Neuroimaging-derived brain endophenotypes link molecular mechanisms to Alzheimer's disease and aging

Ruixi Li[1,†], Ru Feng[1,2,†], Anjing Liu[1], Xuewei Cao[1,2], Philip L. De Jager[3,4,5], David Bennett[6], The Alzheimer's Disease Functional Genomics Consortium, Christos Davatzikos[7,*], Junhao Wen[8,9,*], Gao Wang[1,5,10,*]

[1]Center for Statistical Genetics, The Gertrude H. Sergievsky Center, Columbia University, New York, NY, USA
[2]Computational and Systems Biology, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[3]Center for Translational & Computational Neuroimmunology, Columbia University, New York, NY, USA
[4]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY, USA
[5]Department of Neurology, Columbia University, New York, NY, USA
[6]Rush Alzheimer's Disease Center and Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA
[7]Artificial Intelligence in Biomedical Imaging Laboratory, Center for AI and Data Science for Integrated Diagnostics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[8]Laboratory of AI and Biomedical Science, Department of Radiology, Columbia University, New York, NY, USA
[9]Department of Biomedical Engineering, Columbia University, New York, NY, USA
[10]Department of Biostatistics, Columbia University, New York, NY, USA

[†] These authors contributed equally to this work
***Corresponding authors:**
wang.gao@columbia.edu, jw4745@cumc.columbia.edu, christos.davatzikos@pennmedicine.upenn.edu

ABSTRACT

Alzheimer's disease (AD) genome-wide association studies (GWAS), typically based on clinical phenotypes, have identified numerous risk loci, yet linking these variants to brain changes and molecular processes remains challenging. We developed a DNE-xQTL framework integrating deep learning-derived dimensional neuroimaging endophenotypes (DNEs) with comprehensive brain molecular quantitative trait loci (xQTL) to dissect genetic pathways underlying AD- and aging-related brain variation. By performing GWAS on seven DNEs and applying integrative computational analyses, we biologically annotated each DNE and prioritized xQTL-supported gene targets. This approach both enhanced interpretation of established AD loci through DNE-mediated annotations and revealed underexplored regulatory pathways, organizing 209 candidate genes into evidence-based tiers. We highlight three regulatory clusters: glutamate-receptor and mitochondrial pathways implicating excitatory-neuron vulnerability, SREBP2-associated cholesterol homeostasis linked to vascular dysfunction, and primary-cilia-associated transport implicated in aging. By connecting pre-symptomatic brain

alterations to molecular targets and relevant cell types, this framework may inform earlier risk stratification before clinical neurodegeneration occurs.

INTRODUCTION

Understanding the molecular architecture of pre-symptomatic brain aging is critical for identifying therapeutic targets before irreversible neurodegeneration occurs. Traditional endophenotypes, including cerebrospinal fluid (CSF) biomarkers (amyloid-β, tau), PET imaging, and structural MRI measures (hippocampal volume, cortical thickness), typically quantify single features and often miss subtle early changes that vary across individuals. Beyond practical limitations of CSF invasiveness and PET costs and radiation exposure, these univariate measures fundamentally fail to capture the heterogeneous spatial patterns characteristic of individual aging trajectories and disease presentations. Dimensional neuroimaging endophenotypes (DNEs), derived from deep learning applied to multivariate MRI data, address these limitations by extracting complex brain-wide patterns as quantitative dimensions that capture individual variability and detect subtle pre-symptomatic alterations. However, DNEs remain abstract imaging constructs lacking biological interpretability: the cellular processes driving these brain changes are uncharacterized, limiting molecular target discovery for early intervention.

Genome-wide association studies (GWAS) have identified approximately 100 AD risk loci but cannot reveal how genetic variants manifest as brain changes or disease mechanisms. Limitations in AD phenotyping further complicate gene discovery: clinical AD diagnosis demonstrates limited accuracy, with autopsy studies revealing 10–30% discordance due to phenotypic overlap with vascular, Lewy body, and other pathologies[1]; moreover, recent large-scale GWAS increasingly rely on proxy phenotypes from parental history reports[2,3], which introduce survival bias and nonrandom participation effects that produce misleading genetic correlations[4]. Endophenotypes, as intermediate traits, offer a path forward by enabling genetic variants to serve as instrumental variables for causal inference[5]. Recent studies leveraged this principle, performing GWAS on DNE dimensions followed by Mendelian randomization (MR) against disease GWAS, establishing putative causal relationships from DNE to AD risk[6,7]. While demonstrating that imaging patterns influence endpoints, this approach cannot explain underlying molecular mechanisms or identify therapeutic targets. Separately, molecular quantitative trait loci (xQTL) mapping links genetic variants to molecular phenotypes (e.g. gene expression and protein levels), revealing how variation influences brain molecular states but not how these connect to *in vivo* pre-symptomatic brain alterations. Each approach addresses only part of the pathway: DNE-endpoint integration may establish causality but remains mechanistically opaque; xQTL-endpoint integration identifies molecular associations but lacks intermediate phenotypes to dissect how molecular changes manifest as observable brain alterations. No study has systematically integrated imaging endophenotypes, comprehensive multi-context brain xQTLs, and endpoint GWAS to simultaneously characterize what DNEs represent biologically and connect them to neurodegeneration.

Here, we developed an integrative genetic framework applied to seven MRI- and AI-derived DNEs from the MULTI and iSTAGING consortia (n=32,000–43,100), large-scale AD and aging GWAS (n=462,666–1,958,774), and brain xQTL from the FunGen-xQTL consortium across 48 molecular contexts from 2,328 individuals (**Figure 1**) to systematically map the molecular architecture of pre-symptomatic brain aging and neurodegeneration. This *DNE-xQTL framework* enabled multi-scale evidence synthesis through fine-mapping and multi-GWAS colocalization to identify putative causal genetic loci; brain xQTL mapping to link variants to genes across molecular modalities and cell types; comprehensive transcriptome-wide association studies (TWAS) to establish gene-level evidence; integration of variant-level colocalization with gene-level TWAS to reduce false-positive gene identification, with cell-type-resolved causal analyses applied to genomically complex loci containing multiple candidates; and pathway analysis to biologically annotate each DNE. We revealed that each DNE captures distinct cellular processes (*iAD1*: cytoskeletal dysfunction; *iAD2*: excitotoxicity; *iAging1-5*: lipoprotein transport, synaptic protein degradation, autophagy-extracellular vesicle impairment, mitochondrial adaptation, and neuroinflammation-driven vascular damage). By consolidating loci-level and gene-level evidence linking DNEs to AD and aging (telomere length), we organized 209 genes into evidence-based tiers and identified three gene clusters that may connect DNEs to clinical outcomes: a synaptic dysfunction cluster revealing oligodendrocyte contributions to glutamate receptor trafficking and excitotoxicity (*TNIP1*, *NSF*, *SEZ6L2*); a vascular-cholesterol cluster linking cholesterol dysregulation to cerebrovascular integrity (*NBEAL1*, *SCAP*); and a primary cilia cluster connecting AD to aging through intraflagellar transport and extracellular vesicle pathways (*CBY1*, *MAP4*, *BIN1*). This hypothesis-generating framework enables biologically motivated gene prioritization beyond what conventional GWAS and xQTL studies can provide, revealing candidate molecular targets for early intervention.

RESULTS

Genetic fine-mapping and multi-trait colocalization identify shared architecture between neuroimaging endophenotypes and neurodegeneration

For the seven DNE traits (*iAD1*, *iAD2*, *iAging1*, *iAging2*, *iAging3*, *iAging4*, and *iAging5*), we conducted fine-mapping and multi-GWAS colocalization analyses to pinpoint putative causal variants and assess genetic sharing across traits. *iAD1-2* represents Alzheimer's disease related imaging dimensions trained on Alzheimer's Disease Neuroimaging Initiative (ADNI) case and control data. *iAging1-5* consists of different normative aging dimensions derived from a large multi-cohort population dataset from the iSTAGING consortium[6,8]. The trained models were applied to UKBB populations (n=32,829–33,541) to get continuous scores for each DNE trait, followed by GWAS adjusting for age, sex, total intracranial volume (ICV) and genetic principal components (**Methods**). Sample size of DNE identification and GWAS, total count of the overall and significant variants for each DNE were summarized in **Table S1**. Bivariate linkage disequilibrium score regression (LDSC) quantified genome-wide polygenic sharing

across DNE pairs (**Figure 2**), with strongest genetic correlations occurred between *iAD2* and *iAging2* (r_g = 0.70, P = 1.53×10^{-108}), and between *iAD1* and *iAging3* (r_g = 0.67, P = 6.17×10^{-94}), with additional significant correlations observed among other DNE pairs (**Figure S1**, **Table S3**).

Genome-wide fine-mapping[9] with a new Linkage Disequilibrium (LD) reference panel for European ancestry[10] identified 2,131 variants within 83 95% credible sets (CS) across seven DNEs (**Methods**; **Figure S2**). Multi-GWAS colocalization using *ColocBoost*[11] identified 1,690 variants within 43 95% colocalized confidence sets (CoS) shared across DNE traits (**Figure 2a, Figure S3**). The distribution of variants per credible set (CS) from fine-mapping analysis showed that 93% (40/43) of CS contained fewer than 100 variants, and colocalization analysis revealed that 91% (60/66) of CoS contained fewer than 100 variants, demonstrating that both methods achieve high-resolution genetic mapping for most regions (**Figure S4**). The strongest loci sharing occurred between *iAD2* and *iAging2* (12 CoS), followed by *iAD1* and *iAging3* (7 CoS), consistent with both genome-wide genetic correlation (**Figure 2**) and fine-mapping overlap patterns (3 95% CS and 2 95% CS, respectively) (**Figure S2**). Multi-GWAS colocalization identified proportionally more shared loci (2.5-3 folds increase) than fine-mapping alone. The correlation between variant level fine-mapping posterior inclusion probability (maxPIP) and colocalization probability (maxVCP) was high (r = 0.90) (**Figure S5**), yet only 31% CoS had overlap with fine-mapped 95% CS, demonstrating that multi-trait analysis elevated many lower-confidence signals to 95% threshold through shared evidence across traits, consistent with previously observed power gains[11]. Indeed, 38.8% non-overlapping CoS can be recovered by fine-mapping when we relax the confidence level to 70% CS coverage. Therefore, for downstream analysis we included all 95% CoS as cross-traits effects alongside fine-mapped 95% CS as per trait effects. Hereafter, we refer to these CS and CoS collectively as *loci*.

To assess whether DNE loci disproportionately contribute to disease heritability, we applied stratified linkage disequilibrium score regression (sLDSC) using variants from fine-mapped and colocalized loci as functional annotations, testing whether sparse, mappable signals explain substantial heritability beyond diffuse genome-wide background captured by previous bivariate LDSC analysis. We analyzed six AD GWAS and two aging GWAS (telomere length and mvAge) using variants with maximum variant colocalization probability (maxVCP), an approach demonstrating robust performance[11] (**Methods**). In AD, we observed significant enrichment for *iAging2* (Enrichment = 97.93, SE = 29.51, $p < 0.05$), *iAD2* (Enrichment = 49.32, SE = 33.20, $p < 0.05$), while *iAging5* showed substantial but non-significant enrichment (Enrichment = 42.00, SE = 21.73) (**Figure 2f**). No significant enrichment was detected for aging traits due to the lack of sufficient sample size (number of variants) in 95% CoS genome-wide, although *iAging5*, *iAD1 and iAging3* showed moderate to strong effects (**Figure 2f**). These enrichment patterns are consistent with the bivariate LDSC results, indicating that *iAD2* and *iAging2* jointly contribute to AD pathology, *iAD1* and *iAging3* primarily reflect normal aging processes, and iAging5 appears to influence both. These results establish DNE utility as intermediate phenotypes for gene discovery.

Given DNE loci disproportionately contribute to disease heritability, we leveraged multi-trait colocalization to boost power for identifying endpoint GWAS loci, particularly those shared with DNEs and potentially explained through intermediate imaging phenotypes. We performed

multi-GWAS colocalization across 16 GWAS traits, including seven DNE traits and endpoint GWAS for Alzheimer's disease (AD), aging (telomere length and mvAge), and Parkinson's disease (PD). For endpoint outcomes, we identified 5,530 AD variants across 101 loci (95% CoS or 95% CS), 4,338 aging variants across 283 loci, and 1,748 PD variants across 16 PD loci. The large variant counts reflect complex LD structure, particularly the *MAPT* locus (chr17:45383525-50162864) capturing 1,561 variants across *iAging3*, AD and PD (**Figure S6**). All 95% CoS reached suggestive significance ($p < 1\times10^{-5}$), with 64.94%, 87.50%, and 44.62% reaching genome-wide significance ($p < 5 \times10^{-8}$) for AD, aging, and PD respectively (**Figure S7**). Among these, 2,224 variants within 13 loci (denoted as orange in **Figure 2**) colocalized across endpoint outcomes and DNEs, representing loci where DNE integration enables finer dissection of endpoint genetic architecture. In comparison, naive overlap of 95% CS between endpoints and DNEs identified only 1 loci. We prioritized these 13 loci as providing strong variant-level evidence linking DNEs to clinical outcomes, which will contribute in a multi-tiered gene prioritization framework described later.

Brain molecular QTL integration reveals cellular processes underlying neuroimaging endophenotypes

Having established DNE genetic architecture at the variant level, we next characterized the biological processes underlying each DNE before examining their connections to clinical endpoints. We integrated brain molecular QTL data from the FunGen-xQTL consortium, encompassing 48 molecular contexts across bulk and single-nucleus expression, splicing, proteomic, glycoproteomic, metabolomic, and epigenomic QTLs from ROSMAP, MSBB, and Knight ADRC cohorts (n=1,183 total) **(Table S2)**, with DNE GWAS summary statistics through complementary variant-level and gene-level approaches (**Figure 1**; **Methods**). Our xQTL data demonstrated high coverage: more than 60% of DNE loci overlapped with at least one xQTL locus across the 48 contexts (**Figure S8**). We applied pretrained prediction models from FunGen-xQTL to DNE GWAS for transcriptome-wide association studies (TWAS), then integrated TWAS with xQTL colocalization following the putative causal gene (PCG) analysis framework integrating colocalization and TWAS via the *INTACT* method[12], which requires convergent variant-level and gene-level evidence to reduce false-positives particularly due to linkage disequilibrium (LD) hitchhiking when multiple genes are involved in a locus (**Methods**).

After integrating both lines of evidence, we identified gene sets for each DNE, with the largest for *iAging5* (147 genes), *iAging1* (127 genes), and *iAD2* (97 genes) (**Figure 2**, **Table S8**). High gene-level sharing was observed between *iAD2* and *iAging2* (16 genes), and between *iAD2* and *iAging5* (11 genes), mirroring the loci-level colocalization patterns from the previous section and suggesting these trait pairs capture related neurobiological processes.

We performed Gene Ontology (GO) and KEGG enrichment analysis to annotate DNE (**Methods**). xQTL-derived gene sets provided substantially more specific enrichments than positional mapping of nearby genes (**Figure S9, Table S9–11**), illustrating that functional

genomic integration beyond conventional GWAS loci are biologically informative. We focused interpretation on pathways reaching statistical significance (FDR adjusted $p < 0.1$).

Pathway enrichment defined distinct cellular processes for each DNE (**Figure 3**, **Table S9**). In summary, *iAD1* genes showed enrichment in axolemma, microtubule-associated complex, and tubulin binding, indicating cytoskeletal and axonal transport dysfunction. *iAD2* genes implicated glutamate receptor-mediated excitotoxicity, neuroinflammation, and blood-brain barrier disruption through adherens junction and immune-related pathways. Joint analysis of *iAD1* and *iAD2* strengthened ciliary basal body and cell projection membrane enrichment (**Figure S10, Table S12**). *iAging1* genes were enriched in high-density lipoprotein particle and transport vesicle membrane, suggesting disrupted lipoprotein vesicle transport under ER stress. *iAging2* genes, genetically correlated with *iAD2*, showed enrichment in protein deubiquitination and glutamatergic synapse pathways in joint enrichment, implicating dysregulated ubiquitin-mediated protein degradation at excitatory synapses. *iAging3* genes were enriched in autophagy, VEGF signaling, and hormone secretion, suggesting autophagy impairment and extracellular vesicle dysfunction. Another pair of DNEs with high genetic coherence, *iAD1* and *iAging3*, show shared-gene enrichment in cell projection. This aligns well with the joint enrichment observed for the two AD-derived DNEs in the GO/KEGG analyses and with the AD enrichment detected in the MSigDB Hallmark analysis (**Figures S10–11**). *iAging4* genes showed suggestive evidence in ubiquinone and ketone biosynthetic processes, and significant mitochondrial pathway enrichment when jointly analyzed with *iAD2*, indicating mitochondrial metabolic adaptation during aging complementary to function captured by *iAD2*. *iAging5* genes demonstrated the strongest enrichments among all DNE in peptide antigen binding, ER-to-Golgi transport vesicle membrane, T cell-mediated immunity, and focal adhesion assembly, indicating immune activation, ER-Golgi transport failure, and blood-brain barrier dysfunction potentially through impaired endothelial cell adhesion.

Collectively, xQTL integration biologically characterized each DNE, establishing them as interpretable intermediate phenotypes for mechanistic dissection of disease. With each DNE now defined by specific cellular processes, we leveraged these biological annotations to interpret gene-disease and gene-endpoint connections through DNE-mediated pathways in subsequent analyses, with a particular focus on genes relevant to excitotoxicity and neuroinflammation-mediated vascular damage as these are biological processes annotated to DNE whose genetic correlation are highest with AD (**Figure 2e–f, Figure S11**).

Multi-tiered evidence framework prioritizes genes linking DNEs to disease endpoints

We next identified which DNE-associated genes influence disease endpoints, by applying a similar variant- and gene-level integration framework to endpoint GWAS combining xQTL colocalization with TWAS across AD, aging (telomere length and multivariate integration of aging traits including healthspan, parental lifespan, exceptional longevity, EAA and frailty), and

PD, to prioritize candidate genes where genetic variation influences both brain structure and clinical outcomes. DNEs exhibited substantial gene-level overlap with disease endpoints. The largest overlap was with telomere length (238 shared genes), with *iAging5* contributing most (64 genes), followed by *iAging3* (41 genes), *iAD2* (27 genes), and *iAging1* (17 genes) (**Figure 2g**). Notably, aging-related DNEs (*iAging1-5*), derived from multiple cohorts including UK Biobank capturing population-level brain aging, showed larger overlaps with AD than AD-specific DNEs (*iAD1*: 0 genes; *iAD2*: 17 genes) derived from clinical cohorts to capture disease-specific atrophy. AD showed more modest overlap (51 genes total), with largest contributions from *iAging3* (26 genes), *iAD2* (18 genes), and *iAging5* (9 genes). A three-way intersection of AD, telomere length, and DNEs encompassed 14 genes, suggesting shared genetic architecture between AD and aging dissectable through DNE-mediated pathways.

While integrating DNEs as intermediate phenotypes enhances power, it introduces complexity in systematically weighing multi-scale evidence across molecular contexts, endophenotypes, and endpoints. We therefore developed empirical prioritization rules categorizing candidate genes into evidence-based tiers (**Table 1**, **Figure 1e**). **Tier 1** (27 genes) included genes supported by both loci-level evidence (colocalization) and gene-level evidence (TWAS) in both endpoints and DNEs, representing the strongest convergent evidence. Among Tier 1 genes, 27 mapped to variants within 10loci colocalized between DNEs and endpoints previously identified, including *TNIP1*, *NBEAL1*, and *CBY1*. **Tier 2** (35 genes) comprised genes with both loci-level and gene-level evidence in DNEs plus gene-level evidence in endpoints, including *SEZ6L2* and *NSF*. **Tier 3** (147 genes) included genes with either loci-level or gene-level evidence in DNEs alongside gene-level evidence in endpoints, such as *MAP4* and *DNAJC5*. While Tier 1 represents highest-confidence targets, Tiers 2–3 capture additional genes that, when connected through shared DNE annotations or pathways, construct comprehensive mechanistic hypotheses. Mendelian randomization (MR) was used on endpoints as a screening criterion for tier assignment and also was applied to complement putative causal gene (PCG) analysis in selected gene targets as additional validation, with results presented in **Table S16–17**.

Integrating evidence across tiers and DNE annotations revealed three gene clusters connecting brain endophenotypes to disease outcomes. The synaptic dysfunction cluster encompasses genes mediating mitophagy suppression (*TNIP1*, *KANSL1*; Tier 1), glutamate receptor dysregulation (*NSF*; Tier 1 and *SEZ6L2*; Tier 2), and calcium imbalance (*DOC2A*; Tier 2), linking *iAD2* and *iAging2* to AD through excitotoxicity. The vascular-cholesterol cluster includes *NBEAL1* (Tier 1) and *SCAP* (Tier 3), connecting *iAging5* to AD and vascular dementia through cholesterol dysregulation. The primary cilia cluster comprises *CBY1* (Tier 1) plus *MAP4*, *BIN1*, *DNAJC5*, *TRIM26*, *GGT7*, and four additional genes (Tiers 2–3), linking *iAD1* and *iAging3* to both AD and aging (telomere length) via autophagy-extracellular vesicle pathways. For genomic loci containing multiple candidate genes, we applied cell-type-specific PCG analysis integrating DNE and endpoint evidence to distinguish putative causal genes from LD hitchhikers. Genes meeting stringent PCG criteria with strong mechanistic support are presented in main text, while genes in complex loci (such as the *MAPT* region on chr17) are detailed in **Supplementary Notes** with locus- and context-specific rationale.

In subsequent sections, we present detailed evidence for these gene clusters, focusing on loci with complete xQTL-DNE-endpoint evidence that reveal underexplored regulatory pathways. Established AD genes where our framework provides DNE-mediated annotations are summarized in **Table S12**. We leverage cell-type-resolved xQTL data to dissect locus-specific mechanisms, examining directional effects across cellular contexts, with *in silico* validation using orthogonal data (AD pathology, cross-species models) to generate mechanistic hypotheses connecting pre-symptomatic brain alterations to disease.

Cell-specific roles of synaptic dysfunction and excitotoxicity in neurodegenerative disorders

Mitophagy suppression in excitatory neurons. Cell-type-resolved DNE-xQTL integration identified *TNIP1* (chr5, Tier1) with excitatory neuron-specific contribution to excitotoxicity connecting *iAD2* and *iAging2* to AD. *TNIP1* encodes a regulator involved in mitophagy and NF-κB mediated inflammatory control, with known roles in cellular stress and mitochondrial quality regulation (**Table 1**). Specifically, *TNIP1* expression is positively associated with AD ($p_{TWAS} = 3.2 \times 10^{-11}$), *iAD2* ($p_{TWAS} = 5.1 \times 10^{-8}$), and *iAging2* ($p_{TWAS} = 8.1 \times 10^{-7}$) in excitatory neurons (**Figure 4a**), and the lead variant rs871269 (chr5:151052827:C>T) colocalized across these traits and two excitatory neuron eQTL datasets (**Figure 4b**). Consistent with its association with lower AD risk and DNE scores, rs871269 also demonstrated a negative association with TDP-43 burden in ROSMAP samples (nominal p = 0.02) (**Figure 4c-will be replaced with stronger validations?**). Complete multi-omic evidence, variant-level colocalization with multiple AD GWAS datasets are provided in **Supplementary Note S4.**

AMPA receptor dysregulation and mitophagy at the MAPT locus. The 17q21.31 region harbors 22 Tiers 1–3 genes in extensive LD due to a large inversion polymorphism[13], making conventional gene prioritization challenging. Previous GWAS suggest pleiotropic effects at this locus across neurodegenerative disorders including AD, PD, and amyotrophic lateral sclerosis (ALS)[5,14,15]. Cell-type-specific putative causal gene (PCG) analysis (**Methods**) integrating DNE and AD evidence identified *KANSL1*, *NSF*, and *CRHR1* as candidate causal genes at this locus, with distinct cell-type contributions (**Supplementary Note S1**). *KANSL1*, part of the NSL chromatin-modifying complex linked to mitochondrial and mitophagy regulation (**Table 1**), showed protective associations with AD and *iAD2* across excitatory neurons and oligodendrocytes. *NSF*, an ATPase required for AMPA receptor trafficking (**Table 1**), displayed oligodendrocyte-specific associations with DNE traits and exhibited opposing directions between *iAD2* and symptomatic AD. *CRHR1*, which modulates glutamatergic and stress-responsive signaling pathways (**Table 1**), showed significant associations in oligodendrocytes and excitatory neurons across DNE and AD endpoints. These genes collectively implicate both neuronal and glial mitochondrial dysfunction and glutamate receptor dysregulation as converging mechanisms at this genomically complex locus.

Oligodendrocyte AMPA receptor scaffolding at chromosome 16. The 16p11.2 region contains 26 Tiers 2–3 genes in high linkage disequilibrium, presenting similar prioritization challenges (**Table S13**). Cell-type-specific PCG analysis identified *SEZ6L2*, *DOC2A*, and *MAPK3* with oligodendrocyte-specific associations linking them to *iAging2* (**Supplementary Note S2**). *SEZ6L2* functions as an AMPA receptor scaffolding protein (**Table 1**) and showed the strongest oligodendrocyte-specific evidence, with opposing directional patterns between *iAging2* and AD that parallel those observed for *NSF*. *DOC2A*, which encodes a calcium sensor modulating glutamate release (**Table 1**), showed complex context-dependent associations across mRNA and protein levels in AC, DLPFC and excitatory neurons. *MAPK3*, relevant to p38-MAPK signaling, showed pleiotropic effects across AD, telomere length, *iAD2*, *iAging2* in oligodendrocytes and microglia (**Figure S12, Table S14**). Together, these genes reveal potential oligodendrocyte contributions to AMPA receptor-mediated excitotoxicity through receptor scaffolding, calcium sensing, and stress signaling mechanisms, expanding traditional neuron-centric models of excitotoxicity in AD.

Collectively, our findings suggest mitochondrial dysfunction and glutamate receptor dysregulation may synergistically contribute to excitotoxicity in neurons and oligodendrocytes, and are potentially relevant to pre-symptomatic changes to AD (**Figure 4d**).

Cell-type-resolved xQTL suggests NBEAL1-mediated vascular-cholesterol dysregulation linking *iAging5* and *iAD1* to AD

Cerebrovascular pathology co-occurs in over 80% of AD cases[16], but the cell-type-specific molecular mechanisms remain poorly characterized. Although our xQTL datasets lack cerebral endothelial cells, the biological annotations of *iAging5* (ER-Golgi transport failure, BBB disruption, vascular inflammation) and *iAD1* (axonal transport dysfunction) suggested vascular mechanisms warranting investigation. Our cell-type-resolved xQTL integration identified *NBEAL1* (Tier 1) and *SCAP* (Tier 3), which have been reported to implicate SREBP2-mediated cholesterol homeostasis as a mechanism connecting DNEs to cerebrovascular contributions to AD pathogenesis[17,18].

NBEAL1 is a regulator of vesicular trafficking and endolysosomal function (**Table 1**), exhibits positive associations between gene expression and AD across six brain regions ($p_{TWAS} < 2.5 \times 10^{-6}$ in AC, DLPFC, PCC, FP, STG, IFG) with stronger associations for *iAging5* in the same regions (**Figure 4e;** also **Table S14** for TWAS p-values and **Table S15** for brain region abbreviations). Cell-type-resolved analysis demonstrated positive associations in excitatory neurons (AD: $p_{TWAS} = 3.60 \times 10^{-7}$; *iAging5*: $p_{TWAS} = 3.87 \times 10^{-18}$), astrocytes (*iAging5*: $p_{TWAS} = 1.3 \times 10^{-6}$), and oligodendrocytes (*iAging5*: $p_{TWAS} = 2.3 \times 10^{-9}$). GTEx data demonstrate abundant *NBEAL1* expression in arterial tissue despite the absence of cerebral endothelial cells in our xQTL datasets (**Figure S13**). The lead variant rs114123510 (chr2:202966489:T>A) colocalized across *iAging5*, AD and brain eQTL contexts including bulk tissues (AC, PCC) **and excitatory neurons** (**Figure 4f**).

SCAP, known to escort SREBPs from the ER to the Golgi (**Table 1**), exhibits positive associations between gene expression and *iAD1* in bulk brain tissue ($p_{TWAS} < 2.5 \times 10^{-6}$ in DLPFC, PCC, FP, PHG, IFG), but negative associations in excitatory neurons ($p_{TWAS} = 2.2 \times 10^{-8}$) and inhibitory neurons ($p_{TWAS} = 5.3 \times 10^{-8}$) (**Figure 4e**). The lead variant rs7628631 (chr3:46,921,725-47,513,711) colocalized across *iAD1*, *iAging5*, and brain eQTL datasets (AC, DLPFC and PCC, **Figure 4f**). Despite weak marginal association with *iAging5*, fine-mapping nonetheless placed these variants in a 70% credible set (**Figure 4g**).

NBEAL1 and *SCAP* encode functionally coupled proteins regulating SREBP2 processing with established roles in endothelial cholesterol homeostasis and vascular disease (**Table 1**). Our integration reveals that both genes connect to AD through DNE capturing vascular dysfunction: *NBEAL1* associates primarily with *iAging5*, annotated for ER-Golgi transport failure, BBB disruption, and vascular inflammation (**Figure 3**); *SCAP* shows associations with both *iAD1* (axonal transport dysfunction) and *iAging5* (vascular dysfunction). The convergent evidence linking both genes to the SREBP2 pathway, combined with *NBEAL1* proximity to cerebral small vessel disease loci[19,20], implicates cerebrovascular cholesterol dysregulation as a mechanism contributing to AD pathogenesis (**Figure 4h**). Additional genes at the *NBEAL1* locus (*CARF*, *ICA1L*, *WDR12*) showed variant-level colocalization at the *NBEAL1* locus but lacked consistent gene-level support from PCG analysis and clear mechanistic connection to vascular dysfunction or cholesterol pathways; these results are detailed in **Supplementary Note S3**.

Primary cilia integrate autophagy and extracellular vesicle pathways linking DNE to telomere length and AD

Primary cilia are sensory organelles that regulate cellular homeostasis through intraflagellar transport and signaling pathways, with established roles in aging. DNE-xQTL integration identified ten genes spanning 7 genomic loci associated with telomere length and/or AD that converge on ciliary function through two possible mechanisms: extracellular vesicle-mediated intercellular communication (*BIN1*, *CBY1*, *DNAJC5*, *TRIM26*, *GGT7* via *iAging3*), and structural regulation of cilia and autophagy (*MAP4*, *KIF9*, *SETD2*, *MON1A*, *USP19* via *iAD1*) (**Figure 5g**). LD structure between lead variants from colocalized sets of genes (**Table S13**) and prioritization rationale are detailed in **Supplementary Note S4**.

Ciliogenesis and extracellular vesicle communication. *CBY1* (Tier 1), an essential gene for primary ciliogenesis (**Table 1**), showed positive associations with telomere length and negative associations with AD in excitatory neurons (**Figure 5e**). The lead variant rs5750633 (chr22:38597462:A>G) colocalized across telomere length, AD, *iAging3*, and brain eQTL datasets including DLPFC, excitatory neurons, inhibitory neurons, and oligodendrocytes (**Figure 5a**), with multi-trait colocalization supported by strong molecular QTL evidence. Despite weaker marginal GWAS *iAging3* association, we consider this DNE relevant because its annotated biological processes involve autophagy impairment and VEGF-related processes (**Figure 3**, **Table 1**). *BIN1*, a tau-interacting protein involved in extracellular vesicle-mediated tau trafficking and the second most strongly associated AD locus after *APOE* (**Table 1**), demonstrated negative associations with *iAging3* in inhibitory neurons and with AD in excitatory neurons, but positive associations with AD in other contexts including microglia, revealing cell-type-dependent roles in EV-mediated tau propagation. Additional genes (*DNAJC5*, *TRIM26*,

GGT7) connect telomere length to neurodegeneration through EV-mediated pathways, extending known mechanisms of tau spreading and EV regulation to aging processes through *iAging3* (**Table 1**). Complete TWAS statistics, variant-level colocalization, and functional details for all EV communication genes are provided in **Supplementary Note S3**.

Structural ciliary regulation and autophagy. *MAP4*, known to regulate primary cilium length through modulation of axonemal microtubule stability (**Table 1**), demonstrated cell-type-specific associations exclusively in oligodendrocytes: negative with telomere length and positive with *iAD1* (**Figure 5e**). *MON1A* and *USP19*, two autophagy-related genes involved in autophagosome maturation and protein quality control, showed distinct associations with telomere length and *iAD1*: *MON1A* showed positive associations with both telomere length and *iAD1* in DLPFC, while *USP19* showed negative associations with telomere length and positive associations with *iAD1* in monocytes (**Figure 5a**). Additional supporting genes (*KIF9*, *SETD2*), along with complete statistical evidence and variant-level colocalization for all structural regulation genes, are detailed in **Supplementary Note S3**.

DISCUSSION

Our study demonstrates that integrating imaging-derived brain endophenotypes with multi-omics molecular QTL profiles and disease GWAS provides a coherent genetic framework for interpreting early neurobiological changes. By linking imaging-derived endophenotypes to gene regulation across multiple molecular modalities, the DNE-xQTL framework identifies mechanisms underlying pre-symptomatic brain alterations and their contributions to neurological disease. Our *DNE-xQTL framework* addresses critical gaps in current approaches for understanding disease mechanisms: whereas GWAS identify disease risk loci without revealing how variants manifest as brain changes, and xQTL studies link genetics to molecular phenotypes without capturing *in vivo* neurobiology, our integration simultaneously characterizes the molecular architecture of pre-symptomatic brain alterations and connects them to disease through convergent genetic evidence.

Specifically, using DNEs as intermediate phenotypes boosts statistical power to detect disease-relevant genes, especially at loci with modest AD signals ($P < 1\times10^{-5}$) that gain support through shared imaging and molecular evidence. Secondly, integrating convergent genetic and functional evidence at variant- and gene-level through systematic statistical framework enables gene prioritization at complex loci to support focused mechanistic hypotheses. Third, cell-type-resolved xQTL integration revealed context-specific mechanisms, demonstrating that the same gene may exhibit divergent effects across cell types or disease stages, provides insights into disease etiology. Fourth, our framework can identify fine-mapped loci across molecular traits, endophenotypes and endpoints, prioritizing from thousands of associated variants a handful candidates as instrumental variables for causal inference and mechanistic studies.

Together, these advances suggests three putative regulatory hubs connecting pre-symptomatic brain changes to neurodegeneration: a synaptic dysfunction hub implicating oligodendrocyte contributions to glutamate receptor-mediated excitotoxicity; a vascular-cholesterol hub linking

BBB dysfunction to cerebrovascular dementia through SREBP2-mediated cholesterol dysregulation; and a primary cilia hub positioning cilia as integrators of cellular aging and intercellular protein propagation.

SREBP2-mediated cholesterol homeostasis. The proposed vascular-cholesterol hub demonstrates how SREBP2-mediated cholesterol homeostasis connects neuroimaging patterns potentially relevant to blood-brain barrier dysfunction and cerebrovascular inflammation. SREBP2 acts as the master regulator of cholesterol biosynthesis and also involved in immune responses and neurodegeneration progression. Within the vascular-cholesterol hub, convergent genetic evidence points to a regulatory locus on chromosome 2 containing *NBEAL1*, *CARF*, *ICA1L*, and *WDR12*, among which *NBEAL1* was prioritized based on convergent lines of evidence: strongest TWAS associations across six brain regions and three cell types (excitatory neurons, astrocytes, oligodendrocytes); genomic proximity to GWAS loci for lacunar stroke and cerebral small vessel disease[19,20]; and involvement in endothelial cholesterol homeostasis (e.g. upregulated heart endothelial expression in obesity[21]). Follow up *in silico* analysis in ROSMAP and GTEx further strengthened these roles. *NBEAL1* enhances SREBP2 proteolysis in endothelial cells, promoting cholesterol biogenesis in HUVECs[17]. Notably, the observed associations with *iAging5* in astrocytes and oligodendrocytes are particularly relevant given that these cell types being the primary sources of brain cholesterol[22]. Under physiological conditions, *SCAP* traffics SREBP2 for proteolytic activation and nuclear signaling that maintains cholesterol homeostasis[22], and also transports NLRP3 for controlled inflammasome activation[18]. We hypothesize that pathologically enhanced *NBEAL1*-mediated SREBP2 proteolysis drives excessive SCAP-SREBP2 activity concurrently with NLRP3-dependent inflammation, ultimately contributing to neurodegenerative outcomes. This cascade may result in hypercholesterolemia and chronic neuroinflammation that compromises vascular integrity of small cerebral vessels, ultimately promoting cognitive decline characteristic of vascular dementia (**Figure 6b**). The cell-type-resolved convergence of genetic evidence from *iAging5* (annotated for ER-Golgi transport failure, BBB disruption, and vascular inflammation) and *iAD1* (microtubule dysfunction), together with vascular disease AD and aging GWAS signals, alongside endothelial cholesterol biology, generates a testable hypothesis with candidate genetic variants as tools for functional validation. Understanding these pathways may deepen our knowledge on vascular dementia, with clinical translation targeting cholesterol dysregulation in early disease stages, with *iAD1* and *iAging5* serving as a screening endophenotype to identify patients at risk for this vascular mechanism.

Glutamate receptor dysregulation and mitochondrial dysfunction in excitotoxicity. The proposed synaptic dysfunction regulatory hub implicates glutamate receptor dysregulation and mitochondrial dysfunction in excitotoxicity — a fundamental mechanisms that initiate and propagate neurodegenerative cascades — with further evidence pointing to oligodendrocyte contributions expanding beyond neuron-centric models. Through DNE-xQTL integration, we prioritized genes within two genomically complex regions where conventional nearest-gene approaches cannot readily distinguish among multiple candidates. At MAPT locus (17q21.31) with extensive LD, we applied PGC analysis (**Methods**) integrating evidence across both DNE traits and disease endpoints, as xQTL and endpoint alone cannot resolve gene prioritization in such complex loci where 54 genes show xQTL-AD association evidence in our tiered genes.

This DNE-mediated dissection identified *KANSL1*, *NSF*, and *CRHR1*, with *KANSL1* showing consistent effects across *iAD2*, *iAging2*, and AD, while others showing opposing directional patterns are examined in **Supplementary Notes S1**. These observations suggest convergence of mitochondrial vulnerability and glutamatergic dysregulation that may influence both early excitotoxicity and later disease progression. Another LD-dense region on chromosome 16 shows a distinct but related regulatory architecture across 32 genes show xQTL-AD association evidences. DNE integration prioritized *DOC2A*, *SEZ6L2*, and *MAPK3* based on oligodendrocyte-specific associations with *iAD2* and *iAging2*. *MAPK3* showed broad pleiotropic effects across multiple contexts, whereas *SEZ6L2* and *DOC2A* displayed directions consistent with compensatory responses in early aging and potential post-transcriptional modulation of calcium homeostasis in glutamatergic synapse dysfunction. Notably, *NSF* and *SEZ6L2* — both AMPA receptor regulators — showed oligodendrocyte-specific associations, suggesting that oligodendrocytes may modulate glutamate receptor trafficking and participate in excitotoxic processes traditionally attributed to neurons. This complements traditional neuron-centric excitotoxicity models and suggests glial-targeted therapeutic strategies. *TNIP1* and *KANSL1* further implicate mitochondrial dysfunction through mitophagy suppression: *TNIP1* shows excitatory neuron specific associations that may increase early AD vulnerability through mitophagy suppression, consistent with the mitochondrial pathway enrichment observed for *iAD2* and *iAging4*; while *KANSL1* additionally shows strong PD associations, indicating that mitochondrial dysfunction may represent a shared mechanism across neurodegenerative disorders. These findings reveal possible *MAPT*-independent therapeutic targets through synaptic transmission functions (**Figure 6a**), demonstrating how intermediate phenotypes may enable mechanistic gene prioritization in genomically complex regions. Additional mechanistic interpretation, including epigenetic regulation through chromatin accessibility and opposing bulk versus single-cell effects, is provided in **Supplementary Notes S3 and S5**.

Intracellular transport and intercellular communication through primary cilia in aging. The proposed primary cilia hub positions cilia as regulators integrating intracellular transport with intercellular communication in aging and neurodegeneration. Primary cilia dysfunction has also been linked to numerous neurological disorders, underscoring their sensitivity to aging and involvement in neurodegenerative processes[23]. Structurally, ciliary axoneme, built from nine microtubule doublets, supports bidirectional intraflagellar transport (IFT) and may contribute to axon transport relevant processes captured by *iAD1*. Functionally, primary cilium is both a broadcaster and a receiver in a complex cellular communication network mediated by EVs, which is implicated by *iAging3*. A total of 145 genes were implicated by DNE-xQTL integration, including intraflagellar transport genes (analogous to axonal transport captured by *iAD1*) and EV pathway genes (implicated in tau propagation captured by *iAging3*), suggest cilia coordinate both cellular homeostasis and intercellular protein propagation. The convergence of *CBY1* (essential for ciliogenesis) and *MAP4* (which increases during oligodendrocyte differentiation as cilia disassemble) situates this genetic architecture within a broader framework connecting ciliary dynamics to axonal transport and myelination biology in aging and AD. In addition, *MON1A* participates in autophagosome maturation and endolysosomal trafficking, whereas *USP19* contributes to *DNAJC5*-dependent EV export of misfolded proteins, illustrating how autophagy-related regulators could interface with the primary cilia hub. This combined statistical

and functional evidence generates the hypothesis that primary cilia dysfunction may represent a nexus linking cellular aging to pathological protein propagation in neurodegeneration, suggesting that cilia function not only as sensory organelles but also as integrators of proteostasis and intercellular signaling **(Figure 6c)**. Detailed interpretation of individual genes, their tissue-specific effects, and the autophagy-EV mechanistic links is detailed in **Supplementary Notes S3 and S4**.

Opposing directional effects in oligodendrocytes. The opposing directional effects observed for several genes (*NSF*, *SEZ6L2*, *DOC2A*, *MAPK3*), *i.e.*, elevated expression with pre-symptomatic traits (*iAD2*, *iAging2*) but reduced expression in symptomatic AD, may reflect stage-dependent regulatory mechanisms. All four genes show their strongest associations in oligodendrocytes, which consistently exhibit early upregulation consistent with their roles in axonal support and buffering excitatory stress. For example, elevated oligodendrocyte-specific expression of *SEZ6L2*, *DOC2A*, *MAPK3* in *iAging2* may represent compensatory responses supporting myelination and synaptic communication under early aging-related stress, in contrast to their reduced expression in symptomatic AD. Similarly, at the 17q21.31 locus, *NSF* showed positive associations with *iAD2* but negative associations with AD. Since *NSF* promotes AMPA receptor surface expression while tau protein inhibits its activity[24], elevated expression in early stages may increase receptor-mediated excitotoxic vulnerability before tau pathology suppresses its function. These bidirectional patterns raise the possibility that DNE traits reflect transitional or compensatory states distinct from the end-stage pathology observed in symptomatic disease.

This study has several limitations that highlight future directions. As an *in silico* gene prioritization framework leveraging multi-scale genetic data, functional validation through experimental models remains essential to test our mechanistic hypotheses. Our xQTL contexts lack certain disease-relevant cell types, particularly cerebral endothelial cells needed to strengthen vascular mechanisms, and are limited to post-mortem brain tissue that may not capture dynamic regulatory states antemortem. In addition, establishing causality across the complete xQTL-DNE-endpoint pathway remains challenging. While Mendelian randomization has been applied to pairs of traits (DNE-endpoint and xQTL-endpoint), extending causal inference across all three studies sequentially with limited instrumental variables poses methodological challenges for future statistical genetics methods development. Beyond these limitations, this framework opens several promising directions. First, incorporating multi-organ neuroimaging endophenotypes and xQTL data would enable systematic investigation of systemic aging mechanisms beyond the brain. Second, multi-ancestry xQTL studies would enable higher-resolution fine-mapping in complex loci where European-ancestry LD patterns limit variant resolution. Third, establishing longitudinal neuroimaging-molecular cohorts that track individuals from pre-symptomatic to symptomatic stages could directly test stage-dependent mechanisms. In the longer term, linking DNE-derived early brain signatures with molecular pathways may support earlier risk stratification, refine target selection, and inform clinical strategies aimed at intervening before irreversible neurodegeneration occurs.

In conclusion, this gene prioritization and hypothesis-generating framework demonstrates that biologically characterized neuroimaging endophenotypes, when integrated with comprehensive

multi-context molecular QTL data, can reveal candidate molecular targets for early intervention in neurodegeneration. By transforming abstract imaging patterns into testable cellular hypotheses with molecular and cell-type resolution, this approach enables biologically motivated gene prioritization in genomically complex regions where conventional GWAS approaches fail. The three regulatory hubs identified provide candidates for experimental validation and therapeutic development targeting pre-symptomatic brain changes before irreversible neurodegeneration occurs.

METHODS

Data description

Summary statistics for Dimensional Neuroimaging Endophenotypes (DNEs) and Alzheimer's disease (AD) from multiple cohorts were obtained from publicly available repositories as described in the original publications. Details of data sources, cohort composition, and sample sizes are provided in **Table S1**. Briefly, DNEs were defined using machine learning models trained on multimodal MRI features to capture disease-related neurobiological dimensions. These imaging-derived traits were developed through a standardized framework and validated across multiple brain disorders, with the trained models applied to UKBB participants (n = 32,829-33,541) to generate continuous DNE scores and GWAS performed adjusting for age, sex, their interactions with age and age-squared, total intracranial volume (ICV), and genetic principal components, yielding associations with common genetic variants. All genomic coordinates were harmonized to the GRCh38 (hg38) reference assembly using LiftOver[25].

This work is based on summary statistics from multiple molecular quantitative trait locus (xQTL) datasets, including bulk and single-nucleus expression, splicing, proteomic, glycoproteomic, metabolomic, and epigenomic (DNA methylation, histone modification, and single-nucleus Assay for Transposase-Accessible Chromatin) QTLs, were generated by the Functional Genomics Consortium xQTL Working Group (FunGen-xQTL) using data from the Religious Orders Study and Memory and Aging Project (ROSMAP), Mount Sinai Brain Bank (MSBB), and Knight Alzheimer Disease Research Center (ADRC). The combined multiple brain regions include anterior cingulate cortex (AC), dorsolateral prefrontal cortex (DLPFC), and posterior cingulate cortex (PCC) from ROSMAP; and frontal pole (FP), inferior frontal gyrus (IFG), superior temporal gyrus (STG), and parahippocampal gyrus (PHG) from MSBB; and primarily parietal cortex from Knight ADRC as summarized in **Table S2** with details available in the corresponding publications[26–32].

Fine-mapping

For DNE and AD GWAS summary statistics described in **Table S1**, we performed the following quality control procedures prior to fine-mapping analysis. Variants with missing alleles, strand mismatches, or duplicates were excluded, and allele orientation was aligned to the linkage disequilibrium (LD) reference panel. The LD reference was estimated from 17,000 unrelated whole-genome-sequenced European individuals from the Alzheimer's Disease Sequencing Project (ADSP), provided by the Genome Center for Alzheimer's Disease (GCAD)[33]. Strand-ambiguous SNPs (A/T or C/G), variants with missingness greater than 1%, minor allele frequency (MAF) below 0.0025, or not shared between GWAS and genotype-based LD reference panels were removed. Additional quality control was performed to detect and remove variants with inconsistent z-scores or abnormal LD patterns. For each LD block, z-scores were jointly modeled with the local LD matrix under the SuSiE[34] framework to identify discrepancies between the statistical signal (z) and the corresponding LD structure (R). Variants exhibiting outlier behavior were excluded from both the summary statistics and LD matrix to ensure internal consistency. Missing or filtered variants were then imputed using the RAISS[35] algorithm, which reconstructs missing z-scores by leveraging LD information and nearby correlated variants (rcond = 0.01, $R^2 > 0.6$, and a minimum LD threshold of 5). Fine-mapping was performed using SuSiE-RSS[9] model on quality-controlled summary statistics, with the number of causal components (L) optimized iteratively (initial = 5, maximum = 10, increment = 5) within each LD block as described in previously published work[36]. Regions without strong signals (PIP < 0.025) are skipped automatically. Posterior inclusion probabilities (PIP) were used to define 95%, 70%, and 50% credible sets, providing fine-mapping results at multiple confidence thresholds.

Fine-mapping diagnostics were then conducted to evaluate the reliability and interpretability of credible sets, providing an automated quality assessment of fine-mapping results. For loci without identified credible sets, the variant with the highest PIP was retained. Correlation matrices were computed to quantify interdependence among credible sets by calculating the maximum and minimum absolute correlations. Each locus was subsequently assigned an optimal fine-mapping update strategy based on the number, significance, and correlation of credible sets, selecting among Bayesian Variable Selection Regression (BVSR), Single Effect Regression (SER), and Bayesian Conditional Regression (BCR) models. Following automated evaluation, results were curated through manual exploratory analysis. Posterior credible set metrics were aggregated at the study-block level, where correlation matrices between credible sets were used to detect redundant or overlapping signals (e.g., $|r| > 0.8$) and to assess variant sharing across studies. Genomic blocks were classified into single- or multi-credible set categories, and models with unusually large credible sets were flagged as potentially unstable. Manual inspection included visualization of PIP distributions, cross-credible set correlation structures, and concordance with marginal P-values, allowing exclusion of outlier or low-confidence regions.

Colocalization analysis

Multi-context colocalization analyses were performed using two complementary strategies with *ColocBoost*[11]:

(1) Multi-GWAS colocalization for gene prioritization. This strategy integrated DNE traits and AD GWAS traits (**Table S1**) for joint modeling of association signals. Both DNE and AD GWAS analyses used the same quality-controlled summary statistics and the same LD reference panel from the ADSP/GCAD consortium, consistent with the fine-mapping analyses.

(2) Molecular QTL colocalization across multi-omic contexts for prioritized gene validation. Multi-GWAS colocalization was also performed across diverse molecular phenotypes, including bulk and single-nucleus expression, splicing, and protein QTLs from the ROSMAP cohort generated by the FunGen-xQTL Consortium (**Table S2**).
All analyses were performed using the ColocBoost framework. Regions with weak signals (PIP < 0.025) were excluded for splicing QTLs, with all other parameters set to default.

TWAS analysis

Transcriptome-Wide Association Studies (TWAS) analyses were performed for xQTL molecular traits to assess their causal associations with DNE, aging, and AD GWAS traits (**Table S2**).

For each molecular trait, the genome was partitioned into LD blocks[36], consistent with the fine-mapping analysis, and within each block, prediction models were trained for all genes whose transcription start sites (TSS) were located within the corresponding region. Six predictive models, including Elastic Net, LASSO, Bayesian Ridge, Bayesian LASSO, Sum of Single Effect (SuSiE), and Mr.ASH[34,37–41], were trained to estimate cis-xQTL weights for candidate genes following the optimized FunGen-xQTL protocol. Elastic Net combines both L1 and L2 regularization penalties, whereas LASSO applies only the L1 norm. Bayesian Ridge and Bayesian LASSO extend linear regression by incorporating Gaussian and double-exponential priors, respectively, to model uncertainty in effect-size estimates. SuSiE captures the joint effects of multiple causal variants within a probabilistic framework, while Mr.ASH uses an empirical Bayes approach to estimate effect-size distributions across traits. Model performance was evaluated using five-fold cross-validation, and models with mean $R^2 > 0.01$ were retained for downstream analyses.

After model training, the corresponding GWAS summary statistics were harmonized with the same ADSP LD reference data. Associations between genetically predicted molecular trait levels and GWAS z-scores were then tested using burden-based statistics to derive TWAS z-scores and P-values. A TWAS signal was considered significant if the best-performing model (1) had a TWAS p-value below (2.5×10^{-6}) in the global landscape analysis, or (2) remained significant after Bonferroni correction (Bonferroni corrected $p < 0.05$).

Mendelian Randomization

Fine-mapped xQTL variants with a cumulative posterior inclusion probability (cPIP) > 0.5 were used as instruments. For each molecular trait, xQTL and GWAS estimates were aligned after allele matching and standard quality control. Exposure effects were scaled by their standard errors, and credible-set level causal effects were obtained by weighting ratio estimates by their posterior inclusion probabilities. Inverse-variance weighted credible-set estimates were integrated across credible sets using fixed-effect meta-analysis to derive gene-level MR effects. Cochran's Q and the corresponding I^2 statistic were used to assess heterogeneity. Gene-level MR associations were considered significant when at least one credible set was present, cPIP > 0.9, the meta-analytic p-value < 0.05, and $I^2 < 0.5$.

Definition of xQTL-associated genes for DNE traits

Genes were prioritized from DNE-associated variants through integration with xQTL data using significant colocalization and TWAS associations, following the strategy outlined below.

(1) Colocalization: Variants supported by the multi-GWAS colocalization analysis were intersected with fine-mapped xQTL variants from the FunGen-xQTL Consortium across bulk and single-nucleus expression, splicing, proteomic, and glycoproteomic datasets. For each xQTL trait, we included variants from credible sets at multiple coverage levels (95%, 70%, and 50%), along with additional variants outside these sets that had posterior inclusion probability of at least 0.25, to broaden the set of candidate variants. Genes linked to overlapping variants were designated as prioritized targets.

(2) Putative causal gene (PCG) analysis: Merged TWAS results across multiple brain cohorts were loaded and filtered for robust models (cross-validation $R^2 > 0.01$, $P_{cv} < 0.05$) and significant associations ($P_{TWAS} < 2.5 \times 10^{-6}$). For each LD block and molecular context, the gene containing the variant with the highest posterior inclusion probability (PIP) within the 95%, 70%, or 50% credible set was selected as the most likely causal candidate for evaluation using DNE GWAS summary statistics. GWAS summary statistics were integrated at the chromosome level, and missing p-values were recalculated from z-scores when necessary. The variant with the smallest GWAS p-value within each region was then identified as the putative causal variant, and its corresponding gene was extracted as the final prioritized target.

Genes were prioritized into three tiers based on integrated evidence from TWAS and colocalization analyses across DNE, aging, and AD traits. Tier 1 included the most confident candidates supported by both TWAS and colocalization signals in all three trait categories (AD, Aging, and DNE). Tier 2 comprised genes showing both TWAS and colocalization evidence in DNE traits but only TWAS support in AD and Aging. Tier 3 included genes with either TWAS or colocalization evidence in DNE traits and only TWAS support in AD and Aging.

Gene-set enrichment analysis

Genes from the union of colocalization-prioritized and causal TWAS-prioritized results were identified as significant xQTL gene targets for each DNE trait and subjected to Gene Ontology

(GO)[42] and Kyoto Encyclopedia of Genes and Genomes (KEGG)[43] enrichment analyses using the R packages *clusterprofiler*[44] to characterize their functional relevance for each DNE trait. Gene lists were compared across DNE traits using either overlapping or union sets when applicable. GO analyses were performed across all three ontology categories: biological process (BP), cellular component (CC), and molecular function (MF). All annotated human genes in the default *clusterProfiler* background (*org.Hs.eg.db*) were used as the reference set for enrichment. Pathways with Benjamini-Hochberg adjusted $p < 0.1$ were considered statistically significant. Results from GO and KEGG enrichment were combined and jointly visualized to highlight functional enrichments across different DNE traits.
Genes reaching significance in both TWAS ($p < 2.5 \times 10^{-6}$) and Mendelian randomization were assessed by over-representation analysis using msigdbr[45] (v25.1.1) to access curated human gene sets from MSigDB. Enrichment analyses were conducted to delineate biological processes pertinent to Alzheimer's disease or telomere-length biology. Pathways with $p < 0.05$ were considered significant.

Heritability enrichment analysis

We quantified the heritability enrichment of functional annotations using the stratified linkage disequilibrium score regression (sLDSC)[46] framework implemented in the *polyfun* package. The sLDSC model estimates the proportion of SNP-based heritability attributable to specific functional categories by regressing GWAS $\square^2$ statistics on LD scores partitioned by annotation, thereby assessing whether variants within annotated regions explain more heritability than expected under a null model of uniform heritability per SNP.

The variants with maximum variant colocalization probability (maxVCP) were extracted from all Colocalized Confidence Sets (CoS) and uncolocalized CoS (UCoS) in multi-GWAS colocalization results for each DNE trait and compiled into annotation files indicating SNP membership within DNE trait-specific regions. AD and aging GWAS summary statistics were used to estimate SNP-based heritability and evaluate functional enrichment across all annotations, after filtering variants with MAF ≤ 0.001 and imputation quality ≤ 0.6, and excluding strand-ambiguous and human leukocyte antigen (HLA) region variants. Baseline LD scores and regression weights derived from the 1000 Genomes Project European[48] reference panel were included to account for genome-wide linkage disequilibrium structure. The model estimated the marginal annotation coefficient (tau), representing the per-SNP contribution of each annotation to heritability, and the enrichment statistic, defined as the proportion of heritability explained by SNPs within a given annotation relative to their genomic coverage.

Genetic correlation analysis

Pairwise genetic correlations between imaging-derived neurodegenerative dimensions (DNEs) were estimated using bivariate LD Score Regression (LDSC)[49,50]. We used the same genome-wide association summary statistics generated in the sLDSC processing, restricted to autosomal HapMap3 SNPs to satisfy LDSC model assumptions. Standard QC filters were

applied to remove variants with missing allele information, duplicated positions, or extreme test statistics.

Bivariate LDSC was performed using LD scores and regression weights from the 1000 Genomes Project Phase 3 European (HapMap3) reference panel. For each pair of DNE traits, LDSC regressed the product of Z-scores from the two GWAS onto the HapMap3 LD scores to estimate genetic covariance and genetic correlation (rg). Pairwise rg values and standard errors were visualized in a genetic correlation matrix to summarize the polygenic similarity structure among DNE, Aging and AD traits, the Bellenguez et al. dataset was treated as a proxy AD cohort and the Kunkle et al. dataset as a clinical AD cohort.

GTEx tissue expression analysis

Gene expression across tissues was quantified using GTEx v10 RNA-seq data downloaded from GTEx Portal (https://gtexportal.org). For each gene of interest, TPM values were extracted and merged with the corresponding tissue assignments. Tissue-level expression distributions were summarized across samples and visualized using boxplots on a $\log_{10}$ TPM scale, ordered by the mean expression across tissues.

Pathology association analysis

Genome-wide association analyses were performed for nine harmonized neuropathological traits in ROSMAP (TDP-43 pathology, Tangles burden, Cerebral atherosclerosis, Neurofibrillary tangle burden, Cerebral amyloid angiopathy, Cognitive decline slope, Neuritic plaque burden, Amyloid burden, Arteriolosclerosis). For each trait, single-variant association testing was conducted using PLINK[51] v1.9 under an additive linear model, adjusting for age at death, sex, and the first four genetic principal components.

Data and Code Availability

All data supporting the findings of this study are provided in the main text and Supplementary Information. Summary statistics generated by the FunGen-xQTL project will be made available upon publication of the accompanying consortium manuscript. The complete analysis codebase, including workflows for fine-mapping, colocalization, TWAS, and functional enrichment, is available at https://github.com/rl3328/DNE_xQTL_paper.

ACKNOWLEDGEMENTS

We thank members of the Wang, Wen, and Davatzikos labs for discussions and support. We thank the members of the Alzheimer's Disease Sequencing Project Functional Genomics Consortium (FunGen-AD) for providing the FunGen-xQTL resource. This work was supported

by NIH/NIA R01AG076901 and R01AG086467(G.W., R.L.), a FunGen-AD Consortium Grant U01AF072572 (G.W., R.F., P.L.D.), a grant from The Urbut Family Foundation (G.W.).

AUTHOR CONTRIBUTIONS

R.L., R.F. contributed equally. G.W., J.W., C.D. jointly supervised research. R.L., R.F., J.W., G.W. conceived and designed the experiments, performed the analyses, and interpreted the results. X.C., A.L. contributed statistical methodology. P.L.D., D.B., G.W., R.F. contributed to and supervised other molecular QTL data production. C.D. and J.W. contributed to and supervised the DNE data production. R.L., R.F., G.W. wrote and revised the manuscript.

COMPLETING INTERESTS

The authors declare no competing interests.

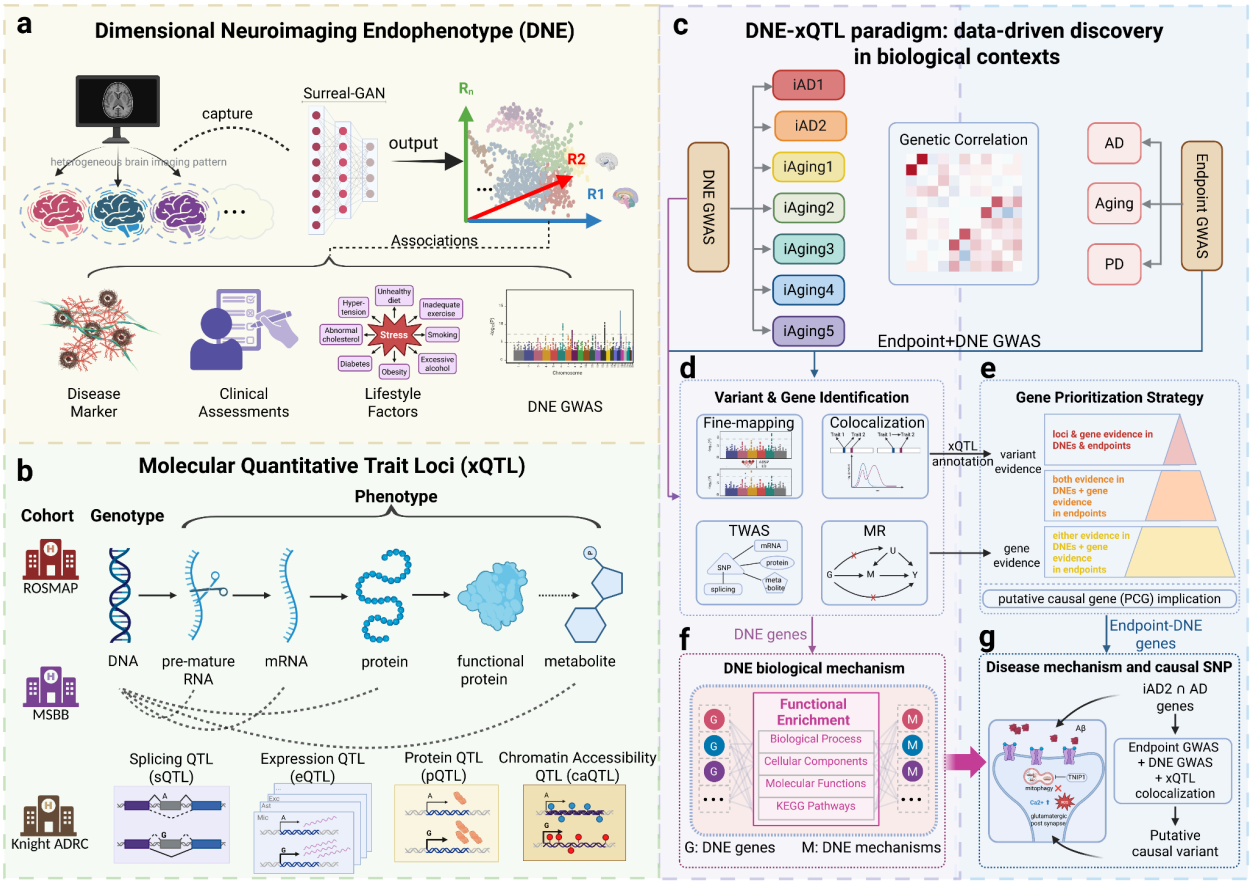

Figure 1. The DNE–xQTL analytical framework bridges genetics, neuroimaging, and molecular mechanisms to reveal disease biology. (a) DNE generation from brain MRI across up to 43,100 individuals using deep learning trained on AD and aging outcomes. **(b)** Comprehensive brain molecular QTL data across multiple modalities (eQTL, sQTL, pQTL, etc.) and diverse cell types from the FunGen-AD consortium. **(c)** Genetic correlations across 16 GWAS traits to assess polygenic similarity and heritability enrichments to pinpoint DNE contributions to endpoints. **(d)** Variant and gene identification: fine-mapping and multi-GWAS colocalization identify putative causal variants; brain xQTL mapping across molecular contexts links variants to regulated genes; TWAS and Mendelian Randomization (MR) identify putative causal genes. (e) Gene prioritization strategy: genes supported by both variant-level and gene-level evidence are organized into tiers, with convergent evidence ranked highest; putative causal gene (PCG) implication addresses loci containing multiple candidate genes. **(f)** Functional annotation via GO/KEGG enrichment assigns biological context for each DNE. **(g)** Integration of DNE biological annotations with colocalization across endpoints, DNEs, and xQTLs identifies putative causal variants linking pre-symptomatic brain changes to disease. This framework enables generating testable hypothesis linking abstract imaging patterns to molecular targets with cellular resolution.

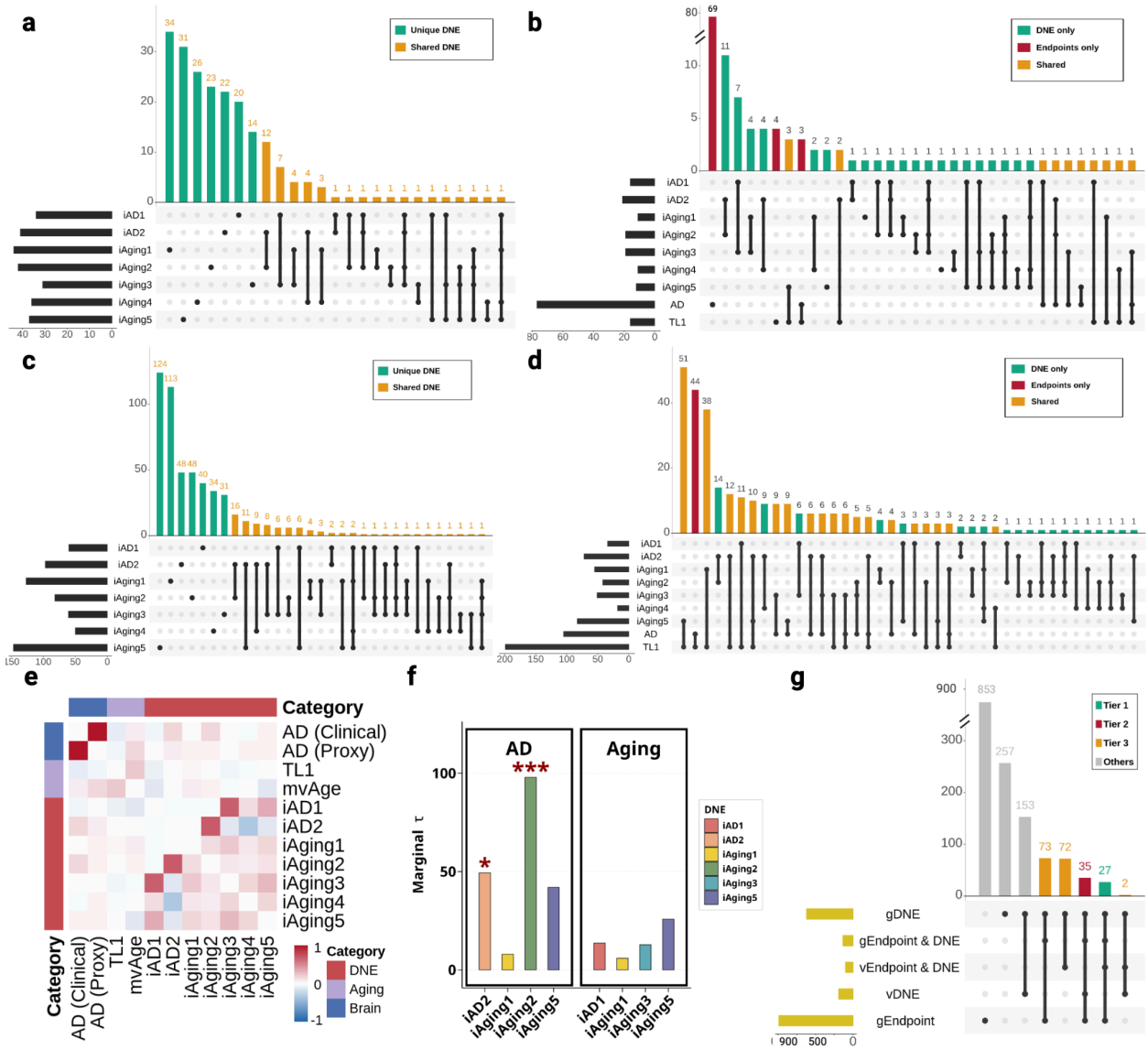

Figure 2. Putative causal variants and genes across DNE, Alzheimer's disease (AD) and aging. (a) Multi-GWAS colocalization identified 95% colocalization sets (CoS) across seven DNEs, showing loci-level patterns of sharing. **(b)** DNE-endpoint colocalization further identifies shared genetic architecture with AD and Aging (telomere length). **(c)** UpSet plot of genes identified by transcriptome-wide association studies (TWAS) integrating brain xQTL across seven DNEs, showing gene-level patterns of sharing. **(d)** DNE-endpoint gene overlaps from TWAS reveal extensive shared genetic architecture: 238 genes with telomere length, genes overlap between DNEs and AD, including 14 genes showing three-way convergence across AD, telomere length, and DNEs. The use of DNE enables fine-grained dissection at both variant-level (A-B) and gene-level (C-D) resolutions. **(e)** Bivariate LDSC across all GWAS used, assessing polygenetic similarity. **(f)** Stratified LD score regression using both maxVCP from colocalization and maxPIP from fine-mapping quantifies heritability enrichment for each DNE in AD and aging, showing positive enrichments with asterisks indicating levels of statistical significance (*iAging2* in AD: τ=0.54, P=3.4×10^{-5}; *iAging5* in AD: τ=0.34, P=6.5×10^{-3}), establishing DNE relevance as intermediate phenotypes. **(g)** UpSet plot of gene-level evidence across DNE-endpoint pairs with a 3-tier evidence framework organizes candidate genes: Tier 1 (causal variant and gene evidence in endpoints and DNEs); Tier 2 (both evidence types in DNEs, and gene evidence in endpoints); Tier 3 (either evidence in DNEs, and gene evidence in endpoints).

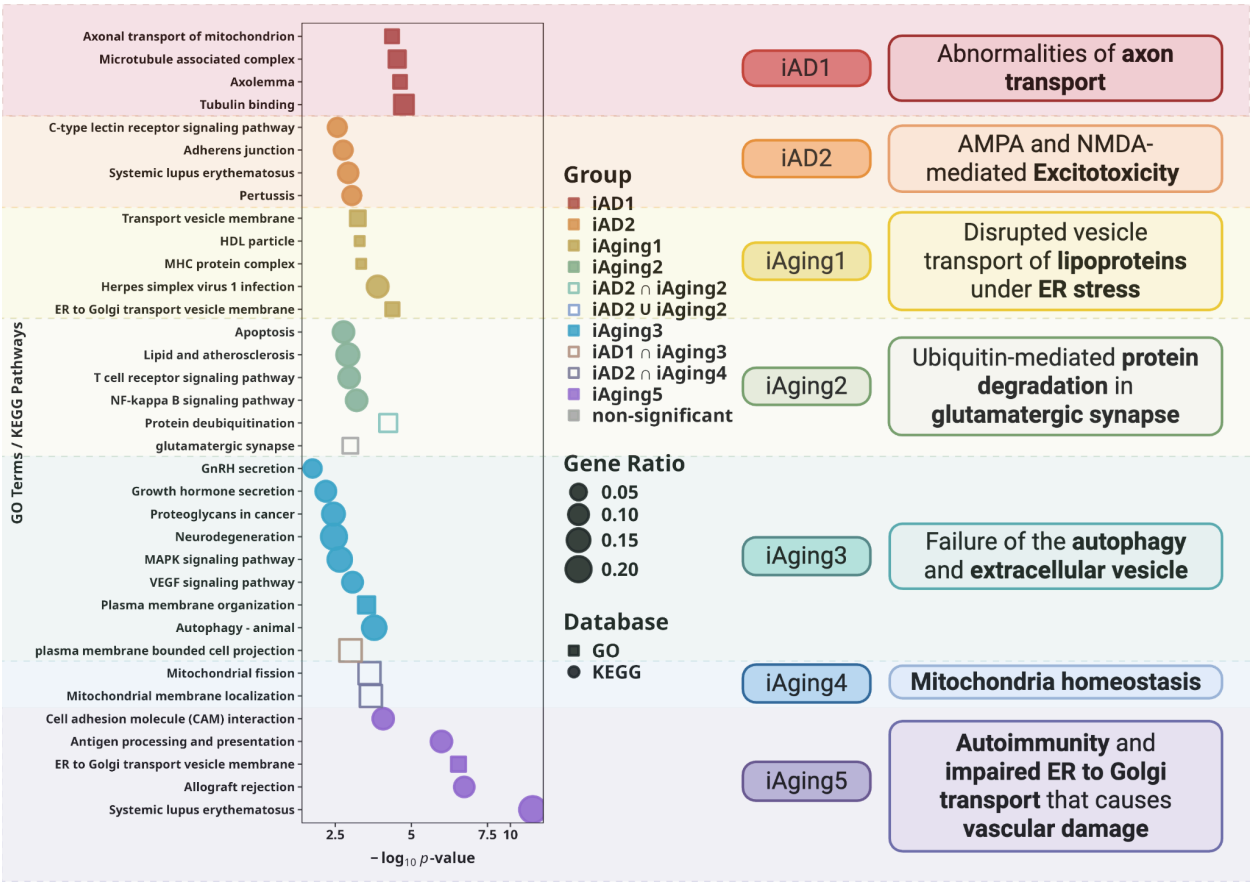

Figure 3. xQTL gene targets dissect distinct cellular processes underlying each DNE. GO and KEGG enrichment of genes from integration of xQTL colocalization and TWAS define biological signatures. Dot size indicates gene ratio; color indicates different DNE with statistical significance (FDR<0.1). Each DNE captures specific cellular responses: *iAD1* represents cytoskeletal and axonal transport dysfunction while *iAD2* reflects glutamate receptor–mediated excitotoxicity; *iAging1* reflects disrupted vesicle transport of lipoproteins under ER stress; *iAging2* indicates dysregulated ubiquitin-mediated protein degradation in excitatory neurons with convergence on glutamatergic synapses; *iAging3* highlights impairments in MAPK signaling, glial differentiation, and autophagy, with extracellular vesicles as potential mediators; *iAging4* points to mitochondrial metabolic adaptation and altered dynamics; and *iAging5* reveals ER-Golgi transport failure and auto-inflammatory states promoting potential vascular damage. Two AD-related DNEs jointly implicate additional pathways involving the ciliary basal body.Collectively, these results show that each DNE represents a distinct but interconnected cellular response underlying aging and AD.

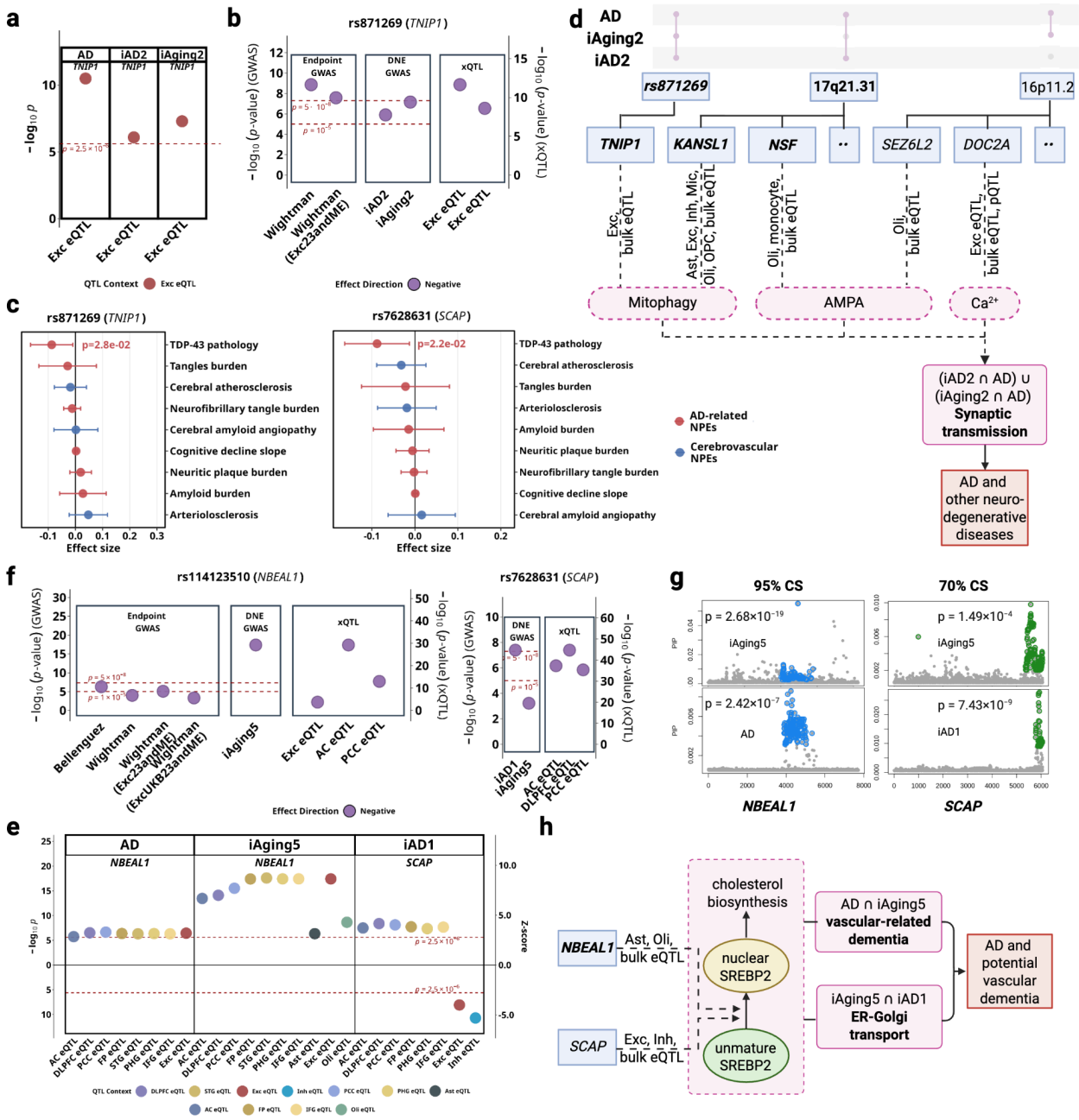

Figure 4. ***iAD1* and *iAD2* implicate excitotoxicity and cholesterol dysregulation in AD pathogenesis.** a-d shows the synaptic transmission regulatory hub; e-h shows the cholesterol/vascular regulatory hub. **(a)** Lead *TNIP1* variant (rs871269) colocalization across datasets with AD. **(b)** TWAS *Z*-scores for *TNIP1* across significant tissues/cell types denoted with different colors. **(c)** Pathology associations for lead variants for *TNIP1* and *SCAP*. **(d)** Possible mechanistic model involving key biological processes in synaptic transmission: mitophagy, glutamate receptor signaling, and calcium imbalance. **(e)** TWAS *Z*-scores for *NBEAL1* and *SCAP* across significant tissues/cell types. (**f**) Lead *NBEAL1* variant (rs114123510) colocalization across AD and *SCAP* variant (rs7268631) across *iAD1*, both involving *iAging5*, with **(g)** fine-mapping 95% and 70% credible sets shown for these variants, respectively. **(h)** Possible mechanistic model involving modulation of SREBP maturation to influence AD and other vascular-related dementia.

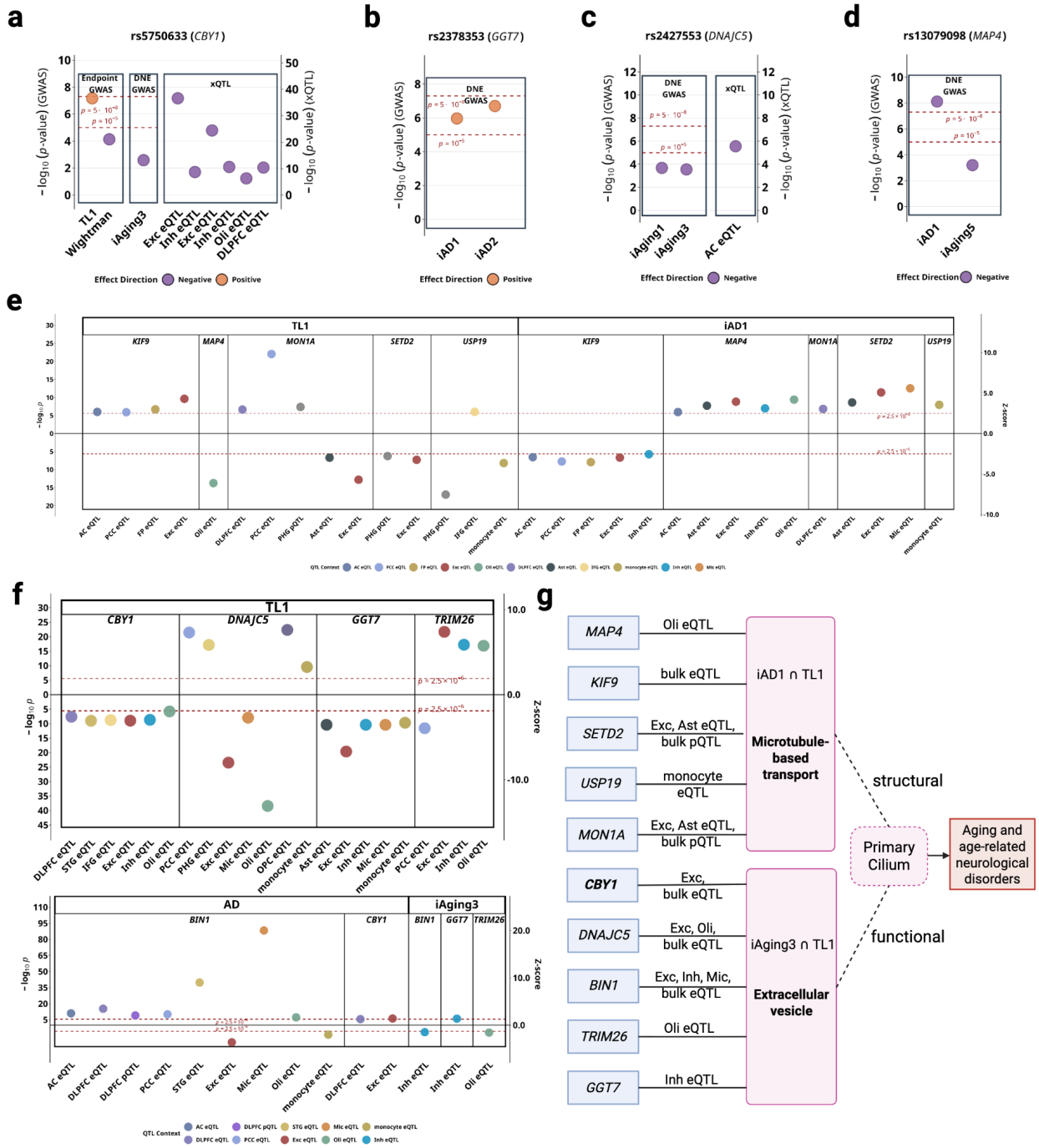

Figure 5. ***iAD1* and *iAging3* implicate primary cilia dysfunction in aging and AD pathology. (a-d)** Lead variants colocalization across endpoint and DNE GWAS, and brain xQTL datasets for *CBY1*, *GGT7*, *DNAJC5* and *MAP4*. **(e)** TWAS *Z*-scores for *iAD1*-telomere length genes (*MAP4*, *KIF9*, *SETD2*, *MON1A*, *USP19*) across significant tissues/cell types denoted by different colors. **(f)** TWAS Z-scores for *iAging3*-telomere length genes (*CBY1*, *DNAJC5*, *BIN1*, *TRIM26*, *GGT7*) across significant tissues/cell types. **(g)** Possible mechanistic model involving 10 telomere length and/or AD-associated genes linked to *iAD1* or *iAging3*, all structurally or functionally related to primary cilia.

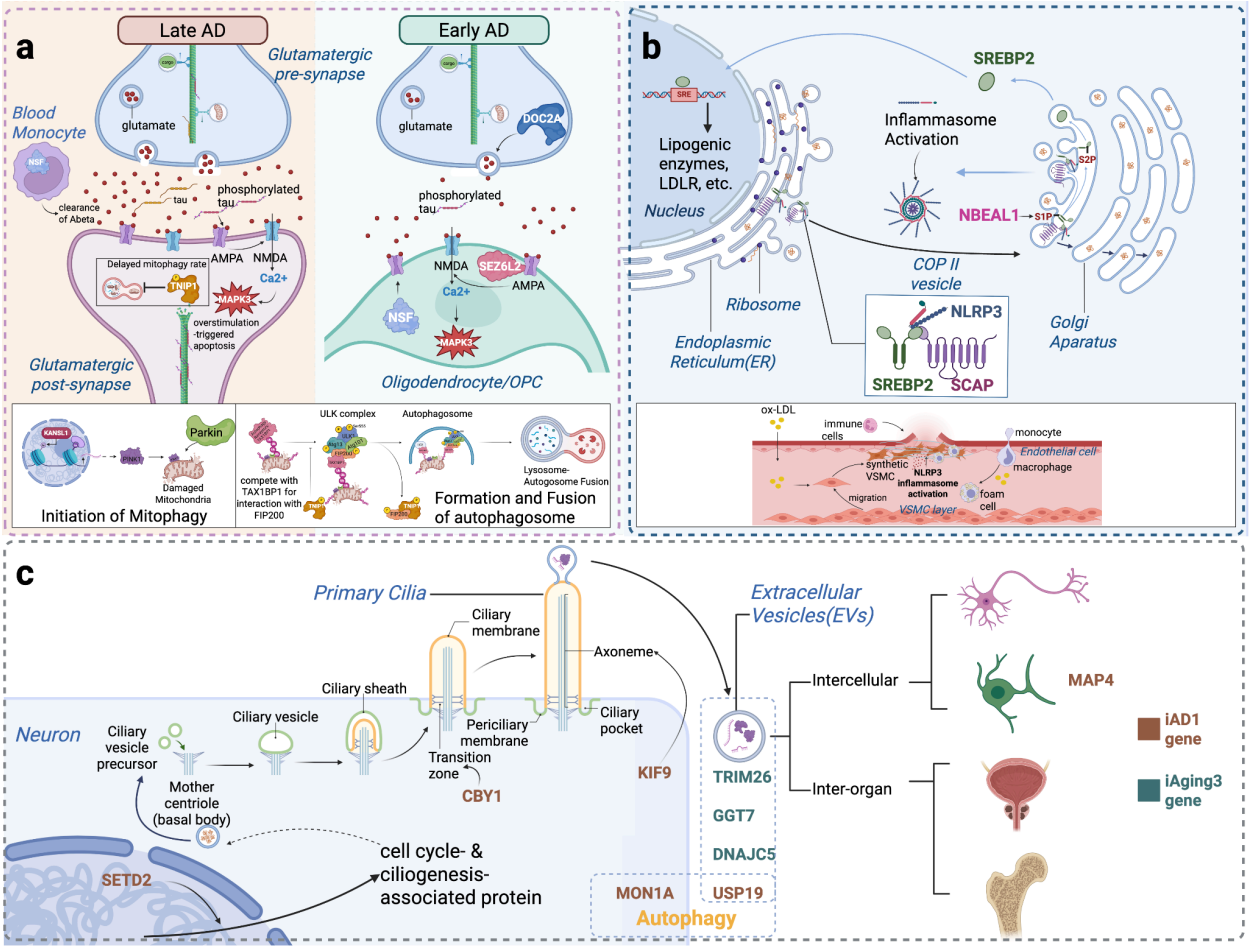

Figure 6. Summary of hypotheses for DNE-mediated pathways in AD and aging.

(a) Oligodendrocytes in excitotoxicity: Hippocampal synaptic transmission declines in aging through reduced ubiquitin-mediated protein regulation (*iAging2*). *DOC2A* promotes synaptic transmission by modulating calcium level, showing a positive association with *iAging2* and a negative association with AD in DLPFC but the effects are reversed in the AC. In AD, glutamate excitotoxicity via AMPA/NMDA receptors (*iAD2*) drives synaptic dysfunction. *KANSL1* modulates AMPA receptor breakdown, while *NSF* regulates AMPA receptor trafficking, and *SEZ6L2* serves as an AMPA receptor scaffolding protein—all showing associations with *iAD2* and AD. Mitochondrial dysfunction exacerbates both aging and AD pathways. *TNIP1* suppresses mitophagy, showing positive associations with *iAD2*, *iAging2*, and AD. Critically, *NSF* and *SEZ6L2* expression in oligodendrocytes reveals essential glial contributions to excitotoxicity regulation, expanding beyond neuron-centric models.

(b) Cholesterol dysregulation in vascular dementia: Under physiological conditions, *SCAP* escorts SREBP2 from ER to Golgi, where S1P and S2P proteases cleave SREBP2 to release its active fragment, which translocates to the nucleus to induce LDL receptor expression and maintain cholesterol homeostasis. *SCAP* also transports NLRP3 for controlled inflammasome activation. In pathological states, *NBEAL1* may enhance SREBP2 proteolysis, potentially driving excessive SCAP-SREBP2 activity that causes aberrant LDL receptor upregulation and cholesterol accumulation. This could trigger NLRP3-dependent inflammation, resulting in hypercholesterolemia and chronic neuroinflammation that may compromise vascular integrity of small cerebral vessels, ultimately promoting cognitive decline and vascular dementia.

(c) Primary cilia as aging regulators: Ten telomere length-associated genes converge on ciliary function through two mechanisms. *MAP4*, *KIF9*, *SETD2*, *MON1A*, and *USP19* participate in intraflagellar transport, which is analogous to axonal transport, and may regulate cilia-autophagy interactions, with *MAP4* exclusively expressed in oligodendrocytes, suggesting potential roles in cilia dynamics during oligodendrocyte maturation. *CBY1*, *DNAJC5*, *BIN1*, *TRIM26*, and *GGT7* mediate extracellular vesicle formation and secretion for intercellular communication. *CBY1* is essential for ciliogenesis; *DNAJC5* and *USP19* regulate EV-mediated misfolded protein release; *TRIM26* functions in oligodendrocyte EVs, while *BIN1* operates in inhibitory neuron EVs; *GGT7* serves as an EV cargo protein. Collectively, these genes suggest primary cilia may function as central regulators integrating intracellular transport with intercellular communication in aging and neurodegeneration.

Table 1. Candidate genes in regulating brain aging with tiers of evidence

Genomic Coordinate	Trait	Selected Genes	Group	Known Function	Novel Insights
Tier 1: Genes with both loci-level and gene-level evidence in endpoint and DNE					
chr2:201677542-203951659	AD-Aging5	*NBEAL1*	Vascular dysfunction	Shared genetic loci for AD, stroke and cerebral vascular disease[2,19,20]	Downstream cholesterol and inflammatory responses
chr5:149895776-152257891	AD-AD2-Aging2	*TNIP1*	Excitatory-neuron vulnerability	TNIP1 emerged as a risk gene for Alzheimer's disease and medial temporal atrophy[2,52]	Biological relevance to both excitotoxicity (through mitophagy suppression) and ubiquitination-dependent proteostasis
chr17:45383525-50162864	PD/AD-Aging3	*KANSL1*	Excitatory-neuron vulnerability	Established role in tauopathy pathogenesis of MAPT[53]	Mitophagy or autophagy effect critical for synaptic transmission
chr17:45383525-50162864	PD/AD-Aging3	*NSF*	Excitatory-neuron vulnerability	Regulate AMPA receptor trafficking[24]	Highlight its role in mechanisms, especially in oligodendrocytes
chr22:37320379-39179317	TL1/AD-Aging3	*CBY1*	Primary-cilia-associated transport	Essential for primary cilia biogenesis, related with autophagy, VEGF signalings and extracellular vesicles[54]	Novel genetic association between ciliogenesis and aging-related neurodegeneration
chr6:26341332-28390030	TL1-AD2	*H2AC11, H4C13*	Excitatory-neuron vulnerability	Core histone proteins essential for DNA packaging and gene regulation in the cell nucleus[55]	Potential regulation via histone modification during aging-related processes
Tier 2: Genes with both loci-level and gene-level evidence in endophenotypes along with gene-level evidence in endpoint					
chr16:29685831-46381513	AD-Aging2	*DOC2A*	Excitatory-neuron vulnerability	Calcium sensor that modulates pre-synaptic asynchronous neurotransmitter release by interacting with SNARE complexes[56]	Reported AD loci with more functional mechanism
chr16:29685831-46381513	AD-Aging2/Aging1	*SEZ6L2*	Excitatory-neuron vulnerability	Auxiliary subunit modulating AMPA receptor signaling and synaptic scaffolding[57]	Novel loci and novel mechanism linking glutamate receptor dysfunction to synaptic transmission alterations in oligodendrocytes in AD
chr20:32060775-36429279	TL1-Aging1/Aging3	*GGT7*	Primary-cilia-associated transport	Gamma-glutamyltransferase enzyme enriched in extracellular vesicles, cancer biomarker[58]	Validation of extracellular vesicle-mediated intercellular communication in brain aging processes

chr3:47545024-52179438	TL1(TWAS)-AD1 (CB)	*MAP4*	Primary-cilia -associated transport	Microtubule-associated protein negatively regulating primary cilium length[59]	Potential impact on oligodendrocyte development along with primary cilia length modulation
Tier 3: Genes with gene-level or loci-level evidence in endophenotypes along with gene-level evidence in endpoint that help support the hypothesis					
chr2:125689597-127728648	AD(MR)-Aging3(TWAS)	*BIN1*	Primary-cilia -associated transport	Functional AD loci, microglia EVs - tau propagation[60]	Validates biological significance of extracellular vesicle signaling in AD
chr20:63558301-64331539	TL1(MR)-Aging1 /Aging3(CB)	*DNAJC 5*	Primary-cilia -associated transport	Encodes CSP, regulates synaptic exocytosis and extracellular vesicle release in neurodegeneration[61]	Validates biological significance of extracellular vesicle signaling in aging-related brain changes
chr6:30136357-31604117	TL1(MR)-AD2/Aging3/Aging5(TWAS)	*TRIM26*	Primary-cilia -associated transport	E3 ubiquitin ligase[62]	Novel candidate with functional similarity to DNAJC5 in EVs regulation
chr3:45108216-47545024	TL1-AD1(TWAS)	*KIF9*	Primary-cilia -associated transport	Regulates primary cilium length[63]	Supports association between cilia structure and aging processes
chr3:45108216-47545024	TL1-AD1(TWAS)	*SETD2*	Primary-cilia -associated transport	Histone methyltransferase regulating cilia function and autophagy[64]	Validates crosstalk between cilia biology and autophagy in aging
chr3:47545024-52179438	TL1-AD1(MR)	*MON1A*	Primary-cilia -associated transport	Forms regulatory complex with Ccz1 as positive autophagy regulator[65]	Supports connection between autophagy regulation and aging processes
chr3:45108216-47545024	TL1-AD1(TWAS)	*SCAP*	Vascular dysfunction	SREBP chaperone controlling cholesterol homeostasis[22]	Orchestrates inflammasome activation through abnormal the post-translational modification of SREBP2 during aging
chr3:47545024-52179438	TL1-AD1(TWAS)	*USP19*	Primary-cilia -associated transport	Collaborates with CSP in misfolded protein secretion (MAPS) pathway[61]	Mechanistic link between AD1 and Aging3 through autophagy-mediated protein clearance

REFERENCE

1. Beach TG, Monsell SE, Phillips LE, Kukull W. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. *J Neuropathol Exp Neurol*. 2012;71(4):266-273. doi:10.1097/NEN.0b013e31824b211b

2. Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet*. 2022;54(4):412-436. doi:10.1038/s41588-022-01024-z

3. Wightman DP, Jansen IE, Savage JE, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat Genet*. 2021;53(9):1276-1282. doi:10.1038/s41588-021-00921-z

4. Wu Y, Sun Z, Zheng Q, et al. Pervasive biases in proxy GWAS based on parental history of Alzheimer's disease. *bioRxiv*. Published online October 17, 2023:2023.10.13.562272. doi:10.1101/2023.10.13.562272

5. Shade LMP, Katsumata Y, Abner EL, et al. GWAS of multiple neuropathology endophenotypes identifies new risk loci and provides insights into the genetic risk of dementia. *Nat Genet*. 2024;56(11):2407-2421. doi:10.1038/s41588-024-01939-9

6. Yang Z, Wen J, Erus G, et al. Brain aging patterns in a large and diverse cohort of 49,482 individuals. *Nat Med*. 2024;30(10):3015-3026. doi:10.1038/s41591-024-03144-x

7. Wen J, Yang Z, Nasrallah IM, et al. Genetic and clinical correlates of two neuroanatomical AI dimensions in the Alzheimer's disease continuum. *Transl Psychiatry*. 2024;14(1):420. doi:10.1038/s41398-024-03121-5

8. Wen J, Yang Z, Nasrallah IM, et al. Genetic and clinical correlates of two neuroanatomical AI dimensions in the Alzheimer's disease continuum. *Transl Psychiatry*. 2024;14(1):420. doi:10.1038/s41398-024-03121-5

9. Zou Y, Carbonetto P, Wang G, Stephens M. Fine-mapping from summary data with the "Sum of Single Effects" model. *PLOS Genet*. 2022;18(7):e1010299. doi:10.1371/journal.pgen.1010299

10. Leung YY, Lee WP, Kuzma AB, et al. Alzheimer's Disease Sequencing Project Release 4 Whole Genome Sequencing Dataset. *medRxiv*. Preprint posted online December 6, 2024:2024.12.03.24317000. doi:10.1101/2024.12.03.24317000

11. Integrative multi-omics QTL colocalization maps regulatory architecture in aging human brain | medRxiv. Accessed September 18, 2025. https://www.medrxiv.org/content/10.1101/2025.04.17.25326042v1

12. Okamoto J, Wang L, Yin X, et al. Probabilistic integration of transcriptome-wide association studies and colocalization analysis identifies key molecular pathways of complex traits. *Am J Hum Genet*. 2023;110(1):44-57.

13. Boettger LM, Handsaker RE, Zody MC, McCarroll SA. Structural haplotypes and recent evolution of the human 17q21. 31 region. *Nat Genet*. 2012;44(8):881-885.

14. Kim JJ, Vitale D, Otani DV, et al. Multi-ancestry genome-wide association meta-analysis of Parkinson's disease. *Nat Genet*. 2024;56(1):27-36.

15. Wainberg M, Andrews SJ, Tripathy SJ. Shared genetic risk loci between Alzheimer's disease and related dementias, Parkinson's disease, and amyotrophic lateral sclerosis. *Alzheimers Res Ther*. 2023;15(1):113.

16. Attems J, Jellinger KA. The overlap between vascular disease and Alzheimer's disease - lessons from pathology. *BMC Med*. 2014;12:206. doi:10.1186/s12916-014-0206-2

17. Bindesbøll C, Aas A, Ogmundsdottir MH, et al. NBEAL1 controls SREBP2 processing and cholesterol metabolism and is a susceptibility locus for coronary artery disease. *Sci Rep*. 2020;10(1):4528. doi:10.1038/s41598-020-61352-0

18. Guo C, Chi Z, Jiang D, et al. Cholesterol Homeostatic Regulator SCAP-SREBP2 Integrates NLRP3 Inflammasome Activation and Cholesterol Biosynthetic Signaling in Macrophages. *Immunity*. 2018;49(5):842-856.e7. doi:10.1016/j.immuni.2018.08.021

19. Chung J, Marini S, Pera J, et al. Genome-wide association study of cerebral small vessel disease reveals established and novel loci. *Brain*. 2019;142(10):3176-3189.

20. Traylor M, Persyn E, Tomppo L, et al. Genetic basis of lacunar stroke: a pooled analysis of individual patient data and genome-wide association studies. *Lancet Neurol*. 2021;20(5):351-361.

21. Bondareva O, Rodríguez-Aguilera JR, Oliveira F, et al. Single-cell profiling of vascular endothelial cells reveals progressive organ-specific vulnerabilities during obesity. *Nat Metab*. 2022;4(11):1591-1610.

22. Zhang J, Liu Q. Cholesterol metabolism and homeostasis in the brain. *Protein Cell*. 2015;6(4):254-264. doi:10.1007/s13238-014-0131-3

23. Ma R, Kutchy NA, Chen L, Meigs DD, Hu G. Primary cilia and ciliary signaling pathways in aging and age-related brain disorders. *Neurobiol Dis*. 2022;163:105607. doi:10.1016/j.nbd.2021.105607

24. Hanley JG, Khatri L, Hanson PI, Ziff EB. NSF ATPase and α-/β-SNAPs disassemble the AMPA receptor-PICK1 complex. *Neuron*. 2002;34(1):53-67.

25. Genovese G, Rockweiler NB, Gorman BR, et al. BCFtools/liftover: an accurate and comprehensive tool to convert genetic variants across genome assemblies. *Bioinformatics*. 2024;40(2):btae038. doi:10.1093/bioinformatics/btae038

26. Najar CFBA, Feng R, Dai C, et al. Genetic and functional analysis of unproductive splicing using LeafCutter2. *bioRxiv*. Preprint posted online April 8, 2025:2025.04.06.646893. doi:10.1101/2025.04.06.646893

27. Fujita M, Gao Z, Zeng L, et al. Cell subtype-specific effects of genetic variation in the Alzheimer's disease brain. *Nat Genet*. 2024;56(4):605-614. doi:10.1038/s41588-024-01685-y

28. Wang M, Beckmann ND, Roussos P, et al. The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease. *Sci Data*. 2018;5(1):180185. doi:10.1038/sdata.2018.185

29. Johnson ECB, Dammer EB, Duong DM, et al. Large-scale Proteomic Analysis of Alzheimer's Disease Brain and Cerebrospinal Fluid Reveals Early Changes in Energy Metabolism Associated with Microglia and Astrocyte Activation. *Nat Med*. 2020;26(5):769-780. doi:10.1038/s41591-020-0815-6

30. De Jager PL, Ma Y, McCabe C, et al. A multi-omic atlas of the human frontal cortex for aging and Alzheimer's disease research. *Sci Data*. 2018;5(1):180142. doi:10.1038/sdata.2018.142

31. Iturria-Medina Y, Adewale Q, Khan AF, et al. Unified epigenomic, transcriptomic, proteomic, and metabolomic taxonomy of Alzheimer's disease progression and heterogeneity. *Sci Adv*. 8(46):eabo6764. doi:10.1126/sciadv.abo6764

32. Fernandez MV, Liu M, Beric A, et al. Genetic and multi-omic resources for Alzheimer disease and related dementia from the Knight Alzheimer Disease Research Center. *Sci Data*. 2024;11(1):768. doi:10.1038/s41597-024-03485-9

33. NIAGADS DSS Releases Additional 17K Whole Genomes. DSS NIAGADS. March 1, 2021. Accessed November 6, 2025. https://dss.niagads.org/niagads-dss-releases-additional-17k-whole-genomes/

34. Wang G, Sarkar A, Carbonetto P, Stephens M. A Simple New Approach to Variable Selection in Regression, with Application to Genetic Fine Mapping. *J R Stat Soc Ser B Stat Methodol*. 2020;82(5):1273-1300. doi:10.1111/rssb.12388

35. Julienne H, Shi H, Pasaniuc B, Aschard H. RAISS: robust and accurate imputation from summary statistics. *Bioinformatics*. 2019;35(22):4837-4839. doi:10.1093/bioinformatics/btz466

36. Berisa T, Pickrell JK. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics*. 2016;32(2):283-285. doi:10.1093/bioinformatics/btv546

37. Zou H, Hastie T. Regularization and Variable Selection Via the Elastic Net. *J R Stat Soc Ser B Stat Methodol*. 2005;67(2):301-320. doi:10.1111/j.1467-9868.2005.00503.x

38. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *J R Stat Soc Ser B Methodol*. 1996;58(1):267-288. doi:10.1111/j.2517-6161.1996.tb02080.x

39. da Silva FA, Viana AP, Correa CCG, et al. Bayesian ridge regression shows the best fit for SSR markers in Psidium guajava among Bayesian models. *Sci Rep*. 2021;11(1):13639. doi:10.1038/s41598-021-93120-z

40. Park T, Casella G. The Bayesian Lasso. *J Am Stat Assoc*. 2008;103(482):681-686. doi:10.1198/016214508000000337

41. stephenslab/mr.ash.alpha. Published online February 14, 2025. Accessed August 12, 2025. https://github.com/stephenslab/mr.ash.alpha

42. Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology. *Nat Genet*. 2000;25(1):25-29. doi:10.1038/75556

43. Kanehisa M, Furumichi M, Sato Y, Matsuura Y, Ishiguro-Watanabe M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res*. 2025;53(D1):D672-D677. doi:10.1093/nar/gkae909

44. Xu S, Hu E, Cai Y, et al. Using clusterProfiler to characterize multiomics data. *Nat Protoc*. 2024;19(11):3292-3320. doi:10.1038/s41596-024-01020-z

45. Dolgalev I. msigdbr: MSigDB gene sets for multiple organisms in a tidy data format. *R Package Version*. 2022;7(1).

46. Finucane HK, Reshef YA, Anttila V, et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet*. 2018;50(4):621-629. doi:10.1038/s41588-018-0081-4

47. Weissbrod O, Hormozdiari F, Benner C, et al. Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat Genet*. 2020;52(12):1355-1363. doi:10.1038/s41588-020-00735-5

48. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393

49. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236-1241. doi:10.1038/ng.3406

50. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291-295. doi:10.1038/ng.3211

51. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi:10.1086/519795

52. Jiang C, van der Lee SJ, Tesi N, van der Flier WM, Tijms BM, Reus LM. Identification of novel candidate loci for Alzheimer's disease and related dementias by leveraging the shared genetic basis with hippocampal volume. *Aging Brain*. 2025;8:100147.

53. Li T, Lu D, Yao C, et al. Kansl1 haploinsufficiency impairs autophagosome-lysosome fusion and links autophagic dysfunction with Koolen-de Vries syndrome in mice. *Nat Commun*. 2022;13(1):931.

54. Lee YL, Santé J, Comerci CJ, et al. Cby1 promotes Ahi1 recruitment to a ring-shaped domain at the centriole–cilium interface and facilitates proper cilium formation and function. *Mol Biol Cell*. 2014;25(19):2919-2933.

55. Chen Y, Liang R, Li Y, et al. Chromatin accessibility: biological functions, molecular mechanisms and therapeutic application. *Signal Transduct Target Ther*. 2024;9(1):340.

56. Pang ZP, Bacaj T, Yang X, Zhou P, Xu W, Südhof TC. Doc2 supports spontaneous synaptic transmission by a Ca2+-independent mechanism. *Neuron*. 2011;70(2):244-251.

57. Yaguchi H, Yabe I, Takahashi H, et al. Sez6l2 regulates phosphorylation of ADD and neuritogenesis. *Biochem Biophys Res Commun*. 2017;494(1-2):234-241.

58. Kawakami K, Fujita Y, Matsuda Y, et al. Gamma-glutamyltransferase activity in exosomes as a potential marker for prostate cancer. *BMC Cancer*. 2017;17(1):316.

59. Zahnleiter D, Hauer NN, Kessler K, et al. MAP 4-Dependent Regulation of Microtubule Formation Affects Centrosome, Cilia, and G olgi Architecture as a Central Mechanism in Growth Regulation. *Hum Mutat*. 2015;36(1):87-97.

60. Crotti A, Sait HR, McAvoy KM, et al. BIN1 favors the spreading of Tau via extracellular vesicles. *Sci Rep*. 2019;9(1):9477.

61. Xu Y, Cui L, Dibello A, et al. DNAJC5 facilitates USP19-dependent unconventional secretion of misfolded cytosolic proteins. *Cell Discov*. 2018;4(1):11.

62. Zhao J, Cai B, Shao Z, et al. TRIM26 positively regulates the inflammatory immune response through K11-linked ubiquitination of TAB1. *Cell Death Differ*. 2021;28(11):3077-3091.

63. Konjikusic MJ, Lee C, Yue Y, et al. Kif9 is an active kinesin motor required for ciliary beating and proximodistal patterning of motile axonemes. *J Cell Sci*. 2023;136(5):jcs259535.

64. González-Rodríguez P, Delorme-Axford E, Bernard A, et al. SETD2 transcriptional control of ATG14L/S isoforms regulates autophagosome–lysosome fusion. *Cell Death Dis*. 2022;13(11):953.

65. Cai CZ, Zhuang XX, Zhu Q, et al. Enhancing autophagy maturation with CCZ1-MON1A complex alleviates neuropathology and memory defects in Alzheimer disease models. *Theranostics*. 2022;12(4):1738.