

## Supplementary Information

### Genotyping, Quality Control (QC), and Imputation, and Population Substructure

The ADNI-1 and ADNI-GO/2 samples were genotyped using the Illumina Human610-Quad, and HumanOmniExpress microarray chips (Illumina, Inc., San Diego, CA), respectively. QC procedures for the genotype data were performed using PLINK [1]. Samples with genotyping call rates < 95% were excluded. Subjects with non-European ancestry were removed based on the self-reported ethnicity. We excluded SNPs with a minor allele frequency (MAF) < 5% or significant deviation from Hardy-Weinberg equilibrium ( $P < 1.0 \times 10^{-6}$ ). A total of 548,010 SNPs for ADNI-1 and 641,106 SNPs for ADNI-GO/2 were retained after cleaning. Genome-wide imputation of SNP allele dosages was performed using the 1000 Genomes reference panel (v37, March, 2012) and IMPUTE2 software after estimating haplotypes using SHAPEIT [2, 3]. Imputed SNPs with an imputation quality estimate ( $R^2 \geq 0.40$ ) were used for association analyses. After QC, the ADNI-1 sample with GWA data consisted of 187 CN, 329 MCI, and 163 AD subjects, and the ADNI-GO/2 contained 118 CN, 252 MCI, and 27 AD subjects with GWA data. Population substructure was evaluated in each dataset with reference populations from the 1000 Genome Project by principal components analysis using EIGENSTRAT [4]. Only individuals clustering with European samples were retained for further analysis. Population substructure in the remaining European ancestry sample was assessed using genotyped SNPs that passed QC. These SNPs were pruned to remove pairs with high linkage disequilibrium (LD) based on pairwise LD ( $r^2 > 0.2$ ) using PLINK, and using a window size of 1,500 SNPs. Principal components (PCs) of ancestry were derived using the smartpca script in EIGENSTRAT. The first three PCs were included as covariates in association test models to minimize spurious associations and maximize power to detect true associations.

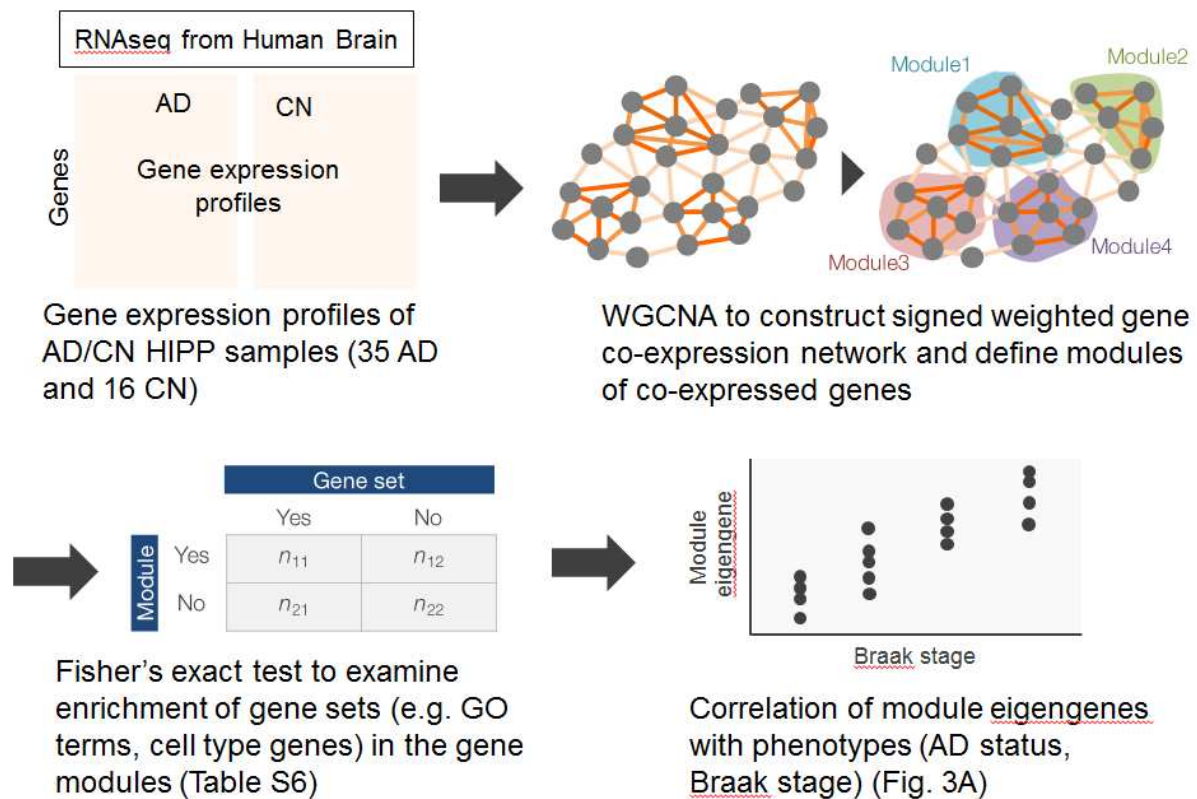
### Eisai Bio Bank Description

The Eisai Bio Bank has human brain samples from University of Miami and McLean Hospital. Total RNA purified from hippocampal brain regions were extracted with Qiagen RNeasy Mini Kits (Qiagen, Hilden, Germany). Poly-A-enriched mRNA was reverse transcribed and amplified using the Nugen Ovation Kit (NuGEN, San Carlos, CA, USA). Paired-end cDNA was sequenced with an Illumina MiSeq at 106 base pair length (Illumina, San Diego, CA, USA). RNA-sequencing reads were checked with FASTX-Toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit](http://hannonlab.cshl.edu/fastx_toolkit)), trimmed with Trimmomatic, and aligned to the GRCh38 assembly transcriptome with Bowtie [5]. Transcript expression levels were estimated in transcripts per million (TPM) using RSEM [6]. Samples of RNA integrity numbers (RIN) larger than 5 were kept for downstream analysis.

Source	Total	Braak Stages								AD cases			Controls		
	N	NA	0	1	2	3	4	5	6	N (Female)	Age (Mean/SD)	RIN (Mean/SD)	N (Female)	Age (Mean/SD)	RIN (Mean/SD)
McLean	26	0	1	6	3	4	3	3	6	16 (7)	76.1 (5.5)	6.7 (0.95)	10 (5)	66.8 (8.9)	6.6 (1.1)
U of Miami	25	6	0	4	1	5	3	5	1	19 (11)	85.8 (9.0)	6.6 (0.89)	6 (3)	85.7 (6.6)	6.8 (1.0)

## Co-expression network analysis

We constructed a gene co-expression network in the human hippocampus transcriptomic data in the Eisai Bio Bank using an R package WGCNA [7]. The Flow chart of co-expression network using RNAseq data from the human hippocampal area is the following.



This software defines module of genes highly co-expressed based on Pearson correlation analysis of gene expression values. The correlation coefficients were applied to the following function for calculating adjacency matrix  $\alpha$ .

$$\alpha_{ij} = |0.5 + 0.5 \times \text{cor}(x_i, x_j)|^\beta$$

where  $x_i$  is a vector including expression levels of gene  $i$  in 38 samples. This power function was used for reducing effects of weak correlation and increasing effect of strong correlation. The

parameter  $\beta$  was 12 in this study to meet scale-free topology of the co-expression networks. The adjacency  $\alpha$  was used to calculate topological overlap matrix (TOM), an indicator assigned for every pair of genes and representing whether the pair is closely located in the co-expression network. Hierarchical clustering of the genes was performed based on TOM with the average linkage method [8]. Co-expression modules were defined from the hierarchical tree by “dynamic tree cut algorithm” [9]. The modules including less than 100 genes were considered as noisy modules and ignored for the further analyses. Expression patterns of the modules were summarized with their first principal components (module eigengenes). Pearson correlations between gene expressions and the module eigengenes were calculated as module membership values [7].

#### **Definition of brain cell type-specific genes**

We used single-cell RNA sequencing data of human brain cells published by Darmanis et al. [10]. Mapping results were downloaded from Gene Expression Omnibus (Accession number: GSE67835) and the mapped read counts were normalized into CPM (count per million) as previously described [10]. Cell type-specific genes were defined when median and mean expression levels (CPM) of genes greater than 4.0.

**Supplementary Table 1.** Sample size and characteristics of three endophenotypes at the baseline in ADNI 1 and ADNI GO/2.

Trait	Subgroup	ADNI 1			ADNI GO/2		
		N (Female %)	Age (Mean/SD)	Trait (Mean/SD)	N (Female %)	Age (Mean/SD)	Trait (Mean/SD)
CSF A $\beta$ <sub>42</sub>	ALL	365 (40)	74.9 (6.9)	167.6 (55.8)	365 (44.1)	72.9 (7.3)	215.6 (73.1)
	CN	96 (47.9)	75.6 (4.9)	201.2 (55.6)	110 (49.1)	74.8 (5.6)	231.2 (68.7)
	MCI	177 (34.5)	74.6 (7.2)	161.4 (54.1)	230 (42.6)	71.7 (7.4)	215.5 (73.9)
	AD	92 (42.4)	74.6 (7.8)	144.3 (41.6)	25 (36)	75.7 (10.8)	147.3 (39.8)
CSF t-Tau	ALL	365 (40)	74.9 (6.9)	98.2 (53.7)	359 (44.6)	73 (7.3)	88.2 (52.7)
	CN	96 (47.9)	75.6 (4.9)	71.9 (30.6)	107 (49.5)	74.8 (5.6)	72.9 (33)
	MCI	177 (34.5)	74.6 (7.2)	101.1 (55.7)	228 (43)	71.7 (7.4)	90.2 (56.4)
	AD	92 (42.4)	74.6 (7.8)	120.1 (57.8)	24 (37.5)	76.3 (10.7)	137.1 (56.4)
CSF p-Tau	ALL	362 (40.1)	74.9 (6.9)	34.4 (18.6)	364 (44)	72.9 (7.4)	25.2 (12.4)
	CN	94 (47.9)	75.6 (5)	25.5 (14.8)	109 (48.6)	74.8 (5.6)	22.6 (8.9)
	MCI	176 (34.7)	74.6 (7.3)	35.4 (17.7)	230 (42.6)	71.7 (7.4)	25.4 (13.3)
	AD	92 (42.4)	74.6 (7.8)	41.3 (20.3)	25 (36)	75.7 (10.8)	34 (12.7)
HPV	ALL	546 (40.3)	74.9 (6.7)	6500.2 (1191.3)	337 (44.5)	72.3 (7.1)	7155.2 (1101.4)
	CN	166 (43.4)	75.6 (4.8)	7297.8 (922)	102 (50)	74.1 (5.4)	7397.6 (848.2)
	MCI	261 (36.4)	74.6 (7.1)	6378.9 (1082.6)	213 (42.3)	71.1 (7.1)	7177.1 (1125.8)
	AD	119 (44.5)	74.6 (7.8)	5653.9 (1065)	22 (40.9)	74.9 (10.4)	5818.8 (996.2)
LMiT	ALL	679 (39.9)	75.4 (6.7)	8.4 (4.9)	397 (43.8)	72.8 (7.3)	10.8 (4.2)
	CN	187 (43.9)	75.9 (4.9)	14.1 (3.4)	118 (49.2)	74.7 (5.6)	14.4 (2.9)
	MCI	329 (35)	75.3 (7.1)	7.2 (3.2)	252 (42.1)	71.6 (7.4)	9.9 (3.3)
	AD	163 (45.4)	75.3 (7.7)	4.2 (2.9)	27 (37)	75.1 (10.5)	4.1 (2.6)
LMdT	ALL	679 (39.9)	75.4 (6.7)	5.8 (5.5)	397 (43.8)	72.8 (7.3)	9.1 (4.4)
	CN	187 (43.9)	75.9 (4.9)	13.3 (3.6)	118 (49.2)	74.7 (5.6)	13.6 (3.1)
	MCI	329 (35)	75.3 (7.1)	3.8 (2.7)	252 (42.1)	71.6 (7.4)	7.7 (2.8)
	AD	163 (45.4)	75.3 (7.7)	1.2 (1.9)	27 (37)	75.1 (10.5)	1.7 (2)

A $\beta$ <sub>42</sub>, t-Tau, and p-Tau: CSF biomarkers of A $\beta$ <sub>42</sub> and total and phosphorylated tau proteins. HPV: the MRI scan of hippocampal volume. LMiT and LMdT: the scores from logical memory tests (immediate and delayed recalls). CN: clinically normal subjects. MCI: subjects with mild cognitive impairment. ALL: a composed group of CN, MCI and AD subjects

**Supplementary Table 2.** Association of *APOE*  $\epsilon 4$  carrier status with CSF levels of  $A\beta_{42}$ , total tau (t-Tau), and phosphorylated tau (p-Tau), with hippocampal volume (HPV), and with logical memory tests of immediate (LMiT) and delayed (LMdT) recall.

Trait	Subgroup	$\epsilon 4$ FREQ	BETA	SE	P
CSF	ALL	0.22	-0.95	0.07	$5.43 \times 10^{-41}$
	CN	0.12	-0.87	0.14	$5.46 \times 10^{-9}$
$A\beta_{42}$	MCI	0.24	-0.89	0.09	$1.48 \times 10^{-20}$
	AD	0.34	-0.73	0.15	$6.18 \times 10^{-6}$
CSF	ALL	0.22	0.69	0.07	$5.31 \times 10^{-21}$
	CN	0.12	0.26	0.14	0.06
t-Tau	MCI	0.24	0.76	0.09	$1.37 \times 10^{-14}$
	AD	0.34	0.23	0.19	0.24
CSF	ALL	0.22	0.63	0.07	$2.19 \times 10^{-17}$
	CN	0.12	0.30	0.15	0.04
p-Tau	MCI	0.24	0.68	0.10	$1.04 \times 10^{-11}$
	AD	0.34	0.23	0.18	0.22
HPV	ALL	0.23	-0.55	0.07	$1.20 \times 10^{-16}$
	CN	0.13	-0.05	0.11	0.66
	MCI	0.26	-0.29	0.09	$8.04 \times 10^{-4}$
	AD	0.35	-0.59	0.14	$5.26 \times 10^{-5}$
LMiT	ALL	0.23	-0.49	0.06	$4.05 \times 10^{-15}$
	CN	0.13	-0.08	0.09	0.36
	MCI	0.25	-0.19	0.06	$2.87 \times 10^{-3}$
	AD	0.33	0.07	0.12	0.59
LMdT	ALL	0.23	-0.55	0.06	$3.77 \times 10^{-19}$
	CN	0.13	-0.08	0.08	0.28
	MCI	0.25	-0.21	0.06	$2.84 \times 10^{-4}$
	AD	0.33	-0.05	0.10	0.61

$\epsilon 4$  FREQ is the  $\epsilon 4$  allele frequency.

**Supplementary Table 3.** Association *P*-values of Alzheimer disease loci previously established by GWAS with AD-related endophenotypes according to cognitive group (Separate spreadsheet)

**Supplementary Table 4.** Effect sizes and P-values for additional SNPs that show a suggestive association ( $P < 1.0 \times 10^{-6}$ ) for CSF  $A\beta_{42}$ , total tau and phosphorylated tau levels; MRI measure of hippocampal volume; and test scores for immediate and delayed recall in cognitively normal, mild cognitively impaired, and AD subjects. (Separate spreadsheet)



**Supplementary Table 5.** Top-ranked genes based on the genetic association tests in the CN and MCI subgroups with  $P < 10^{-6}$ .

<b>ID</b>	<b>Symbol</b>	<b>Entrez Gene Name</b>
ADAMTS18	ADAMTS18	ADAM metallopeptidase with thrombospondin type 1 motif 18
ADAMTS3	ADAMTS3	ADAM metallopeptidase with thrombospondin type 1 motif 3
AKR7A3	AKR7A3	aldo-keto reductase family 7 member A3
ANAPC4	ANAPC4	anaphase promoting complex subunit 4
ANO2	ANO2	anoctamin 2
APOE	APOE	apolipoprotein E
ARHGAP24	ARHGAP24	Rho GTPase activating protein 24
ART3	ART3	ADP-ribosyltransferase 3
BCL2L14	BCL2L14	BCL2 like 14
BEST3	BEST3	bestrophin 3
BRIP1	BRIP1	BRCA1 interacting protein C-terminal helicase 1
C20orf96	C20orf96	chromosome 20 open reading frame 96
C6orf58	C6orf58	chromosome 6 open reading frame 58
CAMK2B	CAMK2B	calcium/calmodulin dependent protein kinase II beta
CCDC3	CCDC3	coiled-coil domain containing 3
CDCP1	CDCP1	CUB domain containing protein 1
CLIC5	CLIC5	chloride intracellular channel 5
CMAHP	CMAHP	cytidine monophospho-N-acetylneuraminic acid hydroxylase, pseudogene
CYTH3	CYTH3	cytohesin 3
DAB1	DAB1	DAB1, reelin adaptor protein
DENND4A	DENND4A	DENN domain containing 4A
DMGDH	DMGDH	dimethylglycine dehydrogenase
ERBB2IP	ERBIN	erbB2 interacting protein
FILIP1	FILIP1	filamin A interacting protein 1
FLJ37035	FLJ37035	uncharacterized LOC399821
FRA10AC1	FRA10AC1	fragile site, folic acid type, rare, fra(10)(q23.3) or fra(10)(q24.2) candidate 1
FSHR	FSHR	follicle stimulating hormone receptor
FSIP1	FSIP1	fibrous sheath interacting protein 1
GOLM1	GOLM1	golgi membrane protein 1
GRIN2B	GRIN2B	glutamate ionotropic receptor NMDA type subunit 2B
IFITM10	IFITM10	interferon induced transmembrane protein 10
JPH3	JPH3	junctionophilin 3
KCNE4	KCNE4	potassium voltage-gated channel subfamily E regulatory subunit 4
KCNK10	KCNK10	potassium two pore domain channel subfamily K member 10
KIAA2026	KIAA2026	KIAA2026
LRRTM4	LRRTM4	leucine rich repeat transmembrane neuronal 4
MCM5	MCM5	minichromosome maintenance complex component 5
MMP28	MMP28	matrix metallopeptidase 28
MTUS1	MTUS1	microtubule associated tumor suppressor 1
NRG1	NRG1	neuregulin 1
INADL	PATJ	PATJ, crumbs cell polarity complex component
PDGFRL	PDGFRL	platelet derived growth factor receptor like
PDS5B	PDS5B	PDS5 cohesin associated factor B

PLD4	PLD4	phospholipase D family member 4
PRKG1	PRKG1	protein kinase, cGMP-dependent, type I
RCAN1	RCAN1	regulator of calcineurin 1
RGS19	RGS19	regulator of G-protein signaling 19
RNF165	RNF165	ring finger protein 165
SDAD1	SDAD1	SDA1 domain containing 1
SNTB2	SNTB2	syntrophin beta 2
SORCS1	SORCS1	sortilin related VPS10 domain containing receptor 1
SRRM4	SRRM4	serine/arginine repetitive matrix 4
SYT9	SYT9	synaptotagmin 9
TEKT5	TEKT5	tektin 5
TNIK	TNIK	TRAF2 and NCK interacting kinase
TRPV2	TRPV2	transient receptor potential cation channel subfamily V member 2
WBSCR17	WBSCR17	Williams-Beuren syndrome chromosome region 17
WDR5	WDR5	WD repeat domain 5
ZCCHC4	ZCCHC4	zinc finger CCHC-type containing 4
ZNF7	ZNF7	zinc finger protein 7
ZNF804B	ZNF804B	zinc finger protein 804B

---

**Supplementary Table 6.** Co-expression network profiles for the modules that contain known AD genes as well as the genes with significant association ( $P < 1.0 \times 10^{-6}$ ) with at least one trait for CN, MCI, or all subjects.

Module (Size)	Gene (MM)	Known AD genes (MM) in a same module	Enriched Gene Ontology (FDR)	Enriched Diseases (FDR)
M1 (1618)	<i>C14orf79</i> (0.49)	<i>MAPT-AS1</i> (0.37)		AD ( $6.2 \times 10^{-3}$ )
M2 (314)	<i>PRKG1</i> (0.88)	<i>GLIS3</i> (0.89) <i>FERMT2</i> (0.80)	Protein binding ( $6.8 \times 10^{-39}$ ) Focal adhesion ( $1.9 \times 10^{-10}$ ) Signal transduction ( $8.8 \times 10^{-10}$ )	AD ( $2.2 \times 10^{-9}$ ) CP ( $3.1 \times 10^{-9}$ ) PD ( $8.4 \times 10^{-7}$ )
M3 (528)	<i>ARHGAP24</i> (0.07)	<i>BZRAP1-AS1</i> (0.59)		
M4 (229)	<i>MTUS1</i> (0.87) <i>ERBIN</i> (0.86)	<i>BACE1-AS</i> (0.72) <i>ZCWPW1</i> (0.67)	Protein binding ( $3.5 \times 10^{-8}$ )	
M5 (879)	<i>AKAP9</i> (0.66)	<i>CELFI</i> (0.82)	DNA binding ( $2.0 \times 10^{-43}$ ) Protein binding ( $7.6 \times 10^{-40}$ ) Transcription ( $3.3 \times 10^{-39}$ )	Schizophrenia ( $3.0 \times 10^{-4}$ ) AD (0.01) PD (0.01)
M6 (843)	<i>APOE</i> (0.65)	<i>APOE</i> (0.65)	Type I interferon signaling (0.03)	
M7 (3885)	<i>GRIN2B</i> (0.95) <i>CAMK2B</i> (0.88) <b><i>SRRM4</i></b> (0.84) <i>NRG1</i> (0.83) <i>JPH3</i> (0.78) <i>DAB1</i> (0.73) <i>ZNF804B</i> (0.55) <i>BRIP1</i> (0.51)	<i>PTK2B</i> (0.89) <i>APP</i> (0.86) <i>PLXNA4</i> (0.84) <i>MAPT</i> (0.75) <i>BZRAP1</i> (0.67) <i>PSEN2</i> (0.64)	Chemical synaptic transmission ( $1.8 \times 10^{-84}$ ) Neuron projection ( $1.7 \times 10^{-30}$ ) Nervous system development Axon guidance ( $1.6 \times 10^{-24}$ ) Postsynaptic density ( $8.6 \times 10^{-22}$ )	AD ( $1.8 \times 10^{-46}$ ) PD ( $1.3 \times 10^{-25}$ ) CP ( $2.4 \times 10^{-23}$ )
M8 (476)	<i>RCANI</i> (0.74)	<i>BIN1</i> (0.87) <i>PSEN1</i> (0.79)	Protein binding ( $2.8 \times 10^{-16}$ ) Microtubule binding ( $4.1 \times 10^{-5}$ ) Regulation cell shape ( $2.3 \times 10^{-4}$ )	AD ( $3.8 \times 10^{-5}$ ) Schizophrenia (0.01) CP (0.01)

Module membership (MM) means a hub-gene score in a module of co-expression network. The FDR indicates false discovery rate estimated using the Benjamin-Hochberg procedure. Enriched diseases were the first three enriched diseases for the module. Bold SNPs denote the study-wide significant association ( $P < 8.33 \times 10^{-9}$ ) with a trait. AD, PD, or CP indicates Alzheimer disease, Parkinson disease, and cognitive performance, respectively.

**Supplementary Table 7.** Correlation expression matrix for genes which were suggestively associated with at least one endophenotypes or were previously identified in GWAS of AD risk (Separate spreadsheet)

**Supplementary Table 8.** Rank of the identified modules in enrichment analysis in co-expression networks from the human hippocampal area

Module	# of Genes in a Module	# of AD-associated Genes in a Module	P-value	FDR
<b>M7</b>	3885	265	1.06x10 <sup>-47</sup>	1.81x10 <sup>-46</sup>
<b>M2</b>	314	32	1.27x10 <sup>-10</sup>	2.17x10 <sup>-9</sup>
Module X1	289	23	3.86x10 <sup>-6</sup>	2.19x10 <sup>-5</sup>
Module X2	379	28	1.57x10 <sup>-6</sup>	2.67x10 <sup>-5</sup>
<b>M8</b>	476	32	2.23x10 <sup>-6</sup>	3.7x10 <sup>-5</sup>
Module X3	379	27	4.71x10 <sup>-6</sup>	4.c01x10 <sup>-5</sup>
Module X4	299	13	1.69x10 <sup>-4</sup>	0.001
<b>M1</b>	1618	66	4.86x10 <sup>-4</sup>	0.006
Module X5	286	18	7.59x10 <sup>-4</sup>	0.006
<b>M5</b>	879	39	0.002	0.011
Module X6	252	16	0.001	0.011
Module X7	244	15	0.003	0.021
Module X8	590	27	0.005	0.022
Module X9	379	12	0.004	0.025

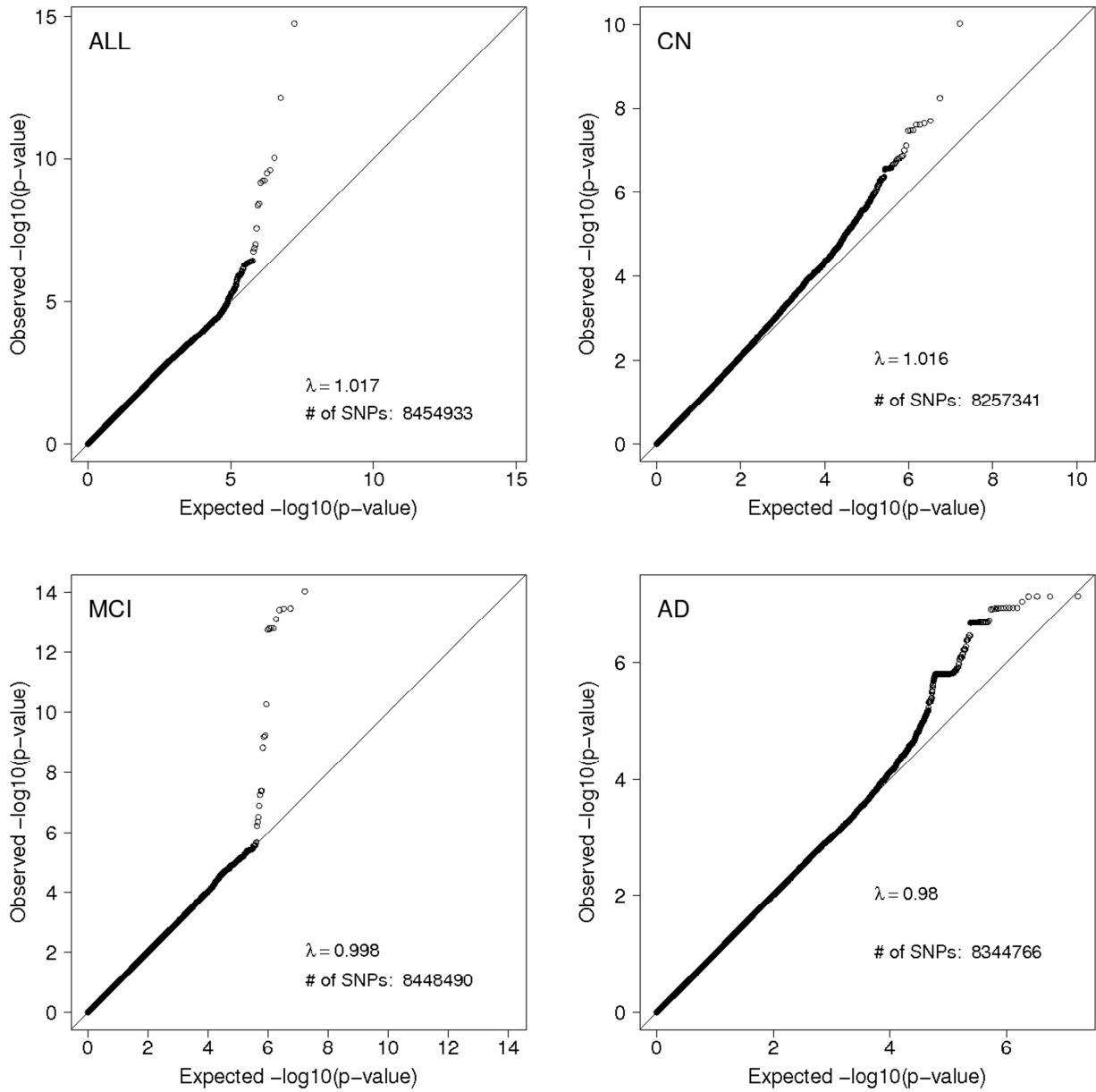
The number of total AD-associated genes is 1,101 obtained from GWASdb2 [11]. The bold modules contain top ranked genes ( $P < 1.0 \times 10^{-6}$ ) in this study.

**Supplementary Table 9.** Genetic associations with CSF biomarkers for the SNPs which were identified in the Deming *et al.* study [12] with genome-wide significant association with CSF biomarkers (Separate spreadsheet)

**Supplementary Table 10.** Previous findings for loci showing at least suggestive evidence for association ( $P < 1.0 \times 10^{-6}$ ) in the current study for AD-related endophenotypes.

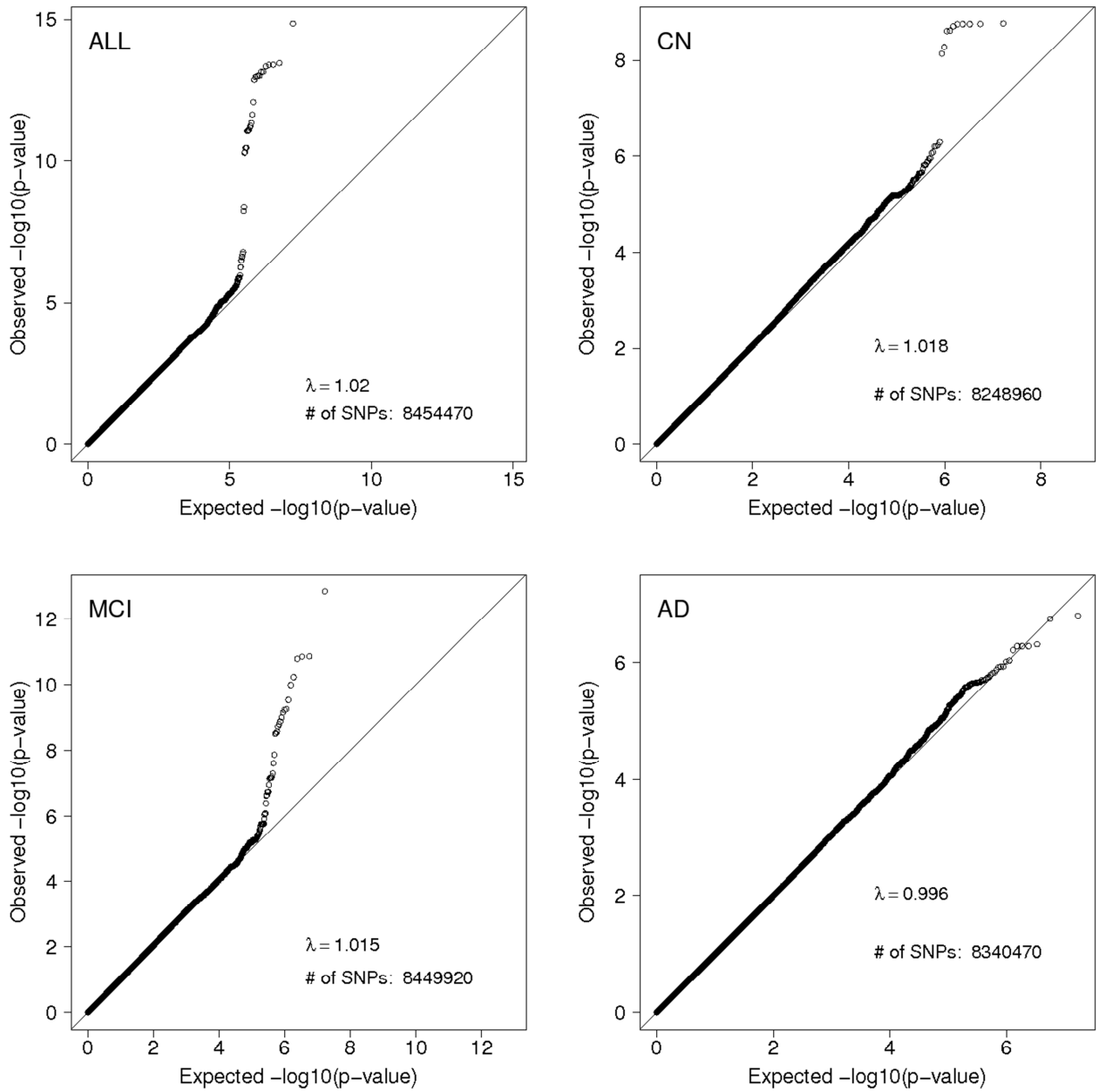
Current Study								Previous Studies			
Trait	Sub	GENE	CH	SNP	Beta	SE	P-value	Statistical Method/Trait	SNP	P-value	Ref
A $\beta$ <sub>42</sub>		<i>AKR7A3</i>	1	rs59348345	0.88	0.18	5.06x10 <sup>-7</sup>	Entire brain (voxel-GWAS)	rs710865	1.10x10 <sup>-7</sup>	[13]
	CN	<i>GRIN2B</i>	12	rs74442473	-1.02	0.18	2.52x10 <sup>-8</sup>	Temporal lobe atrophy	rs11055612	1.20x10 <sup>-7</sup>	[14]
		<i>KCNK10</i>	14	rs17692117	-1.05	0.21	5.30x10 <sup>-7</sup>	Information processing speed	rs17124581	2.71x10 <sup>-6</sup>	[15]
	AD	<i>ERBB4</i>	2	rs839506	-0.60	0.11	1.73x10 <sup>-7</sup>	Copy number variant (AD patients only)	.	.	[16, 17]
t-Tau	CN	<i>SRRM4</i>	12	rs10775009	0.51	0.09	1.59x10 <sup>-9</sup>	GWAS of CSF t-Tau	rs1997111	2.11x10 <sup>-7</sup>	[18]
	AD	<i>RELN</i>	7	rs3819489	-0.78	0.15	2.20x10 <sup>-7</sup>	GWAS of Braak score (from controls)	rs6943822	0.02	[20]
HPV	CN	<i>MTUS1</i>	8	rs4921790	0.61	0.10	4.58x10 <sup>-8</sup>	GWAS of CSF p-Tau	rs7842088	2.12x10 <sup>-7</sup>	[18]
	AD	<i>AKAP9</i>	7	rs149454736	1.91	0.33	4.76x10 <sup>-9</sup>	AD using whole exome sequencing in African Americans	rs144662445 rs149979685	0.002	[20]
LMT	MCI	<i>NRG1</i>	8	rs118130881	-0.61	0.11	1.27x10 <sup>-8</sup>	AD using linkage analysis	rs392499	0.01	[21]

**Supplementary Figure 1.** Quantile-quantile plots of observed (y-axis) vs. expected (x-axis) P-values for CSF A $\beta_{42}$  in the total sample (ALL), Clinically normal (CN), Mild Cognitively Impaired (MCI), and AD subjects.

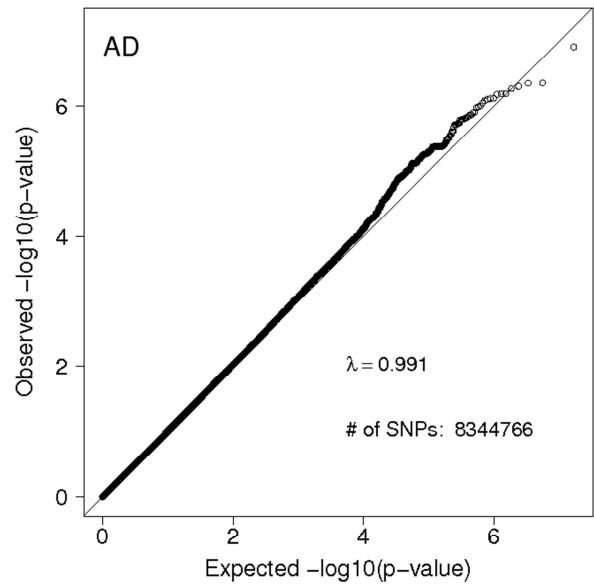
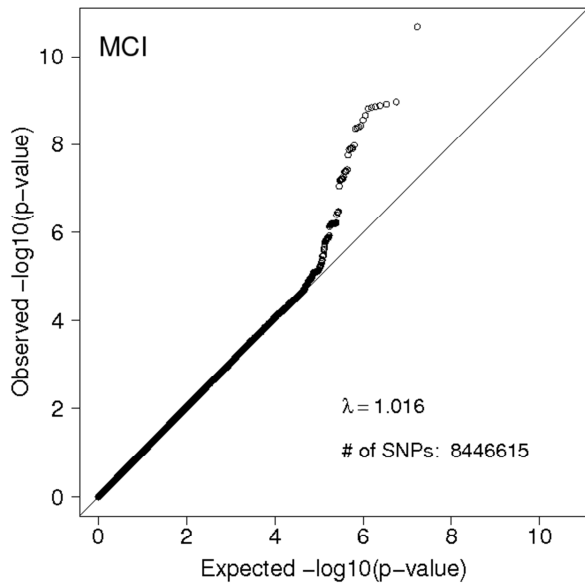
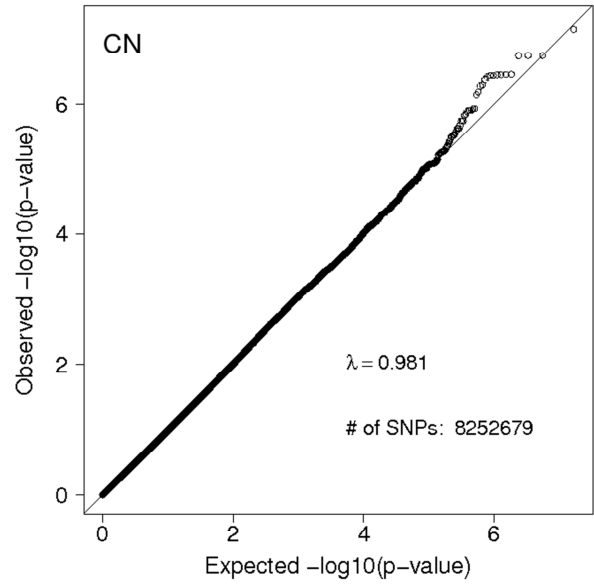
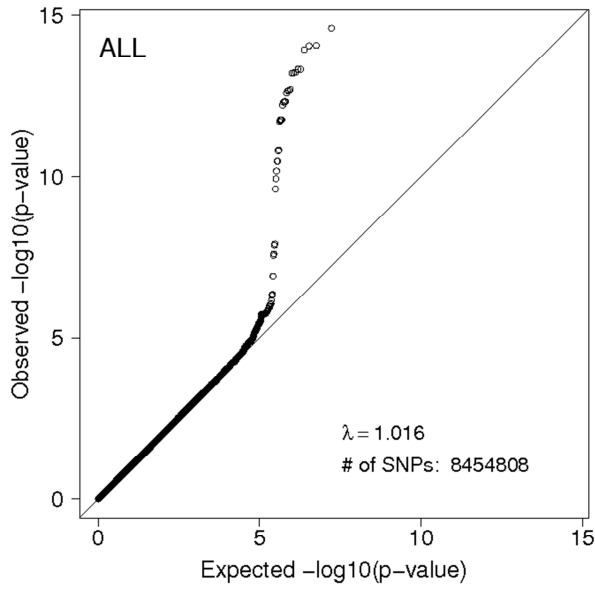




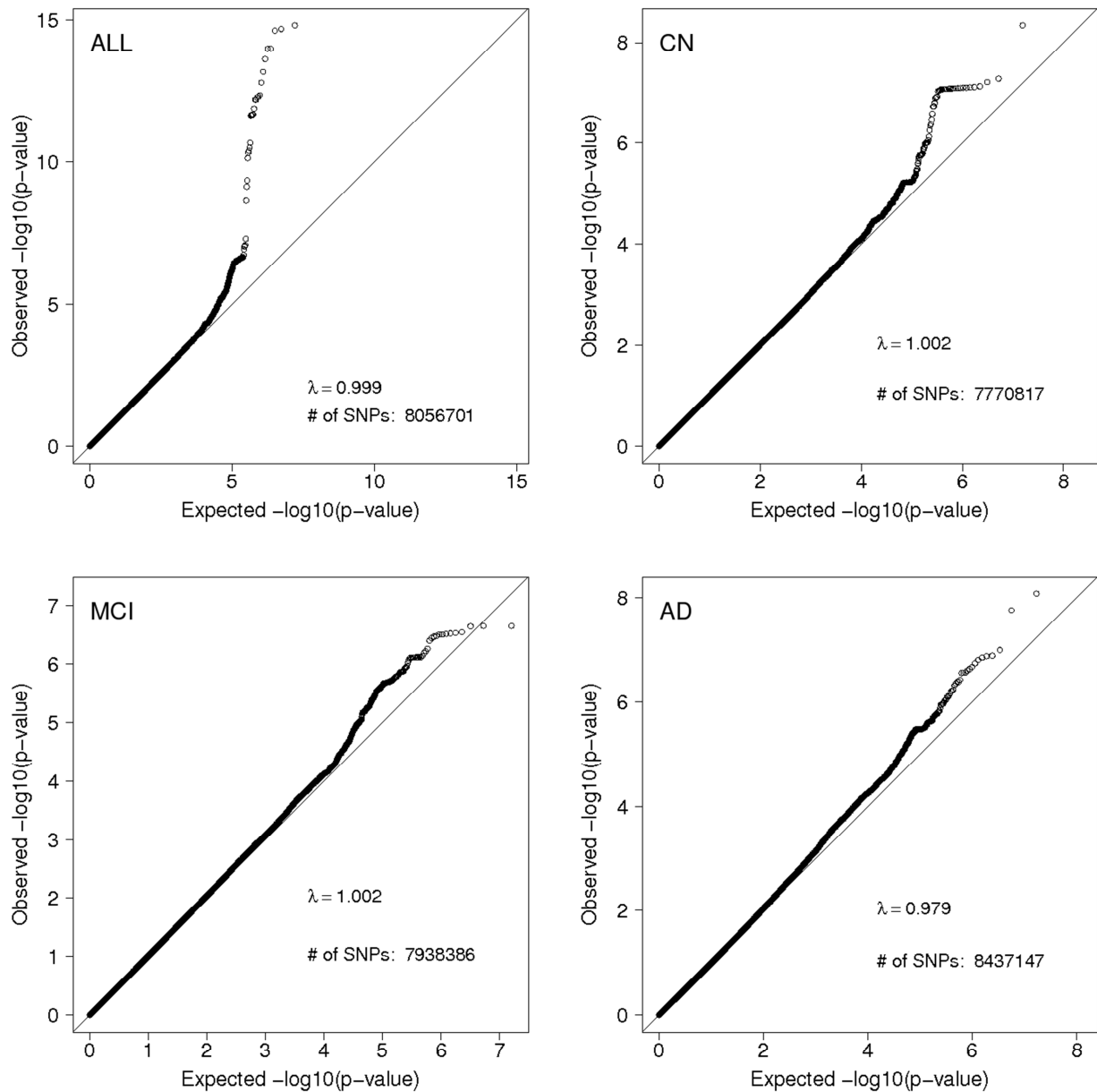
**Supplementary Figure 2.** Quantile-quantile plots of observed (y-axis) vs. expected (x-axis) P-values for CSF t-Tau in the total sample (ALL), Clinically normal (CN), Mild Cognitively Impaired (MCI), and AD subjects.



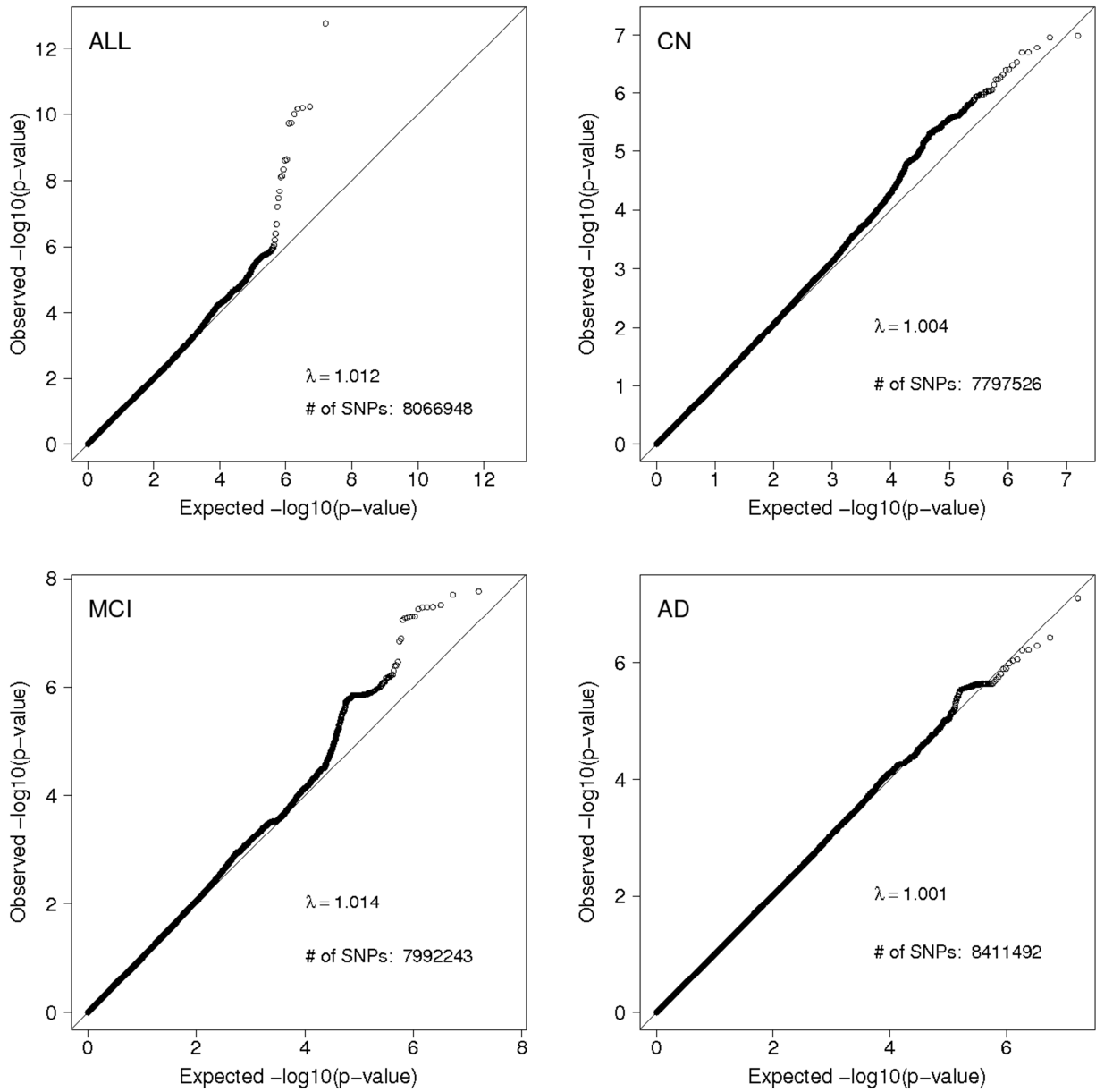
**Supplementary Figure 3.** Quantile-quantile plots of observed (y-axis) vs. expected (x-axis) P-values for CSF p-Tau from the sample in ALL, Clinically normal (CN), Mild Cognitive Impairment (MCI), and AD.



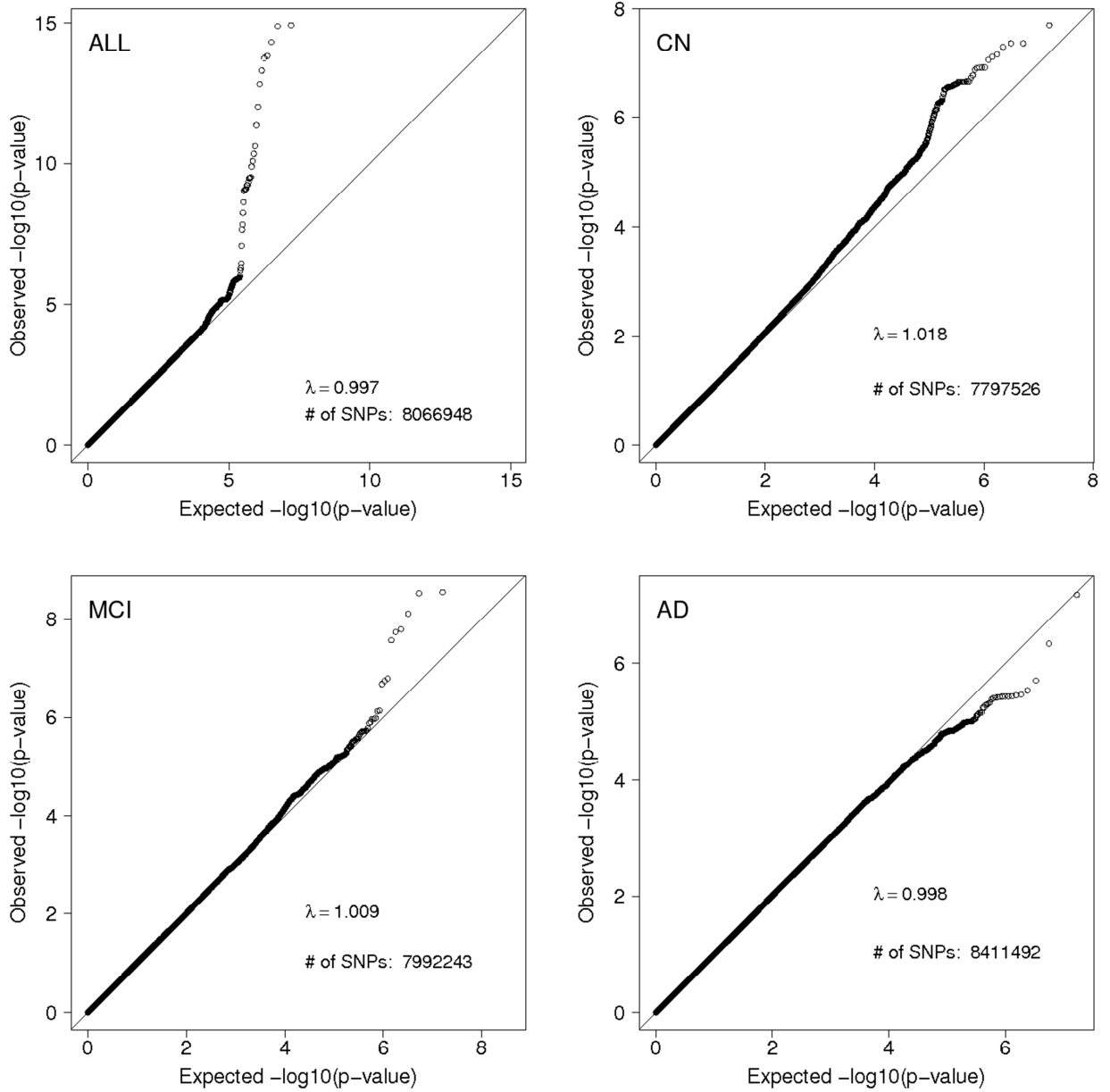
**Supplementary Figure 4.** Quantile-quantile plots of observed (y-axis) vs. expected (x-axis) P-values for Hippocampal volume (HPV) from the sample in ALL, Clinically normal (CN), Mild Cognitive Impairment (MCI), and AD.



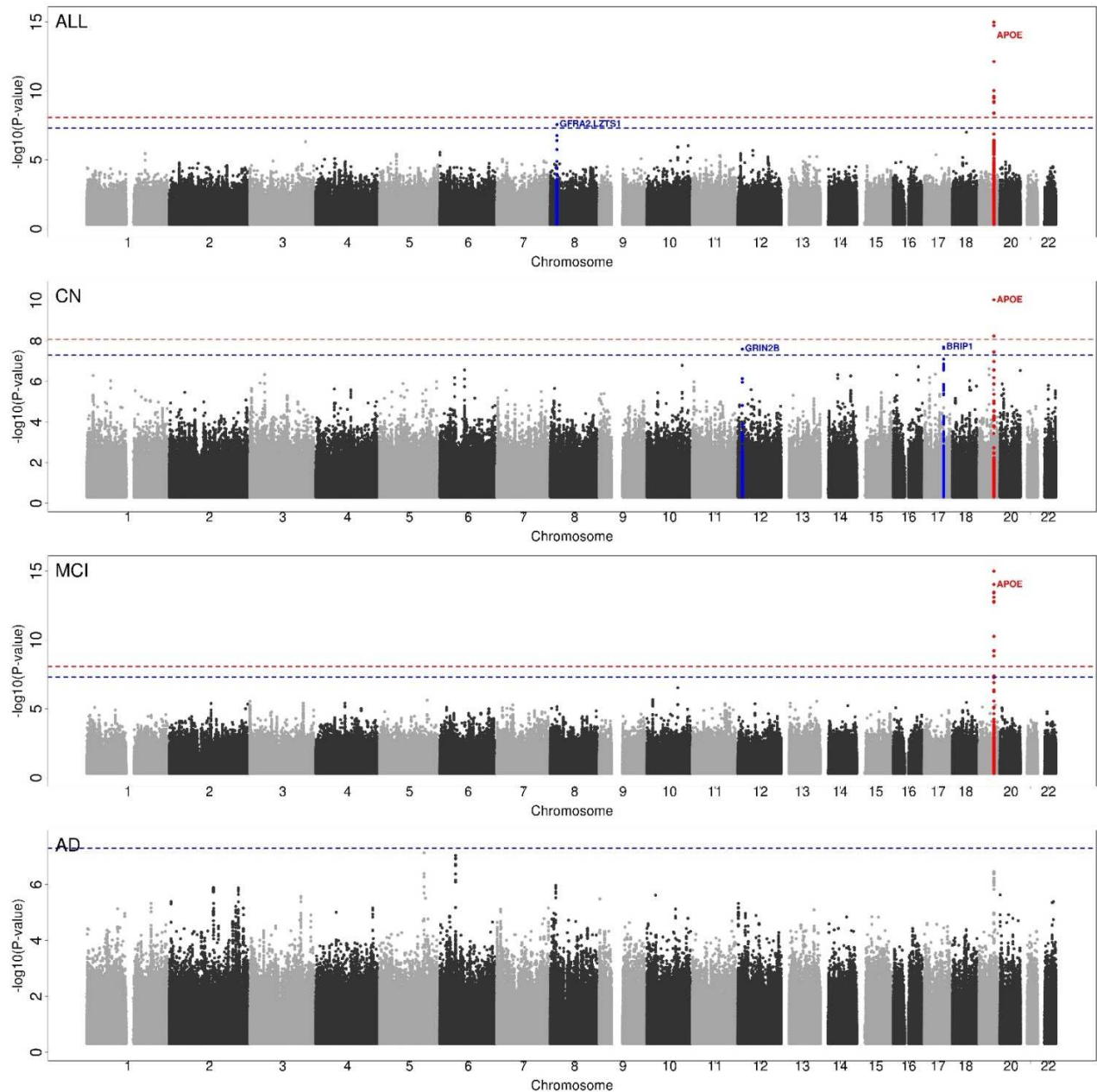
**Supplementary Figure 5.** Quantile-quantile plots of observed (y-axis) vs. expected (x-axis) P-values for logical memory test (LMT)-immediate recalls from the sample in ALL, Clinically normal (CN), Mild Cognitive Impairment (MCI), and AD.



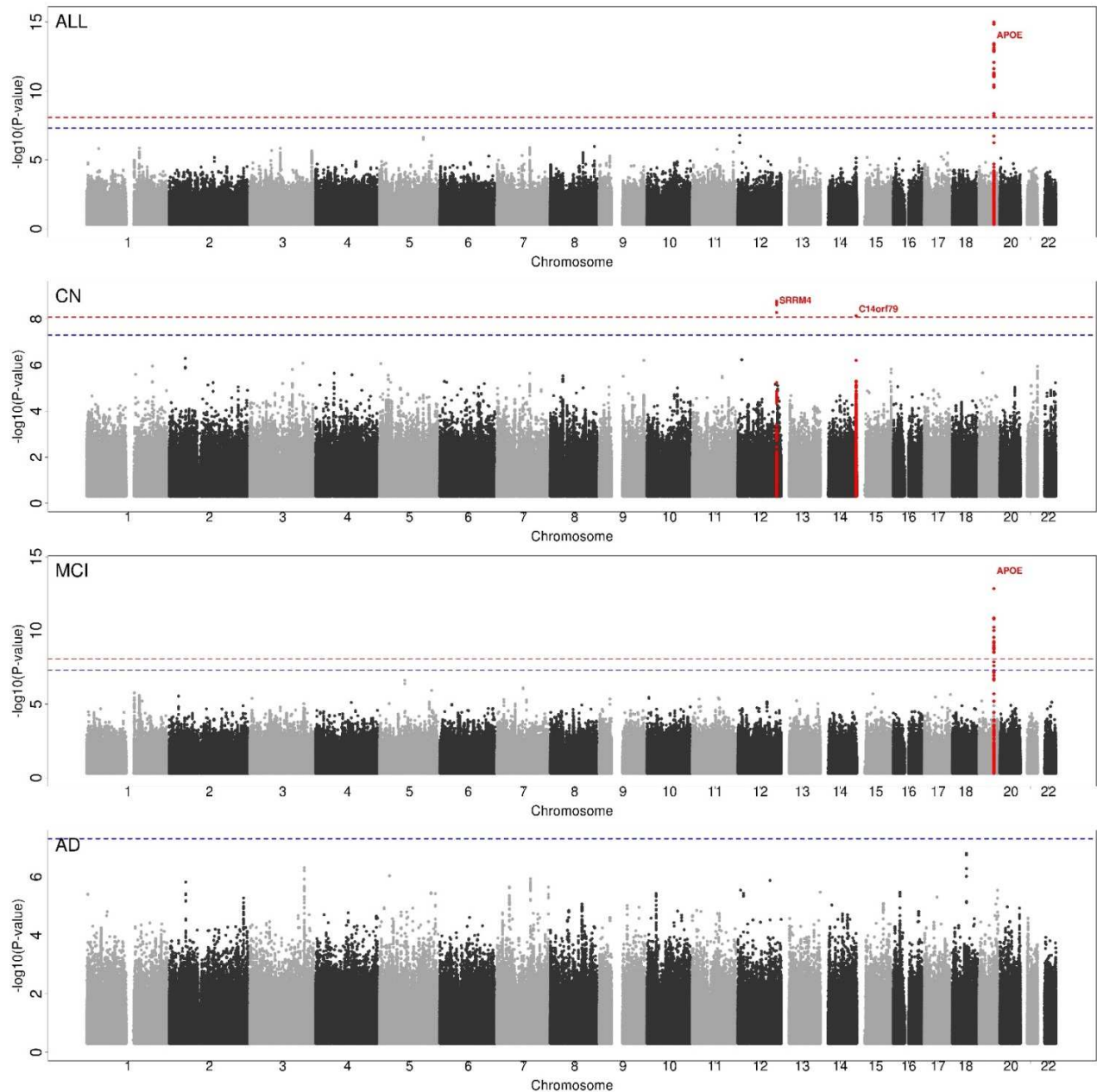
**Supplementary Figure 6.** Quantile-quantile plots of observed (y-axis) vs. expected (x-axis) P-values for logical memory test (LMT)-delayed recalls from the sample in ALL, Clinically normal (CN), Mild Cognitive Impairment (MCI), and AD.



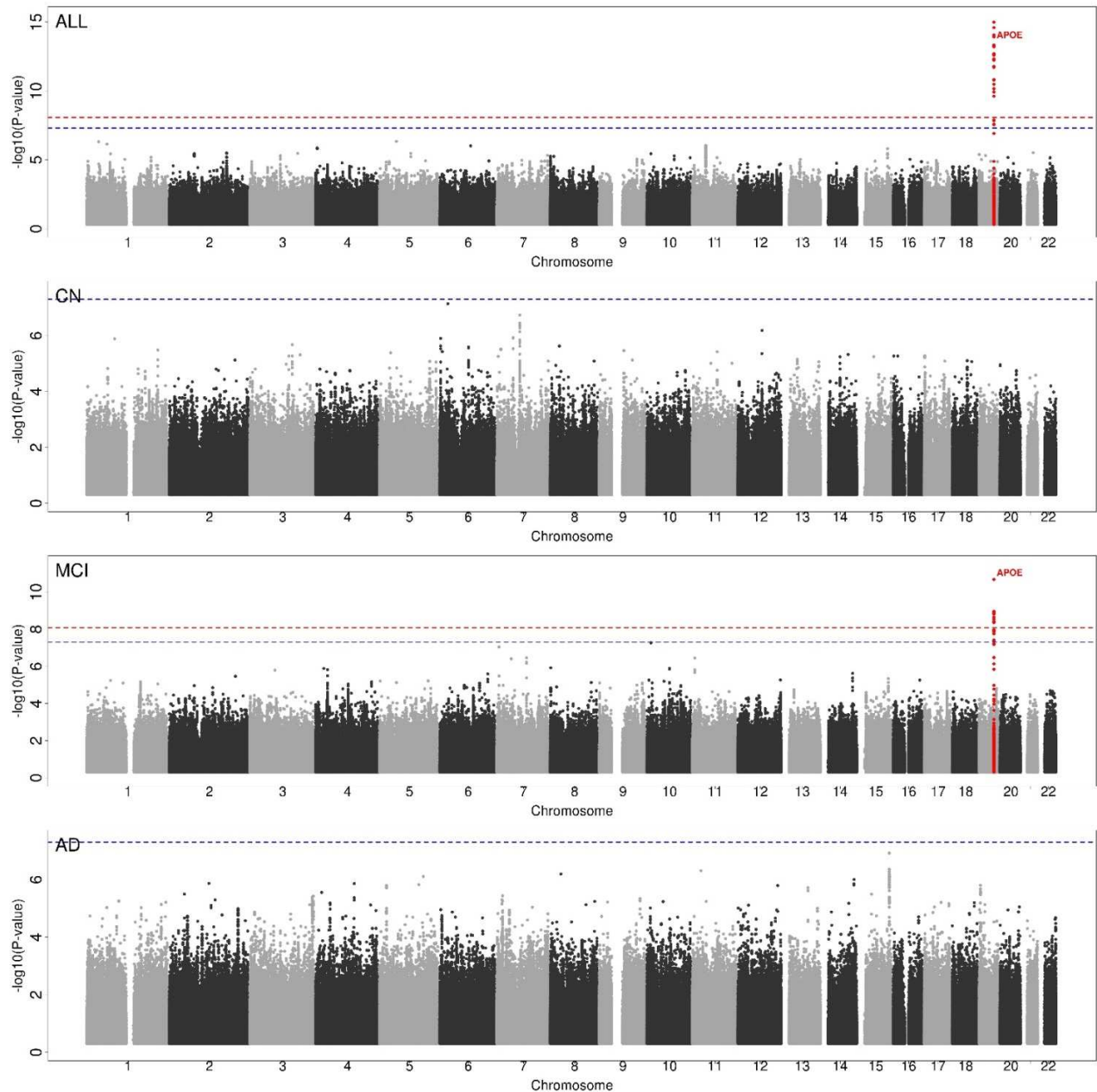
**Supplementary Figure 7.** Manhattan plot showing the meta-analyzed results of CSF A $\beta_{42}$  from all the subjects and three subgroups (CN: clinically normal, MCI: mild cognitive impairment, and AD). Blue dot line represents the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$ , and the red dot line indicates study-wide significant threshold of  $P < 8.33 \times 10^{-9}$ . Loci achieving study-wide significance are highlighted in red, and loci at genome-wide significance are in blue.



**Supplementary Figure 8.** Manhattan plot showing the meta-analyzed results of CSF total Tau from all the subjects and three subgroups (CN: clinically normal, MCI: mild cognitive impairment, and AD). Blue dot line represents the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$ , and the red dot line indicates study-wide significant threshold of  $P < 8.33 \times 10^{-9}$ . Loci achieving study-wide significance are highlighted in red, and loci at genome-wide significance are in blue.

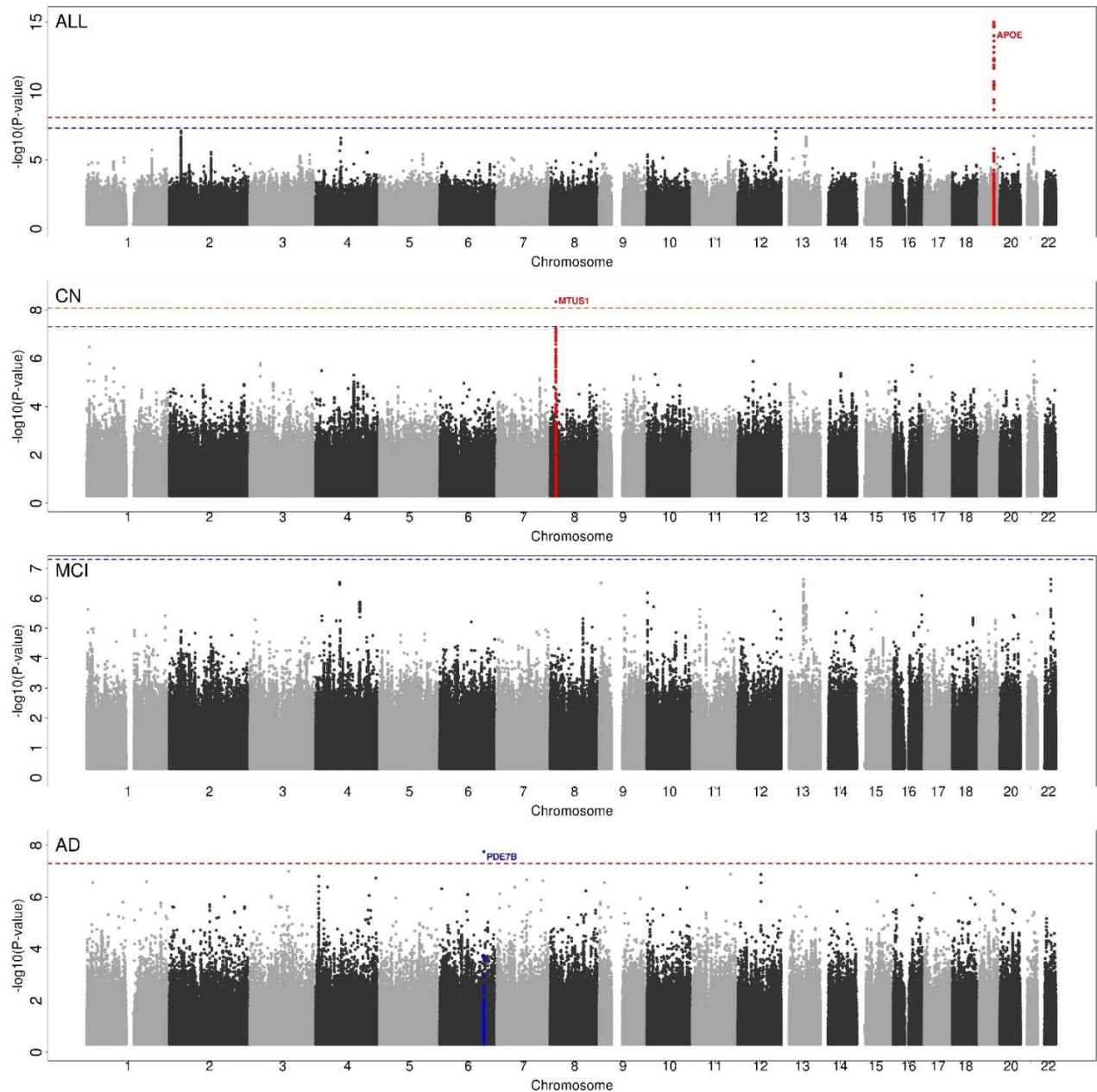


**Supplementary Figure 9.** Manhattan plot showing the meta-analyzed results of CSF phosphorylated tau from all the subjects and three subgroups (CN: clinically normal, MCI: mild cognitive impairment, and AD). Blue dot line represents the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$ , and the red dot line indicates study-wide significant threshold of  $P < 8.33 \times 10^{-9}$ . Loci achieving study-wide significance are highlighted in red, and loci at genome-wide significance are in blue.

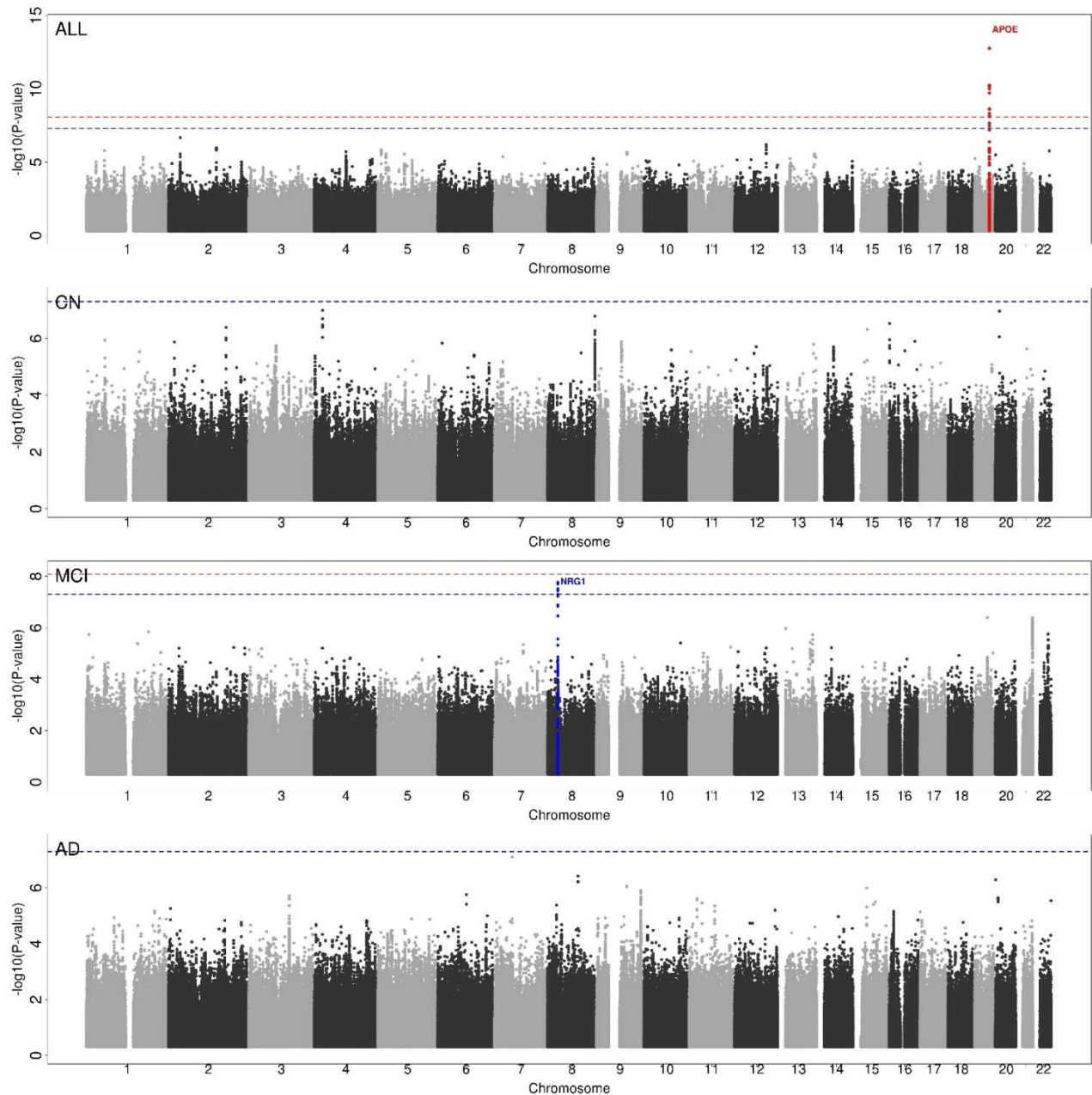




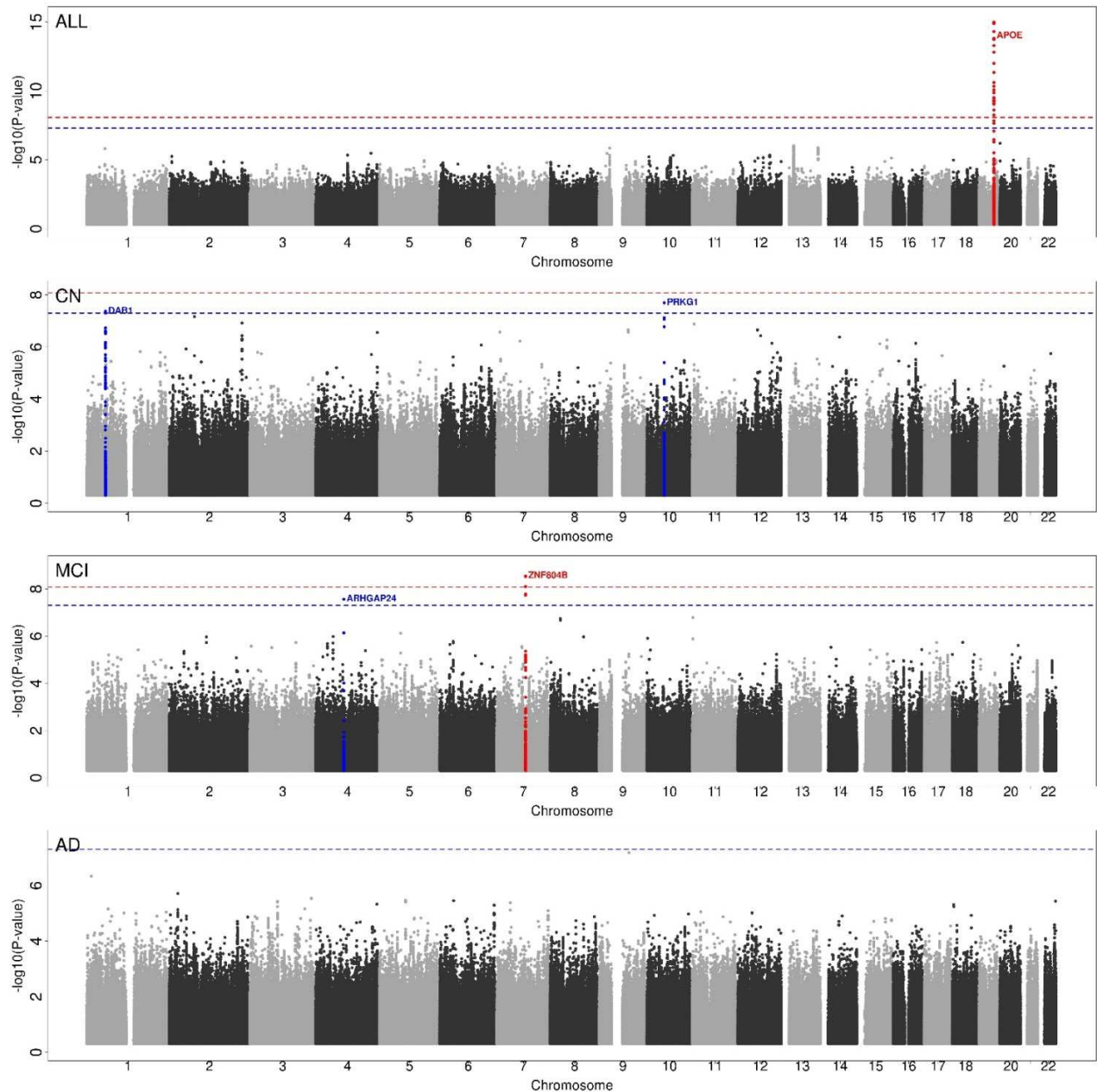
**Supplementary Figure 10.** Manhattan plot showing the meta-analyzed results of hippocampal volume (HPV) in brains from all the subjects and three subgroups (CN: clinically normal, MCI: mild cognitive impairment, and AD). Blue dot line represents the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$ , and the red dot line indicates study-wide significant threshold of  $P < 8.33 \times 10^{-9}$ . Loci achieving study-wide significance are highlighted in red, and loci at genome-wide significance are in blue.



**Supplementary Figure 11.** Manhattan plot showing the meta-analyzed results of logical memory test (LMT)-immediate recall from all the subjects and three subgroups (CN: clinically normal and MCI: mild cognitive impairment). Blue dot line represents the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$ , and the red dot line indicates study-wide significant threshold of  $P < 8.33 \times 10^{-9}$ . Loci achieving study-wide significance are highlighted in red, and loci at genome-wide significance are in blue.

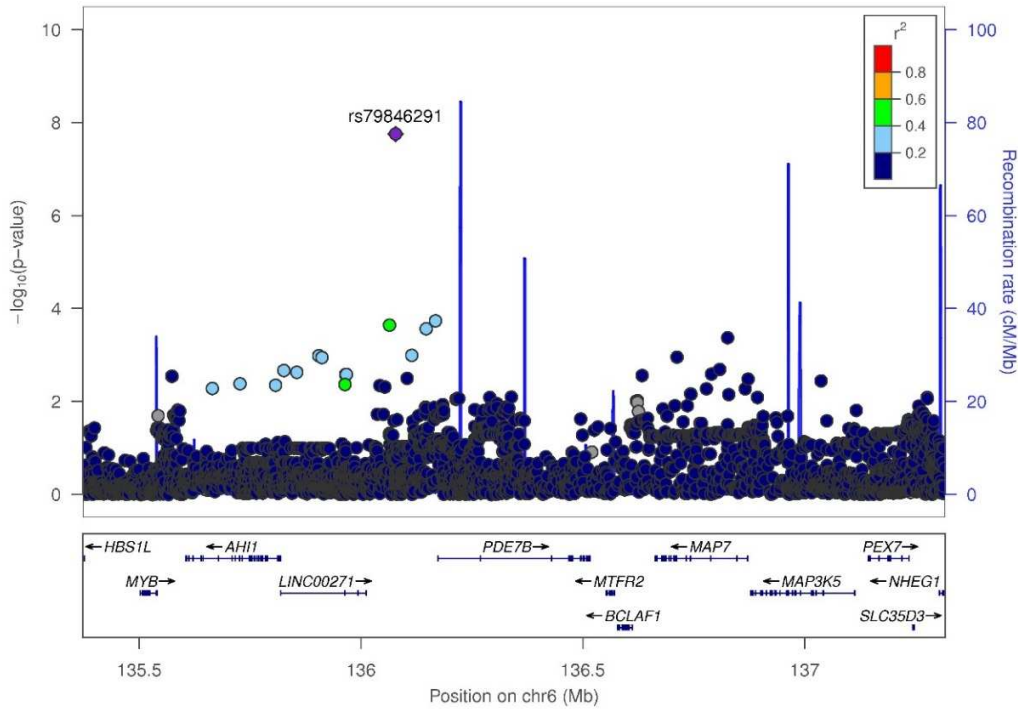


**Supplementary Figure 12.** Manhattan plot showing the meta-analyzed results of logical memory test (LMT)-delayed recall from all the subjects and three subgroups (CN: clinically normal and MCI: mild cognitive impairment). Blue dot line represents the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$ , and the red dot line indicates study-wide significant threshold of  $P < 8.33 \times 10^{-9}$ . Loci achieving study-wide significance are highlighted in red, and loci at genome-wide significance are in blue.

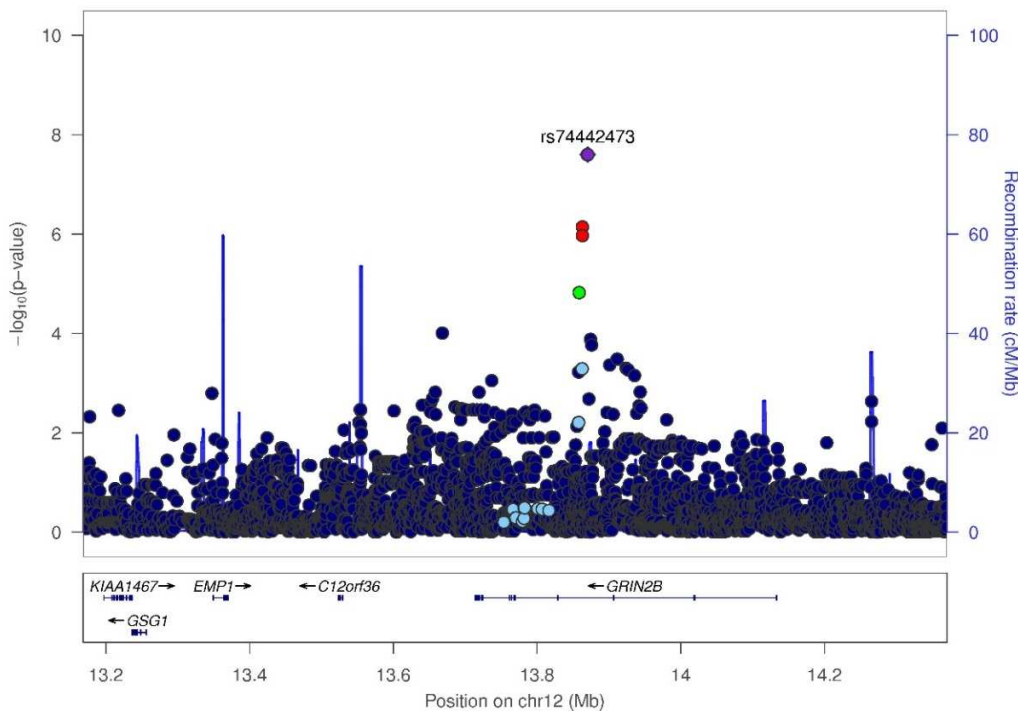


**Supplementary Figure 13.** Regional association plots of (A) *PDE7B* from association test with hippocampal volume in AD subjects, (B) *GRIN2B* from association test with CSF A $\beta_{42}$  in CN subjects, (C) *NRG1* from logical memory test immediate recall in MCI subjects, (D) *AKAP9* from CSF phosphorylated Tau in the MCI subjects, and (E) *BDNF* from CSF phosphorylated Tau using CN subjects

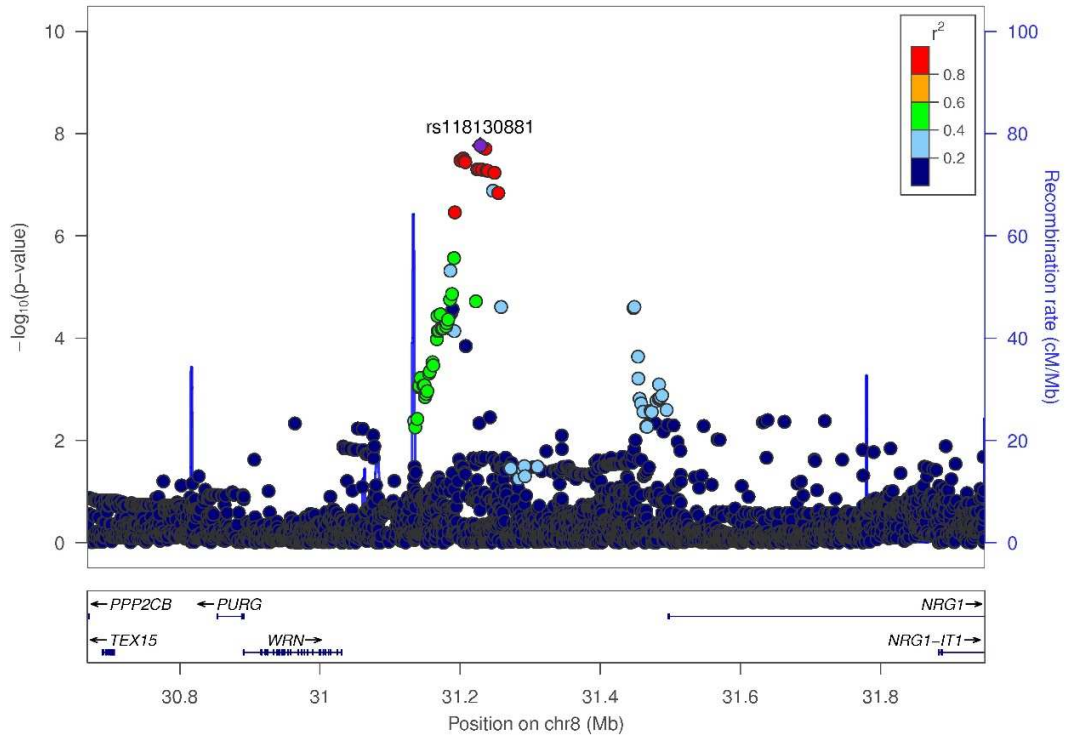
**A.**



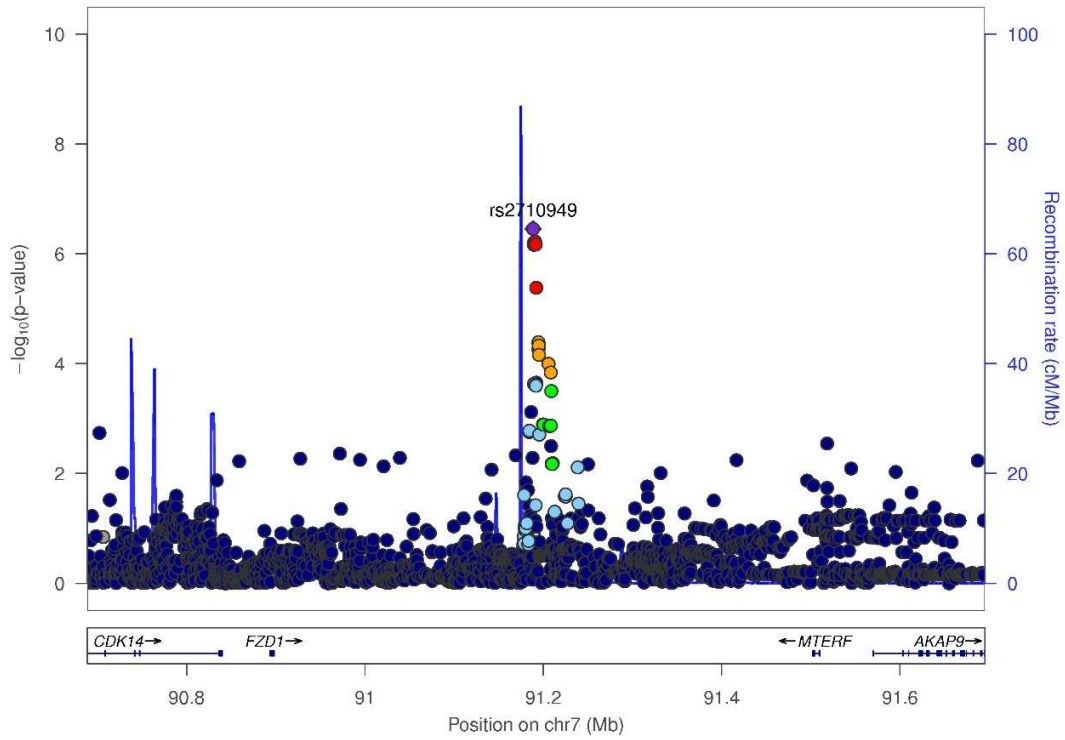
**B.**



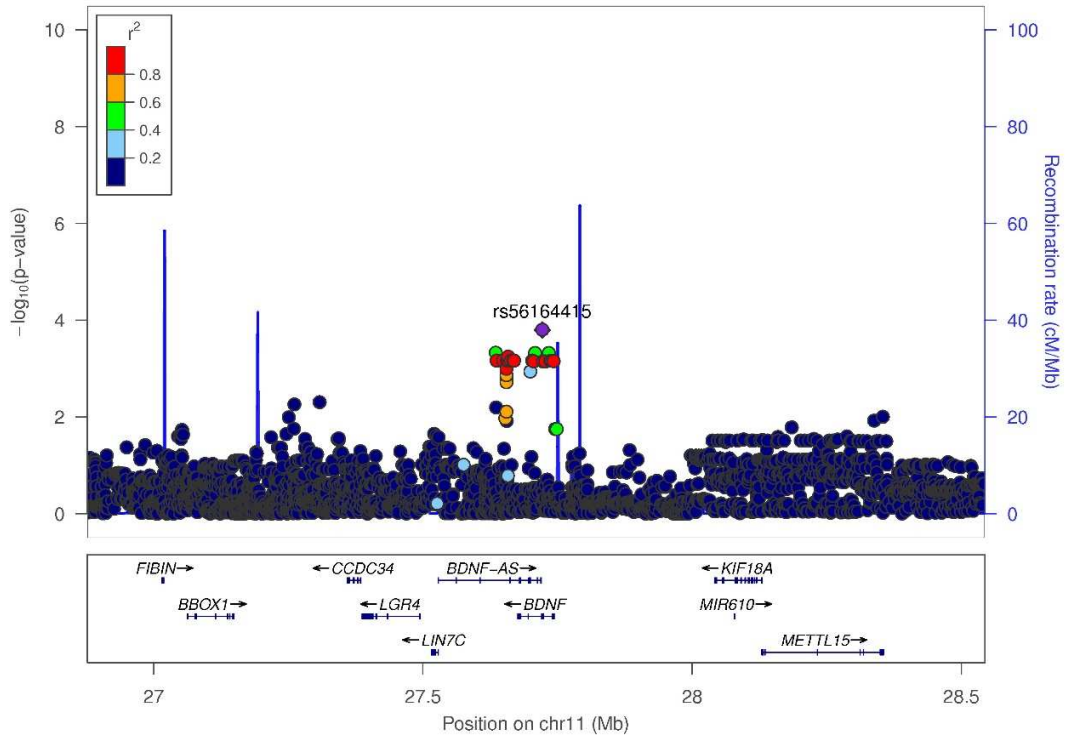
C.



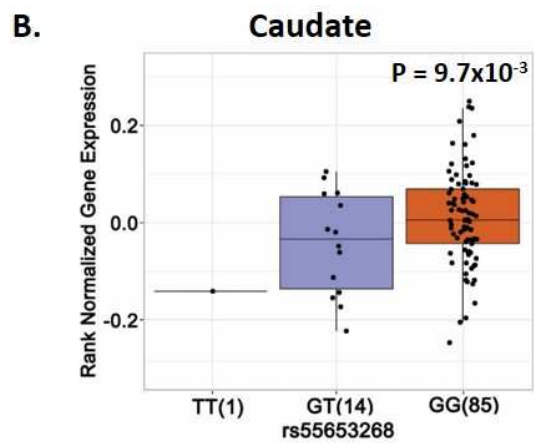
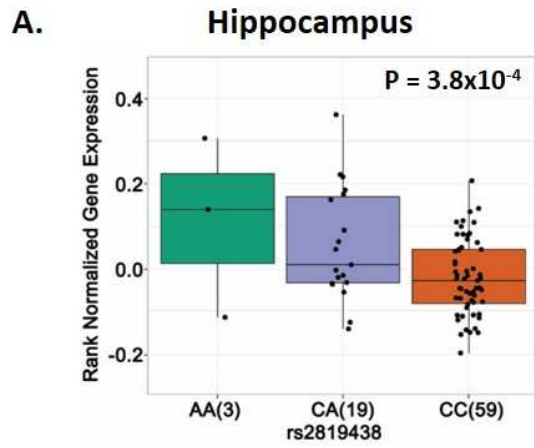
D.



E.

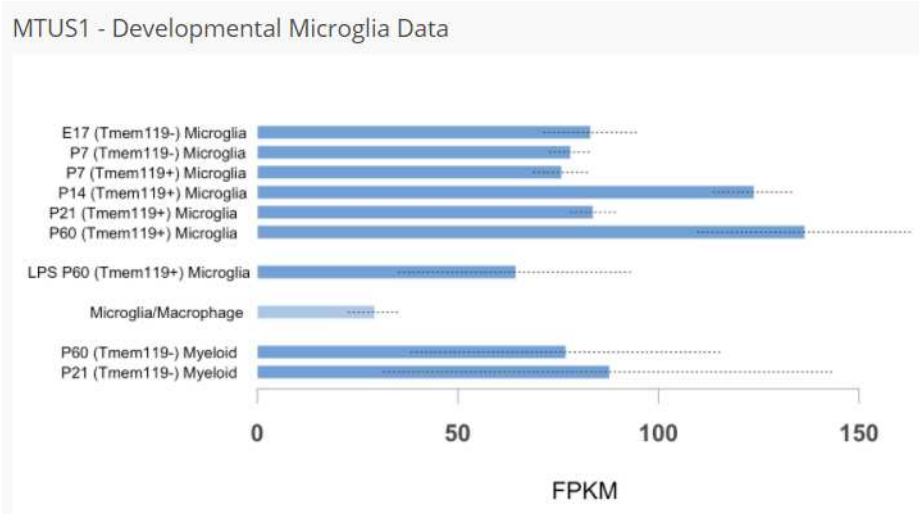


**Supplementary Figure 14.** Association between gene expression levels in brains from the GTEx Portal database and SNPs. **A.** *C14orf79* in hippocampus and rs2819438. **B.** *MTUS1* in caudate and rs55653268.

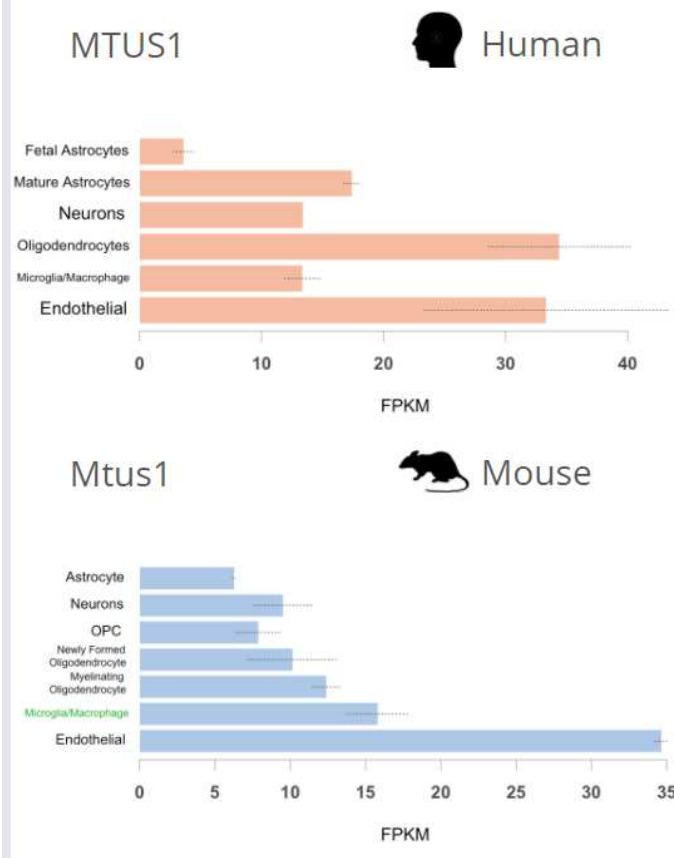


**Supplementary Figure 15.** Expression of *MTUS1* in mouse and human brain. (A) Expression profiles in murine myeloid cells. (B) Cell type specific expression in human and mouse. Data were obtained from: [http://web.stanford.edu/group/barres\\_lab/brainseqMariko/brainseq2.html](http://web.stanford.edu/group/barres_lab/brainseqMariko/brainseq2.html).

**A.**



**B.**





## References

- [1] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81.
- [2] Howie BN, Donnelly P, Marchini J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLoS Genet.* 2009;5:e1000529.
- [3] Delaneau O, Marchini J. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nature Communications.* 2014;5:3934.
- [4] Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006;38:904-9.
- [5] Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology.* 2009;10:R25.
- [6] Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011;12:323.
- [7] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics.* 2008;9:559.
- [8] Sokal RR, Michener CD. A statistical method for evaluating systematic relationships. *Univ. Kans. Sci. Bull.* 1958;28:1409–1438.
- [9] Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics.* 2008;24:719-20.
- [10] Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, et al. A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci.* 2015;112:7285–7290.

[11] Li MJ, Liu Z, Wang P, Wong MP, Nelson MR, Kocher JP, et al. GWASdb v2: an update database for human genetic variants identified by genome-wide association studies. *Nucleic Acids Res.* 2016;44:D869-76.

[12] Deming Y, Li Z, Kapoor M, Harari O, Del-Aguila JL, Black K, et al. Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. *Acta Neuropathol.* 2017;133:839-56.

[13] Stein JL, Hua X, Lee S, Ho AJ, Leow AD, Toga AW, et al. Voxelwise genome-wide association study (vGWAS). *Neuroimage.* 2010;53:1160-74.

[14] Stein JL, Hua X, Morra JH, Lee S, Hibar DP, Ho AJ, et al. Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. *Neuroimage.* 2010;51:542-54.

[15] Luciano M, Hansell NK, Lahti J, Davies G, Medland SE, Raikonen K, et al. Whole genome association scan for genetic polymorphisms influencing information processing speed. *Biol Psychol.* 2011;86:193-202.

[16] Shanker S, Li S, Sungeun K, Mark I, John DW, Kelley MF, et al. Analysis of Copy Number Variation in Alzheimer's Disease: The NIALOAD/ NCRAD Family Study. *Current Alzheimer Research.* 2012;9:801-14.

[17] Swaminathan S, Kim S, Shen L, Risacher SL, Foroud T, Pankratz N, et al. Genomic Copy Number Analysis in Alzheimer's Disease and Mild Cognitive Impairment: An ADNI Study. *International Journal of Alzheimer's Disease.* 2011;2011:10.

[18] Han MR, Schellenberg GD, Wang LS, Alzheimer's Disease Neuroimaging I. Genome-wide association reveals genetic effects on human Abeta42 and tau protein levels in cerebrospinal fluids: a case control study. *BMC Neurol.* 2010;10:90.

[19] Kramer PL, Xu H, Woltjer RL, Westaway SK, Clark D, Erten-Lyons D, et al. Alzheimer disease pathology in cognitively healthy elderly: a genome-wide study. *Neurobiol Aging*. 2011;32:2113-22.

[20] Logue MW, Schu M, Vardarajan BN, Farrell J, Bennett DA, Buxbaum JD, et al. Two rare AKAP9 variants are associated with Alzheimer's disease in African Americans. *Alzheimer's & Dementia*. 2014;10:609-18.e11.

[21] Go RC, Perry RT, Wiener H, Bassett SS, Blacker D, Devlin B, et al. Neuregulin-1 polymorphism in late onset Alzheimer's disease families with psychoses. *Am J Med Genet B Neuropsychiatr Genet*. 2005;139B:28-32.

Supplementary Table 1. Association P-values of Alzheimer disease loci previously established by GWAS with CSF biomarkers of Aβ42, total and phosphorylated tau, MRI hippocampal volume, and test scores for immediate and delayed recall in normal, mild cognitively impaired, and AD subjects.

CHR	BP	SNP	Gene	CSF Aβ42				CSF T-Tau				HPV				LMIT				LMdT																																																										
				RE		AV		RE		AV		RE		AV		RE		AV		RE		AV																																																								
				Beta	SE P	Beta	SE P	Beta	SE P	Beta	SE P	Beta	SE P	Beta	SE P	Beta	SE P	Beta	SE P	Beta	SE P	Beta	SE P																																																							
1	20707049	r5855201	CELI	-0.12	0.07	0.13	-0.06	0.15	0.07	-0.09	0.10	0.35	-0.01	0.13	0.97	0.08	0.08	0.30	0.10	0.13	0.84	0.04	0.10	0.09	0.04	0.15	0.77	0.07	0.07	0.83	0.14	0.14	0.93	0.05	0.10	0.63	-0.10	0.14	0.96	-0.02	0.06	0.74	0.12	0.10	0.25	0.09	0.08	0.31	0.02	0.12	0.89	-0.14	0.06	0.02	-0.20	0.08	0.01	0.01	0.06	0.99	0.06	0.10	0.58	-0.14	0.06	0.02	-0.13	0.07	0.97	0.01	0.06	0.90	0.05	0.08	0.50	0.05	0.08	0.50
2	127892810	r5733839	BIN1	-0.06	0.06	0.26	-0.11	0.10	0.26	-0.03	0.08	0.73	-0.06	0.12	0.60	0.02	0.06	0.74	-0.02	0.09	0.87	0.00	0.08	1.00	0.04	0.14	0.77	0.04	0.06	0.53	-0.02	0.09	0.89	0.01	0.08	0.89	-0.09	0.13	0.49	-0.02	0.05	0.64	0.03	0.07	0.70	-0.06	0.07	0.38	-0.11	0.10	0.26	-0.03	0.05	0.46	-0.04	0.05	0.44	-0.03	0.08	0.68	-0.01	0.05	0.80	-0.01	0.05	0.90	0.00	0.05	0.93	-0.05	0.06	0.43						



Table with columns for ID, Name, and a large grid of numerical data. The table contains multiple rows of data, likely representing a dataset or a list of items with associated values.



**Supplementary Table 7.** Correlation matrix of expressions for the genes which were suggestively associated with at least one or endophenotypes or were previously detected in GWAS of AD risk.

	ERBIN	MTUS1	BACE1.AS	ZCWPW1	BRIP1	CAMK2B	DAB1	GRIN2B	JPH3	NRG1	SRRM4	ZNF804B	BZRAP1	MAPT	MEF2C	PTK2B	APP	PSEN2	PLXNA4
<b>ERBIN</b>	1.00	0.83	0.53	0.52	-0.34	-0.60	-0.44	-0.68	-0.56	-0.52	-0.68	-0.44	-0.70	-0.50	-0.69	-0.55	-0.62	-0.45	-0.55
<b>MTUS1</b>	0.83	1.00	0.57	0.42	-0.42	-0.54	-0.39	-0.67	-0.50	-0.65	-0.71	-0.48	-0.61	-0.44	-0.73	-0.52	-0.63	-0.60	-0.51
<b>BACE1-AS</b>	0.53	0.57	1.00	0.56	-0.26	-0.58	-0.42	-0.69	-0.40	-0.69	-0.35	-0.21	-0.23	-0.45	-0.35	-0.67	-0.53	-0.49	-0.68
<b>ZCWPW1</b>	0.52	0.42	0.56	1.00	-0.28	-0.53	-0.50	-0.62	-0.38	-0.55	-0.39	-0.21	-0.30	-0.46	-0.47	-0.55	-0.47	-0.46	-0.57
<b>BRIP1</b>	-0.34	-0.42	-0.26	-0.28	1.00	0.33	0.20	0.44	0.09	0.35	0.47	0.47	0.25	0.12	0.60	0.34	0.55	0.36	0.36
<b>CAMK2B</b>	-0.60	-0.54	-0.58	-0.53	0.33	1.00	0.74	0.90	0.78	0.68	0.66	0.34	0.65	0.69	0.40	0.93	0.65	0.59	0.89
<b>DAB1</b>	-0.44	-0.39	-0.42	-0.50	0.20	0.74	1.00	0.81	0.73	0.57	0.52	0.29	0.57	0.50	0.32	0.79	0.60	0.49	0.78
<b>GRIN2B</b>	-0.68	-0.67	-0.69	-0.62	0.44	0.90	0.81	1.00	0.74	0.82	0.72	0.50	0.64	0.70	0.56	0.94	0.83	0.63	0.92
<b>JPH3</b>	-0.56	-0.50	-0.40	-0.38	0.09	0.78	0.73	0.74	1.00	0.55	0.72	0.35	0.75	0.71	0.41	0.75	0.52	0.43	0.66
<b>NRG1</b>	-0.52	-0.65	-0.69	-0.55	0.35	0.68	0.57	0.82	0.55	1.00	0.63	0.48	0.46	0.59	0.49	0.76	0.76	0.52	0.71
<b>SRRM4</b>	-0.68	-0.71	-0.35	-0.39	0.47	0.66	0.52	0.72	0.72	0.63	1.00	0.52	0.80	0.65	0.78	0.64	0.71	0.38	0.54
<b>ZNF804B</b>	-0.44	-0.48	-0.21	-0.21	0.47	0.34	0.29	0.50	0.35	0.48	0.52	1.00	0.38	0.46	0.57	0.30	0.58	0.46	0.30
<b>BZRAP1</b>	-0.70	-0.61	-0.23	-0.30	0.25	0.65	0.57	0.64	0.75	0.46	0.80	0.38	1.00	0.49	0.53	0.57	0.54	0.30	0.49
<b>MAPT</b>	-0.50	-0.44	-0.45	-0.46	0.12	0.69	0.50	0.70	0.71	0.59	0.65	0.46	0.49	1.00	0.46	0.63	0.64	0.49	0.60
<b>MEF2C</b>	-0.69	-0.73	-0.35	-0.47	0.60	0.40	0.32	0.56	0.41	0.49	0.78	0.57	0.53	0.46	1.00	0.39	0.63	0.47	0.41
<b>PTK2B</b>	-0.55	-0.52	-0.67	-0.55	0.34	0.93	0.79	0.94	0.75	0.76	0.64	0.30	0.57	0.63	0.39	1.00	0.67	0.50	0.94
<b>APP</b>	-0.62	-0.63	-0.53	-0.47	0.55	0.65	0.60	0.83	0.52	0.76	0.71	0.58	0.54	0.64	0.63	0.67	1.00	0.54	0.66
<b>PSEN2</b>	-0.45	-0.60	-0.49	-0.46	0.36	0.59	0.49	0.63	0.43	0.52	0.38	0.46	0.30	0.49	0.47	0.50	0.54	1.00	0.52
<b>PLXNA4</b>	-0.55	-0.51	-0.68	-0.57	0.36	0.89	0.78	0.92	0.66	0.71	0.54	0.30	0.49	0.60	0.41	0.94	0.66	0.52	1.00



Supplementary table 3.

CHR BP	SNP	Gene	Deming et al	EA RA	CSF Aβ <sub>42</sub>																CSF p-Tau																CSF T-Tau																	
					ALL				LCN				MCI				AD				ALL				LCN				MCI				AD				ALL				LCN				MCI				AD					
					FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P	FREQ	BETA	SE	P						
1	53908219	rs185051519	GLIS1	G	C	0.076	-0.32	0.12	#####	0.08	-0.34	0.2	#####	0.089	-0.26	0.16	0.12	0.091	-0.13	0.23	0.56	0.075	-0.03	0.12	0.82	0.08	0.03	0.2	0.89	0.089	-0.16	0.17	0.33	0.091	0.23	0.25	0.36	0.076	0.01	0.12	0.95	0.08	0.09	0.19	0.64	0.07	-0.1	0.16	0.56	0.092	0.05	0.26	0.84	
6	2838248	rs316341	SERPIN1	G	A	0.29	-0.11	0.06	0.06	0.288	-0.14	0.11	0.23	0.292	-0.12	0.08	0.12	0.295	-0.08	0.12	0.52	0.291	-0.03	0.06	0.65	0.291	-0.02	0.11	0.86	0.292	-0.06	0.08	0.46	0.295	0.1	0.13	0.46	0.291	0.06	0.06	0.33	0.289	0.11	0.1	0.27	0.292	0.03	0.08	0.69	0.297	0.13	0.14	0.33	
13	53504675	rs9527039	PCDH8	p-Tau	C	T	0.071	-0.03	0.11	0.77	0.075	-0.38	0.18	#####	0.069	0.05	0.15	0.74	0.068	0.27	0.22	0.23	0.071	-0.14	0.11	0.13	0.075	-0.06	0.17	0.73	0.069	-0.23	0.15	0.14	0.068	0.02	0.24	0.98	0.071	-0.1	0.11	0.37	0.074	-0.08	0.17	0.64	0.071	-0.13	0.15	0.38	0.069	0.07	0.25	0.78
18	77381649	rs12961169	CTDP1	p-Tau	T	C	0.189	-0.11	0.07	0.14	0.197	-0.12	0.13	0.37	0.187	-0.1	0.1	0.31	0.185	-0.04	0.14	0.79	0.188	0.22	0.07	#####	0.192	0.14	0.13	0.26	0.188	0.23	0.11	#####	0.185	0.13	0.16	0.46	0.187	0.11	0.07	0.15	0.19	0.04	0.12	0.71	0.188	0.13	0.1	0.18	0.183	-0.03	0.16	0.84
3	190663557	rs35055419	GMNC	p-Tau	C	T	0.361	-0.1	0.06	0.08	0.31	0.13	0.11	0.24	0.38	-0.1	0.08	0.2	0.389	-0.18	0.12	0.13	0.361	0.21	0.06	#####	0.309	0.3	0.1	#####	0.381	0.07	0.08	0.37	0.389	0.26	0.13	0.16	0.36	0.27	0.06	#####	0.307	0.31	0.1	#####	0.379	0.14	0.08	0.08	0.389	0.28	0.13	0.08
9	3929424	rs1514716	GLIS3	p-Tau	C	T	0.123	0.03	0.08	0.68	0.149	0.02	0.15	0.89	0.11	0.01	0.11	0.91	0.13	-0.05	0.17	0.78	0.123	-0.24	0.08	#####	0.146	-0.18	0.14	0.2	0.111	-0.17	0.11	0.13	0.13	-0.38	0.18	#####	0.122	-0.19	0.08	#####	0.15	-0.16	0.13	0.24	0.11	-0.08	0.11	0.47	0.126	-0.32	0.19	0.1
19	45410002	rs769449	APUE	All three biomarkers	A	G	0.208	-0.73	0.06	#####	0.098	-0.81	0.15	#####	0.251	-0.66	0.08	#####	0.515	-0.52	0.11	#####	0.209	0.47	0.06	#####	0.099	0.16	0.15	0.29	0.235	0.5	0.08	#####	0.515	0.2	0.13	0.18	0.209	0.48	0.06	#####	0.1	0.15	0.14	0.28	0.235	0.48	0.08	#####	0.509	0.24	0.14	0.1