

Combining cell type-specific genomic/epigenomic analyses and experimental validation identifies genetic drivers across three neurodegenerative diseases

Yuan Hou

Cleveland Clinic

Liam Wetzel

Case Western Reserve University

Xin Chen

Cleveland Clinic

Xiaoyu Yang

Cleveland Clinic

Dina Bugybayeva

Case Western Reserve University

Teresa Thomas

Case Western Reserve University

Wenqiang Song

Cleveland Clinic

Zhigang Liu

Cleveland Clinic

Yichen Li

Cleveland Clinic

Zhibing Tan

Cleveland Clinic

Ming Hu

Cleveland Clinic

Yang Li

Washington University in St. Louis

Xiongwei Zhu

Case Western Reserve University

Jagan Pillai

Cleveland Clinic

Andrew Pieper

Case Western Reserve University

Jeffrey Cummings

University of Nevada, Las Vegas

Tian Liu

Case Western Reserve University

Feixiong Cheng

chengf@ccf.org

Cleveland Clinic

Research Article

Keywords:

Posted Date: March 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8779983/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. Dr. Cummings has provided consultation to AB Science, Acadia, Alkahest, AlphaCognition, ALZPathFinder, Annovis, AriBio, Artery, Avanir, Biogen, Biosplice, Cassava, Cerevel, Clinilabs, Cortexyme, Diadem, EIP Pharma, Eisai, GatehouseBio, GemVax, Genentech, Green Valley, Grifols, Janssen, Karuna, Lexeo, Lilly, Lundbeck, LSP, Merck, NervGen, Novo Nordisk, Oligomerix, Ono, Otsuka, Pharmacotrophix, PRODEO, Prothena, ReMYND, Renew, Resverlogix, Roche, Signant Health, Suven, Unlearn AI, Vaxxinity, VigilNeuro pharmaceutical, assessment, and investment companies. The other authors have declared no competing interests.

**Combining cell type-specific genomic/epigenomic analyses and
experimental validation identifies genetic drivers across three
neurodegenerative diseases**

Yuan Hou^{1-3,*,#}, Liam Wetzel^{4,#}, Xin Chen^{1,2}, Xiaoyu Yang¹⁻³, Dina Bugybayeva⁴, Teresa
M. Thomas⁴, Wenqiang Song¹⁻³, Zhigang Liu^{1,2}, Yichen Li^{1,2}, Zhibing Tan¹⁻³, Ming Hu^{2,3},
Yang E. Li^{5,6}, Xiongwei Zhu⁴, Jagan A. Pillai⁷, Andrew A. Pieper⁸⁻¹², Jeffrey
Cummings¹³, Tian Liu^{4,*}, Feixiong Cheng^{1-3,*}

¹Cleveland Clinic Genome Center, Cleveland Clinic Research, Cleveland Clinic,
Cleveland, OH 44195, USA

²Department of Genomic Sciences and Systems Biology, Cleveland Clinic Research,
Cleveland Clinic, Cleveland, OH 44195, USA

³Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine, Case
Western Reserve University, Cleveland, OH 44195, USA

⁴Department of Pathology, Case Western Reserve University, Cleveland, OH 44106,
USA

⁵Department of Neurosurgery & Genetics, Washington University School of Medicine,
St. Louis, MO 63110, USA

⁶Department of Genetics, Washington University School of Medicine, St. Louis, MO
63110, USA

⁷Center for Brain Health, Neurological Institute, Department of Neurology, Cleveland
Clinic, Cleveland, OH 44195, USA

24 ⁸Brain Health Medicines Center, Harrington Discovery Institute, University Hospitals
25 Cleveland Medical Center, Cleveland, OH 44106, USA
26 ⁹Department of Neurosciences, Case Western Reserve University, Cleveland, OH, USA
27 ¹⁰Geriatric Psychiatry, GRECC, Louis Stokes Cleveland VA Medical Center, Cleveland,
28 OH 44106, USA
29 ¹¹Institute for Transformative Molecular Medicine, School of Medicine, Case Western
30 Reserve University, Cleveland, OH 44106, USA
31 ¹²Department of Pathology, Case Western Reserve University, School of Medicine,
32 Cleveland, OH 44106, USA
33 ¹³Chambers-Grundy Center for Transformative Neuroscience, Department of Brain
34 Health, Kirk Kerkorian School of Medicine, University of Nevada Las Vegas, Las Vegas,
35 NV 89154, USA
36
37 #These authors equally contribute to this work.
38 *Correspondence to houy2@ccf.org (Y.H), txl757@case.edu (T.L.), chengf@ccf.org (F.C.)
39
40 Correspondence to Feixiong Cheng, Ph.D.
41 Cleveland Clinic Genome Center
42 Tel: +1-216-444-7654; Fax: +1-216-636-0009
43 Email: chengf@ccf.org
44
45

Abstract

The genetic underpinnings of neurodegenerative disease risk remain poorly characterized due to the lack of a unified framework for synthesizing causal evidence from diverse statistical and biological sources. Here, we present an integrative approach to analyze Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS) risk that unifies evidence from Mendelian randomization, colocalization, cell type-specific genomic and epigenomic profiling, human induced pluripotent stem cell (iPSC) and mouse model-based functional validation. We identified 173 high confidence risk genes: 105 for AD, 50 for PD, and 27 for ALS. Single-nucleus multi-omics revealed distinct cell type-specific regulatory architectures, with AD risk genes enriched in microglia and PD and ALS risk genes predominantly associated with neurons and oligodendrocytes. Functional mechanisms were elucidated for specific variants, including a non-coding variant (rs1171832) that regulates expression of the DNA repair and cell cycle *CCDC6* gene by modulating chromatin accessibility in microglia and we further validate it in human iPSC derived microglia by CRISPRi experiments. A nonsynonymous mutation (p.P290A) in the mitochondrial metabolism gene *MTCH2* (rs1064608) that increases its neuronal expression and elevates AD risk. We found that AAV-mediated delivery of mutated *MTCH2* gene ameliorated cognitive deficits in a mouse model, indicating a protective role. This work establishes a functional genomic framework that bridges genetic association with molecular mechanism and target validation, uncovering cell type-specific genetic drivers of neurodegenerative diseases.

Introduction

Genetic epidemiology has identified thousands of genetic variants associated with the risk and severity of neurodegenerative disease¹⁻³. However, whether these associations reflect causal effects of the encoded gene products remains largely unresolved^{4,5}. A primary challenge is the difficulty of evaluating variant effects in a cell type-specific manner^{3,6-8}. For example, conventional bulk tissue analyses average signals across all cell types, which can mask critical cell type-specific pathogenic mechanisms⁹.

Neurodegenerative diseases, including Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS), arise in a highly heterogeneous cellular environment^{1,10-13}, where distinct brain cell types contribute uniquely to disease processes. Recent cell type-specific expression quantitative trait loci (eQTL) studies have found that risk variants are enriched in specific cell types across these diseases^{1,7,14}. For instance, astrocytic genetic risk in AD is primarily implicated in amyloid- β pathology, while microglial risk is associated with both amyloid- β and tau pathology¹⁴. Statistical fine-mapping methods, including colocalization, transcriptome wide association studies (TWAS), and Mendelian randomization (MR), have been applied to these datasets^{15,16}. Nonetheless, these approaches rely on statistical model assumptions that are often unverifiable¹⁷. Consequently, they provide only putative causal evidence, which requires subsequent functional validation^{7,18}.

To address these limitations, we developed an integrative analytical framework that aggregates evidence from MR, colocalization, statistical fine-mapping, and functional validation (both human iPSC and mouse models) to comprehensively assess the causal role of gene products in disease risk. To enhance sensitivity and address the

challenge of heterogeneous GWAS datasets, we developed a consensus neurodegenerative MR score (ND-MR score) that integrates evidence across multiple methods and datasets. We further leveraged matched single nucleus assay of transposase accessible chromatin (ATAC)-seq and single-nucleus RNA-seq data to mechanistically validate single nucleotide polymorphism (SNP)-gene regulatory relationships in a cell type-specific manner. Lastly, we used *in vivo* models, including human induced pluripotent stem cell (iPSC)-derived microglia and mouse models, for functional and behavioral validation. Our work establishes a unified framework that bridges genetic association with mechanistic insight and therapeutic gene target validation, enabling identification of cell type-specific genetic drivers of neurodegenerative disease.

Results

Brain cell type-specific gene effect for AD, PD and ALS

We performed comprehensive genome wide MR analysis to systematically identify cell type-specific genes and therapeutic targets for AD, PD, and ALS. Our MR approach utilized human brain cortex cell type-specific eQTLs as instrumental variables to evaluate the causal effects of 20,320 human genes in these 3 diseases (**Fig. 1a** and **Supplementary Fig. 1**). GWAS datasets exhibit significant heterogeneity, particularly due to the inclusion of proxy AD cases from the UK Biobank in recent AD GWAS studies¹⁹⁻²³. To address this challenge, we developed a consensus ND-MR score (see Method) that integrates and normalizes evidence across multiple GWAS and eQTL data sources, enabling us to screen therapeutic targets with higher sensitivity in causal

relationships. Our ND-MR scores are linear combinations of Z-statistics from MR and can be used to assess MR evidence for each gene and create a ranked list of genes by the strength of this evidence. MR was performed on different pairs of cell type-specific eQTL and disease GWAS cohorts, and Z-statistics from MR are equal to the causal effect estimates divided by their standard errors. For each gene, ND-MR scores are statistically independent since the AD GWAS sample sizes are large, even after considering that many AD GWAS cohorts largely overlap as provide in our previous work²⁴.

Our analysis identified 173 genes with significant, likely causal effects ($p < 1 \times 10^{-6}$), including 105 genes for AD, 50 genes for PD, and 27 genes for ALS (**Supplementary Table 1**). Cell type-specific analysis revealed distinct patterns of genetic effects across human brain cell populations in each disease. For AD, 86 genes (*BIN1*, *ACE*, *GRN*, *EPHX2* and *MTCH2*) showed MR effects in brain cell types in total, with 67 genes (78%, *ACE* for excitatory neurons, *CD2AP* and *MS4A6A* for microglia) showing specificity to a single cell type. The most significant MR-associated genes were enriched in microglia and excitatory neurons (**Fig. 1b** and **Supplementary Table 2**). For PD, 37 genes (*GPNNMB*, *CD38* and *LRRK2*) showed cell type-specific MR effects, with 23 genes (62%) restricted to single cell type (**Fig. 1b** and **Supplementary Table 3**). For ALS, 19 genes (*C9orf72*, *SCFD1* and *MOBP*) demonstrated cell type-specific effects, with 14 genes (74%) showing single cell type specificity (**Fig. 1b** and **Supplementary Table 4**). In both PD and ALS, excitatory neurons exhibited the most significant MR associations.

Beyond single cell type-specific effects, we observed that several genes influenced multiple cell types. For example, *USP6NL* showed significant ND-MR scores

across microglia, oligodendrocytes, and inhibitory neurons in AD, while *GPNMB* showed significant ND-MR scores in both microglia and oligodendrocyte precursor cells (OPCs) in PD. Recent studies have demonstrated that GPNMB secretion is a key process in response to lysosomal stress²⁵. GPNMB also physically interacts with alpha-synuclein (aSyn), the main pathological protein in PD, and promotes the internalization of fibrillar aSyn into neurons²⁶. In ALS, *C9orf72* exhibited effects across all cell types except OPC, highlighting its broad cellular impact in disease pathogenesis. These findings illustrate our consensus MR approach for identifying cell type-specific gene signatures and reveal the multi-cellular complexity of neurodegeneration in AD, PD and ALS.

Prioritization of cell type-specific neurodegeneration-associated risk genes

Next, we performed colocalization analysis (PP.H4 > 0.5) and fine-mapping (FOCUS, PIP > 0.5) for AD, PD, and ALS to validate the ND-MR score-prioritized genes (**Supplementary Tables 5-7**). We identified 45 significant genes with ND-MR scores supported by either colocalization or fine-mapping evidence for AD (**Fig. 2** and **Supplementary Figs. 2 and 4**). Among these, nine genes showed robust significance validated by all three methods (MR, colocalization, and FOCUS), including *USP6NL*, *ACE*, *EGFR*, *INPP5D*, *CASS4*, *CCDC6*, *ADAM10*, *NDUFS2*, and *PRSS36*. Notably, four genes (*USP6NL*, *INPP5D*, *CASS4*, and *CCDC6*) illustrated microglia-specific effects for AD in colocalization analysis (**Supplementary Fig. 2**). Some of these genes were validated through FOCUS and MR analysis, including *EPHX2*, *GRN*, *TSPAN14*, and *TOMM40* (**Fig. 2** and **Supplementary Fig. 4**). *TSPAN14* has also been

demonstrated in human WTC11-derived microglia to affect physiological functions in AD⁷.

PD risk genes displayed distinct validation patterns, with 23 genes confirmed through MR with colocalization or FOCUS analysis. Five genes (*FAM47E*, *SH3GL2*, *CD38*, *GPNMB*, and *HIP1R*) were verified by all three methods (**Fig. 2** and **Supplementary Figs. 3 and 4**). We observed that genes validated by colocalization were more enriched in microglia, including *LRRK2*, *TMEM163*, *GPNMB*, and *KANSL1*, while genes such as *SNCA* and *NUP42* were verified only by FOCUS and MR (**Supplementary Fig. 4**). *KANSL1* showed significant negative association in both AD and PD at the *MAPT* locus (encoding tau protein), suggesting shared disease mechanisms related to tau pathology^{27,28}. ALS risk genes showed significant associations primarily in excitatory and inhibitory neurons, exemplified by *C9orf72*, *SCFD1*, and *SLC9A8* (**Fig. 2**). The ALS risk gene *MOBP* was validated by both colocalization and FOCUS in oligodendrocytes, consistent with experimental evidence of its function in stabilizing the myelin sheath in the central nervous system (CNS)²⁹. These cell type-specific cross-verification analyses demonstrate that our ND-MR approach identifies disease relevant genes with distinct cellular patterns across AD, PD and ALS.

Brain cell type-specific risk loci in neurodegeneration

While MR and colocalization analyses can identify putative causal genes and genomic loci, they cannot pinpoint causal SNPs or elucidate regulatory mechanisms at the single genetic variant level. To address this limitation, we utilized matched single-nucleus

ATAC-seq (snATAC-seq) and single-nucleus RNA-seq (snRNA-seq) data from the ROSMAP cohort³⁰ to further evaluate the putative causal SNP-gene associations identified through ND-MR score analysis (**Supplementary Table S8-9**). For AD, 155 SNP-gene pairs were prioritized by ND-MR score. Among these, 29 pairs were in open chromatin regions as evidenced by snATAC-seq, and 14 showed significant peak-to-gene associations (**Fig. 3a**). We also identified 159 SNP-gene pairs verified through colocalization analysis, among which 19 pairs were regulated by open chromatin regions and 16 pairs were validated by peak-to-gene association analysis. We found 11 SNP-gene pairs prioritized by both MR and colocalization analysis and verified by snATAC-seq and peak-to-gene analysis, such as rs1171832 of *CCDC6*, rs117618017 of *APH1B*, rs113706587 of *RASGEF1C*, and rs6733839 of *BIN1* (**Fig. 3a**).

We extended our single-cell multi-omics approach to examine putative causal SNPs for PD and ALS. For PD, ND-MR score analysis identified 66 SNP-gene pairs as high-priority candidates (**Supplementary Tables 8 and 9**). Within this set, 11 pairs were in open chromatin regions detected by snATAC-seq, while 5 pairs exhibited significant peak-to-gene regulatory relationships using matched snATAC-seq and snRNAs-eq data³⁰ (**Fig. 3a**). Colocalization analysis provided additional support for 62 SNP-gene pairs, nine of which displayed evidence of open chromatin-mediated regulation and 13 of which received confirmation through peak-to-gene association analysis. Four SNP-gene pairs were consistently prioritized by both MR and colocalization methods and validated through chromatin accessibility analysis, including rs10847864 of *HIP1R*, rs858239 of *GPNMB*, and rs10935841 of *P2RY12*. For ALS, we identified 6 SNP-gene pairs that were prioritized by ND-MR score and supported by snATAC-seq and snRNA-

seq evidence. These pairs showed regulation within open chromatin regions and significant associations with gene expression, including rs13079368 of *MOBP*, rs10812619 of *C9orf72*, rs6020069 of *SLC9A8*, and rs229141 of *SCFD1*.

Next, we visualized the cell type-specific features of validated SNP-gene pairs to illustrate the complex regulatory architecture underlying AD, PD, and ALS (**Fig. 3b**). Network analysis revealed the ability of individual SNPs to influence gene expression across multiple brain cell types, showing either broad regulatory impact or cell type-specific effects. For example, SNP rs1532278 cis-regulated *CLU* expression in multiple cell types, including inhibitory neurons, endothelial cells, OPC, oligodendrocytes, and astrocytes, and this cis-regulation was validated in iPSC derived astrocytes². We also identified microglia specific SNP-gene pairs, such as rs17783630 of *RIN3*, and rs113706587 of *RASGEF1C* (**Fig. 3b** and **Fig. 4a**). These findings emphasize the central role of microglial dysfunction in AD.

For PD, network analysis revealed distinct cell type specific patterns, with rs858239 of *GPNMB* specifically affecting glial cell types, including microglia, OPC, and astrocytes (**Fig. 3b** and **Fig. 4a**). Similarly, CD38 showed regulatory effects primarily in microglia and excitatory neurons, highlighting the important role of neuroinflammatory pathways in PD pathogenesis. The SNP rs10847864 of *HIP1R* showed predominant effects in astrocytes, oligodendrocytes and OPC (**Supplementary Fig. 5a**), while rs10935841 of *P2RY12* revealed microglia and astrocyte regulation (**Fig. 4a**). For ALS, we observed distinct cell type-specific regulatory patterns, with rs6020069 of *SLC9A8* showing oligodendrocyte specific effects and rs13079368 of *MOBP* demonstrating excitatory neuron regulation (**Fig. 3b** and **Supplementary Fig. 6a**). *SCFD1* also

exhibited excitatory neuron specific effects, while rs10812619 of C9orf72 displayed regulation in both microglia and excitatory neurons (**Supplementary Fig. 6b**), consistent with the multi-cellular pathology of ALS. Collectively, these cell type specific validation analyses demonstrate that different neurodegenerative diseases exhibit distinct cellular regulatory architectures. AD shows prominent microglial involvement, while PD and ALS are characterized predominantly by neuronal and oligodendrocyte effects.

Discovery of microglia specific *cis*-regulatory variant of CCDC6 in AD by CRISPRi

To further evaluate cell type-specific regulatory effects in disease pathologies, we performed correlation analysis between gene expression and amyloid and tau densities across different brain cell types (**Fig. 4b**). We found distinct correlation patterns for amyloid-related genes and tau tangle-related genes across astrocytes, excitatory neurons, inhibitory neurons, microglia, oligodendrocytes, and OPCs. The genes *APH1B*, *RIN3*, and *CCDC6* showed significant positive correlations with both amyloid and tau tangle densities specifically in microglia. In contrast, *BIN1*, *CLU*, *RHOH*, and *POLR2E* genes showed significant positive correlations only with amyloid density in microglia, oligodendrocytes, and OPCs. These findings provide further supporting evidence that AD pathology relevant risk genes exhibit distinct cell type-specific regulatory mechanisms (**Fig. 4**).

Next, we performed CRISPRi experiments using human iPSC-derived microglia (*cf.* Methods) to validate the *cis*-regulatory relationship between rs1171832 and *CCDC6* expression (**Fig. 4c**). We designed two independent gRNAs targeting the genomic

region containing rs1171832 and measured the changes of *CCDC6* expression compared to internal controls. The CRISPRi results demonstrated a significant reduction in *CCDC6* expression ($p = 1 \times 10^{-4}$) in microglia when the rs1171832 region was blocked, confirming the causal regulatory relationship identified by our ND-MR and snATAC-seq approach.

Experimental validation of nonsynonymous p.P290A mitochondrial carrier homolog 2 (MTCH2) mutation in AD

Cell type-specific eQTL analysis (*cf.* Methods) found that the nonsynonymous mutation rs1064608 p.P290A (G/C) was significantly associated with elevated mitochondrial carrier homolog 2 (*MTCH2*) gene expression in excitatory and inhibitory neurons (**Supplementary Fig. 7a**). rs1064608 has another rare p.P290T mutation in general population. Consistently, snRNA-seq analysis revealed that excitatory neurons expressed higher levels of *MTCH2* compared to other cell types in the human brain (**Supplementary Fig. 7b**), especially in the frontal lobe, cerebellum, and temporal lobe (**Supplementary Fig. 7c**). Our MR analysis further revealed that the p.P290A variant in *MTCH2* showed significant association with increased AD risk in both excitatory neurons ($\text{FDR} = 4.43 \times 10^{-8}$) and inhibitory neurons ($p = 3.28 \times 10^{-8}$, **Fig. 5a**).

Colocalization analysis of the *MTCH2* locus revealed that the lead noncoding SNP rs3740688 showed strong colocalization eQTL signals ($\text{PP.H4} = 0.83$, **Fig. 5b**) in excitatory neuron. These findings indicate that elevated *MTCH2*^{p.P290A} expression in excitatory neurons increases the likelihood of AD, while higher expression of wildtype *MTCH2* may mediate a potential compensatory mechanism.

To examine potential alterations of MTCH2 protein levels in AD, we conducted immunoblot analysis of human cortical brain samples from AD patients (n = 10) and matched controls (n = 10), **Fig. 5c-e**). The immunoblot experiment showed significantly reduced soluble MTCH2 protein levels (p = 0.008, **Fig. 5c,5e**) but significantly increased insoluble MTCH2 protein in AD samples relative to controls (p = 0.012, **Fig. 5d,5e**). Our immunofluorescence staining results showed that the p.P290A and p.P290T *MTCH2* mutants formed aggregates and disrupted *MTCH2*-wildtype colocalization with mitochondria (**Fig. 5g**). We further assessed the solubility of wildtype and mutant *MTCH2* using HEK293T cells and found that *MTCH2* wildtype was predominantly identified in the soluble fraction, whereas significantly elevated protein level of p.P290A or p.P290T *MTCH2* mutants were detected in the insoluble fraction (p = 0.002, **Fig. 5h-i**). The p.P290A and p.P290T mutations impair *MTCH2* protein solubility and disrupt its normal mitochondrial localization, suggesting that these genetic variants contribute to AD pathogenesis through mitochondrial dysfunction (**Supplementary Fig. 8a**).

***MTCH2* associates with elevated tau pathology and impacts cognitive behavior in a mouse model**

We evaluated the impact of *MTCH2* variants on AD pathology using Tau--fluorescence resonance energy transfer (FRET) biosensor cells using an over expression vector system and found that *MTCH2* p.P290A or p.P290T mutants significantly enhanced the tau seeding (p.P290A p = 3.35×10^{-9} ; p.P290T p = 5.90×10^{-12} , **Fig. 6a, b**) compared to *MTCH2*-WT group. Both p.P290A and p.P290T *MTCH2* mutants showed consistent effects in tau seeding experiments, but p.P290T exhibited more severe abnormal

phenotypes. Based on these observations, we selected MTCH2-WT and MTCH2-p.P290T for subsequent *in vivo* studies.

We designed an adeno-associated virus 9 (AAV9) delivery system to overexpress *MTCH2*-WT and *MTCH2*-p.P290T in C57BL/6J mouse brain hippocampus through stereotaxic injection (n = 6, female vs. male = 1:1). Immunohistochemistry for HA-tag detection confirmed successful viral transduction and protein expression across the hippocampus (**Fig. 6c**). Specifically, astrocyte (glial fibrillary acidic protein (GFAP) marker) and microglial (Iba1 marker) cell densities were significantly increased in *MTCH2*-p.P290T mutant groups compared to wildtype (**Fig. 6d, e** and **Supplementary Fig. 8b**), suggesting enhanced neuroinflammatory responses induced by mutant MTCH2.

We performed behavioral assessments to evaluate potential phenotypic and cognitive abnormalities between the *MTCH2*-WT and *MTCH2*-p.P290T injection groups. In the novel object recognition (NOR) task, mice were trained with two similar objects, with one object substituted for a novel object with differing in size, color, and texture in subsequent day. *MTCH2*-p.P290T mice exhibited reduced total exploration of both new (p = 0.054) and old objects (p = 0.002, **Fig. 6f**), demonstrating impaired exploratory behavior. The open field test was performed to evaluate locomotor activity, exploratory behavior, and anxiety-like responses in a novel, open environment (**Fig. 6g**). Mice were monitored based on the duration they occupied in the inner, middle, and outer zones of the area. Mice injected with AAV-MTCH2-pP290T exhibited significantly reduced time in the inner and middle zones and increased time in the outer zone relative to those

injected with *MTCH2*-WT ($p = 7.60 \times 10^{-4}$, **Fig. 6g**), indicating heightened anxiety-like behaviors by *MTCH2*-p.P290T.

These comprehensive functional experiments support that our integrated MR platform can translate genetic associations into validated disease risk targets. *MTCH2*-p.P290T and p.P290A mutations exert deleterious effects on brain function through multiple pathological mechanisms. The mutations enhance tau seeding and aggregation in cellular models, elicited neuroinflammatory responses in vivo using an AAV overexpression mouse model system, and induce behavioral deficits including anxiety-like responses.

Discussion

We identified 173 cell type specific, candidate risk genes across AD, PD, and ALS using integrated MR with single-nucleus multi-omics (including RNA-seq and ATAC-seq) validation. The matched snATAC-seq and snRNA-seq results revealed distinct cellular regulatory architectures for each disease. For example, AD showed prominent glial involvement (**Fig. 2** and **Fig. 3b**), particularly in microglia, consistent with the importance of neuroinflammatory mechanisms in AD pathophysiology.³¹ PD and ALS, on the other hand, displayed focused effects in excitatory neurons and oligodendrocytes (**Fig. 2** and **Fig. 3b**).

Cell type-specific correlations between gene expression and AD pathology (amyloid beta and tau densities) further support the disease relevance of our computationally prioritized genes (**Fig. 4b**). For example, *APH1B*, *RIN3*, and *CCDC6* correlated with amyloid and tau pathologies in microglia. *CCDC6* encodes a coiled-coil

domain containing protein involved in DNA repair and cell cycle regulation,³² and accumulated DNA damage increases cytosolic DNA fragments, potentially linking it to microglial inflammation³³. Our CRISPRi experiment supported a cis-regulatory relationship between the AD risk SNP rs1171832 and *CCDC6* in iPSC-derived microglia, reinforcing the involvement of *CCDC6* in microglial responses. *RIN3*, a Ras and Rab interactor that regulates endosomal trafficking and vesicle transport,^{34,35} shows associations with both amyloid and tau pathologies, consistent with a role in microglial clearance mechanisms.

MTCH2 exemplifies the ability of our MR score framework to identify genes with potential causal roles. *MTCH2* showed strong causal evidence in MR, neuronal-specific expression in snRNA-seq data, and functional consequences when mutated (**Fig. 5**). We observed that the *MTCH2*-p.P290A variant shifts from a soluble to insoluble form, which impairs mitochondrial function and enhances tau seeding. As a key mitochondrial membrane protein that regulates mitochondrial fusion,³⁶ lipid metabolism,³⁷ and cellular apoptosis, these findings provide independent validation of the MR result and supports the role for *MTCH2* contributes to neuronal mitochondrial dysfunction in neurodegenerative disease.

We acknowledge several potential limitations. Our snRNA-seq and snATAC-seq validation captures a static view of regulatory relationships that likely change across disease stages, environments, and genetic backgrounds. This temporal and contextual variability may explain why some well-established genes prioritized by colocalization and MR lack strong snATAC-seq signals. Experimental validation was also constrained to specific cell types and conditions which may not fully recapitulate the complex *in vivo*

environment. AAV mediated delivery and overexpression does not confirm naturally occurring, endogenous expression changes. Future work should incorporate longitudinal single cell/nuclear datasets to track regulatory dynamics during disease progression and expand functional assays to multiple cell types, co-culture systems, and disease relevant stimuli (i.e. inflammatory challenges, protein aggregates). Integrating spatial transcriptomics and proteomics will help resolve how genetic variants alter cellular function within complex tissue microenvironments.

Methods

Collection of GWAS Data from Diverse Populations

For all colocalization and MR analyses we used publicly available GWAS summary statistics for each of PD, AD, and ALS. For PD, we used GWAS data from Kim³⁸ et al. (2023), which included 49,049 cases, 18,618 proxy cases and 2,458,063 controls of mixed ancestry (i.e., 62% European, 23% Latin American, 8% African, 7% East Asian). For ALS, we used GWAS data from Van Rheenen et al³⁹. (2021), which included 27,205 European cases and 110,881 controls. For AD, we used seven GWAS data sources: (1) Bellenguez et al¹⁹. (2022, 86K AD and proxy cases, 401K controls), Wightman et al.²³ (2021, 40K cases and 358K controls without UK-biobank and 23 and me data), Jansen et al²¹. (2019, 72K AD and proxy cases and 383K controls), Kunkle et al²⁰. (2019, 30K cases and 52K controls), Schwartzenuber et al²². (2021, 75K AD and proxy cases and 397K controls), Schwartzenuber et al. (2021, 53K proxy cases and 356K controls). A visual description of the multiple AD cohorts and their degree of overlap is presented in **Supplementary Figure S1.**

Multi-source quantitative trait loci (QTL) data

We collected multi-source quantitative trait loci (QTL) data, including protein QTL (pQTL) and brain region and cell type specific expression QTL (eQTL) data in this study. pQTL data were from Religious Orders Study/Memory (ROSMAP) cohort and Banner Health brain project⁴⁰. Bulk eQTL data from cortex tissue were from three cohorts, including ROSMAP (cortex tissue), Mayo Clinic (cortex and cerebral tissues)⁴¹, and MetaBrain⁴² (cortex, cerebral, basal ganglia, hippocampus, and spinal cord tissues). All data sets are publicly available from the AD Knowledge Portal (synapse ID: syn10715233) and MetaBrain. We analyzed 19 cell type specific eQTL datasets from brain tissue. Seven datasets came from DLPFC tissue¹ and eight datasets⁴³ came from prefrontal cortex, temporal cortex, and deep white matter. These datasets covered eight different cell types, including astrocytes, endothelial cells, excitatory neurons, inhibitory neurons, microglia, oligodendrocytes, OPCs, and pericytes. Another five microglia specific eQTL datasets from fresh brain samples across four brain subventricular zone, medial frontal gyrus, superior temporal gyrus thalamus⁴⁴ and dorsolateral prefrontal cortex⁶ were also included.

Mendelian Randomization score

We performed univariable MR genome-wide and treated a gene product as the MR exposure and one of three neurodegenerative diseases (ALS, AD, PD) as the outcome. To satisfy the key MR assumptions of instrument strength and no horizontal pleiotropy^{45,46}, we selected instrumental variables (IVs) that: (i) had an FDR q-value less

than 0.05 in their original GWAS, (ii) had a SNP-specific F-statistic greater than 10, (iii) were not in absolute linkage disequilibrium greater than 0.1 with any other IVs within a 500kb window, and (iv) were not associated with the products of more than two other genes. Since not all MR estimators of interest could be applied when the number of IVs was small, we implemented the following scheme based on IV availability⁴⁷: (i) #IVs=1: Wald ratio, (ii) #IVs=2: fixed-effect IVW, (iii) #IVs>2: random-effect IVW, MR-PRESSO⁴⁸, MR-MaxLik, Weighted Median. When possible, we also calculated the Cochran's Q-statistic and heterogeneity statistic I^2 . All MR analyses were conducted using the TwoSampleMR package in R (<https://mrcieu.github.io/TwoSampleMR/>) and R v4.1.1. Gene products for which no IVs could be identified were skipped.

We develop a score to quantify the magnitude of causal evidence from MR which supports a nonzero causal effect when multiple, potentially non-independent, causal effect estimates are made. Such a score is intended to be largest for genes with the strongest causal effects and for which we have the highest degree of confidence that they are so. We additionally want the score to capture evidence aggregated across different biologically related gene phenotypes (e.g., gene expression across different cell types and brain protein abundance).

$$S(g) = \frac{1}{N_g} \sum_j \sum_k Z_{jkg} \quad (\text{Equation 1})$$

The $S(g)$ is ND-MR score for gene g , which is the normalized of Z-statistics across all k xQTL cohorts and j GWAS cohorts (**Equation 1**), where Z_{jkg} represents the Z-statistic testing the causal null hypothesis for gene g in the k th xQTL cohort and j th GWAS cohort (**Equation 2**). Z-score is the *causal effect* (θ_{jkg} , IVW estimator in **Equation 3**) divided by its *standard error* (s)

$$Z_{jkg} = \frac{\theta_{jkg}}{s_{jkg}} \quad (\text{Equation 2})$$

$$\theta_{jkg} = \beta_{jkg}^T \alpha_{jkg} h_{jkg}^{-1} \quad (\text{Equation 3})$$

We first sum the Z-statistics used to test the causal null hypothesis in each gene product-disease cohort pair and is referred to as ‘sumZscore’. Then, the score was the average of the sumZscore values across all tested gene product-AD cohort pairs and is referred to as ‘normalized-Zscore’ ($S(g)$). The rationale for using both scores is that genes with a large sumZscore value have broad causal effects on risk across multiple tissues and/or gene products, whereas genes with large normalized-Zscore values have evidence of causality for only a relatively small number of gene products. An illustration of how these two scores complement each other is presented in our previous study.

Colocalization and fine mapping

We performed colocalization analyses of xQTLs and fine mapping of their corresponding genes to provide additional evaluations of causality for each gene product. We performed colocalization as an additional evaluation of causality for each gene locus, and fine mapping as additional evidence of causality for a gene product. MR is theoretically a stricter form of colocalization in which nonzero genetic correlation is also required for causal inference, making colocalization and MR not completely independent evaluations. However, these two approaches apply different methods to test the causal null hypothesis and so their causal evidence is complementary. Colocalization was performed using the coloc⁴⁹ package in R and fine mapping was performed using the FOCUS¹⁷ method.

Mapping GWAS loci to open chromatin regions

To identify regulatory elements underlying genetic associations, we mapped cell type specific MR instrumental variables and fine-mapping SNPs from colonization to cell type specific chromatin accessibility regions from processed ROSMAP multiomics data³⁰, including 850K RNA and ATAC cells in 92 AD case control individuals (AD Knowledge Portal ID: syn2580853). For each disease (AD, PD, ALS), we extracted all MR IVs that passed quality control filters ($p < 1 \times 10^{-6}$). SNP positions in GRCh38 were then overlapped with cell type specific chromatin accessibility peaks, defining an overlap when the SNP position fell within the peak boundaries. Besides the open chromatin peak, we also validate our significant SNP in significant peak to gene dataset (FDR < 0.05) to ensure that the identified regulatory elements were functionally relevant to the gene products tested in our MR analysis. All data processing and overlap analyses were performed in R 4.1.1 platform.

Functional validation using CRISPRi in iPSC-derived microglia.

To achieve stable and uniform expression of transgenes during microglial maturation, a dCas9-KRAB construct under the control of the CAG promoter was inserted into the intronic region of the CLYBL safe harbor locus in wild-type WTC11 iPSCs (PMID: 37735198). WTC11-derived microglia were differentiated through staged exposure to mesoderm- and hematopoietic-inducing cytokines in defined serum-free media as previously reported (PMID: 28528700). Briefly, iPSCs were first treated with BMP4 (PeproTech 120-05) for 4 days. The cells were then exposed to FGF2, SCF, and VEGF (PeproTech 100-18, 300-07, 100-20) in StemPro-34 medium for 2 days, followed by

480 sequential cytokine cocktails containing IL-3, SCF, M-CSF, TPO, and FLT3 (PeproTech
481 200-03, 300-25, 300-18, 300-19) to promote hematopoietic progenitor specification.
482 From day 15 onward, cultures were maintained with M-CSF, FLT3, and GM-CSF
483 (PeproTech 300-03), and by days 25-35 the floating cells were matured in RPMI
484 (Thermo Fisher 12633012) with IL-34 (PeproTech 200-34) and GM-CSF to generate
485 microglia-like cells. Single guide RNAs (sgRNAs) directed against prioritized SNP or risk
486 gene promoters were designed using CHOP-CHOP (PMID: 31106371). Potential off-
487 target binding sites with one (MM1), two (MM2), or three (MM3) base mismatches were
488 evaluated, and two distinct sgRNAs with zero MM1/MM2 events and fewer than five
489 MM3 hits were selected to ensure specificity for each genomic region (Supplemental
490 Table x). For cloning, sgRNA oligonucleotides were inserted into the lentiGuide-puro
491 backbone (Addgene 52963) and confirmed with Sanger sequencing. Lentiviral particles
492 were produced in 293T cells by co-transfection of the sgRNA plasmid with pMD2.G
493 (Addgene 12259) and psPAX (Addgene 12260) helper vectors using Lipofectamine
494 2000 (Invitrogen 1166802). Viral supernatants were harvested over 48 h, passed
495 through 0.45 µm filters (Millipore SLHV033RS), and concentrated with Amicon Ultra
496 centrifugal devices (Millipore UFC801024). Viral titers were determined in iPSCs by
497 puromycin selection. For functional CRISPRi experiments, microglia-like cells were
498 seeded and matured for 14 days before lentiviral infection carrying sgRNAs at a
499 multiplicity of infection (MOI) of approximately 5. Following transduction, cultures were
500 maintained under puromycin selection (0.5 µg/mL) for four days. Cells were harvested
501 seven days after infection for RNA extraction. Gene expression changes were

quantified by RT-qPCR from at least 3 biological replicates, and relative expression across replicates was analyzed by two-tailed Student's t-test.

Human Brain Samples

Human cortical tissues employed for immunoblotting assays were obtained from the Brain Bank at Case Western Reserve University and the NIH Neurobiobank. The neuropathology of normal controls (ages 60–92 years, postmortem interval range 4–25 hours) and AD cases (ages 67–90 years, postmortem interval range 3–26 hours) was evaluated. All procedures adhered to authorized IRB protocols and conformed to institutional bioethical guidelines. All methods and protocols used in this study involving mice were approved by the Institutional Animal Care and Use Committee (IACUC), and all methods were carried out following relevant guidelines and regulations, which were also approved by the IACUC and Institutional Biosafety Committees (IBC) at Case Western Reserve University.

Mice

Wild-type C57BL/6J mice were obtained from Jackson Laboratory. Mice were kept in sterile plastic cages with pelleted bedding in an SPF facility until they were euthanized by transcardial perfusion. There was unlimited access to food and water, a 12h cycle of light and dark, and enrichment from a mouse igloo. Mice were selected for stereotaxic injection based on age and sex and divided into experimental groups for the relevant experiments. Direct mouse studies were carried out by individuals who were blind to their experimental injection groups. All mice had been healthy up to the point of euthanasia and had no obvious abnormalities.

525 **Antibodies and reagents**

526 Anti-HA (Cat# 51064-2-AP), anti- β -actin (Cat# 66009-1-Ig), and anti-MTCH2
527 (Cat#16888-1-AP) antibodies were purchased from Proteintech (IL, USA). The anti-
528 GFAP (Cat # 14-9892-82) antibody was purchased from Thermo Scientific (MA, USA).
529 Anti-Iba1(Cat# 019-19741) antibody was purchased from Wako Chemicals USA
530 (Richmond, VA, USA). Secondary fluorescent antibodies Alexa Fluor 488 (Cat # A-
531 11008 1 mg, Cat # A-11001
532 1 mg), Alexa Fluor 594 (Cat # A-11012 1 mg, Cat # A-21207 1 mg) were purchased
533 from Thermo Scientific (MA, USA).

534 **Cell culture**

535 HEK293T, HT22, and Tau RD-FRET biosensor cells were cultured in Dulbecco's
536 modified Eagle's medium (DMEM, Thermo Scientific, Waltham, MA, USA)
537 supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (P/S).
538 All cells were cultured in a humidified atmosphere (5% CO₂) at 37 C.

539 **DNA Constructs, Transfections**

540 pPM-MTCH2^{WT}-HA, pPM MTCH2^{p.P290A}-HA, and pPMMTCH2^{p.P290T}-HA constructs
541 (PC130856) were generated by Applied Biological Materials Inc. (Richmond, BC,
542 Canada). For DNA plasmid transfections, cells were transfected with Eugene HD
543 (Promega, Madison, WI, USA) in Opti-MEM I (Invitrogen, Carlsbad, CA, USA) according
544 to the manufacturer's instructions and harvested 48h post-transfection. For siRNA
545 transfections, Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) was used according
546 to the manufacturer's instructions.

Generation of rAAV9 and Stereotaxic Injections in Mice

Recombinant AAV9 viruses rAAV9-MTCH2^{WT}-HA and rAAV9-MTCH2^{p.P290T}-HA (VC130856) were produced by Applied Biological Materials Inc. (Richmond, BC, Canada). Five-month-old mice, balanced by sex across treatments, were put under anesthesia with isoflurane and secured in a stereotaxic apparatus (RWD, Sugar Land, TX, USA) for stereotaxic injections. Burr holes were precisely drilled into the skull utilizing a dental drill. Bilateral injections aimed at the hippocampus were conducted utilizing a 26-gauge needle attached to a 10-μL Hamilton syringe (Hamilton, Reno, NV, USA). The injection coordinates in relation to bregma were: anteroposterior -2.7 mm, lateral -2.7 mm, and dorsoventral 3.0 mm. 2.5 μL of pure rAAV9 (1×10^{11} vg/mL) was administered per site over a duration of 2 minutes via convection-enhanced delivery. After the injection, incisions were sutured and sanitized, and the animals were observed until recovery, generally within 20 minutes. Mice were individually housed and euthanized three months post-injection for subsequent analysis.

Cell/Tissue Lysis and SDS-PAGE

Cells and brain tissues were lysed using RIPA lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.1% sodium dodecyl sulfate) (citation). Total protein concentrations were measured with a colorimetric detection assay (Quick Start Bradford Protein Assay, Bio-Rad, Hercules, CA, USA). Equal amounts of protein lysates were subjected to sodium dodecyl sulfate (SDS) – polyacrylamide gel electrophoresis (PAGE) and subsequently transferred to a nitrocellulose membrane (Amersham Protran, Cytiva, Marlborough, MA, USA). In filter-trap assays (FTA), equivalent

quantities of RIPA-soluble or sonicated RIPA-insoluble extracts were filtered through 0.2 mm cellulose acetate membranes (CA023001, Sterlitech Corporation, Auburn, WA, USA) utilizing a 96-well vacuum dot blot apparatus (Bio-Rad, Hercules, CA, USA), followed by PBS washing and fixation with 20% methanol. Targeted proteins were analyzed using primary antibodies, subsequently followed by horseradish peroxidase-conjugated secondary antibodies and ECL detection (Jackson ImmunoResearch Inc., West Grove, PA, USA). All immunoblot images were obtained with ImageQuant 800 (Amersham, Chicago, IL, USA) and quantified utilizing ImageJ (NIH, Bethesda, MD, USA).

Immunocytochemistry, Immunohistochemistry, and Microscopy

For immunocytochemistry (ICC), cells were rinsed with PBS and fixed for 15 minutes at ambient temperature using 4% paraformaldehyde (PFA) (J19943-K2, ThermoFisher Scientific, Waltham, MA, USA). Fixed cells were rinsed with PBS, treated with a blocking solution (0.2% Triton X-100, 3% normal goat serum (NGS, Abcam, Cambridge, UK)) in PBS) for 1 hour, incubated with primary antibody at 4°C overnight, washed three times with PBS, and subsequently incubated with Alexa-488 or Alexa-594, conjugated secondary IgG antibodies (ThermoFisher Scientific, Waltham, MA, USA) for 1 hour at room. The slides were subsequently rinsed three times with PBS before being mounted with Fluoromount-G mounting solution (ThermoFisher Scientific, Waltham, MA, USA). For immunohistochemistry (IHC), mice were perfused with PBS, and half of the brains were preserved at -80°C for subsequent biochemical investigation. The remaining half of the brain was preserved in 4% PFA at 4°C for 48 hours, subsequently transferred to

30% sucrose for cryoprotection. Following washing, the tissues were incubated with 3% normal goat serum (NGS) at room temperature for 1 hour, then subjected to primary antibody application overnight at 4°C, and subsequently treated with secondary fluorescent antibodies (Alexa Fluor 488, Alexa Fluor 594) and DAPI for 1 hour at room temperature. Immunohistochemistry on human brain samples was performed similarly, with overnight blocking at 4°C and True Black staining at room temperature for 10 minutes after the application of the secondary antibody. Confocal images were obtained using a Nikon AX Ti2 confocal microscope (Tokyo, Japan), and ImageJ software (18 October 2023, NIH, Bethesda, MD, USA) was employed to quantify the immunoreactivities. In all ICC/IHC studies, the comparison images maintained uniform laser intensity, exposure duration, and filter specifications. During image acquisition and quantification, researchers were blinded to the conditions of the experiment, and regions of interest were selected randomly. Brightness and contrast changes were similarly applied to all comparative images.

Open Field and Novel Object Recognition

Mice aged six to seven months were individually housed for one week before to behavioral testing. Every mouse was handled for a minimum of 2 minutes daily. Mice were relocated to the behavior room for a 30-minute acclimatization period before testing commenced in the open area. Mice were thereafter positioned in the open field arena and let to explore for 10 minutes. Visual cues were affixed to the walls above the arena in various shapes. The test was documented by an overhead camera and evaluated using ANY-maze 6.3 tracking and analysis software. The investigator doing

the test was unaware of the genotype. The overhead lights were subdued, and the investigator was obscured by a black curtain during the examination. Novel object recognition (NOR) was conducted on littermate mice aged 6 to 7 months. Mice were allowed to go acclimatize to the testing environment for 30 minutes before the assessment. Identical objects were positioned in opposing quadrants of the arena, with their locations recorded. On Day 1, animals were permitted a 10-minute habituation period with the two identical objects. The arena and objects were meticulously cleansed between trials with mice using a 30% ethanol solution to eliminate smell cues. Following a 24-hour period, the mice were reintroduced to the behavioral chamber. One object was substituted with a new object that differed in dimensions, form, and texture. Mice were allotted 5 minutes to investigate both the novel and familiar things. The experiment was documented by an overhead camera and evaluated using ANY-maze 6.3 tracking and analytic software. The investigator performing the test was blind to the genotype. The overhead lights were subdued, and the investigator was shadowed by a black curtain during the test.

ACKNOWLEDGMENTS

Funding: This work was primarily supported by the National Institute on Aging (NIA) under Award Number U01AG073323, R01AG066707, R01AG084250, R01AG076448, R01AG082118, R01AG092462, R01AG092591, RF1AG082211, and R33AG083003, and the National Institute of Neurological Disorders and Stroke (NINDS) under Award Number RF1NS133812, the Alzheimer's Association award (ALZDISCOVERY-1051936), and the Alzheimer's Disease Drug Discovery Foundation (ADDF) to F.C. This

639 work was supported in part by the Rebecca E. Barchas, MD, Professorship in
640 Translational Psychiatry, the Valour Foundation, Project 19PABH134580006-AHA/Allen
641 Initiative in Brain Health and Cognitive Impairment, the Elizabeth Ring Mather & William
642 Gwinn Mather Fund, S. Livingston Samuel Mather Trust, and the Louis Stokes VA
643 Medical Center resources and facilities to A.A.P. This work was supported in part by
644 Keep Memory Alive (KMA), NIGMS grant P20GM109025, NINDS grant U01NS093334,
645 NIA grant R01AG053798 and R35AG071476, and the ADDF to J.C. This work was
646 partly supported by the Alzheimer's Association award (ALZDISCOVERY-1051936) to
647 F.C. This work was also supported by grants from the National Institutes of Health
648 (NIH) [R01AG086365 to T.L, R03AG084948 to T.L]. T.L was also supported by the
649 Research Education component of the Cleveland Alzheimer's Disease Research Center
650 (NIA P30 AG072959).

651 Data collection was supported through funding by National Institute on Aging
652 (NIA) grants P30AG10161, R01AG15819, R01AG17917, R01AG30146, R01AG36836,
653 U01AG32984, U01AG46152 and U01AG61356 and by the Illinois Department of Public
654 Health. Mayo RNA-seq study data were provided by the following sources: the Mayo
655 Clinic Alzheimer's Disease Genetic Studies, led by N. Ertekin-Taner and S. G. Younkin
656 (Mayo Clinic, Jacksonville, Florida), using samples from the Mayo Clinic Study of Aging,
657 the Mayo Clinic Alzheimer's Disease Research Center and the Mayo Clinic Brain Bank.
658 Data collection was supported through funding by NIA grants P50 AG016574, R01
659 AG032990, U01 AG046139, R01 AG018023, U01 AG006576, U01 AG006786, R01
660 AG025711, R01 AG017216 and R01 AG003949; by National Institute of Neurological
661 Disorders and Stroke (NINDS) grant R01 NS080820; Mount Sinai Brain Bank data were

generated from postmortem brain tissue collected through the Mount Sinai VA Medical Center Brain Bank and were provided by E. Schadt of the Mount Sinai School of Medicine through funding from NIA grant U01AG046170. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Conflicts of Interest

Dr. Cummings has provided consultation to AB Science, Acadia, Alkahest, AlphaCognition, ALZPathFinder, Annovis, AriBio, Artery, Avanir, Biogen, Biosplice, Cassava, Cerevel, Clinilabs, Cortexyme, Diadem, EIP Pharma, Eisai, GatehouseBio, GemVax, Genentech, Green Valley, Grifols, Janssen, Karuna, Lexeo, Lilly, Lundbeck, LSP, Merck, NervGen, Novo Nordisk, Oligomerix, Ono, Otsuka, Pharmacotrophix, PRODEO, Prothena, ReMYND, Renew, Resverlogix, Roche, Signant Health, Suven, Unlearn AI, Vaxxinity, VigilNeuro pharmaceutical, assessment, and investment companies. The other authors have declared no competing interests.

Author Contributions

F.C. T.L., and Y.H. conceived the study. Y.H. performed all genetic and genomic data analysis and experiments. T.L., L.W., D.B., and T.M.T performed cellular and animal experiments. X.C. and Y.Y. conduct CRISPRi experiments using human iPSC-derived microglia. W.S., A.A.P, J.A.P., Z.L., Y.L., Z.T., M.H., Y.E.L. X.Z., T.L., and J.C. interpreted the data analysis. Y.H., F.C., P.Z. drafted the manuscript. Y.H., L.W., and

T.L. draft the manuscript. Y.H., L.W., T.L., F.C., A.A.P., and J.C. critically revised the manuscript. All authors critically revised and gave final approval of the manuscript.

Data Availability

The GWAS summary statistic data available in GWAS Catalog and NIAGADS include: (1) five AD GWAS summary statistics: Bellenguez, 2022 (GCST90027158); Schwartzenruber, 2021 (GCST90012877 & GCST90012878); Jansen, 2019 (GCST007320); and AD Kunkle Stage 1/2 (NG00075); (2) two PD GWAS summary statistics (GCST009325 and GCST90275127); and (3) two ALS GWAS summary statistics (GCST90027163 and GCST90027164). Cell type-specific eQTL data from Bryois et al. were downloaded from <https://zenodo.org/records/7276971>. Cell type-level eQTL summary statistics from Fujita et al. are available at Synapse (<https://doi.org/10.7303/syn52335732>). Two additional microglia-specific eQTL datasets are available from <https://zenodo.org/records/4118605> and <https://doi.org/10.7303/syn26207321>. The snATAC-seq and snRNAseq data are available at https://personal.broadinstitute.org/bjames/AD_snATAC/.

Code Availability

All codes used in this study is freely available at: <https://github.com/ChengF-Lab/ND-MR>

Declarations

Ethics approval and consent to participate: Not applicable.

References

1. Fujita, M. *et al.* Cell subtype-specific effects of genetic variation in the Alzheimer's disease brain. *Nat Genet* **56**, 605-614 (2024).
2. Liu, Z. *et al.* Astrocytic response mediated by the CLU risk allele inhibits OPC proliferation and myelination in a human iPSC model. *Cell Rep* **42**, 112841 (2023).
3. Jang, B. *et al.* SingleBrain: A Meta-Analysis of Single-Nucleus eQTLs Linking Genetic Risk to Brain Disorders. *medRxiv* (2025).
4. Uffelmann, E. *et al.* Genome-wide association studies. *Nature Reviews Methods Primers* **1**, 59 (2021).
5. Tam, V. *et al.* Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics* **20**, 467-484 (2019).
6. Kosoy, R. *et al.* Alzheimer's disease transcriptional landscape in ex vivo human microglia. *Nature Neuroscience* (2025).
7. Yang, X. *et al.* Functional characterization of Alzheimer's disease genetic variants in microglia. *Nature Genetics* **55**, 1735-1744 (2023).
8. Liu, Z. *et al.* Single-cell multiregion epigenomic rewiring in Alzheimer's disease progression and cognitive resilience. *Cell*.
9. Arvanitis, M., Tayeb, K., Strober, B.J. & Battle, A. Redefining tissue specificity of genetic regulation of gene expression in the presence of allelic heterogeneity. *Am J Hum Genet* **109**, 223-239 (2022).
10. Hawrylycz, M. *et al.* SEA-AD is a multimodal cellular atlas and resource for Alzheimer's disease. *Nature Aging* **4**, 1331-1334 (2024).

11. Mathys, H. *et al.* Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. *Cell* **186**, 4365-4385.e27 (2023).
12. Kamath, T. *et al.* Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease. *Nature Neuroscience* **25**, 588-595 (2022).
13. Pineda, S.S. *et al.* Single-cell dissection of the human motor and prefrontal cortices in ALS and FTLD. *Cell* **187**, 1971-1989.e16 (2024).
14. Yang, H.-S. *et al.* Cell-type-specific Alzheimer's disease polygenic risk scores are associated with distinct disease processes in Alzheimer's disease. *Nature Communications* **14**, 7659 (2023).
15. Haglund, A. *et al.* Cell state-dependent allelic effects and contextual Mendelian randomization analysis for human brain phenotypes. *Nature Genetics* **57**, 358-368 (2025).
16. Song, X. *et al.* MiXcan: a framework for cell-type-aware transcriptome-wide association studies with an application to breast cancer. *Nature Communications* **14**, 377 (2023).
17. Mancuso, N. *et al.* Probabilistic fine-mapping of transcriptome-wide association studies. *Nat Genet* **51**, 675-682 (2019).
18. Tak, Y.G. & Farnham, P.J. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* **8**, 57 (2015).

- 752 19. Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer's disease
753 and related dementias. *Nat Genet* **54**, 412-436 (2022).
- 754 20. Kunkle, B.W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease
755 identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat*
756 *Genet* **51**, 414-430 (2019).
- 757 21. Jansen, I.E. *et al.* Genome-wide meta-analysis identifies new loci and functional
758 pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404-413 (2019).
- 759 22. Schwartzenuber, J. *et al.* Genome-wide meta-analysis, fine-mapping and
760 integrative prioritization implicate new Alzheimer's disease risk genes. *Nat Genet*
761 **53**, 392-402 (2021).
- 762 23. Wightman, D.P. *et al.* A genome-wide association study with 1,126,563
763 individuals identifies new risk loci for Alzheimer's disease. *Nat Genet* **53**, 1276-
764 1282 (2021).
- 765 24. Cheng, F. *et al.* *Combining genetics with real-world patient data enables*
766 *ancestry-specific target identification and drug discovery in Alzheimer's disease*,
767 (2024).
- 768 25. Diaz-Ortiz, M.E. *et al.* GPNMB confers risk for Parkinson's disease through
769 interaction with α -synuclein. *Science* **377**, eabk0637 (2022).
- 770 26. Bogacki, E.C., Longmore, G., Lewis, P.A. & Herbst, S. GPNMB is a biomarker for
771 lysosomal dysfunction and is secreted via LRRK2-modulated lysosomal
772 exocytosis. *bioRxiv*, 2025.01.01.630988 (2025).

- 773 27. Soutar, M.P.M. *et al.* Regulation of mitophagy by the NSL complex underlies
774 genetic risk for Parkinson's disease at 16q11.2 and MAPT H1 loci. *Brain* **145**,
775 4349-4367 (2022).
- 776 28. Wainberg, M., Andrews, S.J. & Tripathy, S.J. Shared genetic risk loci between
777 Alzheimer's disease and related dementias, Parkinson's disease, and
778 amyotrophic lateral sclerosis. *Alzheimers Res Ther* **15**, 113 (2023).
- 779 29. Bacmeister, C.M. *et al.* Motor learning promotes remyelination via new and
780 surviving oligodendrocytes. *Nat Neurosci* **23**, 819-831 (2020).
- 781 30. Xiong, X. *et al.* Epigenomic dissection of Alzheimer's disease pinpoints causal
782 variants and reveals epigenome erosion. *Cell* **186**, 4422-4437.e21 (2023).
- 783 31. Lecca, D. *et al.* Role of chronic neuroinflammation in neuroplasticity and
784 cognitive function: A hypothesis. *Alzheimers Dement* **18**, 2327-2340 (2022).
- 785 32. Merolla, F. *et al.* Loss of CCDC6, the first identified RET partner gene, affects
786 pH2AX S139 levels and accelerates mitotic entry upon DNA damage. *PLoS One*
787 **7**, e36177 (2012).
- 788 33. Maliar, N.L., Talbot, E.J., Edwards, A.R. & Khoronenkova, S.V. Microglial
789 inflammation in genome instability: A neurodegenerative perspective. *DNA*
790 *Repair (Amst)* **135**, 103634 (2024).
- 791 34. Meshref, M. *et al.* The Role of RIN3 Gene in Alzheimer's Disease Pathogenesis:
792 a Comprehensive Review. *Mol Neurobiol* **61**, 3528-3544 (2024).
- 793 35. Shen, R. *et al.* Upregulation of RIN3 induces endosomal dysfunction in
794 Alzheimer's disease. *Transl Neurodegener* **9**, 26 (2020).

- 795 36. Guna, A. *et al.* MTCH2 is a mitochondrial outer membrane protein insertase.
796 *Science* **378**, 317-322 (2022).
- 797 37. Buzaglo-Azriel, L. *et al.* Loss of Muscle MTCH2 Increases Whole-Body Energy
798 Utilization and Protects from Diet-Induced Obesity. *Cell Rep* **14**, 1602-1610
799 (2016).
- 800 38. Kim, J.J. *et al.* Multi-ancestry genome-wide association meta-analysis of
801 Parkinson's disease. *Nature Genetics* **56**, 27-36 (2024).
- 802 39. van Rheenen, W. *et al.* Common and rare variant association analyses in
803 amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic
804 architectures and neuron-specific biology. *Nat Genet* **53**, 1636-1648 (2021).
- 805 40. Wingo, A.P. *et al.* Integrating human brain proteomes with genome-wide
806 association data implicates new proteins in Alzheimer's disease pathogenesis.
807 *Nat Genet* **53**, 143-146 (2021).
- 808 41. Sieberts, S.K. *et al.* Large eQTL meta-analysis reveals differing patterns between
809 cerebral cortical and cerebellar brain regions. *Scientific Data* **7**, 340 (2020).
- 810 42. de Klein, N. *et al.* Brain expression quantitative trait locus and network analyses
811 reveal downstream effects and putative drivers for brain-related diseases. *Nature*
812 *Genetics* **55**, 377-388 (2023).
- 813 43. Bryois, J. *et al.* Cell-type-specific cis-eQTLs in eight human brain cell types
814 identify novel risk genes for psychiatric and neurological disorders. *Nat Neurosci*
815 **25**, 1104-1112 (2022).
- 816 44. Lopes, K.P. *et al.* Genetic analysis of the human microglial transcriptome across
817 brain regions, aging and disease pathologies. *Nat Genet* **54**, 4-17 (2022).

- 818 45. Burgess, S. & Thompson, S.G. Interpreting findings from Mendelian
819 randomization using the MR-Egger method. *Eur J Epidemiol* **32**, 377-389 (2017).
- 820 46. Burgess, S., Foley, C.N., Allara, E., Staley, J.R. & Howson, J.M.M. A robust and
821 efficient method for Mendelian randomization with hundreds of genetic variants.
822 *Nature Communications* **11**, 376 (2020).
- 823 47. Qi, G. & Chatterjee, N. Mendelian randomization analysis using mixture models
824 for robust and efficient estimation of causal effects. *Nature Communications* **10**,
825 1941 (2019).
- 826 48. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread
827 horizontal pleiotropy in causal relationships inferred from Mendelian
828 randomization between complex traits and diseases. *Nature Genetics* **50**, 693-
829 698 (2018).
- 830 49. Wallace, C. A more accurate method for colocalisation analysis allowing for
831 multiple causal variants. *PLoS Genet* **17**, e1009440 (2021).

Figure Legends

Figure 1. Mendelian randomization analysis in neurodegenerative disease. a.

Workflow of genome-wide MR analysis, colocalization studies, and validation using single-nucleus RNA-seq and ATAC-seq data. **b.** Circle plot displaying cell-type specific MR results across Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS). The colors of points denote the significant targets at $p < 1 \times 10^{-6}$ for AD (green), PD (orange) and ALS (purple) and the sizes of point denote normalized MR scores for each gene.

Figure 2. Cell type-specific risk genes for AD, PD, and ALS supported by genetic

evidence. Heatmap showing ND-MR scores, colocalization (Coloc PP.H4 > 0.5), and fine-mapping (FOCUS, PIP > 0.5) evidence for significant genes across 6 brain cell types and two brain tissue eQTL and pQTL data. The green color shows the ND-MR score, represents increased gene expression will increase disease risk and red color represents increased gene expression will reduce disease risk; colocalization from light blue (PP.H4 = 0.5) to dark blue (PP.H4 = 1); FOCUS from light orange (PIP = 0.5) to dark orange (PIP = 1); gray indicates no data. Upset plot shows the number of significant genes and validation overlap for each disease. Bottom stacked barplot shows gene distribution across cell types.

Figure 3. Single-cell multi-omics validation of SNP-gene associations across AD

and PD. a. Upset plot shows snATACseq and snRNAseq validated SNP-gene association pairs across AD and PD. Numbers indicate validated pairs for each disease

and overlap between validation methods. **b.** Network visualization of cell type specific SNP-gene regulatory relationships. Circular pie plots represent different brain cell types (microglia, astrocytes, excitatory neurons, inhibitory neurons, oligodendrocytes, and OPC), with each connected node indicating a validated SNP-gene pair. Circle nodes denote the p value of GWAS datasets.

Figure 4. The case studies of SNP-gene pairs validated by peak-to-gene analysis and iPSC derived microglia. **a.** snATAC-seq and peak-to-gene loop validated SNP-gene association pairs across AD and PD. Different colored arch lines denote cell type specific significant peak-to-gene associations and the colorful peaks indicate chromatin accessibility in snATAC-seq data. **b.** Correlation analysis between gene expression and amyloid and tau densities across different brain cell types. The shape denotes the absolute correlation values and gradient color indicates the $-\log_{10}$ (p-value). **c.** CRISPRi experimental validation of causal relationship between SNP rs1171830 and CCDC6 expression in human iPSC-derived microglia.

Figure 5. *MTCH2* shows strong genetic evidence for AD risk in neurons.

a, Mendelian randomization analysis of *MTCH2* in excitatory neurons and inhibitory neurons. CI: confidence intervals. **b,** Scatter plot showing the *MTCH2* locus from AD GWAS (Wightman et al., 2021, excluding UKBiobank and 23andMe). *MTCH2* eQTLs from excitatory and inhibitory neurons show colocalization with AD GWAS at the lead SNP rs3740688 (red diamond). PP.H4: Posterior Probability of Hypothesis 4. **c,** Western blot shows the reduced *MTCH2* protein level in AD compared to control human

brain samples. **d**, Filter trap assay shows the increased insoluble MTCH2 level in normal controls and AD. **e – f**, Quantification of soluble MTCH2 levels and insoluble MTCH2 aggregation in AD patients' brain (n=10) and controls (n = 10). **g**, Immunostaining analysis shows aggregated MTCH2^{P290A}-HA and MTCH2^{P290T}-HA mutations compared to MTCH2^{WT}-HA and vector control in HT22 cells (an immortalized mouse hippocampus cell line). The Mito-DsRed adenovirus (red color) was utilized for mitochondrial labeling. **h**, Western blots show soluble and insoluble proteins of transfected MTCH2^{WT}-HA, MTCH2^{p.P290A}-HA, MTCH2^{p.P290T}-HA, and vector control in HEK293T cells. **i**, Quantification of MTCH2 soluble/insoluble protein abundance across wildtype and MTCH2 p.P290A and p.P290T mutations (n = 6)

Figure 6. MTCH2 shows strong genetic evidence in vivo in mouse brains

a, Immunostaining analysis of tau seeding activity in TauRD cells transfected with MTCH2-WT-HA, MTCH2-p.P290A-HA, MTCH2-p.P290T-HA, and vector control. **b**, Quantification of tau intensity among MTCH2 variants and vector control. **c**, Immunostaining of MTCH2-HA in the brains of C57BL6/J mice injected with rAAV9-MTCH2^{WT}-HA and rAAV9-MTCH2^{p.P290T}-HA. **d**, Immunostaining of GFAP intensity in the hippocampal region of C57BL6/J mice injected with rAAV9-MTCH2^{WT}-HA and rAAV9-MTCH2^{p.P290T}-HA. **e**, Quantification of GFAP intensity (n = 4 mice/group). **f**, Novel Object Recognition test comparing rAAV9-MTCH2^{WT}-HA and rAAV9-MTCH2^{p.P290T}-HA injected mice regarding time spent with novel objects, habituated object, exploration time, and the discrimination index between time with novel and habituated object. (n = 6 mice/group). **g**, Open field test comparing rAAV9-MTCH2^{WT}-HA and rAAV9-

MTCH2^{p.P290T}-HA injected mice regarding time spent in inner, middle, and outer zones.
(n = 6 mice/group).

Supplementary Figure Legends

Supplementary Figure 1. Workflow of Mendelian Randomization analysis across Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS).

Supplementary Figure 2. Colocalization of brain region eQTLs with Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS) GWAS signals. Three different shapes represent AD, PD, and ALS GWAS data sources. Twelve colors denote eQTL data sources from different brain regions.

Supplementary Figure 3. Colocalization of brain cell type specific eQTLs with Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS) GWAS signals. Three different shapes represent AD, PD, and ALS GWAS data sources. Colors denote eQTL data sources from different cell types.

Supplementary Figure 4. FOCUS fine-mapping analysis for Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS). Colors indicate different GWAS datasets and shapes indicate the QTL data used for analysis.

Supplementary Figure 5. The case studies of SNP-gene pairs validated by peak-to-gene analysis for Parkinson's disease (PD). **a.** snATAC-seq and peak-to-gene loop validated SNP-gene association pairs across PD. Different colored arch lines denote cell type specific significant peak-to-gene associations and the colorful peaks indicate chromatin accessibility in snATAC-seq data.

Supplementary Figure 6. The case studies of SNP-gene pairs validated by peak-to-gene analysis for amyotrophic lateral sclerosis (ALS). **a.** snATAC-seq and peak-to-gene loop validated SNP-gene association pairs across ALS. Different colored arch lines denote cell type specific significant peak-to-gene associations, and the colorful peaks indicate chromatin accessibility in snATAC-seq data.

Supplementary Figure 7. *MTCH2* eQTL data and expression level across different human brain cell types. **a,** Cell type specific eQTL of *MTCH2* nonsynonymous mutation rs10646608 p.P290A (C>G). OPC: oligodendrocyte precursor cells. **b-c,** *MTCH2* expression levels across cell types (**b**) and brain regions (**c**).

Supplementary Figure 8. *MTCH2* shows strong genetic evidence in vivo in mouse brains by AAV experiments. **a,** The mitochondrial genome expression comparing in the groups of *MTCH*-WT, P290A, P290T and controls. **b,** Representative images of IBA1 in the brains of C57BL6/J mice injected with rAAV9-*MTCH2*^{WT}-HA and rAAV9-*MTCH2*^{p.P290T}-HA. Quantification of IBA1 intensity (t-test, ***P < 0.001, n = 4 mice/group).

Supplementary table legends

Table S1. Mendelian Randomization (MR) scores to prioritize likely risk genes for neurodegenerative disease.

Table S2. Summary of cell type specific genes for Alzheimer's disease (AD),

Table S3. Summary of cell type specific genes for Parkinson's disease (PD)

Table S4. Summary of cell type specific genes for amyotrophic lateral sclerosis (ALS).

Table S5. Brain region- and cell type-specific gene colocalization in Alzheimer's disease (AD),

Table S6. Brain region- and cell type-specific gene colocalization in Parkinson's disease (PD).

Table S7. Brain region- and cell type-specific gene colocalization in amyotrophic lateral sclerosis (ALS).

Table S8. Disease risk SNPs validated by chromatin accessibility in snATAC-seq data.

Table S9. Disease risk SNPs validated by peak to gene analysis.

Figure 1

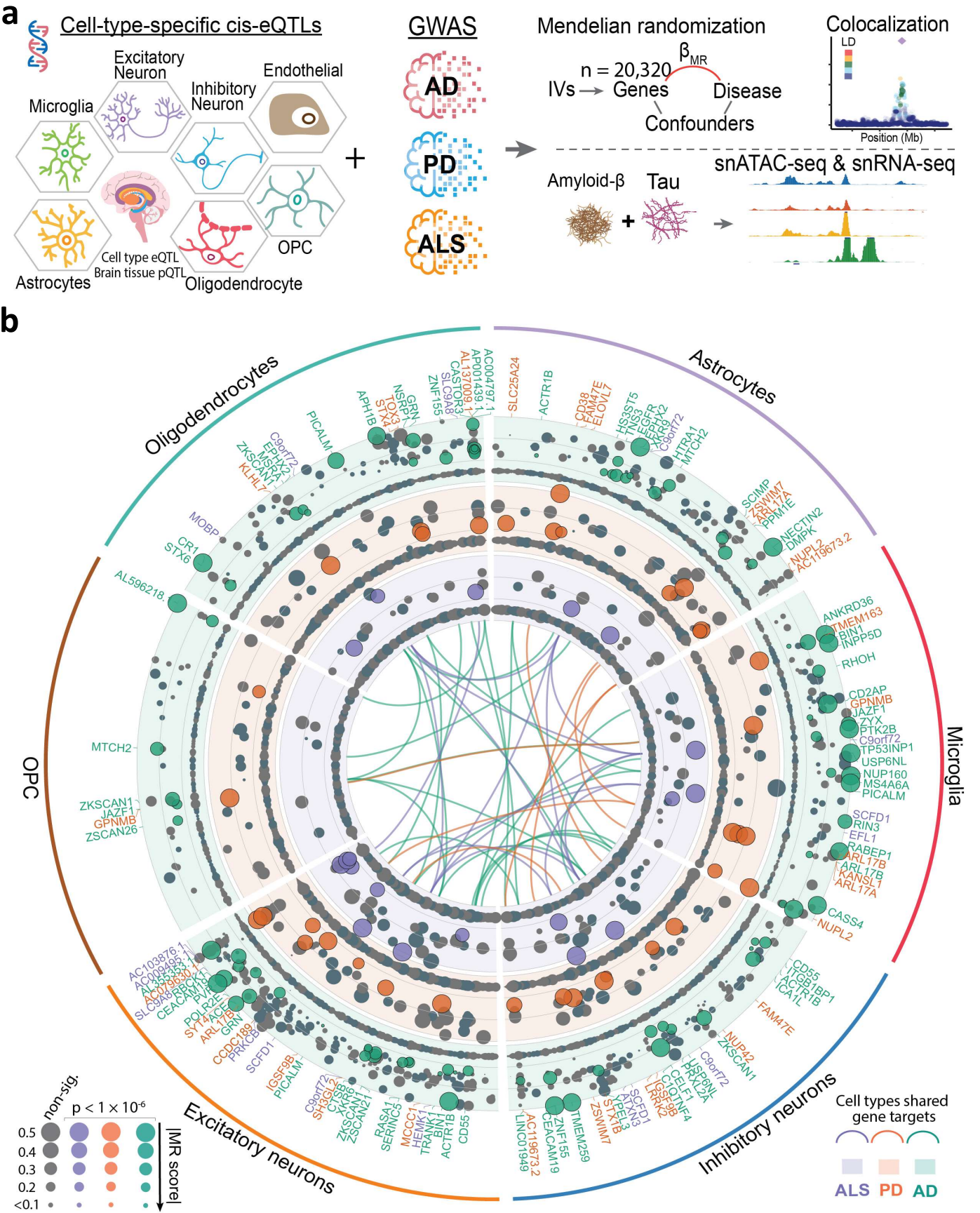


Figure 2

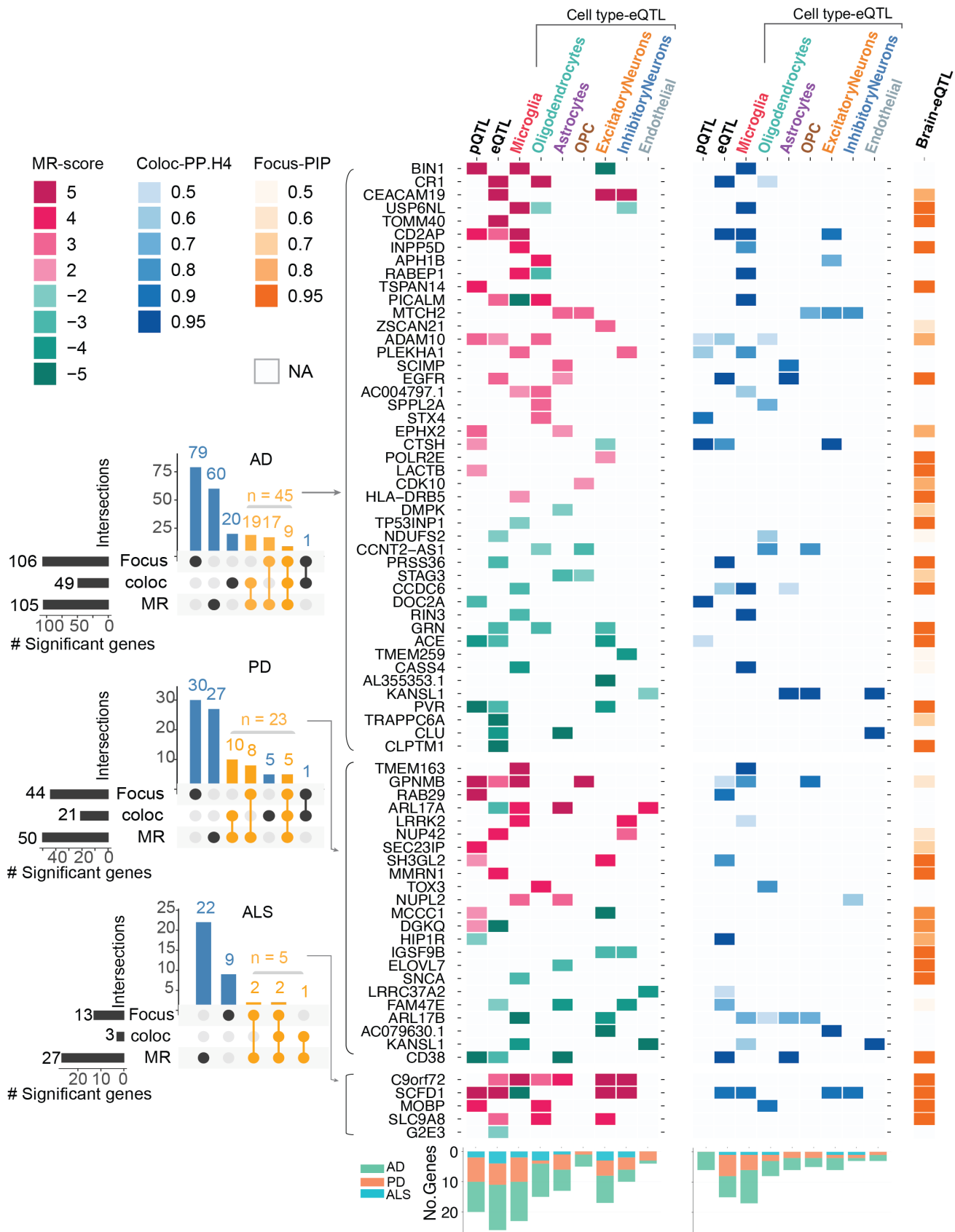


Figure 3

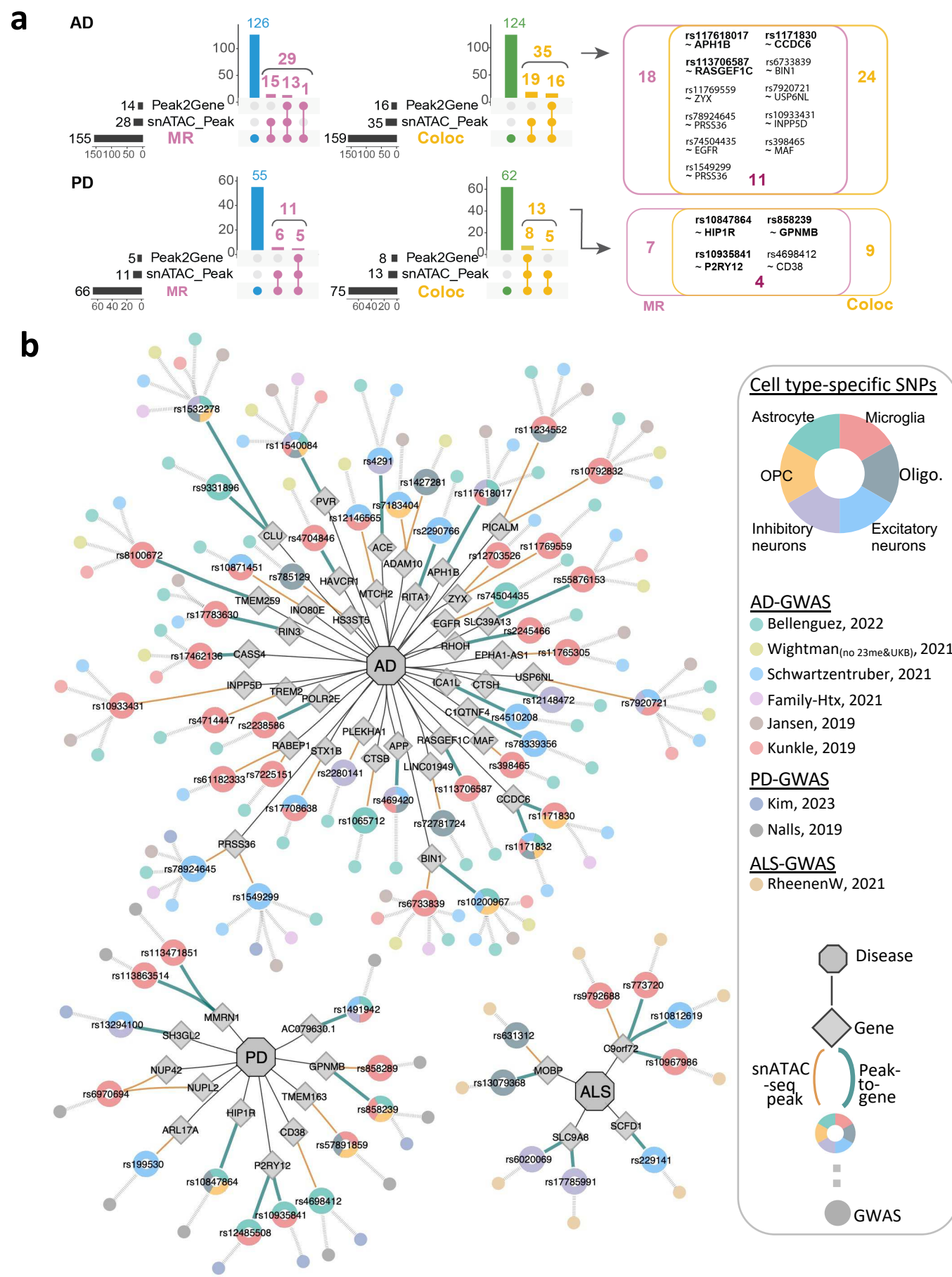


Figure 4

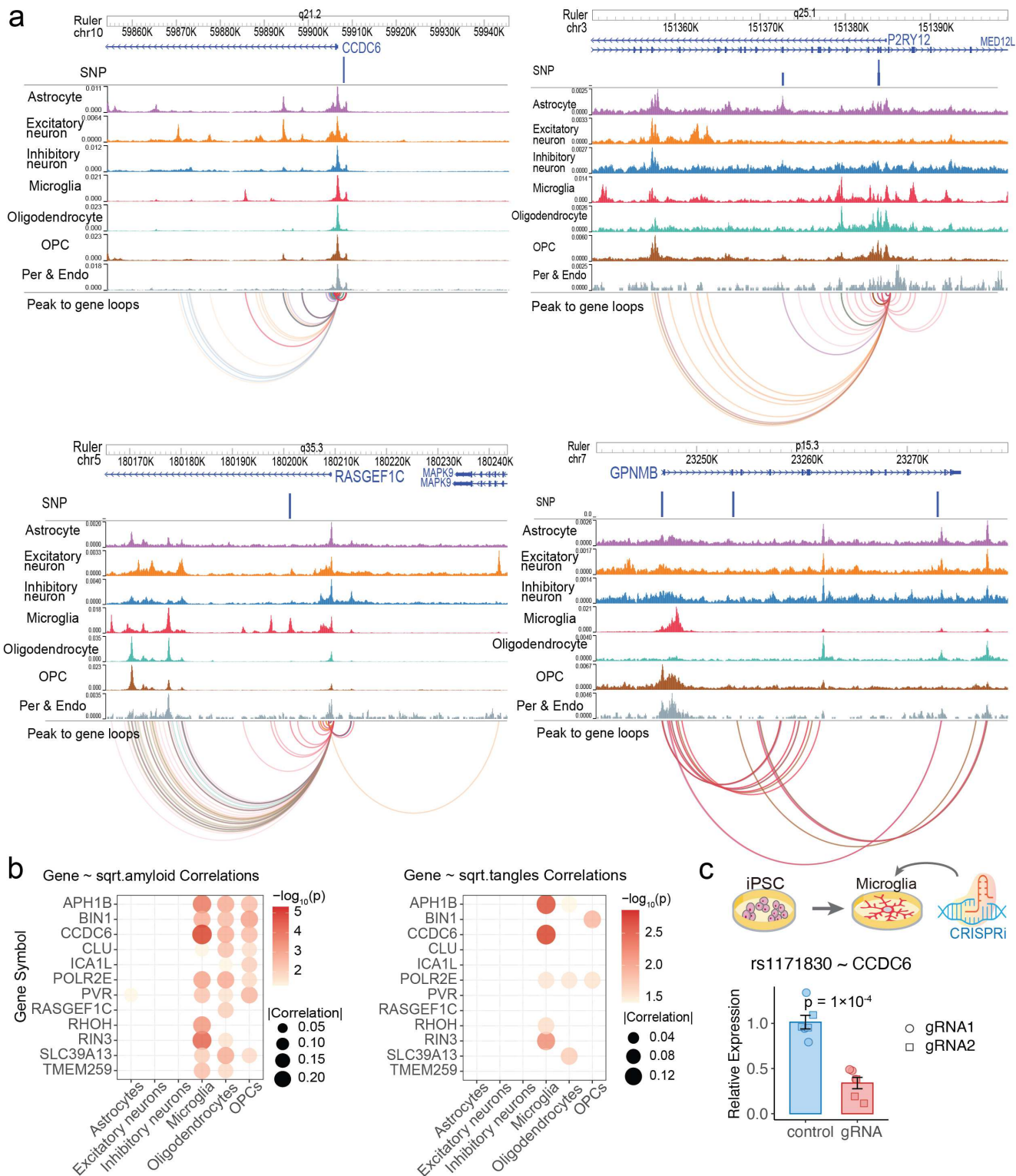


Figure 5

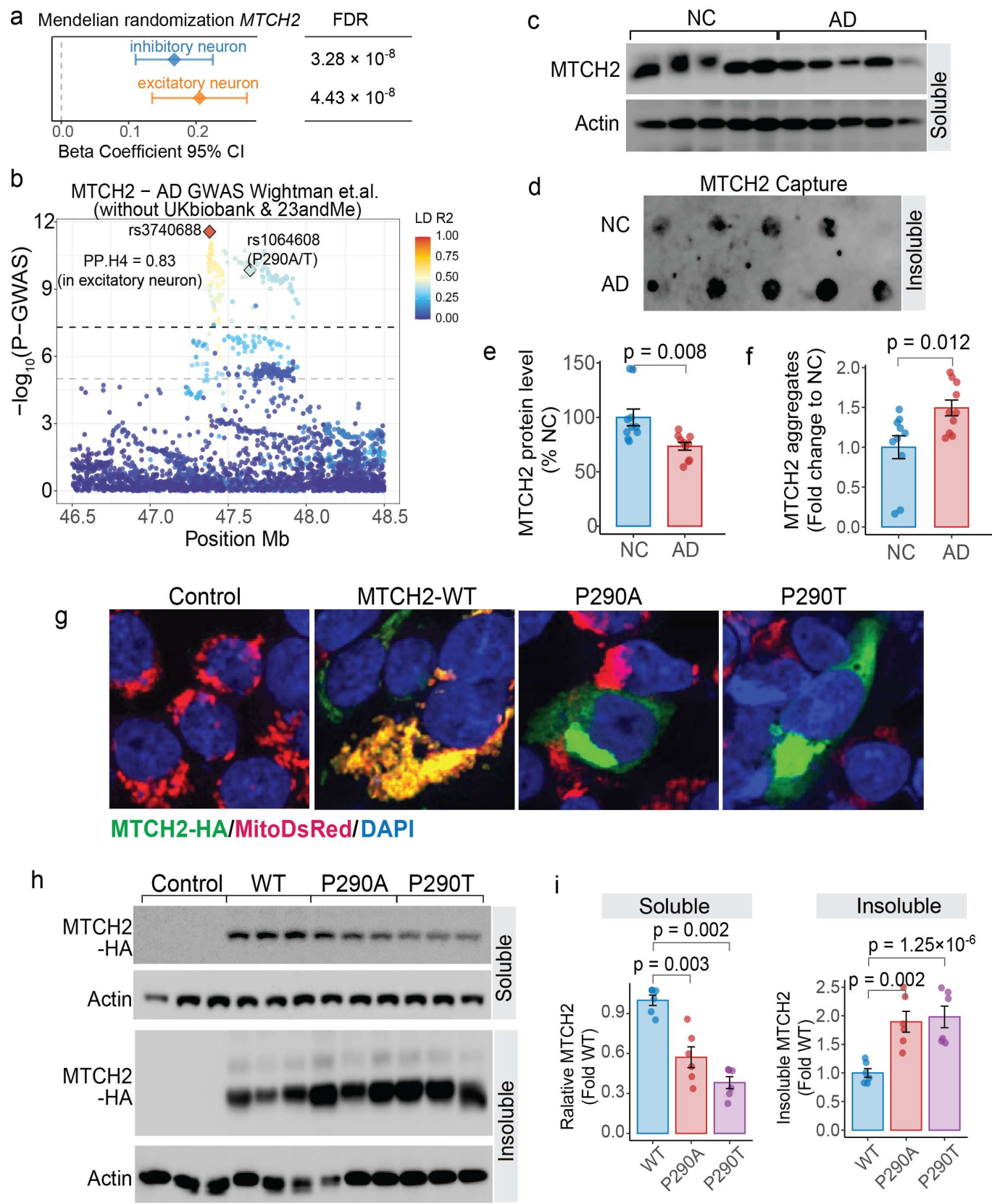
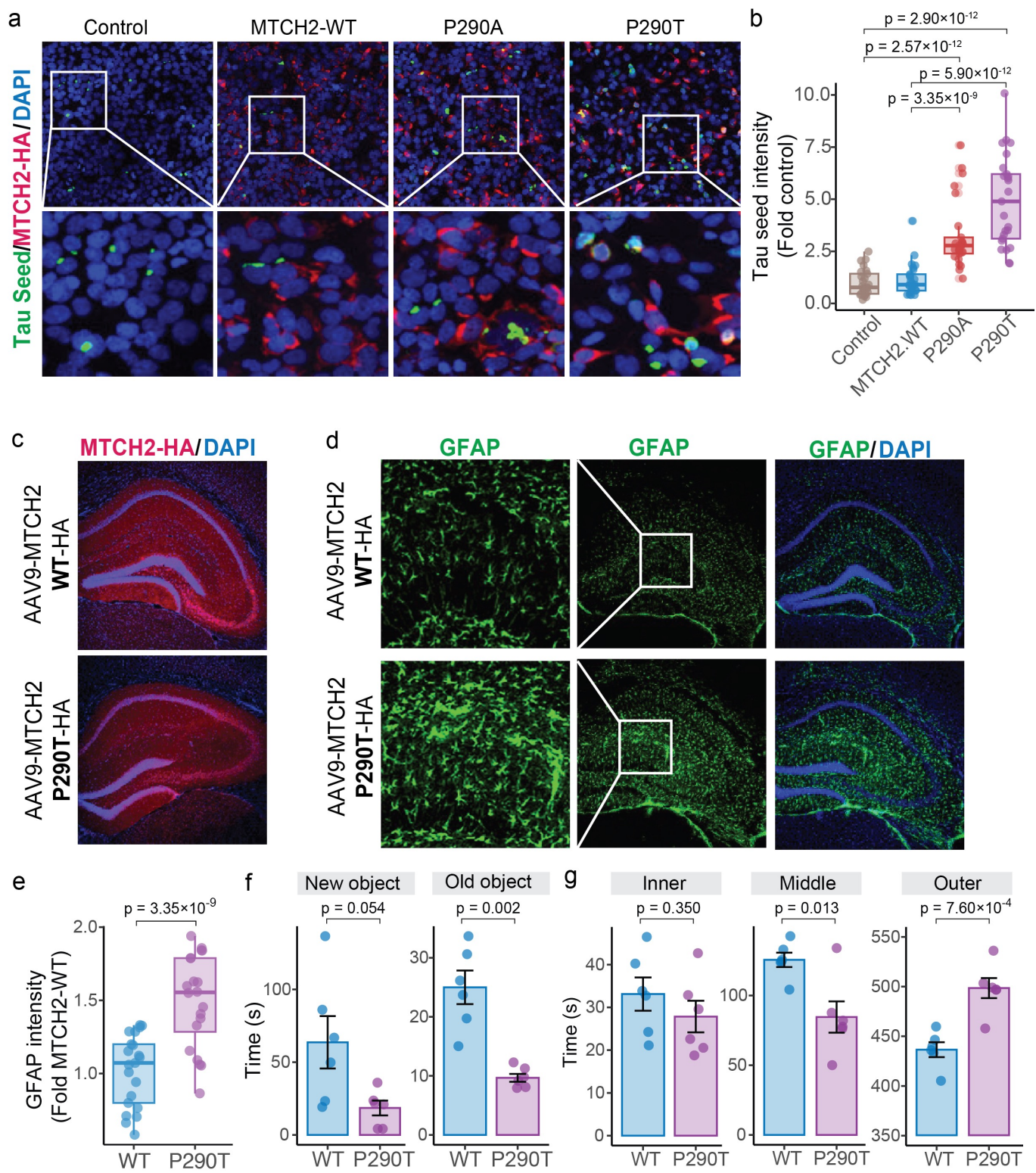


Figure 6



Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFiguresS1S8Cheng.pdf](#)
- [SupplementaryTablesS1S9.zip](#)
- [westernblotdata.zip](#)