Description of Additional Supplementary Files

File Name: Supplementary Data 1

Description: This supplementary file contains information related to the study participants, sample-level and variant-level quality control metrics, and a summary of the whole-genome sequencing dataset.

Table S1A: Characteristics of the study participants.

Table S1B: Quality check for samples and varaints.

Table S1C: Summary of whole genome sequencing study on Alzheimer's disease subjects.

File Name: Supplementary Data 2

Description: This supplementary file provides the results from single-variant and gene-based association analyses related to Alzheimer's disease. It also includes gene prioritization based on eQTL evidence, functional annotations, and literature support.

Table S2A: Quality check metrics of lead SNPs

Table S2B: Summary statistics of gene-based association analysis on clinical diagnosis, two-sided Firth logistic regression was performed using REGENIE, with no correction for multiple testing. (Norminal p value < 0.05)

Table S2C: Summary statistics of gene-based association analysis on amyloid beta, two-sided Firth logistic regression was performed using REGENIE, with no correction for multiple testing. (Norminal p value < 0.05)

Table S2D: Rare coding variants in DRC7 gene (AF<1%)

Table S2E: Summary of colocalizaed variants within the 95% credible set

Table S2F: eQTLs of lead variants

Table S2G: Number of publications resulting from the search for prioritized genes with Alzheimer's or brain

Table S2H: Peak2gene of lead variants from Xiong et al, 2023,cell

Table S2I: Summary of differential expression for prioritized genes based on final consensus cognitive diagnosis from Mathys et al., 2023, Cell. Two-sided statistical tests were performed with multiple testing correction applied locally (p_adj.loc). (|log2FC| > 0.25, p_adj.loc < 0.05)

File Name: Supplementary Data 3

Description: This supplementary file includes the data and results used in the association analysis of rare noncoding variants related to Alzheimer's Disease.

Table S3A: Annotation terms used in association test.

Table S3B: All categories from rare variants and single category enrichment test result.

Table S3C: Comparison results of the significant count for each category by genomic region.

Table S3D: Clusters from network analysis.

Table S3E: Variants with annoation information from five risk clusters.

Table S3F: Excitatory CRE variants with 3D interaction genes from Cluster 25.

File Name: Supplementary Data 4

Description: This supplementary file includes the data and results from the association analysis of structural variants related to Alzheimer's Disease.

Table S4A: Copy number variation association analysis ($P < 0.05$). Two-sided logistic regression was used to assess associations, adjusting for age, sex, and the first three principal components.

Table S4B: Model 1: length test results ($P < 0.05$). Two-sided logistic regression was performed with sample-specific and technical covariates, including APOE ε4 carrier status.

Table S4C: Model 2: threshold test results ($P < 0.05$). Two-tailed Fisher's exact test was applied to evaluate category-specific enrichment.

Table S4D: Model 3: outlier test results.

Table S4E: Linkage disequilibrium analysis between STR outliers and GWAS SNPs.

Table S4F: Gene set enrichment test for STR outliers observed in samples carrying 40 or more outliers. Two-sided Fisher's exact test was used, with multiple testing correction applied using the false discovery rate (FDR) method.