

Identifying Causal Genes That Drive Alzheimer's Disease Progression



Xizhi Xu¹, Yuriko Katsumata^{2,5}, Lincoln M.P. Shade¹, Xian Wu^{2,5}, Khine Zin Aung^{2,5}, Peter T. Nelson,^{3,5} Erin L. Abner^{4,5}, David W. Fardo^{1,5}

¹Department of Biostatistics, ²Division of Biomedical Informatics, ³Department of Pathology, ⁴Department of Epidemiology and Environmental Health, ⁵Sanders-Brown Center on Aging. University of Kentucky.

Background

- Alzheimer's disease (AD) is biologically complex and can be misdiagnosed clinically.
- Neuropathological endophenotypes (NPEs) derived from brain autopsies provide more direct insight into disease mechanisms.
- Braak stage is a key AD-related NPE that indicates burden and location of neurofibrillary tau tangle progression in the brain.
- Genetics studies, including genome-wide association studies (GWAS), have identified many variants associated with AD [cite Bellenguez 2022], but identifying truly causal genes remains challenging.
- Existing methods (e.g., transcriptome-wide association study [TWAS]) are likely confounded by linkage disequilibrium

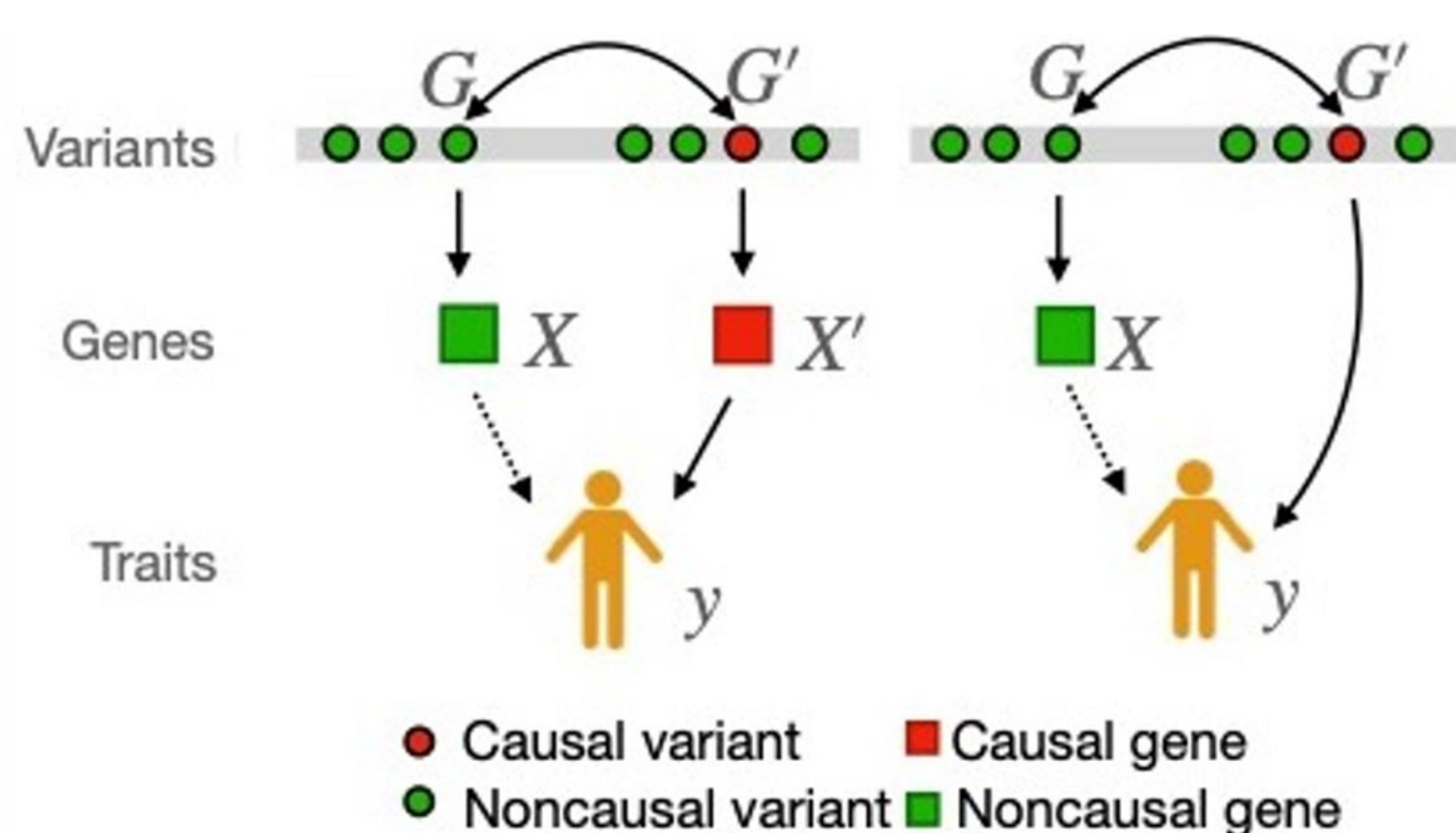


Fig 1. Schematic representing standard TWAS.

Hypothesis & Aims

- Hypothesis:** Genetic regulation influences tangle progression and can reveal key AD-related genes.
- Aim 1:** Identify likely causal genes associated with tangle progression, operationalized by Braak staging, using causal TWAS (cTWAS).
- Aim 2:** Compare performance with standard TWAS methods.

Methods

- The causal-TWAS (cTWAS) method jointly models all genes and variants in a genomic region.

$$Y = \underbrace{\sum_{j=1}^J \tilde{X}_j \beta_j}_{\text{gene effects}} + \underbrace{\sum_{m=1}^M G_m \theta_m}_{\text{variant effects}} + \epsilon$$

Y: latent score for Braak staging severity of an individual.
 ϵ : error term $\sim N(0, \sigma^2)$.

- Key intuition:** causal signals in a genomic region are likely sparse.
- Motivation:** Bayesian fine-mapping.
- Application** to Braak:
 - GWAS summary statistics. N = 7804 autopsied participants from multiple sources: including the National Alzheimer's Coordinating Center (NACC), the Religious Orders Study and Rush Memory and Aging Project (ROSMAP), and the Adult Changes in Thought (ACT) study.
 - Gene expression and splicing prediction models derived from the Genotype-Tissue Expression (GTEx). All available brain regions (total = 26) were included.
 - Reference data: variants information and LD information.

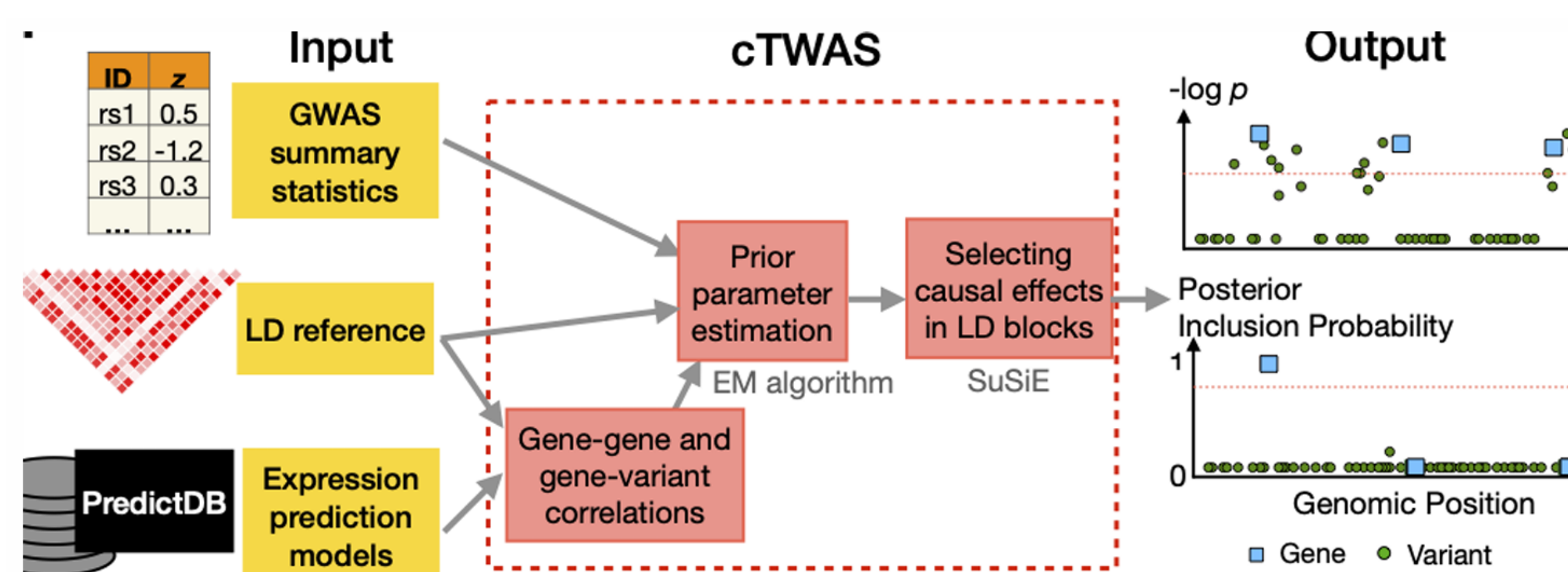


Fig 2. The overall workflow of cTWAS. (Reprinted from Zhao et al., 2024)

Results

- Among ~17k available genes for testing, 160 genes remained after integrating evidence from expression and splicing data. The vast majority showed little evidence of association, with 97% of genes having posterior inclusion probabilities (PIPs) below 0.1.
- We are interested in 2 genes with PIP > 0.8: *YBEY* and *PIK3R5*.

Gene name	Region ID	Combined PIP	Expression PIP	Splicing PIP
YBEY	21_46072312_46699840	1	0.0006	1
PIK3R5	17_8403107_10062604	0.9621	0.9621	< 0.0001
PCNT	21_46072312_46699840	0.6473	0.647	0.0003
CEACAM19	NA	0.0498	0.0051	0.0447
ZNF229	NA	0.005	0.005	< 0.0001

Table 1. Prioritized genes from fine-mapping results.

Results

- We also created gene locus plots for visualization. For example, *PIK3R5*:

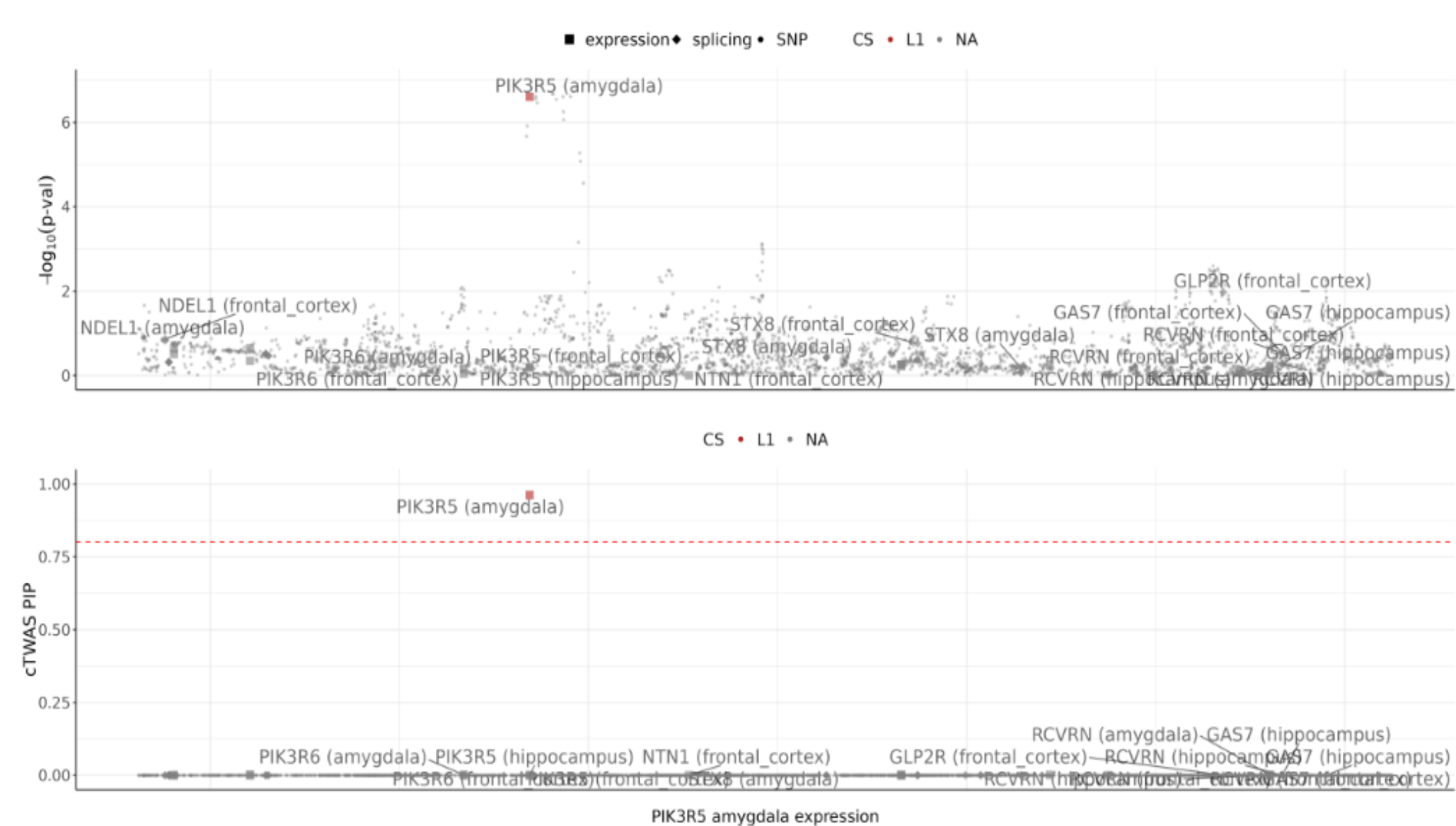


Fig 3. Gene locus plot for *PIK3R5*. Top: standard TWAS plot, where x-axis represents genomic position and y-axis represents $-\log_{10}(p)$. Bottom: cTWAS plot, where x-axis represents genomic position and y-axis represents PIPs.

Discussions

- YBEY*- and *PIK3R5*-related biological mechanisms are plausibly associate with progression of tau pathology.
- YBEY*: Involved in mitochondrial function, a key pathway in early neurodegeneration.
- PIK3R5*: Regulates the PI3Ky pathway, highly active in microglia and linked to immune dysfunction in dementia.
- Limitations:** current analysis was limited to GTEx data. Future work: use a different cohort and prediction model data to perform the replication study.

References

- Siming Zhao, et al. (2024). Adjusting for genetic confounders in transcriptome-wide association studies improves discovery of risk genes of complex traits. Nature genetics. s41588-023-01648-9.
- Lincoln M. P. Shade, et al. (2024). GWAS of multiple neuropathology endophenotypes identifies new risk loci and provides insights into the genetic risk of dementia. Nature genetics. s41588-024-01939-9.