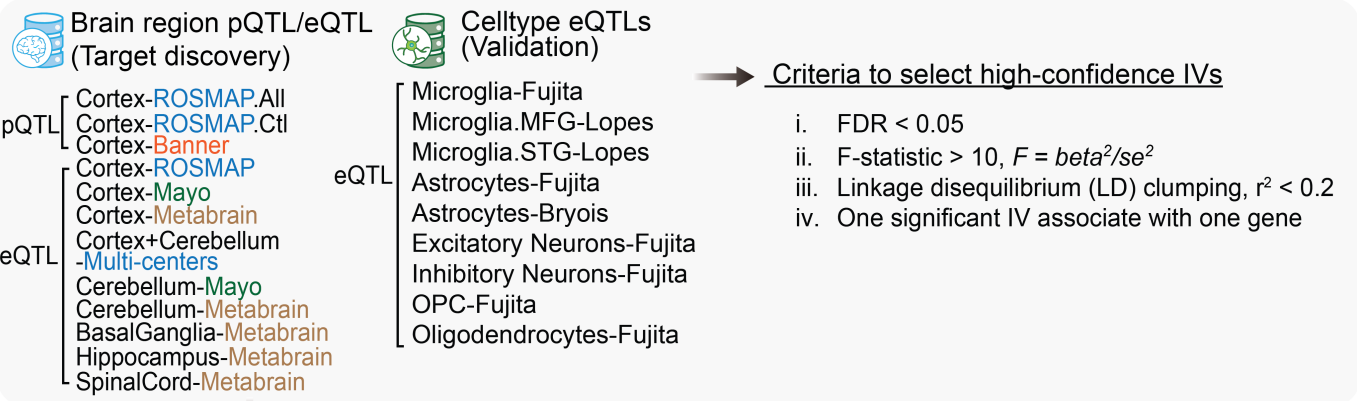


Workflow of Drug Target Mendelian Randomization (MR) in Alzheimer's Disease

1)

Instrument Variables (IVs)



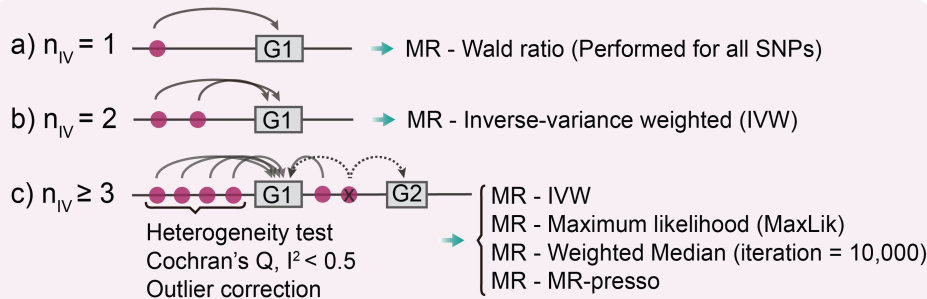
Exposure



Outcomes



MR approaches



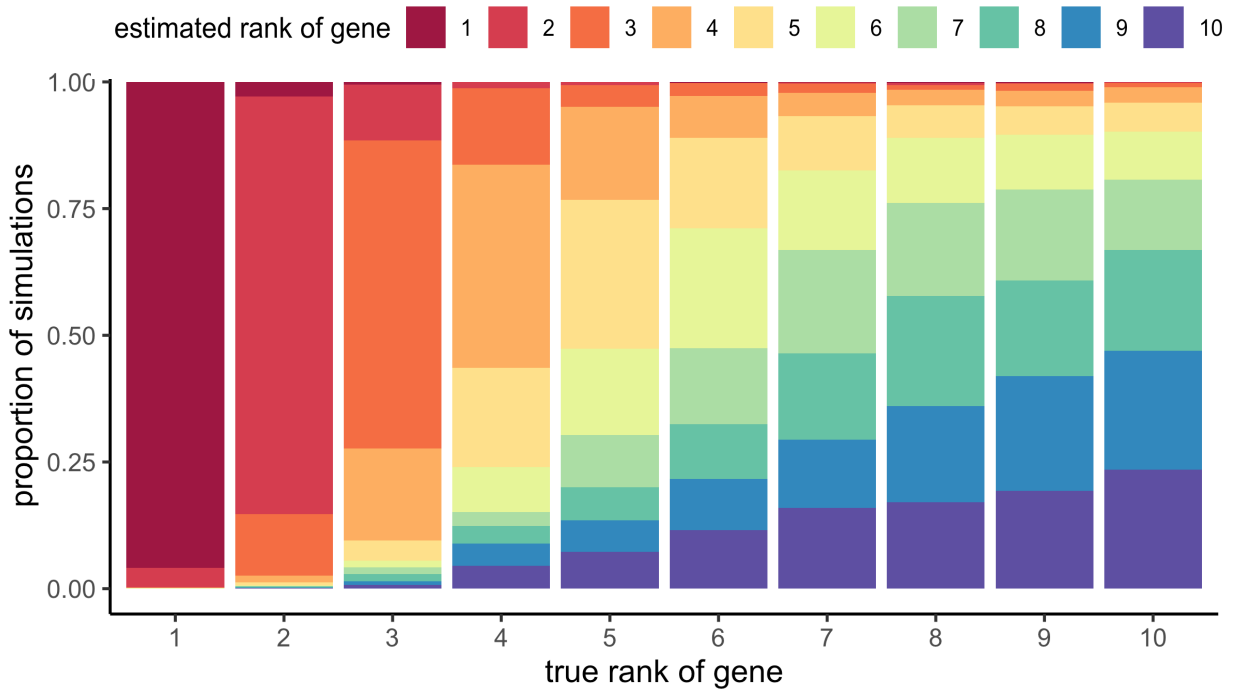
alzMR-score

$$\text{alzMR}(g) = \sum_j \sum_k Z_{jkg}$$

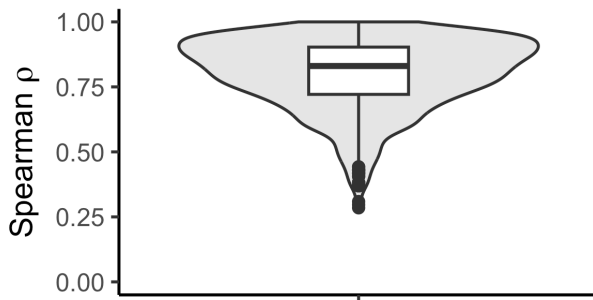
Z, Z-statistic
 $k = 1, \dots, K$ (kth QTLs)
 $j = 1, \dots, J$ (jth AD-GWAS)
 $g = 1, \dots, G$ (gth target)

Supplementary Fig. 1 The workflow of MR analysis. The process begins with the selection of instrumental variables (IVs) from pQTL and eQTL datasets across multiple human brain biobanks. Exposure selection focused on 1,229 druggable targets characterized by strong binding affinities (gray background). The analysis utilizes seven AD GWAS datasets, applying various MR models according to the number of IVs (light red background). The alzMR scores were designed by Z-statistics to prioritize the top targets for AD (blue background).

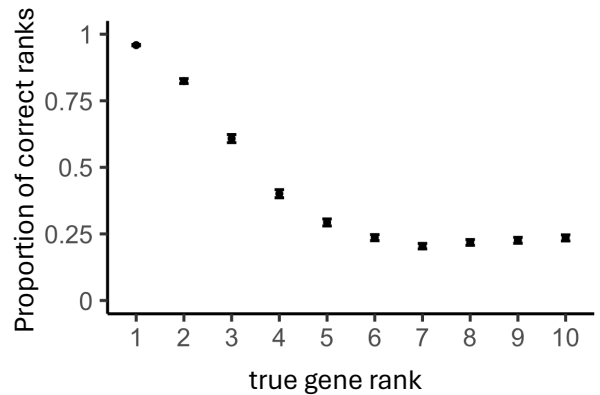
a) Performance of gene ranks using alzMR scores



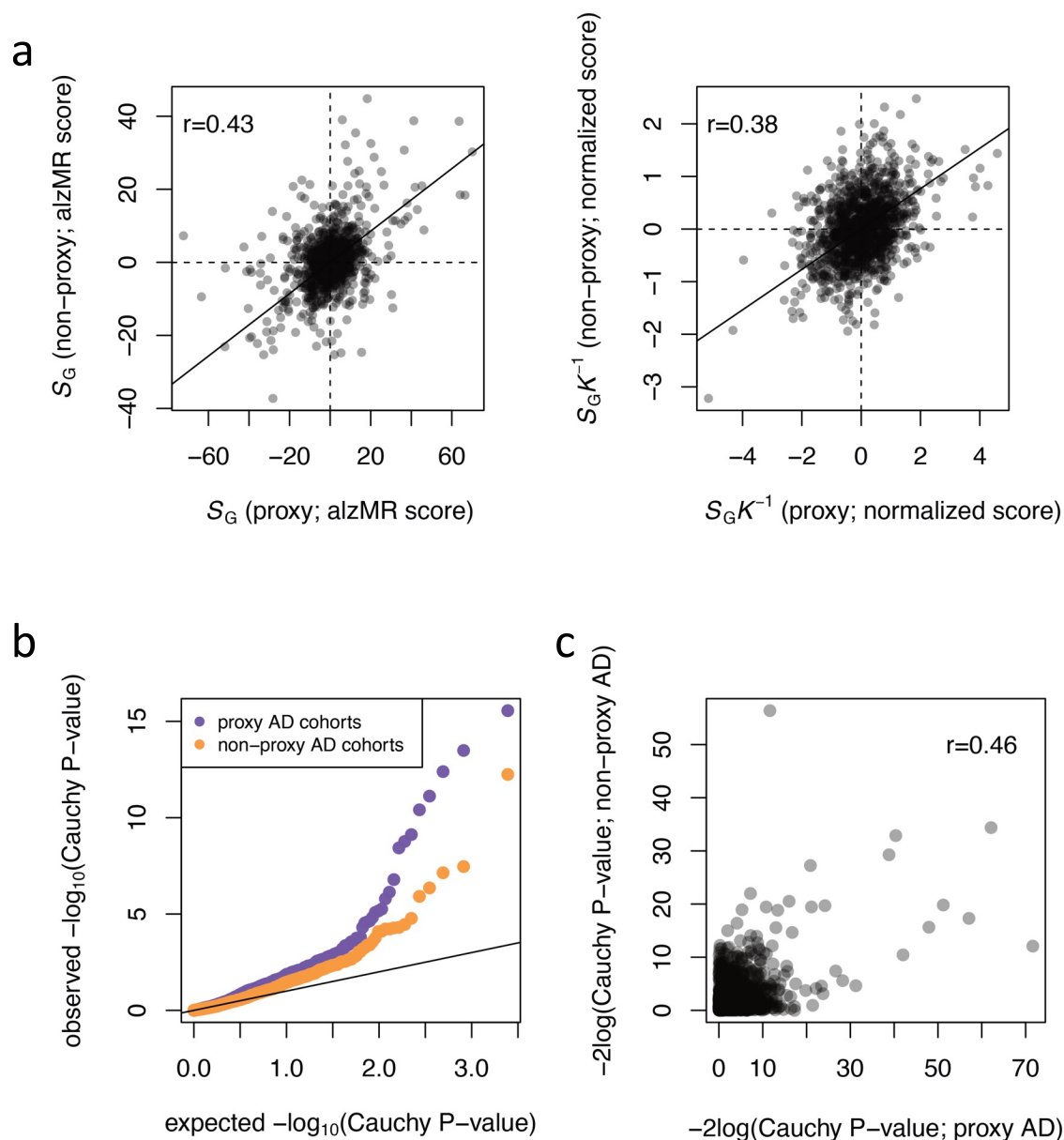
b) Correlation between estimated and true ranks



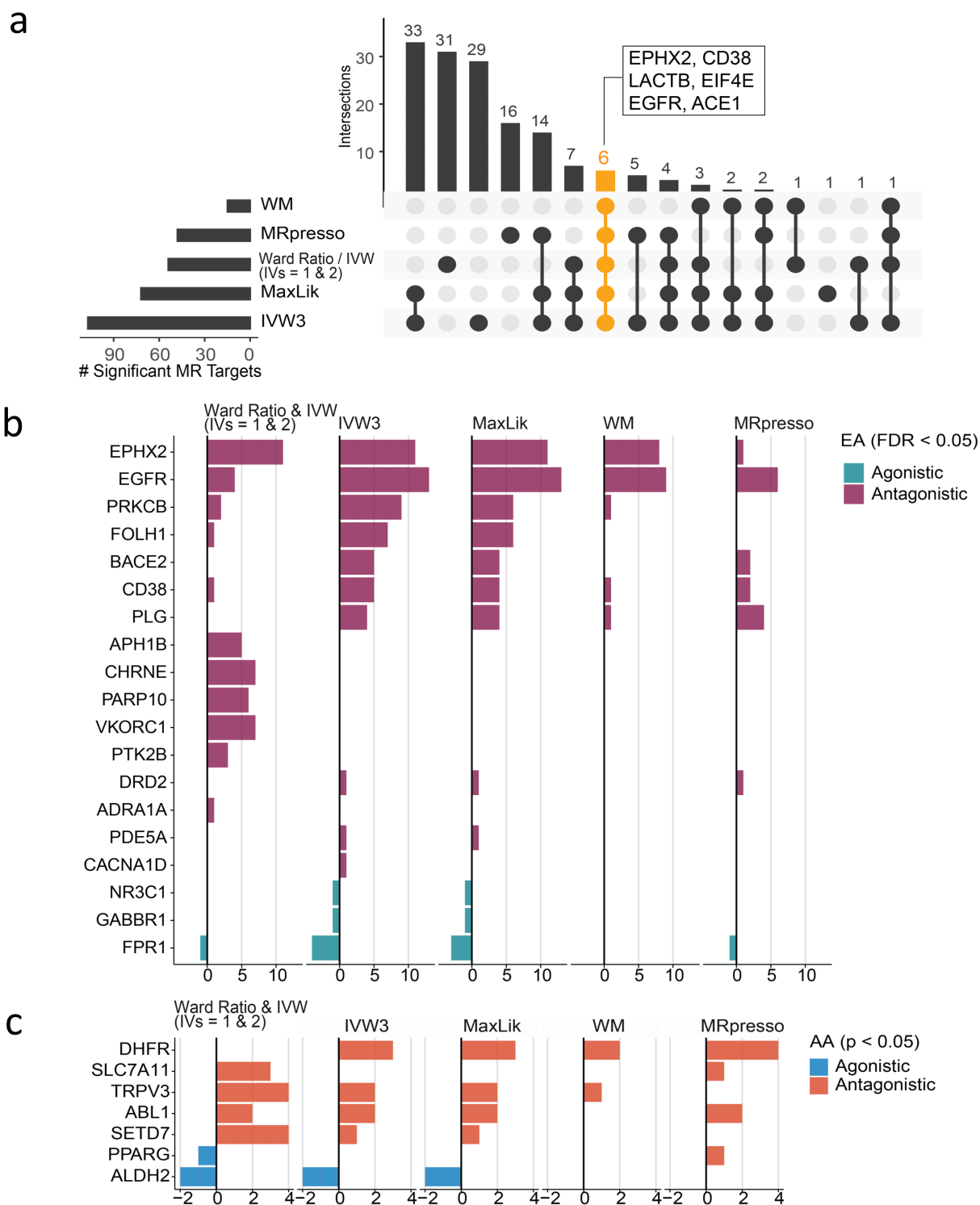
c) Distribution of correctly estimated gene ranks



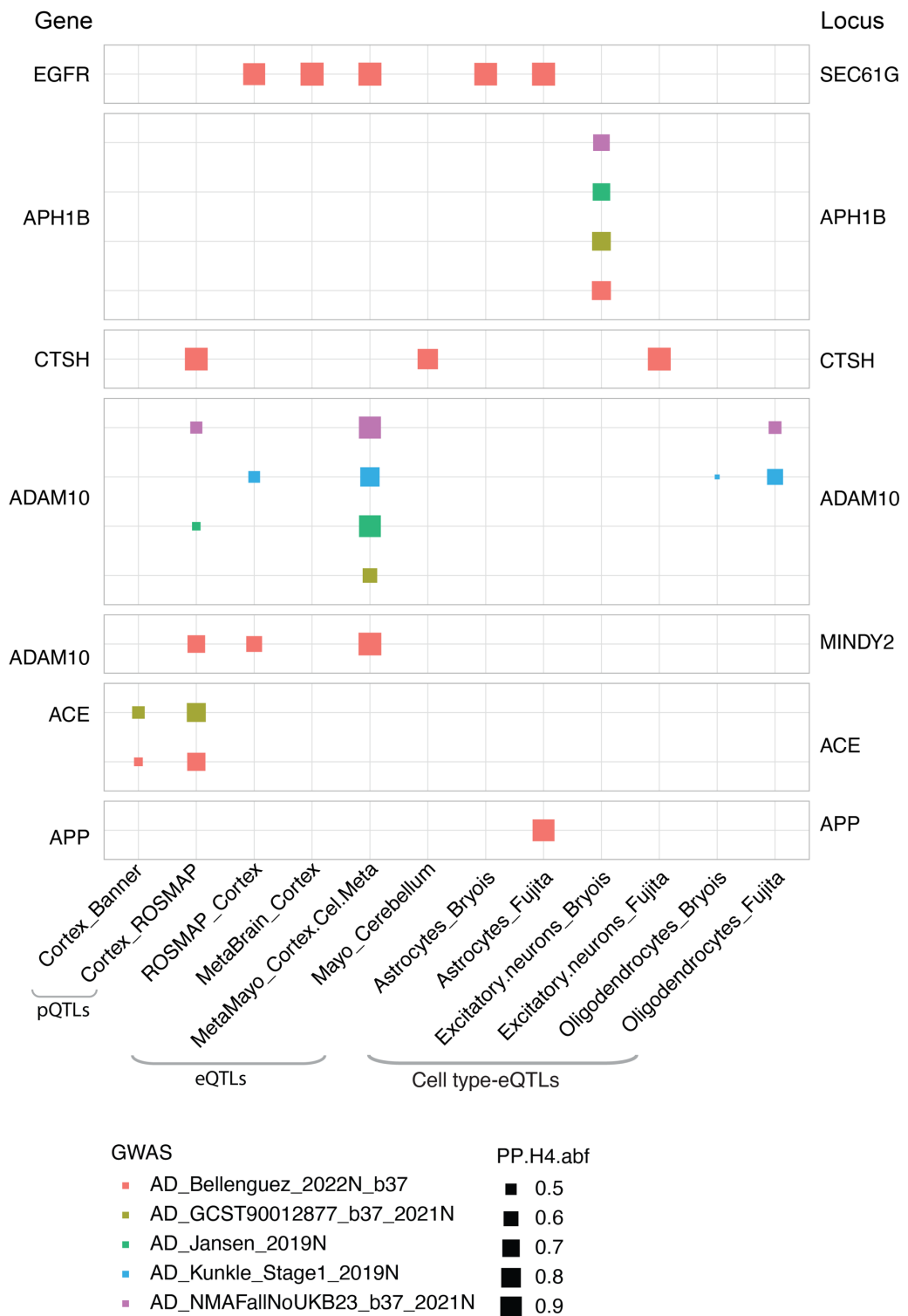
Supplementary Fig. 2 Assessment of gene ranking performance and correlation based on simulations. a. Performance of gene ranking using Z-statistic-based alzMR scores. The heatmap illustrates the proportion of correct estimations (y-axis) for each true rank (x-axis) of a gene, with color intensity indicating the estimated rank. b. Spearman correlation between estimated and true ranks. c. Distribution of correctly estimated gene ranks. The y-axis displays the proportion of correctly estimated ranks, while the x-axis shows the true rank of genes.



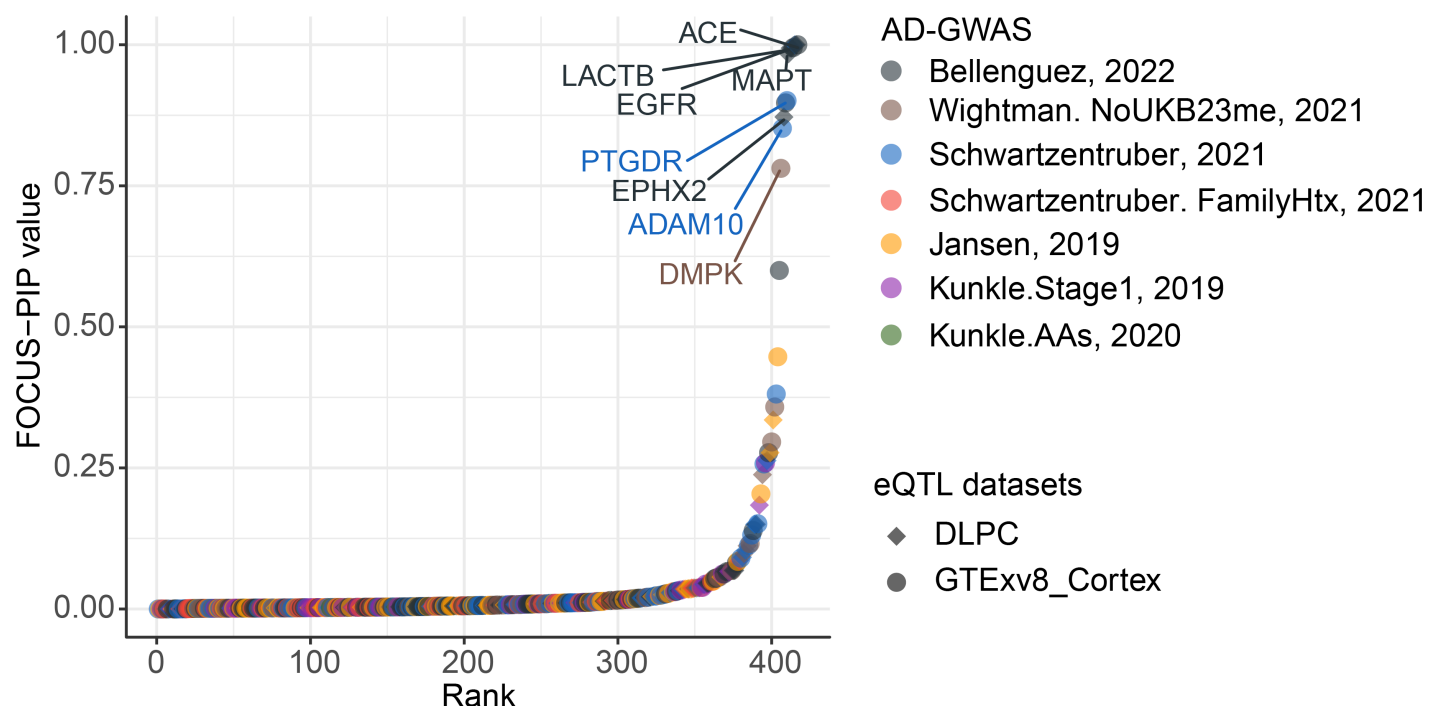
Supplementary Fig. 3 Sensitivity analysis between clinically diagnosed AD and proxy-AD cohorts. **a.** Correlation of the alzMR score (left panel) and the normalized score (right panel) between proxy and non-proxy AD cohorts. **b.** QQ plot compares observed and expected Cauchy P-value for proxy (purple) and non-proxy (orange) cohorts. **c.** The correlation between Cauchy P-values for proxy and non-proxy cohorts.



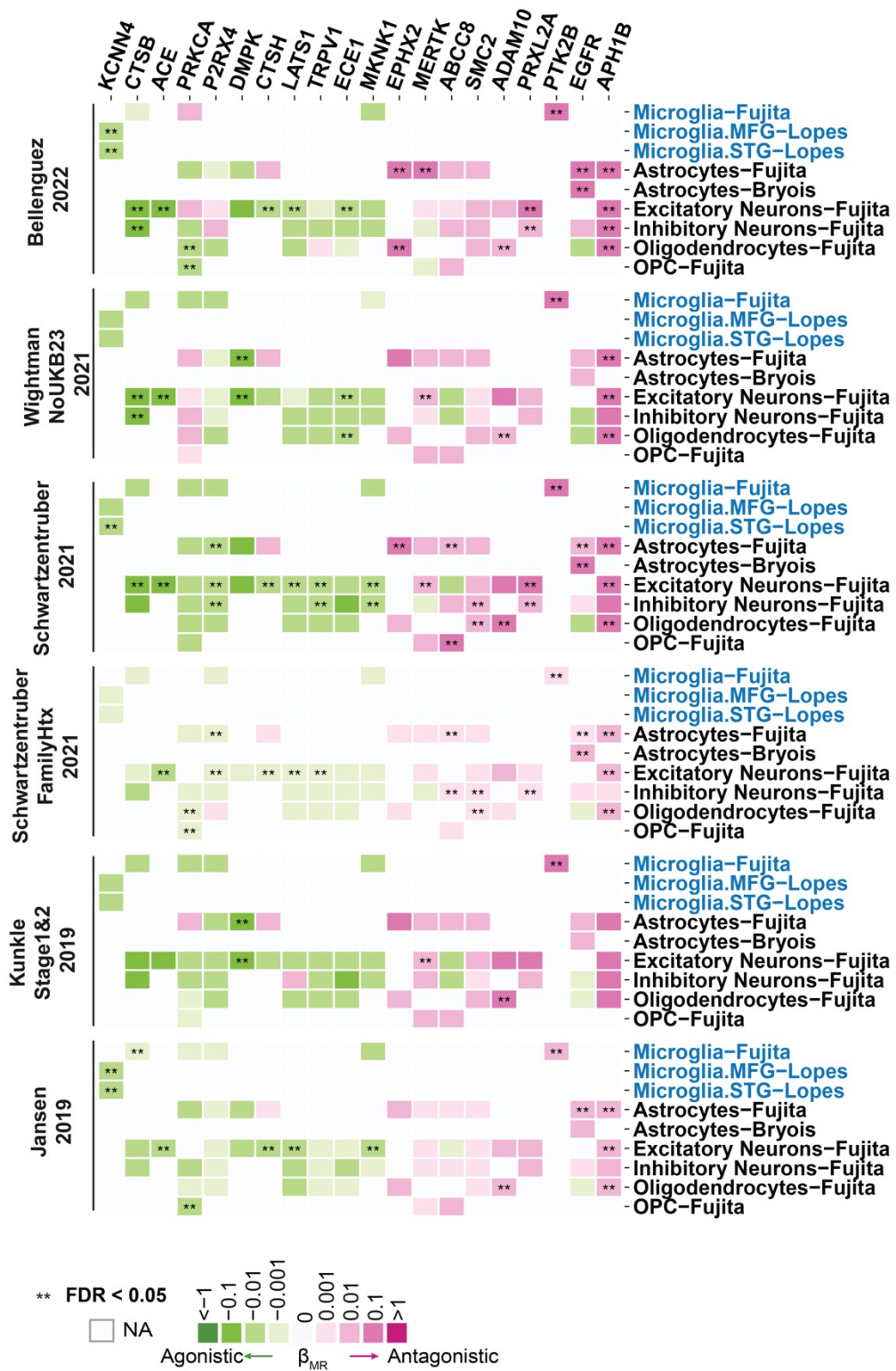
Supplementary Fig. 4 Summary of results across five Mendelian randomization (MR) methods. a. The upset plot showing the intersections of significant drug targets among 5 MR models. **b.** The bar plot shows alzM score for 19 drug targets across 5 MR models for European ancestry (EA). **c.** The bar plot shows alzM score for 7 drug targets across 5 MR models for African ancestry (AA).



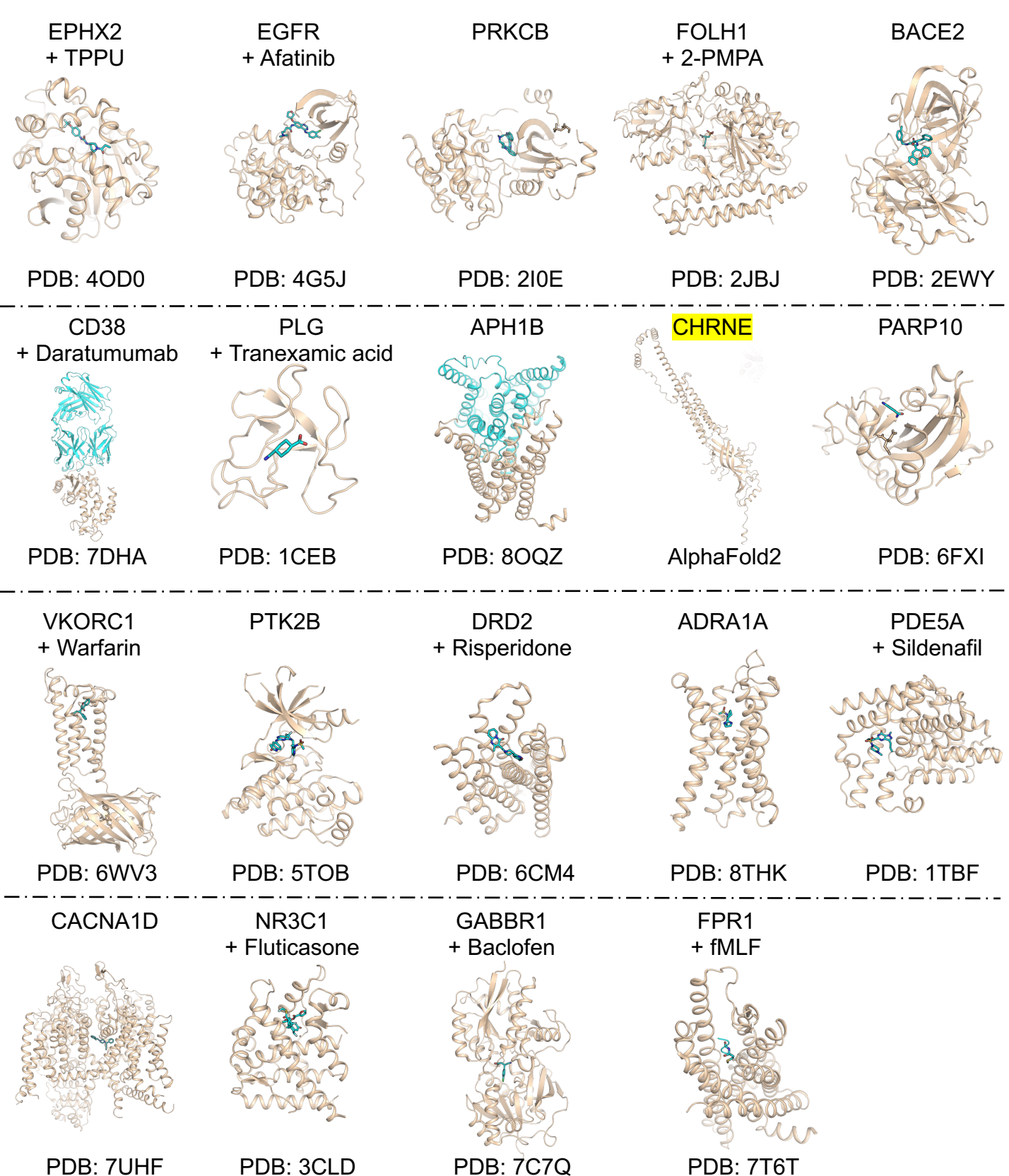
Supplementary Fig. 5 Colocalization analysis between 5 AD GWAS and brain pQTLs and eQTLs. Different colors denote the 5 AD GWAS datasets. The size of the square represents the values of posterior probabilities (PP). H4.abf.



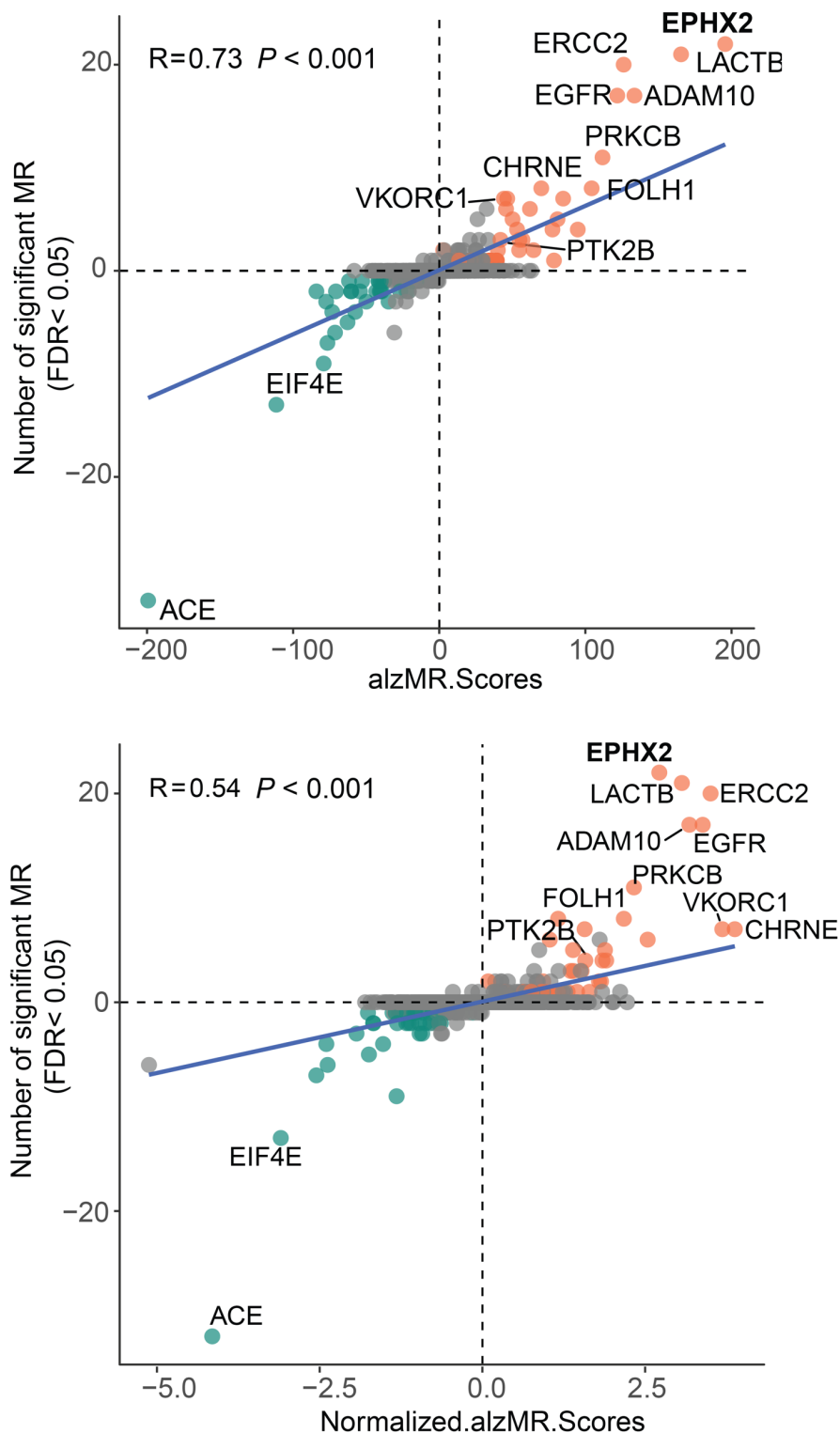
Supplementary Fig. 6 Fine-mapping analysis of druggable targets in AD. The graph showed the fine-mapping causal probability (FOCUS-PIP) values of various druggable gene targets for AD across different datasets. Each point represents a gene. The color for each point denotes one AD-GWAS dataset, and different shapes representing eQTL datasets. The plot highlighted the label text exclusively for the data point with the highest PIP value across the 7 AD GWAS datasets.



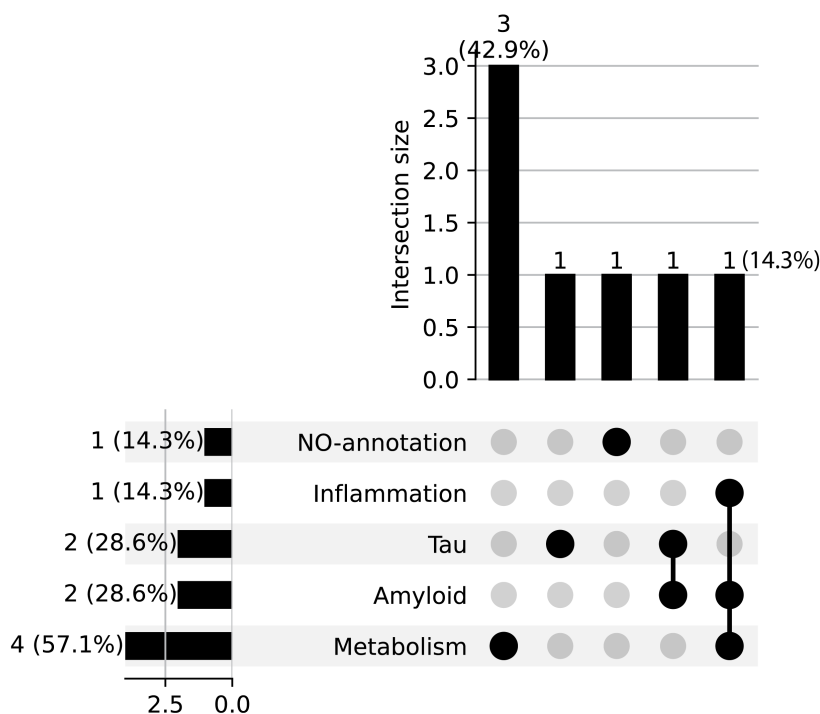
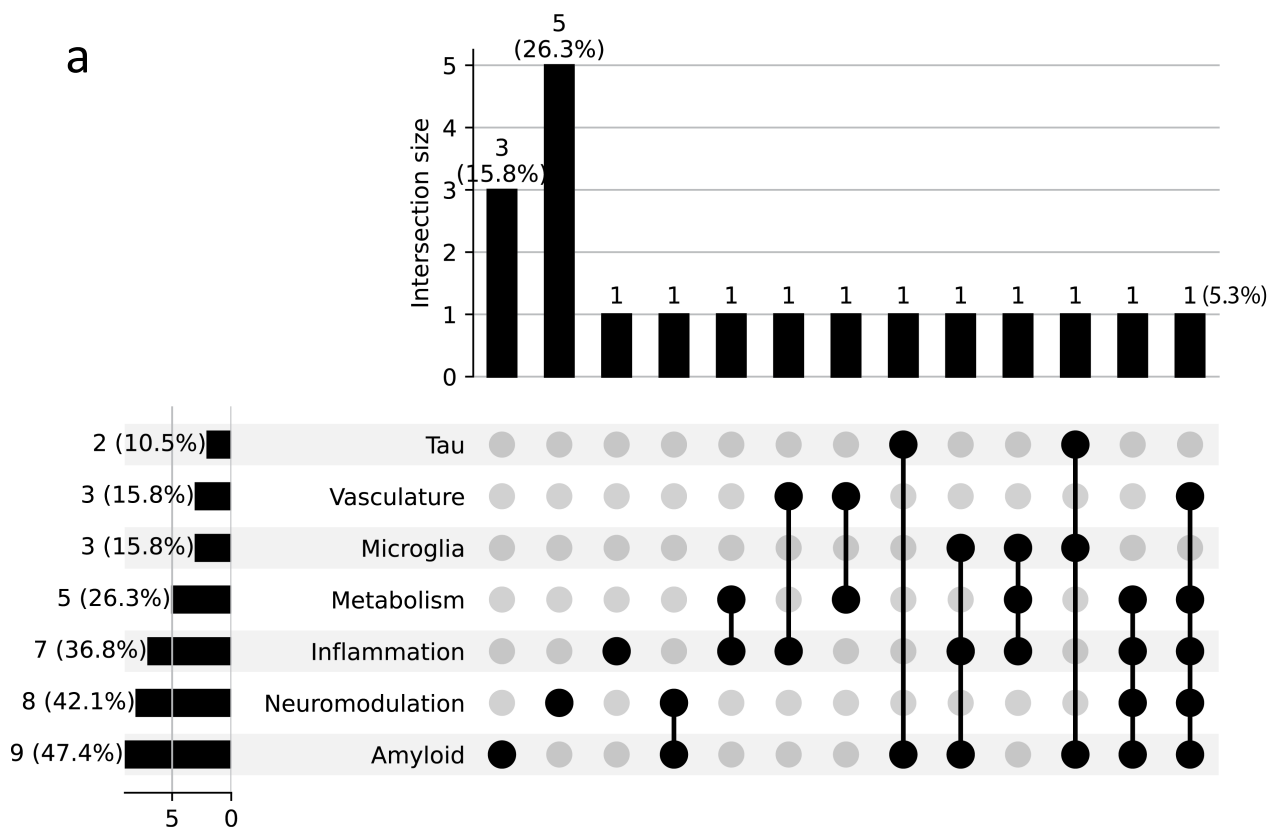
Supplementary Fig. 7 Cell-type specific MR results across 6 AD GWAS datasets with EA.



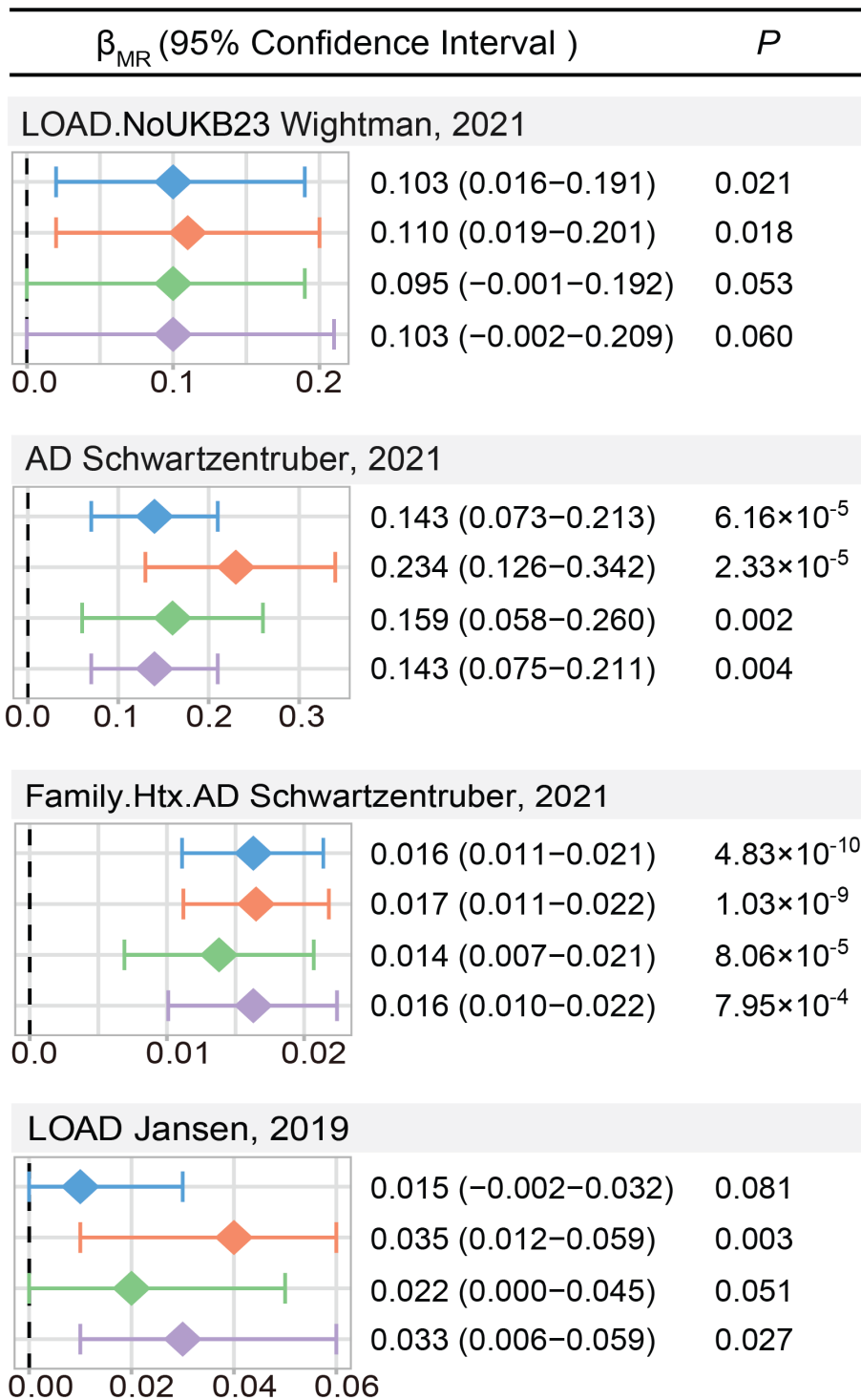
Supplementary Fig. 8. This presents the selected structural complexes of drug targets, highlighting potential druggability for identified targets in EA. The structure complexes show the targets (wheat color) and their corresponding drugs (cyan color) that highlighted in **Figure 3c**.



Supplementary Fig. 9 The correlation between the alzMR score (upper panel), normalized alzMR score (lower panel) and the number of significant MR results at $FDR < 0.05$. The points denote 1,229 druggable targets for AD. Red points denote the antagonistic targets and green point denote agonistic targets.

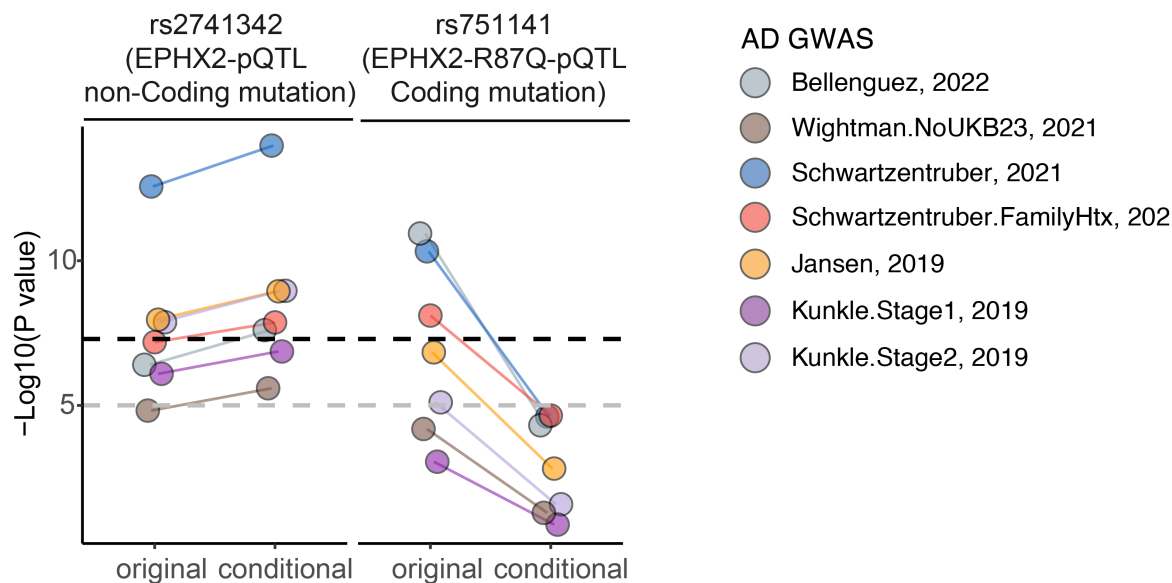


Supplementary Fig. 10 The percentage distribution of 8 mechanism-of-action (MoAs). The upset plot showing the intersections of drug targets among MoAs in (a) European ancestry (EA) and (b) African ancestry (AA).

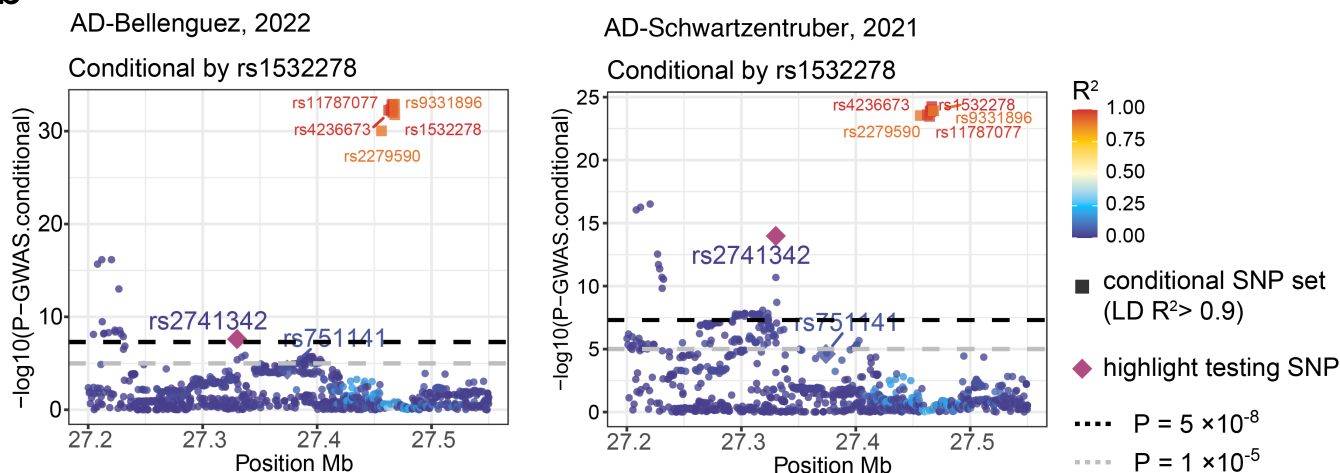


Supplementary Fig. 11 MR results for EPHX2 are validated by four MR methods in human cortex region across four AD GWAS datasets. Seven Instrumental variables (IVs) were extracted from the ROSMAP pQTL dataset.

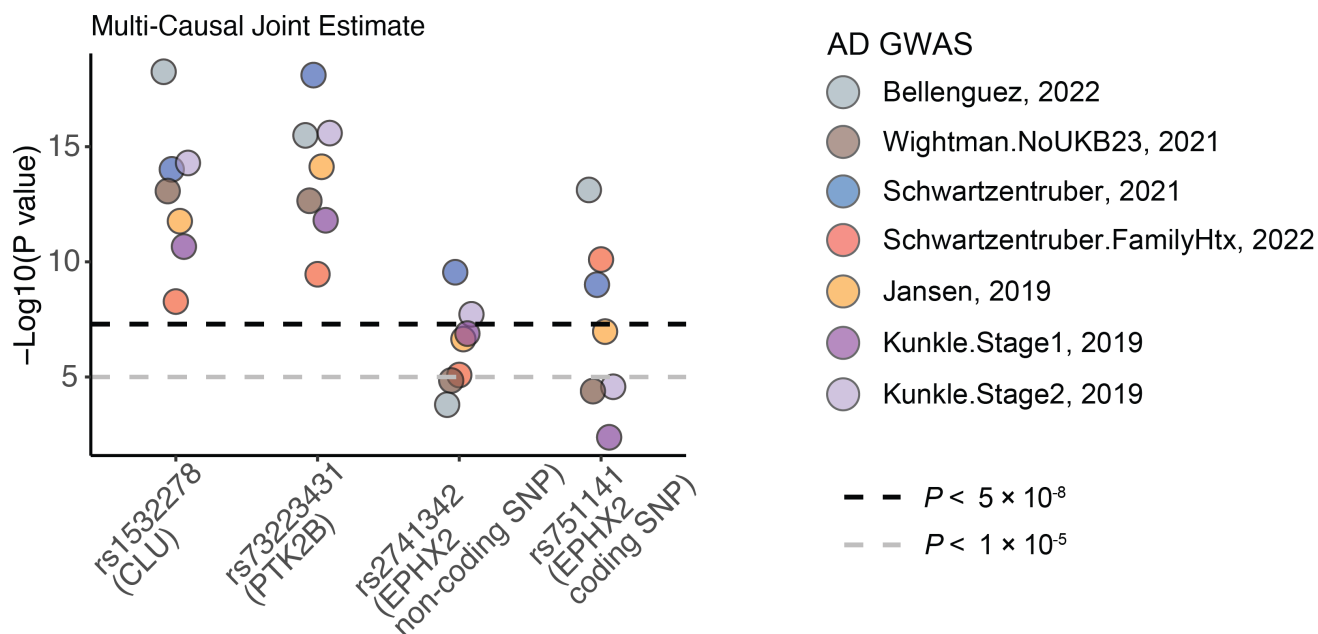
a Conditional SNP-rs1532278 (CLU locus)



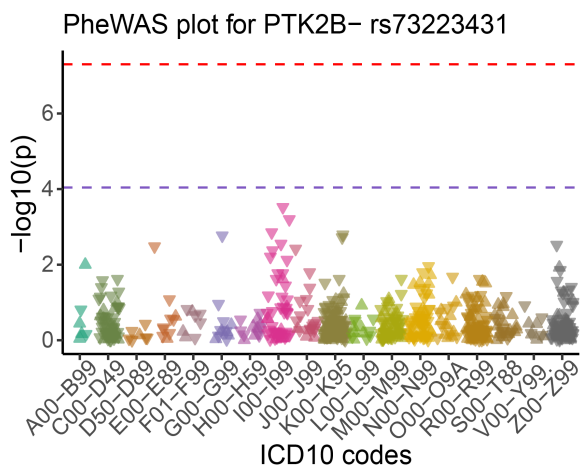
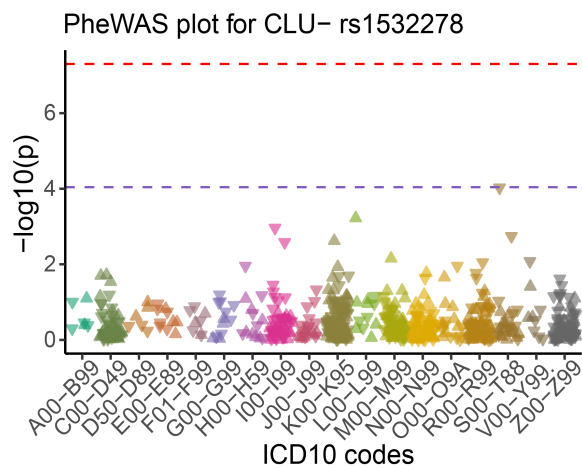
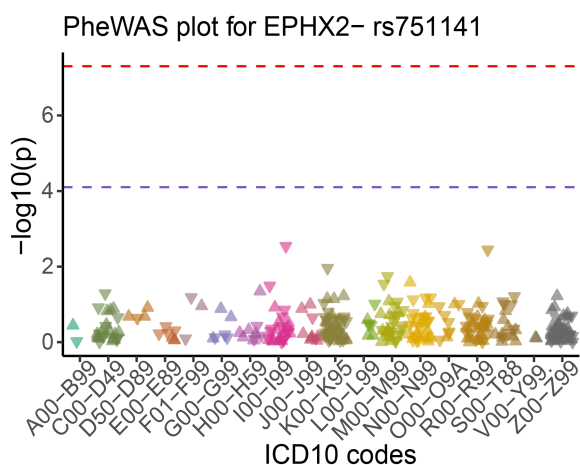
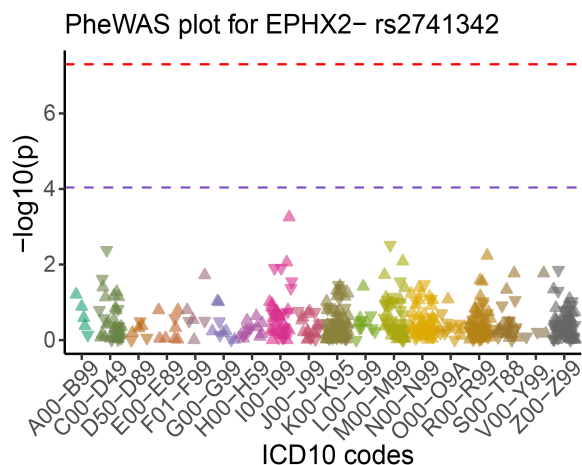
b



Supplementary Fig. 12 Conditional analysis of EPHX2 pQTL mutations. **a** The conditional SNP is rs1532278, which is the lead-SNP associated with CLU. "Original" refers to the $-\log_{10}(\text{p-value})$ from the originally published GWAS summary statistics data; "conditional" refers to the $-\log_{10}(\text{p-value})$ after the conditional analysis using SNP rs1532278. **b** The scatter plot showed conditional P values for two GWAS datasets. We highlight a set of lead SNPs (identified in multiple publications) at the CLU locus, which were used as conditional SNPs for analysis.



Supplementary Fig. 13 Multi-causal joint estimate of four SNPs at CLU locus. The colored circles denote the 7 AD GWAS datasets.

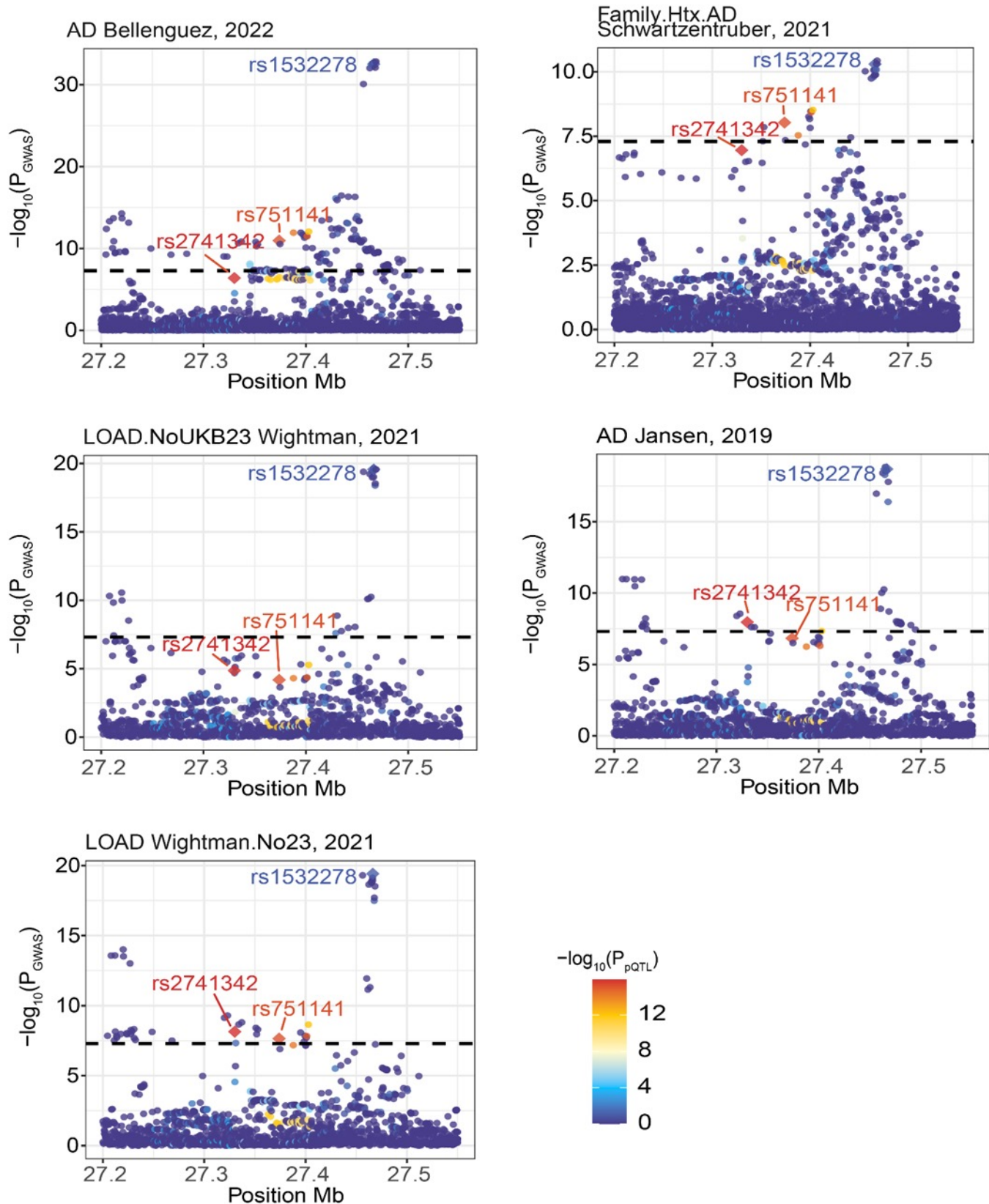


Diagnoses - main ICD10 codes of UK-Biobank

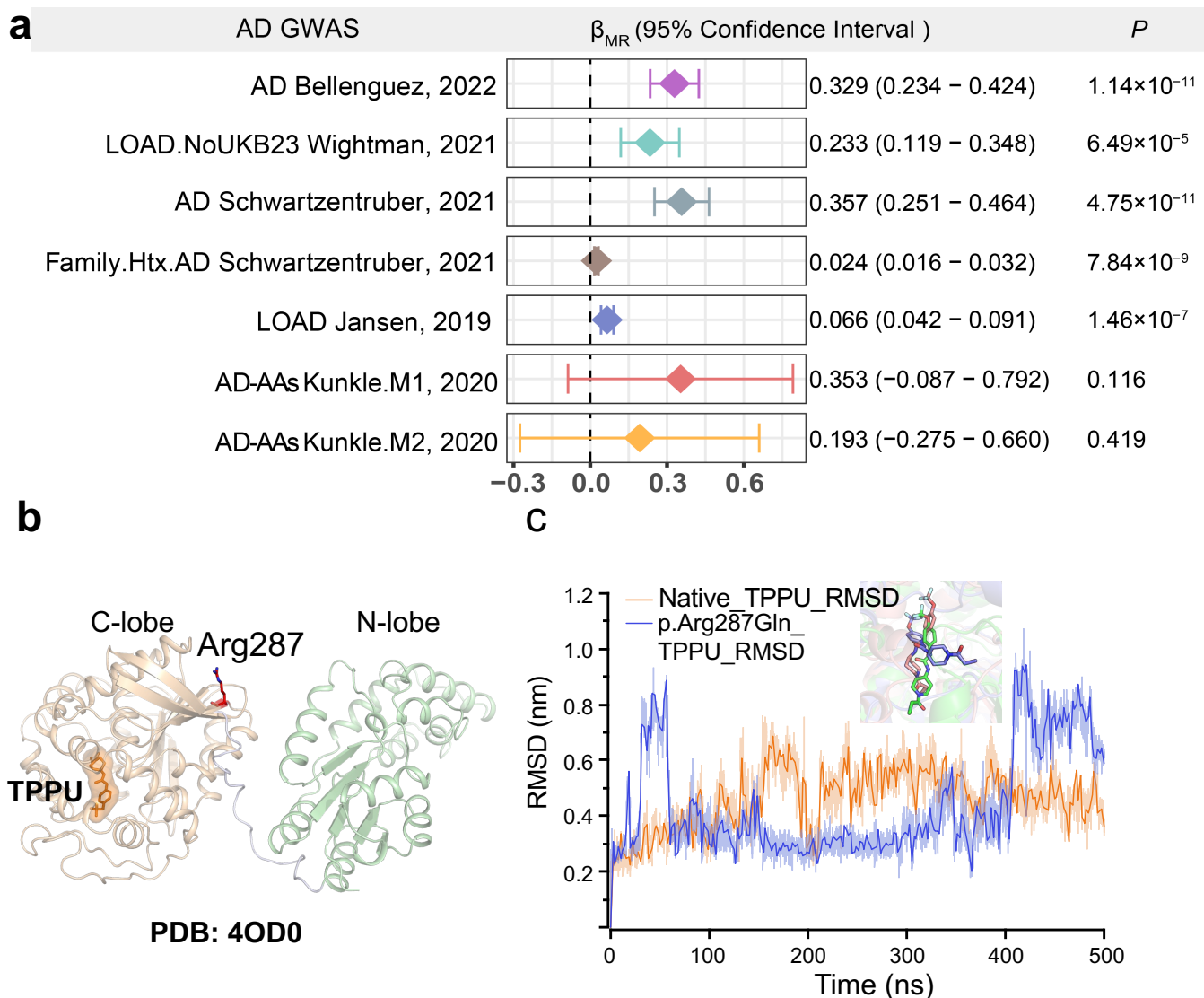
- A00-B99, Certain infectious and parasitic diseases
- C00-D49, Neoplasms
- D50-D89, Diseases of the blood and blood-forming organs
- E00-E89, Endocrine, nutritional and metabolic diseases
- F01-F99, Mental, Behavioral and Neurodevelopmental disorders
- G00-G99, Nervous system
- H00-H59, Eye and adnexa
- I00-I99, Circulatory system
- J00-J99, Respiratory system
- K00-K95, Digestive system
- L00-L99, Skin and subcutaneous tissue
- M00-M99, Musculoskeletal system and connective tissue
- N00-N99, Genitourinary system
- O00-O9A, Pregnancy, childbirth and the puerperium
- R00-R99, Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
- S00-T88, Injury, poisoning and certain other consequences of external causes
- V00-Y99, External causes of morbidity
- Z00-Z99, Factors influencing health status and contact with health services

- △ $\beta_{\text{PheWAS}} > 0$
 ▽ $\beta_{\text{PheWAS}} < 0$

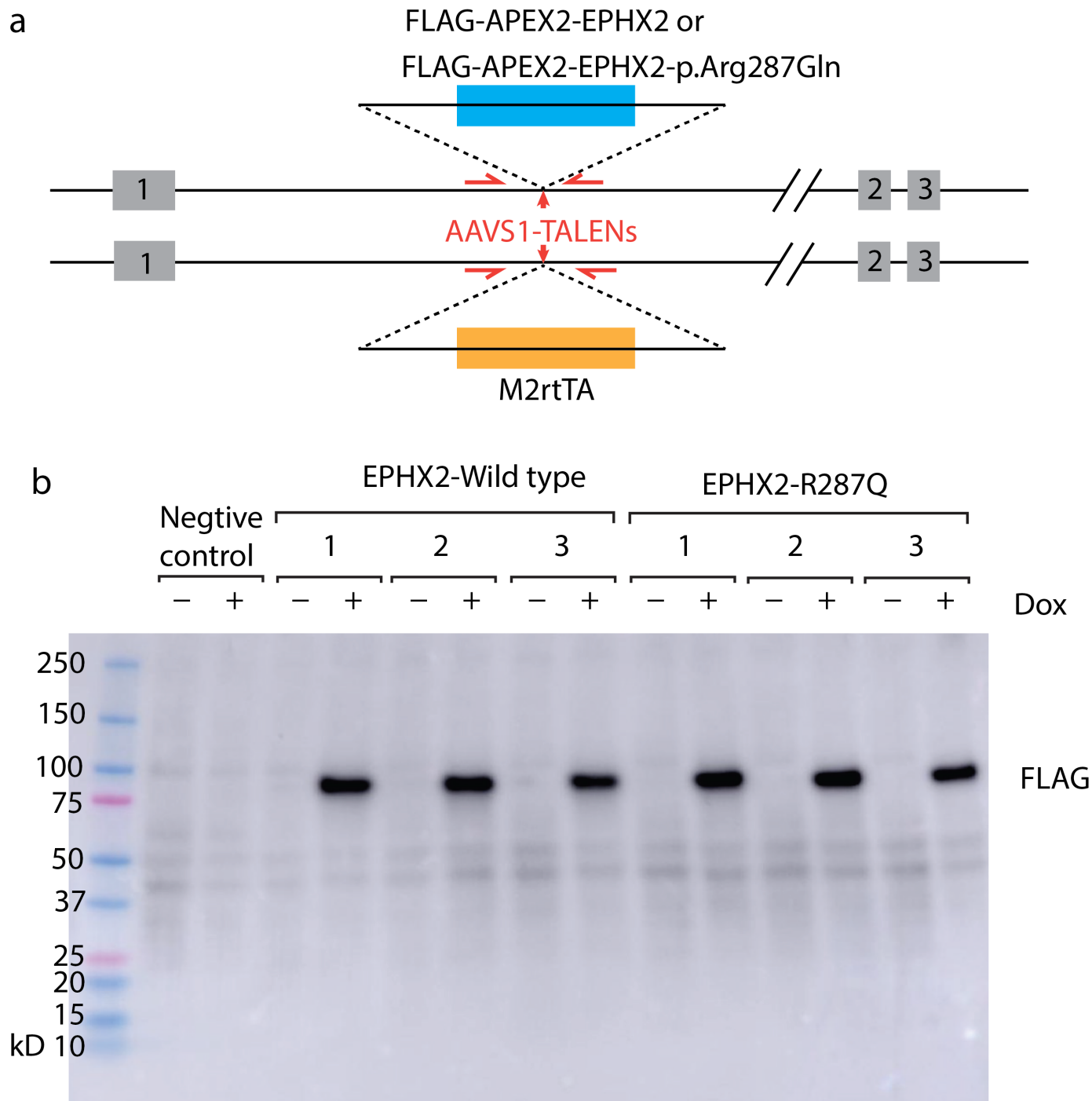
Supplementary Fig. 14 The PheWAS results for EPHX2 IVs rs2741342 and rs751141 are presented. The data points represent 624 human diseases based on ICD10 codes from UK-biobank cohort with colors denoting different ICD10 categories. Up-triangles represent a beta coefficient of PheWAS > 0 , indicating a positive association between EPHX2 IV and a specific disease; the down-triangles denote the beta coefficient of PheWAS < 0 , indicating a negative association between EPHX2 IV and a specific disease. The purple dashed line shows a significant cutoff at $\text{FDR} < 0.05$ ($-\log_{10}(P) = 4.08$); and red dash line showed the genome-wide significant cutoff at $5e-8$.



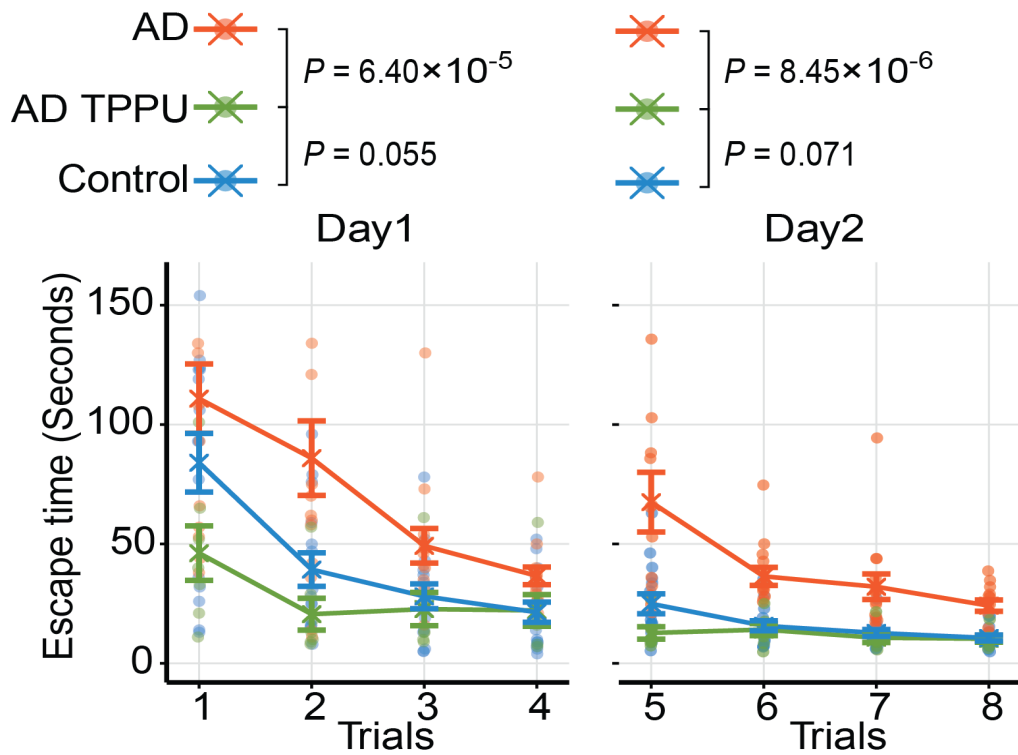
Supplementary Fig. 15 Fine-mapping of lead SNPs at EPHX2 locus. Y axis shows the GWAS p value across 5 AD GWAS datasets. X axis shows the genome position based on hg19. The gradient color denotes the p value alteration in ROSMAP.All pQTL dataset.



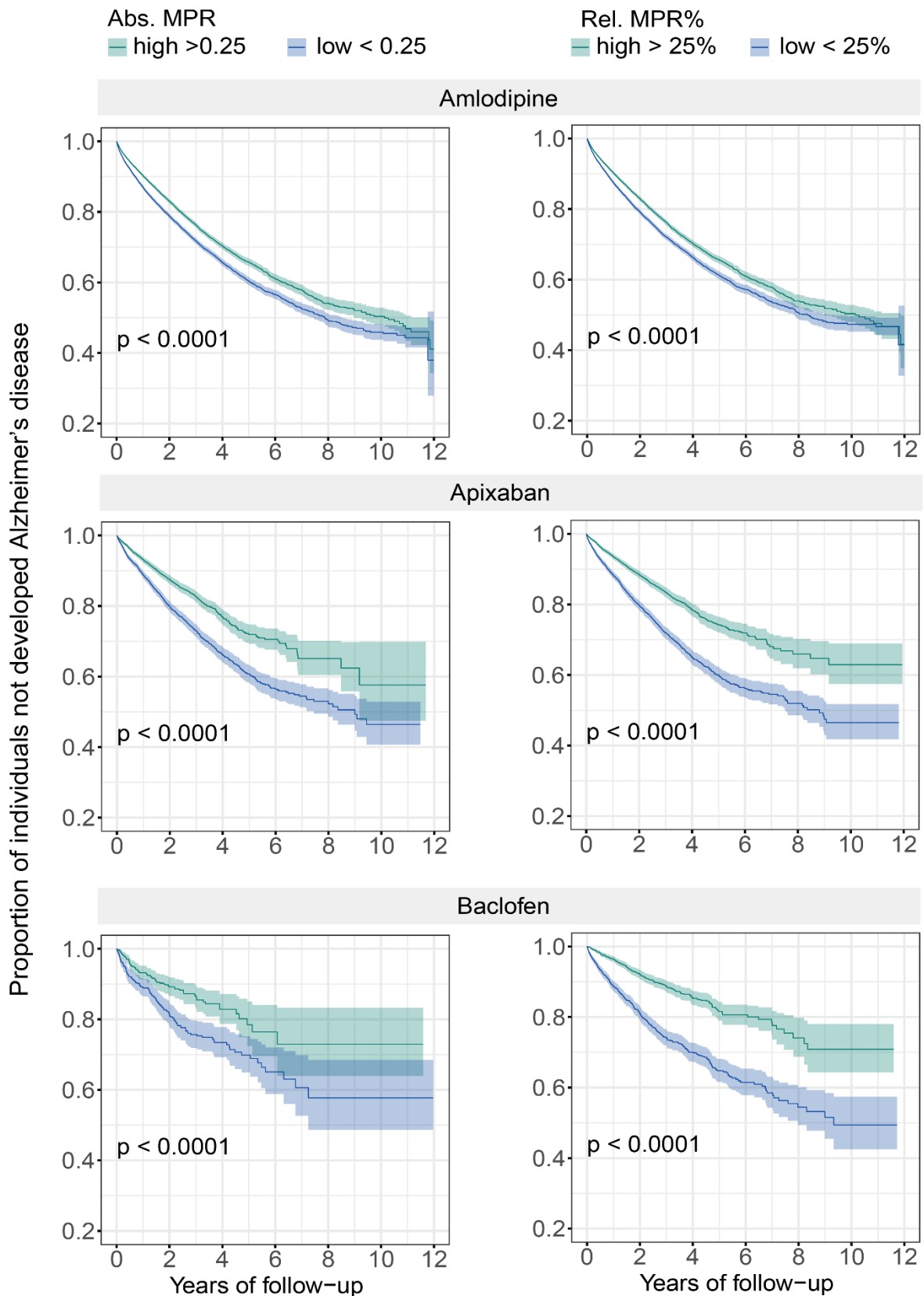
Supplementary Fig. 16 The genome-wide significant loss-of-function variant rs751141 disrupts binding affinity between EPHX2 and TPPU. a. Forest plot shows MR results for EPHX2 based on Wald ratio method. The variant rs751141 pQTL data was used as Instrumental variable (IV). **b** Protein structural model of EPHX2-TPPU complex (PDB: 4OD0). Loss-of-function mutation p.Arg287Gln (rs751141) was highlighted by red color. The orange shadow highlights the binding surface of EPHX2 with TPPU. **c** Root-mean-squared deviation (RMSD) shows the EPHX2-TPPU complex alteration with and without p.Arg287Gln mutation. Color orange and blue indicates native and p.Arg287Gln mutant systems, respectively. Molecular dynamics (MD) simulation of ligand TPPU were shown in right corner. Green color indicates the initial state of TPPU; orange and blue colors indicate the final state of native and p.Arg287Gln mutant in 500-ns MD simulations (see Methods). The structures were plotted by PyMOL v1.8.2.0.



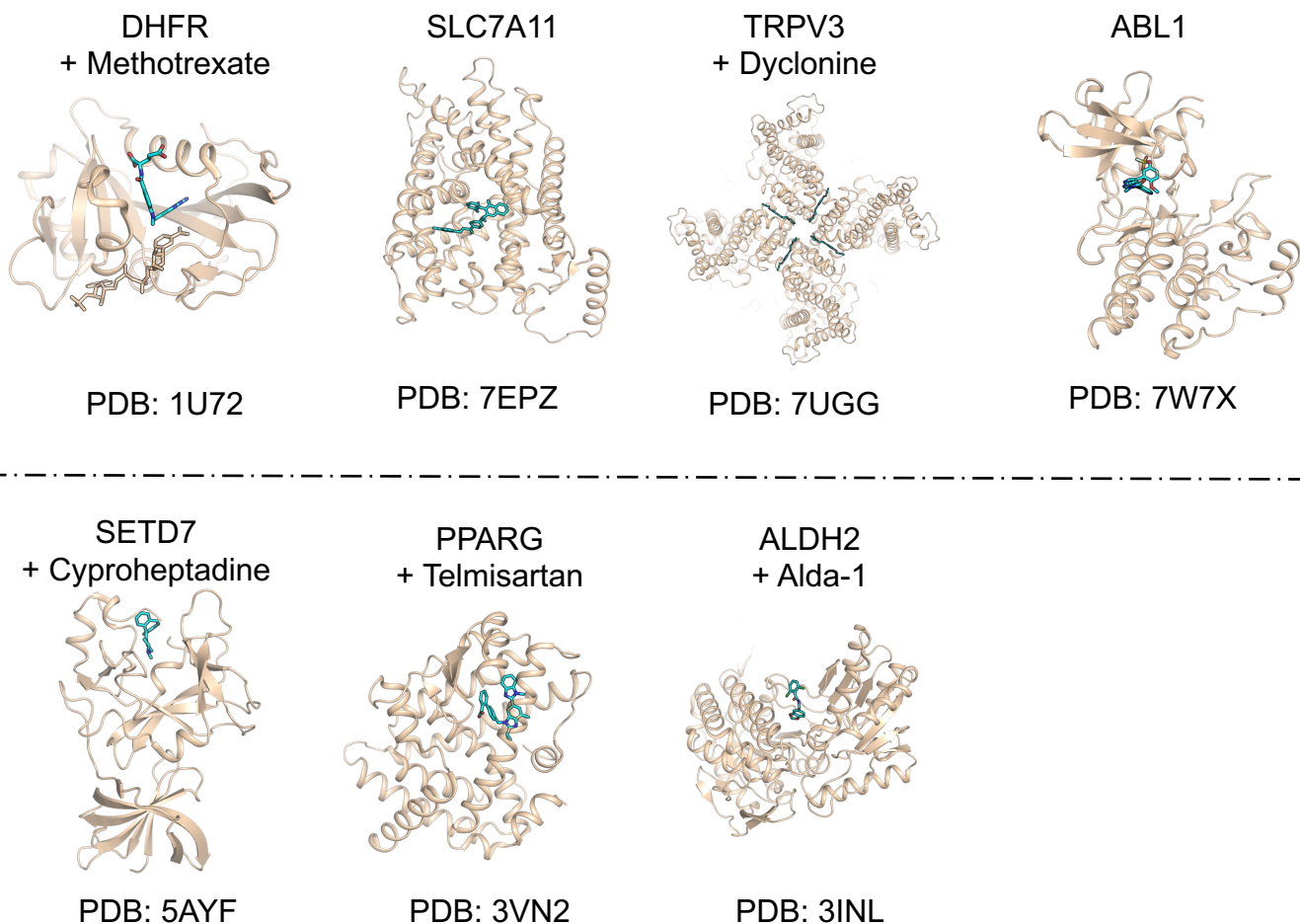
Supplementary Fig. 17 Overexpression system of EPHX2-wildtype and p.Arg287Gln mutants in AD patient induced pluripotent stem cells (iPSCs)- derived neurons. a. The schematic diagram of TALEN system. **b.** Western blotting results showed doxycycline (Dox) inducible overexpression of EPHX2 wild-type and p.Arg287Gln mutants in AD patient iPSC-derived neurons. The “+” and “-” denote the neuron medium with or without Dox added.



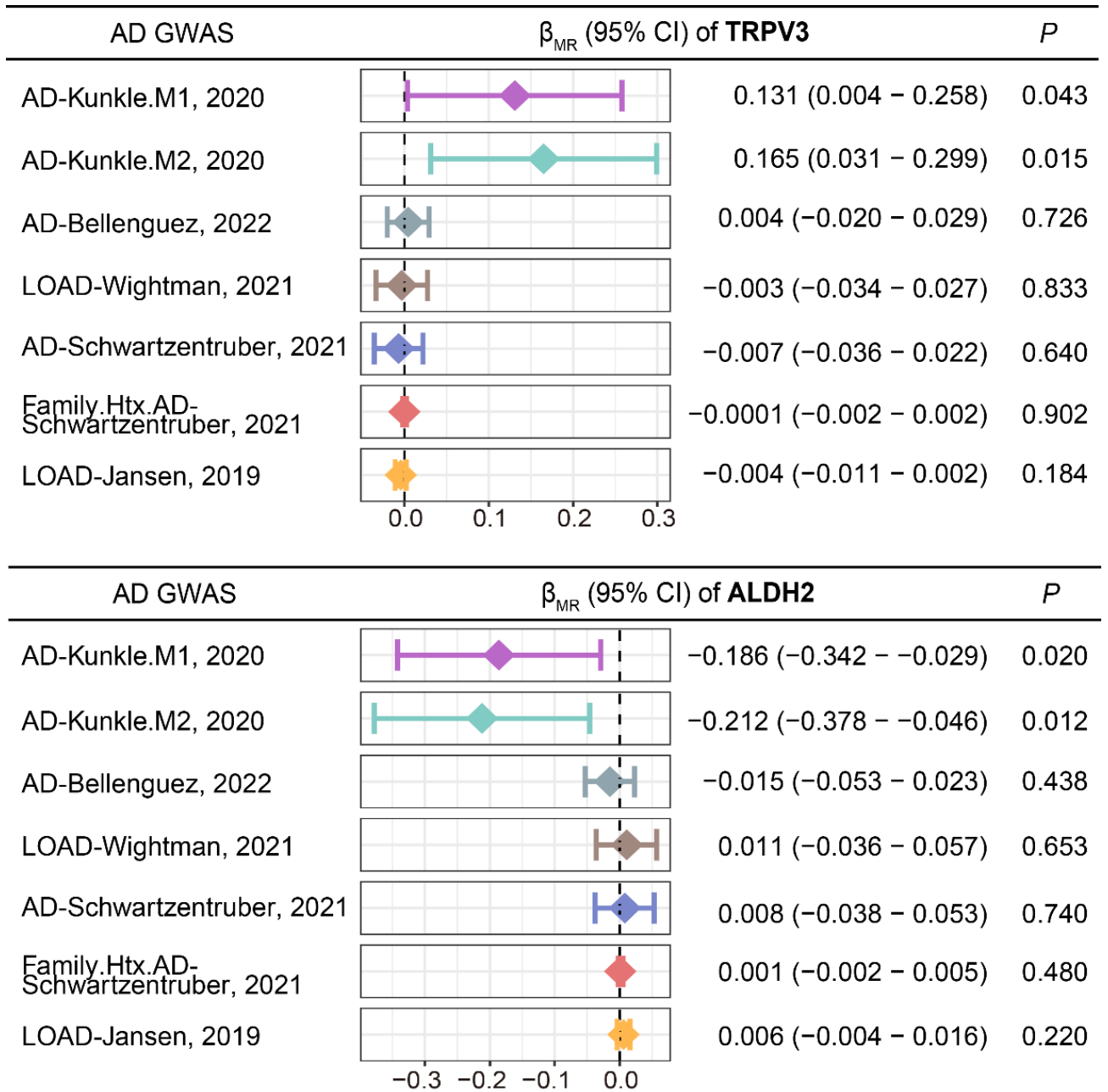
Supplementary Fig. 18 Cognitive function was tested in wildtype rats (n = 14) and TPPU treated (n = 7) rat model of AD (TgF344-AD) using an eight-arm water maze. Four trails were conducted 2 and 24 hrs after the initial training period over two days.



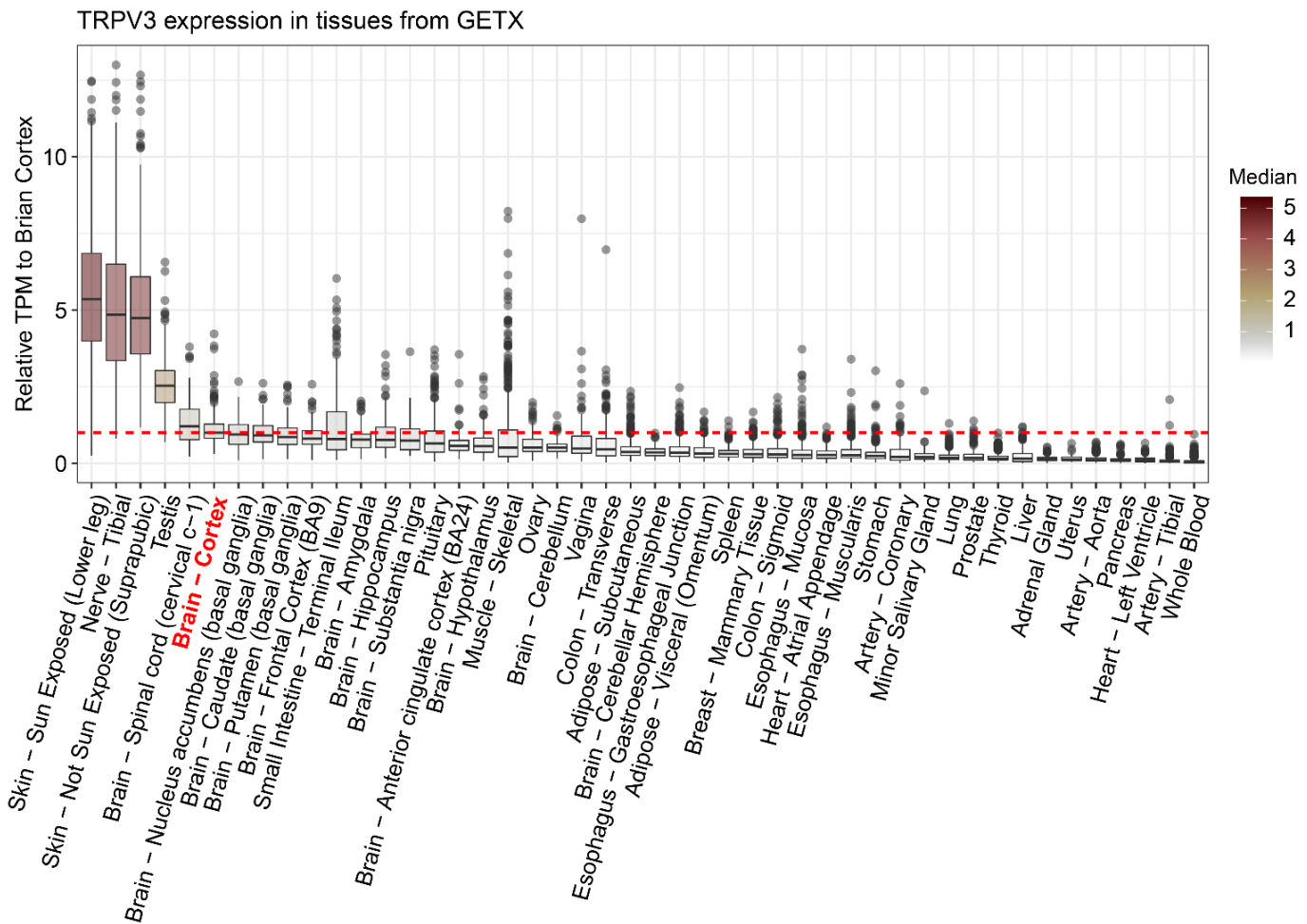
Supplementary Fig. 20 Using propensity score matched survival analyses based on two drugome-wide association studies (DWAS) models. We investigated the associations between medication possession ratios (MPRs) and time to AD. Abs: high exposure (MPR >0.25) vs. low exposure (MPR ≤0.25). Rel: relative high exposure (MPR >25% quantile of the corresponding drug's MPRs) vs. relative low exposure (MPR ≤25% quantile of the corresponding drug's MPRs).



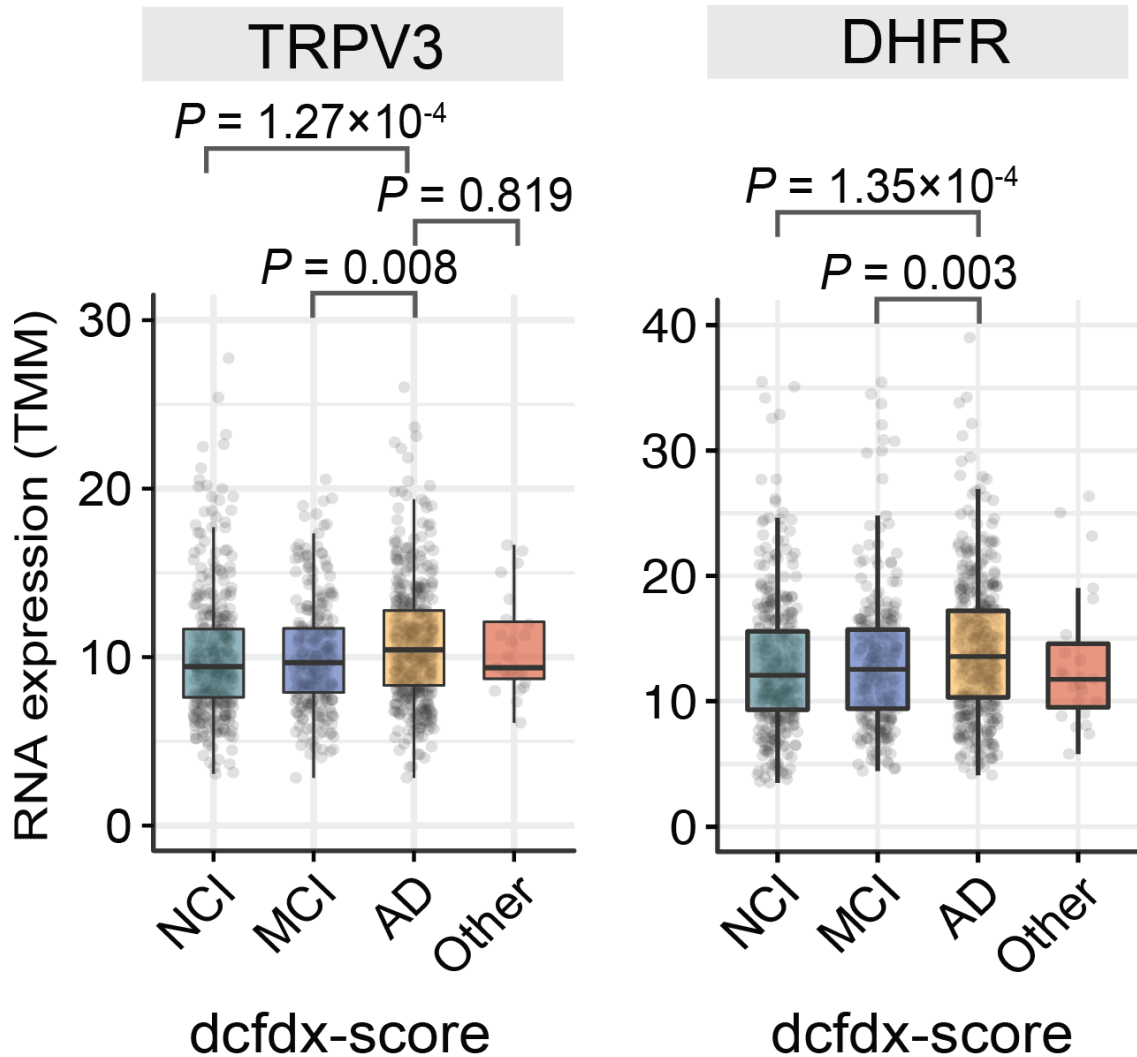
Supplementary Fig. 21 This presents the selected structural complexes of drug targets, highlighting potential druggability for identified targets in **AAs**. The structure complexes show the targets (wheat color) and their reported drugs (cyan color) that highlighted in Figure 6c.



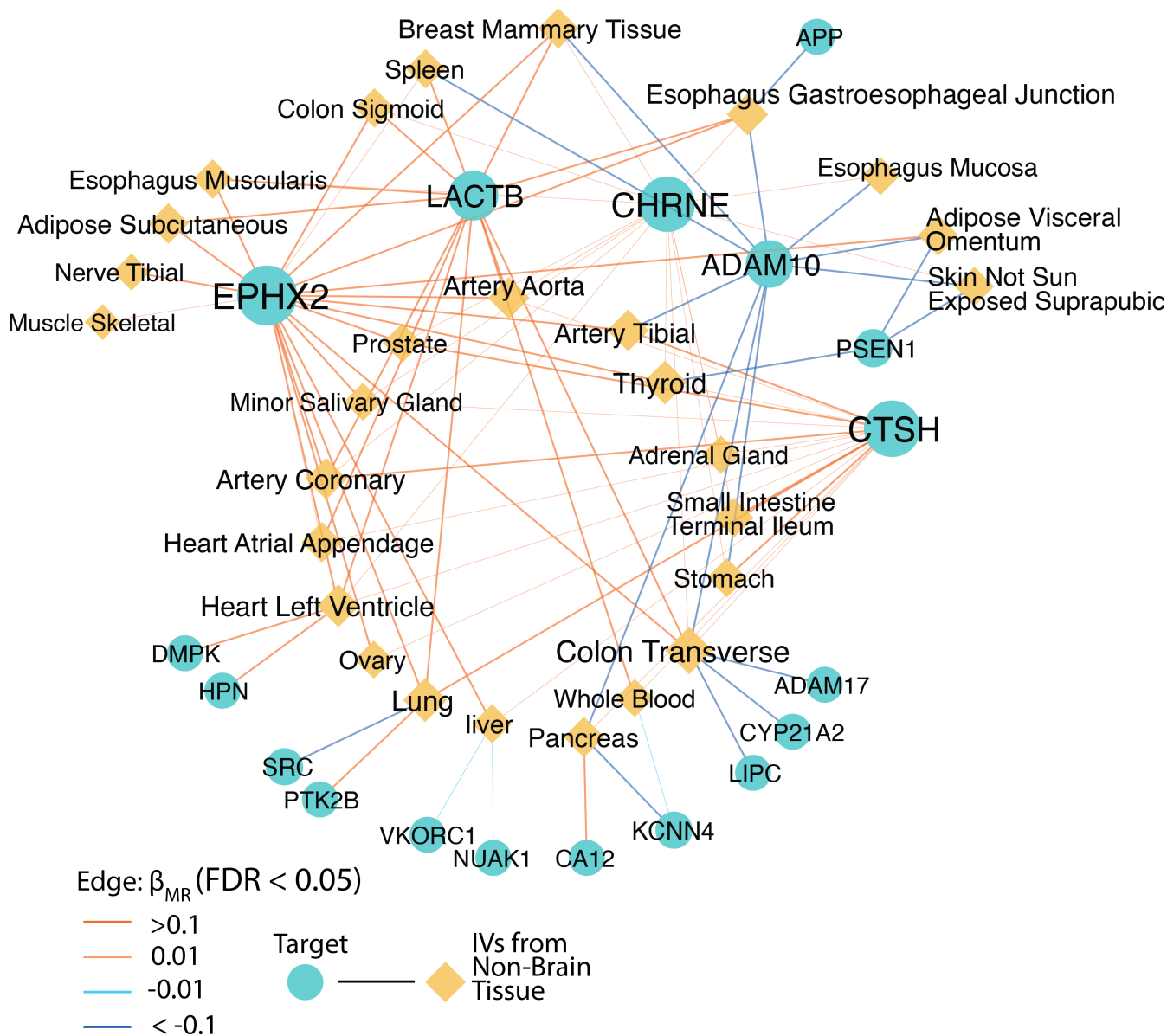
Supplementary Fig. 22 The forest plot shows Mendelian randomization (MR) results of **TRPV3** and **ALDH2** using AA-specific eQTL. 7 AD GWAS datasets (AA-model1 [M1] AA-model2 [M2] and 5 EA) were tested by IVW method.



Supplementary Fig. 23 TRPV3 expression level in 46 human tissues using genotype-tissue expression (GTEx) data. The Y axis shows the relative TPM (transcript per million). Median TPM value of TRPV3 in each tissue was shown by a gradient red color.



Supplementary Fig. 24 The relationship among TRPV3 and DHFR expression and 4 dcfdx score. Mann-Whitney U test was performed for statistical analysis at $P < 0.05$.



Supplementary Fig. 25 The network shows MR results in non-brain tissues for Alzheimer’s disease (AD). The edges indicate significant MR results (FDR < 0.05) when the IVs of the genes (drug targets, circles) from specific non-brain tissues (diamonds) are analyzed for the AD outcome (Bellenguez et al., 2022). The color represents the β -coefficient of the MR results.