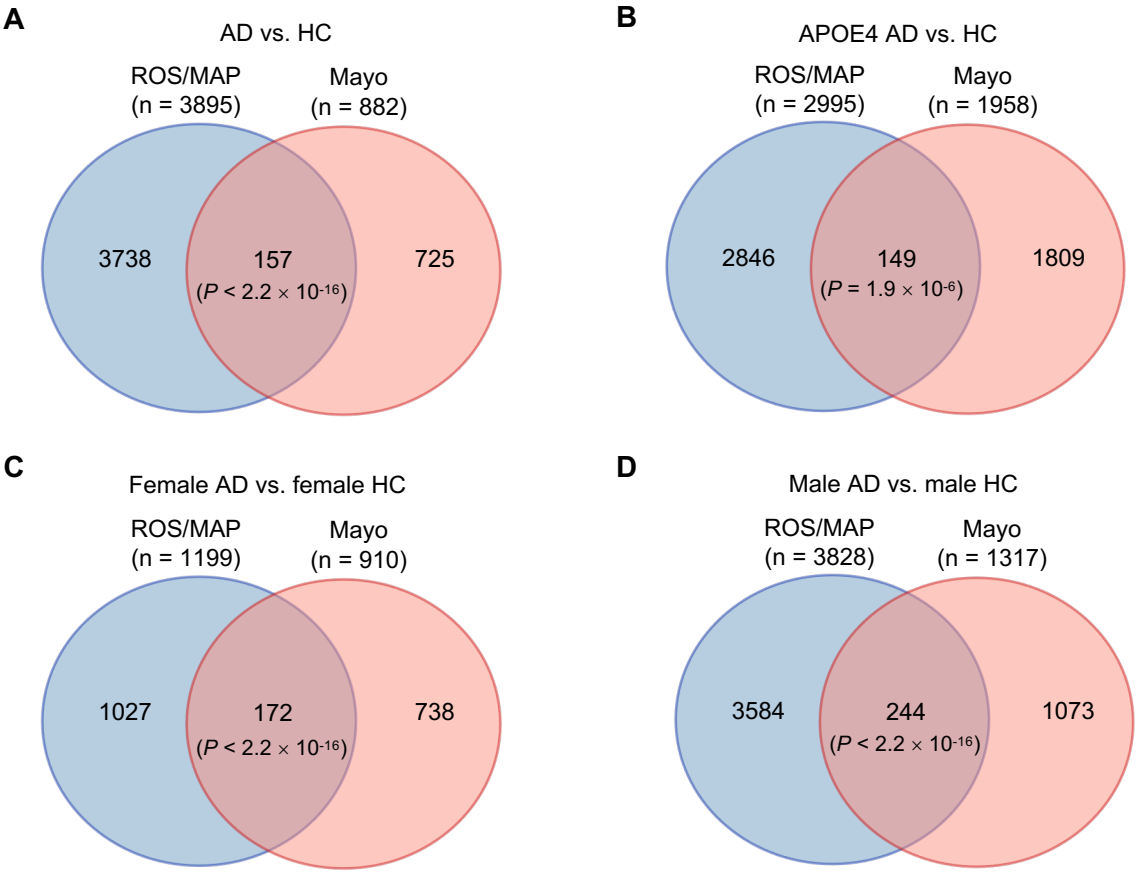

Figure S6. Venn diagrams showing overlap trend of locus-based differentially expressed TEs between ROS/MAP and Mayo brain biobanks in different biological comparison groups. *P* values was calculated using Fisher's exact test.

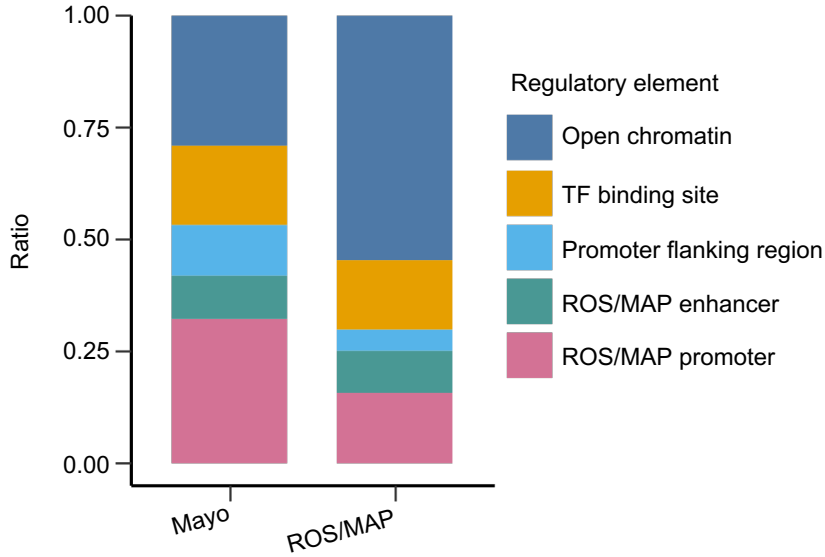

Figure S7. Bar plot showing proportions of genomic locations of locus-based differentially expressed TEs within five categories of functional cis-regulatory elements, including open chromatin regions, TF binding sites, promoter flanking regions, ROS/MAP enhancer regions, and ROS/MAP promoters.

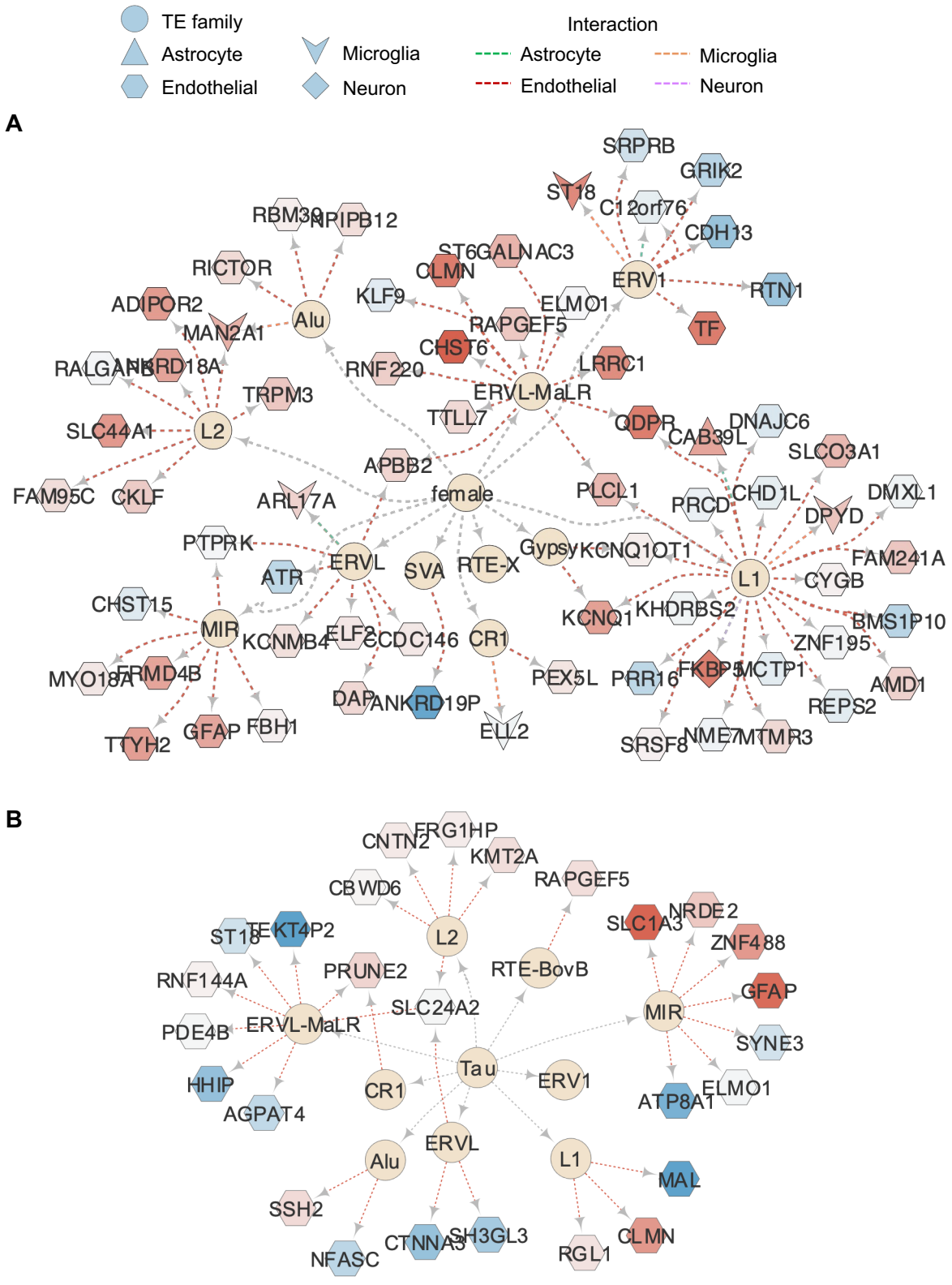

Figure S8. TE-gene regulatory networks depicting cell type-specific differentially expressed TEs affected by female- (A) and male-biased (B) in ROS/MAP brain biobanks. Circles in the inner layer represent TE families. Genes on the outer layer surface refer to the nearest genes of the locus based TEs. Colors on the outer layer panel refer to $\log_2$FC of differentially expressed genes from the RNA-seq Harmonization Study (syn21241740). Genes differentially expressed in microglia are showing as V, neuron as diamond, endothelial as hexagon, and astrocytes as triangle.

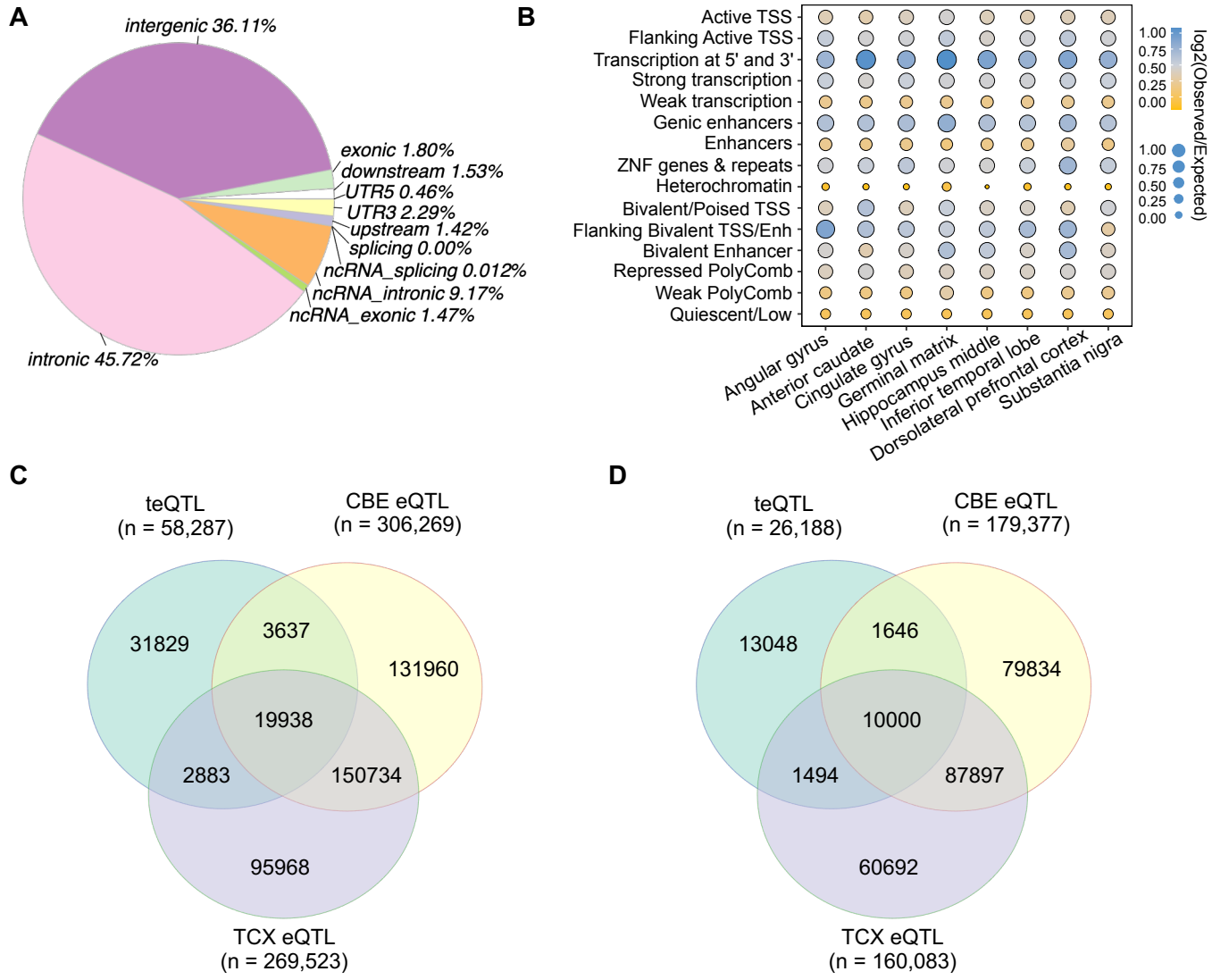

Figure S9. **Transposable element expression quantitative trait loci (teQTL) analysis in human brains, Related to Figure 3.** (**A**) Pie chart showing proportions of teQTLs in specific genomic regions annotated with GRCh37 reference genome using ANNOVAR. (**B**) Enrichment of teQTLs in 15 different chromatin states as defined by the Roadmap Epigenomic Consortium projects in eight human brain cell lines, including brain angular gyrus (E067), brain anterior caudate (E068), brain cingulate gyrus (E069), brain germinal matrix (E070), brain hippocampus middle (E071), brain inferior temporal lobe (E072), brain dorsolateral prefrontal cortex (E073), and brain substantia nigra (E074). (**C** and **D**) Venn diagrams showing overlap trend between teQTLs and eQTLs from Mayo brain biobank under sub-GWAS threshold (**C**) and genome-wide significant threshold (**D**).

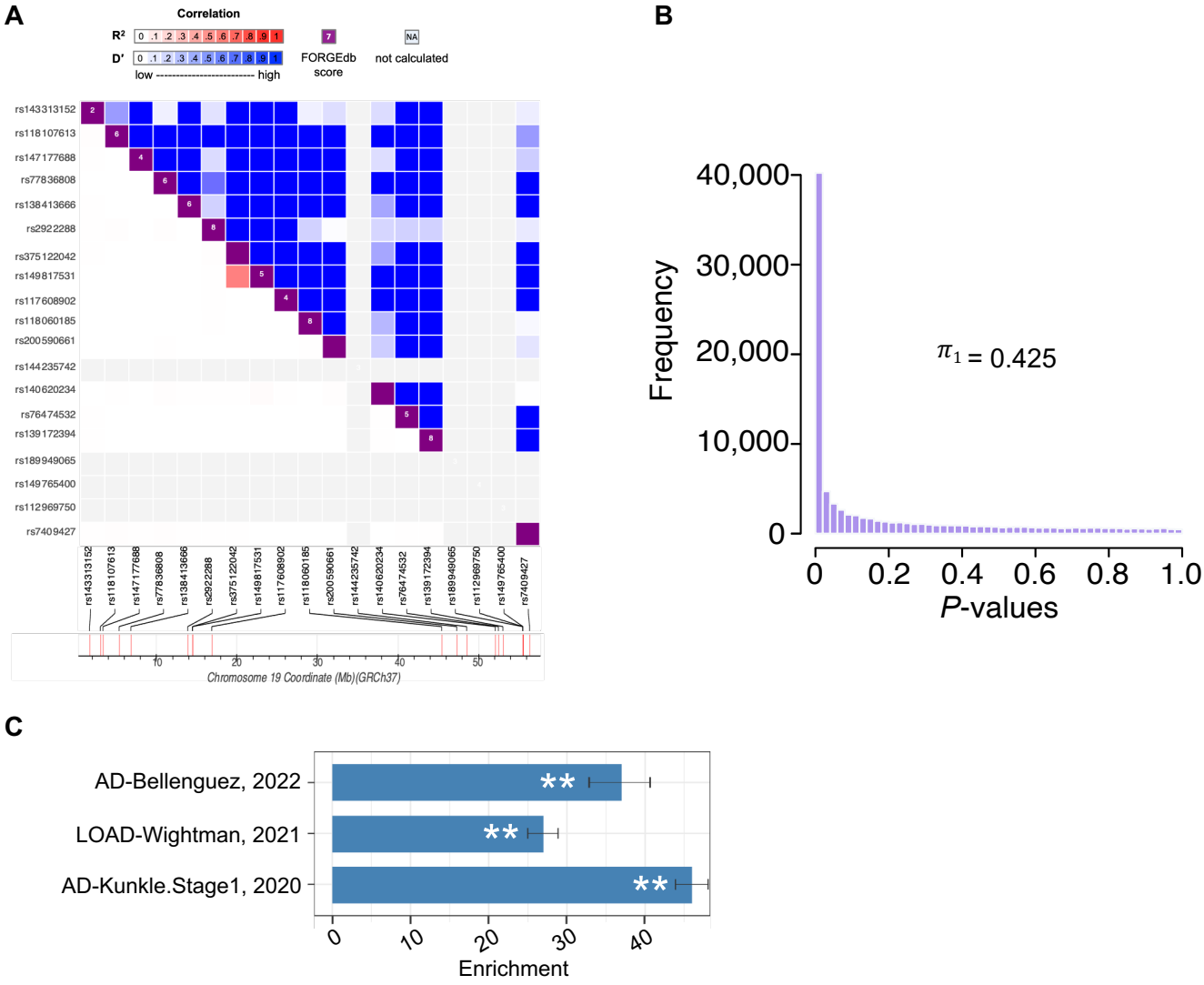

Figure S10. **Transposable element expression quantitative trait loci (teQTL) analysis in human brains, Related to Figure 3.** (**A**) LD matrix showing relationship between teQTLs and eQTLs of Mayo meta-analysis for the 1,134 common regulated genes. This figure shows results from chromosome 19 using LDmatrix tool. (**B**) *P* values distribution of Mayo teQTLs that are significant in ROS/MAP haQTLs ($q < 0.05$). (**C**) LD score regression analysis revealed enrichment of teQTLs across the three AD GWAS studies.

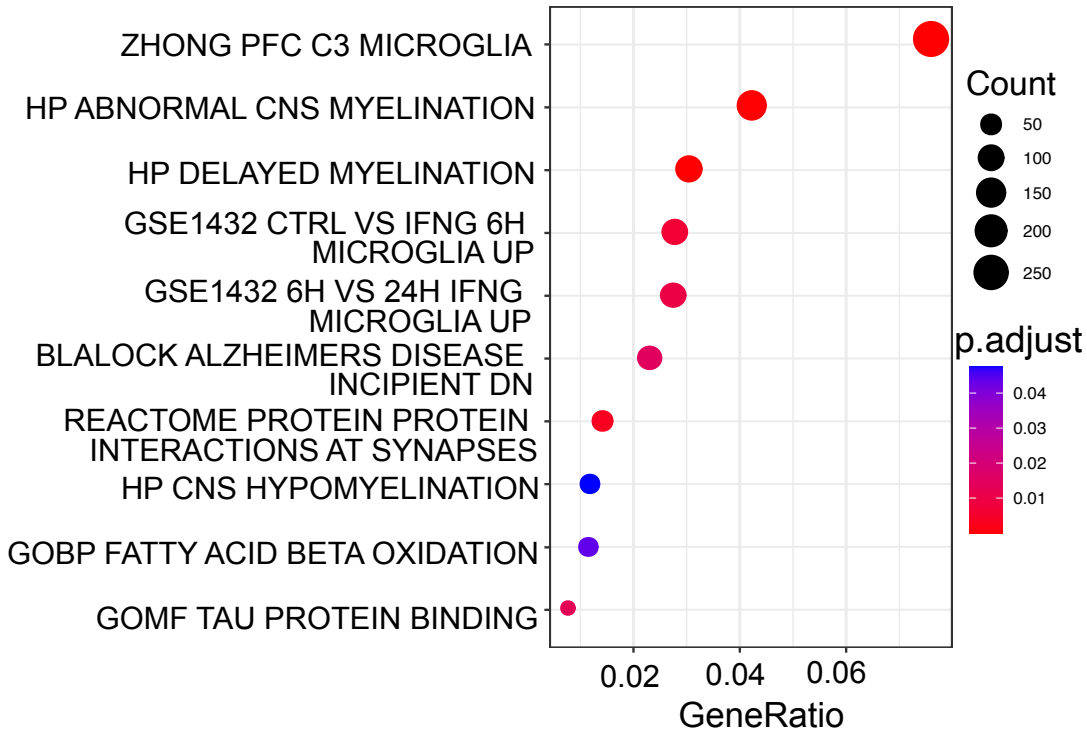

Figure S11. **Gene set enrichment analysis of genes regulated by human brain teQTLs.** Fisher's exact test was used for statistical analysis. Terms with Benjamini-Hochberg corrected *P* values (*q*) < 0.05 were defined as significantly enriched pathways. Circle size represents number of genes enriched in each pathway. Circle color represents *q* value.

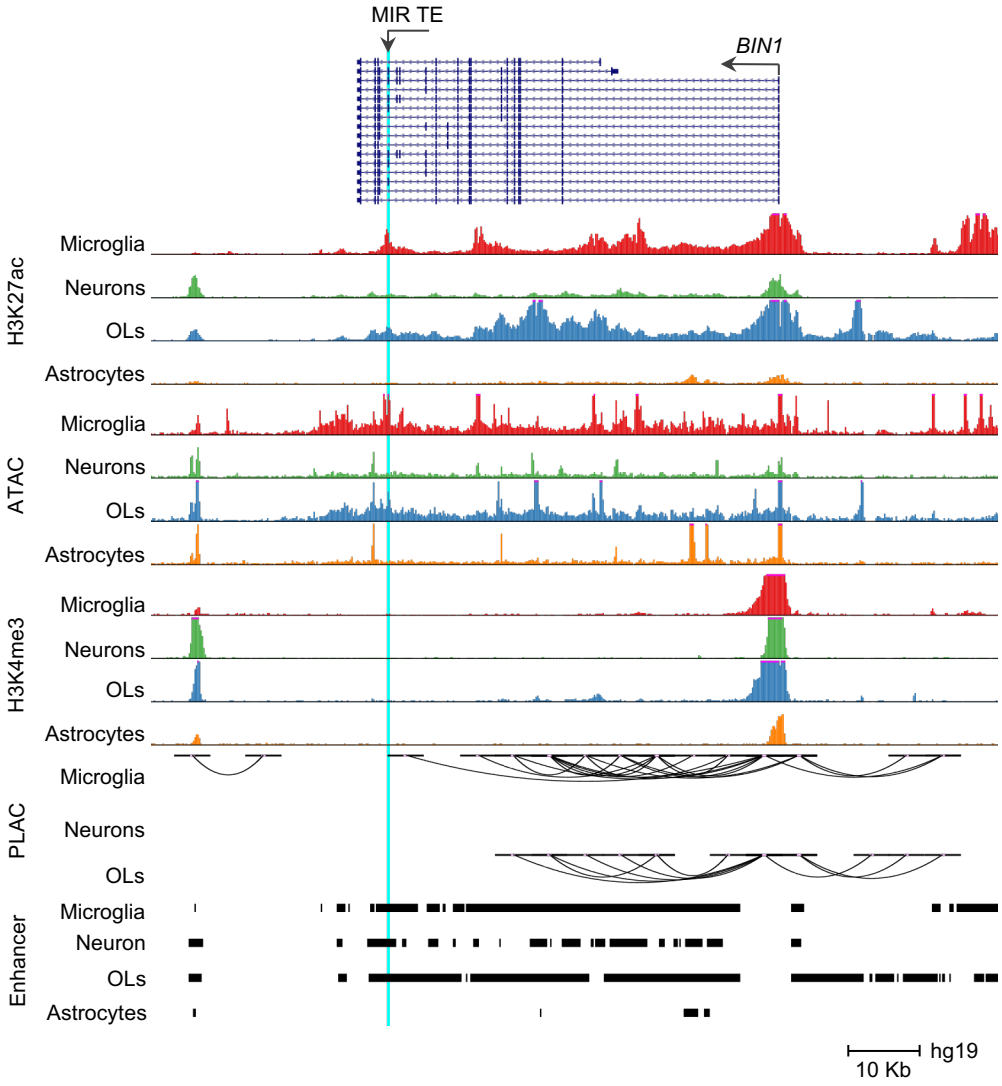

Figure S12. UCSC genome browser visualization of chromatin interactions between MIR family TE-located neuron-specific enhancer and promoter regions of *BIN1*. OLs: oligodendrocytes.

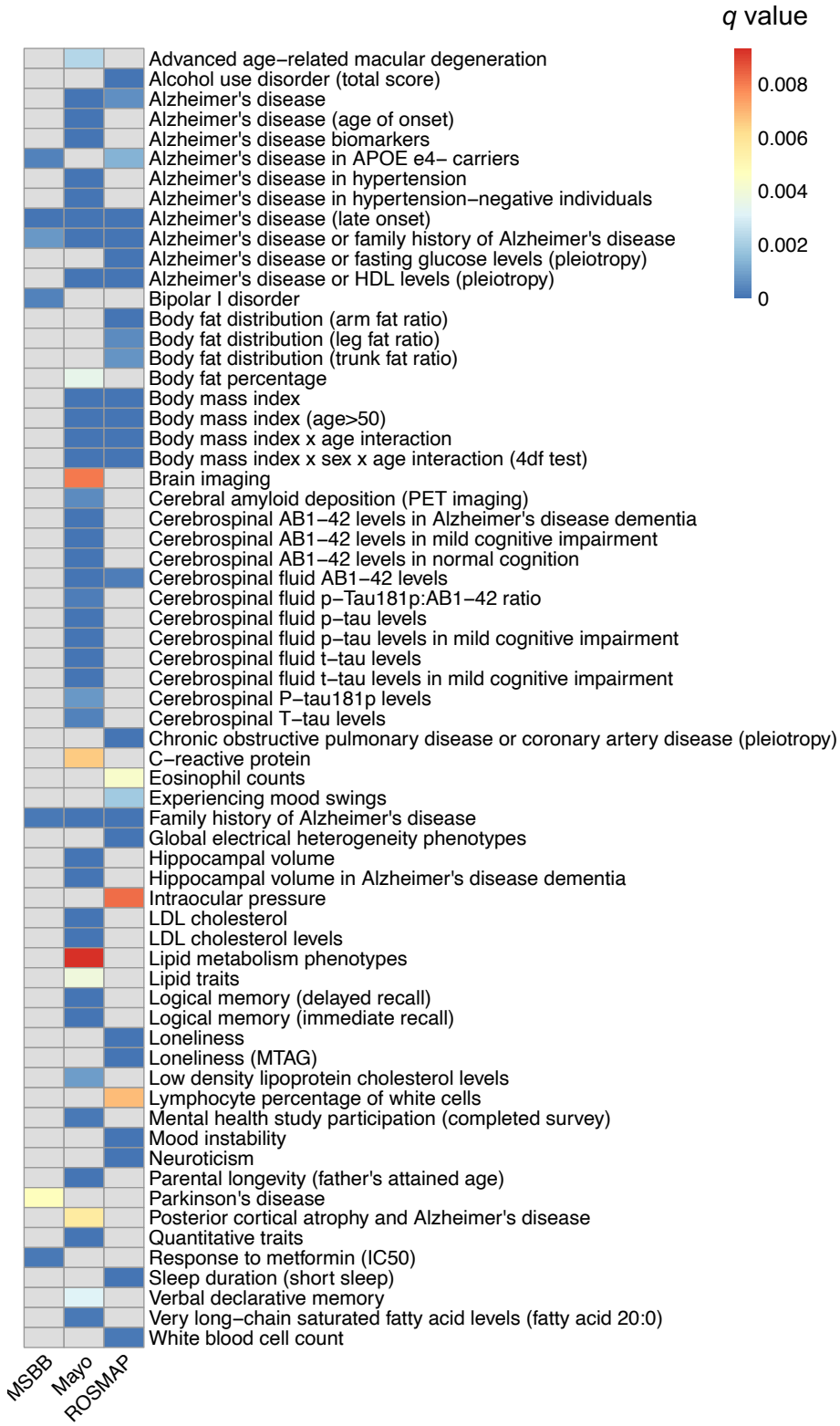

Figure S13. Functional mapping and annotation of GWAS SNPs using SNP2GENE and Gene2FUNC functions in Functional Mapping and Annotation (FUMA), Related to Figure 4. Those SNPs were in 1Mb flanking regions of differential expressed TEs.

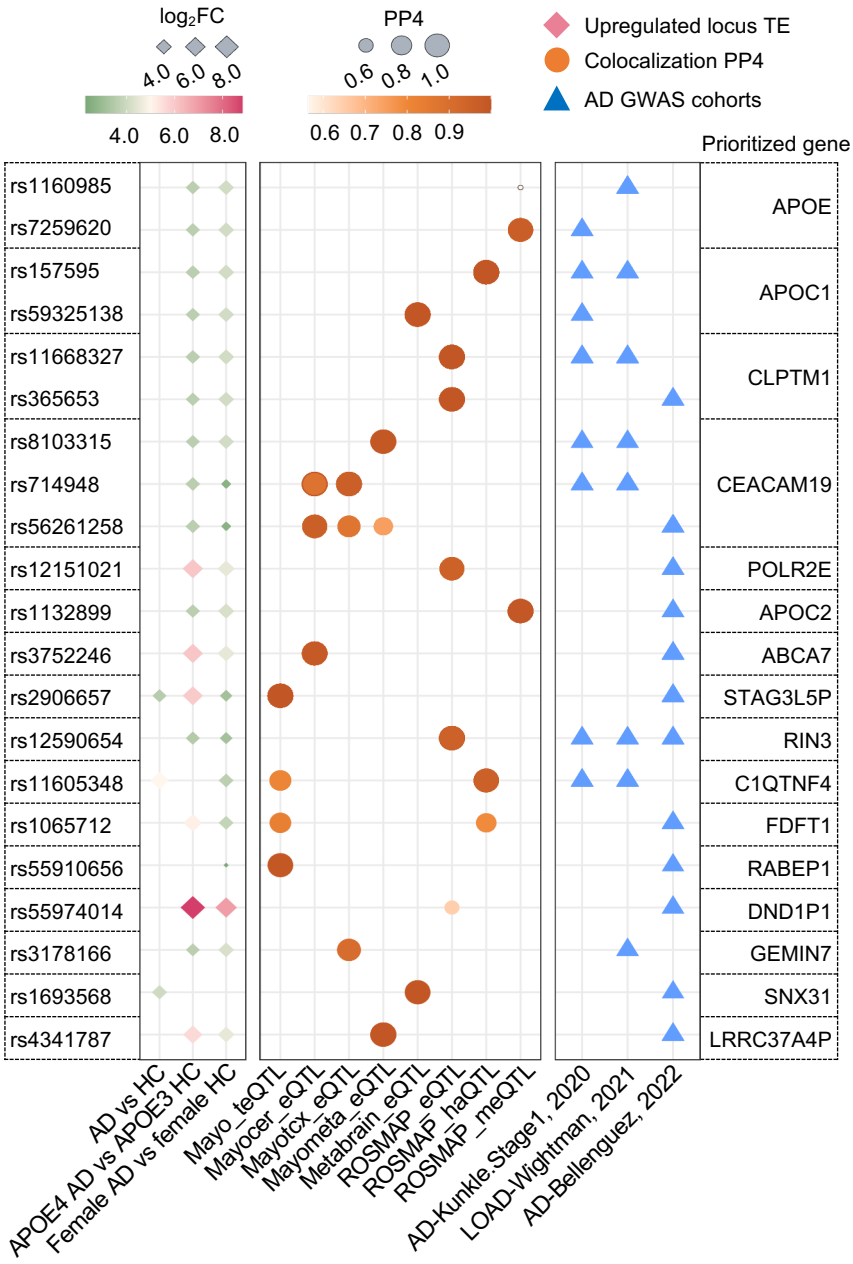

Figure S14. **Colocalization analysis around 500 Kb flanking regions of locus-based differentially expressed TEs between AD GWAS SNPs and xQTLs in Mayo brain biobank.** xQTLs include teQTL from Mayo, meQTL from ROS/MAP, haQTL from ROS/MAP, and eQTL from ROS/MAP, Mayo, and MetaBrain datasets. We showed SNPs with genome-wide significant AD GWAS $P < 5.0 \times 10^{-8}$. Loci are named for their nearest protein-coding genes or the regulated eGene. All genes and loci shown here with PP4 $\geq$ 0.5 in at least one xQTL dataset. Circles size and color refer to PP4 value. Diamonds refer to locus-based differentially expressed TEs from different biological comparison groups. Diamonds size and color refer to $\log_2$FC of locus-based differentially expressed TEs under consideration. Triangles refer to the six AD GWAS datasets. PP4: posterior probability of colocalization hypothesis 4.

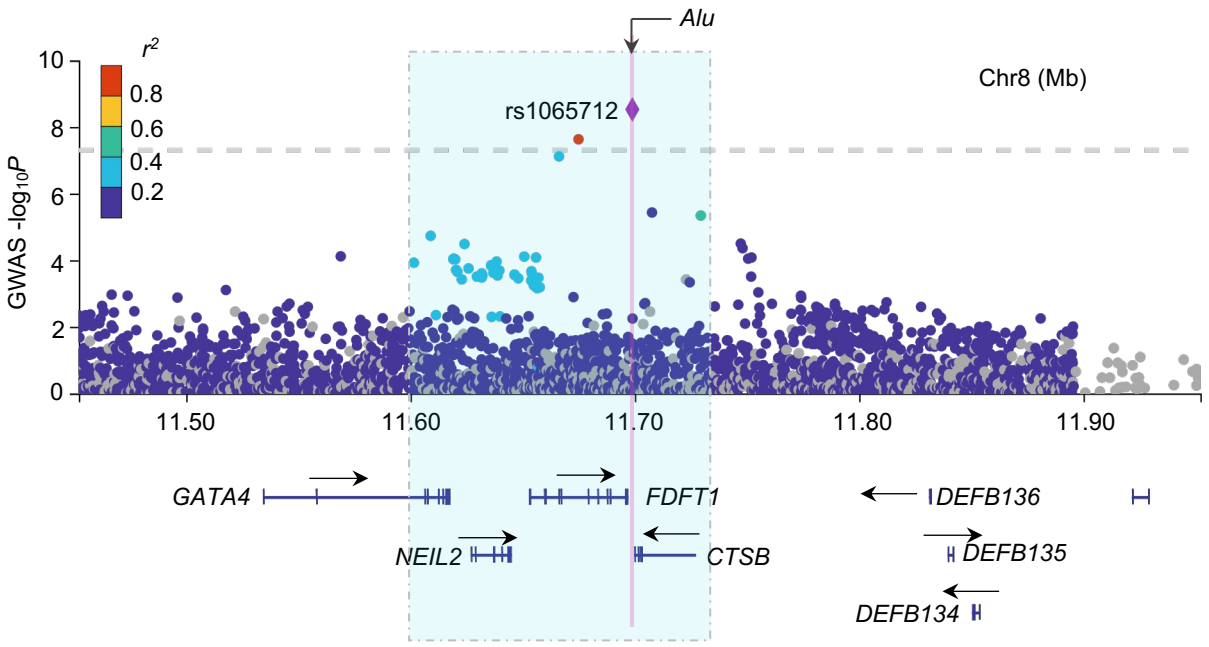

Figure S15. Regional SNP association plots with TE from *AluJb* subfamily (chr8: 11840915-11841089) shown in LD blocks of rs1065712 in AD. Plots are colored based on linkage disequilibrium (LD) bins relative to the lead SNPs rs1065712 (red, ≥ 0.8; orange, 0.6-0.8; green, 0.4-0.6; light blue, 0.2-0.4; and dark blue, < 0.2). The SNP pairwise LD data were calculated based on the 1000 Genomes Phase 3 (ALL) reference panel. Gene annotations: the NCBI RefSeq Select database. Assembly GRCh37, scale in Mb.

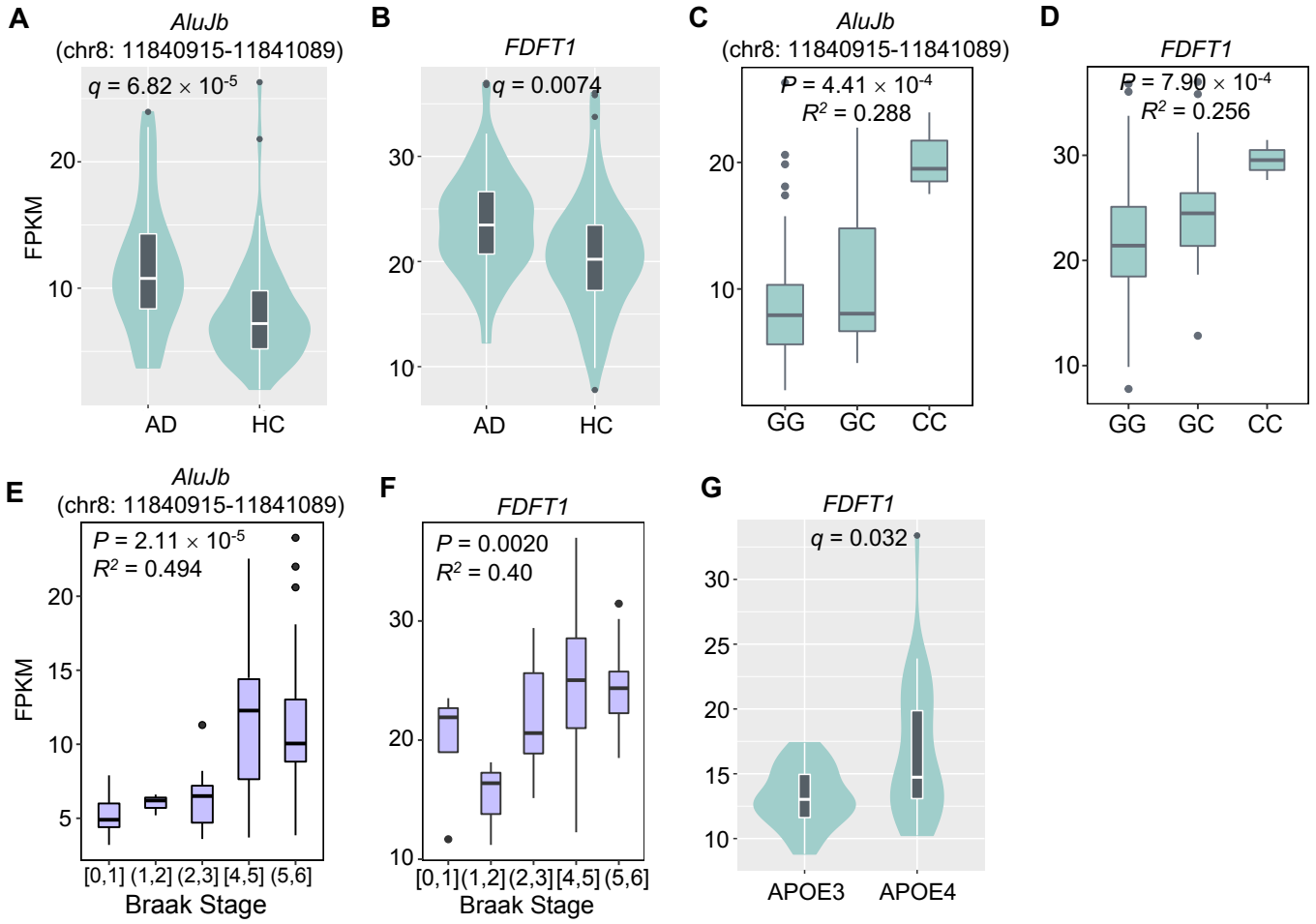

Figure S16. Expression comparison between *Alu* subfamily TE (chr8: 11840915-11841089**) and its nearest gene *FDFT1*, Related to Figures 5 and 6.** (**A-D**) mRNA abundance comparison between AD cases and cognitive healthy controls in *AluJb* subfamily TE (chr8: 11840915-11841089) (**A**), and *FDFT1* (**B**). (**C, E**) Expression level of *AluJb* subfamily TE (chr8: 11840915-11841089) (**C**) and its nearby gene *FDFT1* (**E**) showing significantly positive correlation with the presence of risk allele of rs1065712. *P* values are calculated using Spearman correlation test. R^2 values showing the strength of expression correlation by Spearman correlation test. (**D, F**) Expression level of *Alu* subfamily TE (chr8: 11840915-11841089) (**D**) and its nearby gene *FDFT1* (**F**) showing significantly positive correlation with Braak staging score. *P* values are calculated using Spearman correlation test. R^2 values showing the strength of expression correlation by Spearman correlation test. (**G**) Comparison of *FDFT1* expression between heterozygous *APOE4* AD patients and heterozygous *APOE3* AD patients from Mayo brain biobank.

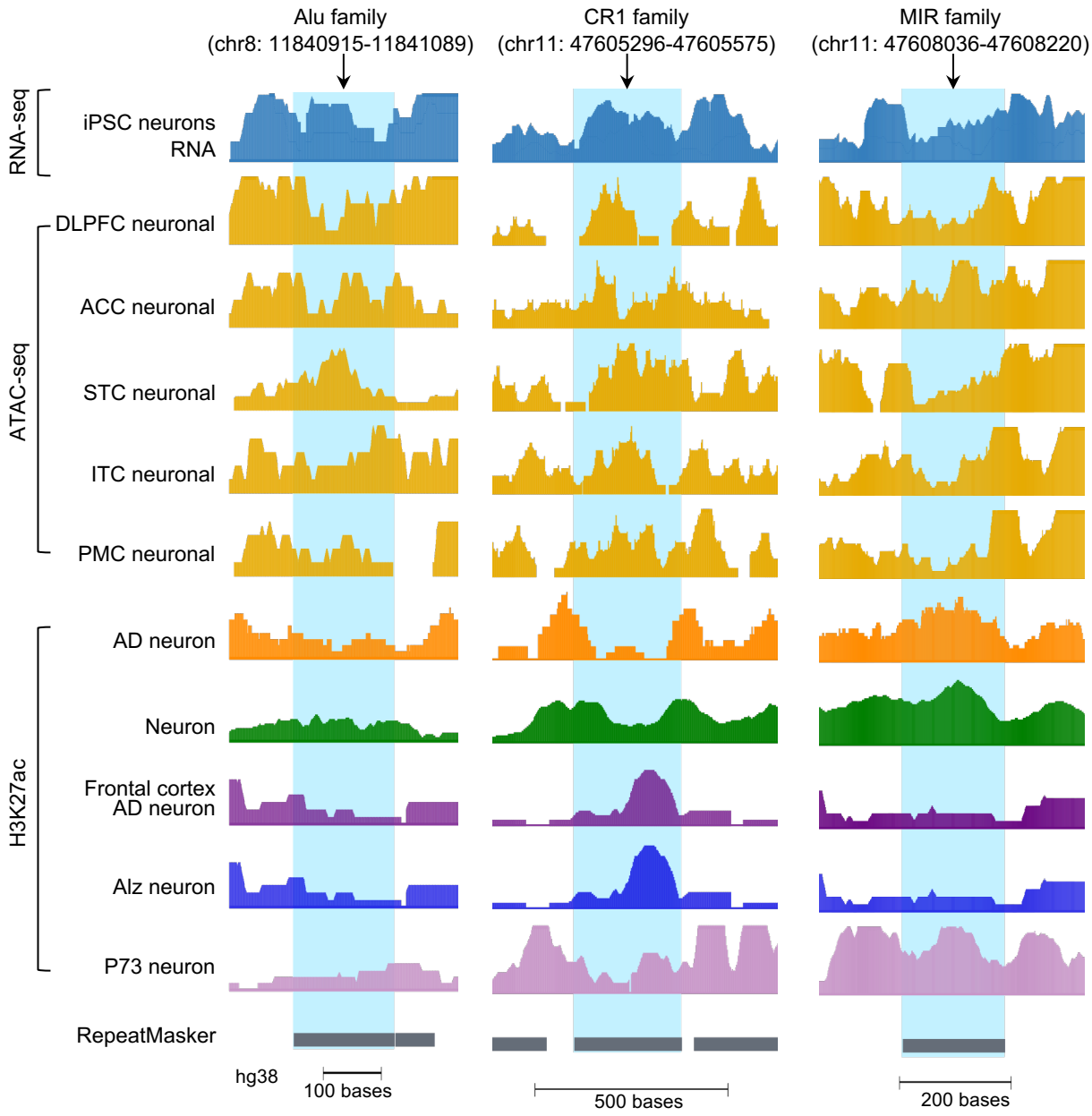

Figure S17. **UCSC genome browser visualization showing candidate TEs, RNA-seq read counts in iPSC-derived neurons, neuron chromatin accessibility data across different brain regions, and H3K27ac chromatin mark in different biological context of neurons**. All the TEs shown in this study are annotated with assembly GRCh38.

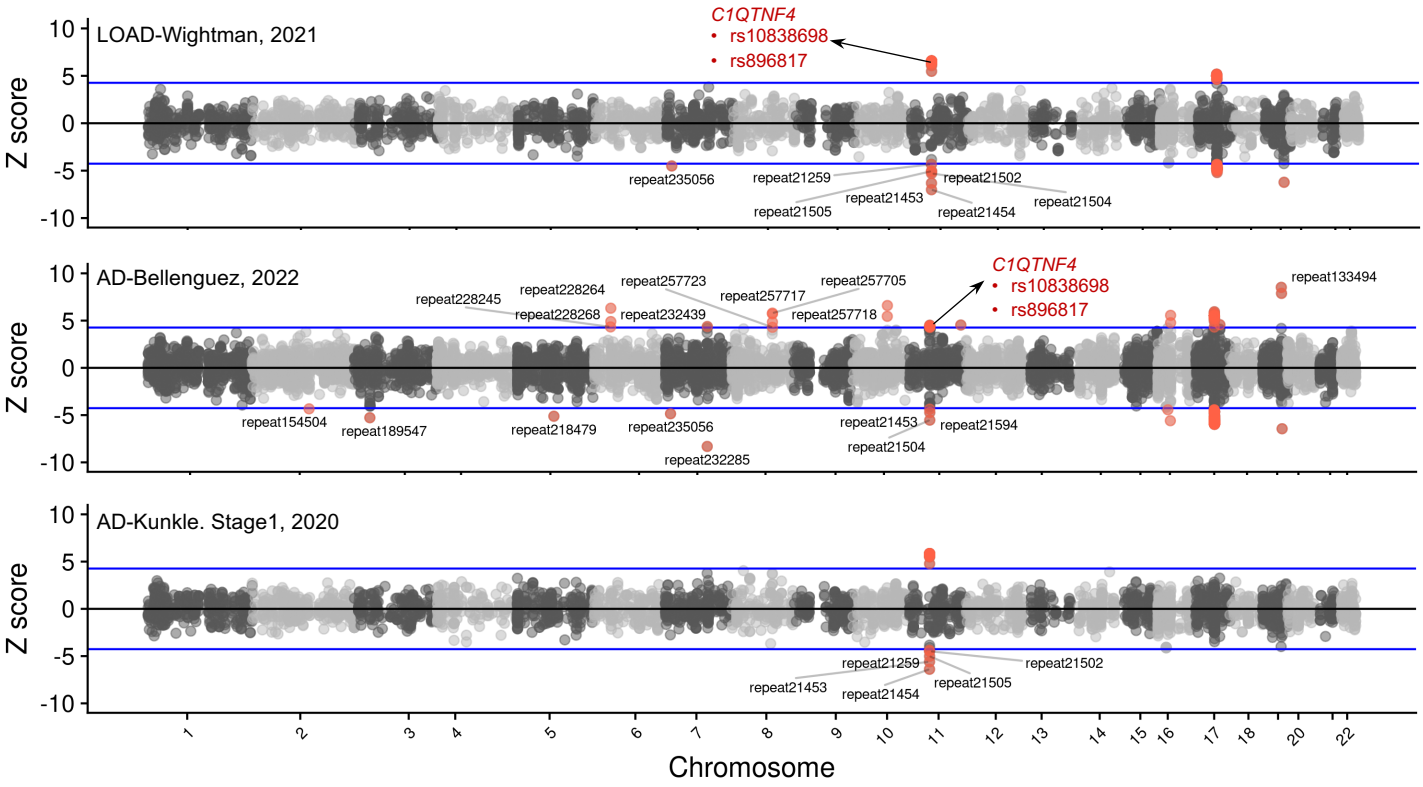

Figure S18. Manhattan plots illustrating the z-scores of TE transcriptome-wide association study using six AD GWAS summary statistic datasets, Related to Figure 6. 152 samples from Mayo brain biobank were used in TE transcriptome-wide association analysis. Key TEs associated with those teQTLs were highlighted in manhattan plots. Genes in the parentheses refer to the nearby genes for the TEs under consideration.

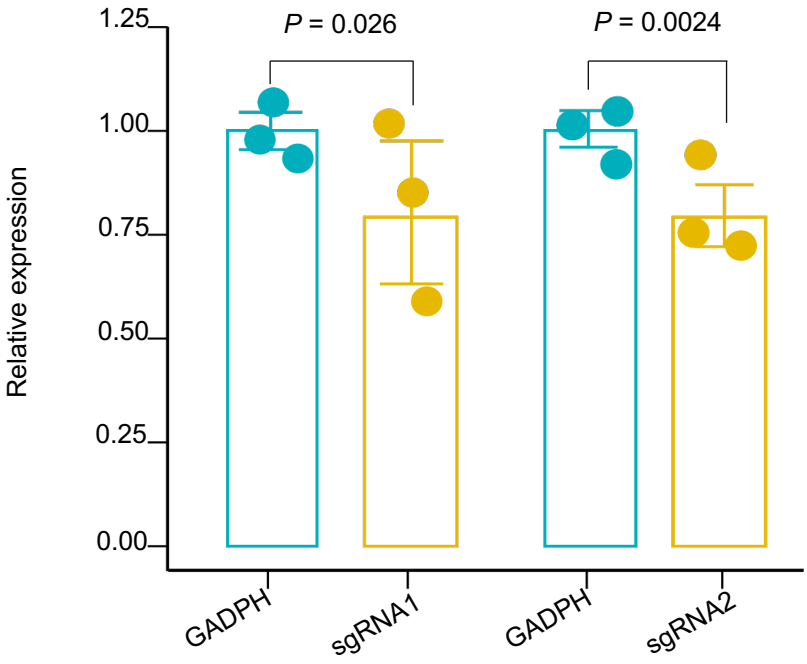

Figure S19. CRISPR inference validation of Plat_L3 family TE (chr11: 47605296-47605575) in iPSC-derived neurons. NDUFS3 expression relative to GAPDH in neurons with control and with CRISPRi targeted Plat_L3 subfamily TE (chr11: 47605296-47605575). Error bars, 95% CI. All data are shown as mean (SEM) of three independent experiments. *P* values are calculated using Student's t test.

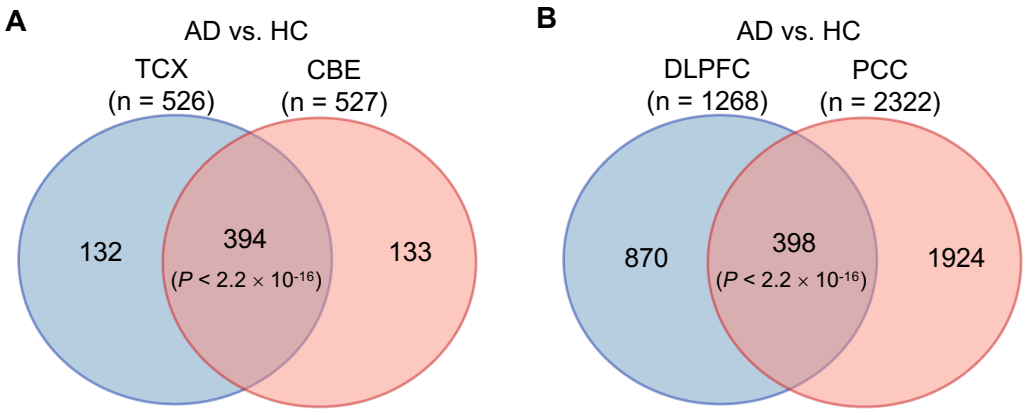

Figure S20. Overlap analyses of locus-based differentially expressed TEs between different brain regions in Mayo (A) and ROS/MAP (B) brain biobanks when comparing AD brains with cognitively heathy controls, using a well-established data harmonization approach established by the AMP-AD consortium (syn21241740). *P* values was calculated using Fisher's exact test.

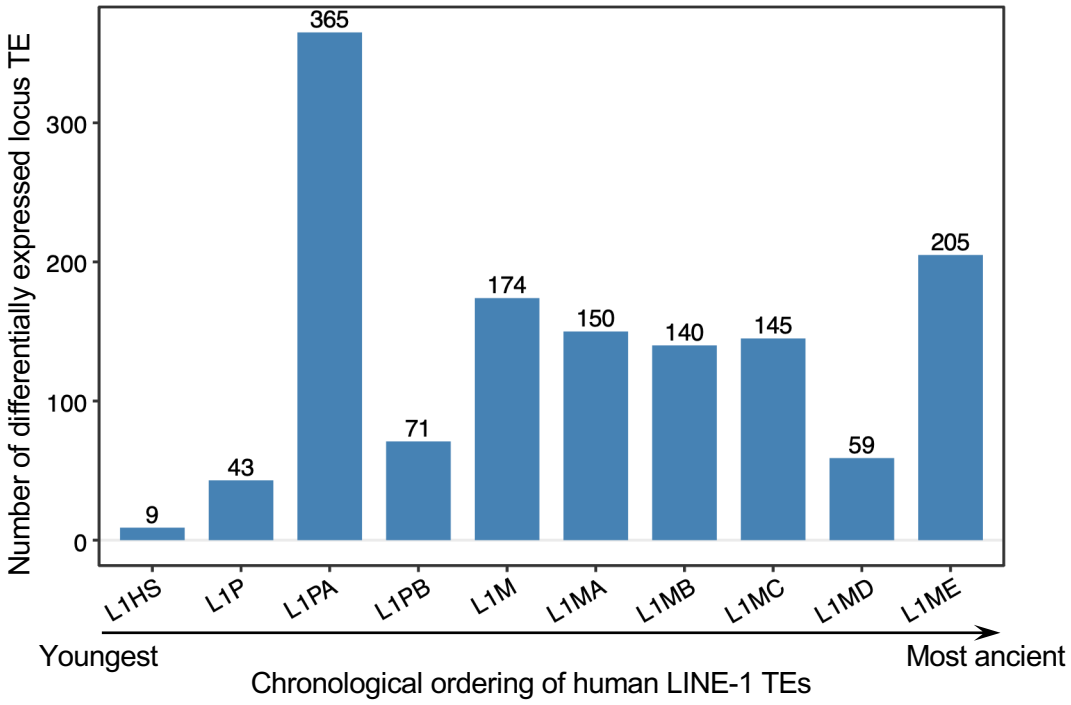

Figure S21. Number of locus-based differentially expressed TEs across chronological ordering of human LINE-1 family TE.

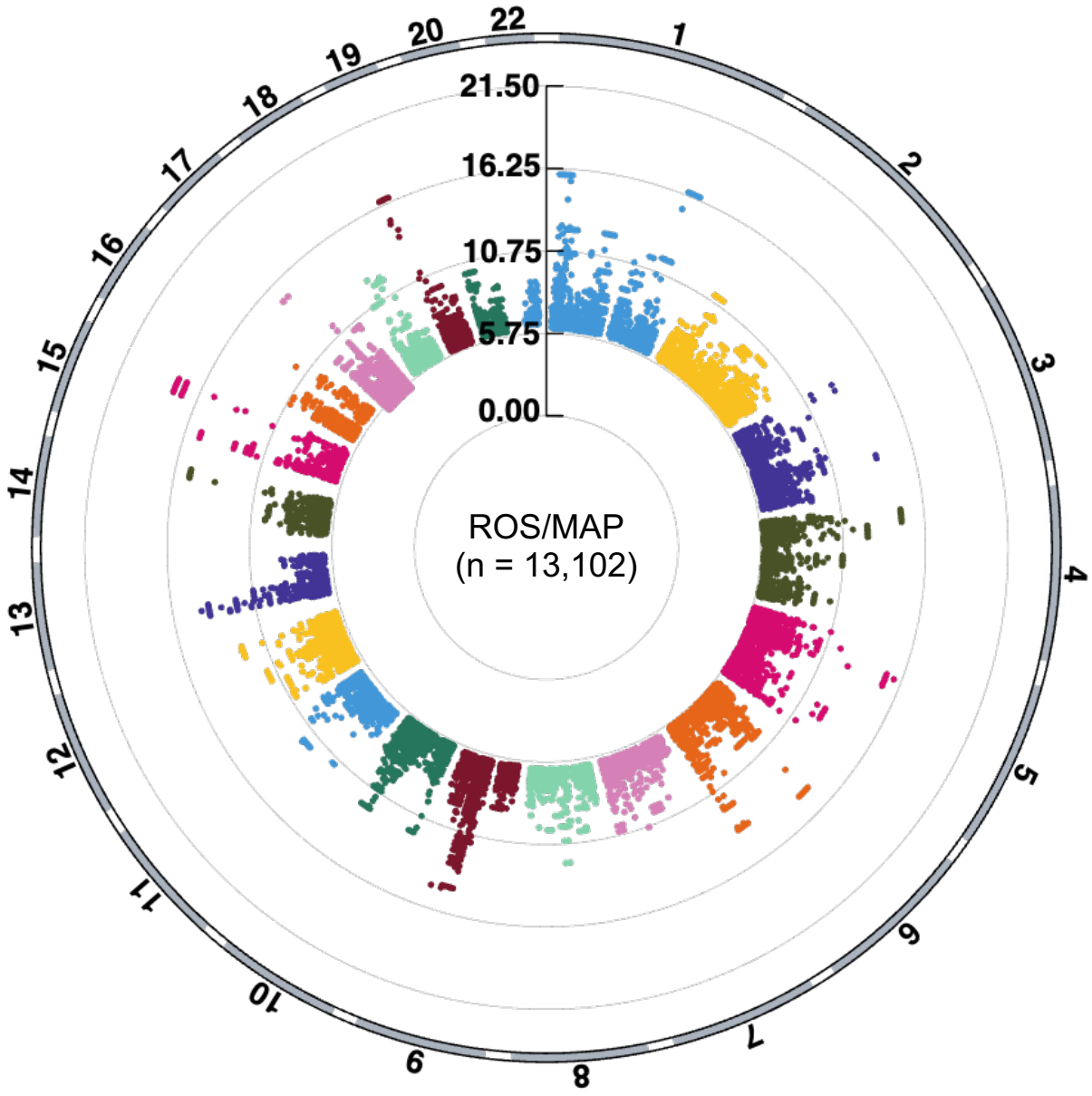

Figure S22. Circle plot showing the *P* value distribution of teQTLs identified in the ROS/MAP brain biobank, of which 45 samples have both RNA-seq and WGS data, Related to Figure 3.

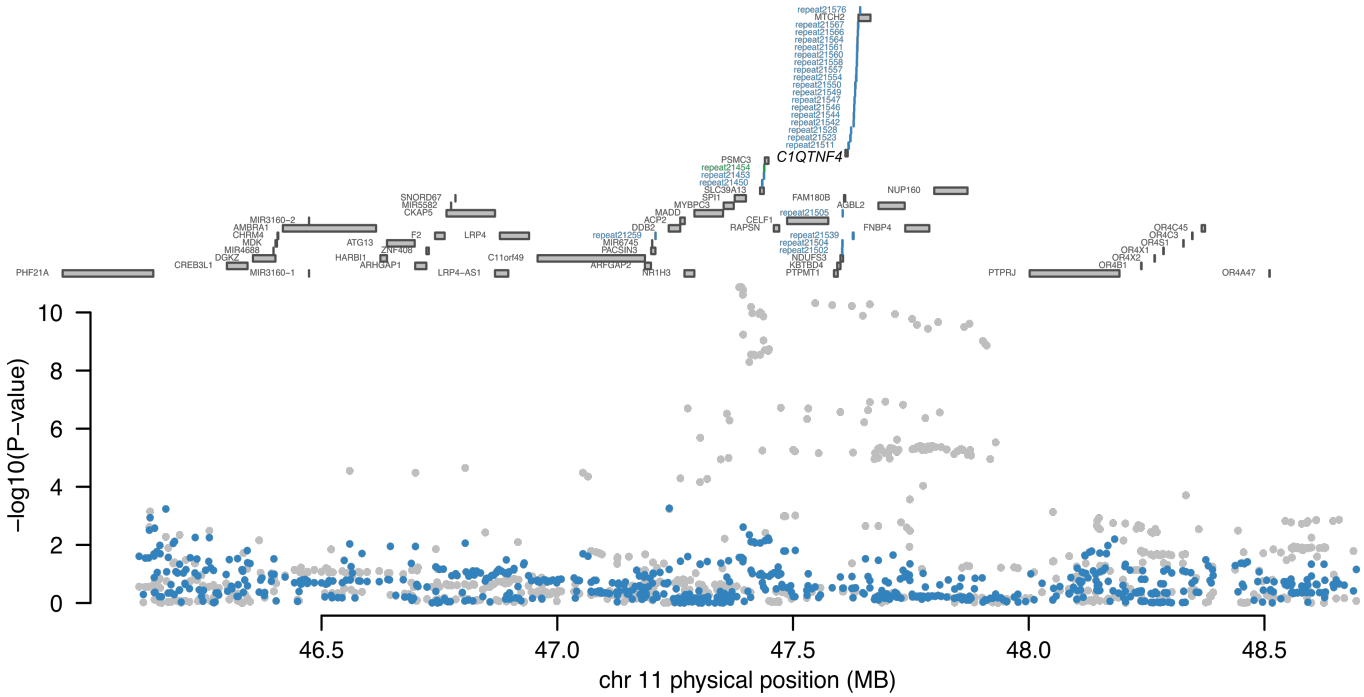

Figure S23. FUSION conditional analysis using AD GWAS as a reference panel. The top panel shows all the TEs in the locus. The marginally TWAS associated TEs are highlighted in blue, and those that are jointly significant highlighted in green, Related to Figure 6.

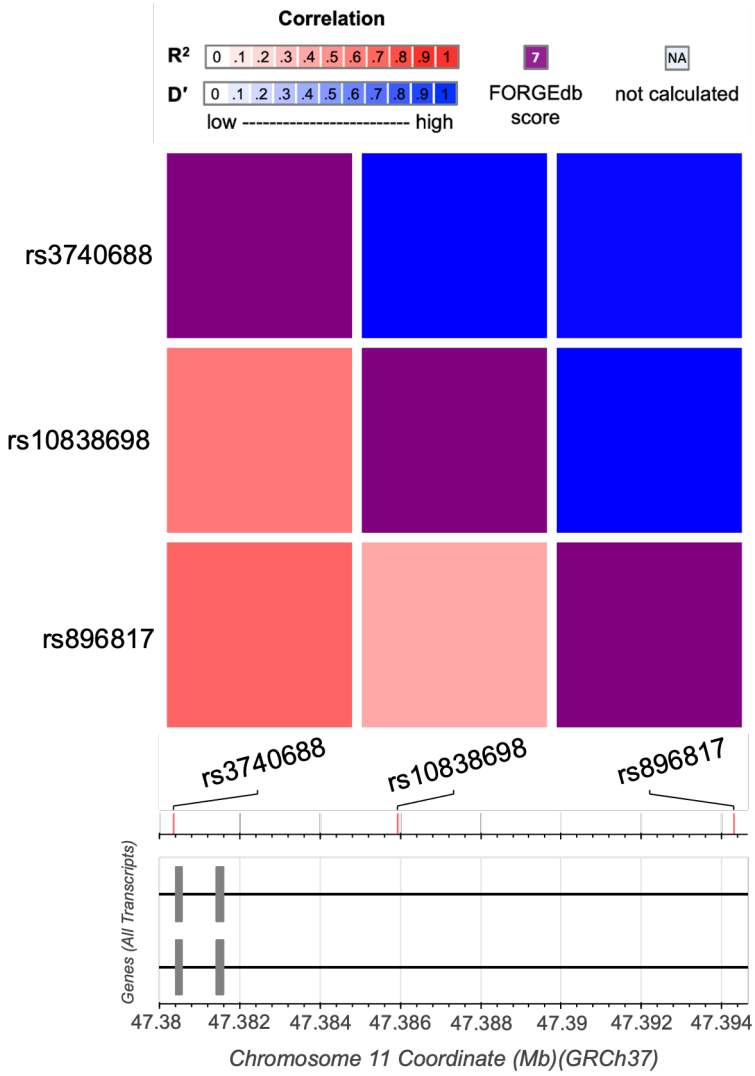

Figure S24. LD matrix showing linkage disequilibrium value between teQTLs rs3740688 and another two teQTLs rs10838698 and rs896817, Related to Figure 6.