2020

Tissue-dependent analysis of common and rare genetic variants for Alzheimer's disease using multi-omics data

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

AND

COLLEGE OF ENGINEERING

Dissertation

TISSUE-DEPENDENT ANALYSIS OF COMMON AND RARE GENETIC VARIANTS FOR ALZHEIMER'S DISEASE USING MULTI-OMICS DATA

by

DEVANSHI PATEL

B.S., Tufts University, 2012
M.S., Boston University, 2014

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2020

Approved by

First Reader

Lindsay A. Farrer, Ph.D.
Distinguished Professor of Genetics
Professor of Medicine
Professor of Neurology
Professor of Ophthalmology
Boston University, School of Medicine

Professor of Biostatistics
Professor of Epidemiology
Boston University, School of Public Health

Second Reader

Xiaoling Zhang, MD, PhD
Assistant Professor of Biostatistics
Boston University, School of Public Health

ACKNOWLEDGMENTS

This dissertation and doctoral work would not be possible without: my advisor Dr. Lindsay Farrer and his immense knowledge and patience throughout my PhD; my supervisor Dr. Xiaoling Zhang and her guidance and great ideas during these years; my committee- Dr. Paola Sebastiani, Dr. Honghuang Lin, and Dr. Kathy Lunetta — and all their support and valuable suggestions in getting to this point; the faculty and members of the Farrer Lab, and the advice and help of Dr. John Farrell and Dr. Jaeyoon Chung; the Research Computing Services (RCS) group, specifically Katia Bulekova; the Bioinformatics program and the continued administrative support of Dave King through my masters and PhD; and most importantly my family and friends, and their unconditional love and unwavering support through all these years.

TISSUE-DEPENDENT ANALYSIS OF COMMON AND RARE GENETIC VARIANTS FOR ALZHEIMER'S DISEASE USING MULTI-OMICS DATA

Boston University, Graduate School of Arts and Sciences

and

College of Engineering, 2020

Major Professor: Lindsay A. Farrer, Ph.D., Distinguished Professor of Genetics, Professor of Medicine, Professor of Neurology, Professor of Ophthalmology, Boston University, School of Medicine; Professor of Biostatistics, Professor of Epidemiology, Boston University, School of Public Health

ABSTRACT

Alzheimer's disease (AD) is a complex neurodegenerative disease characterized by progressive memory loss and caused by a combination of genetic, environmental, and lifestyle factors. AD susceptibility is highly heritable at 58-79%, but only about one third of the AD genetic component is accounted for by common variants discovered through genome-wide association studies (GWAS). Rare variants may contribute to some of the unexplained heritability of AD and have been demonstrated to contribute to large gene expression changes across tissues, but conventional analytical approaches pose challenges because of low statistical power even for large sample sizes. Recent studies have demonstrated by expression quantitative trait locus (eQTL) analysis that changes in gene expression could play a key role in the pathogenesis of AD. However, regulation of gene expression has been shown to be context-specific (e.g., tissue and cell-types), motivating a context dependent approach to achieve more

precise and statistically significant associations. To address these issues, I applied a strategy to identify new AD risk or protective rare variants by examining mutations occurring only in cases or only controls, observing that different mutations in the same gene or variable dose of a mutation may result in distinct dementias. I also evaluated the impact of rare variation on expression at the gene and gene pathway levels in blood and brain tissue, further strengthening the rare variant findings with functional evidence and finding evidence for a large immune and inflammatory component to AD. Lastly, I identified cell-type specific eQTLs in blood and brain tissue to explain underlying genetic associations of common variants in AD, and also discovered additional evidence for the role of myeloid cells in AD risk and potential novel blood and brain AD biomarkers. Collectively, these findings further explain the genetic basis of AD risk and provide insight about mechanisms leading to this disorder.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD = Alzheimer's disease

ADGC = Alzheimer's disease Genetic Consortium

ADNI = Alzheimer Disease Neuroimaging Initiative

ADSP = Alzheimer's Disease Sequencing Project

CADASIL = Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy

CADD = Combined Annotation Dependent Depletion

ct-eQTL = Cell-type specific eQTL

EA = European ancestry

EOAD = Early-onset autosomal dominant AD

eQTL = Expression quantitative trait locus

FHS = Framingham Heart Study

FTD = Frontotemporal dementia

GWAS = Genome-wide association studies

HLA = Human leukocyte antigen

Indels = Insertion-deletion polymorphisms

LOAD = Late-onset AD

MAC = Minor allele counts

MAF = Minor allele frequency

MCI = Mild Cognitive Impairment

MHC = Major histocompatibility

NHD = Nasu-Hakola disease

PMI = Post-mortem interval

QC = Quality control

ROSMAP = Religious Orders Study/Memory & Aging Project

SNPs = Single-nucleotide polymorphisms

SNVs = Single nucleotide variants

SVA = Surrogate variable analysis

UPDB = Utah Population Database

VEP = Variant Effect Predictor

WES = Whole exome sequencing

WGCNA = Weighted Gene Co-expression Network Analysis

WGS = Whole genome sequencing

CHAPTER 1: Introduction

Alzheimer's disease (AD) is a complex neurodegenerative disease characterized by progressive memory loss with no current effective prevention or treatment [1]. It is the most common type of dementia that affects an estimated 5.7 million individuals just in the United States, with the number projected to rise to 14 million by 2050 [1] as well as the 6th leading cause of death in the United States [1]. Abnormal deposits of proteins form amyloid plaques and tau tangles throughout the brain leading to brain cell death [1]. Brain damages seems to begin in the hippocampus (which controls memory) and spreads throughout the brain as more neurons die, resulting in significantly shrunken brain tissue [1].

AD is caused by a combination of genetic, environmental, and lifestyle factors [2]. AD susceptibility is highly heritable at 58-79% [2]. Early-onset autosomal dominant AD (EOAD) is caused by highly penetrant variants in a few known risk genes, whereas late-onset AD (LOAD), which is diagnosed in patients over 65, is caused by many different low penetrance mutations [3]. About 30 LOAD susceptibility loci are known, and risk is very polygenic [3]. Only about one third of the AD genetic component is accounted for by common variants discovered through genome-wide association studies (GWAS), with a sizable portion of the genetic risk unexplained [2].

Rare variants may contribute to some of the unexplained heritability of AD and have been demonstrated to contribute to large gene expression changes across tissues [4]. Genome-wide searches have identified robust AD

associations with rare variants in relatively few genes and methods to evaluate rare variants are still under development [3]. **Chapter 2** will focus on a strategy to identify new AD risk or protective rare variants by looking at mutations in only cases or only controls.

Genetic variation in AD risk is mostly studied through association analysis of single-nucleotide polymorphisms (SNPs). AD risk may also be associated with SNPs influencing changes in gene expression. Recent studies have demonstrated by expression quantitative trait locus (eQTL) analysis that changes in gene expression could play a key role in the pathogenesis of AD [5, 6]. **Chapter 3** will show the impact of rare variation on gene expression in both blood and brain tissue, also further strengthening the rare variant findings from Chapter 2 with functional evidence.

However, regulation of gene expression, as well as many biological functions, has been shown to be context-specific (e.g., tissue and cell-types, developmental time point, sex, disease status, and response to treatment or stimulus) [7-10]. These findings motivate a context dependent analysis to achieve more precise associations that may help elucidate regulatory networks. **Chapter 4** will identify cell-type specific eQTLs in blood and brain tissue to explain underlying genetic associations of common variants in AD. One of the largest studies of this kind classified 12% of over 23,000 eQTLs as cell-type specific [9] and much of the protocol used in this study will be followed.

Chapters 3 and 4 compare gene expression in blood and brain tissue from

AD cases and cognitively and pathologically normal controls, noting that brain studies have several limitations. Specifically, large samples of brain specimens are difficult to obtain. There is also concern that brains undergo changes post-mortem that may affect gene expression [11]. Results of previous brain eQTL studies in AD are inconsistent; two studies that examined expression in prefrontal cortex showed less than 10% of overlapping eQTLs [5]. Also, although neurons are arguably the most relevant cell type for AD research, several studies demonstrated that peripheral monocytes play crucial roles in AD pathogenesis [12, 13]. Most eQTL and genetic studies to date have considered only a single tissue, typically blood [5].

The overall aim of this dissertation is to further explain undiscovered AD genetic risk with rare and common variants and increase our understanding of the genetic architecture of AD though a series of analyses of multi-omics data. Novel findings can be identified by integrating multiple types of ‘omics data. In addition, the proposed multi-tissue analysis will likely yield more consistent findings and identify tissue-specific differences. Tissue-dependent and cell-type dependent association patterns will improve our understanding about pathophysiology underlying AD risk and provide insight about strategies to slow progression and for prevention. Finally, a reliable AD-associated biomarker found in both brain and blood would be significant for predicting future development of AD.

CHAPTER 2: Identification of novel AD risk variants and genes in Alzheimer Disease Sequencing Project (ADSP) data

2.1 Abstract

Background: Much of the unexplained heritability of Alzheimer disease (AD) may be due to rare variants whose effects are not captured in GWAS because very large samples are needed to observe statistically significant associations.

Methods: Rare variants were identified by whole exome sequencing (WES) in unrelated samples, including 5,396 AD cases and 4,415 controls of European ancestry (EA), from the Alzheimer's Disease Sequencing Project (ADSP). Minor alleles genome-wide and in 95 genes previously associated with AD, AD-related traits, or other dementias were tabulated and filtered for predicted functional impact and occurrence in AD cases, but not controls. Top findings were further evaluated using multiple analytical approaches. Support for several findings was sought in a WES dataset comprised of 19 affected relative pairs from Utah high-risk pedigrees and several whole genome sequencing (WGS) datasets from the ADSP and Alzheimer Disease Neuroimaging Initiative (ADNI).

Results: A total of 24 variants with moderate or high functional impact from 19 genes were observed in 10 or more AD cases but not in controls. These include a missense mutation (rs149307620, n=10) in *NOTCH3*, coding mutations in which are associated with CADASIL, a diagnostically distinct disorder marked by severe headaches in young adulthood followed by stroke and dementia later in life. Evaluation of clinical and autopsy data from these AD cases revealed classic

AD symptoms with progressive memory loss, moderate to severe amyloid and tau pathology at autopsy and limited evidence of stroke or microvascular disease. The rare rs149307620 was identified in one AD case and one subject with mild cognitive impairment in the WGS datasets. Two additional rare *NOTCH3* missense mutations that may be disease-causal were found in one Utah family. The *TREM2* Q33X high-impact mutation was identified in four AD cases that in homozygous form causes Nasu-Hakola disease, a very rare disorder characterized by early-onset dementia and multifocal bone cysts, suggesting an intermediate inheritance model for the mutation with respect to dementia outcome. In addition, AD cases have a significantly higher burden of deleterious rare coding variants in known AD, AD-related or other dementia genes compared to controls.

Conclusions: We identified several novel rare variants predicted to have high impact on protein structure which may alter AD risk and be potential therapeutic targets.

2.2 Background

AD susceptibility is highly heritable (h^2=58-79%) [2], but only about one third of the genetic component is accounted for by common variants discovered through genome-wide association studies (GWAS) [2]. Some of the unexplained heritability of AD may be due to rare variants, which remain challenging to discover in genomic studies because of statistical power limitations, despite large

sample sizes [3]. Genome-wide searches have identified robust AD associations with rare variants in relatively few genes including *TREM2*, *AKAP9*, *UNC5C*, *ZNF655, IGHG3*, and *CASP7* [14-18], and methods to evaluate rare variants are still under development [3]. We applied a strategy focused on rare variants occurring only in cases to identify and characterize additional high penetrance risk variants in AD that would be otherwise undetected in analyses that do not render results when a variant is not observed in the control group.

2.3 Methods

2.3.1 *Study Population and Data Processing*

The Alzheimer's Disease Sequencing Project (ADSP) performed whole-exome sequencing (WES) on DNA samples obtained from non-Hispanic subjects of European ancestry (EA) and Caribbean Hispanics (CH). All subjects were at least 60 years old. AD cases met NINCDS-ADRDA criteria for possible, probable or definite AD after clinical and/or neuropathologic examination [19] and controls were cognitively normal and showed no signs of AD pathology if examined at autopsy. ADSP participants were selected using a risk score based on age, sex, and *APOE* ε4 carrier status to maximize cases most likely to have AD risk variants and controls most likely to have AD protective variants. Subject characteristics are provided in Table S2.1. DNA sequencing was performed at the Broad Institute, the Baylor College of Medicine's Human Genome Sequencing Center, and Washington University's McDonnell Genome Institute.

Genotypes for bi-allelic single nucleotide variants (SNVs) and short insertion-deletion polymorphisms (indels) were called using ATLAS2 [20]. Only variants that overlapped the target regions captured by kits used by the three sequencing centers (Illumina and Nimblegen) were included. Variants that showed extreme departure from Hardy-Weinberg equilibrium (HWE, $P<10^{-6}$) among unrelated controls, were monomorphic, or had call rates <80% or average read depth <10 were excluded. In addition, samples that were outliers according to population substructure analysis or had a genetically determined sex that was inconsistent with the reported sex were excluded [17]. Cryptic relatedness was estimated using pairwise identity-by-descent (IBD) in PLINK [21] and one member from each of 69 pairs of individuals with $\hat{\pi}$ (proportion of alleles shared IBD)>0.4 was excluded [17]. After quality control (QC), there remained for analysis a total of 10,211 EA (5,617 AD cases, 4,594 controls) and 400 CH (221 cases, 179 controls) subjects. The CH sample was deemed too small for in-depth studies of rare variants and thus excluded from the analysis.

2.3.2 *Variant Selection and Annotation*

Minor allele counts (MAC) genome-wide and in a group of 95 genes previously associated with AD, AD-related traits, or dementia by genetic association or experimental approaches (Table S2.2) were tabulated for rare variants occurring only in AD cases. Variants were annotated with the Variant Effect Predictor tool (VEP) [22] and Combined Annotation Dependent Depletion (CADD) scores [23] as having high functional impact (includes splice acceptor, splice donor, stop

gained, frameshift, stop lost, start lost, or transcript amplification variants) or moderate functional impact (includes in-frame insertion, in-frame deletion, missense, or protein altering variants). A scaled CADD score of 20, for example, means that a variant is among the top 1% of deleterious variants in the human genome, and a scaled CADD score of 30 means that the variant is in the top 0.1%. Synonymous mutations and variants with a MAC ≤ 2 were excluded. The study design is illustrated in Figure S2.1 and the filtering of variants at each step in the pipeline is shown in Table S2.3.

2.3.3 *Haplotype Analysis*

PLINK was used to find common SNPs near the rare variant of interest within a specified kilobase (kb) window. The wildcard option was used to infer haplotypes and estimate haplotype frequencies [21]. Haploview was used to visualize regional LD and confirm haplotypes and frequencies among different SNP combinations using multi-marker haplotype association tests [24].

2.3.4 *Gene-set Burden Test*

A gene-set burden test was performed using the Combined and Multivariate Collapsing (CMC) method [25] in R to compare the burden of high and moderate impact mutations occurring in only cases for known AD/dementia genes between AD cases and controls. Logistic regression models included covariates for sex, age, sequencing center, and principal components (PCs) of ancestry.

2.3.5 *Protein Homology Modeling and Pathway Analysis*

Protein homology modeling was performed for several high-impact variants in *NOTCH3* with BLAST-P [26], SWISS-MODEL [27] and Maestro software [28]. Since the *NOTCH3* structure could not be found in any protein databases, homologous proteins were considered for modeling the region containing the *NOTCH3* mutations. Considering the AD-associated mutation on exon 6 (rs149307620), the closest matching sequence was identified in *NOTCH1* which had 44% homology of amino acid sequence identity suggesting that these proteins may have similar functions. SWISS-MODEL was used to create a model for the *NOTCH3* region using homologous Protein Data Bank (PDB) structure 5UK5 in *NOTCH1* [29]. The same procedure was applied for other mutations in *NOTCH3* that are associated with Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL). A high-confidence (confidence score >0.7) human protein-protein interaction network was then created with the latest version (v10) of the STRING database for *NOTCH3* and its ligand *JAG1* [30]. The set of genes forming the protein network was tested for gene-set enrichment using PANTHER pathways using the Fisher's Exact test with FDR multiple test correction [31].

2.3.6 *Rare Variant Analysis in an Independent Dataset*

To extend and enhance discovery of novel associations, we evaluated WES data obtained from 19 AD-affected first- or second-cousin pairs identified in the Utah Population Database (UPDB) belonging to a pedigree with a statistical excess of

AD risk. Genealogical records indicate that these pairs do not share any common ancestors as far back as the early 1800s. Details of the UPDB, case classification, and identification of high-risk pedigrees are published elsewhere [32]. We retained 564 rare variants that were shared between at least one cousin pair and had a minor allele frequency (MAF) < 0.1% in one of several public whole exome and whole genome sequencing databases including 1000 Genomes [33], Exome Aggregation Consortium (ExAC) [34], the Genome Aggregation Database (gnomAD) [34], and NHLBI GO Exome Sequencing Project (ESP) [35] resulting in 400 candidate variants. Variants were then assessed for pathogenicity based on (1) American College of Medical Genetics and Genomics criteria [36] for “pathogenic” or “likely pathogenic”, (2) association with loss or gain of gene function, and (3) prediction of haploinsufficiency. The genes containing the remaining 389 variants were screened for relevance to AD or β-amyloid pathology, after which 118 variants remained. Twelve variants were added to the group of 118 variants for further evaluation that were (1) reported in the literature as an AD risk variant (n=2), (2) had a p value < 0.09 for association of increased risk of AD in the ADGC GWAS [37] (n=2), (3) absent in whole genome sequences of 1,354 participants of the Wellderly cohort who are >80 years old with no chronic diseases and who are not taking chronic medications [38], suggesting that the variant is never seen in healthy elderly population [n=3], or (4) observed in multiple cousin pairs (n=5).

2.4 Results

2.4.1 *Rare Variants in NOTCH3*

Evaluation of high and moderate impact rare variants in genes that were previously established as genetically or functionally related to AD or dementia revealed a missense mutation in *NOTCH3* (rs149307620; A284T) that was present in 10 AD cases, but no controls (Table 2.1). This variant is extremely rare in EAs (MAF= 0.0005) [34] and was verified by Sanger sequencing in eight of these subjects for whom DNA was available. Because several other high or moderate impact *NOTCH3* mutations have been associated with CADASIL, a diagnostically distinct disorder marked by severe headaches in young adulthood followed by strokes and dementia later in life [39], we sought clinical and autopsy data from the AD cases with the rs149307620 mutation to determine whether they are enriched for cerebrovascular risk factors. Neuropathological information that was available for one of these subjects revealed moderate atherosclerosis but no arteriosclerosis, lacunes or microinfarcts, which are hallmarks of CADASIL (Table S2.4). Autopsy also confirmed the presence of AD pathology (CERAD neuritic plaques: moderate; CERAD diffuse plaques: moderate; Braak neurofibrillary degeneration: stage VI). The mean age of symptom onset for the 10 *NOTCH3* mutation carriers was approximately 80 years, similar to that for the entire sample of ADSP EA cases (mean=76.4 years) and much greater than age at onset of cognitive impairment among individuals with CADASIL (usually <50 years). None of the *NOTCH3* mutation carriers had a history of clinical strokes

(though one had multiple infarcts on MRI and a history of diabetes and cardiovascular disease) and all had prominent memory impairment as the initial presentation with a progressive course. One other *NOTCH3* mutation (rs114447350) was observed in four AD cases but not in controls (Table 2.1). Unlike rs149307620, this variant is not particularly rare in EAs (MAF=0.024) or in persons of African ancestry (MAF=0.091) [33], suggesting it is unlikely to be pathogenic.

Because rs149307620 is very rare, we investigated the possibility that this mutation occurred once or only a few times by performing a haplotype analysis with common SNPs. This analysis revealed a 5-SNP haplotype with a frequency of 15% in the AD cases and 14% in controls that is common to all 10 cases with the *NOTCH3* mutation (Figure S2.2). Although the degree of IBD sharing across the genome is higher among the 10 cases ($\hat{\pi}$= 0.028) than among the entire ADSP sample ($\hat{\pi}$= 0.013), this small difference indicates that they are not more closely related to each other than to all subjects. Taken together, these results suggest that the mutation in these subjects originated in a common ancestor who lived many generations ago.

To investigate the possibility that the mutation carriers belong to a particular subpopulation, we plotted the first 2 principal components of ancestry that were derived previously for the entire sample [17] and observed that 8 of the 10 mutation carriers were clustered in a distinct minor portion of the sample (Figure

S2.3). Analysis of mitochondrial DNA variants revealed that most individuals in this cluster had mitochondrial haplogroups K1a1b1a or K1a9 that are common among Ashkenazi Jewish individuals [40]. Moreover, *NOTCH3* mutations carriers accounted for 4.0% of the participants with AD who have either the K1a1b1a or K1a9 haplogroup. The proportion of mutation carriers in this cluster was significantly greater among participants with AD (8 of 358 [2.2%]) than controls (0 of 337) ($Z=2.76$, $P=.006$). The frequency of the rs149307620 mutation is about 25 times higher in Ashkenazi Jewish individuals (MAF, 0.0046) compared with other EA groups (MAF, 0.00019) [34].

Analysis of the 130 variants that met the filtering criteria in the 19 affected cousin pairs from the Utah high-risk pedigrees provided additional support for a role of *NOTCH3* in AD. Both affected subjects in one family who are half-first cousins had two rare *NOTCH3* missense mutations, rs141402160 and rs140914494, each with a frequency of 0.0002 [33]. Review of available clinical and family history information for this family did not indicate findings consistent with CADASIL in the probands or relatives. The pedigree of the carriers of the rs141402160 and rs140914494 mutations includes seven additional members who had AD or dementia listed on death certificates. All but one of the individuals in the line of descent from the common ancestor of the probands died before age 60 and thus may have developed AD type dementia had they lived longer. Two affected cousins in another family had the *NOTCH3* missense variant

rs112197217 that has a frequency of 0.010 in EAs [33], but is much rarer in other populations. The evidence of AD among other relatives in this family was inconclusive.

To further distinguish *NOTCH3* variants that may be causally related to AD, we screened for these variants in whole genome sequence (WGS) data obtained from a multi-ethnic sample of unrelated 1,432 AD cases and 1,660 controls in the ADSP Extension Study, 550 AD cases and 283 controls in the ADSP multi-ethnic Family Study [41] and 809 participants of the ADNI Study (239 AD cases, 321 MCI cases, 249 controls) [42] . Characteristics of subjects in these datasets are provided in Table S2.5. The minor alleles for rs11219217 (n=70) and rs114447350 (n=286) were observed appreciably in AD cases, MCI cases and controls of multiple ethnicities suggesting that they are not associated with AD risk. The rs149307620 variant was found in one AD case (age at onset = 89 years) and one MCI case (age at onset = 76 years), but not in controls. The rare rs141402160 and rs140914494 variants were not detected in any of the WGS samples.

Protein modeling showed that the rs149307620 mutation is located in the EGF repeat region between EGF10 and EGF11 and more precisely in the EGF calcium binding (EGF_CA) domain, near the Jagged-1 (JAG1) - NOTCH3 binding site [29]. Modeling predicted that the major allele for rs149307620 results in wild type notch-3 with a corresponding amino acid alanine (Figure 2.1A).

Alanine is non-polar and would not be predicted to have any intra or inter-protein interactions. The minor allele for rs149307620 results in mutant notch-3 with a corresponding amino acid threonine (Figure 2.1B). Threonine is polar and will form hydrogen bonds where possible with itself and with polar arginine at the site of JAG1-NOTCH3 binding (Figure 2.1C). These results suggest that the mutant Notch-3 causes greater interaction with the ligand, possibly changing downstream processes. The other AD-associated NOTCH3 mutations, rs141402160 and rs140914494 also involve either the gain or loss of an alanine. In both instances, the mutation change leads to increased polarity and hydrogen bonding with possible increased interactions, to a greater or lesser extent that observed with rs149307620.

Unlike the AD-associated rs149307620 and rs141402160 mutations, but similar to the AD-associated rs140914494 mutation, most of the >25 reported distinct *NOTCH3* mutations causing CADASIL are located in exons 3 and 4 (Table 2.2) [43]. However, one CADASIL-associated variant, rs137852641, is a missense mutation in codon 332 in exon 6, resulting in the replacement of an arginine residue with a cysteine [44] that is proximate to rs149307620 (codon 284 in exon 6) (Figure 2.2).

To further explore the biological functions and pathways for *NOTCH3* in AD, a high confidence protein-protein interaction network was constructed including *NOTCH3* and *JAG1* (Figure 2.3). The resulting 30-gene interaction network

contains several AD related genes including *BACE1, PSEN1, PSEN2,* and *APP* (Figure 2.3). Gene-set enrichment analysis revealed the network genes were significantly enriched in the Notch signaling pathway ($P = 6.48\times10^{-49}$), angiogenesis ($P = 1.61\times10^{-12}$), and two AD-related pathways involving secretase mediated amyloid precursor protein cleavage ($P = 3.50\times10^{-16}$) and presenilin γ-secretase complex ($P = 5.78\times10^{-26}$) (Table 2.3).

2.4.2 *TREM2 Q33X*

We also identified the high impact *TREM2* Q33X mutation (rs104894002) in four out of 5,617 (0.071%) AD cases (Table 2.1), a frequency that is slightly lower than that observed in a *TREM2* sequencing study of AD cases and controls in 2013 (2/1,084 = 0.17% of AD cases) [14]. Because this mutation in homozygous state causes Nasu-Hakola disease, a rare autosomal recessive disorder characterized by early-onset dementia and multifocal bone cysts [45], we evaluated clinical data obtained from the four *TREM2* Q33X mutation carriers to assess potential pleiotropic effects (Table S2.6). All of these participants met the criteria for probable AD and none had reported bone cysts or unusual behavioral symptoms.

2.4.3 *Other Rare Mutations*

A total of 32 moderate or high impact variants in 24 previously established genes for AD or other dementias were each observed in four or more AD cases and no controls (Table 2.1). Five of these variants were previously reported as associated with AD and include missense mutations in *PSEN1* (rs63749824, n=7

[46, 47]; rs63750592, n=4 [48, 49]), *SORL1* (rs139710266, n=5) [50], and *MAPT* (rs63750424, n=4) [51-53] and a stop-gain mutation in *ABCA7* (E1769X, n=4) [54]. A previously unreported mutation observed only in AD cases in *PSEN1* (rs375376095, n=5) is only 2,779 bp from the known AD-associated rs63749824 variant. Genome-wide, 24 variants in 19 genes with moderate/high functional impact were observed in 10 or more AD cases but absent in controls (Table 2.4). Further examination of the genes represented by multiple variants revealed that 10 subjects have three *ABCD4* missense variants (rs57773157, rs34992370, rs58272575) that co-occur in a rare 8-SNP haplotype spanning 12.9 kb with a frequency of 0.3% in cases and 0% in controls (Figure S2.4A). Another 10 subjects have two *CELSR1* missense variants (rs61741871 and rs75983687), and eight of these subjects also have two *GTSE1* missense variants. One subject is homozygous for all four variants. The subjects who have the *CELSR1* and *GTSE1* variants (rs61741871, rs75983687, rs34404175, rs35503220) share a rare 12-SNP haplotype spanning 77.6 kb with a frequency of 0.1% in cases and 0.1% in controls (Figure S2.4B). IBD estimates for the 10 cases with the *ABCD4* variants ($\hat{\pi}$=0.028) and the eight AD cases with the *CELSR1* and *GTSE1* variants ($\hat{\pi}$=0.029) are only slightly higher than genome-wide IBD sharing and indicate that these individuals are not more closely related to each other than to all subjects. Of note, there are very few common SNPs in the 500 kb region including the rare *ABCD4* variants, suggesting high sequence conservation in this region.

To identify additional genes which may have over-representation of deleterious AD-related variants, we filtered genes which contained at least three distinct variants each occurring in at least five AD cases but absent in controls (Table S2.7). The *ABCD4* rs61744947 variant appears on the same haplotype containing the other three *ABCD4* variants. Three *LAMC3* variants were observed in the same seven subjects. *TTN* had the greatest number of distinct variants (n=6) that were observed in AD cases only. Genome-wide, nine genes not previously associated with AD contained a high functional impact variant that was present in at least seven AD cases but absent in controls (Table S2.8).

2.4.4 *Rare Variant Burden*

To test if disease is associated with greater burden of rare deleterious variants, gene burden tests were performed for models including high impact variants and high and moderate impact variants for MAF$\leq$0.01 and MAF$\leq$0.5 in genes previously associated with AD risk, AD-related traits, or other dementias (Table 2.5). These analyses showed that AD cases have a significantly higher burden of moderate and high impact rare deleterious variants in this group of genes compared to controls (p=0.006).

2.5 Discussion

We identified several rare variants that are predicted to have high impact on protein structure and may increase AD risk. These variants were not detected in previous analyses of the same ADSP WES dataset that were agnostic with

respect to functional impact of the variants and conducted using current statistical testing approaches [17]. Our focus on variants observed in AD cases but not controls yielded results that are often undetected by traditional genetic association methods that cannot evaluate empty cells, regardless of sample size or frequency of variants among cases. Several of our top-ranked results confirm previously identified AD associations with rare variants including *PSEN1* rs63749824 [46] and rs63750592 [49], *SORL1* rs139710266 [50], *MAPT* rs63750424 [51], and *ABCA7* E1769X [54] which suggest that novel findings identified by our approach may be robust. Two of our novel findings suggest evidence of shared genetic mechanisms between AD and other rare dementia syndromes, namely CADASIL and Nasu-Hakola disease. Our study also showed that AD cases have a significantly higher burden of deleterious rare coding variants in known AD, AD-related or other dementia genes compared to controls. This observation generalizes previous findings in *SORL1* [50, 55], *MAPT* [56], *TREM2* [57, 58] and *ABCA7* [59] that both common and rare variants in the same gene may independently contribute to AD risk.

We observed the rare *NOTCH3* rs149307620 allele in 11 AD cases and one MCI case, but not in controls, in the combined ADSP WES, ADSP WGS and ADNI WGS datasets. The most remarkable finding from analysis of the Utah high-risk pedigree WES dataset were rare *NOTCH3* rs140914494 and rs141402160 alleles in a pair of affected half first-cousins. These mutations in exons 4 (rs140914494), 5 (rs141402160) and 6 (rs149307620) are located in the JAG1

binding site and involve the gain or loss of an alanine residue (Table 2.2). In contrast, the rs114447350 and rs112197217 variants are located near the end of the coding sequence (exons 33 and 21, respectively) and predicted to be clinically benign [60], and are thus unlikely to be causally related to AD. Many other *NOTCH3* variants have been associated with CADASIL which typically replace the wild-type amino acid with a cysteine residue or replace a highly conserved cysteine residue with another amino acid [39, 43, 44] although there are several exceptions [61]. Available clinical and autopsy data for the individuals with *NOTCH3* mutations were consistent with the diagnosis of AD and not CADASIL. Our protein modeling demonstrated that the AD-associated *NOTCH3* mutations in exons 4-6 result in quantitative changes in hydrogen bonding causing increased ligand interaction, whereas CADASIL *NOTCH3* mutations lead to qualitative changes involving disrupted di-sulfide bonding that affect overall protein structure and receptor maturation, and differ with respect to their consequences both on ligand binding and ligand-induced signaling [43, 44].

Our protein-protein interaction network and gene-set enrichment analyses demonstrated that *NOTCH3* is closely tied to AD pathways and biological processes. Notch-3 signaling can be triggered by both Delta-JAG-type ligands and requires ADAM10 and presenilin-1 or -2, making it part of the AD related presenilin pathway [62]. Previous studies showed JAG1-Notch signaling and subsequent hippocampal neurogenesis and astrogenesis are regulated by cleavage by BACE1, a promising AD drug target [62]. This process, which is

more active during early development and decreases in adulthood, affects normal neuronal development and alters neurogenesis and thus can have long-term effects [62]. In addition, Notch-3 is a substrate for γ-secretase (presenilin) inhibition, which when dysregulated, can cause mis-processing of the amyloid precursor protein resulting in accumulation of the toxic amyloid-β peptide [63].

Our collective genetic and bioinformatics findings provide the strongest pathogenic link to date between *NOTCH3* and AD. A previous study reported association of AD with a distinct *NOTCH3* mutation (p.R1231C) in a Turkish family [64], however this variant was detected in only one affected member and there is conflicting information about its pathogenicity [65]. Sassi et al. tested the hypothesis that genes associated with Mendelian adult onset leukodystrophy are also associated with AD in a sample including 332 sporadic AD cases and 676 controls, and found significant gene-based association with *NOTCH3*, a result driven primarily by a common synonymous coding variant [66].

The *TREM2* Q33X mutation that was observed in four AD cases in our sample and in four AD cases and one unaffected relative with an unspecified age of an AD Q33X carrier in gene resequencing studies targeting *TREM2* [14, 67, 68] is much rarer than the well-documented R47H variant that has been associated with increased risk of AD in several studies [14, 69] including the ADSP cohort [7,8]. Homozygosity of this mutation causes Nasu-Hakola disease (NHD), a very rare disorder characterized by early-onset dementia and multifocal bone cysts

[45], and has also been observed in a member of a Turkish family with frontotemporal dementia (FTD)-like syndrome including the appearance of aggressive behavior and generalized tonic-clonic seizures before age 30 but without bone involvement [70]. Because persons with NHD and the FTD syndrome case have a more severe phenotype overall and much earlier onset of dementia symptoms than late-onset AD cases who are heterozygous for Q33X, the behavior of this mutation may more resemble an intermediate inheritance than an autosomal dominant model. This idea is consistent with the observation that an NHD patient's living the parents, who were obligate Q33X heterozygotes, both had evidence of β-amyloid deposition by cerebrospinal fluid analysis and florbetapir-PET imaging [71].

Furthermore, unlike the *TREM2* R47H mutation and rare coding variants at other loci that have been associated with AD [4-8], the *NOTCH3* rs149307620 and *TREM2* Q33X mutations appear to be fully penetrant among persons surviving to late age, which perhaps would be the first examples of causative mutations for late-onset AD. Obviously, this assertion is somewhat speculative given the small number of AD cases documented to have these mutations.

Our study also implicated multiple functional variants in several novel genes as risk factors for AD. Mutations in *ABCD4* cause an inborn error of vitamin B12 metabolism [72]. B12 deficiency is associated with cognitive impairment and the level of circulating vitamin B12 has been associated with AD risk [73]. *ABCD4*

encodes an ATP binding cassette transporter that is in the same family as well-established AD gene *ABCA7* [37, 58, 74]. The AD-associated *CELSR1* rs61741871 (P2983A) missense variant has also been associated with craniorachischisis, a severe neural tube defect [75], and other *CELSR1* variants have been identified as ischemic stroke risk factors in Japanese individuals [76]. The *CELSR1-3* family of genes has multiple functions in the nervous system and has distinct roles in brain development and maintenance [77]. *GTSE1* regulates G1/S cell cycle transition and microtubule stability and is involved in pivotal neurodegeneration pathways [78]. It is not clear which of these *CELSR1* and *GTSE1* mutations may directly influence AD risk. *LAMC3* encodes laminin subunit gamma 3 and multiple experimental studies have linked laminins to AD [79, 80]. *LAMC3* has been significantly associated with age at onset of AD [81].

Our study has several noteworthy limitations. Because we focused on very rare variants, our sample of more than 10,000 subjects was inadequate to establish statistical significance. Thus, our findings require replication in independent samples. Unfortunately, we were unable to replicate these findings in the ADGC GWAS dataset because of low and inconsistent imputation quality for these very rare variants, despite the use of the large HRC reference panel [82]. In addition, our genome-wide MAC cutoffs for focusing on particular variants are arbitrary and, therefore, some important findings may have been overlooked. Finally, although one of the explicit goals of the ADSP is to identify variants that protect against AD [41], the design corresponding to the one we employed to identify risk

variants (i.e., a “controls-only” analysis) is less rigorous because in the absence of statistically significant tests it is difficult to demonstrate a protective effect if the variant has reduced penetrance.

Conclusions and Future Directions. We identified associations with novel variants in previously established AD genes and with several novel potential AD genes that did not emerge in previous analyses of a large WES dataset using conventional statistical thresholds [17]. Our findings with the *NOTCH3* and *TREM2* variants demonstrate that mutations in the same gene can result in very dissimilar types of dementia. Moreover, variable dose of a particular mutation (i.e., *TREM2* Q33X) can cause different types of dementia. These findings suggest that minor differences in protein structure or amount of wild type protein can result in different clinical outcomes. Understanding these genotype-phenotype relationships may provide further insight into the pathogenic nature of the mutations, as well as offer clues for developing new therapeutic targets.

Table 2.1 High and moderate impact rare variants in previously established AD genes occurring in ≥ 4 AD cases and no controls.

SNP (Chromosome: position: major allele: minor allele) *	MAC Cases	Gene	ID	Mutation Type	Disease Impact	CADD Score	Previously Associated with AD
19:15302421:C:T	10	NOTCH3	rs149307620	Missense	Moderate	26.6	No
14:73637653:C:T	7	PSEN1	rs63749824	Missense	Moderate	18.8	[26, 27]
9:4118361:G:C	6	GLIS3	rs200263979	Missense	Moderate	15.1	No
14:93154397:G:A	6	RIN3	Novel	Missense	Moderate	10	No
17:44143925:C:T	6	KANSL1	rs138698439	Missense	Moderate	21.8	No
11:121384991:A:G	5	SORL1	rs139710266	Missense	Moderate	26.3	[30]
14:73640432:G:A	5	PSEN1	rs375376095	Missense	Moderate	10.5	No
16:31193959:ATTC:A	5	FUS	Novel	In-frame deletion	Moderate	11.9	No
17:56385997:G:A	5	BZRAP1	rs61732758	Missense	Moderate	15.4	No
19:1054190:A:G	5	ABCA7	rs376824416	Splice acceptor	High	21.7	No
2:234113197:C:T	4	INPP5D	rs532718867	Missense	Moderate	1.387	No
4:7731386:G:A	4	SORCS2	rs371407070	Missense	Moderate	25.4	No
6:41129295:G:A	4	TREM2	Q33X, rs104894002	Stop Gained	High	5.4	[4]
7:143088779:C:T	4	EPHA1	rs201365734	Missense	Moderate	20	No
7:51096830:C:G	4	COBL	rs112568753	Missense	Moderate	0.319	No
8:17478643:A:G	4	PDGFRL	rs144384825	Missense	Moderate	17.96	No
9:4118324:C:A	4	GLIS3	rs200959196	Missense	Moderate	2.739	No
11:59949160:A:G	4	MS4A6A	Novel	Missense	Moderate	22.9	No
14:105396368:T:G	4	PLD4	Novel	Missense	Moderate	23.6	No
14:73637521:G:A	4	PSEN1	rs63750592, R35Q	Missense	Moderate	15.68	[28, 29]
14:93154353:T:TGTGCGCGCA	4	RIN3	Novel	In-frame insertion	Moderate	16.91	No
15:102192524:A:G	4	TM2D3	rs201415552	Missense	Moderate	8.863	No
16:81972496:A:G	4	PLCG2	Novel	Missense	Moderate	22.4	No
17:42429772:C:T	4	GRN	Novel	Stop Gained	High	18.4	No
17:44101427:C:T	4	MAPT	rs63750424, R406W	Missense	Moderate	18.49	[31-33]
17:61568688:G:A	4	ACE	rs143507892	Missense	Moderate	33	No

17:61574683:C:T	4	ACE	Novel	Missense	Moderate	5.394	No
19:1043794:G:A	4	ABCA7	rs147846250	Missense	Moderate	8.041	No
19:1055151:T:C	4	ABCA7	rs145987355	Missense	Moderate	26	No
19:1058154:G:T	4	ABCA7	E1769X, rs770510230	Stop Gained	High	37	[34]
19:1059029:G:A	4	ABCA7	rs143615723	Missense	Moderate	0.011	No
19:15272218:G:A	4	NOTCH3	rs114447350	Missense	Moderate	16.05	No

* Chromosome position according to GRCh38.p7 assembly

Table 2.2 *NOTCH3* mutations associated with AD

SNP	Position (chr 19) *	Exon	Protein Position	Residue Change	Observed Mutation Carriers
rs140914494	15,192,046	4	198	Ala>Glu	AD-affected relative pair
rs141402160	15,191,804	5	248	Gly>Ala	11 AD cases, 1 MCI case
rs149307620	15,191,610	6	284	Ala>Thr	AD-affected relative pair

* Chromosome position according to GRCh37.p13 assembly

Table 2.3 Gene-set enrichment analysis of *NOTCH3/JAG1* protein-protein interaction network

PANTHER Pathway	# Genes Annotated to Pathway	# Genes in Network	Expected P value *	Fold Enrich-ment	Direction	Unadjusted P value	FDR
Notch signaling	42	22	0.06	> 100	+	3.98E-51	6.48E-49
Presenilin	123	14	0.18	> 100	+	7.09E-28	5.78E-26
Amyloid secretase	69	7	0.10	> 100	+	8.59E-18	3.50E-16
Angiogenesis	173	10	0.25	40.54	+	4.92E-14	1.61E-12

PANTHER = **P**rotein **AN**alysis **TH**rough **E**volutionary **R**elationships) Classification System

*Expected probability of observing at least x number of genes out of the total n genes in the PANTHER list annotated to a particular pathway, given the proportion of genes in the reference Homo Sapiens whole genome that are annotated to that pathway

Table 2.4 High and moderate impact rare variants genome-wide occurring in > 10 AD cases and no controls.

SNP (Chromosome: position: major allele: minor allele) *	MAC Cases	Gene	ID	Mutation Type	Disease Impact	CADD Score
11:5474894:T:C	12	OR51I2	rs74049540	Missense	Moderate	15.5
19:38742032:G:A	12	PPP1R14A	rs140507040	Missense	Moderate	15.5
14:45374714:CTTG:C	11	C14orf28	Novel	In-frame deletion	Moderate	20.5
17:42433931:C:G	11	FAM171A2	rs190723348	Missense	Moderate	22.2
17:7164266:G:A	11	CLDN7	rs149308129	Missense	Moderate	21.9
20:62198583:G:A	11	HELZ2	rs35691275	Missense	Moderate	5.6
22:46759981:G:C	11	CELSR1	rs61741871	Missense	Moderate	3.6
22:46762988:C:T	11	CELSR1	rs75983687	Missense	Moderate	17
1:93682193:A:T	10	CCDC18	rs191574433	Missense	Moderate	15.1
7:150389837:TC:T	10	GIMAP2	Novel	Frameshift	High	5.9
9:33941796:C:T	10	UBAP2	rs150194348	Missense	Moderate	23.2
11:67957408:C:T	10	SUV420H1	rs138431226	Missense	Moderate	17.4
12:88456506:T:C	10	CEP290	Novel	Missense	Moderate	22.1
14:31119819:C:A	10	SCFD1	rs35187633	Missense	Moderate	32
17:74288944:A:G	10	QRICH2	Novel	Missense	Moderate	9.1
14:74754936:G:A	10	ABCD4	rs57773157	Missense	Moderate	18.4
14:74763064:C:T	10	ABCD4	rs34992370	Missense	Moderate	8.7
14:74766360:T:C	10	ABCD4	rs58272575	Missense	Moderate	15.7
17:58260659:A:C	10	USP32	rs201933998	Missense	Moderate	11
19:15302421:C:T	10	NOTCH3	rs149307620	Missense	Moderate	26.6
19:49123796:G:A	10	SPHK2	Rs158184205	Missense	Moderate	15.4
22:41605776:G:C	10	L3MBTL2	rs143455680	Missense	Moderate	27.4
22:46704734:C:T	10	GTSE1	rs34404175	Missense	Moderate	5.9

22:46708152:G:A	10	GTSE1	rs35503220	Missense	Moderate	0.001

* Chromosome position according to GRCh38.p7 assembly

Table 2.5 Rare variant burden for established AD genes.

Model	Beta	SE	P-value
High Impact, MAF≤ 0.01	0.005	0.166	0.977
High/Moderate Impact, MAF≤ 0.01	0.062	0.023	0.006*
High Impact, MAF≤ 0.05	0.005	0.166	0.977
High/Moderate Impact, MAF≤ 0.05	0.061	0.022	0.006*

* Significant at .05 alpha p-value threshold

Figure 2.1 Notch-3 protein model highlighting position of the AD-associated SNP rs149307620. Predicted model for: **A.** wild type allele with alanine at mutation site, **B.** mutant allele with threonine at mutation site, and **C.** binding of notch-3 (in red) to JAG1 ligand (in yellow). Possible hydrogen bonding is displayed which would likely cause greater interaction with the ligand.

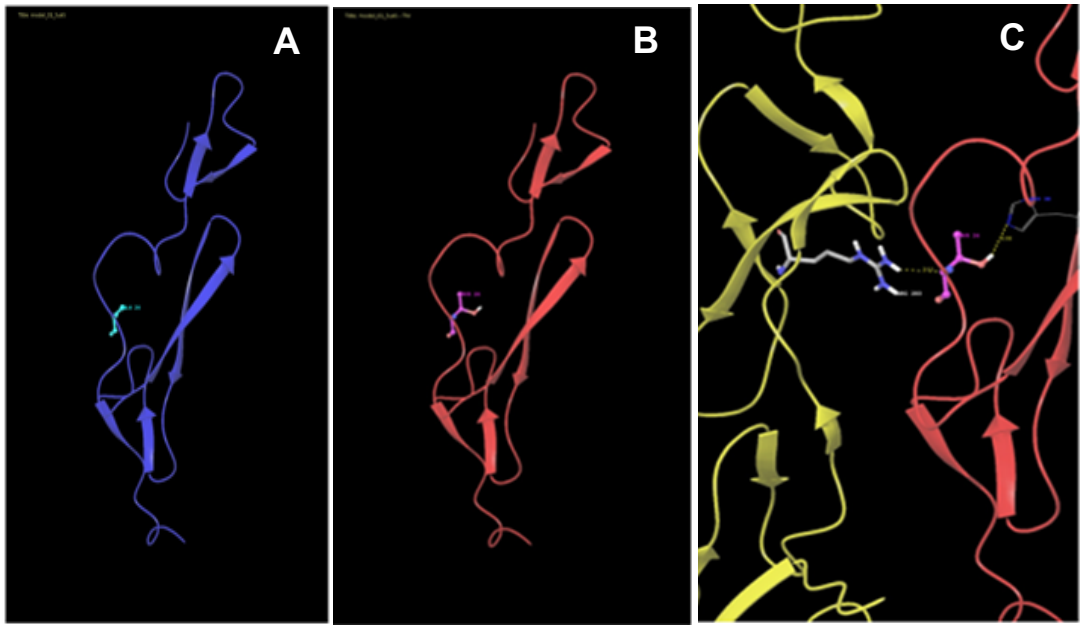


Figure 2.2 Homologous Protein Modeling of CADASIL *NOTCH3* rs137852641 mutation. Predicted model for: **A.** wild type allele with arginine at the mutation site and **B.** mutant allele with cysteine at the mutation site. Gain of a cysteine residue disrupts di-sulfide bonding within the protein, affecting overall protein structure.

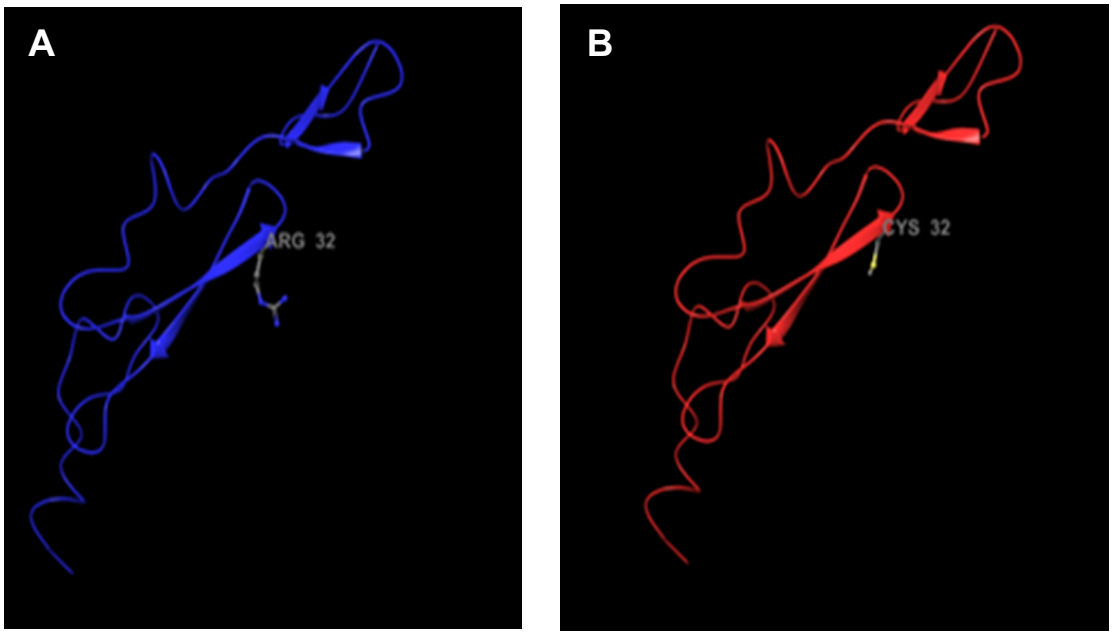
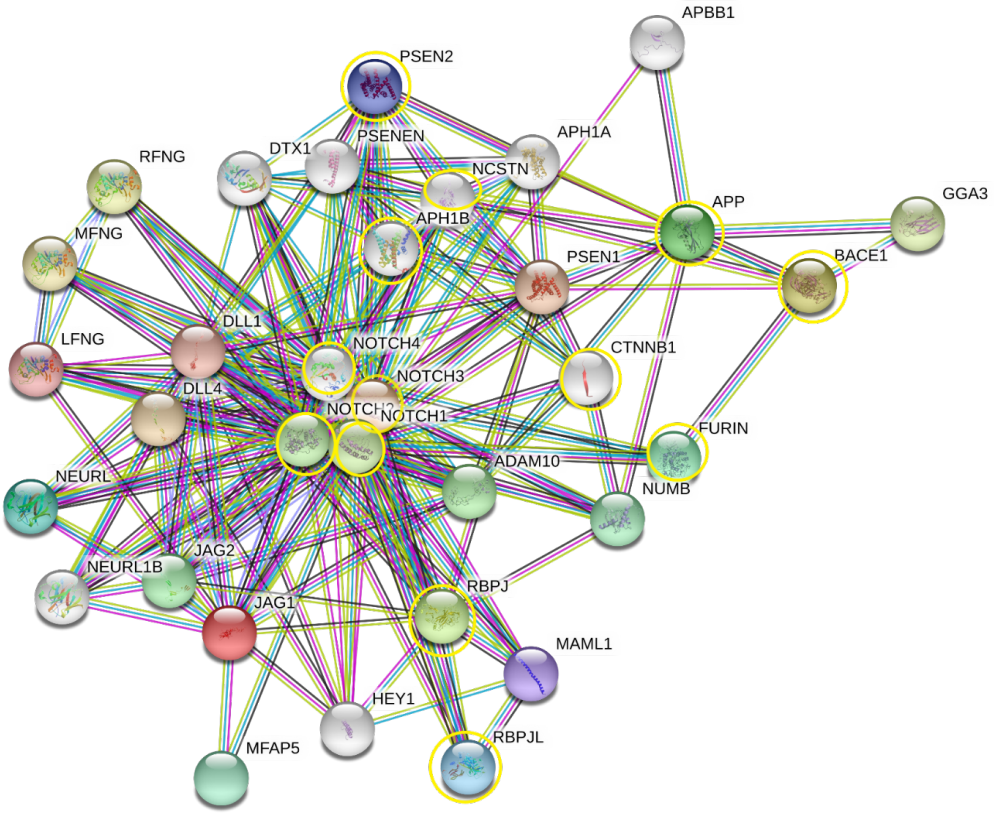


Figure 2.3 Protein-protein interaction network including *NOTCH3* and *JAG1*. Network includes 30 genes showing high-confidence interactions in humans according to the STRING database [18]. The colors of the edges refer to the type of evidence linking the corresponding proteins: red=gene fusion, dark blue=co-occurrence, black=co-expression, magenta=experiments, cyan=databases, light green=text mining, mauve=homology. Genes present in AD pathways determined by gene-set enrichment analysis (Table 2.3) are highlighted by yellow circles.



.

CHAPTER 3: Identification of rare variants that influence AD gene expression in blood and brain tissue data

3.1 Abstract

Background: Some of the unexplained heritability of Alzheimer's disease (AD) may be due to rare variants, and AD risk is also associated with changes in gene expression. Rare variants have been to have a large impact on gene expression changes across tissues. However, gene expression studies face the same challenges as genome-wide association studies (GWAS) studies for rare variants, so we investigated set-based methods for rare expression quantitative trait loci (eQTL) analysis.

Methods: Gene-level and pathway-level cis rare-eQTL mapping was performed genome-wide in blood from 713 Alzheimer's Disease Neuroimaging Initiative (ADNI) participants and in brain donated by 475 Religious Orders Study/Memory & Aging Project (ROSMAP) participants. The association of gene or pathway expression with a set of all cis potentially regulatory low frequency and rare variants within 1 Mb of genes was evaluated using linear regression models in blood and brain tissues to further assess AD risk.

Results: A total of 65 genes were significant brain gene-level rare eQTLs and a total of 307 genes were significant blood gene-level rare eQTLs. Significant brain gene-level results contain two established AD genes: *HLA-DRB1* and *HLA-DRB5*. Interestingly, 17% (11/65) of all significant brain gene-level results were on chromosome 6, including two other HLA genes- *HLA-A* and *HLA-DOB*. There

also appear to be a number of cytoskeleton related genes in blood gene-level findings such as *NUMA1*, *ACTG1*, and *DNAAF3*. 5 genes were seen in both blood and brain gene-level results: *GNMT*, *LDHC*, *RBPMS2*, *DUS2*, and *HP* (Table 3.1). Of the matching genes, 4 out of 5 had similar p-values but *RBPMS2* was much more significant in blood (Blood $p=1.69x10^{-36}$ vs. brain $p=9.90x10^{-08}$). In brain, an aggregate of low frequency and rare variants in *CCL7* and *CCL8* were each associated with the Inflammation mediated by chemokine and cytokine signaling pathway. Of the 22 significant genes associated with pathway expression in blood, 6 (*ALOX5AP*, *CXCR2*, *FPR2*, *GRB2*, *IFNAR1*, *RAF1*) were associated with the inflammation pathway as well, and 5 out of the 6 have previously been associated with AD.

Conclusions: This study identified a number of significant gene-level and pathway level significant rare eQTL results, found additional evidence for the importance of the immune/inflammatory system in AD, and highlighted the advantages of using a set-based eQTL method for low frequency and rare variants.

3.2 Background

Late-onset Alzheimer's disease (AD) is the most common type of dementia that affects an estimated 5.7 million individuals older than 65 years old just in the United States, with the number projected to rise to 14 million by 2050 [1]. AD susceptibility is highly heritable (h^2=58-79%) [2], but common variants

discovered through genome-wide association studies (GWAS) can only explain one third of its heritability [2]. Rare variants with large effects may contribute to some of the missing heritability of AD. [3]. Recent sequencing studies have identified robust AD associations with rare variants in gene loci including *TREM2*, *AKAP9*, *UNC5C*, *ZNF655, IGHG3*, *CASP7* and *NOTCH3* [14-18, 83], and more AD-related rare variants will be identified from large-scale whole-genome sequencing (WGS) studies including rare variants in non-coding regions. However, the underlying mechanism by which these rare variants impact AD risk and their target/functional genes remain challenging to discover.

AD risk variants are associated with gene expression levels, as demonstrated by recent expression quantitative trait locus (eQTL) studies [5, 6]. Rare variants contribute to extreme gene expression within single tissues and multi-tissues [4, 84-86]. However, eQTL studies face the same challenges as GWAS studies for rare variants. Although gene-based tests, which test aggregate effects of rare variants, are commonly used in rare variant association analysis in complex diseases, only a few studies have applied this type of model for eQTL studies of rare variants. Recent eQTL studies that have used a set-based eQTL approach include testing gene expression with multiple SNPs chosen by variable selection [87, 88], using a gene-based partial least squares method to correlate multiple gene transcript probes with multiple SNPs [89], and identifying variants associated with transcript and protein modules [90]. These methods were not focused on rare variants but still had the advantages of higher

power with a potential to find significant variants with lower frequency.

Few studies have applied a set-based eQTL method for rare variants. Lutz et al. very recently applied burden and SKAT tests to normalized read counts in RNA-seq studies [91]. In this study, a set-based eQTL method focusing on low frequency (MAF<0.05) and rare (MAF <0.005) variants using SKAT-O [92] is implemented with sets defined as genes and gene modules/pathways. A gene-based cis eQTL approach will be applied to identify genes which have a set of potentially regulatory rare variants significantly associated with their expression. A pathway-based approach is applied to identify which gene with its set of regulatory rare variants most contributes to the overall gene expression profile of a significant pathway containing a set of co-expressed genes/functional-related genes. These methods will identify rare-eQTLs in human blood and brain to further assess their effect on AD risk and prioritize function genes for AD.

3.3 Methods

3.3.1 *Study Cohorts*

Alzheimer's Disease Neuroimaging Initiative (ADNI). ADNI is a longitudinal multisite study that aims to develop clinical, imaging, genetic, and biochemical biomarkers to improve the prevention and treatment of AD [93]. The study began in 2004 and is currently funded until 2021 with 63 sites in the US and Canada enrolling subjects across three stages: AD dementia, mild cognitive impairment (MCI), and normal cognitive functioning [93]. Affymetrix Human Genome U219

array gene expression data from whole blood and whole-genome sequencing (WGS) genotype data are used in this study.

Religious Orders Study (ROS)/ Memory and Aging Project (MAP). ROS enrolled older nuns and priests from across the US, without known dementia for longitudinal clinical analysis and brain donation. MAP enrolled older subjects without dementia from retirement homes who agreed to brain donation at the time of death [94, 95]. RNA-sequencing brain gene expression data generated from DLPFC region and WGS genotype data were obtained from the AMP-AD knowledge portal (https://www.synapse.org/#!Synapse:syn3219045) [96]. Characteristics of these subjects are provided in Table S3.1.

3.3.2 *Data Processing*

Downloaded ADNI microarray gene expression data were normalized, background-correction and log-transformed. ROSMAP RNA-seq data were normalized by library size (i.e., FPKM), then log-transformed. Then we applied surrogate variable analysis (SVA) [97] on the log-transformed data to obtain surrogate variables for global technical effects and hidden effects which were included as covariates in analysis models for eQTL discovery. Additional filtering steps of ADNI and ROSMAP GWAS and gene expression data included eliminating subjects and values with missingness, restricting gene expression data to protein coding genes (12,971 genes in ROSMAP and 16,025 genes in ADNI), and selecting only bi-allelic low-frequent and rare variants (MAF ≤ 0.05).

3.3.3 *Set-based eQTL Analysis*

A flowchart of the analysis in blood and brain tissues to identify set-based eQTLs is shown in Figure 3.1, including the total number of tests performed and the Bonferroni-corrected P-value thresholds used for defining gene-level and pathway-level significant eQTL results. Rare eQTLs throughout will refer to eQTLs with an aggregate of low frequency and rare variants (MAF<0.05).

Gene-level cis-eQTL Analysis

For common variants, eQTL analysis tests the association between the expression of one gene and one variant. Gene-level eQTL analysis will test the association between the expression of one gene and many rare variants in and near the gene, comparable to how gene-based analysis in genetics tests the association of a trait such as AD status with more than one variant within one gene. The gene-based test SKAT-O combines the best features of variance component (SKAT) and burden tests into one test with optimal power [92] and is implemented in this study design for set-based eQTLs by setting the outcome as the gene expression value. The gene expression value for each gene is evaluated for association in a linear regression model with the aggregation of all low frequent and rare *cis*-regulatory SNPs within 1 Mb of the gene. In ROSMAP, the regression model also includes covariates for age, sex, post-mortem interval (PMI), study (ROS or MAP), and a term for a surrogate variable (SV1) derived from the gene expression data matrix to account for unmeasured/hidden technical effects on gene expression using SVA [97]. In ADNI, covariates are

baseline age, sex, RIN, year of collection, and also SV1. RIN refers to the RNA Integrity Number, an indicator of the quality of the RNA. SKAT-O was implemented with groupwise tests using EPACTS software(https://genome.sph.umich.edu/wiki/EPACTS). A total of 12,971 tests were performed in ROSMAP and a total of 16,025 tests were performed in ADNI, examining the associations between each of the genes and low frequency and rare variants within 1 Mb of the gene in the datasets. After a Bonferroni correction for multiple testing, the significance thresholds to define significant gene-level rare eQTLs were as follows: $P < 3.86 \times 10^{-6}$ (0.05/12,971) in ROSMAP and $P < 3.12 \times 10^{-6}$ (0.05/16,024) in ADNI.

To determine which particular variants are driving significant results, eQTL tests using linear regression models in R [98] were performed afterwards for all significant genes and each individual *cis*-regulatory bi-allelic low frequency and rare variant (MAF ≤ 0.05) within 1Mb of the gene that was included in the aggregate group of SNPs from the gene-level tests. After a Bonferroni correction for multiple testing, the significance thresholds to define significant individual gene-SNP eQTLs were as follows: $P < 1.83 \times 10^{-6}$ (0.05/ 27,393) in ROSMAP and $P < 1.17 \times 10^{-7}$ (0.05/ 425,995) in ADNI.

Pathway-level cis-eQTL Analysis

Pathway-level eQTL analysis will test the association of a pathway, containing many genes, with sets of variants in each of the genes in the pathway one at a time. The first step in the pathway level analysis involved finding gene modules

of co-expressed genes with Weighted Gene Co-expression Network Analysis (WGCNA) in R [99]. With all protein coding genes in ADNI and ROSMAP datasets, a co-expression network was created in WGCNA and gene modules(clusters) of highly correlated genes were identified. WGCNA network construction and module detection was run with the following default parameters: soft threshold power beta = 6.00, deepSplit = 2(medium sensitivity), minimum module size of 20, and a merge cut height of 0.15. These default parameters were chosen because they have worked well in several applications as referenced by the WGCNA methods paper [99] and similar AD studies [100]. Each gene module can be summarized with its module eigengene (ME) value, the first principal component of a module. The ME is considered to be representative of the gene expression profiles in a gene module.

Each gene module was then input into PANTHER for pathway enrichment analysis; only the significantly enriched pathways in the gene modules were analyzed further. Pathway-level eQTL analysis was performed on each significantly enriched pathway with the ME of the gene module as the outcome. Associations between the ME and each gene in the gene module were tested individually using all cis SNP genotypes with MAF<0.05 within 1 Mb of the gene that were annotated as potentially regulatory. The same covariates were used for ADNI and ROSMAP datasets as in the gene-level eQTL tests and analyses were also implemented in EPACTS. A total of 77 tests (9 enriched pathways vs. genes in each pathway=77 genes) were performed in ROSMAP and a total of 100 tests

(16 enriched pathways vs. genes in each pathway=100 genes) were performed in ADNI, examining the associations between each significant pathway expression and the potentially regulatory low frequency and rare variants within 1 Mb of each gene from the gene module in the pathway. After a Bonferroni correction for multiple testing, the significance thresholds to define significant gene-level rare eQTLs were: $P < 6.49 \times 10^{-4}$ (0.05/77) in ROSMAP and $P < 5.0 \times 10^{-4}$ (0.05/100) in ADNI. The same covariates for ROSMAP and ADNI used in the gene-level analyses are applied to the pathway-level analyses.

3.3.4 *Functional Annotation of Variants*

Annotations were collected CADD v1.6 (hg19) and gwava v1.0 (hg19). Genomic coordinates of the ADNI GRCh38 genome build dataset were converted using liftOver software (https://genome.ucsc.edu/cgi-bin/hgLiftOver) to hg19 to use EPACTS software which only supports VCFs aligned with NCBI build 37. ADNI and ROSMAP WGS SNPs were annotated by matching chromosome, position, reference, and alternative alleles. To detect potentially regulatory variants, variants having a CADD score > 15 or a gwava region score > 0.5 were selected.

3.3.5 *Pathway Enrichment Analysis*

Gene-set pathway enrichment analysis was performed using the PANTHER (Protein ANalysis THrough Evolutionary Relationships) software tool [31] to determine which pathways are significantly enriched in the gene modules identified from the WGCNA for pathway-level eQTL analysis. Significance of the enriched pathways was determined by the Fisher's Exact test with false

discovery rate (FDR) < 0.05.

3.4 Results

3.4.1 *Gene-level rare cis-eQTL associations*

A total of 65 genes ($P < 3.86 \times 10^{-6}$) were significant brain gene-level low frequency and rare eQTLs (Table S.2) and a total of 307 genes ($P < 3.12 \times 10^{-6}$) were significant blood gene-level rare eQTLs (Table S.3). The average number of cis bi-allelic variants used in significant result tests is 416 in brain and 678 in blood, which is a somewhat larger dataset. Significant brain gene-level results contain two established AD genes: *HLA-DRB1* [101] and *HLA-DRB5* [58] (Table S.2). Interestingly, 17% (11/65) of all significant brain gene-level results were on chromosome 6, including two other HLA genes- *HLA-A* and *HLA-DOB*. Significant blood gene-level results contain three established AD genes: *ABCA7 [58]*, *ECHDC3 [101]*, and *MS4A6A* [58] (Table S.3). 5 genes were seen in both blood and brain gene-level results: *GNMT*, *LDHC*, *RBPMS2*, *DUS2*, and *HP* (Table 3.1). Of the matching genes, 4 out of 5 had similar p-values but *RBPMS2* was much more significant in blood (Blood $p=1.69 \times 10^{-36}$ vs. brain $p=9.90 \times 10^{-08}$). Overall, all 5 genes have many more variants in blood gene-level tests. When testing each significant gene with each individual low frequency and rare variant within 1 Mb of the gene, there were 61 significant eGene-eSNP eQTL pairs (with 22 unique eGenes) in brain and 832 significant eQTL pairs (with 185 unique eGenes) in blood, as seen in Tables S3.4 and S3.5. *RP11-529K1* has the

most significant eQTL p-value in brain with one eSNP ($p = <1.0 \times 10^{-314}$), but its gene-level rare eQTL p-value is not as strong ($p = 3.54 \times 10^{-11}$). There are also a total of 7 eSNPs for eGenes *COPZ1* and *TMPRSS6*, with most of these variants having strong p-values. Significant eGene-eSNP eQTL results in blood has numerous variants associated with expression of the following genes: *KRT79* (n=36), *TAC3* (n=32), *CDK12* (n=24), and *SOS1* (n=20). In brain, there were two variants that were both significantly associated with the expression of two different zinc finger protein genes (*ZNF101* and *ZNF253*). There were 28 eSNPs that were duplicates in blood results, with each eSNP associated with two different eGenes that are located close together (*GSDMA* and *IKZF3*, *DHRS4* and *DHRS4L2*, *CMTM2* and *ATP6V0D1*).

3.4.2 *Significant gene module pathways*

There were 17 co-expressed gene modules in brain data with 4,481 genes in the largest module M1 and 38 genes in the smallest module M17 (Table S3.6), while there were 34 gene modules in blood data with 2,065 genes in the largest module M1 and 29 genes in the smallest M34 (Table S3.7). These modules and their eigengenes are summarized in Figure S3.1. In brain, module 14 seems to have very different expression compared to the rest of the modules and modules 1 and 2 seem to be highly related, as evidenced by their lowest merging heights. In blood, more ‘meta-modules’ exist of groups of eigengene clustering with related expression. The proportion of variance explained by the module eigengene was on average 49% in brain (Table S3.8) and 45% in blood (Table

S3.9), suggesting the module eigengenes captured a significant proportion of gene module variance.

Pathway enrichment analysis of each gene module revealed 9 significant enriched pathways in brain (Table 3.2) and 16 in blood (Table 3.3). The apoptosis signaling pathway, the CCKR signaling map, and inflammation mediated by chemokine and cytokine signaling pathway (appears 3 times separately in different gene modules in blood) all appear in both brain and blood enriched pathways and have all been previously linked to AD. When looking closer at the genes within each enriched pathway, *HSPA6* and *TNFRS10C* are in the apoptosis signaling pathway and *MMP-9* is in CCKR Signaling Map in both tissues (Figure S3.2). Of these shared genes in both tissues, only *TNFRSF10C* is significantly contributing to pathway expression at a Bonferroni threshold, though all shared genes contribute more significantly to their pathway in blood compared to brain (Table S3.10). Though there is no overlap in genes between the enriched inflammation pathways in both tissues, there is an intersect of 4 genes between the gene modules in which these pathways occur: *CCL3*, *CCL4*, *SPOCD1*, and *CXCL5* (Table S3.11). The module eigengenes for the gene modules for the brain enriched pathways all occur in the rightmost branch of the eigengene clustering tree (Table 3.2, Figure S3.1-A2). A large proportion (5/16) of the blood enriched pathways occur in gene module 5 (Table 3.3).

3.4.3 *Pathway-level rare cis-eQTL associations*

A total of 2 genes had rare eQTLs significantly associated ($P < 6.49 \times 10^{-4}$) with

pathway expression in brain (Table 3.4) and a total of 22 genes had rare eQTLs significantly associated ($P < 5.0 \times 10^{-4}$) with pathway expression in blood (Table 3.5). The average number of cis bi-allelic variants used in significant result tests is 329 in brain and 740 in blood, which is a somewhat larger dataset. In brain, the aggregated variants in *CCL7* and *CCL8* were each associated with the Inflammation mediated by chemokine and cytokine signaling pathway (Table 3.4). Of the 22 significant genes associated with pathway expression in blood, 6 (*ALOX5AP*, *CXCR2*, *FPR2*, *GRB2*, *IFNAR1*, *RAF1*) were associated with the inflammation pathway as well, including the top 2 most significant, and 5 out of the 6 have previously been associated with AD. There were also 3 genes (*CFLAR*, *TMBIM6*, and blood/brain pathway shared gene *TNFRSF10C*) significantly associated with the apoptosis signaling pathways and 1 heat shock protein gene (*HSPA5*) significantly associated with the Parkinson's disease (PD) pathway. Both gene-level and pathway-level significant results contain rare eQTLs with the gene *ALOX5AP* (Table S3.3, Table 3.5).

3.5 Discussion

Low frequency and rare variants have a significant impact on gene-level and pathway-level expression, as evidenced by these results. By applying set-based rare eQTL tests, many novel significant associations were identified that would be missed with single rare variant analysis, which highlights the importance of this method. If a set-based method for rare variant eQTLs is more

widely used, numerous new findings could be found such as there have been with gene-based genetics tests for complex disease. These findings can also be utilized to add functional support and strengthen previous single rare variant findings

Results in this study are biased to over-select for AD because of the large number of AD cases in the dataset, making the findings especially noteworthy for AD. At the brain gene-level (Table S3.2), the top most significant genes include a number of genes previously linked with AD. *RNF39* ($p=7.71 \times 10^{-27}$) is an AD-associated differentially methylated region in multiple previous studies [102, 103]. A zinc finger protein *ZNF253* ($p= 7.45 \times 10^{-19}$) is an upregulated gene and transcription factor (TF) correlated with initial stages of AD [104], while *POLD2* ($p= 1.56 \times 10^{-13}$) is an up-regulated gene in a mouse tau pathology study [105], and lipid metabolism gene *ACOT1* ($p= 2.34 \times 10^{-12}$) is upregulated in the parietal cortex of patients with AD [106]. *FAM154B* ($p=7.71 \times 10^{-27}$), also known as *SAXO2*, a microtubule-stabilizing protein, is seen in gene-based testing for shared genes between AD and ischemic stroke (IS) [107]. Gene level significant rare eQTLs also include known AD genes *HLA-DRB1* and *HLA-DRB5* [58, 101] which are located in the human leukocyte antigen (HLA) region, a key susceptibility locus in many immunological diseases [108]. Moreover, as discussed 17% (11/65) of significant brain gene-level results are on chromosome 6 including other HLA genes- *HLA-DOB* and *HLA-A*, *RNF39* which lies within the major histocompatibility complex (MHC) class I region, blood and brain results

shared gene *GNMT*, and *RPL10A* (Table S3.2). *HLA-DOB* was previously found as an eGene with cell-type specific eQTLs in microglia and monocytes/macrophages [109], which are both considered myeloid cells. It is believed that a large proportion of AD genetic risk can be explained by genes expressed in myeloid cells and not by genes expressed in other cell-types [110]. Meanwhile, *RPL10A* was one of top 50 proteins that had significantly increased expression in rpAD (rapidly progressive) plaques [111]. Besides the immune related genes on chromosome 6, *IL27* ($p=1.69 \times 10^{-30}$) is also part of the cytokine family and *CARD17* ($p= 6.73 \times 10^{-13}$) is a regulatory protein of inflammasomes which are responsible for the activation of inflammatory responses [112].

In blood gene-level significant findings (Table S3.3), the most significant result was for the nuclear mitotic apparatus protein 1 gene *NUMA1* ($p= 6.01 \times 10^{-76}$) and was much more significant than the top brain gene-level significant finding (*C5orf17*, $p=4.56 \times 10^{-49}$). *NUMA1* is a common gene in hippocampus upregulated in an AD study [113], and other top significant gene *GAD1* ($p= 1.49 \times 10^{-58}$) was downregulated with reduced neuronal activity [114]. There were several top genes linked to AD, aside from known AD genes *MS4A6A*, *ECHDC3*, and *ABCA7* [58, 101]. *SOS1*(also known as *SOS-1*, $p= 3.58 \times 10^{-48}$) is increased in the pyramidal neurons of AD patients [115]. Another ribosomal protein similar to *RPL10A*, *RPS23* regulates beta-amyloid (Aβ) levels and tau phosphorylation in mice [116] and regulates synaptic plasticity in humans [117]. *KIF1B* ($p= 4.49 \times 10^{-21}$) expression at gene and protein level is significantly increased in AD, and is

associated with accelerated progression in neurodegenerative diseases [118, 119]. A potential new therapeutic target, *SFRP1* ($p= 2.16 \times 10^{-20}$) is secreted by established AD gene *ADAM10* and is significantly increased in the brain and cerebrospinal fluid (CSF) of AD patients [120]. There also appear to be a number of cytoskeleton related genes such as *NUMA1*, actin gene *ACTG1*, Kinesin Family Member 1B gene *KIF1B*, and dynein gene *DNAAF3* [121].

Five genes (*GNMT*, *LDHC*, *RBPMS2*, *DUS2*, and *HP*) were shared in gene-level results in brain and blood (Table 3.1) and of note, 4 out of 5 have clear associations to AD. *GNMT* expression is detected in the hippocampus and its deficiency results in reduced neurogenic capacity, spatial learning, and memory impairment [122]. *LDHC* has differentially methylated regions in blood in AD [123]. Overexpression of flavoprotein *DUS2* reduces Aβ42 toxicity [124]. There is a significantly higher mean serum *HP* or haptoglobin (*Hpg*) level among the AD patients compared to healthy controls [125]. *RBPMS2*, a RNA-binding protein, has not been linked to AD but was very significant in blood compared to the rest of these in both tissues ($p=1.7 \times 10^{-36}$) and has been seen in a leukocyte signature of Traumatic Brain Injury [126].

There a large number of unique low frequency and rare variants in each gene for these gene-level tests. Particular variants seem to be driving significant results in blood based on the large number of significant individual eGene-eSNP eQTL results (n=832) and because all 10 of the most significant eGenes in gene-level results are also a significant eGene in individual eGene-eSNP results.

Whereas, there are much fewer significant eGene-eSNP eQTL results in brain (n=61) and only 6/10 of the most significant gene-level results are eGenes, so the aggregate of SNPs may be more important for testing the association of low frequency and rare variants in brain.

The significant enriched pathways are representative of the AD enriched gene expression datasets (Table 3.2 + 3.3). All shared pathways in blood and brain- Wnt signaling, Apoptosis signaling, and Inflammation mediated by chemokine and cytokine signaling pathways- function in AD. In brain, the Wnt signaling pathway and in blood, the PD pathway also can be connected to AD. In addition, it is not surprising that significant pathways in blood include blood coagulation and heme biosynthesis. Examining the shared genes in the shared pathways, *HSPA6*, which appeared in the apoptosis signaling pathways, is involved in neuronal responses to proteotoxic stress in neurodegenerative diseases that have been characterized as protein misfolding disorders [127]. *TNFR10C*, also from the apoptosis signaling pathways and significantly associated in blood(Table 3.5), is inducible by DNA damage, is not expressed in the brain, but its receptors are found on neurons, astrocytes and oligodendrocytes, and has been seen in a protein signature from an AD dataset [128]. Increased levels of plasma *MMP-9*, the shared gene from the CCKR signaling maps, have been observed in AD patients and CSF levels of *MMP-9* have correlated with both CSF T-tau and P-tau in elderly controls, suggesting MMPs could be associated with neuronal degeneration [129]. These shared

genes appear to be making a larger impact on the apoptosis and CCKR signaling pathways in blood due to the more significant p-values compared to brain (Table S3.10). Looking at the complete gene modules which were enriched in the Inflammation pathway, the shared genes included *CCL3*, *CCL4*, *SPOCD1*, and *CXCL5* (Table S3.11). *CCL3* and *CCL4* could be key to microglial function in aging and disease possibly through recruiting peripheral immune cells for the CNS [130]. Specifically, *CCL3* increases *BACE1* and Aβ deposition in the brain and is expressed by peripheral T cells to enhance transmigration through the blood-brain-barrier in AD [131]. *CCL4* has been expressed in astrocytes and microglia in studies evaluating inflammation with AD, is secreted increasingly by macrophages when treated with Aβ, and has genetic associations with AD related CSF levels [132]. *SPOCD1* was a down-regulated gene in the hippocampus and entorhinal cortex tissues in a previous study [133] and *CXCL5* is a chemokine, having a large role in inflammatory diseases [134].

The significance of the inflammation pathway continues to be illustrated with the pathway level results (Table 3.4 + 3.5). In brain, 2 genes- *CCL7* and *CCL8*- were each associated with the Inflammation mediated by chemokine and cytokine signaling pathway. Both genes are chemokines which regulate immune cells involved in inflammatory responses, specifically monocytes/macrophages for *CCL7* [121]. Chemokine levels have been found be significantly changed in AD patients in serum, CSF and brain tissue [135]. There were 22 genes with rare eQTLs associated with pathway expression in blood, with a large proportion of

genes, including the most significant, in the inflammation pathway. This includes *GRB2*, which was the most significant of pathway level results ($p= 6.04 \times 10^{-07}$), and plays a preventive role in AD by protecting cytoskeletal architecture [136]. The chemokine receptor *CXCR2* produces Aβ peptides [137] while Aβ activation of fyn disrupts neuronal *PAK1* kinase and its cytoplasmic levels are decreased in moderate to severe AD [138]. Removal of another significant gene *IFNAR1* provides neuro-protection after Aβ1-42 insult, decreasing type-1 interferon (IFN) production and apoptosis, presenting type-I IFNs as a potential therapeutic target for AD [139]. FPR2 is another potential AD target that is involved in the uptake and clearance of Aβ and contributes to innate immunity and inflammation [140]. *ALOX5AP* is a microglial gene upregulated in both pathologic and normal aging [141], and is significant in both gene-level and pathway-level results in blood. Though the inflammation pathway appears 3 separate times in different gene modules in blood, only the enriched inflammation pathway in gene module 14 has genes that are significantly contributing to its pathway expression (Table 3.5). An endoplasmic reticulum (ER) chaperone heat shock protein, *HSPA6* is involved in amyloid precursor protein metabolism and neuronal death in AD [142], and is seen associated with the PD pathway, a neurodegenerative disease with clinical similarities with AD [143]. Several heat shock proteins have also previously been seen in AD [144]. An AD-related gene significantly associated with the apoptosis signaling pathway is *CFLAR*, which is upregulated at the beginning of AD and downregulated during the deterioration [145]. Most gene-

based significant results are not seen in pathway results, suggesting the regulation of gene expression is not affected by pathways and instead are downstream targets.

It is clear there is a large immune and inflammatory component to the rare eQTL gene-level and pathway-level results, and thus providing further evidence for importance of the immune system in AD. Though the inflammation pathway is seen in results in both brain and blood, the genes that are significantly contributing to pathway expression differ between the tissues. There is no overlap in the genes in the enriched gene module pathway either, regardless of significance. This suggests that different genes and biological mechanisms are affecting the inflammatory pathology of AD in blood and brain. The significant pathway-level genes have the potential to be target genes for the AD inflammatory response in blood and brain tissues.

This study has limitations. Most comparisons between brain and blood are based on independent groups in the ROSMAP brain and ADNI blood datasets. Our validation of top findings in recently released ROSMAP blood data is meant to alleviate some of this bias. Also, a limitation with brain expression findings is that it may reflect post-mortem changes unrelated to disease or cell-type different expression [11]. In addition, some findings may not be AD-related because we did not compare expression between AD cases and controls or test the interaction effect of AD and low frequency/rare variants on gene expression.

To conclude, this study identified of a number of significant gene-level and pathway level significant rare eQTL results, found additional evidence for the importance of the immune/inflammatory system in AD, and highlighted the advantages of using a set-based eQTL method for low frequency and rare variants. Future directions would involve comparing results among AD cases and controls, as well as validation with functional experiments.

Table 3.1. Intersection of significant genes between blood and brain gene-level results

ROSMAP Brain							
CHR	**BEGIN**	**END**	**NUM PASS VARS**	**NUM SING VARS**	**STATRHO**	**P-VALUE**	**GENE**
6	41942338	43929364	671	437	0	1.85E-06	GNMT
11	17434230	19468040	429	273	0	2.07E-07	LDHC
15	64039999	66063761	404	249	0	9.90E-08	RBPMS2
16	67034867	69106452	714	482	0	1.98E-06	DUS2
16	71090452	73094829	741	461	0	2.28E-09	HP
ADNI Blood							
CHR	**BEGIN**	**END**	**NUM PASS VARS**	**NUM SING VARS**	**STATRHO**	**P-VALUE**	**GENE**
6	41933605	43930372	1006	640	0	2.87E-07	GNMT
11	17434219	19472162	762	473	0	2.25E-10	LDHC
15	64039217	66064168	648	417	0	1.69E-36	RBPMS2
16	67022599	69106642	1085	723	0	6.41E-08	DUS2
16	71089779	73094855	1206	750	0	2.43E-11	HP

BEGIN POS: Beginning position of range for rare variants within 1 Mb of gene to be tested
END POS: End position of range for rare variants within 1 Mb of gene to be tested
NUM PASS VARS: Number of variants passing all thresholds for EPACTS software
NUM SING VARS: Number of singletons among variants in NUM PASS VARS
P-VALUE: P-value of burden tests
STATRHO: represents the RHO value from SKAT-O test, rho = 1 (burden) and rho = 0 (SKAT)

Table 3.2. Significant pathway enrichment in gene modules in brain

	All ROSMAP genes (REF)		Gene module genes					
PANTHER Pathways	**Gene module**	**#**	**# of genes in pathway**	**Expected Value***	**Fold enrichment**	**+/-**	**Raw P value**	**FDR**
Cadherin signaling pathway	4	127	14	3.97	3.53	+	9.59E-05	7.77E-03
Angiogenesis	4	126	12	3.94	3.05	+	9.95E-04	3.22E-02
Gonadotropin-releasing hormone receptor pathway	4	152	14	4.75	2.95	+	5.28E-04	2.14E-02
Wnt signaling pathway	4	235	21	7.35	2.86	+	3.49E-05	5.65E-03
Apoptosis signaling pathway	7	77	12	1.64	7.3	+	3.09E-07	5.01E-05
p53 pathway	7	62	7	1.32	5.29	+	5.76E-04	2.33E-02
CCKR signaling map	7	111	10	2.37	4.22	+	2.22E-04	1.20E-02
Toll receptor signaling pathway	8	32	6	0.46	12.97	+	1.45E-05	2.36E-03
Inflammation mediated by chemokine and cytokine signaling pathway	16	173	5	0.51	9.89	+	1.54E-04	2.50E-02

PANTHER = Protein ANalysis THrough Evolutionary Relationships) Classification System;
*The expected value is the # of genes expected in the gene module list for a pathway based on the reference list.

Table 3.3. Significant pathway enrichment in gene modules in blood

	All ADNI genes (REF)		Gene module genes					
PANTHER Pathways	**Gene module**	**#**	**# of genes in pathway**	**Expected Value***	**Fold enrichment**	**+/-**	**Raw P value**	**FDR**
Histamine H2 receptor mediated signaling pathway	5	24	5	0.67	7.41	+	1.03E-03	3.37E-02
Angiotensin II-stimulated signaling through G proteins and beta-arrestin	5	33	6	0.93	6.47	+	6.14E-04	3.34E-02
Ras Pathway	5	64	8	1.8	4.44	+	7.61E-04	3.10E-02
Apoptosis signaling pathway	5	112	12	3.15	3.81	+	1.51E-04	2.47E-02
CCKR signaling map	5	164	14	4.61	3.04	+	3.94E-04	3.21E-02
Heme biosynthesis	6	11	4	0.25	15.93	+	2.74E-04	4.47E-02
JAK/STAT signaling pathway	7	17	4	0.28	14.33	+	3.21E-04	2.62E-02
Integrin signalling pathway	7	180	11	2.96	3.72	+	2.84E-04	4.62E-02
B cell activation	12	66	6	0.66	9.08	+	7.89E-05	1.29E-02
PDGF signaling pathway	12	127	7	1.27	5.5	+	3.79E-04	2.06E-02
Inflammation mediated by chemokine and cytokine signaling pathway	14	237	11	2.27	4.84	+	2.50E-05	4.08E-03
Parkinson disease	15	85	7	0.79	8.82	+	2.28E-05	3.72E-03
Inflammation mediated by chemokine and cytokine signaling pathway	20	237	8	1.78	4.48	+	5.07E-04	8.26E-02
Blood coagulation	24	43	8	0.27	29.9	+	8.22E-10	1.34E-07
Inflammation mediated by chemokine and cytokine signaling pathway	24	237	7	1.47	4.75	+	7.94E-04	4.32E-02
T cell activation	32	73	4	0.18	22.02	+	3.87E-05	6.30E-03

PANTHER = Protein ANalysis THrough Evolutionary Relationships) Classification System;
*The expected value is the # of genes expected in the gene module list for a pathway based on the reference list.

Table 3.4. Pathway-level cis rare eQTL significant results in brain

CHR	BEGIN POS	END POS	NUM PASS VARS	NUM SING VARS	STATRHO	P-VALUE	GENE MODULE	PATHWAY	GENE
17	31600172	33592552	340	206	0	1.84E-05	16	Inflammation mediated by chemokine and cytokine signaling pathway	CCL7
17	31648819	336655	319	195	0	4.50E-04	16	Inflammation mediated by chemokine and cytokine signaling pathway	CCL8

BEGIN POS: Beginning position of range for rare variants within 1 Mb of gene to be tested
END POS: End position of range for rare variants within 1 Mb of gene to be tested
NUM PASS VARS: Number of variants passing all thresholds for EPACTS software
NUM SING VARS: Number of singletons among variants in NUM PASS VARS
P-VALUE: P-value of burden tests
STATRHO: represents the RHO value from SKAT-O test, rho = 1 (burden) and rho = 0 (SKAT)

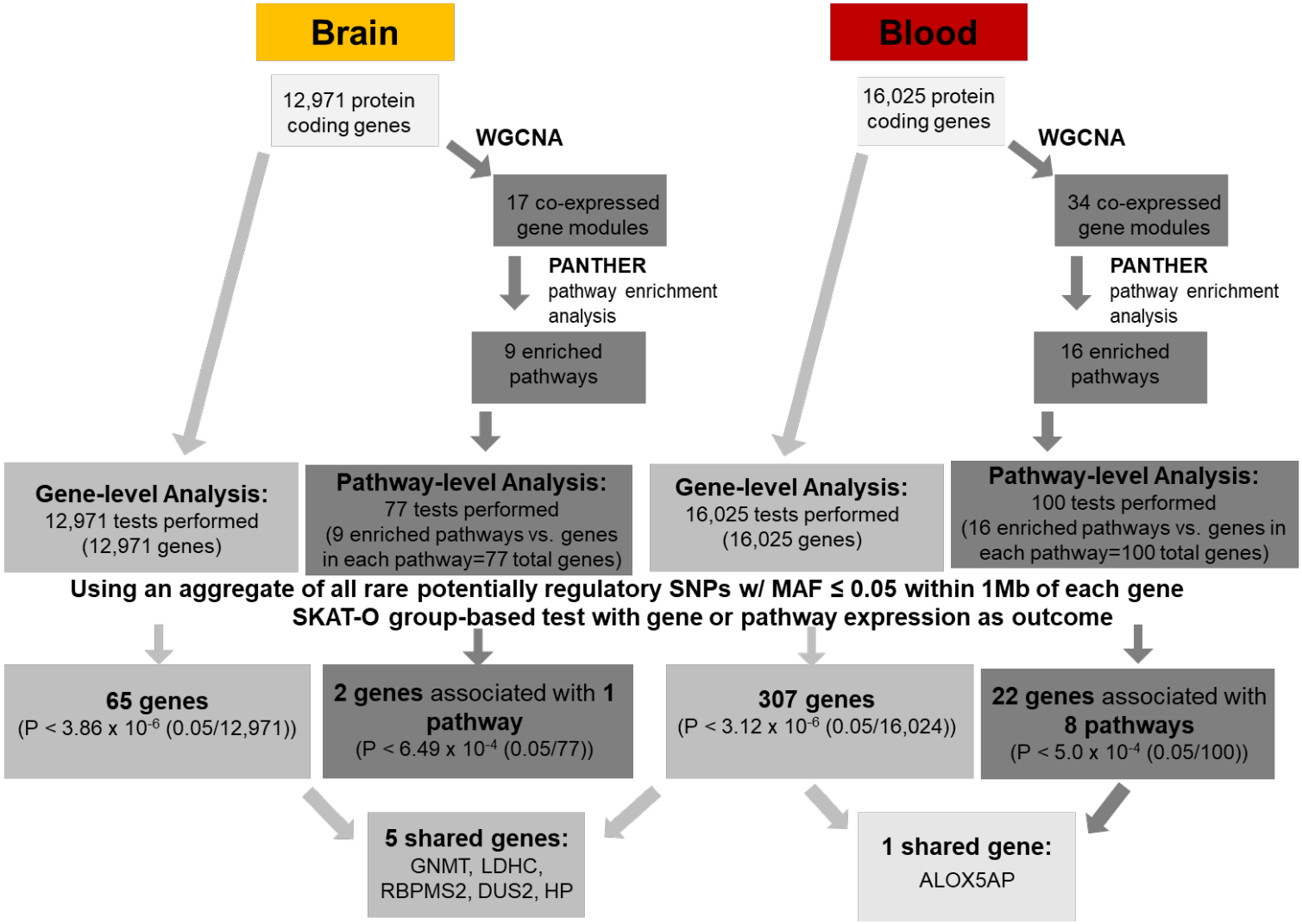
Table 3.5. Pathway-level cis rare eQTL significant results in blood

CHR	BEGIN POS	END POS	NUM PASS VARS	NUM SING VARS	STATRHO	P-VALUE	GENE MODULE	PATHWAY	GENE
17	72322351	74401630	1108	717	0.1	6.04E-07	14	Inflammation mediated by chemokine and cytokine signaling pathway	GRB2
2	217992496	220001949	943	564	0	1.53E-06	14	Inflammation mediated by chemokine and cytokine signaling pathway	CXCR2
5	174085268	176108976	335	196	0.1	9.07E-06	5	Histamine H2 receptor mediated signaling pathway	HRH2
1	25859096	27901441	1208	790	0.1	1.01E-05	5	Ras Pathway	RPS6KA1
1	25859096	27901441	1208	790	0.1	1.01E-05	5	CCKR signaling map	RPS6KA1
11	76033278	78180311	565	355	0.1	1.83E-05	5	Ras Pathway	PAK1
11	76033278	78180311	565	355	0.1	1.83E-05	5	CCKR signaling map	PAK1
21	33696834	35718581	525	332	0	1.98E-05	14	Inflammation mediated by chemokine and cytokine signaling pathway	IFNAR1
3	11628812	13702170	431	255	0	2.11E-05	14	Inflammation mediated by chemokine and cytokine signaling pathway	RAF1
1	83964144	85961982	589	336	0	3.43E-05	5	Histamine H2 receptor mediated signaling pathway	GNG5

9	115150150	117160754	620	425	0	4.97E-05	6	Heme biosynthesis	ALAD
1	44478672	46476606	1061	649	0	5.92E-05	6	Heme biosynthesis	UROD
19	13202507	15228794	767	501	0.1	7.60E-05	5	Histamine H2 receptor mediated signaling pathway	PRKACA
19	13202507	15228794	767	501	0.1	7.60E-05	5	CCKR signaling map	PRKACA
9	127005465	128998618	1235	692	0.1	7.91E-05	15	Parkinson disease	HSPA5
8	21946761	23968794	817	520	0.1	8.77E-05	5	Apoptosis signaling pathway	TNFRSF10C
19	51273985	53272173	440	280	0.1	1.23E-04	14	Inflammation mediated by chemokine and cytokine signaling pathway	FPR2
13	30317837	32332540	297	197	0.1	1.26E-04	14	Inflammation mediated by chemokine and cytokine signaling pathway	ALOX5AP
2	200984212	203030077	637	404	0	2.42E-04	5	Apoptosis signaling pathway	CFLAR
17	39458200	41463831	1093	708	0	3.08E-04	12	PDGF signaling pathway	STAT5A
14	50190597	52294891	516	323	1	3.65E-04	12	PDGF signaling pathway	NIN
12	49108257	51158233	1082	716	0.1	4.48E-04	5	Apoptosis signaling pathway	TMBIM6

BEGIN POS: Beginning position of range for rare variants within 1 Mb of gene to be tested
END POS: End position of range for rare variants within 1 Mb of gene to be tested
NUM PASS VARS: Number of variants passing all thresholds for EPACTS software
NUM SING VARS: Number of singletons among variants in NUM PASS VARS
P-VALUE: P-value of burden tests
STATRHO: represents the RHO value from SKAT-O test, rho = 1 (burden) and rho = 0 (SKAT)

Figure 3. 1. Overview of set-based rare eQTL analysis. Flowchart of the gene-level and pathway-level rare eQTL analysis in brain and blood tissues. Starting with all protein coding genes, for the gene-level analysis, tests are performed for each gene using an aggregate of all low frequency and rare potentially regulatory SNPs w/ MAF ≤ 0.05 within 1Mb of each gene. For the pathway-level analysis, WCGNA is first performed to find co-expressed gene modules and PANTHER pathway enrichment analysis is applied next to find the significant enriched pathways in these gene modules. The pathway-level tests are then performed on each enriched pathway for the aggregate of SNPs for each gene in the gene module that exists in the pathway. Significant results are found at the Bonferroni corrected threshold ($\alpha = 0.05$). Among the brain and blood gene-level results, there were 5 shared genes and among the blood gene-level and pathway-level results, there was 1 shared gene.


Brain
12,971 protein coding genes
WGCNA
17 co-expressed gene modules
PANTHER pathway enrichment analysis
9 enriched pathways
Blood
16,025 protein coding genes
WGCNA
34 co-expressed gene modules
PANTHER pathway enrichment analysis
16 enriched pathways
Gene-level Analysis: 12,971 tests performed (12,971 genes)
Pathway-level Analysis: 77 tests performed (9 enriched pathways vs. genes in each pathway=77 total genes)
Gene-level Analysis: 16,025 tests performed (16,025 genes)
Pathway-level Analysis: 100 tests performed (16 enriched pathways vs. genes in each pathway=100 total genes)
Using an aggregate of all rare potentially regulatory SNPs w/ MAF ≤ 0.05 within 1Mb of each gene
SKAT-O group-based test with gene or pathway expression as outcome
65 genes (P < 3.86 x 10-6 (0.05/12,971))
2 genes associated with 1 pathway (P < 6.49 x 10-4 (0.05/77))
307 genes (P < 3.12 x 10-6 (0.05/16,024))
22 genes associated with 8 pathways (P < 5.0 x 10-4 (0.05/100))
5 shared genes: GNMT, LDHC, RBPMS2, DUS2, HP
1 shared gene: ALOX5AP

CHAPTER 4: Identification of cell-type specific eQTLs of common variants in blood and brain tissue data

4.1 Abstract

Background: Changes in gene expression are associated with Alzheimer disease (AD) risk. Because regulation of gene expression is heritable and context-dependent, we investigated AD-related gene expression patterns in multiple cell-types in blood and brain.

Methods: Cis-expression quantitative trait locus (eQTL) mapping was performed genome-wide in blood from 5,257 Framingham Heart Study (FHS) participants and in brain donated by 475 Religious Orders Study/Memory & Aging Project (ROSMAP) participants. The association of gene expression with genotypes for all cis SNPs within genes was evaluated using linear regression models for unrelated subjects and linear mixed models for related subjects. Cell type-specific eQTL (ct-eQTL) models included an interaction term for expression of "proxy" genes that discriminate the particular cell type.

Results: A total of 173,857 eQTLs and 51,098 ct-eQTLs in brain, and 847,429 eQTLs and 30,405 ct-eQTLs in blood were significant after Bonferroni correction. Ct-eQTL analysis identified 11,649 and 2,533 significant gene-SNP eQTL pairs in brain and blood, respectively, that were not detected in generic eQTL analysis. Of note, 24,028 significant gene-SNP eQTL pairs were shared between blood and brain results and the 386 unique eGenes in these shared eQTL pairs are enriched in the apoptosis and Wnt signaling pathway. Five of these shared genes

(*HLA-DRB5*, *HLA-DRB1*, *ECHDC3*, *CR1*, and *WWOX*) are established AD loci. *HLA-DRB1* and *HLA-DRB5* are ct-eQTLs in both monocytes/macrophages and microglia, providing further evidence for the role of the immune pathway in AD. The potential importance and relevance to AD of eQTLs and ct-eQTLs in myeloid cell-types is supported by the observation that a large portion of GWS ct-eQTLs map within 1Mb of established AD loci and 58% (23/40) of the most significant eGenes in these eQTLs have previously been implicated in AD. Also, we identified several novel candidate AD genes and variants.

Conclusions: This study identified cell-type specific expression patterns for established and potentially novel AD genes, found additional evidence for the role of myeloid cells in AD risk, and discovered potential novel blood and brain AD biomarkers that highlight the importance of cell-type specific analysis.

4.2 Background

Recent expression quantitative trait locus (eQTL) analysis studies suggest that changes in gene expression have a role in the pathogenesis of AD [5, 6]. However, regulation of gene expression, as well as many biological functions, has been shown to be context-specific (e.g., tissue and cell-types, developmental time point, sex, disease status, and response to treatment or stimulus) [7-10]. One study of 500 healthy subjects found over-representation of T cell-specific eQTLs in susceptibility alleles for autoimmune disease and AD risk alleles polarized for monocyte-specific eQTL effects [13]. In addition, disease and trait-

associated cis-eQTLs were more cell type specific than average cis-eQTLs [13]. Another study classified 12% of more than 23,000 eQTLs in blood as cell-type specific [9]. Large eQTL studies across multiple human tissues have been conducted by the GTEx consortium, with a study on genetic effects on gene expression levels across 44 human tissues collected from the same donors characterizing patterns of tissue specificity recently published [146].

Microglia, monocytes and macrophages share a similar developmental lineage and are all considered to be myeloid cells [147]. It is believed that a large proportion of AD genetic risk can be explained by genes expressed in myeloid cells and not other cell-types [110]. Several established AD genes are highly expressed in microglia [147, 148] and a variant in the AD-associated locus *CELF1* has been associated with lower expression of *SPI1* in monocytes and macrophages [110]. AD risk alleles have been shown to be enriched in myeloid specific epigenomic annotations and in active enhancers of monocytes, macrophages, and microglia [149], and to be polarized for cis-eQTL effects in monocytes [13]. These findings suggest that a cell-type specific analysis in blood and brain tissue may identify novel and more precise AD associations that may help elucidate regulatory networks. In this study, we performed a genome-wide *cis* ct-eQTL analysis in blood and brain, respectively, then compared eQTLs and cell-type specific eQTLs (ct-eQTLs) between brain and blood with a focus on genes, loci, and cell-types previously implicated in AD risk by genetic approaches.

4.3 Methods

4.3.1 *Study Cohorts*

Framingham Heart Study (FHS). The FHS is a multigenerational study of health and disease in a prospectively followed community-based and primarily non-Hispanic white sample. Procedures for assessing dementia and determining AD status in this cohort are described elsewhere [150]. Clinical, demographic, and pedigree information, as well as 1000 Genomes Project Phase 1 imputed SNP genotypes and Affymetrix Human Exon 1.0 ST array gene expression data from whole blood, were obtained from dbGaP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000007.v31.p12) . Requisite information for this study was available for 5,257 participants. Characteristics of these subjects are provided in Table S4.1.

Religious Orders Study (ROS)/ Memory and Aging Project (MAP). ROS enrolled older nuns and priests from across the US, without known dementia for longitudinal clinical analysis and brain donation and MAP enrolled older subjects without dementia from retirement homes who agreed to brain donation at the time of death [94, 95]. RNA-sequencing brain gene expression and whole-genome sequencing (WGS) genotype data were obtained from the AMP-AD knowledge portal (https://www.synapse.org/#!Synapse:syn3219045) [96].

4.3.2 *Data Processing*

Generation and initial quality control (QC) procedures of the FHS GWAS and expression data are described elsewhere and include all genotype QC and pre-adjustment of gene expression levels for batch effects and other technical covariates [150]. ROSMAP gene expression data were log-normalized and adjusted for known and hidden variables detected by surrogate variable analysis (SVA) [97] in order to determine which of these variables should be included as covariates in analysis models for eQTL discovery. Additional filtering steps of FHS and ROSMAP GWAS and gene expression data included eliminating subjects with missing data, restricting gene expression data to protein coding genes, and retaining common variants (MAF ≥ 0.05) with good imputation quality ($R^2 \geq 0.3$).

4.3.3 *Cis eQTL Mapping*

Cis-eQTL mapping was performed using a genome-wide design (Figure S4.1). The association of gene expression with SNP genotypes for all cis SNPs within1 Mb of protein-coding genes was evaluated using linear mixed models adjusting for family structure in FHS and linear regression models for unrelated individuals in ROSMAP. In FHS, lmekin function in the R kinship package (version 1.1.3) [151] was applied assuming an additive genetic model with covariates for age and sex, and family structure modeled as a random-effects term for kinship - a matrix of kinship coefficients calculated from pedigree structures. The linear model for analysis of FHS can data be expressed as follows:

$$Y_i = I + \beta_1 G_j + \beta_2 A_{ij} + \beta_3 S_{ij} + U_{ij} + \varepsilon_{ij}$$

where Y_i is the expression value for gene *i*, G_j is the genotype dosage for cis SNP j, *Aij* and S_{ij} are the covariates for age and sex respectively, U_{ij} is the random effect for family structure, and β_1, β_2, and β_3 are regression coefficients.

ROSMAP data were analyzed using the lm function in the base stats package in R [98]. The regression model, which included covariates for age, sex, post-mortem interval (PMI), study (ROS or MAP), and a term for a surrogate variable (SV1) derived from analysis of high dimensional data, can be expressed as:

$$Y_i = I + \beta_1 G_j + \beta_2 A_{ij} + \beta_3 S_{ij} + \beta_3 S_{ij} + \beta_4 PM_{ij} + \beta_{5ij} S2 + \beta_{6ij} SV1 + \varepsilon_{ij}$$

where Y_i is the expression value for gene *i*, G_j is the genotype dosage for cis SNP j, *Aij,* S_{ij}, PM_{ij}, $S2_{ij}$, and $SV1_{ij}$ are the covariates for age, sex, PMI, study and SV1 respectively, ε_{ij} is the residual error, and the βs are regression coefficients.

4.3.4 *Cis ct-eQTL Mapping*

Models testing associations with cell type-specific eQTLs (ct-eQTLs) included an interaction term for expression levels of “proxy” genes that represent cell types. Proxy genes representing 10 cell types in whole blood [9] and five cell types in brain [152-154] were incorporated in cell type-specific models (Table 4.1). These proxy genes for cell types in blood were established previously using BLUEPRINT expression data to validate cell-type-specific expression in each cell-type module [9] and the proxy genes for brain cell types have been

incorporated in several studies [152-154]. Cell type-specific expression analyses in blood of FHS participants were conducted using the following model:

$$Y_i = I + \beta_1 G_j + \beta_2 P + \boldsymbol{\beta_3 (P^* G_j)} + \beta_4 A_{ij} + \beta_5 S_{ij} + U_{ij} + \varepsilon_{ij}$$

where in each $eQTL_{ij}$ pair, Y_i is the eQTL expression value for gene *i*, G_j is the genotype dosage for cis SNP j, P is the proxy gene, $\mathbf{P^* G_j}$ is the interaction term representing the effect of genotype in a particular cell type, *Aij* and S_{ij} are covariates for age and sex respectively, U_{ij} is the random effect for family structure, and βs are regression coefficients. Models with significant interaction terms indicate cell type specific eQTLs.

The following model was used to evaluate cell type-specific expression in brain in ROSMAP:

$$Y_i = I + \beta_1 G_j + \beta_2 P + \boldsymbol{\beta_3 (P^* G_j)} + \beta_4 A_{ij} + \beta_5 S_{ij} + \beta_6 PM_{ij} + \beta_{7ij} S2 + \beta_{8ij} SV1 + \varepsilon_{ij}$$

where in each $eQTL_{ij}$ pair, variables Y_i, G_j, P, *Aij,* S_{ij}, P_{ij}, ε_{ij} and βs are as described above, and PM_{ij}, $S2_{ij}$, and $SV1_{ij}$ are covariates for PMI, study, and SV1 respectively.

A Bonferroni correction was applied to determine the significance threshold for each analysis (Table S4.3).

4.3.5 *Selection of AD-related eQTLs and Gene-set Pathway Enrichment Analysis*

To assess significant eQTLs and ct-eQTLs for relevance to AD, eGenes (genes whose expression levels are associated with variation at a particular eSNP) were

matched to 88 genes located near 80 distinct uncorrelated SNPs that have been associated with AD or AD-related traits by genetic association or experimental approaches (Table S4.2) and eSNPs (SNPs that significantly influence gene expression) under the 80 significant association peaks. Gene-set enrichment analysis was performed using the PANTHER (Protein ANalysis THrough Evolutionary Relationships) software tool [31] to determine if the unique genes in the significant eQTL/ct-eQTL pairs shared by both brain and blood datasets are associated with a specific biological process or molecular function. Significance of the pathways was determined by the Fisher's Exact test with False discovery rate (FDR) multiple test correction.

4.3.6 *Colocalization Analyses*

Assessment of causal variants shared by adjacent GWAS and eQTL signals was performed using a Bayesian colocalization approach implemented in the R package *coloc* [155]. This analysis incorporated information about significantly associated variants for AD risk obtained from a recent large GWAS [101] and lead eQTL variants each defined as the eSNP showing the strongest association with gene expression. Following recommended guidelines, the variants were deemed to be colocalized by a high posterior probability that a single shared variant is responsible for both signals (PP4 > 0.8) [155, 156]. A lower threshold for statistical significance with a false discovery rate (FDR) < 0.05 for eQTL significant results was applied to maximize detection of colocalized pairs. Regional plots were constructed with LocusZoom [157].

4.4 Results

4.4.1 *Shared eQTLs and ct-eQTLs in blood and brain*

A total of 173,857 eQTLs and 51,098 ct-eQTLs in brain, and 847,429 eQTLs and 30,405 ct-eQTLs in blood were significant after Bonferroni correction (Table S4.3). Additional significant gene-SNP eQTLs pairs in brain (n=11,649) and blood (n=2,533) were observed in ct-eQTL analysis that were not detected in eQTL analysis (Figure 4.1A). Of note, 24,028 significant gene-SNP eQTL pairs were shared between blood and brain. The 386 distinct eGenes among these shared eQTL pairs (Table S4.4) are most enriched in the apoptosis signaling (p=0.023) and Wnt signaling (p=0.036) pathways (Table S4.5). Five of these eGenes (*HLA-DRB5*, *HLA-DRB1*, *ECHDC3*, *CR1*, and *WWOX*) were previously associated with AD [58, 101]. Three eSNPs in eQTLs involving *HLA-DRB1*/*HLA-DRB5* (rs9271058) and *ARL17A*/*LRRC37A2* (rs2732703 and rs113986870) which are near *KANSL1* and *MAPT* were previously associated with AD risk at the genome-wide significance level [56, 101] (Table 4.2).

Notably, the eQTLs involving *CR1, ECHDC3* and *WWOX* were much more significant in brain than blood and their tissue-specific effects were in opposite directions. *ECHDC3* is a significant eGene in blood and brain eQTLs (specifically in neurons). *HLA-DRB5* and *HLA-DRB1* are the only eGenes ascribed to significant ct-eQTLs in both blood and brain noting that of the 10 distinct lead eSNPs, five are unique to each tissue (Table 4.2). Although the eQTLs involving these genes with the largest effect were observed in blood across multiple cell

types, the total number of significant eSNP-eGene combinations was far greater in brain (particularly in microglia and neurons). The only instance in which the lead eSNP is also associated with AD risk at the GWS level was observed in the blood eQTL pair of *HLA-DRB1* with eSNP rs9271058 (Table 4.2A). Among the AD-associated SNPs at the GWS level, rs9271058 is a significant eSNP for *HLA-DRB1* in both blood and brain cell types (the most significant association by p-value was observed in anti-bacterial cells and microglia) and rs9271192 is a significant ct-eQTL for the gene in multiple brain cell types (Table 4.2). Both of these SNPs are also eSNPs for *HLA-DB5* in the brain in neurons only.

There were 657 gene-SNP eQTL pairs comprising 16 unique eGenes that were significant in blood and brain overall as well as in specific cell types in both blood and brain (Table S4.6). None of these eGenes were observed in significant pathways enriched for AD genes, however, they included AD-associated genes *HLA-DRB1* and *HLA-DRB5*.

4.4.2 *AD associated and co-localized results*

Slightly more than half (42/80 = 52.5%) of the established AD associations (Table S4.2) are eGene targets for significant eQTLs in blood (Table S4.7). By comparison, only seven established AD loci were eGene targets for significant eQTLs in brain, among which *OARD1* was significant in endothelial cells only (Table S4.7). Many GWS SNPs for AD risk are eSNPs affecting expression of the nearest gene, which is usually recognized as the causative gene, but several GWS SNPs target other genes (Table S4.8). For example, AD-associated eSNPs

rs113986870 and rs2732703 in the *MAPT/KANSL1* region target *ARL17A* in blood, but are paired in seven of eight eQTLs and ct-eQTLs with *LRRC37A2* in brain (Table S4.8). *HLA-DRB1* is the only AD gene with a significant ct-eQTL in blood, whereas many AD genes have significant blood eQTLs. In brain, only four AD loci (*CR1, HLA-DRB1/DRB5, IQCK* and *MAPT/KANSL1*) have significant brain eQTLs of which *HLA-DRB1/DRB5* and *MAPT/KANSL1* are the only brain ct-eQTLs, noting that all are significant in microglia, neurons and endothelial cells.

Next, we evaluated whether the most significant eSNPs and SNPs genome-wide significantly associated with AD status (i.e., AD-SNPs) co-localize and thus to identify a single shared variant responsible for both signals (Posterior probability of shared signals (PP4) > 0.8). This analysis revealed eight eQTL/ct-eQTL signals that colocalized with seven AD GWAS signals and half of the co-localized signals involved a ct-eQTL (Table 4.3 and Figure S4.2). Two different eSNPs for *CD2AP*, rs4711880 (eQTL $p=1.4x10^{-104}$) and rs13201473 (ct-eQTL $p= 1.47x10^{-9}$), flank *CD2AP* GWAS SNP rs10948363 which is also the second most significant eQTL ($p=2.32x10^{-104}$) and the second most significant ct-eQTL in NK cells / CD8+ T-Cells ($p=2.66x10^{-9}$). These three SNPs span a 9.0 kb region in intron 2 and are in complete linkage disequilibrium (LD, $r^2 =1.0$), indicating that any one or more of them could affect the function of target gene CD2AP. Rs6557994 is the most significant eSNP for and located in *PTK2B* (blood ct-eQTL $p=2.58x10^{-9}$) and is moderately correlated with the *PTK2B* GWAS SNP

(rs28834970, r^2=0.78, p=1.58x10^{-9}). Thus, it is not surprising that rs6557994 is also significantly associated with AD risk (p=8.19x10^{-7}). Rs6557994 is also correlated with a GWAS SNP in *CLU*, located approximately 150 kb from PTK2B, that is not significantly associated with expression of any gene. Because *PTK2B* and *CLU* are independent AD risk loci [58], it is possible that this eSNP has an effect on AD pathogenesis through independent pathways (Figure S4.2). The most significant eSNP in *MADD* (rs35233100, p=2.88x10^{-10}) was predicted to have functional consequences because it is a stop-gained mutation. This brain eQTL is colocalized (PP4=0.95) and weakly correlated with a GWAS SNP (p=1.91x10^{-5}) in *CELF1* rs10838725 (r^2=0.12).

4.4.3 *Significant results in myeloid cells*

Examination of the distribution of the significant ct-eQTL results genome-wide showed that nearly two-thirds of the ct-eQTLs in blood occurred in interferon response/anti-bacterial cells in blood, whereas brain ct-eQTLs are highly represented in endothelial cells, neurons and microglia (Figure 4.1B, Table S4.9). Further examination of significant results within myeloid cell lineages (i.e., microglia and monocytes/ macrophages) which account for a large proportion of the genetic risk for late-onset AD [110] revealed that 3,234 or 10.6% of all significant ct-eQTLs in blood were in monocytes/macrophages. This subset includes 128 unique eGenes which are significantly enriched in the AD amyloid secretase pathway (FDR p=0.013, Table S4.10). A total of 974 or 30.1% of ct-eQTLs including 4 of the 20 most significant eGenes in monocytes/macrophages

are located within 1 Mb of established AD loci (Table 4.4A). One of the eGenes in this top-ranked group (*HLA-DRB5*) is an established AD locus, and three others that are near established AD loci (*DLG2* near *PICALM* [158], *C4BPA* near *CR1* [159], and *MYO1E* near *ADAM10* [160]) are reasonable AD gene candidates based on evidence using non-genetic approaches. Microglia accounted for 15,560 (30.5%) of significant ct-eQTLs in the brain (Table 4.9) which involved 304 unique eGenes. Approximately 52% of significant ct-eQTLs in microglia are located in AD regions including five of the 20 most significant ct-eQTLs in this group (Table 4.4B). One of these five eGenes is an established AD locus (*HLA-DRB1*) and two others (*ALCC* [161] and *WNT3* [162]) have been linked to AD in previous studies.

Considering significant eGene-eSNP pairs in myeloid cell types, 251 pairs including five distinct eGenes (*BTNL3*, *FAM118A*, *HLA-DOB*, *HLA-DRB1*, and *HLA*-DRB5) are shared between microglia and monocytes/macrophages (Table 4.5A and Figure 4.2A). Three of these pairs involving eSNPs rs3763355, rs3763354, and rs1183595100 have the same target gene *HLA-DOB* and occur only in microglia and monocytes/macrophages (Table 4.5B). Among the significant ct-eQTLs in brain, the cell types with the largest proportion that were also significant in monocytes/macrophages were microglia (1.6%) and neurons (1.3%) (Table 4.5C). Conversely, among the significant ct-eQTLs in blood, the cell types with the largest proportion that were also significant in microglia were NK/CD+ T-cells (12.9%) and monocytes/macrophages (7.8%). Among ct-eQTLs

which are significant only for one cell-type each in blood and one in brain, monocytes/ macrophages shared three ct-eQTLs with microglia but with no other brain cell-types (Figure 4.2B, Table 4.5C). By comparison, microglia shared 63 ct-eQTLs with interferons/ anti-bacterial cells, but with no other blood cell types. The much larger number of ct-eQTLs in microglia that were common with interferons/bacterial cells than monocytes/ macrophages may reflect the substantially greater proportion of significant eQTLs in blood involving interferons/antibacterial cells (64%) than monocytes/macrophages (10.6%) (Table S4.9). The only other ct-eQTLs that were unique to a pair of cell types in brain and blood cell type involved neurons paired with neutrophils (n=3) and with interferons/anti-bacterial cells (n=65) (Figure 4.2B).

4.5 Discussion

We identified several novel AD-related eQTLs that highlight the importance of cell-type dependent context. It is noteworthy that there were more significant ct-eQTLs in brain (n=51,098) than blood (n=30,405) even though the dataset containing expression data from blood (FHS) is several times larger than the brain expression dataset (ROSMAP). This could be due to greater cell type heterogeneity in brain, the enrichment of AD cases in the ROSMAP dataset who may show different patterns of gene expression compared to persons without AD, or highly variable gene expression across cell-types in the nervous system [163]. Because expression studies in brain are often constrained by the small

number specimens compared to studies in other tissues, post-mortem changes that may affect gene expression in brain [11], and the growing recognition that AD is a systemic disease [164-166] , incorporating expression data from multiple tissues can enhance discovery of additional genetic influences on AD risk and pathogenesis.

Although most significant findings were tissue-specific, the 386 distinct eGenes among more than 24,000 significant gene-SNP eQTL pairs that were shared between blood and brain were enriched in the apoptosis signaling pathway that contributes to much of the underlying pathology associated with AD [167, 168]. Five established AD genes (*CR1*, *ECHDC3*, *HLA-DRB1*, *HLA-DRB5*, and *WWOX* [58, 101]) were shared eGenes in brain and blood and could be playing a key role in the systemic AD mechanisms. The complement receptor 1 (*CR1*) gene encodes a transmembrane glycoprotein functioning in the innate immune system by promoting phagocytosis of immune complexes, cellular debris, and Aβ [169]. *CR1* is an eGene for several eSNPs, including AD GWAS peak rs6656401 located within the gene, in brain and blood eQTLs and the effects on *CR1* expression are opposite in blood and brain. There are multiple possible explanations for the effect direction differences across tissues. The effect of the eSNP rs6656401 on *CR1* expression may be developmental, noting that the average age of the FHS subjects (group with expression data in blood) is more than 30 years younger than the ROSMAP subjects (group with expression data in brain). The difference between brain and blood may also reflect post-

mortem changes in brain that are not indicative of expression *in vivo*. Alternatively, these effects may be related to AD because few FHS subjects were AD cases at the time of blood draw whereas 60% of subjects in the ROSMAP sample are AD cases. This idea is supported by the observation of a larger and positive effect of rs6656401 on *CR1* expression in AD (β=0.020) compared to control brains (β=-0.0086). GWS variants located in the region spanning *ECHDC3* and *USP6NL* have previously been associated with AD [170]. We found that *ECHDC3* is the target gene for eSNP rs866770710 located in its promoter region, and this eQTL was significant in brain and specifically in neurons. Altered *ECHDC3* expression in AD brains [171] supports the idea that this gene has a role in AD. Knockout of *WWOX* in mice leads to aggregation of amyloid-β (Aβ) and Tau, and subsequent cell death [172, 173].

The human leukocyte antigen (HLA) region is the key susceptibility locus in many immunological diseases and many associations have been reported between neurodegenerative diseases and HLA haplotypes [108]. In addition, the most widely used marker to examine activated microglia in normal and diseased human brains is *HLA-DR* and microglia activation increases with the progression of AD [174, 175]. *HLA-DRB5* and *HLA-DRB1* have been implicated in numerous GWAS studies as significantly associated with AD risk [58, 101] and appeared frequently among significant results in blood and brain in this study. Rs9271058, which is located approximately 17.8 kb upstream of *HLA-DRB1*, is significantly associated with AD risk ($p=5.1 \times 10^{-8}$ [101]) and when paired with *HLA-DRB1* is

a significant eQTL and ct-eQTL in multiple cell types in blood and brain including myeloid lineage cells (i.e., monocytes/macrophages and microglia). This eSNP is also a significant eQTL in brain and specifically in neurons when paired with *HLA-DRB5*. Rs9271192, which is adjacent to rs9271058 and also significantly associated with AD risk ($p=2.9 \times 10^{-12}$ [58]), is a significant eQTL and ct-QTL with multiple cell types in brain but not blood when paired with *HLA-DRB5 and HLA-DRB1*.

Significant associations for AD have been reported with variants spanning a large portion of the major histocompatibility (MHC) region in HLA class I, II and III loci [108, 176, 177]. While the strongest statistical evidence for association in this region is with variants in *HLA-DRB1* [101], fine mapping in this region suggests that a class I haplotype (spanning the HLA-A and HLA-B loci) and a class II haplotype (including variants in *HLA-DRB1, HLA-DQA1* and *HLA-DQB1*) are more precise markers of AD risk. Given the complexity of the MHC region and extensive LD, further work is needed to confirm this is a true eQTL or a signal generated from a specific HLA allele or HLA haplotype. Although functional studies may be required to discern which HLA variants have functional consequences relevant to AD and methods accounting for HLA polymorphisms would be required to detect the differential gene expression between the HLA alleles, our findings support a role for the immune system in AD [164, 178] and the hypothesis that a large proportion of AD risk can be explained by genes expressed in myeloid cells [110].

The potential importance and relevance to AD of eQTLs and ct-eQTLs in myeloid cell-types is supported by the observation that a large portion of GWS ct-eQTLs we identified map within 1 Mb of established AD loci, and 58% (12/20 in monocytes/ macrophages and 11/20 in microglia) of the most significant eGenes have been previously implicated in AD (Table 4.4). *DLG2* encodes a synaptic protein whose expression was previously reported as down-regulated in an AD proteome and transcriptome network [179] and inversely associated with AD Braak stage [158]. Genome-wide significant associations of AD risk with *PTPRG* was observed in a family-based GWAS [180] and with *CLNK* in a recent large GWAS for which the evidence was derived almost entirely with a proxy AD phenotype in the UK Biobank [181]. *NFXL1* is a novel putative substrate for *BACE1*, an important AD therapeutic target [182]. *FCRL5* may interact with the *APOE*E2* allele and also modifies AD age of onset [183]. *FMOD* is a putative CSF autoantibody biomarker for AD [184]. *INPPF* has been linked to AD and Parkinson disease (PD) [185]. *C4BPA* was shown to be a consistently down-regulated in MCI and AD patients, and the protein encoded by this gene accumulates in Aβ plaques in AD brains [159, 186]. Lower levels of the *PAM* have been observed in the brains and CSF of AD patients compared to healthy controls [187] and *MYO1E* is expressed by anti-inflammatory disease associated microglia [160]. As a calcium channel protein, *CACNB2* may affect AD risk by altering calcium levels which could cause mitochondrial damage and then induce apoptosis [188, 189].

Likewise, several eGenes of top-ranked ct-eQTLs in microglia that are not established AD loci (Table 4.4) may have a role in the disease. It was shown that copy number variants (CNVs) near *HNRNPCL1* overlapped the coding portion of the gene in AD cases but not controls [190]. A region of epigenetic variation in *ALLC* was associated with AD neuropathology [161]. *FAM21B*, a retromer gene in the endosome-to-Golgi retrieval pathway, was associated with AD in a candidate gene study [191]. Vacuolar sorting proteins genes in this pathway including *SORL1* have been functionally linked to AD through trafficking of Aβ [55]. One study demonstrated that *WNT3* expression in the hippocampus was increased by exercise and alleviated AD-associated memory loss by increasing neurogenesis [162]. Expression of *RPL9* is downregulated in severe AD [192] and significantly differs by sex among persons with the *APOE* ε4 allele [193]. *XRCC2* was one of the candidate genes associated by linkage to AD [194]. GWS associations of variants in *DEFB121* with regional brain volumes have been observed [195]. Significant evidence of association with a *TRIM49B* SNP was found in a genome-wide pleiotropy GWAS of AD and major depressive disorder (MDD) [196]. *TMPRSS9* is a differentially expressed gene in mitochondrial function in AD [197] and is also associated with PD [198].

HLA-DOB, which is one of the five distinct eGenes (*BTNL3*, *FAM118A*, *HLA-DOB*, *HLA-DRB1*, and *HLA-DRB5)* for significant ct-eQTLs shared between microglia and monocytes/macrophages, and is the target gene for three eSNPs (rs3763355, rs3763354, and rs1183595100) that were evident only in these

myeloid cell types. These eSNPs have similar eQTL p-values in both cell types, but have slightly larger effect sizes in monocytes (Figure 4.2). The effect of rs3763355, which though not GWS, has previously been associated with AD (p= 1.19E-03) [101], on expression is in opposite directions in monocytes and microglia which suggests *HLA-DOB* may be acting in different immune capacitates in AD in blood and brain. Though the functions of the genes *BTNL3* and *FAM118A* are unknown, a *BTNL8-BTNL3* deletion has been correlated with TNF and ERK1/AKT pathways, which have an important role in immune regulation inducing inflammation, apoptosis, and proliferation, suggesting the deletion could be correlated to inflammatory disease [199]. This suggests that the majority of the shared myeloid cell types genes- the *HLA* genes and possibly *BTNL3*, are all immune-related. Ct-eQTLs involving microglia and monocytes/macrophages had a larger proportion of total intersection, an isolated set interaction and a statistically significant overlap (p<1.0E-314), demonstrating a stronger connection than other brain/blood cell types in this study and thus providing further evidence for importance of the immune system in AD.

The proportions of significant ct-eQTLs in NK cells/CD8+T cells, monocytes/macrophages, and eosinophils are comparable to those observed in reference blood tissue [200, 201]. Similarly, significant eQTL distributions in endothelial cells, neurons, and glia are consistent with reference brain tissue [202]. The majority of significant blood eQTLs were type I interferon response cells which cross-regulate with pro-inflammatory cytokines that drive

pathogenesis of autoimmune diseases including AD and certain heart diseases [203-205] and the enrichment of interferon ct-eQTLs in this study could possibly be due to the high proportion of subjects these diseases in the FHS dataset. In contrast, the proportion of significant ct-eQTLs among glial cells is much lower in astrocytes and oligodendrocytes and much higher (3 times as many) in microglia than in reference brain tissue [202].. Because many AD risk genes are expressed in myeloid cells including microglia [110], the large number of microglia ct-eQTLs is consistent with the high proportion of AD subjects in the ROSMAP dataset.

Several SNPs previously reported to be associated with AD at the GWS level were associated with eGenes that differ from genes ascribed to AD loci and thus may have a role in AD (Table S4.8). Karch et al. observed that the expression of *PILRB*, which is involved in immune response and is the activator receptor to its inhibitory counterpart *PILRA,* an established AD gene [206, 207], was highest in microglia [148]. *CNN2*, the eGene for eSNP rs4147929 located near the end of *ABCA7,* significantly alters extracellular Aβ levels in human induced pluripotent stem cell-derived neurons and astrocytes [208]. Rs4147929 also targeted *HMHA1* which plays several roles in the immune system in an HLA-dependent manner [209]. The eSNP/GWAS SNP rs3740688 located in *SPI1* also affects expression of *MYBPC3* and *MADD*. *MYBPC3* was recently identified as a target gene for eSNPs located in *CELF1* and *MS64A6A* in a study of eQTLs in blood for GWS AD loci [210]. *MADD* is expressed in neurons [148], is involved in neuronal cell death in the hippocampus [211], and was shown to be a tau toxicity

modulator [212]. Although eSNP rs113986870 in *KANSL1* when paired with the nearby eGene *LRRC37A2* was a significant brain eQTL and ct-eQTL, *LRRC37A2* encodes a leucine rich repeat protein that is expressed primarily in testis and has no apparent connection to AD. However, rs113986870 also significantly influenced expression of another gene in this region, *ARL17A*, that was previously linked to progressive supranuclear palsy by analysis of gene expression and methylation [213].

Our study has several noteworthy limitations. The proxy genes individually or collectively may not accurately depict cell-type specific context. In addition, the comparisons of gene expression in blood and brain may yield false results because they are based on independent groups ascertained from a community-based longitudinal study of health (FHS – blood) and multiple sources for studies of AD (ROSMAP – brain) which were not matched for age, sex, ethnicity and other factors which may affect gene expression. Also, findings in brain may reflect post-mortem changes unrelated to disease or cell-type different expression [11]. Moreover, some cell types are vastly under-represented compared to others. Because myeloid cell types are represented in only a small proportion of the total cell populations in brain and blood (generally ~10%), it is difficult to identify myeloid-specific signals [149]. Despite this limitation, many of the most significant and noteworthy results of this study involved monocytes/macrophages and microglia. Finally, some findings may not be AD-related because we did not compare expression between AD cases and controls.

In summary, our observation of cell-type specific expression patterns for established and potentially novel AD genes, finding of additional evidence for the role of myeloid cells in AD risk, and discovery of potential novel blood and brain AD biomarkers highlight the importance of cell-type specific analysis. Future studies that compare cell type specific differential gene expression among AD cases and controls using single cell RNA-sequencing or cell count data, as well as functional experiments, are needed to validate and extend our findings.

Table 4.1: Proxy genes for cell types in blood and brain

Proxy gene	Tissue	Cell-type	Function/GO biological process
STX3	Blood	Neutrophils 1	Detection of bacterium
FAM102A	Blood	CD4+ T-Cells	T cell selection
SAMD3	Blood	NK cells / CD8+ T-Cells	Cellular defense response
TSPAN5	Blood	Erythrocytes	Hemoglobin metabolic process
PATL1	Blood	Monocytes / Macrophages	Defense response to virus
PACS1	Blood	Unknown	Unknown/ Nerve growth factor receptor signaling pathway
SP140	Blood	Interferon response(+)/ Anti-bacterial(-)	Type 1 interferon response/Anti-bacterial, Regulation of defense response
FBXL5	Blood	Neutrophils 2	Detection of bacterium
STRBP	Blood	B-cells	B cell receptor signaling pathway
SIGLEC8	Blood	Eosinophils	Regulation of myeloid leukocyte mediated immunity
CD34	Brain	Endothelial Cells	Lining blood vessels
RELN	Brain	Neurons	Main signaling units
CD68	Brain	Microglia	Brain immune cells
GFAP	Brain	Astrocytes	Support neuronal growth, function and neurotransmitter recycling
OLIG2	Brain	Oligodendroglia	Insulating neuronal axons

Table 4.2: eQTLs and ct-eQTLs in established AD loci appearing in both blood and brain

A. eQTLs and ct-eQTLS in established AD genes in both blood and brain

eGene	Tissue	Cell-type	Lead eSNP	Position	MAF	Beta	Std Error	P-value	# of total significant eSNPs in gene/cell-type	AD GWAS peaks
CR1	Blood	NA	rs7533408	1:207673631	0.25	0.059	0.006	3.60E-22	169	NA
HLA-DRB5	Blood	NA	rs9269008	6:32436217	0.17	-2.580	0.057	<1.0E-314	72	NA
HLA-DRB1	Blood	NA	rs9271058	6:32575406	0.14	-2.950	0.028	<1.0E-314	630	Lead eSNP
ECHDC3	Blood	NA	rs11257290	10:11780324	0.28	0.041	0.005	2.91E-19	115	NA
WWOX	Blood	NA	rs7202722	16:78282458	0.40	0.023	0.003	2.60E-14	45	NA
HLA-DRB5	Blood	Interferon response(+)/ Anti-bacterial(-)	rs9269047	6:32438783	0.12	-7.120	0.335	3.04E-100	9 [all (-)]	NA
HLA-DRB5	Blood	Monocytes/ Macrophages	rs9269047	6:32438783	0.12	-11.600	1.030	2.02E-29	1	NA
HLA-DRB5	Blood	NK cells / CD8+ T-Cells	rs9269047	6:32438783	0.12	-7.660	0.994	1.30E-14	1	NA
HLA-DRB1	Blood	NK cells / CD8+ T-Cells	rs9270928	6:32572461	0.15	-4.070	0.377	3.60E-27	287	rs9271058
HLA-DRB1	Blood	Eosinophils	rs9270994	6:32574250	0.14	-2.700	0.415	7.72E-11	42	NA
HLA-DRB1	Blood	Interferon response(+)/ Anti-bacterial(-)	rs9271147	6:32577385	0.14	-5.510	0.250	1.19E-107	346 [260(-)/86(+)]	rs9271058
HLA-DRB1	Blood	Monocytes/ Macrophages	rs9271148	6:32577442	0.13	-6.110	0.709	6.83E-18	222	rs9271058
CR1	Brain	NA	rs12037841	1:207684192	0.17	-0.096	0.007	9.25E-44	64	rs6656401
HLA-DRB5	Brain	NA	rs3117116	6:32367017	0.12	-2.780	0.070	<1.0E-314	10537	rs9271058, rs9271192
HLA-DRB1	Brain	NA	rs73399473	6:32538959	0.26	-2.050	0.058	8.78E-272	10792	rs9271058, rs9271192
ECHDC3	Brain	NA	rs866770710	10:11784320	0.0002	-0.252	0.018	4.61E-44	45	NA
WWOX	Brain	NA	rs12933282	16:78124987	0.45	-0.133	0.017	1.13E-15	75	NA
HLA-DRB5	Brain	Microglia	rs67987819	6:32497655	0.14	-1.900	0.137	9.82E-44	754	NA

HLA-DRB5	Brain	Endothelial Cells	rs67987819	6:32497655	0.14	-2.410	0.220	6.32E-28	343	NA
HLA-DRB1	Brain	Microglia	rs72847627	6:32538512	0.28	-2.130	0.125	4.15E-65	2305	rs9271058, rs9271192
HLA-DRB1	Brain	Neurons	rs115480576	6:32538570	0.26	-2.210	0.153	2.72E-47	3263	rs9271058, rs9271192
HLA-DRB1	Brain	Endothelial Cells	rs9269492	6:32542924	0.30	-2.250	0.243	2.06E-20	351	rs9271192
HLA-DRB5	Brain	Neurons	rs9270035	6:32553446	0.14	-2.520	0.137	1.46E-75	2540	rs9271058, rs9271192
ECHDC3	Brain	Neurons	rs866770710	10:11784320	0.0002	0.328	0.045	3.13E-13	2	NA

B. eQTLs and ct-eQTLs involving AD GWAS association peak SNPs in both brain and blood

eGene	Tissue	Cell-type	eSNP+GWAS SNP	Position	MAF	Beta	Std Error	P-value
HLA-DRB1	Blood	NA	rs9271058	6:32575406	0.27	-2.950	0.028	<1.0E-314
ARL17A	Blood	NA	rs2732703	17:44353222	0.21	0.147	0.023	5.95E-11
ARL17A	Blood	NA	rs113986870	17:44355683	0.09	0.166	0.025	2.30E-11
HLA-DRB1	Blood	Interferon response(+)/ Anti-bacterial(-)	rs9271058	6:32575406	0.27	-3.010	0.159	6.36E-80
HLA-DRB1	Blood	NK cells / CD8+ T-Cells	rs9271058	6:32575406	0.27	-4.090	0.464	1.20E-18
HLA-DRB1	Blood	Monocytes / Macrophages	rs9271058	6:32575406	0.27	-3.540	0.497	1.06E-12
HLA-DRB1	Brain	NA	rs9271058	6:32575406	0.27	-1.690	0.054	1.94E-213
HLA-DRB5	Brain	NA	rs9271058	6:32575406	0.27	-1.770	0.081	2.28E-106
LRRC37A2	Brain	NA	rs2732703	17:44353222	0.21	1.370	0.053	4.13E-150
LRRC37A2	Brain	NA	rs113986870	17:44355683	0.09	1.260	0.068	1.98E-76
ARL17A	Brain	NA	rs113986870	17:44355683	0.09	-0.326	0.047	4.96E-12
HLA-DRB1	Brain	Microglia	rs9271058	6:32575406	0.27	-1.400	0.111	1.80E-36
HLA-DRB1	Brain	Neurons	rs9271058	6:32575406	0.27	-1.650	0.135	2.37E-34
HLA-DRB5	Brain	Neurons	rs9271058	6:32575406	0.27	-1.550	0.201	1.24E-14
LRRC37A2	Brain	Neurons	rs2732703	17:44353222	0.21	1.520	0.140	1.84E-27
LRRC37A2	Brain	Microglia	rs2732703	17:44353222	0.21	1.480	0.147	7.65E-24
LRRC37A2	Brain	Endothelial Cells	rs2732703	17:44353222	0.21	1.750	0.233	5.88E-14
LRRC37A2	Brain	Microglia	rs113986870	17:44355683	0.09	1.530	0.195	4.29E-15
LRRC37A2	Brain	Neurons	rs113986870	17:44355683	0.09	1.400	0.184	2.77E-14

*Position according to GRCh37 assembly; MAF = minor allele frequency of variant in 1000 Genomes Combined European Population; Cell-type specific result rows shaded in gray;

Table 4.3: Colocalized AD GWAS/lead eQTL SNP pairs

Region *	AD GWAS Variant							Lead eQTL Variant						eQTL type	PP4	r^2
	rsID	Nearest Gene	MAF	p-value	eQTL p-value	eGene	Cell-type	rsID	MAF	eGene	eQTL p-value	Cell-type	GWAS p-value			
6:46487762-48487762	rs10948363	CD2AP	0.72	1.77E-07	2.32E-104	CD2AP	NA	rs4711880	0.23	CD2AP	1.36E-104	NA	2.57E-07	Blood eQTL	0.909	1.00
6:46487762-48487762	rs10948363	CD2AP	0.72	1.77E-07	2.66E-09	CD2AP	NK cells / CD8+ T-Cells	rs13201473	0.27	CD2AP	1.47E-09	NK cells / CD8+ T-Cells	2.74E-07	Blood ct-eQTL	0.917	1.00
8:26195121-28195121	rs28834970	PTK2B	0.63	1.58E-09	9.15E-09	PTK2B	Interferon response/ Anti-bacterial Cells	rs6557994	0.41	PTK2B	2.58E-09	Interferon response/ Anti-bacterial Cells	8.19E-07	Blood ct-eQTL	0.990	0.78
8:26467686-28467686	rs9331896	CLU	0.61	3.62E-16	Not an eSNP			***rs6557994***	0.45	PTK2B	2.58E-09	Interferon response/ Anti-bacterial Cells	8.19E-07	Blood ct-eQTL	0.990	0.00
1:206692049-208692049	rs6656401	CR1	0.19	2.17E-15	1.05E-43	CR1	NA	rs12037841	0.19	CR1	9.25E-44	NA	1.77E-15	Brain eQTL	0.993	1.00
11:46557871-48557871	rs10838725	CELF1	0.68	1.91E-05	Not an eSNP			rs35233100	0.068	MADD	2.88E-10	NA	1.25E-03	Brain eQTL	0.954	0.12
11:58923508-60923508	rs983392	MS4A6A	0.59	4.76E-15	Not an eSNP			rs11230563	0.35	CD6	2.31E-113	NA	0.48	Brain eQTL	0.854	0.00
19:44411941-46411941	rs429358	APOE	0.78	0.00E+00	Not an eSNP			rs74253343	0.47	RELB	1.9E-14	Oligodendroglia	0.23	Brain ct-eQTL	0.971	0.00

* Map position within 1 Mb of AD GWAS SNP according to GRCh37 assembly; MAF = minor allele frequency; NA = not available; PP4 = posterior probability of colocalization; r^2 = correlation of AD and eQTL variants

Table 4.4: 20 most significant ct-eQTLs in **(A)** monocytes/ macrophages and **(B)** in microglia

A)

eGene	Lead eSNP	Position*	MAF	Beta	Std Err	P-value	# significant eSNPs in gene/cell type
SLC12A1	rs8037626	15:48606346	0.17	-3.340	0.219	1.62E-52	126
DLG2	rs75798025	11:84018349	0.01	5.350	0.364	6.66E-49	597
ABCA9	rs4147976	17:66925923	0.44	0.872	0.068	1.97E-37	48
PTPRG	rs116497321	3:62245373	0.01	2.650	0.221	3.96E-33	10
CLNK	rs5028371	4:10452986	0.50	1.060	0.092	7.66E-31	272
NFXL1	rs10938499	4:47848377	0.33	-1.270	0.112	8.38E-30	73
FCRL5	rs12760587	1:157526021	0.23	2.140	0.19	1.99E-29	93
HLA-DRB5	rs9269047	6:32438783	0.12	-11.600	1.03	2.02E-29	1
FMOD	NA	1:203263699	NA	2.110	0.2	5.08E-26	42
ABCA6	rs144031521	17:67162715	0.01	8.620	0.833	4.27E-25	9
INPP5F	rs181735165	10:121555618	0.02	7.150	0.701	1.99E-24	11
RBMS3	rs192885607	3:29612955	0.00	2.570	0.257	1.52E-23	34
ARHGAP44	NA	17:12750576	NA	1.760	0.177	2.69E-23	54
C4BPA	rs74148971	1:207275799	0.07	-2.300	0.234	8.44E-23	24
DCLK2	rs114930380	4:150954757	0.03	1.630	0.169	5.16E-22	39
PAM	NA	5:102153433	NA	-0.691	0.073	2.00E-21	47
MYO1E	rs146483144	15:59422810	0.03	4.300	0.453	2.26E-21	17
DSP	rs4960328	6:7495948	0.42	0.554	0.061	9.30E-20	6
ROR1	rs1557596882	1:64453767	0.01	3.570	0.393	1.05E-19	31
CACNB2	rs117299889	10:18404550	0.06	1.510	0.168	2.52E-19	61

B)

eGene	Lead eSNP	Position*	MAF	Beta	Std Err	P-value	# significant eSNPs in gene/cell-type
AC142381.1	rs199931530	16:33047273	0.45	-0.401	0.012	3.89E-233	43
MLANA	rs201480524	9:68457329	0.50	-0.176	0.007	4.82E-124	18
AC015688.3	rs62058902	17:25303954	0.50	-0.213	0.009	1.94E-113	11
HNRNPCL1	rs75627772	1:13182567	0.00	0.186	0.008	4.31E-113	8
AL050302.1	rs3875276	21:14472722	0.50	-0.890	0.040	2.62E-111	142
ALLC	rs9808287	2:3624799	0.11	0.833	0.044	1.41E-79	12
FAM21B	NA	10:47917284	NA	-2.210	0.118	2.88E-78	22
WNT3	rs9904865	17:44908263	0.37	-1.570	0.084	3.90E-78	1
RPL9	rs1458255	4:39446549	0.28	-2.370	0.137	4.76E-67	37
HLA-DRB1	rs72847627	6:32538512	0.32	-2.130	0.125	4.15E-65	2305
XRCC2	rs80034602	7:152104360	0.50	2.120	0.128	1.30E-61	5
WI2-3308P17.2	rs4067785	1:120576209	0.50	-0.184	0.011	1.33E-58	9
DEFB121	rs117541536	20:29422202	0.49	-7.420	0.460	1.56E-58	1
GINS1	rs75374582	20:26109209	0.50	-33.200	2.060	1.95E-58	5
EXOSC10	rs2580511	1:121113600	0.50	-4.390	0.276	5.78E-57	5
TRIM49B	rs202086299	11:48363026	0.50	0.205	0.013	2.16E-54	5
TMPRSS9	rs7248384	19:23936403	0.48	-1.410	0.093	1.83E-52	1
LDHC	NA	11:18432033	NA	0.428	0.030	1.07E-47	73
HLA-DOB	rs201194354	6:32796857	NA	1.190	0.084	4.39E-46	70
DEFB119	rs78099404	20:29617870	0.50	-0.377	3.51E-15	<1.0E-314	142

* chromosome and map position according to GRCh37 assembly
MAF = minor allele frequency; NA = not available

Table 4.5: **Overlap of ct-eQTLs in myeloid cell types in brain and blood**

A. Unique eGenes shared in significantly associated ct-eQTLs in monocytes/ macrophages and microglia. Number below each gene represents significant eGene-eSNP eQTL pairs in each gene

BTNL3	FAM118A	HLA-DOB	HLA-DRB1	HLA-DRB5
1	43	6	200	1

B. eSNP-eGene pairs among ct-eQTLs significant in both monocytes/ macrophages and microglia

				Monocytes/Macrophages		Microglia		
eGene	**eSNP**	**Position**	**MAF**	**Beta**	**P-value**	**Beta**	**P-value**	**AD GWAS P-value ***
HLA-DOB	rs3763355	6:32786882	0.06	-2.02	9.98E-15	0.938	3.89E-14	0.001
HLA-DOB	rs3763354	6:32786917	0.15	-1.11	1.40E-10	-0.642	2.80E-13	0.652
HLA-DOB	rs1183595100	6:32768232	NA	-1.13	8.34E-11	-0.605	1.98E-11	NA

* results from 2019 IGAP Summary Statistics [101]

C. Overlap of significant eQTLs in brain and blood with ct-eQTLs in myeloid cell types

Cell-types	**Monocytes/ Macrophages**		**Microglia**	
Blood	# ct-eQTLs common to cell type pair	# ct-eQTLs unique to cell type pair	# ct-eQTLs common to cell type pair	# ct-eQTLs unique to cell type pair
Neutrophils1			3 (0.3%) *	0
CD4+T-Cells			3 (0.5%)	0
NK/CD8+T-Cells			337 (12.9%)	0
Erythrocytes			119 (5.6%)	0
Monocytes/ Macrophages			251 (7.8%)	3
Unknown			0	0
Interferon/ Anti-bacterial			628 (3.3%)	63
Neutrophils2			0	0
B-cells			0	0
Eosinophils			38 (5.2%)	0
Brain				
Endothelial Cells	55 (0.5%)	0		
Neurons	250 (1.3%)	0		
Microglia	251 (1.6%)	3		
Astrocytes	0	0		
Oligodendroglia	0	0		

* number in parentheses represent the proportion of ct-eQTLs for each cell type on the left that were also observed in either microglia or monocytes/macrophages

Figure 4.1 Significant gene-SNP eQTLs and ct-eQTLs in blood and brain tissue genomewide. **A)** Venn diagram shows the number of overlapping eQTLs and ct-eQTLs in blood and brain. Gold color indicates significant eQTLs that are cell-type specific. Orange color indicates significant eQTLs that are shared between blood and brain. **B)** Cell-type distributions of significant genome-wide ct-eQTL results in blood and brain

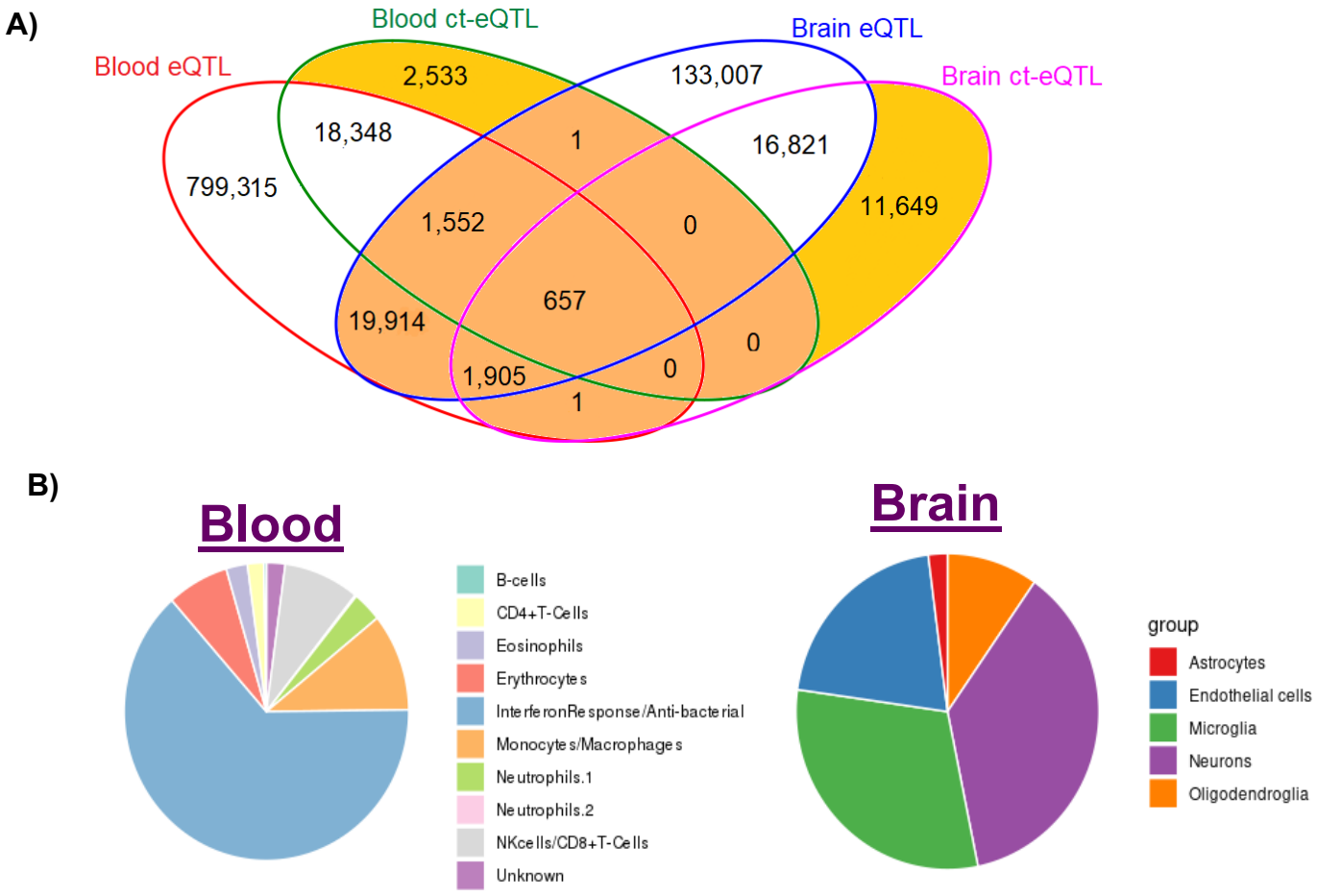
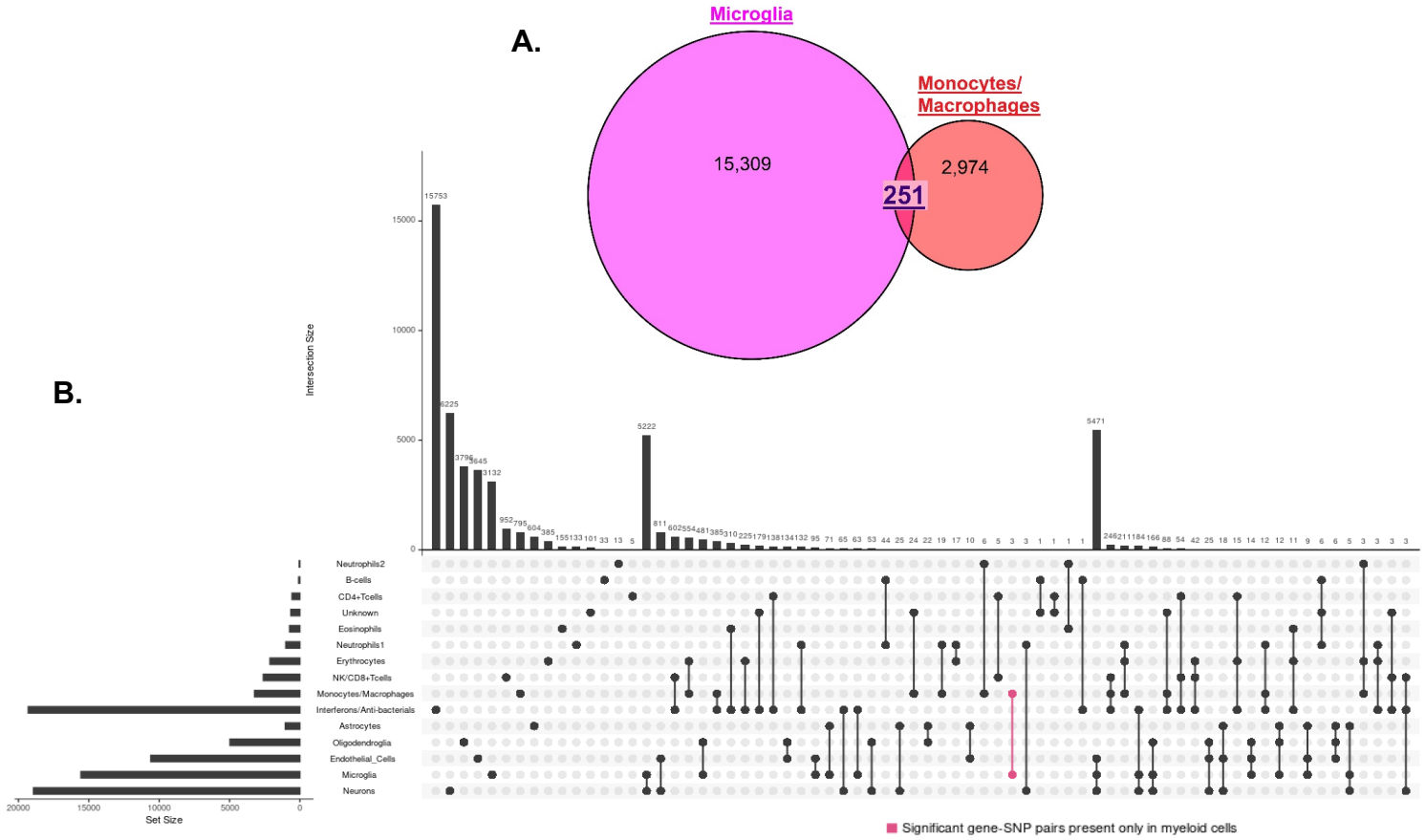


Figure 4.2 Intersection of significant gene-SNP eQTL pairs between cell-types in blood and brain tissue. A) Venn diagram showing overlap of ct-eQTL pairs in myeloid cell types (microglia and monocytes/macrophages). B) Number of significant eQTLs unique to and that overlap cell-types in blood and brain. Bar chart on the left side indicates the number of significant eQTLs involving each cell-type and bar chart above the matrix indicates the number of significant eQTLs that are unique to each cell type and set of cell-types. Pink colored bar indicates the number of eQTLs pairs that are unique to microglia and monocytes/macrophages



CHAPTER 5: Conclusions

To conclude, this dissertation further explains undiscovered AD genetic risk with novel findings by integrating multiple types of 'omics data, yields similarities and differences in AD mechanisms in brain and blood cell types with the multi-tissue and cell-type analysis, and also highlights the importance of applying these different methods to analysis to improve our understanding of AD. Not only were numerous novel findings found with the different strategies implemented in these chapters, but when examining top findings within all results in this work, more interesting associations are seen.

Rare variant results for mutations occurring only in cases or only in controls, as found in Chapter 2, for findings in other chapters are seen Table 5.1. The table includes top findings and results with a MAC ≥7 from Chapters 3 and 4. The top MAC in cases only (n=9) was for zinc finger protein gene *ZNF839*, a shared blood and brain eQTL gene from Chapter 4. The next top MAC(n=8) gene include *RNF39* which has a rare mutation in only 8 cases and zero controls. *RNF39* was a top brain gene-level rare eQTL(p=7.71E-27) and has previously been seen as an AD-associated differentially methylated region [102]. These rare variant results include also include blood gene-level gene *EXOC2*(n=7), which has been identified as an AD age of onset modifier and also as one of the largest co-expressed modules in a weighted gene co-expression analysis of posterior cingulate (PC) astrocytes in AD [183]. Multiple significant pathway-level genes associated with the Inflammation mediated by chemokine and cytokine signaling

pathway are also seen: *IFNAR*, *CXCR2*, and *RAF1*. Top genes in monocyte and microglia ct-eQTLs had a high MAC in cases only, including *TMPRSS9*(n=8), *ABCA6* (n=7), and genes that were potential AD candidates in Chapter 4 such as *DLG2*, *MYO1E*, and *INPP5*, further suggesting AD association with immune systems. The top results from Chapter 2, *NOTCH3* and *TREM2*, did not have significant rare-eQTLs though as seen in Table S5.1.

There were multiple shared pathway genes between brain and blood in Chapter 3, which were also present in common and ct-eQTL results (Table 5.2). *CXCL5* had a very significant eQTL in blood (p= 9.88E-144) and also had cell-type specific eQTLs in microglia and oligodendroglia. Microglial activation and the subsequent changes in cytokines and chemokines such as *CXCL5* are keys steps in the development of neuroinflammation [214]. *CXCL5*, *CCL3*, and *CCL4* are all cell-type specific in oligodendroglia, and both *CXCL5* and *CCL3* are present in both blood and brain eQTLs. In addition, pathway-level rare eQTL findings from Chapter 3 were seen in common and ct-eQTL results as well (Table 5.3). *ALOX5AP* is seen in both gene-level and pathway-level rare eQTLs and common eQTLs all in blood. *FPR2* has a very significant eQTL in blood (p= 1.22E-240) and is also cell-type specific in blood interferon and anti-bacterial cells (p= 3.81E-17). *IFNAR1* also has an eQTL in blood and is cell-type specific in the brain in endothelial cells. *FPR2*, *ALOX5AP*, *IFNAR1*, *RAF1*, *GRB2*, and *CCL8* in this table are all associated with the Inflammation pathway. *TNFRSF10C*, associated with the apoptosis pathway, appears in both blood (p=

7.35E-73) and brain (p= 4.30E-36) eQTLs, though much more significant in the blood.

Throughout this dissertation, I have discovered rare and common variant eQTLs and cell-type specific eQTLs in both brain and blood. I explored the overlap of significant genes between these three types of eQTLs in each tissue (Figure 5.1). There are a total of 203 genes in blood and 40 genes in brain that have both common and rare-eQTLs (Table S5.2). There are 29 genes in blood and 28 genes in brain that have significant rare eQTLs and are cell-type specific (Table S5.3). Interestingly, there are 29 genes in blood and 21 genes in brain that are significant in rare, common, and ct-eQTLs (Table S5.4). When comparing these total intersects in both tissues, there is only one gene in both-*HP*. *HP* or haptoglobin was seen as shared gene in blood and brain gene-level rare-eQTL findings and has previously been found to have a much higher mean serum level in AD patients compared to healthy controls [125].

When examining all results and top findings together, there is further evidence for AD association for many top genes, bringing together genetic and genomic evidence. Collectively, these findings further explain the genetic basis of AD risk and provide insight about mechanisms leading to this disorder.

Table 5.1 Rare variants in cases only for other top findings

Gene	SNP Position	MAC in Cases*	Mutation Type	Disease Impact	CADD score	Top findings or Results MAC≥7 in
ZNF839	14:102805285	9	Missense	Moderate	10.83	Shared blood/brain eQTL genes
RNF39	6:30043301	8	Missense	Moderate	26.1	Brain gene-level rare eQTLs
TAC3	12:57406650	8	Missense	Moderate	17.66	Blood gene-level rare eQTLs
RHCE	1:25718599	8	Missense	Moderate	9.565	Blood gene-level rare eQTLs
TMPRSS9	19:2405464	8	Missense	Moderate	25.7	Top Microglia ct-eQTLs
RARRES1	3:158428589	8	Missense	Moderate	27.6	Shared blood/brain eQTL genes
EXOC2	6:564588	7	Missense	Moderate	25.1	Blood gene-level rare eQTLs
ABCA6	17:67075361	7	Missense	Moderate	26.9	Top Monocyte ct-eQTLs
CNDP2	18:72168601	7	Missense	Moderate	14.14	Shared blood/brain eQTL genes
SPEF2	5:35740256	7	Missense	Moderate	12.44	Shared blood/brain eQTL genes
IFNAR1	21:34697424	5	In-frame deletion	Moderate	21.7	Blood pathway-level rare eQTLs
WWOX	16:78466539	5	Missense	Moderate	22.4	Shared blood/brain AD gene
CXCR2	2:218999633	4	Missense	Moderate	0.455	Blood pathway-level rare eQTLs
RAF1	3:12660099	4	Missense	Moderate	25.2	Blood pathway-level rare eQTLs
INPP5F	10:121551531	4	Missense	Moderate	25.1	Top Monocyte ct-eQTLs
INPP5F	10:121556381	4	Missense	Moderate	23	Top Monocyte ct-eQTLs
DLG2	11:84996317	4	Missense	Moderate	24.8	Top Monocyte ct-eQTLs
SLC12A1	15:48518739	4	Missense	Moderate	21.6	Top Monocyte ct-eQTLs
MYO1E	15:59464106	4	Missense	Moderate	17.8	Top Monocyte ct-eQTLs

*MAC in Controls = 0 for all variants

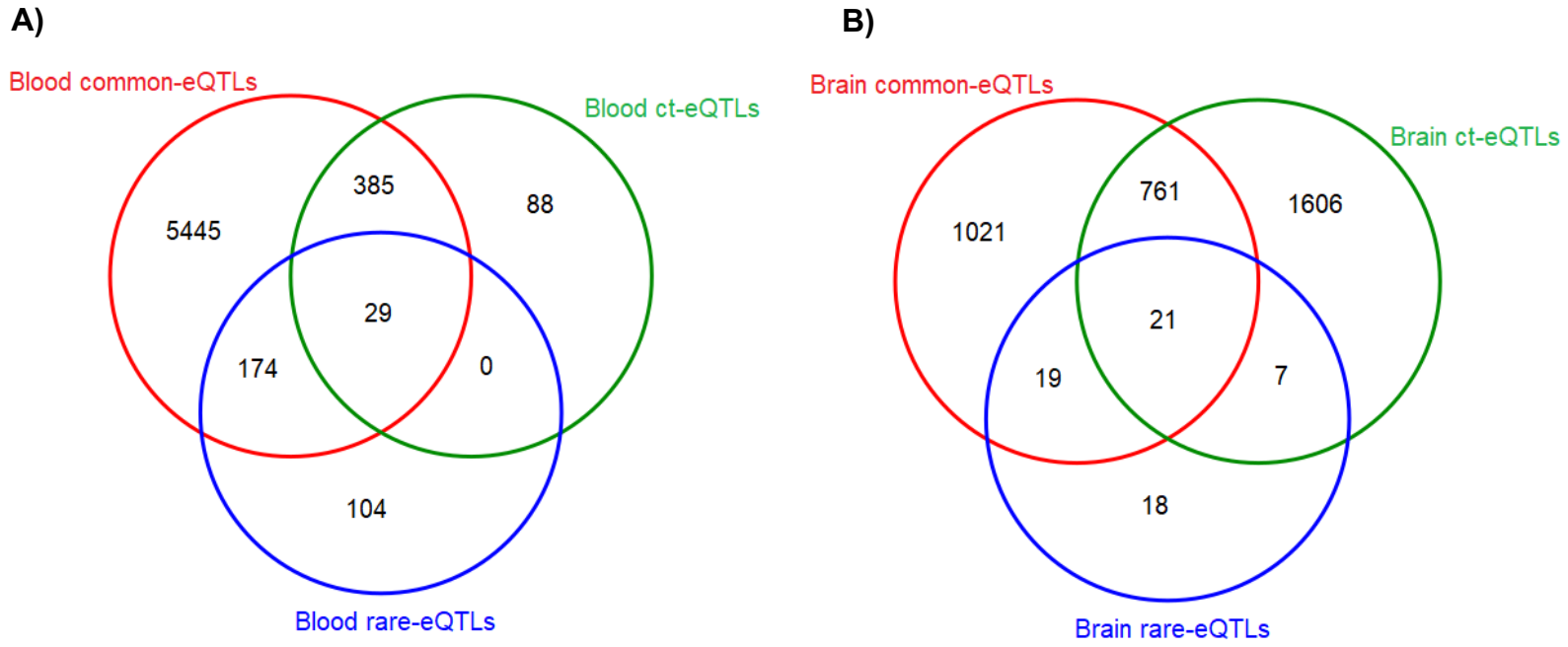
Table 5.2 Common and ct-eQTLs for shared pathway genes in blood and brain

Gene	Tissue	SNP Position	Cell-Type	P-value
CXCL5	Blood	4:74857970	NA	9.88E-144
CCL4	Blood	17:34431403	NA	1.35E-26
CXCL5	Brain	4:7540518	Oligodendroglia	3.03E-24
CXCL5	Brain	4:74601312	Microglia	1.23E-13
CCL3	Brain	17:34591375	Oligodendroglia	3.77E-13
HSPA6	Blood	1:161482520	NA	3.99E-12
CCL4	Brain	17:34762804	Oligodendroglia	4.46E-12
MMP9	Blood	20:44645339	NA	2.28E-11

Table 5.3 Pathway-level rare eQTL findings in common and ct-eQTL results

Gene	Lead eSNP	Cell-Type	eQTL/ct-eQTL Tissue	P-value	Pathway-level Tissue
FPR2	19:52276665	NA	Blood	1.22E-240	Blood
ALOX5AP	13:31312178	NA	Blood	1.76E-158	Blood
TNFRSF10C	8:22957606	NA	Blood	7.35E-73	Blood
TNFRSF10C	8:22941873	NA	Brain	4.30E-36	Blood
IFNAR1	21:34670242	NA	Blood	6.58E-34	Blood
CFLAR	2:201982693	NA	Blood	1.58E-26	Blood
STAT5A	17:40456010	NA	Blood	1.11E-24	Blood
ALAD	9:116167156	NA	Blood	1.76E-22	Blood
CCL8	17:3188133	Oligodendroglia	Brain	3.71E-21	Brain
UROD	1:45498181	NA	Blood	3.65E-20	Blood
GRB2	17:73271758	NA	Blood	9.53E-20	Blood
PAK1	11:77060092	NA	Blood	1.32E-17	Blood
FPR2	19:52269634	Interferon/Anti-bacterial cells	Blood	3.81E-17	Blood
IFNAR1	21:34370749	Endothelial Cells	Brain	7.49E-14	Blood
RPS6KA1	1:26883511	NA	Blood	1.20E-13	Blood
RAF1	3:12690343	NA	Blood	1.33E-11	Blood
GNG5	1:84953765	NA	Blood	2.18E-11	Blood

Figure 5.1 Overlap of significant genes in rare gene-level, common, and cell-type eQTLs. The venn diagrams show the intersects of significant genes in all three types of eQTLs in **A)** blood and **B)** brain.



APPENDIX

Table S2.1 Characteristics of subjects in the ADSP WES Case-Control dataset

Group	N	AD Cases	Controls	Males	Females	Mean Age (SD) *	% *APOE* ε4 carrier
European ancestry	10,367	5,566 (54%)	4,574 (44%)	4,369 (42%)	5,998 (58%)	81.0 (9.1)	30%
Caribbean Hispanic	395	218 (55%)	177 (45%)	149 (38%)	246 (62%)	74.4 (7.8)	40%

Age at onset for AD cases, age at exam for controls

Table S2.2 Known AD/Dementia genes

Gene	Disease	Reference
ABCA7	AD	[74, 215, 216]
ABCG1	AD	[217]
ABI3	AD	[3]
ACE	AD	[218]
ADAM10	AD	[219, 220]
ADAMTS1	AD	[221]
AKAP9	AD	[15]
APOE	AD	[222, 223]
APP	AD	[224]
BIN1	AD	[225]
BZRAP1	AD	[170]
C1QTNF4	AD	[226]
CASP8	AD	[227]
CASS4	AD	[58]
CD2AP	AD	[58]
CD33	AD	[58]
CELF1	AD	[58]
CHCHD10	FTD	[228]
CHMP2B	FTD	[229, 230]
CLU	AD	[228]
COBL	AD	[231]
CR1	AD	[58]
CSF1R	FTD	[232]
DSG2	AD	[58]
ECHDC3	AD	[170]
ECRG4	NP+NFT	[233]
EIF2B1	Leukodystrophy	[234]
EIF2B2	Leukodystrophy	[235]
EIF2B3	Leukodystrophy	[235]
EIF2B4	Leukodystrophy	[234]
EIF2B5	Leukodystrophy	[235]
EPHA1	AD	[37, 236]
FBXL7	AD	[237]
FERMT2	AD	[58]
FRMD4A	AD	[238]
FUS	FTD	[239]
GALNT7	AD	[217]
GLIS3	AD	[57]
GRN	FTD	[240, 241]
HBEGF	AD	[170]
HDAC9	NFT + CAA	[233]
HLA-DRB1	AD	[58]
HLA-DRB5	AD	[242]
IGHV1-67	AD	[243]
INPP5D	AD	[58]
ITM2B	FBD/FDD	[244-246]
KANSL1	AD	[56]
KCNMB2	AD	[217]

LMNB1	Leukodystrophy	[247]
LMX1B	AD	[248]
MAPT	AD	[56]
MEF2C	AD	[58]
MS4A4A	AD	[37]
MS4A6A	AD	[58]
MTUS1	AD	[249]
MVB12B	AD	[250]
NME8	AD	[58]
NOTCH3	CADASIL	[39, 43]
OSTN	AD	[57]
PDGFRL	AD	[251]
PFDN1	AD	[170]
PICALM	AD	[58]
PILRA	AD	[206]
PLCG2	AD	[3]
PLD3	AD	[252]
PLD4	AD	[249]
PLXNA4	AD	[253]
PPP2CB	AD	[254]
PRNP	CJD/GSS	[255]
PSEN1	AD	[256]
PSEN2	AD	[256]
PTK2B	AD	[58]
RIN3	AD	[257]
SLC10A2	AD	[231]
SLC24A4	AD	[58]
SLC2A4A	AD	[249]
SORCS1	AD	[258, 259]
SORCS2	AD	[259]
SORCS3	AD	[259]
SORL1	AD	[50, 242]
SRRM4	CSF Tau	[249]
TARDBP	FTD	[260]
TM2D3	AD	[261]
TP53INP1	AD	[243]
TPBG	AD	[170]
TRAPPC12	NFT + CAA	[233]
TREM2	AD	[14, 69]
TREML2	AD	[57]
TRIP4	AD	[262]
UNC5C	AD	[16]
USP6NL	AD	[170]
VCP	FTD	[263]
ZCWPW1	AD	[58, 207]
ZNF804B	AD	[249]

AD = Alzheimer disease; CAA = cerebral amyloid angiopathy; CADASIL = Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy; CJD = Creutzfeldt-Jakob disease; FBD = Familial British dementia; FDD = Familial Danish dementia; FTD = fronto-temporal dementia; GSS = Gerstmann-Straussler syndrome; HPV = hippocampal volume; HS = hippocampal sclerosis; LMdT = logical memory – delayed recall; NFT = neurofibrillary tangle; NP = neuritic plaque

Table S2.3 Filtering pipeline of rare variants

# of Variants (# of Genes)	Filtering Steps
1,015,329 (29,157)	Total rare variants genome-wide after counts are calculated with PLINK
75,716 (15,702)	Total rare variants genome-wide after filtering (to remove synonymous mutations, those with a count less than or equal to 2) and restricted to those with a high or moderate impact on disease (as classified by VEP annotation).
4,854 (3,619)	Total rare variants genome-wide after filtering and with high impact on disease.
10,249 (95)	Total rare variants in known AD genes
503 (83)	Total rare variants in known AD genes after filtering and with high and moderate impact on disease.
24 (19)	Total rare variants in known AD genes after filtering and with high impact on disease.

Table S2.4 Characteristics of AD Subjects in the ADSP WES dataset with the *NOTCH3* rs149307620 mutation

Study (Dataset)	Sex	Age at Onset	*APOE* Genotype	Braak Stage
ADGC (ADC2)	F	71	33	NA
ADGC (ADC3)	M	76	33	6
ADGC (ADC3)	M	78	34	NA
ADGC (ADC4)	F	74	34	NA
ADGC (ADC5)	M	78	24	NA
ADGC (Memory and Aging Project)	M	84	33	NA
ADGC (Texas Alzheimer's Research and Care Consortium)	F	80	33	NA
CHARGE	M	79	33	NA
CHARGE	M	95	33	NA
CHARGE	F	84	33	NA

NA = Not autopsied

Table S2.5 Characteristics of subjects in the WGS replication datasets

Dataset	N	AD Cases	MCI Cases	Controls	% Female	Mean Age (SD)	% *APOE* ε4 carrier
European ancestry ADSP Extension	1097	485	0	612	63.5%	78.4(7.4)	38.5%
Caribbean Hispanic ADSP Extension	1018	493	0	525	55.1%	73.9(7.4)	29.4%
African American ADSP Extension	977	454	0	523	55.1%	79.2(7.1)	43.3%
European ancestry ADSP families	499	320	0	179	64.5%	73.4(10.6)	39.3%
Caribbean Hispanic ADSP families	290	200	0	90	62.8%	70.6(10.1)	32.5%
African American ADSP families	44	30	0	14	65.3%	69.1(12.1)	51.0%
ADNI	809	239	321	249	44.8	76.3(8.1)	44.0%
Total	**4,734**	**2,221**	**321**	**2,192**			

Table S2.6 Characteristics of AD Subjects with the *TREM2* rs104894002 mutation (Q33X)

Source (Study)	Sex	Onset Age	APOE Genotype	Braak Stage	Stroke Risk Factors
ADGC (ADC)	M	75	23	5	5
ADGC (Mayo Clinic)	M	63.3	44	NA	NA
National Cell Repository for Alzheimer's Disease	F	69	34	4	4
CHARGE (Rotterdam Study)	M	73.5	34	NA	NA

NA = not autopsied

Table S2.7 Genes with at least 3 distinct high/moderate disease impact rare variants each with a MAC ≥ 5 and occurring in only cases

SNP	MAC	ID	Gene	Mutation Type	Disease Impact	CADD Score
14:74754936:G:A	10	rs57773157	ABCD4	Missense	Moderate	18.1
14:74766360:T:C	10	rs58272575	ABCD4	Missense	Moderate	15.7
14:74763064:C:T	10	rs34992370	ABCD4	Missense	Moderate	8.7
14:74764675:G:A	8	rs61744947	ABCD4	Missense	Moderate	22.4
3:38125659:C:T	6	rs143610524	DLEC1	Missense	Moderate	20.4
3:38101245:C:T	6	rs34012183	DLEC1	Missense	Moderate	18.1
3:38101313:G:A	5	rs149190717	DLEC1	Missense	Moderate	13.0
1:225239279:G:A	6	Novel	DNAH14	Missense	Moderate	14.5
1:225512438:G:A	5	rs566891789	DNAH14	Splice Acceptor	High	25.7
1:225458497:C:T	5	rs530417418	DNAH14	Missense	Moderate	35
7:111430638:G:A	5	rs377187510	DOCK4	Missense	Moderate	28.8
7:111617307:G:A	5	rs201242965	DOCK4	Missense	Moderate	23.1
7:111503433:T:C	5	Novel	DOCK4	Missense	Moderate	22.3
9:133932442:C:T	7	rs113443891	LAMC3	Missense	Moderate	17.7
9:133945155:G:A	7	rs113785045	LAMC3	Missense	Moderate	16.9
9:133936490:A:G	6	rs36030184	LAMC3	Missense	Moderate	12.0
9:133948124:G:A	5	rs144118534	LAMC3	Missense	Moderate	4.0
14:64898328:A:G	6	rs139264994	MTHFD1	Missense	Moderate	21
14:64884726:C:T	6	rs199501976	MTHFD1	Missense	Moderate	17.1
14:64916188:C:T	6	rs17857382	MTHFD1	Missense	Moderate	11.2
1:228526693:C:T	5	rs370778898	OBSCN	Missense	Moderate	27.4
1:228558959:G:A	5	rs191931829	OBSCN	Missense	Moderate	17.3
1:228437675:C:T	5	rs199655077	OBSCN	Missense	Moderate	0.1
14:88904567:G:A	6	rs10139784	SPATA7	Missense	Moderate	11.7

14:88892697:G:A	6	rs17124662	SPATA7	Missense	Moderate	2.3
14:88883100:A:G	6	rs61747004	SPATA7	Missense	Moderate	2.1
14:88895750:G:A	6	rs17124677	SPATA7	Missense	Moderate	0.002
14:88883173:T:G	5	rs35137272	SPATA7	Missense	Moderate	0.02
1:1269488:G:A	6	rs112507608	TAS1R3	Missense	Moderate	6.3
1:1268954:C:T	5	rs200679891	TAS1R3	Missense	Moderate	11.0
1:1268661:G:A	5	rs373494410	TAS1R3	Missense	Moderate	10.7
2:179499179:A:G	9	rs34706299	TTN	Missense	Moderate	22.1
2:179412829:C:T	6	rs72648251	TTN	Missense	Moderate	16.2
2:179452061:C:T	5	rs199505416	TTN	Missense	Moderate	21.6
2:179457733:G:A	5	rs72646839	TTN	Missense	Moderate	20.1
2:179440609:A:G	5	rs201836227	TTN	Missense	Moderate	18.8
2:179613715:C:G	5	rs145919543	TTN	Missense	Moderate	10.2

Table S2.8 High impact rare variants genome-wide with a MAC ≥ 7 and occurring only in AD cases.

SNP	MAC	Gene	ID	Mutation Type	Disease Impact	CADD Score
7:150389837:TC:T	10	GIMAP2	Novel	Frameshift	High	5.92
19:39926486:A:G	8	RPS16	rs139109626	Splice donor	High	22.4
11:5758174:CAATT:C	8	OR56B1	Novel	Frameshift	High	14.8
10:16979600:TGGTA:T	7	CUBN	rs556462218	Frameshift	High	40.0
10:25313667:T:A	7	THNSL1	rs150653385	Stop gained	High	37.0
1:63063592:GAACTC:G	7	ANGPTL3	Novel	Frameshift	High	21.0
17:72350672:C:T	7	KIF19	rs200837156	Stop gained	High	20.6
1:21904139:TG:T	7	ALPL	Novel	Frameshift stop lost	High	12.9
2:111918994:A:G	7	BCL2L11	rs76245002	Splice acceptor	High	5.26

Figure S2.1 Study design

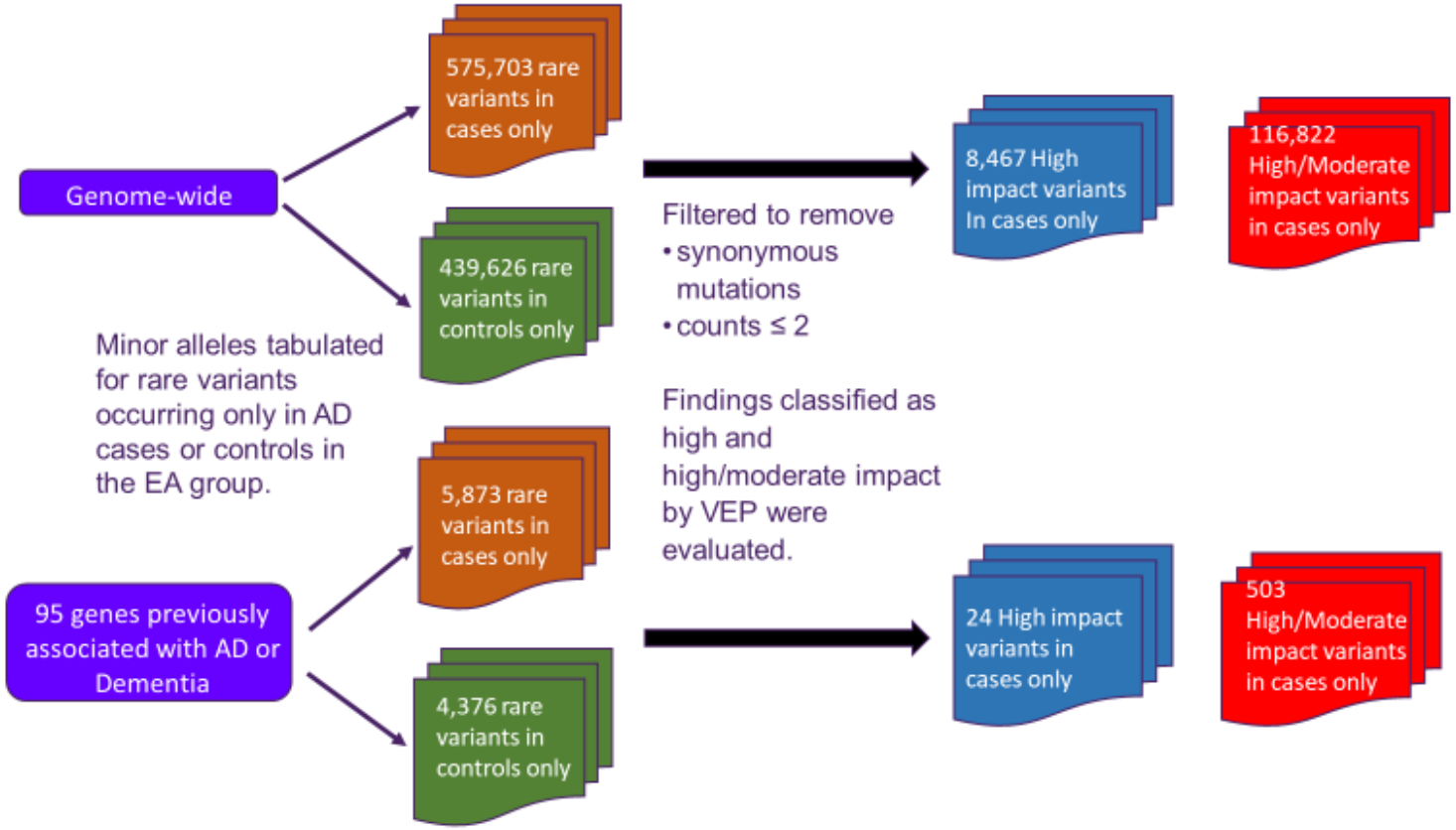


Figure S2.2 Haplotype Analysis of the rare *NOTCH3* rs149307620 variant. All 10 subjects who have the rs149307620 mutation (marker 269) possess the GCCGC haplotype derived from five common SNPs (rs1548555, rs1044006, rs1043997, rs1043996, and rs1043994) spanning a region of ~24.8 kb. The haplotype has a frequency of 15.4% in AD cases and 14.7% in controls. Haplotypes and their corresponding frequencies are shown in the left panel. A measure of linkage disequilibrium (D') between each pair of SNPs is shown in the right panel.

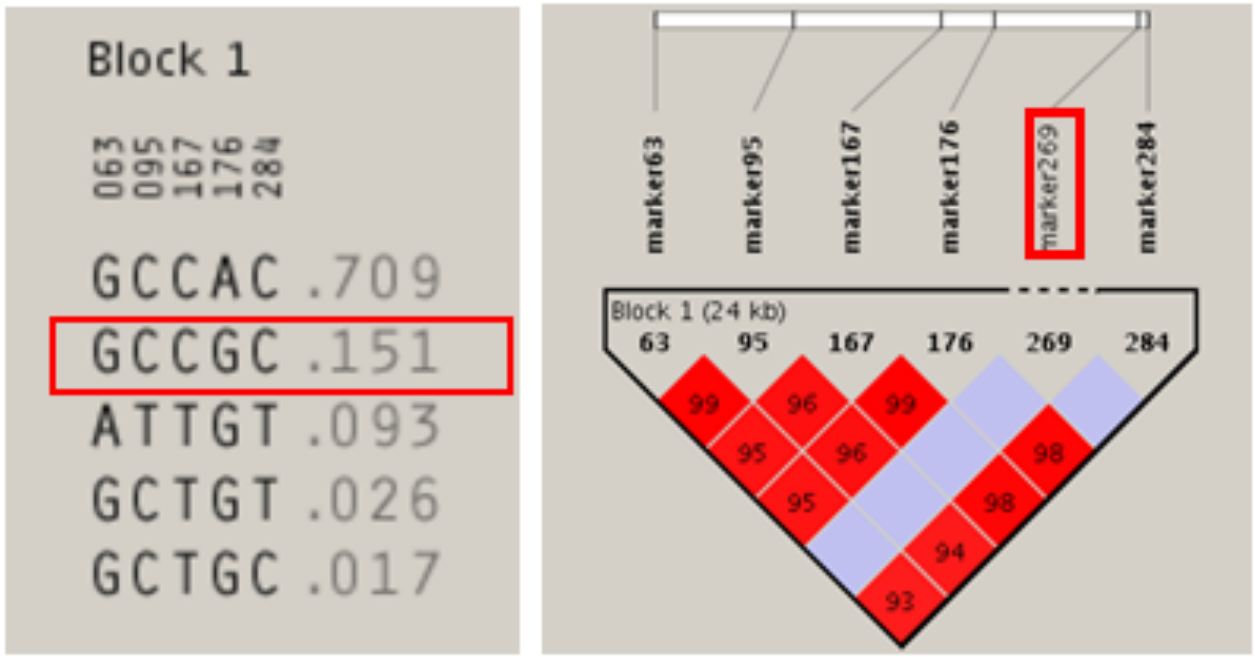


Figure S2.3: Population substructure of the ADSP discovery sample. Population substructure in the non-Hispanic European ancestry portion of the ADSP sample was evaluated by principal components (PC) analysis implemented in EIGENSTRAT[79,80] using the smartpca program as described previously.[2] For this analysis, variants were selected based on the following filtering criteria: MAF≥5%, call rate≥99%, and only one from each pair of variants with linkage disequilibrium (LD) of r^2>0.5 in a 50-variant window. Participants with individual call rate<90% were excluded, and only one participant from every pair with estimated $\hat{\pi}$>0.2 was selected. PCs were computed using the 1000 Genomes Phase 3 reference panel[6] and a subset of 12,351 variants that met the above criteria and were found in both ADSP and 1000G subjects. PCs were computed using all unrelated ADSP and 1000G subjects and projected on individuals that were omitted from the PC computation due to high IBD or low call rate. The plot of the first two PCs (PCA1 and PCA2) shows that participants (each represented by a small dot) are distributed along a population substructure gradient. A large portion of the persons in the cluster on the right side of the plot (PC1>0.025) have a mitochondrial haplogroup (K1a1b1a = light blue dots, K1a9 = dark blue dots) that is common among Ashkenazi Jews suggesting that this cluster includes primarily Ashkenazi Jews. Eight of the ten AD participants who have the *NOTCH3* rs149307620 mutation, including four persons with the K1a1b1a or K1a9 haplogroup (dark red + symbol) and four persons lacking either Ashkenazi haplogroup (brown + symbol), are included in this cluster.

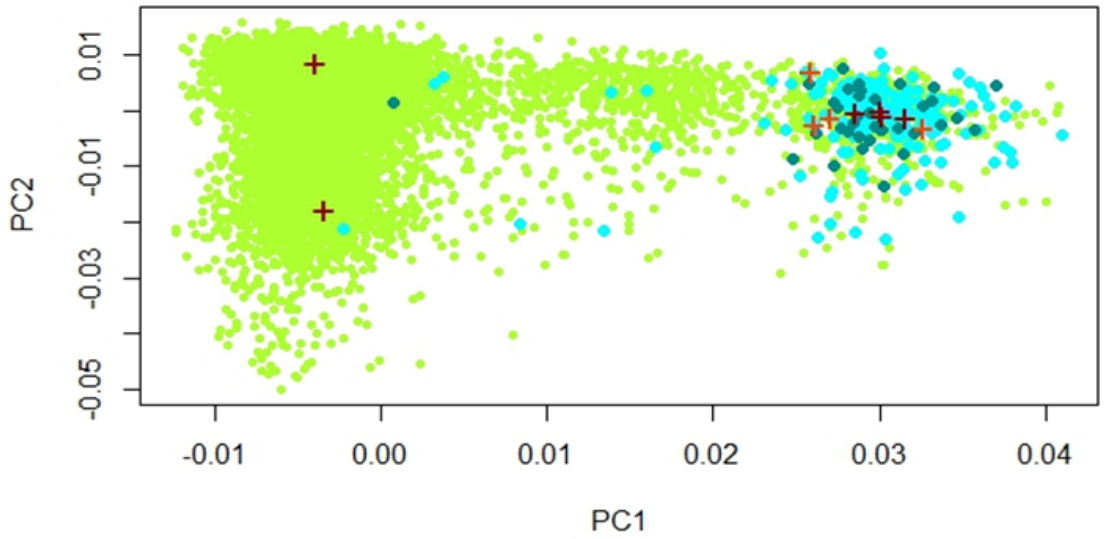
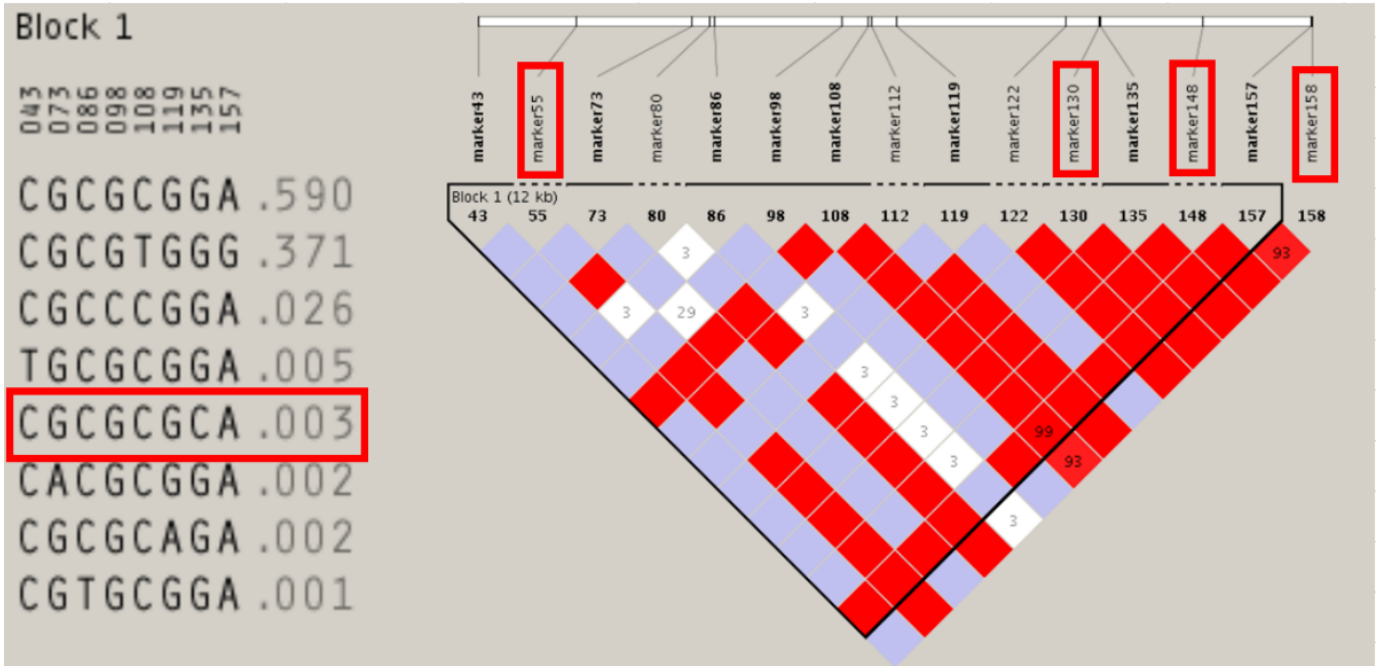


Figure S2.4 Haplotype Analysis of *ABCD4* and *CELSR1/GTSE1*. **A.** All 10 subjects with ABCD4 rare variants rs57773157 (marker 55), rs34992370 (marker 130), and rs58272575 (marker 158) share an 8-SNP haplotype spanning ~12.9 kb that has a frequency of 0.3% in AD cases and 0% in controls. Eight of these subjects also have a fourth *ABCD4* rare variant, rs61744947 (marker148) that is part of the same shared haplotype. **B.** All eight subjects with *CELSR1* variants rs61741871 (marker 268) and rs75983687 (marker345) and *GTSE1* variants rs34404175 (marker 74) and rs35503220 (marker84) share a 12 SNP-haplotype spanning ~77.6 kb that has a frequency of 0.1% in the ADSP population (0.1% in cases, 0.1% in controls). Haplotypes and their corresponding frequencies are shown on the left side of each panel. A measure of linkage disequilibrium (D') between each pair of SNPs is shown on the right side of each panel.

A.



B.

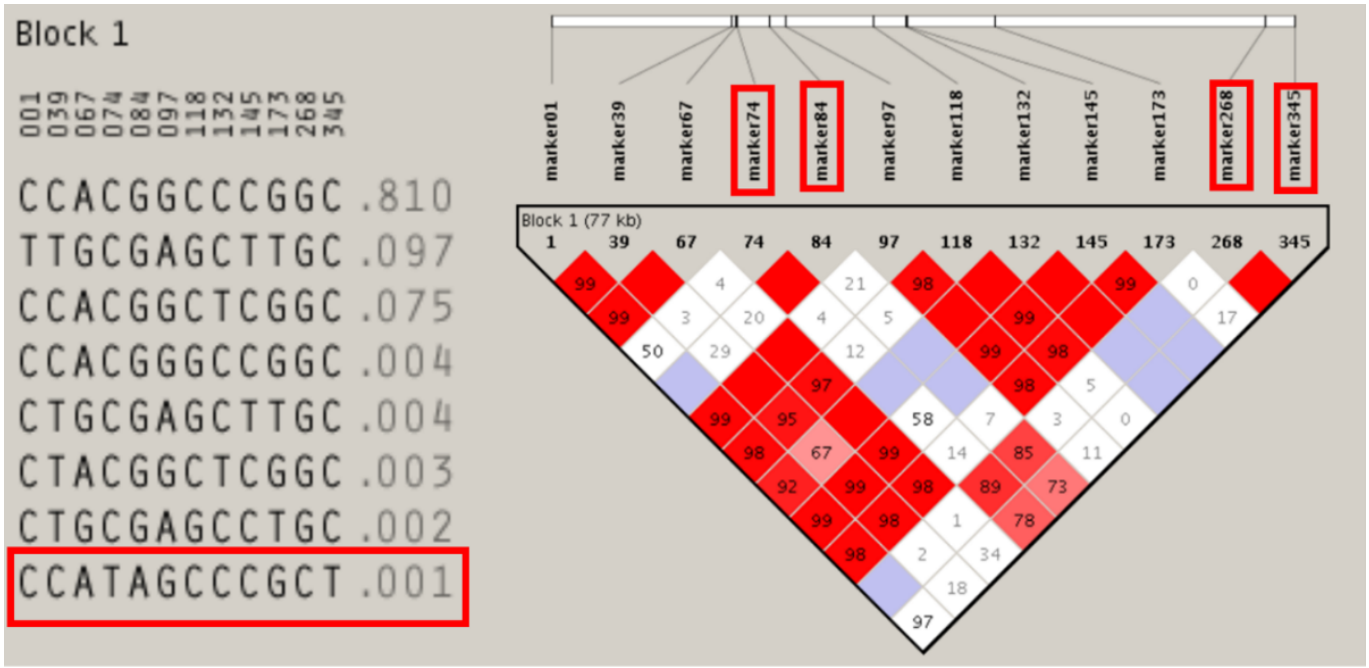

Block 1
001 039 067 074 084 097 118 132 145 173 268 345
CCACGGCCCGGC .810
TTGCGAGCTTGC .097
CCACGGCTCGGC .075
CCACGGGCCGGC .004
CTGCGAGCTTGC .004
CTACGGCTCGGC .003
CTGCGAGCCTGC .002
CCATAGCCCGCT .001
marker01
marker39
marker67
marker74
marker84
marker97
marker118
marker132
marker145
marker173
marker268
marker345
Block 1 (77 kb)
1 39 67 74 84 97 118 132 145 173 268 345

Table S3.1. Characteristics of subjects in the ROSMAP and ADNI datasets

Dataset	Race	N	AD Cases	Controls	Males	Females	Mean Age (SD) *	% *APOE* ε4 carrier
ROSMAP (Brain)	98% Caucasian, 2% African American, < 0.01 % other	475	281(59%)	194 (41%)	175 (37%)	300 (63%)	85.9 (4.8)	26%
ADNI (Blood)	93% Caucasian, 4% African American, 3 % other	713	207(29%)	222 (31%)	396 (56%)	317 (44%)	76.3 (8.1)	40.5%

Table S3.2. Gene-level cis rare eQTL significant results in brain

CHR	BEGIN POS	END POS	NUM PASS VARS	NUM SING VARS	STATRHO	P-VALUE	GENE
5	22985688	25142892	177	113	0	4.56E-49	C5orf17
16	27512608	29486403	269	169	0	1.69E-30	IL27
6	29051606	31043360	426	220	0	7.71E-27	RNF39
19	18779631	20736620	248	173	0	2.19E-22	ZNF101
1	147741039	149755790	186	69	0	2.42E-21	NBPF16
4	89823539	91848018	153	96	0	1.49E-19	MMRN1
19	18977711	20990019	221	155	0	7.45E-19	ZNF253
1	25378602	27385983	482	320	0	1.96E-16	TRIM63
9	138266509	140255330	553	352	0	7.16E-16	DNLZ
7	43155135	45148545	252	162	0	1.56E-13	POLD2
2	169503264	171547270	343	205	0	1.85E-13	CCDC173
11	103966694	105967614	107	68	0	6.73E-13	CARD17
14	73016786	75002697	423	267	0	2.34E-12	ACOT1
16	13847527	15857759	298	191	0	1.64E-11	NPIPA2
1	149958668	151965171	540	352	0	1.83E-11	ANXA9
4	43689060	45457518	69	40	0	2.49E-11	GUF1
16	69333666	71398570	474	316	0	3.54E-11	RP11-529K1.3
15	81558119	83561556	184	106	0	6.95E-11	FAM154B
1	53492087	55481065	460	281	0	1.06E-10	LDLRAD1
7	127418124	129414750	519	310	0	1.07E-10	OPN1SW
6	27300777	29323838	369	221	0	1.22E-10	ZSCAN31
14	73783107	75825914	482	303	0	1.87E-10	VRTN
19	9406430	11373256	444	293	0	1.87E-10	ICAM4
6	29051606	31002490	425	219	0	2.38E-10	ZNRD1
2	24400738	26360191	286	180	0	2.72E-10	POMC
16	17417023	19440804	267	149	0	4.61E-10	NPIPA8
6	31783750	33784748	926	480	0	1.24E-09	HLA-DOB
2	60375560	62373811	538	337	0	1.65E-09	C2orf74
12	53696863	55726368	853	532	0	1.81E-09	COPZ1
16	71090452	73094829	741	461	0	2.28E-09	HP
11	112263796	114271056	464	266	0	2.34E-09	ANKK1
19	1941163	3942240	400	276	0	1.41E-08	ZNF77
6	31496752	33476802	1053	565	0	3.58E-08	HLA-DRB5
1	247920924	249234675	158	88	0	8.53E-08	LYPD8
15	64039999	66063761	404	249	0	9.90E-08	RBPMS2
12	9980142	11951867	155	98	0	1.02E-07	TAS2R9
2	95333119	96830428	153	88	0	1.10E-07	ZNF514
1	227399406	229357415	255	176	0	1.20E-07	C1orf145
6	28909389	30908847	458	238	0	2.02E-07	HLA-A
15	42622894	44630360	522	332	0	2.04E-07	ADAL

11	17434230	19468040	429	273	0	2.07E-07	LDHC
22	41523734	43525652	369	245	0	2.49E-07	CYP2D6
16	71090452	73110685	755	468	0	2.80E-07	HPR
1	149349077	151447295	508	334	0	3.36E-07	RPRD2
12	49721015	51789242	429	284	0	4.37E-07	FAM186A
8	26682344	28667180	197	127	0	4.71E-07	ESCO2
7	63347619	65355768	132	51	0	5.07E-07	ZNF273
6	34447206	36438023	417	258	0	5.55E-07	RPL10A
12	20680554	22677495	165	112	0	6.47E-07	C12orf39
12	9062912	11038219	163	106	0	7.21E-07	KLRF2
2	157116105	159166135	545	334	0	8.96E-07	GALNT5
2	131958523	134009777	395	157	0	1.04E-06	ANKRD30BL
1	39240400	41236581	492	310	0	1.69E-06	OXCT2
14	44615125	46540143	233	140	0	1.73E-06	PRPF39
6	82920251	85135804	241	160	1	1.79E-06	ME1
1	44966476	46963603	600	383	0	1.85E-06	CCDC163P
6	41942338	43929364	671	437	0	1.85E-06	GNMT
10	62424318	64520825	669	426	0	1.95E-06	C10orf107
6	31548925	33554557	1027	545	0	1.98E-06	HLA-DRB1
16	67034867	69106452	714	482	0	1.98E-06	DUS2
21	46707826	48118123	184	132	0	2.06E-06	YBEY
6	30104753	32098106	830	448	0	2.49E-06	PSORS1C1
14	74522315	76532783	500	308	0	2.53E-06	ACYP1
22	36513231	38505189	446	289	0	2.64E-06	TMPRSS6
7	35121219	37109559	238	133	0	3.24E-06	PP13004

BEGIN POS : Beginning position of range for rare variants within 1 Mb of gene to be tested

END POS : End position of range for rare variants within 1 Mb of gene to be tested

NUM PASS VARS : Number of variants passing all thresholds for EPACTS software

NUM SING VARS : Number of singletons among variants in NUM PASS VARS

P-VALUE : P-value of burden tests

STATRHO: represents the RHO value from SKAT-O test, rho = 1 (burden) and rho = 0 (SKAT)

Table S3.3. Gene-level cis rare eQTL significant results in blood

CHR	BEGIN POS	END POS	NUM PASS VARS	NUM SING VARS	STATRHO	P-VALUE	GENE
11	70732846	72788214	555	359	0	6.01E-76	NUMA1
2	170670625	172715732	637	363	0	1.49E-58	GAD1
2	38209429	40345697	564	330	0	3.58E-48	SOS1
7	55086655	57078144	158	102	0	2.29E-46	PSPH
15	64039217	66064168	648	417	0	1.69E-36	RBPMS2
2	85255429	87306752	528	333	0	3.04E-33	POLR1A
17	78478854	80486162	791	535	0	5.77E-33	ACTG1
5	51776562	53780492	582	357	0	4.02E-30	FST
5	80571352	82573932	456	258	0	4.59E-29	RPS23
12	123127921	125140714	635	411	0	9.87E-25	GTF2H3
16	61524	1818783	808	529	0	1.63E-23	MSLN
1	235681939	237713976	298	190	0	4.18E-23	LGALS8
11	58939168	60942636	376	235	0	1.77E-22	MS4A6A
8	132594914	134686827	307	192	0	2.55E-22	LRRC6
1	9273158	11431909	1133	680	0	4.49E-21	KIF1B
16	66485047	68513436	1116	744	0	9.74E-21	ATP6V0D1
8	40131993	42165247	339	209	0	2.16E-20	SFRP1
19	11036366	13088848	764	481	0	2.77E-20	ZNF763
19	54671887	56671530	753	470	0	3.58E-20	DNAAF3
17	71548271	73539560	784	472	0	6.33E-20	CD300C
3	45557584	47620971	534	333	0	1.06E-19	LRRC2
8	16021574	18080407	471	269	0	2.35E-19	ZDHHC2
1	16393711	18445005	861	505	0	2.39E-19	PADI2
17	79202744	81187821	773	520	0	7.46E-19	CSNK1D
12	56404058	58422458	1451	906	0	1.29E-18	TAC3
19	15989593	17988883	604	378	0	1.48E-18	SIN3B
12	52215270	54223128	1347	802	0	4.65E-18	KRT79
6	42023280	44026915	1038	652	0	2.20E-17	MRPL2
6	41847740	43856400	1027	650	0	2.73E-17	RPL7L1
1	24690676	26756608	745	472	0	3.35E-17	RHCE
2	134813827	136932159	453	299	0	9.16E-17	RAB3GAP1
1	201322470	203559602	728	478	0	1.22E-16	PPP1R12B
16	1004369	2990074	1034	705	0	2.07E-16	RPL3L
16	3522828	5543890	732	481	0	3.29E-16	NMRAL1
17	28295809	30326755	480	298	0	1.33E-15	RNF135
19	50630649	52626166	718	444	0	4.39E-15	SIGLEC9
11	64647848	66650805	1295	869	0	5.37E-15	CTSW
14	23440850	25472286	1326	849	0	7.75E-15	DHRS4L2
22	37207715	39192988	740	462	0	8.14E-15	GCAT

11	70829916	72849142	553	361	0	1.24E-14	FOLR3
9	110708672	112769529	599	366	0	1.24E-14	CTNNAL1
1	9063435	11072775	1017	603	0	1.27E-14	RBP7
4	5696136	7691291	292	190	0	2.98E-14	S100P
15	67500977	69549021	704	452	0	4.46E-14	CLN6
12	90496808	92504168	316	197	0	4.48E-14	LUM
10	98528249	100529570	806	502	0	7.89E-14	SFRP5
17	4195862	6288615	836	493	0	1.66E-13	RABEP1
2	131234265	133248784	299	154	0	3.40E-13	MZT2A
14	23424341	25438339	1335	851	0	4.37E-13	DHRS4
16	85563797	87574899	545	338	0	4.92E-13	MTHFSD
18	10989050	13030533	208	132	0	8.70E-13	IMPA2
14	22034663	24048442	1162	730	0	8.93E-13	DAD1
6	142384452	144659265	640	383	0	1.48E-12	AIG1
5	56787482	58785678	296	177	0	2.15E-12	GAPT
10	10786278	12803023	418	261	0	2.21E-12	ECHDC3
2	84822866	86820639	537	346	0	2.46E-12	RNF181
15	32058221	34486766	535	313	0	3.45E-12	FMN1
4	56678367	58679994	501	295	0	3.49E-12	SPINK2
6	137727849	139741886	561	347	0	4.07E-12	HEBP2
6	28641615	30648131	571	303	0	4.25E-12	ZFP57
1	16639761	18688663	821	473	0	4.77E-12	PADI4
8	58499893	60572308	638	405	0	4.91E-12	NSMAF
1	248112311	249234888	192	107	0	4.95E-12	SH3BP5L
16	88786822	90291311	527	364	0	5.33E-12	ZNF276
3	108047845	110054553	288	167	0	1.36E-11	DPPA4
17	72937650	74962382	1013	668	0	1.56E-11	ACOX1
10	59029408	61044084	351	203	0	1.67E-11	CISD1
13	49236382	51263464	433	280	0	2.18E-11	EBPL
14	22356192	24354883	1116	697	0	2.23E-11	REM2
16	71089779	73094855	1206	750	0	2.43E-11	HP
1	45869064	47878616	675	422	0	3.11E-11	FAAH
19	51125939	53117574	543	342	0	3.20E-11	SIGLEC5
5	111871020	113925487	405	243	0	4.23E-11	YTHDC2
17	45623285	47623015	1361	872	0	4.44E-11	HOXB2
19	47551714	49613137	887	559	0	6.68E-11	PLA2G4C
4	73981642	76158280	209	135	0	7.03E-11	MTHFD2L
1	166886211	168890090	377	228	0	7.80E-11	MPC2
19	54104530	56098746	761	477	0	8.53E-11	LILRA2
14	95858598	97952756	587	349	0	9.50E-11	AK7
19	39358860	41439248	890	573	0	1.22E-10	FCGBP
6	154159735	156577519	690	413	0	1.79E-10	TIAM2

6	30380084	32375324	1233	679	0	1.82E-10	MICA
1	149768998	151780099	936	581	0	1.87E-10	CTSK
7	74626276	76669255	392	240	0	1.90E-10	STYXL1
1	172873123	174883270	257	166	0	2.15E-10	SERPINC1
1	19923475	21944440	460	272	0	2.15E-10	CDA
19	10407094	12432117	704	470	0	2.16E-10	TSPAN16
13	30317837	32332540	297	197	0	2.20E-10	ALOX5AP
11	17434219	19472162	762	473	0	2.25E-10	LDHC
17	37119254	39123468	1169	766	0	2.26E-10	GSDMA
1	26323663	28326247	1184	776	0	2.26E-10	TRNP1
10	91985080	94041672	410	267	0	2.28E-10	PCGF5
17	44908286	46908429	1201	762	0	2.30E-10	MRPL10
1	27696376	29823421	865	543	0	2.31E-10	PHACTR4
1	24603413	26656184	711	436	0	2.32E-10	RHD
4	94375046	96572455	673	395	0	2.32E-10	PDLIM5
8	143956369	145948440	1209	826	0	2.40E-10	EPPK1
3	125247802	127261444	317	208	0	2.47E-10	CHST13
1	45161902	47213565	840	526	0	2.50E-10	IPP
6	31136994	33145709	1239	619	0	2.55E-10	AGPAT1
8	104345075	106478376	634	403	0	2.61E-10	DPYS
2	69198931	71188824	509	312	0	2.70E-10	ASPRV1
8	27461825	29603179	454	293	0	2.71E-10	EXTL3
14	23780665	25780153	1063	669	0	2.90E-10	LTB4R
1	202148429	204146674	594	373	0	3.13E-10	CHI3L1
11	57296270	59344358	477	287	0	3.41E-10	LPXN
2	160129032	162348252	912	534	0	3.42E-10	RBMS1
6	30804471	32797847	1199	609	0	3.63E-10	HSPA1B
6	9732498	11723589	611	368	0	3.67E-10	TMEM14C
1	115186162	117240389	538	297	0	4.12E-10	VANGL1
14	100008504	102050746	495	316	0	4.31E-10	BEGAIN
11	119899661	121959886	692	433	0	5.25E-10	TBCEL
1	205811842	207857291	559	337	0	5.47E-10	DYRK3
4	87236280	89236066	352	230	0.1	6.33E-10	HSD17B13
17	57125174	59155836	569	365	0.1	7.05E-10	HEATR6
13	99260793	101548733	710	450	0	8.03E-10	CLYBL
17	36619104	38719889	1321	853	0	1.53E-09	CDK12
12	12134087	14147562	553	341	0	1.85E-09	HEBP1
2	41990355	43989525	848	550	0.1	1.91E-09	OXER1
17	73671194	75695368	747	477	0	2.24E-09	MXRA7
6	159559020	161575294	277	176	0	2.31E-09	SLC22A1
16	87881845	89877462	610	420	0	2.58E-09	APRT
1	246579712	248605362	326	199	0	3.37E-09	NLRP3

17	72258136	74262187	1074	697	0	3.57E-09	MRPS7
6	51767239	53772775	414	260	0	3.93E-09	GSTA3
6	70386105	72558800	365	223	0	4.05E-09	SMAP1
11	65101255	67092273	1198	806	0	4.14E-09	RIN1
1	92914779	95018518	493	307	0	5.56E-09	FNBP1L
6	95509	1689616	331	208	0	6.19E-09	EXOC2
10	45330925	47149308	284	163	0	6.59E-09	AGAP4
22	40982244	42983560	667	434	0	6.98E-09	PMM1
2	25624804	27679481	765	513	0	6.99E-09	DRC1
4	109967852	112118290	612	378	0	7.14E-09	ELOVL6
10	124767496	126850278	691	446	0	7.45E-09	CHST15
22	18171776	20267960	610	373	0	9.78E-09	CLTCL1
19	57982233	59117009	332	223	0	9.98E-09	ZNF324
17	4402946	6515456	856	502	0	9.99E-09	NLRP1
6	31165563	33157187	1240	622	0	1.10E-08	PBX2
5	68370276	70370070	155	99	0	1.14E-08	SMN2
3	47641477	49644176	779	538	0	1.18E-08	UQCRC1
1	63063031	65124457	901	540	0	1.23E-08	PGM1
5	110478344	112742221	502	314	0	1.23E-08	EPB41L4A
17	76071253	78084624	639	405	0	1.25E-08	ENGASE
3	47727545	49776626	933	637	0	1.30E-08	IP6K2
16	65613470	67606351	904	543	0	1.33E-08	CMTM2
3	93531524	94824356	202	117	0	1.39E-08	NSUN3
14	22386204	24381055	1113	696	0	1.45E-08	RBM23
8	144162639	146157862	1223	836	0	1.47E-08	MAF1
5	130148476	132342594	474	288	0	1.54E-08	ACSL6
17	57228408	59246103	625	390	0.1	1.55E-08	CA4
11	9595099	11694454	663	385	0	1.79E-08	MRVI1
10	110972056	113044163	567	322	0	1.79E-08	MXI1
19	38293794	40288129	827	546	0	2.29E-08	LGALS4
6	33760232	35848956	702	422	0	2.41E-08	UHRF1BP1
12	120205104	122340955	867	537	0	2.68E-08	SPPL3
4	185325054	187341177	353	225	0	2.70E-08	UFSP2
7	129002356	130996151	533	353	0	2.70E-08	CPA5
2	238758634	240770102	259	167	0	2.71E-08	TWIST2
3	40291822	42996549	516	323	0	3.16E-08	ULK4
17	73559874	75569998	802	528	0	3.27E-08	ST6GALNAC2
4	73741064	75730068	213	142	0	3.35E-08	CXCL1
1	150594650	152671502	825	497	0	3.36E-08	SNX27
6	107882857	109999117	674	398	0	3.43E-08	FOXO3
1	160555781	162646901	644	408	0	3.48E-08	FCGR2B
9	33332474	35343197	701	417	0	3.63E-08	NUDT2

17	43594358	45656854	375	252	0.1	3.69E-08	ARL17A
17	47622779	49627679	1059	684	0	3.80E-08	SPATA20
7	27339266	29865343	1008	571	0	4.00E-08	CREB5
11	73529551	75657980	592	373	0	4.13E-08	XRRA1
17	29814935	31814120	536	331	0	4.20E-08	CDK5R1
15	54507848	56588718	359	227	0	4.29E-08	RAB27A
17	71466756	73477465	720	436	0	5.25E-08	CD300A
2	218135421	220200383	1140	691	0	5.41E-08	PNKD
12	93963526	96040830	461	286	0.1	5.71E-08	TMCC3
16	67022599	69106642	1085	723	0	6.41E-08	DUS2
14	66767726	68826014	668	399	0	7.13E-08	ATP6V1D
17	3618219	5622505	866	520	0	7.16E-08	ARRB2
8	143390000	145325695	952	630	0	7.36E-08	TOP1MT
5	51089439	53090866	561	358	0	7.58E-08	PELO
11	64591825	66627942	1340	902	0.1	7.69E-08	CFL1
1	27099837	29149659	924	584	0	8.19E-08	STX12
1	5484853	7519049	719	443	0	8.78E-08	ESPN
14	22433584	24425349	1108	692	0	8.89E-08	HAUS4
22	31018975	33057577	584	357	0.1	9.31E-08	PISD
11	87017	1756583	705	467	0	1.03E-07	TALDO1
17	65256227	67417142	461	287	0	1.05E-07	ARSG
12	80187578	82318213	427	261	0	1.05E-07	LIN7A
6	143263546	145384478	486	290	0	1.18E-07	PLAGL1
1	6906518	8909025	566	346	0	1.20E-07	UTS2
4	145414020	147472940	727	462	0	1.22E-07	SMAD1
3	121097698	123094128	484	315	0	1.25E-07	CCDC58
6	29691139	31691972	840	448	0	1.30E-07	TUBB
22	24616059	26622220	397	230	0	1.39E-07	CRYBB2
8	16916427	18942485	422	259	0	1.46E-07	ASAH1
14	64382441	66410464	632	397	0	1.48E-07	CHURC1
5	140338131	142362129	1483	888	0	1.55E-07	RNF14
8	27752679	29921408	443	283	0	1.58E-07	HMBOX1
15	63448627	65449169	592	369	0	1.62E-07	SNX22
6	34268704	36287380	677	413	0.1	1.64E-07	DEF6
4	102792767	104802130	260	153	0	1.65E-07	CISD2
1	206223875	208226181	599	362	0.1	1.67E-07	YOD1
6	31782661	33805117	1338	711	0	1.71E-07	TAP2
13	42600260	44682911	537	311	0	1.84E-07	DNAJC15
1	204197394	206239482	788	505	0	1.90E-07	TMCC2
4	355156	2349969	607	410	0.1	1.99E-07	UVSSA
20	32826904	34862608	801	528	0	2.04E-07	MMP24
6	51845212	53858060	383	238	0	2.06E-07	GSTA4

9	122366515	124476684	763	445	0.1	2.14E-07	MEGF9
2	134214696	136473148	448	287	0	2.15E-07	TMEM163
20	43644056	45644897	761	468	0	2.38E-07	MMP9
1	177994997	180044282	426	274	0	2.45E-07	FAM20B
20	42882351	44881207	776	480	0	2.54E-07	SLPI
10	72976939	74995025	650	376	0	2.55E-07	ANAPC16
2	190212419	192235016	455	269	0	2.65E-07	INPP1
6	41933605	43930372	1006	640	0	2.87E-07	GNMT
4	100946158	103267117	256	149	0	2.97E-07	PPP3CA
12	95367236	97389685	452	268	0	3.21E-07	HAL
20	352623	2370071	516	308	0	3.38E-07	FKBP1A
15	39736969	41758372	821	505	0	3.40E-07	BAHD1
16	83734456	85813446	604	402	0.1	3.50E-07	USP10
19	35399151	37398939	727	493	0	3.95E-07	TYROBP
6	40205095	42206259	725	442	0	4.00E-07	TREML4
11	129746599	131784377	801	481	0	4.02E-07	SNX19
1	32790316	34896636	818	492	0	4.13E-07	PHC2
12	132498057	133812414	294	184	0	4.29E-07	ZNF605
7	131940825	134743940	832	495	0	4.36E-07	EXOC4
19	38227885	40225031	839	554	0	4.39E-07	CAPN12
7	140610651	142797254	521	329	0	4.40E-07	MGAM
2	196636499	198664059	398	251	0	4.51E-07	GTF3C3
7	75027314	77069054	463	285	0	4.62E-07	ZP3
17	35337731	37357506	1003	604	0	4.76E-07	TBC1D3
11	73041373	75107265	622	378	0	4.77E-07	PGM2L1
6	132067998	134083309	668	391	0	4.82E-07	VNN2
18	11309303	13338617	228	142	0.1	4.89E-07	TUBB6
15	34509923	36838075	1135	670	0	5.06E-07	DPH6
7	79006735	81304532	336	206	0	5.18E-07	CD36
10	96480229	98636823	518	312	0	5.87E-07	ENTPD1
5	175735135	177733908	734	455	0	6.02E-07	PRELID1
16	68379570	70384938	760	474	0	6.25E-07	TMED6
11	66071832	68076009	988	631	0	6.50E-07	SSH3
3	10319976	12598571	492	295	0	6.59E-07	ATG7
22	41909029	43914449	574	353	0	6.62E-07	RRP7A
1	35622514	37612017	829	504	0	6.73E-07	TRAPPC3
1	27218969	29241192	869	544	0	7.24E-07	RPA2
1	149570886	151527784	889	548	0	7.32E-07	ADAMTSL4
14	20757422	22802797	753	498	0	7.32E-07	RPGRIP1
17	72205444	74223068	1081	702	0	7.52E-07	NUP85
17	6483131	8479918	1589	1010	0	7.97E-07	CD68
1	152347789	154335845	559	364	1	8.11E-07	S100A12

2	27615681	29637653	625	397	0	8.13E-07	FOSL2
1	9097709	11240589	1095	654	0	8.28E-07	UBE4B
1	153934779	155946585	1104	742	0.1	8.31E-07	SHC1
18	27649902	29679343	402	260	0	8.36E-07	DSC2
15	90427393	92438801	463	312	0	8.85E-07	FES
16	1080758	3088511	1096	747	0	9.06E-07	SLC9A3R2
7	126672632	128669786	916	580	0.1	9.14E-07	LRRC4
10	71437956	73521781	659	379	0	9.49E-07	ADAMTS14
12	68082264	70134035	301	190	0	9.55E-07	NUP107
20	41091953	43090684	566	334	0	9.59E-07	SRSF6
6	2154405	4151062	352	207	0	9.62E-07	TUBB2A
17	78527249	80615149	813	545	0	1.00E-06	NPLOC4
12	73945391	75920901	258	176	0	1.03E-06	ATXN7L3B
18	2069112	4219138	251	154	0	1.13E-06	MYOM1
6	29716711	31712032	854	453	0.1	1.21E-06	IER3
1	118574322	120678835	580	357	0	1.26E-06	WARS2
19	60658	2061656	1011	666	0	1.46E-06	ABCA7
14	69517619	71641508	513	316	0	1.52E-06	SLC8A3
17	45632482	47681054	1409	906	0.1	1.54E-06	HOXB3
16	56507237	58519628	607	372	0	1.59E-06	DOK4
1	9460183	11474259	1077	646	0	1.59E-06	PGD
7	74163379	76368084	364	231	0	1.62E-06	HIP1
22	19076970	21097036	557	357	0	1.63E-06	DGCR8
20	59758487	61777231	448	279	0	1.80E-06	MTG2
15	73836546	75890284	695	441	0.1	1.83E-06	ARID3B
9	69994568	72144275	439	256	0	1.85E-06	PGM5
18	8481285	10532788	353	232	0	1.86E-06	RALBP1
3	156825879	159260741	886	540	0	1.89E-06	RSRC1
7	75761306	77958641	498	316	0	1.91E-06	CCDC146
3	43756435	45761010	351	228	0	1.95E-06	ZNF502
2	23301174	25306970	730	458	0	2.01E-06	TP53I3
19	54105293	56113555	768	479	0	2.04E-06	LILRA1
1	224999010	227009199	398	234	1	2.07E-06	EPHX1
6	41933605	43941175	1010	641	0	2.10E-06	PEX6
3	49653871	51680484	962	627	0	2.11E-06	MAPKAPK3
3	194947440	197014817	434	255	0	2.11E-06	PCYT1A
7	63475980	65447211	195	108	0	2.16E-06	ZNF117
20	24246411	26248963	154	89	0	2.21E-06	PYGB
1	14487336	16546747	646	429	0	2.24E-06	TMEM51
6	159237508	161238481	275	175	0	2.27E-06	PNLDC1
1	20072025	22110160	439	261	0	2.37E-06	HP1BP3
8	19061309	21084014	443	271	0	2.42E-06	ATP6V1B2

5	175912923	177923756	701	435	0	2.45E-06	PDLIM7
1	170222420	172253389	459	293	0	2.49E-06	FMO1
17	36921986	39020028	1232	802	0	2.51E-06	IKZF3
19	3358556	5358897	727	481	0	2.53E-06	MPND
1	39542280	41562630	883	545	0	2.57E-06	PPT1
12	55614925	57615493	1157	716	0	2.58E-06	RNF41
16	19636366	21709125	516	295	0	2.64E-06	ACSM1
13	110295378	112345853	199	128	0	2.66E-06	CARS2
5	174786329	176787075	507	303	0	2.78E-06	KIAA1191
13	43398858	45449542	478	281	0	2.81E-06	CCDC122
22	44559753	46567121	336	209	0.1	2.83E-06	NUP50
22	23590811	25581023	428	260	0	2.89E-06	SUSD2
8	144637367	146293075	1098	759	0	2.94E-06	SLC39A4
14	50002945	52097847	565	360	0	2.98E-06	ATL1
7	28038024	30228621	760	431	0	3.09E-06	CPVL
15	88632760	90744964	697	441	0	3.11E-06	ABHD2

BEGIN POS : Beginning position of range for rare variants within 1 Mb of gene to be tested

END POS : End position of range for rare variants within 1 Mb of gene to be tested

NUM PASS VARS : Number of variants passing all thresholds for EPACTS software

NUM SING VARS : Number of singletons among variants in NUM PASS VARS

P-VALUE : P-value of burden tests

STATRHO: represents the RHO value from SKAT-O test, rho = 1 (burden) and rho = 0 (SKAT)

Table S3.4. Individual Gene-SNP eQTL significant results in brain

Gene	SNP	P-value
RP11-529K1	16:69600184	<1.0E-314
ZNF101	19:19120659	3.96E-61
COPZ1	12:54520272	1.30E-54
COPZ1	12:53853136	1.30E-54
COPZ1	12:54367799	1.30E-54
COPZ1	12:54511341	1.30E-54
COPZ1	12:54986339	1.30E-54
TMPRSS6	22:38340452	1.68E-52
TMPRSS6	22:38206144	1.68E-52
TMPRSS6	22:38093050	1.68E-52
LDLRAD1	1:55481065	1.02E-42
ZNF253	19:19120659	7.47E-36
VRTN	14:73943427	1.88E-35
TMPRSS6	22:37823469	1.91E-34
GALNT5	2:157192762	1.85E-32
ICAM4	19:10742166	6.67E-30
OPN1SW	7:128499675	1.04E-27
RP11-529K1	16:70512234	6.26E-23
TAS2R9	12:11802124	7.77E-23
RPL10A	6:34462967	5.50E-18
RPL10A	6:35543534	5.50E-18
TMPRSS6	22:38412776	5.99E-17
TRIM63	1:26869683	1.51E-16
TRIM63	1:27028031	1.51E-16
ZNF101	19:19529415	2.13E-16
COPZ1	12:54623909	3.80E-15
TRIM63	1:26872452	3.80E-15
TRIM63	1:26485547	3.80E-15
TRIM63	1:26560692	3.80E-15
KLRF2	12:9268393	1.12E-12
CARD17	11:104480278	1.19E-12
VRTN	14:75707465	1.30E-12
VRTN	14:75686993	1.30E-12
ICAM4	19:10781827	4.81E-12
ICAM4	19:10952607	2.62E-11
ICAM4	19:11319636	7.40E-11
TMPRSS6	22:36678827	8.40E-11
ZNF253	19:19529415	3.43E-10
TMPRSS6	22:36745146	7.53E-10

FAM186A	12:51566637	1.59E-09
DNLZ	9:139694595	1.92E-09
C5orf17	5:23132477	3.25E-09
GALNT5	2:157861592	5.11E-09
LDLRAD1	1:54200597	6.58E-09
OPN1SW	7:127668180	6.68E-09
RPL10A	6:35436943	7.99E-09
OPN1SW	7:127560949	2.14E-08
C5orf17	5:24322422	3.23E-08
POLD2	7:43489734	5.54E-08
POMC	2:24583125	8.95E-08
GALNT5	2:157716598	9.43E-08
POLD2	7:44646209	1.23E-07
C2orf74	2:60683420	2.32E-07
C2orf74	2:60753967	2.32E-07
DNLZ	9:140082242	6.08E-07
OPN1SW	7:127696479	7.98E-07
COPZ1	12:54673746	1.21E-06
ICAM4	19:11034582	1.30E-06
OPN1SW	7:128494922	1.41E-06
LDHC	11:17867845	1.60E-06
OPN1SW	7:127953296	1.66E-06

Table S3.5. Individual Gene-SNP eQTL significant results in blood

Gene	SNP	P-value	Gene	SNP	P-value
MTHFD2L	4:74241201	4.09E-163	DAD1	14:23041489	2.70E-53
MTHFD2L	4:75284312	4.09E-163	DAD1	14:23041518	2.70E-53
FST	5:52778262	1.20E-147	RBPMS2	15:64991778	3.17E-53
GAD1	2:171673464	1.09E-145	DPH6	15:34647703	2.61E-52
GAD1	2:171679654	4.09E-145	DPH6	15:35333208	2.61E-52
NUMA1	11:71714526	1.17E-133	DPH6	15:35514161	2.61E-52
MSLN	16:824837	1.04E-114	DPH6	15:36699896	2.61E-52
NUMA1	11:71822560	1.25E-112	GAD1	2:172144563	4.73E-48
GAD1	2:171695387	4.06E-111	MS4A6A	11:59863104	4.36E-47
SFRP1	8:41200160	5.46E-108	LRRC2	3:46878319	6.89E-47
NUMA1	11:71706544	3.08E-72	ACTG1	17:79477190	1.49E-42
RPL7L1	6:42857315	3.59E-71	MS4A6A	11:59863253	6.38E-42
KRT79	12:52291229	5.45E-71	ADAMTS14	10:72509641	2.59E-39
KRT79	12:52390683	5.45E-71	ADAMTS14	10:72699774	3.12E-39
KRT79	12:52938512	5.45E-71	ADAMTS14	10:72898175	3.12E-39
KRT79	12:53575608	5.45E-71	GAD1	2:172641848	8.57E-39
KRT79	12:53808590	5.45E-71	PLAGL1	6:144385777	2.53E-38
KRT79	12:53926760	5.45E-71	SFRP5	10:99644019	5.62E-38
KRT79	12:53952253	5.45E-71	SFRP1	8:41792185	1.02E-37
KRT79	12:54145225	5.45E-71	SFRP1	8:41154024	8.03E-37
KRT79	12:52652265	1.10E-70	AK7	14:97272564	8.48E-37
GAPT	5:57652956	5.01E-68	AK7	14:97399418	8.48E-37
GAD1	2:171674182	5.88E-68	ACTG1	17:79477575	8.65E-37
NUMA1	11:71726168	2.06E-64	ACTG1	17:79478617	8.65E-37
TAC3	12:57398085	4.76E-62	CTSW	11:65631361	1.72E-36
TAC3	12:57420802	4.76E-62	CTSW	11:65652177	1.72E-36
TAC3	12:57464587	4.76E-62	CTSW	11:65659093	2.26E-36
TAC3	12:57472766	4.76E-62	FST	5:52949030	2.98E-35
TAC3	12:57483769	4.76E-62	KRT79	12:53724748	4.79E-35
TAC3	12:57490491	4.76E-62	CTSW	11:65640259	1.28E-34
TAC3	12:57492307	4.76E-62	FAAH	1:46508496	2.14E-34
TAC3	12:57114418	1.60E-61	FAAH	1:46672262	2.14E-34
TAC3	12:57115196	1.60E-61	SMAP1	6:71570880	7.55E-33
TAC3	12:57595605	1.60E-61	SFRP1	8:41196024	1.04E-32
TAC3	12:57620836	1.60E-61	KRT79	12:52388470	1.06E-32
TAC3	12:58240986	1.60E-61	KRT79	12:52388505	1.06E-32
TAC3	12:58347548	1.60E-61	TAC3	12:57865841	4.12E-31
NUMA1	11:71645655	2.51E-59	DNAAF3	19:54675653	4.99E-30
NUMA1	11:71701763	2.51E-59	DNAAF3	19:55912805	4.99E-30

TRAPPC3	1:36602202	8.28E-59	DNAAF3	19:56189999	4.99E-30
DAD1	14:23035455	2.70E-53	LGALS4	19:39616553	9.60E-30
KRT79	12:54219108	2.51E-29	SLC39A4	8:145639654	1.16E-21
ADAMTSL4	1:150524878	5.50E-29	IKZF3	17:37957632	1.25E-21
LGALS8	1:236706278	2.85E-28	TWIST2	2:239772983	1.84E-21
LGALS8	1:236706300	2.85E-28	KIF1B	1:10408850	2.03E-21
MTHFSD	16:86603266	6.30E-27	RPL3L	16:1470825	2.25E-21
FOLR3	11:71720030	1.22E-26	RPL3L	16:2086798	2.25E-21
KRT79	12:52398857	1.93E-26	RPL3L	16:2223816	2.25E-21
NMRAL1	16:4536354	3.76E-26	RPL3L	16:2284726	2.25E-21
LRRC2	3:46140074	1.57E-25	RPL3L	16:2301527	2.25E-21
KRT79	12:52824349	1.64E-25	CDK12	17:37755725	3.28E-21
TIAM2	6:155577913	3.31E-25	POLR1A	2:86529625	4.54E-21
TIAM2	6:155597134	3.31E-25	POLR1A	2:86293004	4.88E-21
LGALS4	19:39162511	3.55E-25	RBPMS2	15:65042560	5.85E-21
LGALS4	19:39390531	3.55E-25	PNLDC1	6:159420734	6.31E-21
LGALS4	19:39804846	3.55E-25	LDHC	11:18627931	8.26E-21
MPND	19:4359191	4.00E-25	PNLDC1	6:160372070	1.11E-20
FAAH	1:46032307	6.50E-25	MZT2A	2:132290912	1.17E-20
KRT79	12:52779003	6.65E-25	ZNF276	16:89644001	1.17E-20
KRT79	12:52783593	6.65E-25	CSNK1D	17:80202204	1.27E-20
KRT79	12:53652706	6.65E-25	CCDC146	7:76889524	1.68E-20
GSDMA	17:38077938	9.23E-25	KIF1B	1:10271628	1.70E-20
KRT79	12:53846595	1.20E-24	KIF1B	1:10273615	1.70E-20
KRT79	12:53268234	3.76E-24	KIF1B	1:10314057	1.70E-20
GTF2H3	12:124068015	3.89E-24	KIF1B	1:10334544	1.70E-20
PDLIM5	4:95550822	4.48E-24	KIF1B	1:10352228	1.70E-20
PDLIM5	4:95551261	4.48E-24	KIF1B	1:10357868	1.70E-20
HSPA1B	6:31852604	6.06E-24	ZNF763	19:12059467	2.09E-20
TRNP1	1:27272864	6.26E-24	RPS23	5:81013700	3.86E-20
DPH6	15:34610561	7.39E-24	DNAAF3	19:55670651	4.18E-20
ZNF276	16:89358050	1.91E-23	DNAAF3	19:55677886	4.18E-20
FMO1	1:172062702	2.46E-23	GAD1	2:170675959	4.39E-20
FMO1	1:172062773	2.46E-23	AK7	14:97906815	9.65E-20
FMO1	1:172094477	2.46E-23	DHRS4L2	14:23468904	1.12E-19
FMO1	1:172108518	2.46E-23	DHRS4L2	14:23537199	1.12E-19
IKZF3	17:37938977	2.92E-23	DHRS4L2	14:23572193	1.12E-19
CLYBL	13:100518634	6.14E-23	DHRS4L2	14:23857458	1.12E-19
UQCRC1	3:48687271	6.44E-23	DHRS4L2	14:23990308	1.12E-19
RPL7L1	6:42981633	8.03E-23	DHRS4L2	14:24015983	1.12E-19
PSPH	7:55759098	9.72E-23	DHRS4L2	14:24523950	1.12E-19
DPH6	15:35046585	1.32E-22	DHRS4L2	14:24841978	1.12E-19

RPL3L	16:2813177	1.48E-22	DHRS4L2	14:24900671	1.12E-19
RBPMS2	15:64792896	1.82E-22	DHRS4L2	14:24123454	1.31E-19
FST	5:52386315	5.79E-22	TMED6	16:69304048	1.84E-19
LDHC	11:18550693	1.12E-21	MSLN	16:1245453	3.37E-19
GAD1	2:171568928	3.41E-19	KRT79	12:53899633	8.53E-16
ADAMTS14	10:72687084	5.05E-19	FOLR3	11:71923660	8.73E-16
CDK12	17:37557599	5.30E-19	RPL3L	16:2161117	8.97E-16
GSDMA	17:38062503	7.17E-19	RPL3L	16:2240341	8.97E-16
MTHFD2L	4:75579573	9.81E-19	LRRC2	3:46033545	1.00E-15
ZNF276	16:89642150	1.11E-18	CTNNAL1	9:111679940	1.12E-15
KIF1B	1:10218472	1.29E-18	CTNNAL1	9:111692163	1.12E-15
SMN2	5:68898565	1.34E-18	CTNNAL1	9:111696795	1.12E-15
GSDMA	17:37938977	1.37E-18	ABCA7	19:1009413	1.22E-15
KRT79	12:52913822	1.71E-18	KRT79	12:53207721	1.24E-15
CDK12	17:37557517	2.50E-18	UTS2	1:8137400	1.29E-15
TAC3	12:58405892	2.67E-18	DNAAF3	19:54809236	1.38E-15
KRT79	12:52311687	4.12E-18	SFRP5	10:100008317	1.64E-15
SOS1	2:40117437	4.41E-18	PSPH	7:55902230	1.73E-15
GSDMA	17:37957632	5.11E-18	GSDMA	17:38251933	1.80E-15
FMO1	1:170431347	6.05E-18	PSPH	7:56327785	1.96E-15
FMO1	1:170443845	6.05E-18	SFRP5	10:100248732	2.12E-15
LRRC2	3:46759204	7.97E-18	DGCR8	22:20068723	2.18E-15
CDK12	17:36720974	9.47E-18	RPL7L1	6:43182906	2.28E-15
LRRC2	3:46786162	1.42E-17	CDK12	17:37792293	2.37E-15
RPS23	5:82096819	1.49E-17	TAC3	12:57637604	2.39E-15
RPS23	5:82100144	1.49E-17	TAC3	12:57654608	2.39E-15
DHRS4L2	14:24647334	1.70E-17	TAC3	12:58120775	2.39E-15
LRRC2	3:45987980	2.41E-17	KIF1B	1:9976862	3.53E-15
PELO	5:52074164	3.44E-17	SOS1	2:39187008	3.61E-15
GSDMA	17:37933468	3.50E-17	TAC3	12:57610460	3.64E-15
GSDMA	17:37975856	3.50E-17	TAC3	12:57635879	3.64E-15
DUS2	16:68072004	5.12E-17	TAC3	12:57637127	3.64E-15
LIN7A	12:81331258	5.31E-17	LRRC2	3:45942686	4.01E-15
LGALS8	1:236767288	5.94E-17	IKZF3	17:38062503	4.15E-15
KRT79	12:52284538	6.34E-17	ELOVL6	4:110000229	4.98E-15
PDLIM5	4:94574633	9.26E-17	ELOVL6	4:110127740	4.98E-15
SOS1	2:38301803	1.70E-16	ELOVL6	4:110303078	4.98E-15
IKZF3	17:37933468	2.04E-16	SFRP5	10:100245620	6.68E-15
IKZF3	17:37975856	2.04E-16	TRNP1	1:27321880	8.18E-15
TIAM2	6:155649950	2.06E-16	DHRS4L2	14:24457160	1.23E-14
LPXN	11:58034597	2.86E-16	DPH6	15:35766600	1.48E-14
LPXN	11:58034651	2.86E-16	SFRP5	10:99991408	1.80E-14

POLR1A	2:86563488	3.04E-16	POLR1A	2:86333375	1.91E-14
POLR1A	2:86832394	3.13E-16	POLR1A	2:87302514	2.27E-14
MZT2A	2:132218737	3.30E-16	TAC3	12:57589784	2.30E-14
CCDC122	13:44455493	4.64E-16	TAC3	12:57590869	2.30E-14
FOLR3	11:71850152	5.31E-16	KIF1B	1:10271490	2.56E-14
TAC3	12:58235544	7.22E-16	ZFP57	6:29638474	2.65E-14
GCAT	22:38379774	3.66E-14	LPXN	11:58125774	5.77E-13
IKZF3	17:37912268	3.79E-14	DNAAF3	19:55995299	5.87E-13
RPL3L	16:1705955	4.00E-14	RAB3GAP1	2:135960616	6.11E-13
DNAAF3	19:56369727	4.00E-14	RPL3L	16:1536140	6.24E-13
DNAAF3	19:56631221	4.00E-14	RHCE	1:24792324	6.34E-13
DUS2	16:67596482	4.26E-14	ZNF324	19:59082368	6.71E-13
ZDHHC2	8:17017140	4.39E-14	RHCE	1:26203733	6.92E-13
DUS2	16:67696365	4.40E-14	CD300C	17:73221439	7.13E-13
KRT79	12:53729630	4.48E-14	DHRS4	14:23468904	7.25E-13
RPL7L1	6:43188600	5.41E-14	DHRS4	14:23537199	7.25E-13
MSLN	16:1560989	5.49E-14	DHRS4	14:23572193	7.25E-13
UQCRC1	3:48506279	5.68E-14	DHRS4	14:23857458	7.25E-13
PSPH	7:55944443	5.93E-14	DHRS4	14:23990308	7.25E-13
TAC3	12:57472762	6.38E-14	DHRS4	14:24015983	7.25E-13
KRT79	12:53700848	6.38E-14	DHRS4	14:24841978	7.25E-13
DUS2	16:67860637	6.68E-14	DHRS4	14:24900671	7.25E-13
DUS2	16:67867739	6.68E-14	CDK12	17:36945701	7.25E-13
UQCRC1	3:48463799	7.93E-14	MSLN	16:846041	8.78E-13
RNF181	2:85824251	8.38E-14	CSNK1D	17:79871981	1.05E-12
RPL3L	16:1262014	9.33E-14	DYRK3	1:207793338	1.07E-12
DPH6	15:35581602	1.04E-13	NUP107	12:69080770	1.10E-12
DPH6	15:35673786	1.04E-13	NUP107	12:69080771	1.10E-12
MTHFSD	16:86539482	1.10E-13	ZNF763	19:12089805	1.16E-12
SNX22	15:64392388	1.24E-13	GCAT	22:38328597	1.16E-12
SNX22	15:64392389	1.24E-13	GCAT	22:38221141	1.22E-12
LGALS4	19:38719246	1.48E-13	DHRS4L2	14:24179039	1.34E-12
PNLDC1	6:160174463	1.67E-13	DHRS4L2	14:24179040	1.34E-12
PNLDC1	6:160183445	1.67E-13	SOS1	2:38750325	1.43E-12
CDK12	17:37826249	1.96E-13	KRT79	12:52984733	1.63E-12
CDA	1:20960385	2.15E-13	FMO1	1:171605526	1.68E-12
CDK12	17:36831189	2.31E-13	CSNK1D	17:79892586	1.74E-12
DHRS4	14:24457160	2.72E-13	DNAAF3	19:56303806	1.75E-12
GAPT	5:58735345	2.78E-13	IMPA2	18:11987767	2.01E-12
DHRS4L2	14:24676498	3.07E-13	GTF2H3	12:123593218	2.27E-12
RPS23	5:81830253	3.14E-13	LGALS4	19:39421274	2.28E-12
PSPH	7:56125757	3.31E-13	LGALS4	19:39423001	2.28E-12

LUM	12:91316180	3.59E-13	POLR1A	2:86297105	2.45E-12
DUS2	16:67516945	4.11E-13	DPYS	8:106337440	2.56E-12
CTNNAL1	9:111678508	4.44E-13	DPYS	8:106349393	2.56E-12
RAB3GAP1	2:136170213	4.50E-13	CDK12	17:38312249	2.93E-12
TAC3	12:58141938	4.80E-13	S100P	4:6548580	2.95E-12
ZDHHC2	8:16885062	5.07E-13	S100P	4:6728934	2.95E-12
DHRS4	14:24123454	5.08E-13	DNAAF3	19:56111573	3.27E-12
DHRS4	14:24523950	5.64E-13	RHCE	1:25773026	3.49E-12
RHCE	1:25803162	3.49E-12	CHURC1	14:65322673	1.07E-11
GTF2H3	12:124135814	3.67E-12	MSLN	16:1279694	1.07E-11
POLR1A	2:86564490	3.70E-12	MMP24	20:33879965	1.22E-11
CD68	17:7469015	3.87E-12	LPXN	11:58377817	1.24E-11
FST	5:53236938	4.57E-12	ENTPD1	10:97583002	1.36E-11
ATP6V0D1	16:67262775	5.02E-12	IPP	1:46126179	1.79E-11
ATP6V0D1	16:67264629	5.02E-12	DUS2	16:67386135	1.82E-11
ATP6V0D1	16:67329327	5.02E-12	SPATA20	17:48603434	1.87E-11
ATP6V0D1	16:67393428	5.02E-12	RPS23	5:82395106	1.92E-11
PBX2	6:32143223	5.12E-12	IPP	1:46145295	2.02E-11
PBX2	6:32149260	5.12E-12	CTSK	1:150789864	2.06E-11
ZDHHC2	8:16925729	5.44E-12	PBX2	6:32147478	2.07E-11
PPP1R12B	1:202573179	5.81E-12	CDK12	17:38137213	2.15E-11
CDK12	17:37427426	6.08E-12	TRNP1	1:27468318	2.27E-11
CDK12	17:37780786	6.08E-12	TRNP1	1:27472451	2.27E-11
CDK12	17:37816702	6.08E-12	PGM2L1	11:74550031	2.31E-11
CDK12	17:37822739	6.08E-12	FST	5:51794332	2.62E-11
CDK12	17:37910592	6.08E-12	KRT79	12:52652088	2.62E-11
CDK12	17:37912069	6.08E-12	SOS1	2:38735689	2.86E-11
CDK12	17:38020470	6.08E-12	CLYBL	13:100513203	3.00E-11
CDK12	17:38051979	6.08E-12	ENTPD1	10:97516401	3.18E-11
CDK12	17:38122591	6.08E-12	TAC3	12:57396842	3.39E-11
CDK12	17:38471182	6.08E-12	ASAH1	8:17731611	3.45E-11
CDK12	17:38519592	6.08E-12	POLR1A	2:86668341	3.47E-11
CDK12	17:38597395	6.08E-12	NMRAL1	16:4417160	3.78E-11
CDK12	17:38601823	6.08E-12	DUS2	16:67269998	3.80E-11
CDK12	17:38703987	6.08E-12	SIN3B	19:17355669	3.80E-11
SOS1	2:38366689	6.16E-12	DUS2	16:68395522	3.91E-11
MTHFSD	16:86510183	6.62E-12	GAD1	2:171547270	4.14E-11
DHRS4L2	14:24042936	6.80E-12	SOS1	2:38525660	4.26E-11
POLR1A	2:85981797	7.08E-12	DUS2	16:67357938	4.36E-11
FMO1	1:171177988	7.29E-12	GTF2H3	12:123282708	4.74E-11
GSDMA	17:37912268	7.59E-12	GTF2H3	12:123378754	4.74E-11
RPL3L	16:2590730	8.53E-12	AK7	14:97601546	4.93E-11

GAD1	2:171260887	8.86E-12	SPINK2	4:58397567	5.44E-11
DPYS	8:105405411	9.27E-12	SPINK2	4:58397610	5.44E-11
SIN3B	19:17654384	9.31E-12	FNBP1L	1:94792971	5.59E-11
DHRS4L2	14:23862997	9.86E-12	SIN3B	19:16038111	5.85E-11
MMP24	20:33575613	1.02E-11	NMRAL1	16:4422744	5.91E-11
MMP24	20:34526720	1.02E-11	ATP6V0D1	16:67475439	6.04E-11
MMP24	20:34755884	1.02E-11	ATP6V0D1	16:67492248	6.04E-11
MMP24	20:34767630	1.02E-11	RHCE	1:24902925	6.12E-11
MMP24	20:34833404	1.02E-11	DHRS4	14:24647334	6.27E-11
MMP24	20:34862450	1.02E-11	FMO1	1:170561948	6.60E-11
ALOX5AP	13:31127975	6.68E-11	EXOC2	6:598588	2.33E-10
SLC22A1	6:160557643	6.85E-11	AK7	14:96195958	2.35E-10
RNF181	2:85846329	7.07E-11	FMO1	1:172053810	2.43E-10
GAPT	5:57653771	7.28E-11	CREB5	7:29045305	2.43E-10
SOS1	2:39462104	7.35E-11	KRT79	12:52312870	2.62E-10
CTSK	1:150550924	8.16E-11	SPINK2	4:58405566	2.78E-10
SIN3B	19:17007363	8.33E-11	SPINK2	4:58407594	2.78E-10
TMEM14C	6:10687746	8.42E-11	SPINK2	4:58420910	2.78E-10
DNAAF3	19:56111377	9.23E-11	SOS1	2:38709757	2.83E-10
SLC8A3	14:70653758	1.10E-10	SPATA20	17:48614426	2.83E-10
LRRC6	8:133621081	1.11E-10	LRRC2	3:46916051	3.07E-10
GSDMA	17:37884176	1.19E-10	TMEM14C	6:9820723	3.31E-10
RBP7	1:11072691	1.20E-10	RAB27A	15:55489250	3.31E-10
FNBP1L	1:93579198	1.24E-10	RPS23	5:81952164	3.41E-10
ULK4	3:41715859	1.30E-10	MSLN	16:1359037	3.41E-10
BAHD1	15:40760444	1.39E-10	GTF2H3	12:123970370	3.59E-10
DOK4	16:58011727	1.41E-10	LILRA1	19:55998281	3.89E-10
SIN3B	19:17727401	1.54E-10	TRNP1	1:26394301	3.93E-10
SNX22	15:65158047	1.57E-10	DYRK3	1:205902130	4.01E-10
TSPAN16	19:11436221	1.61E-10	FOLR3	11:71800166	4.10E-10
RHCE	1:24742263	1.62E-10	RPS23	5:81472669	4.26E-10
LILRA1	19:54275574	1.62E-10	TAC3	12:57912041	4.45E-10
LILRA1	19:54677876	1.62E-10	DHRS4	14:24676498	4.49E-10
LILRA1	19:54678025	1.62E-10	KRT79	12:52946594	4.54E-10
LILRA1	19:54866992	1.62E-10	LUM	12:91643350	4.68E-10
LILRA1	19:55598927	1.62E-10	LRRC2	3:45786881	4.71E-10
LILRA1	19:55659519	1.62E-10	TMED6	16:69377470	4.74E-10
LILRA1	19:56011880	1.62E-10	DNAAF3	19:54848685	4.82E-10
CHI3L1	1:202301584	1.66E-10	DNAAF3	19:54848686	4.82E-10
CHI3L1	1:202537552	1.66E-10	RHCE	1:25573797	4.97E-10
RBPMS2	15:64753826	1.70E-10	PGM2L1	11:74842610	5.26E-10
MSLN	16:1292120	1.71E-10	PGM2L1	11:75062736	5.26E-10

ZFP57	6:29691582	1.77E-10	RPS23	5:81741089	5.60E-10
ACOX1	17:73958795	1.79E-10	RPS23	5:81763218	5.60E-10
LILRA1	19:54974607	1.83E-10	GAD1	2:171361376	5.66E-10
RPS23	5:82154463	1.91E-10	LILRA2	19:54420959	5.75E-10
CAPN12	19:38719964	2.02E-10	SFRP5	10:100234305	5.87E-10
CAPN12	19:38958397	2.02E-10	POLR1A	2:85383138	6.45E-10
CAPN12	19:39039058	2.02E-10	CD300C	17:73263012	6.68E-10
CAPN12	19:39416898	2.02E-10	RNF135	17:29061941	7.05E-10
SFRP5	10:99523005	2.06E-10	TUBB6	18:12378679	7.17E-10
AK7	14:97580202	2.12E-10	SPINK2	4:57215989	7.21E-10
CAPN12	19:39340412	2.20E-10	SPINK2	4:58104786	7.21E-10
CAPN12	19:39875359	2.20E-10	SPINK2	4:58459198	7.21E-10
RNF135	17:29894457	7.47E-10	TAC3	12:57619341	1.97E-09
RHCE	1:25942070	7.65E-10	CTSK	1:150943873	2.07E-09
MRPL2	6:43040666	7.93E-10	FOSL2	2:27975744	2.13E-09
SMAD1	4:146480111	7.95E-10	PHACTR4	1:28830349	2.27E-09
ZNF324	19:59074429	8.13E-10	FOSL2	2:27746230	2.31E-09
MRPL2	6:42933526	8.16E-10	NUDT2	9:34400980	2.38E-09
GTF2H3	12:124092033	8.25E-10	SOS1	2:39204965	2.49E-09
GTF2H3	12:124144359	8.25E-10	SOS1	2:39695138	2.49E-09
GTF2H3	12:124499652	8.28E-10	SOS1	2:39731853	2.49E-09
GTF2H3	12:124856569	8.28E-10	CTSW	11:64882506	2.50E-09
MRPS7	17:73472561	8.41E-10	SIN3B	19:17450038	2.51E-09
ZNF763	19:12790506	8.46E-10	PNLDC1	6:159688848	2.55E-09
ZNF763	19:12792377	8.90E-10	LUM	12:91346671	2.58E-09
DUS2	16:67424412	9.60E-10	LUM	12:91347999	2.58E-09
CAPN12	19:39904557	9.60E-10	TRNP1	1:26861894	2.66E-09
LGALS8	1:236723074	9.89E-10	NLRP3	1:247614657	2.69E-09
LGALS8	1:236737983	9.89E-10	TRNP1	1:27024226	2.88E-09
RPS23	5:82388211	1.01E-09	DHRS4	14:24042936	2.89E-09
GSDMA	17:37911048	1.06E-09	TAC3	12:57872983	2.94E-09
HP	16:72088431	1.12E-09	ZNF324	19:58907700	2.96E-09
ECHDC3	10:12644023	1.13E-09	CTSW	11:64944795	2.97E-09
POLR1A	2:86371233	1.18E-09	FNBP1L	1:93913579	3.05E-09
LRRC6	8:133242503	1.22E-09	SOS1	2:39403006	3.12E-09
GAD1	2:172110323	1.30E-09	RBPMS2	15:65369947	3.17E-09
POLR1A	2:86848893	1.37E-09	PDLIM5	4:95442673	3.22E-09
SOS1	2:38893450	1.39E-09	DHRS4L2	14:23743265	3.43E-09
GTF2H3	12:123568888	1.39E-09	FCGBP	19:40502943	3.71E-09
POLR1A	2:86323339	1.40E-09	FCGR2B	1:161882189	3.90E-09
CPA5	7:129938259	1.46E-09	BAHD1	15:40759347	3.99E-09
AGAP4	10:46285954	1.46E-09	RPL3L	16:1412320	4.25E-09

SFRP1	8:40330425	1.48E-09	RPL3L	16:1638013	4.25E-09
GSDMA	17:38460198	1.48E-09	RPL3L	16:2835376	4.25E-09
ZNF763	19:12812169	1.48E-09	MRPL2	6:42980218	4.31E-09
SUSD2	22:24544631	1.50E-09	RHCE	1:24743223	4.40E-09
AGAP4	10:46258869	1.57E-09	AGAP4	10:46321905	4.45E-09
HP1BP3	1:21106416	1.75E-09	ATL1	14:51186527	4.65E-09
RHCE	1:25664799	1.83E-09	ESPN	1:6520668	4.74E-09
S100P	4:5857909	1.86E-09	DYRK3	1:207221096	4.77E-09
DPYS	8:105601451	1.88E-09	MSLN	16:737278	4.84E-09
PCYT1A	3:195965344	1.90E-09	MSLN	16:747032	4.84E-09
CTSK	1:150894082	1.97E-09	MSLN	16:1255223	4.84E-09
CTSK	1:150939883	1.97E-09	DPH6	15:36165056	4.91E-09
CTSK	1:150940308	1.97E-09	RABEP1	17:5037281	4.96E-09
ANAPC16	10:73964243	1.97E-09	RBPMS2	15:65790090	5.01E-09
PPP1R12B	1:201952244	5.02E-09	ECHDC3	10:11789382	1.13E-08
CISD1	10:59956041	5.02E-09	PADI2	1:17594457	1.16E-08
SOS1	2:38842988	5.03E-09	PLA2G4C	19:48558159	1.18E-08
ZDHHC2	8:16767217	5.23E-09	RHCE	1:24706965	1.19E-08
AK7	14:96181934	5.33E-09	RBPMS2	15:64458405	1.20E-08
RABEP1	17:5044794	5.53E-09	DPH6	15:35008509	1.21E-08
HEATR6	17:58177769	5.66E-09	CMTM2	16:67262775	1.22E-08
ACSM1	16:20587568	5.93E-09	CMTM2	16:67264629	1.22E-08
SPINK2	4:57425674	6.13E-09	CMTM2	16:67329327	1.22E-08
MRPL2	6:43128519	6.49E-09	CMTM2	16:67393428	1.22E-08
MRPL2	6:43147861	6.49E-09	IP6K2	3:47933439	1.31E-08
MRPL2	6:43155458	6.49E-09	PADI2	1:16950389	1.32E-08
RABEP1	17:4849384	6.56E-09	LRRC4	7:126738011	1.32E-08
LRRC2	3:47322700	7.06E-09	DUS2	16:67058614	1.32E-08
STX12	1:28028261	7.14E-09	GTF2H3	12:124102329	1.33E-08
MICA	6:31378864	7.15E-09	ELOVL6	4:111134733	1.34E-08
SPINK2	4:58407327	7.32E-09	DPH6	15:34678870	1.34E-08
AGAP4	10:46282089	7.34E-09	PPT1	1:40425935	1.38E-08
MRPL2	6:42307884	7.70E-09	RNF135	17:28749880	1.38E-08
MRPL2	6:42308855	7.70E-09	RNF135	17:28853723	1.38E-08
MRPL2	6:43019429	7.70E-09	FNBP1L	1:93128629	1.39E-08
MRPL2	6:43472947	7.70E-09	FNBP1L	1:93489164	1.39E-08
MRPL2	6:43501753	7.70E-09	FNBP1L	1:93586068	1.39E-08
MRPL2	6:43815319	7.70E-09	FNBP1L	1:94091316	1.39E-08
TAC3	12:57997777	8.18E-09	FNBP1L	1:94399522	1.39E-08
MRPL2	6:43214786	8.38E-09	FNBP1L	1:94514104	1.39E-08
TAP2	6:32782112	8.64E-09	CTSW	11:65407886	1.39E-08
DPPA4	3:109065503	8.77E-09	ANAPC16	10:73972957	1.41E-08

MRPL2	6:42072698	8.92E-09	HSPA1B	6:32041661	1.46E-08
TAC3	12:58011859	9.04E-09	DPYS	8:105800979	1.46E-08
DPPA4	3:109068774	9.24E-09	DPYS	8:106155586	1.46E-08
MXRA7	17:74730843	9.45E-09	RPS23	5:81848038	1.50E-08
MXRA7	17:74734323	9.45E-09	RPS23	5:82026884	1.50E-08
ELOVL6	4:111543944	9.69E-09	RPS23	5:82026923	1.50E-08
PMM1	22:42305583	1.02E-08	RPS23	5:82026956	1.50E-08
PMM1	22:42318560	1.02E-08	RPS23	5:82027004	1.50E-08
LGALS8	1:236852150	1.03E-08	PCGF5	10:93976268	1.54E-08
PPP1R12B	1:203198939	1.05E-08	RHCE	1:24945855	1.59E-08
DYRK3	1:207197770	1.07E-08	RSRC1	3:157882744	1.60E-08
ZNF763	19:12807066	1.07E-08	DHRS4	14:23743265	1.61E-08
DPH6	15:36505553	1.08E-08	MRPL2	6:43029288	1.65E-08
FES	15:91427692	1.09E-08	MRPL2	6:43705287	1.66E-08
FMN1	15:33384440	1.10E-08	RNF14	5:140353867	1.72E-08
RABEP1	17:5271763	1.10E-08	MAF1	8:145169485	1.83E-08
LGALS4	19:38876263	1.86E-08	SOS1	2:39948037	3.58E-08
LGALS4	19:39797885	1.86E-08	LRRC2	3:46037405	3.64E-08
ZNF763	19:12668495	1.87E-08	RBP7	1:10582127	3.67E-08
AGPAT1	6:32136771	1.95E-08	RNF135	17:29710070	3.71E-08
FOLR3	11:72140415	1.96E-08	CISD2	4:102796668	3.84E-08
ZDHHC2	8:16132332	1.97E-08	CXCL1	4:74431049	3.97E-08
IP6K2	3:49455909	2.04E-08	HOXB2	17:46484307	3.99E-08
ACOX1	17:73511629	2.04E-08	RAB3GAP1	2:135676246	4.12E-08
RBP7	1:10848564	2.05E-08	KRT79	12:53681883	4.22E-08
PNLDC1	6:160174540	2.14E-08	LRRC6	8:133953740	4.23E-08
DHRS4L2	14:24034818	2.14E-08	LGALS4	19:38713687	4.25E-08
LRRC6	8:133621091	2.18E-08	GTF2H3	12:124069213	4.27E-08
PADI2	1:16403208	2.20E-08	GTF2H3	12:124086620	4.27E-08
PBX2	6:31976833	2.20E-08	PHACTR4	1:28864435	4.31E-08
MXI1	10:111819970	2.22E-08	PCGF5	10:92035440	4.32E-08
CISD1	10:60082379	2.28E-08	DAD1	14:23475868	4.34E-08
ATP6V0D1	16:67964361	2.34E-08	DAD1	14:23479347	4.34E-08
SIN3B	19:17844227	2.34E-08	DHRS4	14:24179039	4.39E-08
PGM2L1	11:73936474	2.39E-08	DHRS4	14:24179040	4.39E-08
PGM2L1	11:75005087	2.39E-08	DYRK3	1:207200372	4.47E-08
RPS23	5:81493931	2.47E-08	ZNF763	19:12813683	4.51E-08
HOXB2	17:46626924	2.48E-08	MRPS7	17:73520485	4.61E-08
DNAAF3	19:55727510	2.48E-08	SOS1	2:39924139	5.24E-08
FMO1	1:170312936	2.54E-08	SOS1	2:39938065	5.24E-08
PLA2G4C	19:47743147	2.54E-08	VANGL1	1:116185512	5.32E-08
DPH6	15:35958756	2.55E-08	FCGR2B	1:161500895	5.63E-08

DPH6	15:35958803	2.55E-08	RPL7L1	6:43225363	5.64E-08
IP6K2	3:49063555	2.60E-08	PMM1	22:41996481	5.66E-08
IP6K2	3:49318208	2.60E-08	CAPN12	19:38880107	5.68E-08
IP6K2	3:49456758	2.60E-08	TMCC3	12:95299263	5.74E-08
RPL3L	16:2940941	2.60E-08	CHST15	10:126632419	5.86E-08
MRPL2	6:43173907	2.64E-08	RNF14	5:141308037	5.88E-08
MRPL2	6:43173908	2.64E-08	POLR1A	2:85965280	5.94E-08
PBX2	6:32052255	2.72E-08	S100P	4:5908481	5.95E-08
TBCEL	11:120450321	2.79E-08	S100P	4:7435996	5.95E-08
SNX27	1:152486860	3.01E-08	MXRA7	17:74730188	5.95E-08
DNAAF3	19:55652584	3.07E-08	MXRA7	17:74738016	5.95E-08
HAUS4	14:23402422	3.10E-08	FOXO3	6:108516159	5.96E-08
AIG1	6:142618616	3.11E-08	STX12	1:28038660	6.10E-08
POLR1A	2:86364653	3.35E-08	UQCRC1	3:48501929	6.21E-08
SOS1	2:38491910	3.36E-08	GSDMA	17:37822757	6.26E-08
ZNF763	19:12781358	3.37E-08	GSDMA	17:37840860	6.26E-08
KRT79	12:53671264	3.44E-08	GCAT	22:37962646	6.46E-08
PADI2	1:16955106	3.48E-08	SERPINC1	1:173703928	6.53E-08
SERPINC1	1:173779018	6.53E-08	LRRC6	8:134107312	9.10E-08
SERPINC1	1:173839435	6.53E-08	CFL1	11:65624490	9.17E-08
LUM	12:90725563	6.98E-08	ECHDC3	10:11797467	9.40E-08
DNAAF3	19:55597178	7.26E-08	AK7	14:96342629	9.41E-08
FNBP1L	1:93250370	7.37E-08	CSNK1D	17:79879743	9.57E-08
DAD1	14:22787449	7.43E-08	ADAMTSL4	1:151006406	9.64E-08
DYRK3	1:207196544	7.66E-08	MSLN	16:685313	9.64E-08
SOS1	2:38830131	7.68E-08	IPP	1:46383020	9.91E-08
DYRK3	1:206099780	7.79E-08	MMP24	20:33146457	1.02E-07
TOP1MT	8:144864231	7.92E-08	TBCEL	11:120418441	1.03E-07
TSPAN16	19:11461486	8.03E-08	IPP	1:46595615	1.05E-07
AGPAT1	6:32145399	8.06E-08	NUMA1	11:71725144	1.05E-07
DPH6	15:34985075	8.10E-08	NUMA1	11:71726122	1.05E-07
CTSK	1:150059869	8.27E-08	SUSD2	22:25334184	1.06E-07
CTSK	1:150318987	8.27E-08	MPC2	1:167744168	1.14E-07
CTSK	1:151735504	8.27E-08	SHC1	1:154938235	1.14E-07
SLPI	20:43687355	8.36E-08	SHC1	1:154946517	1.14E-07
ANAPC16	10:74014728	8.50E-08	SHC1	1:154947496	1.14E-07
ELOVL6	4:109979306	8.52E-08	RBMS1	2:161080519	1.14E-07
PNKD	2:219083353	8.55E-08			
CLYBL	13:100645581	8.72E-08			
RPS23	5:82281007	8.85E-08			
NUMA1	11:71816766	8.92E-08			
DHRS4	14:23877553	9.09E-08			

Table S3.6. Brain WGCNA co-expressed gene modules

Gene module	Module color*	# of genes
M0**	grey	2851
M1	turquoise	4481
M2	blue	1305
M3	brown	960
M4	yellow	566
M5	green	413
M6	red	390
M7	black	373
M8	pink	262
M9	magenta	260
M10	purple	223
M11	greenyellow	176
M12	tan	169
M13	salmon	168
M14	cyan	159
M15	midnightblue	125
M16	lightcyan	52
M17	grey60	38

*Module colors are the colors WGCNA uses for labeling the modules
**M0 module designates non-module genes that are not co-expressed with others

Table S3.7. Blood WGCNA co-expressed gene modules

Gene module	Module color*	# of genes
M0**	grey	7433
M1	turquoise	2065
M2	blue	1235
M3	brown	911
M4	yellow	447
M5	green	439
M6	red	358
M7	black	260
M8	pink	232
M9	magenta	188
M10	purple	171
M11	greenyellow	162
M12	tan	158
M13	salmon	154
M14	cyan	149
M15	midnightblue	146
M16	lightcyan	132
M17	grey60	126
M18	lightgreen	124
M19	lightyellow	122
M20	royalblue	121
M21	darkred	115
M22	darkgreen	103
M23	darkturquoise	100
M24	darkgrey	99
M25	orange	90
M26	darkorange	90
M27	white	83
M28	skyblue	70
M29	saddlebrown	68
M30	steelblue	56
M31	paleturquoise	40
M32	violet	39
M33	darkolivegreen	31
M34	darkmagenta	29

*Module colors are the colors WGCNA uses for labeling the modules

**M0 module designates non-module genes that are not co-expressed with others

Table S3.8. Proportion of variance explained by the module eigengenes in brain

Module Eigengene (ME)	Proportion of Variance
ME1	0.574
ME2	0.593
ME3	0.516
ME4	0.440
ME5	0.547
ME6	0.510
ME7	0.431
ME8	0.473
ME9	0.489
ME10	0.445
ME11	0.493
ME12	0.453
ME13	0.447
ME14	0.522
ME15	0.530
ME16	0.401
ME17	0.524

Table S3.9. Proportion of variance explained by the module eigengenes in blood

Module Eigengene (ME)	Proportion of Variance
ME1	0.280
ME2	0.441
ME3	0.450
ME4	0.426
ME5	0.484
ME6	0.482
ME7	0.417
ME8	0.501
ME9	0.488
ME10	0.420
ME11	0.481
ME12	0.481
ME13	0.440
ME14	0.405
ME15	0.448
ME16	0.428
ME17	0.442
ME18	0.443
ME19	0.459
ME20	0.429
ME21	0.424
ME22	0.424
ME23	0.435
ME24	0.469
ME25	0.450
ME26	0.454
ME27	0.444
ME28	0.540
ME29	0.441
ME30	0.466
ME31	0.427
ME32	0.493
ME33	0.498
ME34	0.539

Table S3.10. Shared genes in shared enriched pathways in brain and blood

BRAIN

CHR	BEGIN POS	END POS	NUM PASS VARS	NUM SING VARS	STATRHO	P-VALUE*	GENE MODULE	PATHWAY	GENE
1	160523750	162496199	426	268	0	1.00E+00	7	Apoptosis signaling pathway	HSPA6
8	21946761	23968674	494	330	0	6.03E-01	7	Apoptosis signaling pathway	TNFRSF10C
20	43645152	45637945	450	278	0	8.32E-01	7	CCKR signaling map	MMP9

BLOOD

1	160518049	162494496	616	397	0.2	3.66E-02	5	Apoptosis signaling pathway	HSPA6
8	21946761	23968794	817	520	0.1	8.77E-05	5	Apoptosis signaling pathway	TNFRSF10C
20	43644056	45644897	761	468	0	7.73E-04	5	CCKR signaling map	MMP9

BEGIN POS : Beginning position of range for rare variants within 1 Mb of gene to be tested

END POS : End position of range for rare variants within 1 Mb of gene to be tested

NUM PASS VARS : Number of variants passing all thresholds for EPACTS software

NUM SING VARS : Number of singletons among variants in NUM PASS VARS

P-VALUE : P-value of burden tests

STATRHO: represents the RHO value from SKAT-O test, rho = 1 (burden) and rho = 0 (SKAT)

*Bonferronni significant threshold: P < 6.49 x 10-4 in brain, P < 5.0 x 10-4 in blood

Table S3.11. Genes shared between the inflammation pathway gene modules in brain and blood

Gene	Chr
CCL3	17
CCL4	17
SPOCD1	1
CXCL5	4

Figure S3.1. WGCNA gene module and eigengene network plots. A describes WGCNA networks in brain and **B** in blood. **1)** Gene hierarchical clustering dendrogram with the color row underneath indicating module membership with module colors. The branches correspond to true modules. **2)** Hierarchical clustering of module eigengenes that summarize gene modules. Branches group together eigengenes that are positively correlated. **3)** Eigengene networks shown as heatmaps labeled with red denoting high adjacency (positive correlation) and blue denoting low adjacency (negative correlation).

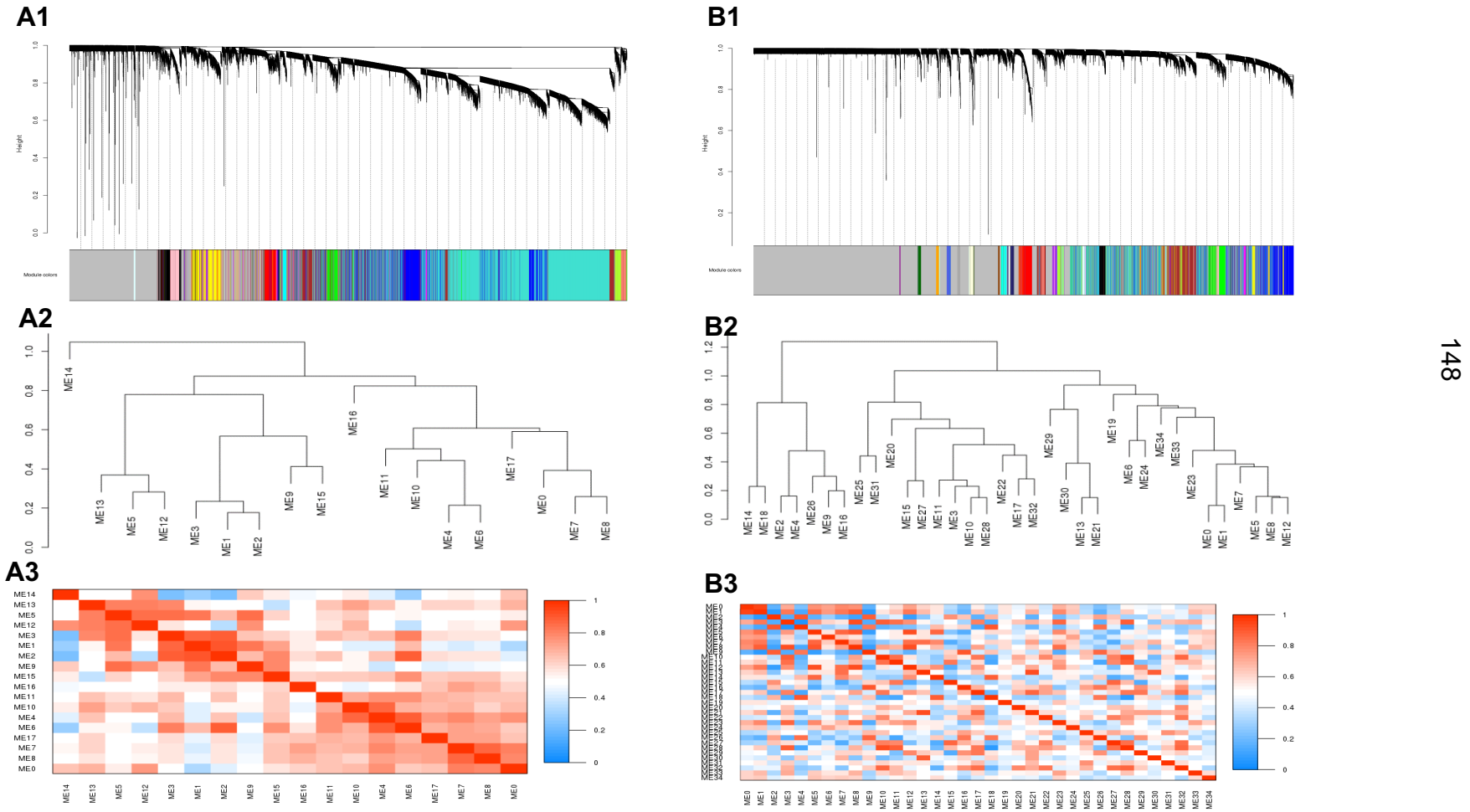
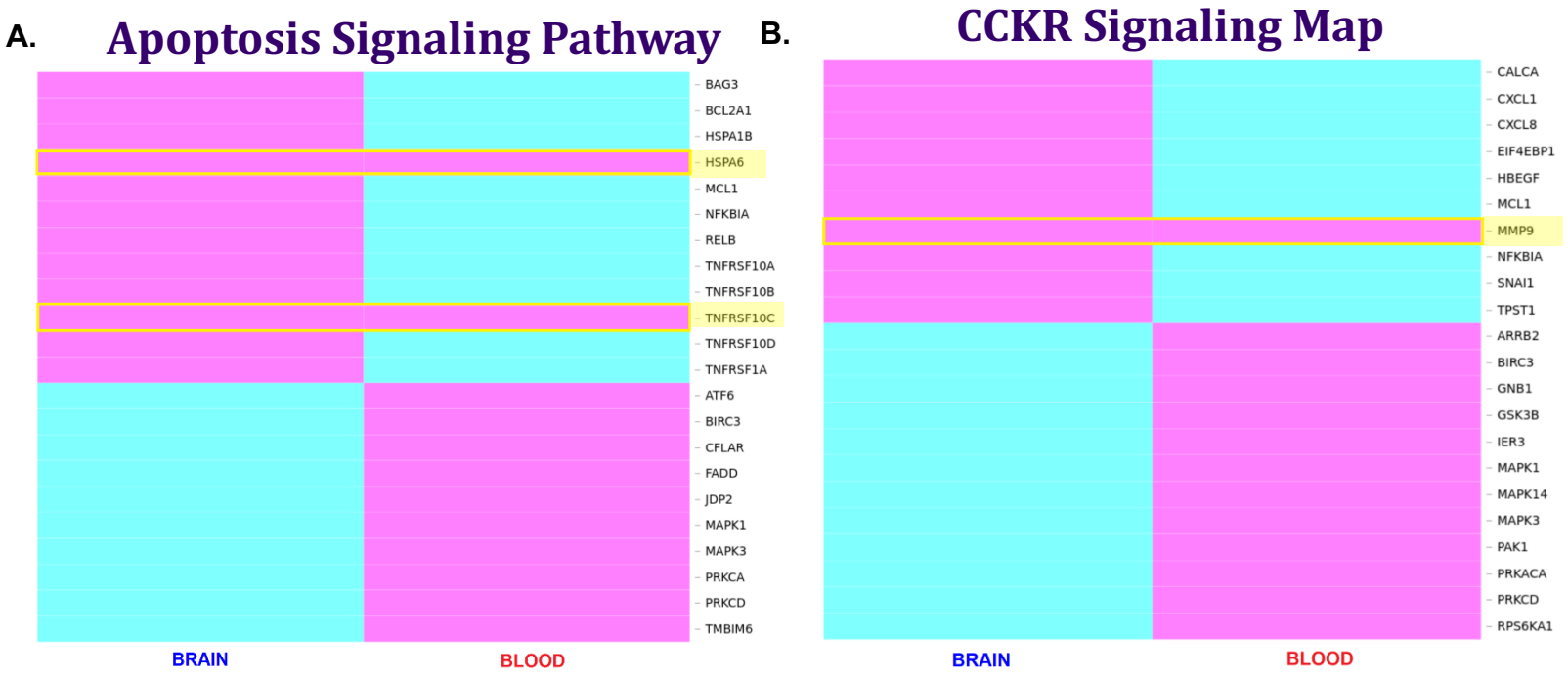


Figure S3.2. Heatmaps for gene overlap in enriched brain and blood shared pathways. The 3 shared pathways, **A)** Apoptosis Signaling Pathway, **B)** CCKR Signaling Map, and **C)** the Inflammation Signaling pathway are seen in blood and brain in heatmaps with the magenta blocks indicting gene membership in pathways. The yellow highlighted blocks and highlighted gene names illustrate the shared genes in shared pathways.



C. **Inflammation Signaling Pathway**

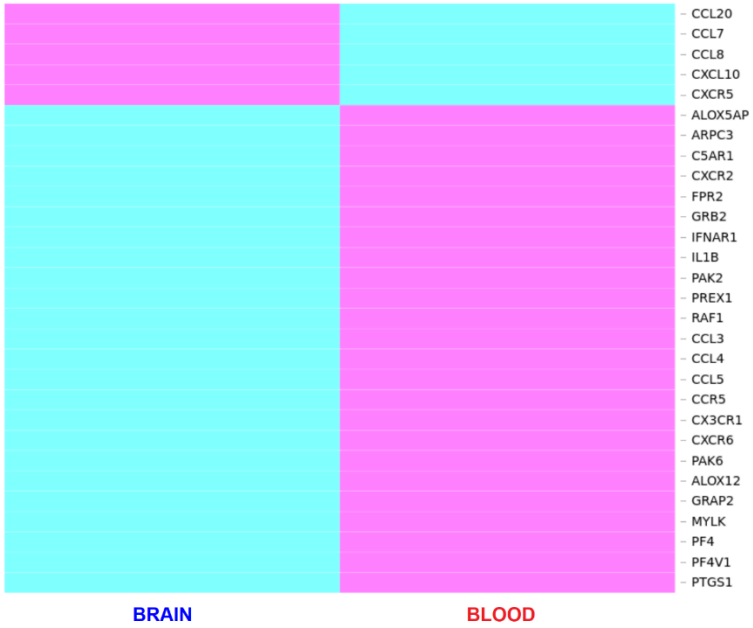


Table S4.1 Characteristics of subjects in the Framingham Heart Study (FHS) and Religious Orders Study (ROS)/ Memory and Aging Project (MAP) datasets

Dataset	Race	N	AD Cases	Controls	Males	Females	Mean Age (SD) *	% *APOE* ε4 carrier
FHS	European Ancestry	5,257	NA	NA	2,421 (46%)	2,836 (54%)	54.9 (13.3)	NA
ROSMAP	98% Caucasian, 2% African American, < 0.01 % other	475	281(59%)	194 (41%)	175 (37%)	300 (63%)	85.9 (4.8)	26%

*Age at onset for AD cases, age at exam for controls

Table S4.2 Established AD loci

Chr	Pos (hg37)	rsID	Nearest Gene	Method*	Trait†	p-value	Ref
1	1:207692049	rs6656401	CR1	G	AD	5.7×10^{-24}	[58]
1	1:227058273	N/A	PSEN2	L,C	AD	NA	[264, 265]
2	2:106642554	rs34487851	ECRG4	G	NP + NFT	2.4×10^{-8}	[233]
2	2:127892810	rs6733839	BIN1	G	AD	6.9×10^{-44}	[58]
2	2:202149628	rs146286958	CASP8	RVG	AD	8.6×10^{-5} [a]	[227]
2	2:234068476	rs35349669	INPP5D	G	AD	3.2×10^{-8}	[58]
2	2:3474085	rs35067331	*TRAPPC12*	G	NFT + CAA	5.8×10^{-8}	[233]
2	NA	NA	ADI1	GB	NFT + CAA	$<1 \times 10^{-6}$ [b]	[233]
3	3:178257562	rs9637454	KCNMB2	G	AD	7.1×10^{-8}	[217]
3	3:190951729	rs9877502	OSTN	G	CSF tau	4.9×10^{-9}	[57]
4	4:174094940	rs62341097	GALNT7	G	NP	6.0×10^{-9}	[217]
4	4:7353052	rs13110208	SORCS2	CG,F	AD	9.0×10^{-3}	[259]
4	4:95170280	rs137875858	UNC5C	L,CG,F	AD	9.5×10^{-3} [c]	[16]
5	5:139707439	rs11168036	PFDN1, HBEGF	G	AD	7.1×10^{-9}	[170]
5	5:15669858	rs75002042	FBXL7	G	AD	6.2×10^{-9}	[237]
5	5:88223420	rs190982	MEF2C	G	AD	3.2×10^{-8}	[58]
6	6: 83072923	rs547721664	TPBG	G	AD	1.8×10^{-6}	[170]
6	6:32575406	rs9271058	HLA-DRB1	G	AD	5.1×10^{-8}	[101]
6	6:32578530	rs9271192	HLA-DRB5	G	AD	2.9×10^{-12}	[58]
6	6: 41034000	rs114812713	OARD1	G	AD	2.1×10^{-13}	[6]
6	6:41129252	rs75932628 (R47H)	TREM2	G,F	AD	2.9×10^{-12}	[57, 101]
6	6:41186912	rs9381040	TREML2	G,F	AD	6.3×10^{-7}	[57, 58]
6	6:41368363	rs6922617	NCR2	G	AD	3.6×10^{-8}	[57]
6	6:47487762	rs10948363	CD2AP	G	AD	5.2×10^{-11}	[58]
7	7:100004446	rs1476679	ZCWPW1	G	AD	5.6×10^{-10}	[58]
7	7:132110923	rs277470	PLXNA4	G	AD	4.1×10^{-8}	[253]
7	7:143110762	rs11771145	EPHA1	G	AD	1.1×10^{-13}	[58]
7	7:154988675	NA	AC099552.4	RVG	AD	1.2×10^{-7} [d]	[17]
7	7:18698331	rs79524815	HDAC9	G	NFT + CAA	1.1×10^{-8}	[233]
7	7:37841534	rs2718058	NME8	G	AD	4.8×10^{-9}	[58]
7	7:51578022	rs112404845	COBL	G	AD	3.8×10^{-8}	[231]
7	7:88406552	rs73705514	ZNF804B	G	LMdT	2.9×10^{-9}	[249]
7	7:91709085	rs144662445	AKAP9	RVG,F	AD	2.2×10^{-3} [c]	[15]
7	NA	NA	PILRA	GB	AD	2.3×10^{-6} [e]	[206, 207]
8	8:145138063	rs138412600	GPAA1	G	AD	7.8×10^{-8} [f]	[266]
8	8:17496561	rs4921790	PDGFRL / MTUS1	G	HPV	4.6×10^{-9}	[249]
8	8:27195121	rs28834970	PTK2B	G	AD	7.4×10^{-14}	[58]
8	8:27467686	rs9331896	CLU	G	AD	2.8×10^{-25}	[58]
8	8:95958637	rs1713669	*TP53INP1*	GB	AD	1.4×10^{-6} [g]	[243]

9	9: 3929424	rs514716	GLIS3	G	CSF tau	3.2×10^{-9}	[57]
10	10:107013252	rs7920533	SORCS3	CG,F	AD	1.0×10^{-2}	[259]
10	10:108562008	rs12248379	SORCS1	CG,F	AD	3.0×10^{-3}	[259]
10	10:115489177	rs1116437863	CASP7	G	AD	2.4×10^{-10}	[18]
10	10:11720308	rs7920721	ECHDC3	G	AD	2.3×10^{-9}	[101]
10	10:11720308	rs7920721	USP6NL	G	AD	3.0×10^{-8}	[170]
10	10:13966445	rs2446581	FRMD4A	G	AD	1.1×10^{-10}	[238]
11	11:121435587	rs11218343	SORL1	G,CG	AD	2.7×10^{-8}	[55, 58, 101]
11	11:47380340	rs3740688	SPI1	G	AD	9.7×10^{-11}	[101]
11	11:47557871	rs10838725	CELF1	G	AD	1.1×10^{-8}	[58]
11	11:59923508	rs983392	MS4A6A	G	AD	6.1×10^{-16}	[58]
11	11:59936926	rs7933202	MS4A2	G	AD	2.2×10^{-15}	[101]
11	11:60021948	rs1582763	MS4A4A	G	AD	1.15×10^{-15}	[267]
11	11:85867875	rs10792832	PICALM	G	AD	9.3×10^{-26}	[58]
11	NA	NA	OR8G5	GB	AD	4.7×10^{-7} [h]	[266]
12	12:119390525	rs10775009	SRRM4	G	CSF Tau	1.6×10^{-9}	[249]
13	13:103663945	rs16961023	SLC10A2	G	AD	4.6×10^{-8}	[231]
14	14:105385352	rs2819438	PLD4	G	CSF Tau	6.9×10^{-9}	[249]
14	14:106236128	rs12890621	IGHG3	G	AD	9.8×10^{-7}	[17]
14	14:106680831	rs2011167	IGHV1-67	G	AD	7.9×10^{-8} [g]	[243]
14	14:53400629	rs17125944	FERMT2	G	AD	7.9×10^{-9}	[58]
14	14:73637653	rs63749824	PSEN1	L,C	AD	NA	[268]
14	14:92926952	rs10498633	SLC24A4 / RIN3	G	AD	5.5×10^{-9}	[58]
14	14:92932828	rs12881735	SLC24A4	G	AD	7.4×10^{-9}	[101]
14	NA	NA	IGHV3-7	GB	AD	9.75×10^{-16}	[266]
15	15:101646763	rs139709573	TM2D3	RVG	AD	6.6×10^{-9}	[269]
15	15:59045774	rs593742	ADAM10	G	AD	6.8×10^{-9}	[101]
15	15:64433291	rs74615166	TRIP4	G	AD	9.7×10^{-9}	[262]
16	16:19808163	rs7185636	IQCK	G	AD	2.4×10^{-8}	[101]
16	16:79355857	rs62039712	WWOX	G	AD	3.7×10^{-8}	[101]
16	16:81942028	rs72824905	PLCG2	G	AD	5.4×10^{-10}	[3]
17	17:44353222	rs2732703	MAPT	G	AD	5.8×10^{-9}	[56]
17	17:44355683	rs113986870	KANSL1	G	AD	1.3×10^{-8}	[56]
17	17:47297297	rs616338	ABI3	G	AD	4.6×10^{-10}	[3]
17	17:56409089	rs2632516	BZRAP1	G	AD	4.4×10^{-8}	[170]
17	17:61569732	rs4351	ACE	G	AD	5.3×10^{-9}	[270]
19	19:1063443	rs4147929	ABCA7	G	AD	1.1×10^{-15}	[58]
19	19:15302421	rs149307620	NOTCH3	RVG,F	AD	NA	[83]
19	19:18546678	rs2303697	ISYNA1	G	AD	4.6×10^{-7} [f]	[266]
19	19:40877595	rs145999145	PLD3	RVG,F	AD	1.4×10^{-11}	[252]
19	19:45411941	rs429358	APOE	L,CG	AD	2.1×10^{-47}	[271]
19	19:51727962	rs3865444	CD33	G	AD	1.6×10^{-9}	[37]

20	20:55018260	rs7274581	CASS4	G	AD	2.5×10^{-8}	[58]
20	NA	NA	SLC24A3	GB	AD	2.7×10^{-12}	[266]
21	21:28915457	rs1487586185	APP	CG	AD	NA	[224]
21	21:28156856	rs2830500	ADAMTS1	G	AD	2.6×10^{-8}	[101]
21	21:43678066	rs142544282	ABCG1	G	NP	8.0×10^{-9}	[217]

***Method:** G = GWAS; RVG = Rare Variant GWAS; C =Cloning; CG = Candidate Gene; L = Linkage; F =Functional evidence; GB = Gene-based;

***Trait:** AD = Alzheimer disease, HPV = hippocampal volume, LMdT = logical memory – delayed recall, NFT = neurofibrillary tangle, NP = neuritic plaque

Study-wide significance level: [a] $8x10^{-4}$, [b] $2.7x10^{-6}$; [c] 0.05, [d] $3.1x10^{-7}$, [e] 0.0014, [f] $2.8x10^{-7}$, [g] $2.5x10^{-6}$, [h] $6.4x10^{-7}$,

Table S4.3 Number of significant cis eQTLs and ct-eQTLs in blood and brain tissue data as seen in genome-wide analysis

	Blood eQTL	**Blood ct-eQTL**	**Brain eQTL**	**Brain ct-eQTL**
Total # of tests	8,662,143	86,621,430	304,732,096	1,523,660,480
# significant eQTL pairs	847,429	30,405	173,857	51,098
p-value threshold [1]	5.77E-09	5.77E-10	1.64E-10	3.28E-11
# unique eGenes in eQTL pairs	6033	502	1,301	799
# unique eSNPs in eQTL pairs	727,384	22,226	144,258	26,757

[1]Bonferroni corrected

Table S4.4 Significant eGenes shared between blood and brain tissues.
386 distinct eGenes among 24,028 significant gene-SNP pairs shared between permutations of blood and brain eQTLs/ct-eQTLs

ACACA	C9orf72	CYSTM1	FOPNL	HPR	LPIN2	NARS2	RAD51C	SPECC1	UFSP2
ACCS	C9orf89	DCAKD	FOXRED1	HSD17B12	LRPAP1	NDUFAF1	RARRES1	SPEF2	UHRF1BP1
ACOT4	CAB39L	DCBLD2	FOXRED2	HSD17B13	LRRC2	NIPSNAP3A	RBL2	SPPL3	ULK4
ADAL	CAT	DDT	FRA10AC1	HTATIP2	LRRC27	NMRK1	RBPMS2	SPSB2	VN1R1
ADAT1	CBLN3	DHRS1	FUT2	HYAL3	LRRC61	NMUR1	RCBTB1	SSR1	WARS2
ADHFE1	CBR3	DNAJC15	FXN	IFI27	LRRCC1	NOP10	RFWD3	STYXL1	WBP2NL
ADORA2B	CCBL2	DPM2	GAA	IFI27L1	LRRIQ3	NPHP3	RMI2	SULT1A1	WBSCR27
AGA	CCDC13	DSC1	GALC	IFI27L2	LSG1	NRBF2	RNF166	SUMF1	WDR27
AHI1	CCDC170	DSCC1	GATC	IFIT5	LXN	NSUN2	RNPEP	SUPT3H	WDR52
AIFM2	CCDC23	DUSP14	GBP3	IFT46	MAEL	NSUN6	RPA2	SUPT4H1	WDYHV1
AK5	CCDC25	**ECHDC3** *	GCLM	IGFLR1	MAN2B2	NUP107	RPL36AL	SUSD1	WFDC3
ALDH8A1	CCDC82	EFCAB13	GGNBP2	IL10RB	MANBA	NUP210L	RPP21	TAF1C	WIPI1
AMACR	CD274	EFCAB2	GIMAP5	IL18R1	MCFD2	NUPL2	RRP1B	TAS2R4	**WWOX** *
ANAPC4	CD6	EFHB	GLIPR1L2	IL1RL1	MCM8	NUSAP1	RTN4	TBC1D9B	XRCC6BP1
ARHGEF3	CDA	EIF2A	GNL3	IL32	MCOLN2	OSCP1	RWDD2B	TC2N	XRRA1
ARHGEF35	CDC25B	ENDOG	GNLY	INPP1	MCPH1	PAAF1	RWDD3	TCEA3	YEATS4
ARL17A	CDC7	EPG5	GOSR1	IPMK	MEI1	PADI4	SAAL1	TCF19	YWHAB
AS3MT	CDK10	ERAP2	GP6	IQCB1	METTL18	PAX8	SCIMP	TDRD6	ZADH2
ASRGL1	CENPP	ERCC3	GPNMB	IQCG	METTL21B	PCNXL4	SCLY	TECPR1	ZFP82
ATF6	CENPV	ERMAP	GSR	IQGAP1	MGMT	PDCD1LG2	SELL	TEN1	ZMAT3
ATP5S	CEP19	ETFDH	GSTA1	ISCU	MGRN1	PDHB	SETD4	TESK2	ZMYND12
ATP6V1E2	CERS5	ETS2	GSTM3	ITGB2	MICB	PDLIM5	SHISA4	TFB1M	ZNF155
ATXN7L3B	CHAF1A	ETV7	GSTT1	ITIH4	MLH3	PDZK1IP1	SLC18A1	THNSL2	ZNF354C
B3GNTL1	CHCHD2	EXOSC6	GTSF1	KANSL2	MON1B	PEX6	SLC22A18	TIMM10	ZNF467
BATF3	CHKB-CPT1B	EYA3	GUF1	KIAA1430	MPHOSPH6	PIGN	SLC25A1	TMEM156	ZNF471
BLMH	CHRNA5	F2RL1	HAUS4	KLC3	MPPE1	PISD	SLC25A24	TMEM245	ZNF501
BLOC1S2	CHRNE	F5	HBS1L	KLHDC4	MRPL10	PLEKHH2	SLC25A51	TMOD3	ZNF502
BLVRA	CKS2	FADD	HEATR3	KLKB1	MRPL18	POLR1B	SLC26A8	TNFRSF10C	ZNF514
BNIP1	CLTCL1	FAH	HIATL1	KRR1	MRPL19	POLR1E	SLC35A1	TNFRSF13C	ZNF577
BTN2A2	CNDP2	FAM118A	HIBCH	L3HYPDH	MRPL21	POLR2J	SLFN5	TNNI3	ZNF593
BTNL3	CNGA1	FAM3B	HIST1H3E	L3MBTL3	MRPL53	POP5	SMARCB1	TOMM7	ZNF670,
C12orf60	CPVL	FANK1	HIST1H4C	LACTB	MRPS10	PPA2	SMC1B	TOP3B	ZNF839
C14orf166	**CR1** *	FAS	HIST1H4F	LDHC	MSH3	PPFIA1	SMC2	TRAPPC4	ZNRD1
C15orf57	CRIPT	FBN2	HLA-DOB	LIG3	MSH4	PPIL3	SNX19	TREML4	ZP3
C1QTNF6	CRLF3	FEM1A	HLA-DQA2	LILRB1	MTHFS	PPM1N	SNX32	TRIM35	ZXDC
C21orf128		FEZ2	HLA-DQB2	LIMD1	MTRR	PPP2R1B			

C22orf34 C7orf13 C8orf31	CRYZ CSGALNACT1 CSGALNACT2 CTNS	FIGNL1 FLT3 FLYWCH1	**HLA-DRB1*** **HLA-DRB5*** HP	LNPEP LONP1 LPIN1	MUL1 MZT<2A NAPRT1	PTGR1 QRSL1 RABEP1	SOHLH2 SPAG7 SPATA5L1 SPATA7	TRIM58 TTC21B TVP23C UBALD2	

Table S4.5: Pathway analysis of 386 distinct eGenes shared in blood and brain

PANTHER Pathway	# Genes Annotated to Pathway	# Genes in Network	Expected P value *	Fold Enrich-ment	Direction	Unadjusted P value
Apoptosis signaling pathway	115	6	2.12	2.83	+	2.28E-02
Wnt signaling pathway	317	1	5.84	0.17	-	3.36E-02
General transcription by RNA polymerase I	17	2	0.31	6.39	+	4.55E-02

PANTHER = Protein ANalysis THrough Evolutionary Relationships) Classification System; *Expected probability of observing at least x number of genes out of the total n genes in the PANTHER list annotated to a particular pathway, given the proportion of genes in the reference Homo Sapiens whole genome that are annotated to that pathway

Table S4.6: Distinct eGenes shared by all blood and brain eQTLs and ct-eQTLs
16 distinct eGenes among 657 significant gene-SNP pairs shared by all blood and brain eQTLs and ct-eQTLs

AS3MT	EFCAB2	HLA-DOB	MRPL21
ATXN7L3B	ERAP2	**HLA-DRB1 ***	NMRK1
BTNL3	FAM118A	**HLA-DRB5 ***	TIMM10
DNAJC15	GSTT1	HP	XRRA1

* Gene previously associated with AD risk by GWAS

Table S4.7 AD loci genes in significant results

eGene	Tissue	Cell-type	Lead eSNP	Position*	MAF	Beta	Std Error	P-value	# of total significant eSNPs in gene/cell-type	Distance (bp) between eSNP and eGene [Nearest gene]
CR1	Blood	NA	rs7533408	1:207673631	0.25	0.059	0.006	3.60E-22	169	0
PSEN2	Blood	NA	rs1289395	1:227042462	0.37	-0.020	0.003	2.77E-12	49	15423
TRAPPC12	Blood	NA	rs71281795	2:3487331	NAV	-0.040	0.005	1.10E-15	104	0
ADI1	Blood	NA	rs57139325	2:3519283	0.18	0.127	0.008	8.45E-58	123	0
BIN1	Blood	NA	rs1060743	2:127826533	0.29	-0.060	0.003	2.48E-99	355	0
CASP8	Blood	NA	rs7560328	2:202164837	0.39	0.028	0.003	1.06E-18	54	12403 [In FLACC1]
INPP5D	Blood	NA	rs7581787	2:234077240	0.44	0.026	0.003	5.79E-16	158	0
GALNT7	Blood	NA	rs1006003	4:174092791	0.50	-0.034	0.003	5.07E-24	102	0
HLA-DRB5	Blood	NA	rs9269008	6:32436217	0.17	-2.580	0.057	<1.0E-314	72	48903 [HLA-DRB9 (5060)]
HLA-DRB1	Blood	NA	rs9270815	6:32569859	0.14	-5.430	0.044	<1.0E-314	630	12234
TREML2	Blood	NA	rs6933231	6:41163700	0.37	-0.034	0.004	5.39E-16	34	0
CD2AP	Blood	NA	rs4711880	6:47480676	0.25	-0.146	0.007	1.36E-104	331	0
NME8	Blood	NA	rs71527594	7:37875986	NAV	0.117	0.006	1.84E-81	304	12213
EPHA1	Blood	NA	rs3935067	7:143104331	0.33	0.039	0.003	5.29E-34	50	0
MTUS1	Blood	NA	rs117496663	8:17660151	0.02	-0.290	0.009	9.94E-218	325	1725
PTK2B	Blood	NA	rs28834970	8:27195121	0.34	-0.068	0.003	1.48E-100	339	0
CLU	Blood	NA	rs9331950	8:27454682	0.22	-0.059	0.010	3.15E-09	1	0
TP53INP1	Blood	NA	rs6987752	8:95966531	0.46	-0.035	0.004	8.43E-19	145	4892 [In NDUFAF6]
USP6NL	Blood	NA	rs968455032	10:11697424	NAV	0.066	0.009	5.12E-12	61	43671
ECHDC3	Blood	NA	rs11257290	10:11780324	0.28	0.041	0.005	2.91E-19	115	4041
FRMD4A	Blood	NA	rs1409327	10:13747195	0.41	0.023	0.003	3.40E-13	11	0
SORCS3	Blood	NA	rs1404786	10:106740404	0.05	-0.067	0.007	9.04E-20	223	0
CASP7	Blood	NA	NAV	10:115439640	NAV	-0.072	0.005	8.00E-52	201	0
MS4A2	Blood	NA	rs514266	11:59877697	0.46	-0.094	0.011	5.74E-18	63	14253
MS4A6A	Blood	NA	rs667897	11:59936979	0.47	0.087	0.006	1.91E-50	145	2102

MS4A4A	Blood	NA	rs2162254	11:60039917	0.40	0.109	0.016	2.62E-12	44	8097
PICALM	Blood	NA	rs7131120	11:85690012	0.22	-0.051	0.006	7.47E-15	41	0
PSEN1	Blood	NA	rs214260	14:73662629	0.16	0.076	0.004	1.28E-80	251	0
RIN3	Blood	NA	rs17783630	14:92955385	0.46	0.037	0.002	3.73E-50	67	24733 [In SLC24A4]
SLC24A4	Blood	NA	rs17783630	14:92955385	0.46	0.063	0.004	2.82E-45	73	0
TM2D3	Blood	NA	rs12907459	15:102225621	0.44	0.062	0.004	3.60E-63	134	33027 [In TARS3]
WWOX	Blood	NA	rs7202722	16:78282458	0.40	0.023	0.003	2.60E-14	45	0
PLCG2	Blood	NA	rs7187863	16:81964977	0.23	0.038	0.004	1.61E-19	68	0
KANSL1	Blood	NA	rs2732716	17:44323046	0.36	-0.087	0.006	4.75E-43	1767	20313 [MAPK8IP1P1(636)]
ABCA7	Blood	NA	rs12462842	19:1100976	0.49	0.017	0.002	2.88E-13	29	35405 [GPX4(2960)]
PLD3	Blood	NA	rs201739636	19:40851678	0.02	0.019	0.003	3.58E-12	8	2685 [In C19orf47]
CD33	Blood	NA	rs200656	19:51724326	0.22	0.038	0.005	1.42E-15	5	3994
SLC24A3	Blood	NA	rs3827978	20:19281291	0.36	0.043	0.004	1.42E-34	182	0
CASS4	Blood	NA	rs6014740	20:55045843	0.35	0.040	0.005	1.50E-17	78	11447 [In RTF2]
APP	Blood	NA	rs8131895	21:27503527	0.36	-0.044	0.004	1.52E-22	126	0
ADAMTS1	Blood	NA	rs373460567	21:28215827	0.22	0.072	0.006	9.18E-38	27	0
ABCG1	Blood	NA	rs9976024	21:43641657	0.14	-0.030	0.004	2.01E-11	43	0
HLA-DRB5	Blood	Interferon response(+)/ Anti-bacterial(-)	rs9269047	6:32438783	0.12	-7.120	0.335	3.04E-100	9 [all (-)]	46337 [HLA-DRB9 (2494)]
HLA-DRB5	Blood	Monocytes/ Macrophages	rs9269047	6:32438783	0.12	-11.60	1.030	2.02E-29	1	46337 [HLA-DRB9 (2494)]
HLA-DRB5	Blood	NK cells / CD8+ T-Cells	rs9269047	6:32438783	0.12	-7.660	0.994	1.30E-14	1	46337 [HLA-DRB9 (2494)]
HLA-DRB1	Blood	NK cells / CD8+ T-Cells	rs9270928	6:32572461	0.15	-4.070	0.377	3.60E-27	287	14836
HLA-DRB1	Blood	Eosinophils	rs9270994	6:32574250	0.14	-2.700	0.415	7.72E-11	42	16625
HLA-DRB1	Blood	Interferon response(+)/ Anti-bacterial(-)	rs9271147	6:32577385	0.14	-5.510	0.250	1.19E-107	346 [260(-)/86(+)]	19760 [HLA-DQA1(18571)]
HLA-DRB1	Blood	Monocytes/ Macrophages	rs9271148	6:32577442	0.13	-6.110	0.709	6.83E-18	222	19817 [HLA-DQA1(18514)]
CR1	Brain	NA	rs12037841	1:207684192	0.18	-0.096	0.007	9.25E-44	64	0
HLA-DRB5	Brain	NA	rs3117116	6:32367017	0.12	-2.780	0.070	<1.0E-314	10537	118103 [In TSBP1-AS1]

HLA-DRB1	Brain	NA	rs73399473	6:32538959	0.26	-2.050	0.058	8.78E-272	10792	7587
ECHDC3	Brain	NA	rs866770710	10:11784320	0.0004	-0.252	0.018	4.61E-44	45	45
WWOX	Brain	NA	rs12933282	16:78124987	0.45	-0.133	0.017	1.13E-15	75	8323
MAPT	Brain	NA	rs2950011	17:43666385	0.23	-0.134	0.020	1.65E-11	186	305363 [DND1P1(2090)]
OARD1	Brain	Endothelial Cells	rs17825664	6:405873	0.08	-0.812	0.120	1.32E-11	6	40595493 [In IRF4]
HLA-DRB5	Brain	Microglia	rs67987819	6:32497655	0.14	-1.900	0.137	9.82E-44	754	0
HLA-DRB5	Brain	Endothelial Cells	rs67987819	6:32497655	0.14	-2.410	0.220	6.32E-28	343	0
HLA-DRB1	Brain	Microglia	rs72847627	6:32538512	0.28	-2.130	0.125	4.15E-65	2305	8034
HLA-DRB1	Brain	Neurons	rs115480576	6:32538570	0.26	-2.210	0.153	2.72E-47	3263	7976
HLA-DRB1	Brain	Endothelial Cells	rs9269492	6:32542924	0.23	-2.250	0.243	2.06E-20	351	3622
HLA-DRB5	Brain	Neurons	rs9270035	6:32553446	0.14	-2.520	0.137	1.46E-75	2540	55382 [In HLA-DRB1]
ECHDC3	Brain	Neurons	rs866770710	10:11784320	0.0004	0.328	0.045	3.13E-13	2	45

NA = not applicable; NAV = not available; * Chromosome and map position according to GRCh37 assembly; MAF = minor allele frequency; Cell-type specific result rows shaded in gray;

Table S4.8 AD association peaks in significant results

eGene	Tissue	Cell-type	eSNP+GWAS SNP	Position *	MAF	Beta	Std Error	P-value	Nearest AD Gene
TRAPPC12	Blood	NA	rs35067331	2:3474085	0.29	0.021	0.003	1.23E-10	TRAPPC12
BIN1	Blood	NA	rs6733839	2:127892810	0.38	-0.040	0.003	7.42E-38	BIN1
INPP5D	Blood	NA	rs35349669	2:234068476	0.46	0.025	0.003	7.88E-14	INPP5D
HLA-DRB1	Blood	NA	rs9271058	6:32575406	0.27	-2.950	0.028	<1.0E-314	HLA-DRB1
CD2AP	Blood	NA	rs10948363	6:47487762	0.25	-0.146	0.007	2.32E-104	CD2AP
NME8	Blood	NA	rs2718058	7:37841534	0.37	0.078	0.006	1.64E-43	NME8
PILRB	Blood	NA	rs1476679	7:100004446	0.30	0.109	0.008	1.15E-46	ZCWPW1
TAS2R60	Blood	NA	rs11771145	7:143110762	0.36	-0.376	0.010	3.29E-274	EPHA1
EPHA1	Blood	NA	rs11771145	7:143110762	0.36	-0.029	0.004	1.90E-14	EPHA1
PTK2B	Blood	NA	rs28834970	8:27195121	0.34	-0.068	0.003	1.48E-100	PTK2B
TRIM35	Blood	NA	rs28834970	8:27195121	0.34	0.018	0.003	3.72E-10	PTK2B
TP53INP1	Blood	NA	rs1713669	8:95958637	0.36	-0.029	0.004	1.21E-12	TP53INP1
MADD	Blood	NA	rs3740688	11:47380340	0.46	0.017	0.003	4.06E-09	SPI1
MYBPC3	Blood	NA	rs3740688	11:47380340	0.46	-0.020	0.002	4.47E-18	SPI1
MS4A6A	Blood	NA	rs983392	11:59923508	0.41	0.072	0.006	1.22E-33	MS4A6A
MS4A6A	Blood	NA	rs7933202	11:59936926	0.39	0.079	0.006	4.40E-40	MS4A2
MS4A4A	Blood	NA	rs1582763	11:60021948	0.36	0.107	0.016	1.44E-11	MS4A4A
SLC24A4	Blood	NA	rs10498633	14:92926952	0.21	-0.044	0.006	2.60E-14	SLC24A4/RIN3
SLC24A4	Blood	NA	rs12881735	14:92932828	0.22	-0.044	0.006	9.19E-15	SLC24A4
FAM63B	Blood	NA	rs593742	15:59045774	0.32	-0.086	0.006	8.18E-45	ADAM10
ARL17A	Blood	NA	rs2732703	17:44353222	0.21	0.147	0.023	5.95E-11	MAPT
ARL17A	Blood	NA	rs113986870	17:44355683	0.09	0.166	0.025	2.30E-11	KANSL1
SUPT4H1	Blood	NA	rs2632516	17:56409089	0.47	-0.036	0.004	5.14E-22	BZRAP1
CNN2	Blood	NA	rs4147929	19:1063443	0.19	-0.121	0.011	7.17E-28	ABCA7
HMHA1	Blood	NA	rs4147929	19:1063443	0.19	-0.030	0.004	1.62E-12	ABCA7
LRRC25	Blood	NA	rs2303697	19:18546678	0.35	0.043	0.003	8.37E-57	ISYNA1
HLA-DRB1	Blood	Interferon response(+)/ Anti-bacterial(-)	rs9271058	6:32575406	0.27	3.010	0.159	6.36E-80	HLA-DRB1

HLA-DRB1	Blood	NK cells / CD8+ T-Cells	rs9271058	6:32575406	0.27	-4.090	0.464	1.20E-18	HLA-DRB1
HLA-DRB1	Blood	Monocytes / Macrophages	rs9271058	6:32575406	0.27	-3.540	0.497	1.06E-12	HLA-DRB1
CR1	Brain	NA	rs6656401	1:207692049	0.17	-0.096	0.007	1.05E-43	CR1
HLA-DRB5	Brain	NA	rs9271058	6:32575406	0.27	-1.690	0.081	2.28E-106	HLA-DRB1
HLA-DRB1	Brain	NA	rs9271058	6:32575406	0.27	-1.770	0.054	1.94E-213	HLA-DRB1
HLA-DRB5	Brain	NA	rs9271192	6:32578530	0.27	-1.680	0.080	5.27E-107	HLA-DRB5
HLA-DRB1	Brain	NA	rs9271192	6:32578530	0.27	-1.760	0.054	2.82E-213	HLA-DRB5
KNOP1	Brain	NA	rs7185636	16:19808163	0.16	0.513	0.031	2.26E-60	IQCK
LRRC37A2	Brain	NA	rs2732703	17:44353222	0.21	1.370	0.053	4.13E-150	MAPT
ARL17A	Brain	NA	rs113986870	17:44355683	0.09	1.260	0.047	4.96E-12	KANSL1
LRRC37A2	Brain	NA	rs113986870	17:44355683	0.09	-0.326	0.068	1.98E-76	KANSL1
HLA-DRB1	Brain	Neurons	rs9271058	6:32575406	0.27	-1.400	0.135	2.37E-34	HLA-DRB1
HLA-DRB5	Brain	Neurons	rs9271058	6:32575406	0.27	-1.650	0.201	1.24E-14	HLA-DRB1
HLA-DRB1	Brain	Microglia	rs9271058	6:32575406	0.27	-1.550	0.111	1.80E-36	HLA-DRB1
HLA-DRB1	Brain	Endothelial Cells	rs9271192	6:32578530	0.27	-1.400	0.245	2.96E-12	HLA-DRB5
HLA-DRB1	Brain	Neurons	rs9271192	6:32578530	0.27	-1.650	0.135	2.37E-34	HLA-DRB5
HLA-DRB5	Brain	Neurons	rs9271192	6:32578530	0.27	-1.550	0.201	1.24E-14	HLA-DRB5
HLA-DRB1	Brain	Microglia	rs9271192	6:32578530	0.27	-1.710	0.110	4.17E-37	HLA-DRB5
LRRC37A2	Brain	Microglia	rs2732703	17:44353222	0.21	1.520	0.147	7.65E-24	MAPT
LRRC37A2	Brain	Neurons	rs2732703	17:44353222	0.21	1.480	0.140	1.84E-27	MAPT
LRRC37A2	Brain	Endothelial Cells	rs2732703	17:44353222	0.21	1.750	0.233	5.88E-14	MAPT
LRRC37A2	Brain	Neurons	rs113986870	17:44355683	0.09	1.530	0.184	2.77E-14	KANSL1
LRRC37A2	Brain	Microglia	rs113986870	17:44355683	0.09	1.400	0.195	4.29E-15	KANSL1

Known AD gene

NA = not applicable; * Chromosome and map position in base pairs; MAF = minor allele frequency; Cell-type specific result rows shaded in gray;

Table S4.9 Cell-type distribution in significant ct-eQTL results

Tissue / cell-type	# significant ct-eQTLs	% significant ct-eQTLs	Reference % in tissue[1]
Blood			
Neutrophils.1	1,007	3.3	53.8±6.1 of leukocytes[2]
Neutrophils.2	50	2.4	53.8±6.1 of leukocytes
CD4+ T-cells	562	1.8	14.6±3.5 of leukocytes
NK cells / CD8+T-cells	2,611	8.6	NK cells=4.4±2.4 of leukocytes CD8+ T cells=6.8±1.3 of leukocytes
Erythrocytes	2,130	7.0	93-96% of blood cells
Monocytes/macrophages	3,234	10.6	8.4±1.3 of leukocytes
Interferon response / Anti-bacterial cells	19,331	63.6	NA
B-cells	89	0.3	5.2±2.3 of leukocytes
Eosinophils	735	2.4	3.2±1.6 of leukocytes
Unknown	646	2.1	NA
Brain			
Endothelial cells	10,597	20.7	~2:1 glia[3] to endothelial cells
Neurons	18,930	37.0	1:1 ratio of glia to neurons
Microglia	15,560	30.4	10% of glial cells
Astrocytes	1,040	2.0	19–40% of glial cells
Oligodendrocytes	4,971	9.7	45–75% of glial cells

[1] References in blood [200, 201] and in brain [202, 272]

[2] Leukocytes: 0.1-0.2% of blood cells

[3] Glial cells include microglia, oligodendrocytes and astrocytes

Table S4.10 Gene-set enrichment analysis of 128 distinct genes in significant results in monocytes/macrophages

PANTHER Pathway	# Genes Annotated to Pathway	# Genes in Network	Expected P value *	Fold Enrichment	Direction	Unadjusted P-value	FDR
Alzheimer disease-amyloid secretase pathway	67	5	0.41	12.06	+	8.18E-05	1.34E-02
Oxytocin receptor mediated signaling pathway	58	4	0.36	11.15	+	5.78E-04	4.74E-02
Thyrotropin-releasing hormone receptor signaling pathway	60	4	0.37	10.78	+	6.52E-04	3.56E-02
5HT2 type receptor mediated signaling pathway	67	4	0.41	9.65	+	9.64E-04	3.95E-02

PANTHER = Protein ANalysis THrough Evolutionary Relationships) Classification System; * Expected probability of observing at least x number of genes out of the total n genes annotated to the pathway; FDR = false discovery rate

Figure S4.1 Study design

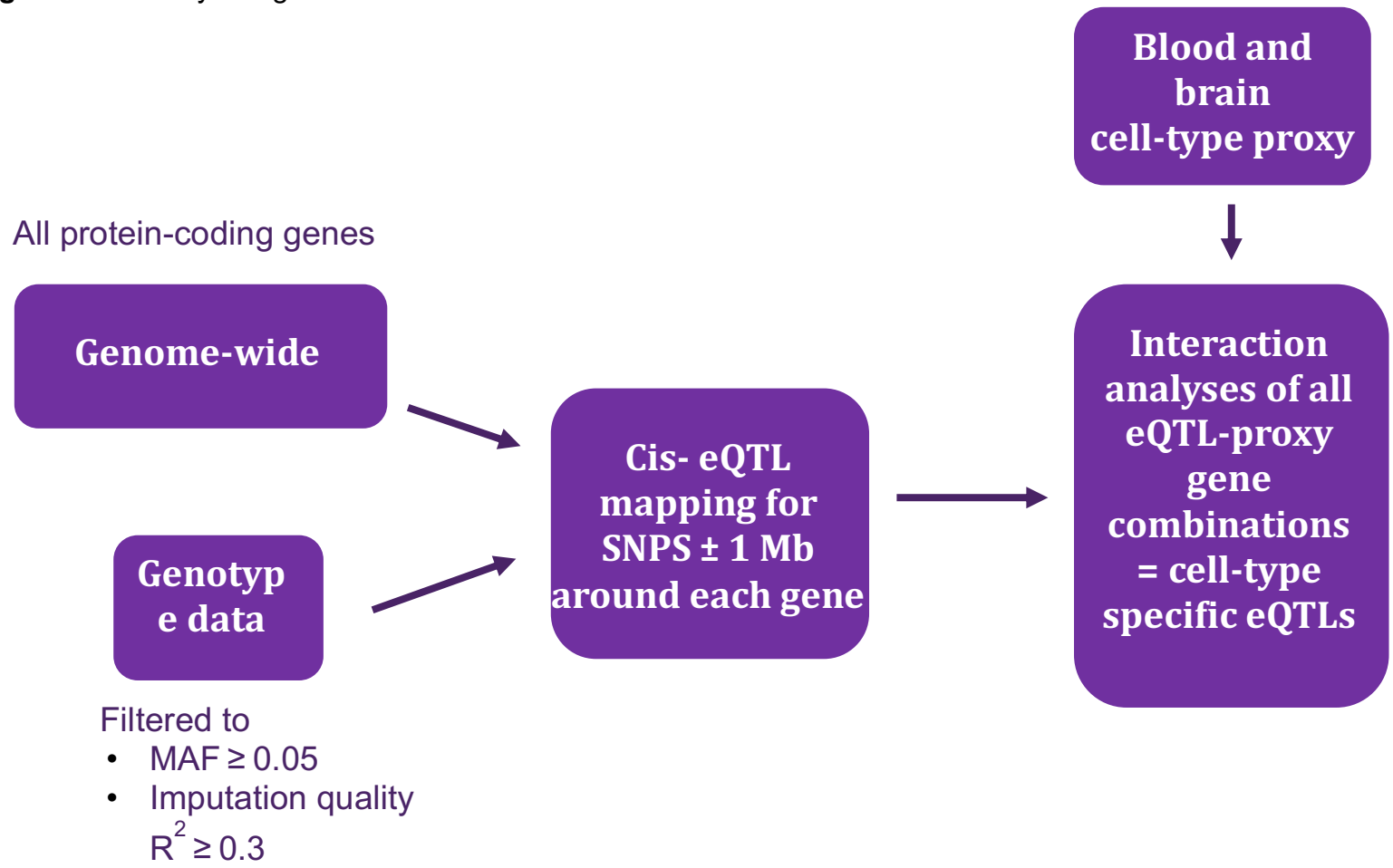
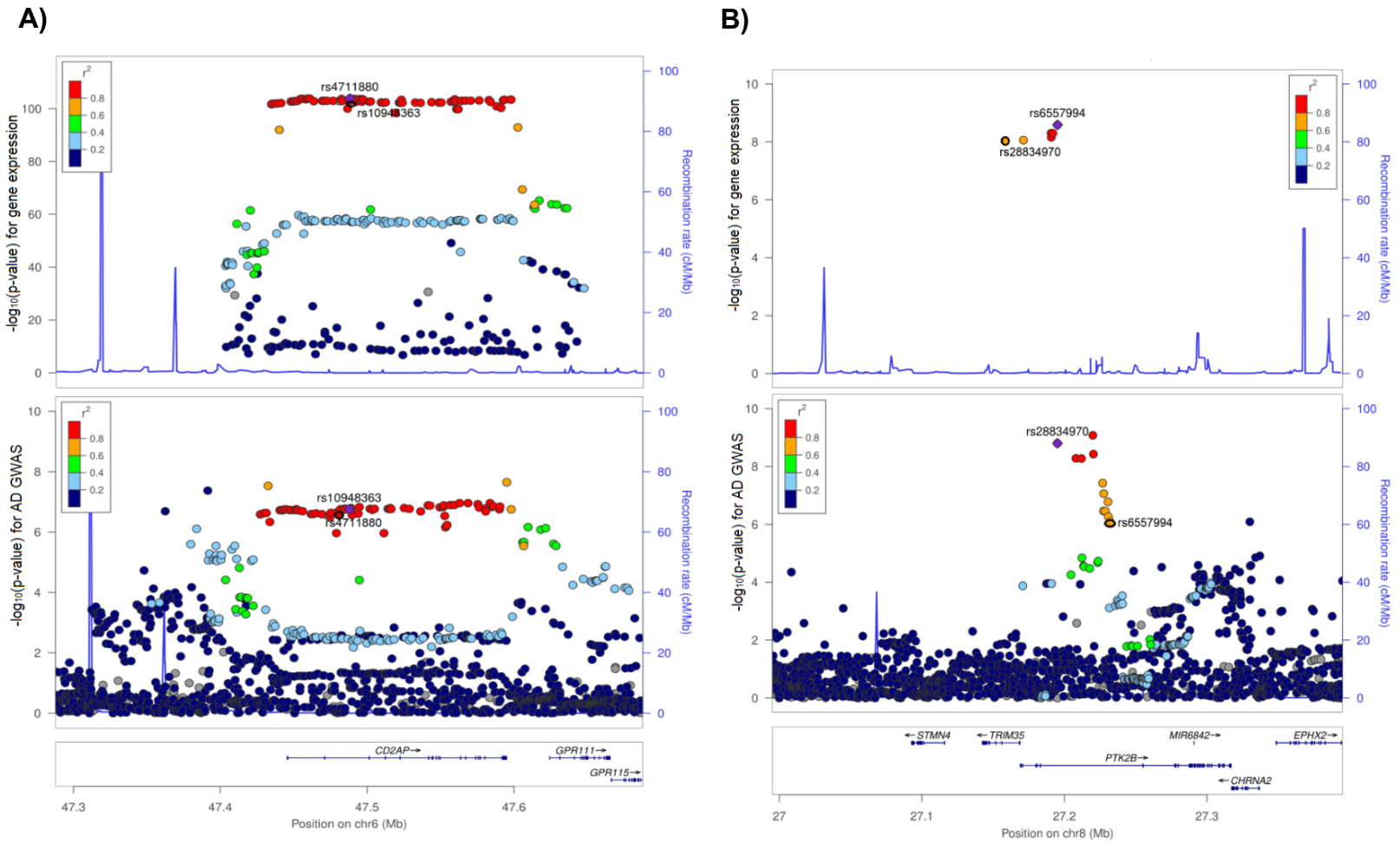
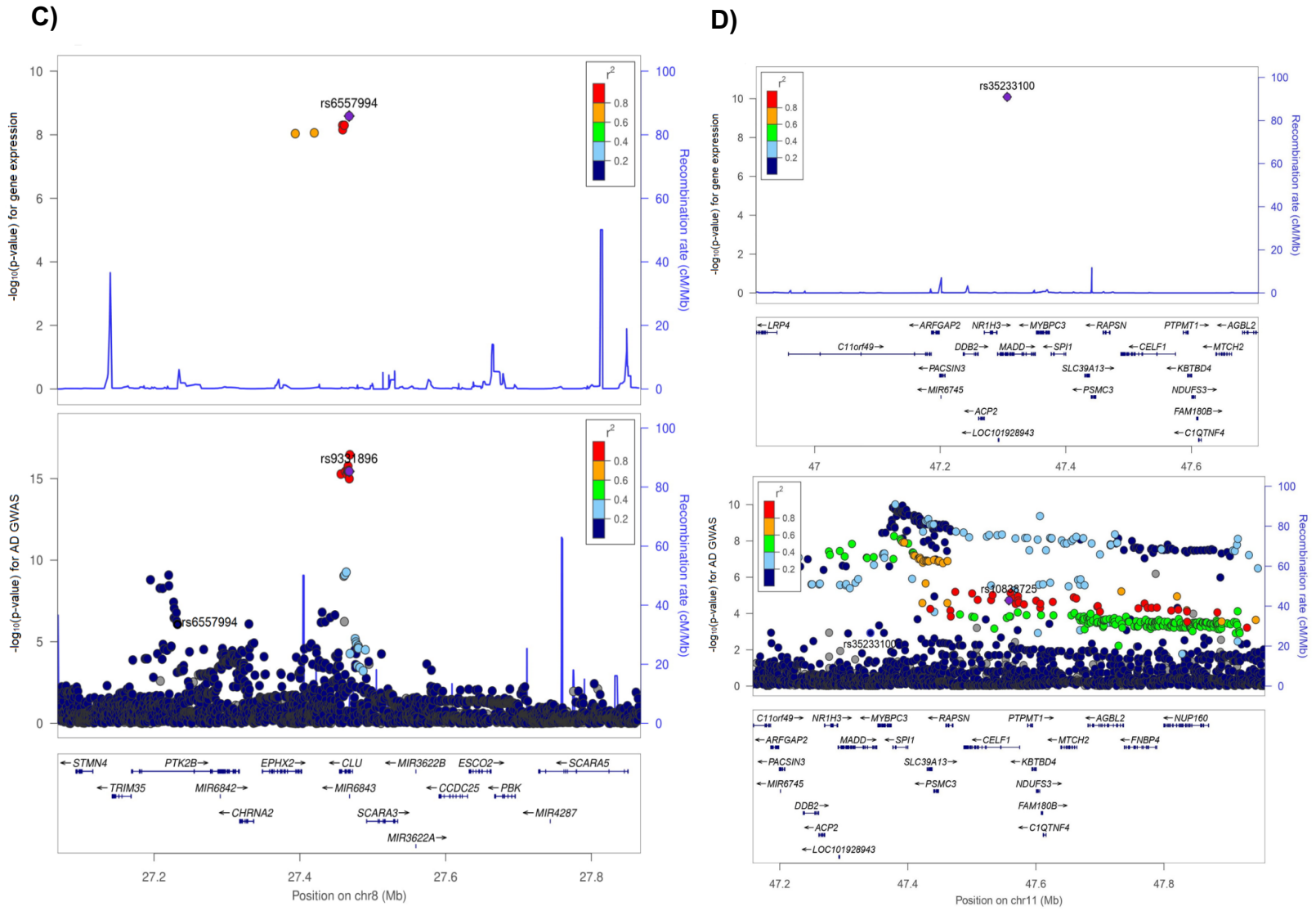


Figure S4.2 Regional plots for colocalized AD GWAS/lead eQTL variant pairs. **A)** rs10948363 *CD2AP*/ rs4711880 *CD2AP;* **B)** rs28834970 *PTK2B*/ rs6557994 *PTK2B;* **C)** rs9331896 *CLU*/ rs6557994 *PTK2B;* **D)** rs10838725 *CELF1*/ rs35233100 *MADD;* **E)** rs983392 *MS4A6A*/ rs11230563 *CD6;* **F)** rs429358 *APOE*/ rs74253343 *RELB*.




C)
D)
rs6557994
rs9331896
rs35233100
rs10828729
-log10(p-value) for gene expression
-log10(p-value) for AD GWAS
Recombination rate (cM/Mb)
r2
0.8
0.6
0.4
0.2
STMN4
PTK2B
EPHX2
CLU
MIR3622B
ESCO2
SCARA5
TRIM35
MIR6842
MIR6843
CCDC25
PBK
CHRNA2
SCARA3
MIR4287
MIR3622A
LRP4
ARFGAP2
NR1H3
MYBPC3
RAPSN
PTPMT1
AGBL2
C11orf49
DDB2
MADD
SPI1
CELF1
MTCH2
PACSIN3
SLC39A13
KBTBD4
MIR6745
PSMC3
NDUFS3
ACP2
FAM180B
LOC101928943
C1QTNF4
NUP160
FNBP4
Position on chr8 (Mb)
Position on chr11 (Mb)

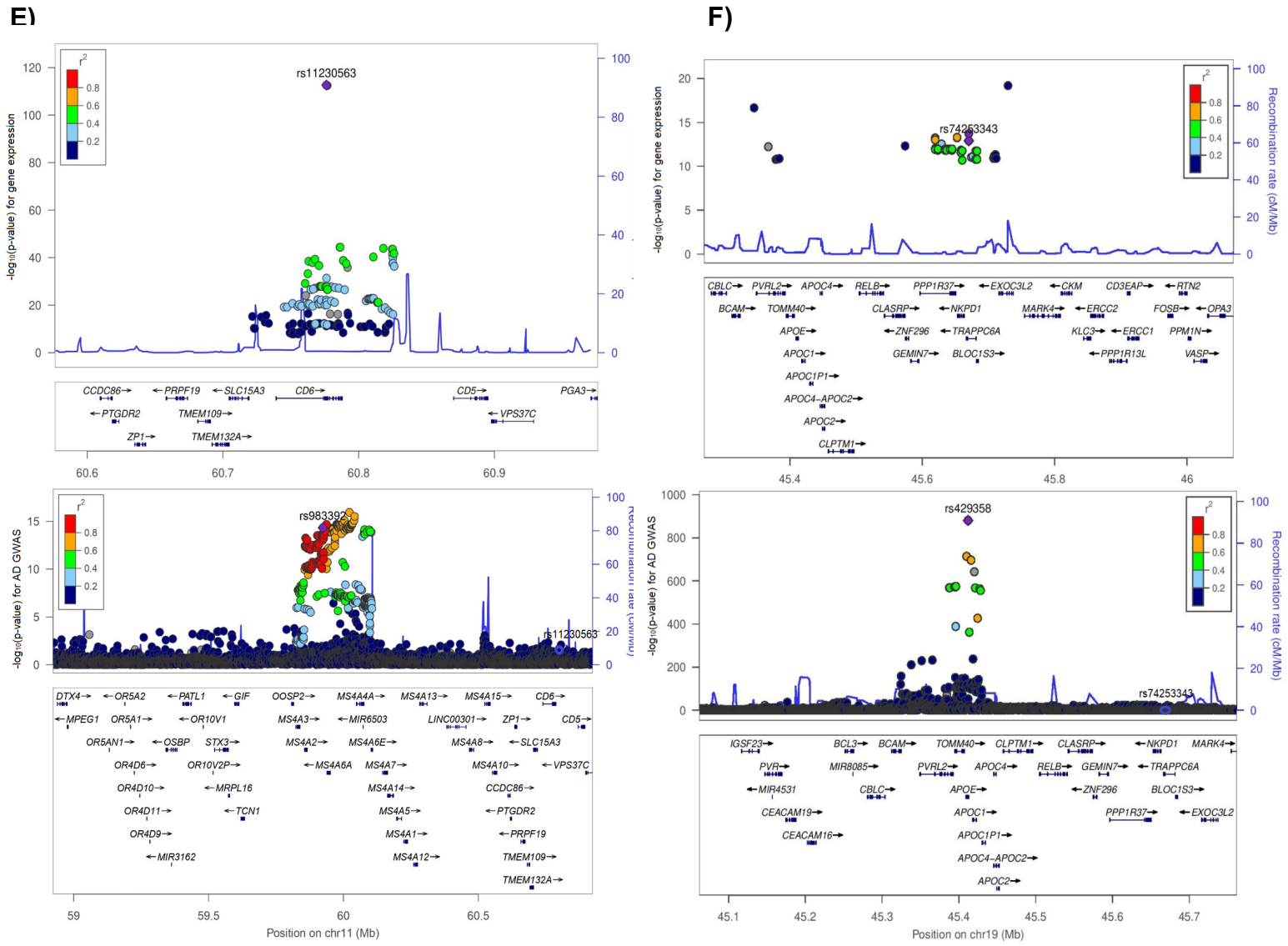

E)
F)
rs11230563
rs74253343
rs983392
rs429358
-log10(p-value) for gene expression
-log10(p-value) for AD GWAS
Recombination rate (cM/Mb)
Position on chr11 (Mb)
Position on chr19 (Mb)

Table S5.1 Rare eQTLs for top rare variant genes in *NOTCH3* and *TREM2*

CHR	BEGIN POS	END POS	NUM PASS VARS	NUM SING VARS	P-VALUE	STATRHO	GENE	TISSUE
19	14270543	16308917	659	416	0.1309	0	NOTCH3	BLOOD
6	40133725	42130917	716	436	0.12376	0	TREM2	BLOOD
6	40133952	42130904	452	288	0.39207	0	TREM2	BRAIN

BEGIN POS : Beginning position of range for rare variants within 1 Mb of gene to be tested
END POS : End position of range for rare variants within 1 Mb of gene to be tested
NUM PASS VARS : Number of variants passing all thresholds for EPACTS software
NUM SING VARS : Number of singletons among variants in NUM PASS VARS
P-VALUE : P-value of burden tests
STATRHO: represents the RHO value from SKAT-O test, rho = 1 (burden) and rho = 0 (SKAT)

Table S5.2 Intersect of rare and common eQTLs in blood and brain

Blood				Brain
ABCA7	DPPA4	LRRC6	RNF181	ACOT1
ABHD2	DSC2	MAF1	RPA2	ADAL
ACOX1	EBPL	MAPKAPK3	RPS23	ANKK1
ACSL6	ECHDC3	MEGF9	RSRC1	ANKRD30BL
ACSM1	ENGASE	MGAM	S100A12	ANXA9
ADAMTSL4	ENTPD1	MMP24	S100P	C10orf107
AGPAT1	EPB41L4A	MMP9	SHC1	C2orf74
AIG1	ESPN	MPC2	SIGLEC5	C5orf17
ALOX5AP	EXOC2	MRPL10	SIGLEC9	CCDC163P
APRT	EXOC4	MRPS7	SIN3B	CCDC173
ARID3B	FAAH	MRVI1	SLC22A1	CYP2D6
ARL17A	FCGBP	MS4A6A	SLPI	DNLZ
ARRB2	FCGR2B	MXI1	SMAD1	FAM154B
ARSG	FKBP1A	MYOM1	SMAP1	GNMT
ASAH1	FNBP1L	MZT2A	SNX19	GUF1
ASPRV1	FOLR3	NLRP1	SPATA20	HLA-A
ATG7	FST	NLRP3	SPPL3	HLA-DOB
ATL1	GAD1	NMRAL1	SSH3	HLA-DRB1
ATP6V0D1	GCAT	NSMAF	ST6GALNAC2	HLA-DRB5
ATP6V1D	GSTA3	NUDT2	STYXL1	HP
ATXN7L3B	GSTA4	NUMA1	TAC3	HPR
BEGAIN	GTF3C3	NUP107	TAP2	IL27
CARS2	HAL	NUP50	TIAM2	LDHC
CCDC122	HAUS4	PADI2	TMCC3	LYPD8
CCDC146	HEATR6	PADI4	TMED6	MMRN1
CD300A	HEBP1	PBX2	TMEM163	NBPF16
CD300C	HEBP2	PCYT1A	TMEM51	NPIPA2
CD36	HIP1	PDLIM5	TOP1MT	NPIPA8
CDA	HMBOX1	PEX6	TP53I3	OPN1SW
CDK5R1	HOXB2	PGM1	TREML4	POMC
CFL1	HOXB3	PGM2L1	TSPAN16	PSORS1C1
CHI3L1	HP	PGM5	TUBB	RBPMS2
CHURC1	HP1BP3	PISD	TUBB2A	RNF39
CISD1	HSD17B13	PLA2G4C	TUBB6	RPRD2
CISD2	HSPA1B	PLAGL1	UBE4B	TRIM63
CLN6	IER3	PNKD	UFSP2	YBEY
CLTCL1	IKZF3	PNLDC1	UHRF1BP1	ZNF253
CLYBL	IMPA2	POLR1A	ULK4	ZNF514
CPA5	INPP1	PPP3CA	UQCRC1	ZNRD1
CPVL	IP6K2	PPT1	USP10	ZSCAN31
CREB5	KIAA1191	PYGB	UTS2	
CTNNAL1	KIF1B	RAB27A	UVSSA	
CTSK	LDHC	RABEP1	VNN2	
CTSW	LGALS4	RALBP1	WARS2	
CXCL1	LGALS8	RBM23	XRRA1	
DAD1	LILRA1	RBMS1	YTHDC2	
DEF6	LILRA2	RBP7	ZDHHC2	
DGCR8	LIN7A	RBPMS2	ZNF502	
DHRS4	LPXN	RHD	ZNF605	
DNAJC15	LRRC2	RIN1	ZP3	
DOK4	LRRC4	RNF14		

Table S5.3 Intersect of rare and cell-type specific eQTLs in blood and brain

Blood	Brain
ASAH1	ACOT1
ATXN7L3B	C2orf74
CHI3L1	CCDC163P
CISD1	CCDC173
CLN6	COPZ1
DAD1	CYP2D6
DNAJC15	ESCO2
FNBP1L	FAM186A
FOLR3	GALNT5
HAUS4	GNMT
HEATR6	GUF1
HP	HLA-DOB
HSD17B13	HLA-DRB1
KIF1B	HLA-DRB5
LRRC6	HP
PADI2	HPR
PEX6	LDHC
PGM5	NPIPA2
PISD	POMC
PPT1	RNF39
RBPMS2	RP11-529K1.3
S100A12	RPRD2
S100P	YBEY
SPATA20	ZNF101
TREML4	ZNF253
TUBB2A	
UHRF1BP1	
UTS2	
XRRA1	

Table S5.4 Intersect of rare, common, and cell-type specific eQTLs in blood and brain

Blood	Brain
ASAH1	ACOT1
ATXN7L3B	C2orf74
CHI3L1	CCDC163P
CISD1	CCDC173
CLN6	CYP2D6
DAD1	GNMT
DNAJC15	GUF1
FNBP1L	HLA-DOB
FOLR3	HLA-DRB1
HAUS4	HLA-DRB5
HEATR6	HP
HP	HPR
HSD17B13	LDHC
KIF1B	NBPF16
LRRC6	NPIPA2
PADI2	POMC
PEX6	RNF39
PGM5	RPRD2
PISD	TRIM63
PPT1	YBEY
RBPMS2	ZNF253
S100A12	
S100P	
SPATA20	
TREML4	
TUBB2A	
UHRF1BP1	
UTS2	
XRRA1	

Blood + Brain intersect

BIBLIOGRAPHY

1. Facts and Figures [Internet].; 2020 [updated 19 Mar.; cited Jun 20, 2020]. Available from: https://alz.org/alzheimers-dementia/facts-figures.

2. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes and environments for explaining Alzheimer disease. Archives of General Psychiatry. 2006 Feb;63(2):168-174.

3. Sims R, Lee, Sven J van der, Naj AC, Bellenguez C, Badarinarayan N, Jakobsdottir J, et al. Rare coding variants in *PLCG2*, *ABI3*, and *TREM2* implicate microglial-mediated innate immunity in Alzheimer's disease. Nature Genetics. 2017 -09;49(9):1373-1384.

4. Li X, Kim Y, Tsang EK, Davis JR, Damani FN, Chiang C, et al. The impact of rare variation on gene expression across tissues. Nature. 2017 -10;550(7675): 239-243.

5. Rao S, Ghani M, Guo Z, Deming Y, Wang K, Sims R, et al. An APOE-independent cis-eSNP on chromosome 19q13.32 influences tau levels and late-onset Alzheimer's disease risk. Neurobiology of Aging. 2018 Jun;66: 178.e1,178.e8.

6. Zou F, Carrasquillo MM, Pankratz VS, Belbin O, Morgan K, Allen M, et al. Gene expression levels as endophenotypes in genome-wide association studies of Alzheimer disease. Neurology. 2010 -2-9;74(6):480-486.

7. Jonkers IH, Wijmenga C. Context-specific effects of genetic variants associated with autoimmune disease. Human molecular genetics. 2017 Oct 1,;26(R2):R185-192.

8. Dobbyn A, Huckins LM, Boocock J, Sloofman LG, Glicksberg BS, Giambartolomei C, et al. Landscape of Conditional eQTL in Dorsolateral Prefrontal Cortex and Co-localization with Schizophrenia GWAS. American Journal of Human Genetics. 2018 -6-07;102(6):1169-1184.

9. Daria V Zhernakova, Patrick Deelen, Martijn Vermaat, Maarten van Iterson, Michiel van Galen, Wibowo Arindrarto, et al. Identification of context-dependent expression quantitative trait loci in whole blood. Nature Genetics. 2017 Jan 1,; 49(1):139-145.

10. Microglia in Alzheimer's Disease: It's All About Context [Internet].: Hindawi; 2012 [cited Sep 17, 2017]. Available from: https://www.hindawi.com/journals/ijad/2012/314185/abs/.

11. Blair JA, Wang C, Hernandez D, Siedlak SL, Rodgers MS, Achar RK, et al. Individual Case Analysis of Postmortem Interval Time on Brain Tissue Preservation. PLoS One. 2016 -3-16;11(3).

12. Tajuddin SM, Schick UM, Eicher JD, Chami N, Giri A, Brody JA, et al. Large-Scale Exome-wide Association Analysis Identifies Loci for White Blood Cell Traits and Pleiotropy with Immune-Mediated Diseases. American Journal of Human Genetics. 2016 Jul 07,;99(1):22-39.

13. Raj T, Rothamel K, Mostafavi S, Ye C, Lee MN, Replogle JM, et al. Polarization of the Effects of Autoimmune and Neurodegenerative Risk Alleles in Leukocytes. Science. 2014 -5-2;344(6183):519-523.

14. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. TREM2 variants in Alzheimer's disease. The New England Journal of Medicine. 2013 Jan 10,;368(2):117-127.

15. Logue MW, Schu M, Vardarajan BN, Farrell J, Bennett DA, Buxbaum JD, et al. Two rare AKAP9 variants are associated with Alzheimer disease in African Americans. Alzheimer’s & Dementia. 2014 -11;10(6): 609,618.e11.

16. Wetzel-Smith MK, Hunkapiller J, Bhangale TR, Srinivasan K, Maloney JA, Atwal JK, et al. A rare mutation in UNC5C predisposes to late-onset Alzheimer's disease and increases neuronal cell death. Nature Medicine. 2014 Dec;20(12): 1452-1457.

17. Bis J, Jian X, Kunkle B, Chen Y, Hamilton-Nelson K, et al. Whole Exome Sequencing Study Identifies Novel Rare and Common Alzheimer’s-Associated Variants Involved in Immune Response and Transcriptional Regulation. Molecular Psychiatry. 2020; 25:1859–1875.

18. Zhang X, Zhu C, Beecham G, Vardarajan BN, Ma Y, Lancour D, et al. A rare missense variant of CASP7 is associated with familial late-onset Alzheimer's disease. Alzheimer’s & Dementia. 2019 03;15(3):441-452.

19. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984 Jul;34(7):939-944.

20. Challis D, Yu J, Evani U, Yu F. Atlas 2. BCM-HGSC. 2014 -03-06T09:32:29-06:00;Version 1.4.1.

21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics. 2007 Sep;81(3):559-575.

22. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biology. 2016 06 06,;17(1):122.

23. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nature Genetics. 2014 -03;46(3):310-315.

24. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005 Jan 15,;21(2):263-265.

25. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. American Journal of Human Genetics. 2008 Sep;83(3):311-321.

26. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990 Oct 05,;215(3):403-410.

27. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 2006 Jan 15,;22(2):195-201.

28. Schrödinger L. Maestro. 2018;Schrödinger Release 2018-1.

29. Luca VC, Kim BC, Ge C, Kakuda S, Wu D, Roein-Peikar M, et al. Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. Science. 2017 03 24,;355(6331):1320-1324.

30. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Research. 2017 -1-04;45(Database issue):D362-368.

31. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Research. 2017 /01/04;45(D1):D183-189.

32. Kauwe JSK, Ridge PG, Foster NL, Cannon-Albright LA. Strong Evidence for a Genetic Contribution to Late-Onset Alzheimer's Disease Mortality: A Population-Based Study. PLoS ONE. 2013 Oct 8,;8(10):e77087.

33. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. Nature. 2015 Oct 01,;526(7571):68-74.

34. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016 - 08;536(7616):285-291.

35. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP) [Internet].; 2018. Available from: http://evs.gs.washington.edu/EVS/.

36. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine. 2015 May; 17(5):405-424.

37. Naj AC, Jun G, Beecham GW, Wang L, Vardarajan BN, Buros J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nature Genetics. 2011 May; 43(5):436-441.

38. Erikson GA, Bodian DL, Rueda M, Molparia B, Scott ER, Scott-Van Zeeland AA, et al. Whole-Genome Sequencing of a Healthy Aging Cohort. Cell. 2016 May 05,;165(4):1002-1011.

39. Joutel A, Corpechot C, Ducros A, Vahedi K, Chabriat H, Mouton P, et al. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. Nature. 1996 Oct 24,;383(6602):707-710.

40. Costa MD, Pereira JB, Pala M, Fernandes V, Olivieri A, Achilli A, et al. A substantial prehistoric European ancestry amongst Ashkenazi maternal lineages. Nature Communications. 2013;4:2543.

41. Beecham GW, Bis JC, Martin ER, Choi S-, DeStefano AL, van Duijn CM, et al. The Alzheimer's Disease Sequencing Project: Study design and sample selection. Neurology Genetics. 2017;3(5).

42. Nho K, Horgusluoglu E, Kim S, Risacher SL, Kim D, Foroud T, et al. Integration of bioinformatics and imaging informatics for identifying rare PSEN1

variants in Alzheimer's disease. BMC Medical Genomics. 2016 08 12,;9 Suppl 1:30.

43. Joutel A, Vahedi K, Corpechot C, Troesch A, Chabriat H, Vayssière C, et al. Strong clustering and stereotyped nature of Notch3 mutations in CADASIL patients. Lancet. 1997 Nov 22,;350(9090):1511-1515.

44. Oliveri RL, Muglia M, De Stefano N, Mazzei R, Labate A, Conforti FL, et al. A novel mutation in the Notch3 gene in an Italian family with cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy: genetic and magnetic resonance spectroscopic findings. Archives of Neurology. 2001 Sep;58(9):1418-1422.

45. Ghezzi L, Carandini T, Arighi A, Fenoglio C, Arcaro M, De Riz M, et al. Evidence of CNS β-amyloid deposition in Nasu-Hakola disease due to the TREM2 Q33X mutation. Neurology. 2017 Dec 12,;89(24):2503-2505.

46. Kauwe JSK, Jacquart S, Chakraverty S, Wang J, Mayo K, Fagan AM, et al. Extreme cerebrospinal fluid amyloid beta levels identify family with late-onset Alzheimer's disease presenilin 1 mutation. Annals of Neurology. 2007 May;61(5): 446-453.

47. Scheltens P. Clinicopathological concordance and discordance in three monozygotic twin pairs with familial Alzheimer's disease. Journal of Neurology, Neurosurgery, and Psychiatry. 2007 -10;78(10):1039.

48. Influence of low frequency PSEN1 variants on familial Alzheimer’s disease risk in Brazil. Neuroscience Letters. 2017 /07/13;653:341-345.

49. Guerreiro RJ, Baquero M, Blesa R, Boada M, Brás JM, Bullido MJ, et al. Genetic screening of Alzheimer's disease genes in Iberian and African samples yields novel mutations in presenilins and APP. Neurobiology of Aging. 2010 -5;31(5):725-731.

50. Fernández MV, Black K, Carrell D, Saef B, Budde J, Deming Y, et al. SORL1 variants across Alzheimer's disease European American cohorts. European Journal of Human Genetics. 2016 12;24(12):1828-1830.

51. Carney RM, Kohli MA, Kunkle BW, Naj AC, Gilbert JR, Züchner S, et al. Parkinsonism and distinct dementia patterns in a family with the MAPT R406W mutation. Alzheimer’s & Dementia. 2014 May;10(3):360-365.

52. Lindquist SG, Holm IE, Schwartz M, Law I, Stokholm J, Batbayli M, et al. Alzheimer disease-like clinical phenotype in a family with FTDP-17 caused by a

MAPT R406W mutation. European Journal of Neurology. 2008 Apr;15(4):377-385.

53. Ikeuchi T, Kaneko H, Miyashita A, Nozaki H, Kasuga K, Tsukie T, et al. Mutational analysis in early-onset familial dementia in the Japanese population. The role of PSEN1 and MAPT R406W mutations. Dementia and Geriatric Cognitive Disorders. 2008;26(1):43-49.

54. Vardarajan BN, Ghani M, Kahn A, Sheikh S, Sato C, Barral S, et al. Rare coding mutations identified by sequencing of Alzheimer disease genome-wide association studies loci. Annals of Neurology. 2015 Sep;78(3):487-498.

55. Rogaeva E, Meng Y, Lee JH, Gu Y, Kawarai T, Zou F, et al. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer's Disease. Nature Genetics. 2007 -2;39(2):168-177.

56. Jun G, Ibrahim-Verbaas CA, Vronskaya M, Lambert J-, Chung J, Naj AC, et al. A novel Alzheimer disease locus located near the gene encoding tau protein. Molecular Psychiatry. 2016 Jan;21(1):108-117.

57. Cruchaga C, Kauwe JSK, Harari O, Jin SC, Cai Y, Karch CM, et al. GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. Neuron. 2013 Apr 24,;78(2):256-268.

58. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nature Genetics. 2013 Dec;45(12):1452-1458.

59. Farrer LA. Expanding the genomic roadmap of Alzheimer's disease. The Lancet Neurology. 2015 /08/01;14(8):783-785.

60. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research. 2007 Jan;35(Database issue):5.

61. Muiño E, Gallego-Fabrega C, Cullell N, Carrera C, Torres N, Krupinski J, et al. Systematic Review of Cysteine-Sparing NOTCH3 Missense Mutations in Patients with Clinical Suspicion of CADASIL. International Journal of Molecular Sciences. 2017 Sep 13,;18(9).

62. Hu X, He W, Luo X, Tsubota KE, Yan R. BACE1 regulates hippocampal astrogenesis via the Jagged1-Notch pathway. Cell Reports. 2013 Jul 11,;4(1):40-49.

63. Konishi J, Kawaguchi KS, Vo H, Haruki N, Gonzalez A, Carbone DP, et al. Gamma-secretase inhibitor prevents Notch3 activation and reduces proliferation in human lung cancers. Cancer Research. 2007 Sep 01,;67(17):8051-8057.

64. Guerreiro RJ, Lohmann E, Kinsella E, Brás JM, Luu N, Gurunlian N, et al. Exome sequencing reveals an unexpected genetic cause of disease: NOTCH3 mutation in a Turkish family with Alzheimer's disease. Neurobiology of Aging. 2012 May;33(5):1008.e17-23.

65. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Research. 2016 Jan 04,;44(D1):862.

66. Sassi C, Nalls MA, Ridge PG, Gibbs JR, Lupton MK, Troakes C, et al. Mendelian adult-onset leukodystrophy genes in Alzheimer's disease: critical influence of CSF1R and NOTCH3. Neurobiology of Aging. 2018 Jun;66: 179.e17,179.e29.

67. Cuyvers E, Bettens K, Philtjens S, Van Langenhove T, Gijselinck I, van der Zee J, et al. Investigating the role of rare heterozygous TREM2 variants in Alzheimer's disease and frontotemporal dementia. Neurobiology of Aging. 2014 Mar;35(3):726.e11-19.

68. Jin SC, Benitez BA, Karch CM, Cooper B, Skorupa T, Carrell D, et al. Coding variants in TREM2 increase risk for Alzheimer's disease. Human Molecular Genetics. 2014 Nov 01,;23(21):5838-5846.

69. Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. The New England Journal of Medicine. 2013 Jan 10,;368(2):107-116.

70. Guerreiro RJ, Lohmann E, Brás JM, Gibbs JR, Rohrer JD, Gurunlian N, et al. Using Exome Sequencing to Reveal Mutations in TREM2 Presenting as a Frontotemporal Dementia–like Syndrome Without Bone Involvement. JAMA Neurology. 2013 /01/01;70(1):78-84.

71. Guerreiro R, Bilgic B, Guven G, Brás J, Rohrer J, Lohmann E, et al. A novel compound heterozygous mutation in TREM2 found in a Turkish frontotemporal dementia-like family. Neurobiology of Aging. 2013 -12;34(12):2890.e1,2890.e5.

72. Coelho D, Kim JC, Miousse IR, Fung S, du Moulin M, Buers I, et al. Mutations in ABCD4 cause a new inborn error of vitamin B12 metabolism. Nature Genetics. 2012 Oct;44(10):1152-1155.

73. Chen H, Liu S, Ji L, Wu T, Ma F, Ji Y, et al. Associations between Alzheimer's disease and blood homocysteine, vitamin B12, and folate: a case-control study. Current Alzheimer Research. 2015;12(1):88-94.

74. Reitz C, Jun G, Naj A, Rajbhandary R, Vardarajan BN, Wang L, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ϵ4,and the risk of late-onset Alzheimer disease in African Americans. JAMA: The Journal of the American Medical Association. 2013 Apr 10,;309(14):1483-1492.

75. Robinson A, Escuin S, Doudney K, Vekemans M, Stevenson RE, Greene NDE, et al. Mutations in the planar cell polarity genes CELSR1 and SCRIB are associated with the severe neural tube defect craniorachischisis. Human Mutation. 2012 Feb;33(2):440-447.

76. Yamada Y, Fuku N, Tanaka M, Aoyagi Y, Sawabe M, Metoki N, et al. Identification of CELSR1 as a susceptibility gene for ischemic stroke in Japanese individuals by a genome-wide association study. Atherosclerosis. 2009 Nov; 207(1):144-149.

77. Boutin C, Goffinet AM, Tissir F. Chapter Seven - Celsr1–3 Cadherins in PCP and Brain Development. In: Yang Y, editor. Current Topics in Developmental Biology. Academic Press; 2012. p. 161-183.

78. Raghavendra Prasad HS, Qi Z, Srinivasan KN, Gopalakrishnakone P. Potential effects of tetrodotoxin exposure to human glial cells postulated using microarray approach. Toxicon. 2004 November 1,;44(6):597-608.

79. Narindrasorasak S, Lowery DE, Altman RA, Gonzalez-DeWhitt PA, Greenberg BD, Kisilevsky R. Characterization of high affinity binding between laminin and Alzheimer's disease amyloid precursor proteins. Laboratory Investigation. 1992 /11;67(5):643-652.

80. Palu E, Liesi P. Differential distribution of laminins in Alzheimer disease and normal human brain tissue. Journal of Neuroscience Research. 2002 -07-15;69(2):243-256.

81. Saad M, Brkanac Z, Wijsman EM. Family-based genome scan for age at onset of late-onset Alzheimer's disease in whole exome sequencing data. Genes, Brain and Behavior. 2015 Nov;14(8):607-617.

82. Consortium tHR, McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nature Genetics. 2016 -10;48(10):1279-1283.

83. Patel D, Mez J, Vardarajan BN, Staley L, Chung J, Zhang X, et al. Association of Rare Coding Mutations With Alzheimer Disease and Other Dementias Among Adults of European Ancestry. JAMA Network Open. 2019 /03/01;2(3):e191350-.

84. Zhao J, Akinsanmi I, Arafat D, Cradick TJ, Lee CM, Banskota S, et al. A Burden of Rare Variants Associated with Extremes of Gene Expression in Human Peripheral Blood. The American Journal of Human Genetics. 2016 -02-04;98(2):299-309.

85. Montgomery SB, Lappalainen T, Gutierrez-Arcelus M, Dermitzakis ET. Rare and Common Regulatory Variation in Population-Scale Sequenced Human Genomes. PLoS Genetics. 2011 Jul 21,;7(7):e1002144.

86. Zeng Y, Wang G, Yang E, Ji G, Brinkmeyer-Langford CL, Cai JJ. Aberrant Gene Expression in Humans. PLoS Genetics. 2015 Jan 24;11(1):e1004942.

87. Daye ZJ, Chen J, Li H. High-Dimensional Heteroscedastic Regression with an Application to eQTL Data Analysis. Biometrics. 2012 -3;68(1):316-326.

88. Sun W, Ibrahim JG, Zou F. Genomewide Multiple-Loci Mapping in Experimental Crosses by Iterative Adaptive Penalized Regression. Genetics. 2010 -5;185(1):349-359.

89. Yang H, Lin C, Chen C, Chen JJ. Applying genome-wide gene-based expression quantitative trait locus mapping to study population ancestry and pharmacogenetics. BMC Genomics. 2014 -4-29;15:319.

90. Yang MQ, Li D, Yang W, Zhang Y, Liu J, Tong W. A Gene Module-Based eQTL Analysis Prioritizing Disease Genes and Pathways in Kidney Cancer. Computational and Structural Biotechnology Journal. 2017 -10-10;15:463-470.

91. Lutz SM, Thwing A, Fingerlin T. eQTL mapping of rare variant associations using RNA-seq data: An evaluation of approaches. PLoS ONE. 2019 Oct 3,; 14(10):e0223273.

92. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, et al. Optimal Unified Approach for Rare-Variant Association Testing with Application to Small-Sample Case-Control Whole-Exome Sequencing Studies. American Journal of Human Genetics. 2012 -8-10;91(2):224-237.

93. ADNI | Alzheimer's Disease Neuroimaging Initiative [Internet].; 2017 [cited Dec 11, 2018]. Available from: http://adni.loni.usc.edu/.

94. Bennett DA, Schneider JA, Arvanitakis Z, Wilson RS. Overview and findings from the religious orders study. Current Alzheimer Research. 2012 Jul;9(6):628-645.

95. Overview and Findings from the Rush Memory and Aging Project [Internet].; 2012 [updated /06/30; cited Sep 7, 2018]. Available from: http://www.eurekaselect.com/99959/article.

96. AMP-AD Knowledge Portal [Internet].; 2018. Available from: https://www.synapse.org/.

97. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012 -3-15;28(6):882-883.

98. R: A language and environment for statistical computing. R Foundation for Statistical Computing [Internet].; 2020 [cited Apr 8, 2020]. Available from: http://www.R-project.org/.

99. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008 Dec 29,;9:559.

100. Zhang B, Gaiteri C, Bodea L, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell. 2013 Apr 25,;153(3):707-720.

101. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer’s disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nature Genetics. 2019 -3;51(3):414-430.

102. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nature Neuroscience. 2014 Sep;17(9):1156-1163.

103. Smith RG, Lunnon K. DNA Modifications and Alzheimer’s Disease. In: Delgado-Morales R, editor. Neuroepigenomics in Aging and Disease. Cham: Springer International Publishing; 2017. p. 303-319.

104. Blalock EM, Geddes JW, Chen KC, Porter NM, Markesbery WR, Landfield PW. Incipient Alzheimer's disease: Microarray correlation analyses reveal major transcriptional and tumor suppressor responses. Proceedings of the National Academy of Sciences of the United States of America. 2004 -2-17;101(7):2173-2178.

105. Maphis NM, Jiang S, Binder J, Wright C, Gopalan B, Lamb BT, et al. Whole Genome Expression Analysis in a Mouse Model of Tauopathy Identifies MECP2 as a Possible Regulator of Tau Pathology. Frontiers in Molecular Neuroscience. 2017;10.

106. Nagata Y, Hirayama A, Ikeda S, Shirahata A, Shoji F, Maruyama M, et al. Comparative analysis of cerebrospinal fluid metabolites in Alzheimer's disease and idiopathic normal pressure hydrocephalus in a Japanese cohort. Biomarker Research. 2018;6:5.

107. Wei C, Cui P, Li H, Lang W, Liu G, Ma X. Shared genes between Alzheimer's disease and ischemic stroke. CNS Neuroscience & Therapeutics. 2019 -3-11;25(8):855-864.

108. Steele NZR, Carr JS, Bonham LW, Geier EG, Damotte V, Miller ZA, et al. Fine-mapping of the human leukocyte antigen locus as a risk factor for Alzheimer disease: A case–control study. PLoS Medicine. 2017 Mar 28,;14(3):e1002272.

109. Patel D, Zhang X, Farrell J, Chung J, Stein T.D., Lunetta K.L., Farrer L.A. Cell-type Specific Expression Quantitative Trait Loci Associated with Alzheimer Disease in Blood and Brain Tissue. Manuscript submitted for publication. 2020.

110. Huang K, Marcora E, Pimenova AA, Di Narzo AF, Kapoor M, Jin SC, et al. A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. Nature Neuroscience. 2017 -8;20(8):1052-1061.

111. Drummond E, Nayak S, Faustin A, Pires G, Hickman R, Askenazi M, et al. Proteomic Differences in Amyloid Plaques in Rapidly Progressive and Sporadic Alzheimer's Disease. Acta Neuropathologica. 2017 -6;133(6):933-954.

112. Latz E, Xiao TS, Stutz A. Activation and regulation of the inflammasomes. Nature Reviews. Immunology 2013 -6;13(6).

113. Moradifard S, Hoseinbeyki M, Ganji SM, Minuchehr Z. Analysis of microRNA and Gene Expression Profiles in Alzheimer's Disease: A Meta-Analysis Approach. Scientific Reports. 2018 -03-19;8(1):1-17.

114. Gleichmann M, Zhang Y, Wood WH, Becker KG, Mughal MR, Pazin MJ, et al. Molecular changes in brain aging and Alzheimer's disease are mirrored in experimentally silenced cortical neuron networks. Neurobiology of Aging. 2012 -1;33(1):205.e1,205.e18.

115. McShea A, Zelasko DA, Gerst JL, Smith MA. Signal transduction abnormalities in Alzheimer's disease: evidence of a pathogenic stimuli. Brain Research. 1999 January 9,;815(2):237-242.

116. Huang X, Chen Y, Li W, Cohen SN, Liao F, Li L, et al. The Rps23rg gene family originated through retroposition of the ribosomal protein s23 mRNA and encodes proteins that decrease Alzheimer's β-amyloid level and tau phosphorylation. Human Molecular Genetics. 2010 -10-1;19(19):3835-3843.

117. Yan L, Chen Y, Li W, Huang X, Badie H, Jian F, et al. RPS23RG1 reduces Aβ oligomer-induced synaptic and cognitive deficits. Scientific Reports. 2016 -01-06;6.

118. Kreft KL, van Meurs M, Wierenga-Wolf AF, Melief M, van Strien ME, Hol EM, et al. Abundant kif21b is associated with accelerated progression in neurodegenerative diseases. Acta Neuropathologica Communications. 2014 -10-3;2.

119. Hares K, Miners JS, Cook AJ, Rice C, Scolding N, Love S, et al. Overexpression of Kinesin Superfamily Motor Proteins in Alzheimer's Disease. Journal of Alzheimer’s Disease. 2017;60(4):1511-1524.

120. Esteve P, Rueda-Carrasco J, Mateo MI, Martin-Bermejo MJ, Draffin J, Pereyra G, et al. Elevated levels of Secreted-Frizzled-Related-Protein 1 contribute to Alzheimer’s disease pathogenesis. Nature Neuroscience. 2019 -08;22(8):1258-1268.

121. GeneCards®: The Human Gene Database [Internet].; 2018.

122. Carrasco M, Rabaneda LG, Murillo-Carretero M, Ortega-Martínez S, Martínez-Chantar ML, Woodhoo A, et al. Glycine N-methyltransferase expression in the hippocampus and its role in neurogenesis and cognitive performance. Hippocampus. 2014 -7;24(7):840-852.

123. Lardenoije R, Roubroeks JAY, Pishva E, Leber M, Wagner H, Iatrou A, et al. Alzheimer's disease-associated (hydroxy)methylomic changes in the brain and blood. Clinical Epigenetics. 2019 11 27,;11(1):164.

124. Chen X, Ji B, Hao X, Li X, Eisele F, Nyström T, et al. FMN reduces Amyloid-β toxicity in yeast by regulating redox status and cellular metabolism. Nature Communications. 2020 -02-13;11(1):1-16.

125. Song I, Kim Y, Chung S, Cho H. Association between serum haptoglobin and the pathogenesis of Alzheimer's disease. Internal Medicine. 2015;54(5):453-457.

126. Meng Q, Zhuang Y, Ying Z, Agrawal R, Yang X, Gomez-Pinilla F. Traumatic Brain Injury Induces Genome-Wide Transcriptomic, Methylomic, and Network Perturbations in Brain and Blood Predicting Neurological Disorders. EBioMedicine. 2017 -2-01;16:184-194.

127. Deane CAS, Brown IR. Intracellular Targeting of Heat Shock Proteins in Differentiated Human Neuronal Cells Following Proteotoxic Stress. Journal of Alzheimer's Disease. 2018 /01/01;66(3):1295-1308.

128. Squillario M, Barla A. A computational procedure for functional characterization of potential marker genes from molecular data: Alzheimer's as a case study. BMC Medical Genomics. 2011 -7-5;4:55.

129. Stomrud E, Björkqvist M, Janciauskiene S, Minthon L, Hansson O. Alterations of matrix metalloproteinases in the healthy elderly with increased risk of prodromal Alzheimer's disease. Alzheimer’s Research & Therapy. 2010 -6-24;2(3):20.

130. Kang SS, Ebbert MTW, Baker KE, Cook C, Wang X, Sens JP, et al. Microglial translational profiling reveals a convergent APOE pathway from aging, amyloid, and tau. Journal of Experimental Medicine. 2018 -9-03;215(9):2235-2245.

131. Domingues C, Cruz e Silva, Odete A. B., Henriques AG. Impact of Cytokines and Chemokines on Alzheimer’s Disease Neuro-pathological Hallmarks. Current Alzheimer Research. 2017 -8;14(8):870-882.

132. Kauwe JSK, Bailey MH, Ridge PG, Perry R, Wadsworth ME, Hoyt KL, et al. Genome-Wide Association Study of CSF Levels of 59 Alzheimer's Disease Candidate Proteins: Significant Associations with Proteins Involved in Amyloid Processing and Inflammation. PLoS Genetics. 2014 Oct 23,;10(10):e1004758.

133. Yan T, Ding F, Zhao Y. Integrated identification of key genes and pathways in Alzheimer’s disease via comprehensive bioinformatical analyses. Hereditas. 2019 -7-16;156.

134. Cheng Y, Ma X, Wei Y, Wei X. Potential roles and targeted therapy of the CXCLs/CXCR2 axis in cancer and inflammatory diseases. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2019 April 1,;1871(2):289-312.

135. Liu C, Cui G, Zhu M, Kang X, Guo H. Neuroinflammation in Alzheimer's disease: chemokines produced by astrocytes and chemokine receptors. International Journal of Clinical and Experimental Pathology. 2014 -12-01;7(12): 8342-8355.

136. Majumder P, Roy K, Singh BK, Jana NR, Mukhopadhyay D. Cellular levels of Grb2 and cytoskeleton stability are correlated in a neurodegenerative scenario. Disease Models & Mechanisms. 2017 -5-1;10(5):655-669.

137. Zuena AR, Casolini P, Lattanzi R, Maftei D. Chemokines in Alzheimer's Disease: New Insights Into Prokineticins, Chemokine-Like Proteins. Frontiers in Pharmacology. 2019;10.

138. Ma Q, Yang F, Frautschy SA, Cole GM. PAK in Alzheimer disease, Huntington disease and X-linked mental retardation. Cell Logistics. 2012 -4-01; 2(2):117-125.

139. Moore Z, Mobilio F, Walker FR, Taylor JM, Crack PJ. Abrogation of type-I interferon signalling alters the microglial response to Aβ 1–42. Scientific Reports. 2020 -02-21;10(1):1-14.

140. Zhang H, Wang D, Gong P, Lin A, Zhang Y, Ye RD, et al. Formyl Peptide Receptor 2 Deficiency Improves Cognition and Attenuates Tau Hyperphosphorylation and Astrogliosis in a Mouse Model of Alzheimer's Disease. Journal of Alzheimer's Disease. 2019;67(1):169-179.

141. Bonham LW, Sirkis DW, Yokoyama JS. The Transcriptional Landscape of Microglial Genes in Aging and Neurodegenerative Disease. Frontiers in Immunology. 2019;10.

142. Hsu W-, Wang H-, Lee L-, Fung H-, Lin J-, Hsu H-, et al. Promoter polymorphisms modulating HSPA5 expression may increase susceptibility to Taiwanese Alzheimer's disease. Journal of Neural Transmission (Vienna). 2008 Nov;115(11):1537-1543.

143. Hj S. Clinical similarities and differences between Alzheimer's disease and Parkinson's disease. Journal of Neural Transmission Suppl. 1987 /01/01;24:87-99.

144. Booth L, Roberts JL, Dent P. HSPA5/Dna K May Be a Useful Target for Human Disease Therapies. DNA Cell Biology. 2015 -3-01;34(3):153-158.

145. Kong W, Mou X, Deng J, Di B, Zhong R, Wang S, et al. Differences of immune disorders between Alzheimer's disease and breast cancer based on transcriptional regulation. PLoS ONE. 2017 Jul 18,;12(7):e0180337.

146. Battle A, Brown CD, Engelhardt BE, Montgomery SB. Genetic effects on gene expression across human tissues. Nature. 2017 10 11,;550(7675):204-213.

147. Efthymiou AG, Goate AM. Late onset Alzheimer's disease genetics implicates microglial pathways in disease risk. Molecular Neurodegeneration. 2017 -5-26;12.

148. Karch CM, Ezerskiy LA, Bertelsen S, Goate AM. Alzheimer's Disease Risk Polymorphisms Regulate Gene Expression in the ZCWPW1 and the CELF1 Loci. PLoS One. 2016 -2-26;11(2).

149. Novikova G, Kapoor M, TCW J, Abud EM, Efthymiou AG, Cheng H, et al. Integration of Alzheimer's disease genetics and myeloid genomics reveals novel disease risk mechanisms. bioRxiv. 2019:694281.

150. Zhang X, Joehanes R, Chen BH, Huan T, Ying S, Munson PJ, et al. Identification of common genetic variants controlling transcript isoform variation in human whole blood. Nature Genetics. 2015 /04;47(4):345-352.

151. Therneau T. The Lmekin Function. 2018(Rochester, MN: Mayo Clinic).

152. Penney J, Ralvenius WT, Tsai L. Modeling Alzheimer's disease with iPSC-derived brain cells. Molecular Psychiatry. 2020 -01;25(1):148-167.

153. McKenzie AT, Wang M, Hauberg ME, Fullard JF, Kozlenkov A, Keenan A, et al. Brain Cell Type Specific Gene Expression and Co-expression Network Architectures. Scientific Reports. 2018 -06-11;8(1):8868.

154. Ren Y, Blitterswijk M, Allen M, Carrasquillo M, Reddy J, Wang X, et al. TMEM106B haplotypes have distinct gene expression patterns in aged brain. Molecular Neurodegeneration. 2018 December 1,;13.

155. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLoS Genetics. 2014 May 15,;10(5):e1004383.

156. Wu Y, Broadaway KA, Raulerson CK, Scott LJ, Pan C, Ko A, et al. Colocalization of GWAS and eQTL signals at loci with multiple signals identifies

additional candidate genes for body fat distribution. Human Molecular Genetics. 2019 /12/15;28(24):4161-4172.

157. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010 /09/15;26(18):2336-2337.

158. Hondius DC, van Nierop P, Li KW, Hoozemans JJM, van der Schors, Roel C., van Haastert ES, et al. Profiling the human hippocampal proteome at all pathologic stages of Alzheimer's disease. Alzheimer’s & Dementia. 2016 -06; 12(6):654-668.

159. Sjöberg AP, Trouw LA, Blom AM. Complement activation and inhibition: a delicate balance. Trends in Immunology. 2009 Feb;30(2):83-90.

160. Rangaraju S, Dammer EB, Raza SA, Rathakrishnan P, Xiao H, Gao T, et al. Identification and therapeutic modulation of a pro-inflammatory subset of disease-associated-microglia in Alzheimer's disease. Molecular Neurodegeneration. 2018 05 21,;13(1):24.

161. Smith AR, Smith RG, Pishva E, Hannon E, Roubroeks JAY, Burrage J, et al. Parallel profiling of DNA methylation and hydroxymethylation highlights neuropathology-associated epigenetic variation in Alzheimer's disease. Clinical Epigenetics. 2019 03 21,;11(1):52.

162. Kim D, Jung S, Kim K, Kim C. Treadmill exercise ameliorates Alzheimer disease-associated memory loss through the Wnt signaling pathway in the streptozotocin-induced diabetic rats. Journal of Exercise Rehabilitation. 2016 -8-31;12(4):276-283.

163. Xu X, Wells AB, O'Brien DR, Nehorai A, Dougherty JD. Cell Type-Specific Expression Analysis to Identify Putative Cellular Mechanisms for Neurogenetic Disorders. The Journal of Neuroscience. 2014 -01-22;34(4):1420-1431.

164. Le Page A, Dupuis G, Frost EH, Larbi A, Pawelec G, Witkowski JM, et al. Role of the peripheral innate immune system in the development of Alzheimer's disease. Experimental Gerontology. 2018 07 01,;107:59-66.

165. Goldeck D, Witkowski JM, Fülop T, Pawelec G. Peripheral Immune Signatures in Alzheimer Disease. Current Alzheimer Research. 2016;13(7):739-749.

166. Harris SA, Harris EA. Herpes Simplex Virus Type 1 and Other Pathogens are Key Causative Factors in Sporadic Alzheimer's Disease. Journal of Alzheimer’s Disease. 2015;48(2):319-353.

167. Obulesu M, Lakshmi MJ. Apoptosis in Alzheimer's disease: an understanding of the physiology, pathology and therapeutic avenues. Neurochemical Research. 2014 Dec;39(12):2301-2312.

168. Fu H, Possenti A, Freer R, Nakano Y, Hernandez Villegas NC, Tang M, et al. A tau homeostasis signature is linked with the cellular and regional vulnerability of excitatory neurons to tau pathology. Nature Neuroscience. 2019 01;22(1):47-56.

169. Crehan H, Holton P, Wray S, Pocock J, Guerreiro R, Hardy J. Complement receptor 1 (CR1) and Alzheimer's disease. Immunobiology. 2012 February 1,; 217(2):244-250.

170. Jun GR, Chung J, Mez J, Barber R, Beecham GW, Bennett DA, et al. Transethnic genome-wide scan identifies novel Alzheimer's disease loci. Alzheimer’s & Dementia. 2017 Jul;13(7):727-738.

171. Desikan RS, Schork AJ, Wang Y, Thompson WK, Dehghan A, Ridker PM, et al. Polygenic Overlap Between C-Reactive Protein, Plasma Lipids, and Alzheimer Disease. Circulation. 2015 Jun 09,;131(23):2061-2069.

172. Chang J-, Chang N-. WWOX dysfunction induces sequential aggregation of TRAPPC6AΔ, TIAF1, tau and amyloid β, and causes apoptosis. Cell Death Discovery. 2015;1:15003.

173. Liu C, Ho P, Lee I-, Chen Y, Chu C, Teng C, et al. WWOX Phosphorylation, Signaling, and Role in Neurodegeneration. Frontiers in Neuroscience. 2018;12: 563.

174. Xiang Z, Haroutunian V, Ho L, Purohit D, Pasinetti GM. Microglia activation in the brain as inflammatory biomarker of Alzheimer's disease neuropathology and clinical dementia. Disease Markers. 2006;22(1-2):95-102.

175. Walker DG, Lue L. Immune phenotypes of microglia in human neurodegenerative disease: challenges to detecting microglial polarization in human brains. Alzheimer’s Research & Therapy. 2015 Aug 19,;7(1):56.

176. Neill D, Curran MD, Middleton D, Mawhinney H, Edwardson JA, McKeith I, et al. Risk for Alzheimer's disease in older late-onset cases is associated with HLA-DRB1*03. Neuroscience Letters. 1999 Nov 12,;275(2):137-140.

177. Lehmann DJ, Wiebusch H, Marshall SE, Johnston C, Warden DR, Morgan K, et al. HLA class I, II & III genes in confirmed late-onset Alzheimer's disease. Neurobiology of Aging. 2001 Jan-Feb;22(1):71-77.

178. Lambert J, Grenier-Boley B, Chouraki V, Heath S, Zelenika D, Fievet N, et al. Implication of the immune system in Alzheimer's disease: evidence from genome-wide pathway analysis. Journal of Alzheimer’s Disease. 2010;20(4): 1107-1118.

179. Hallock P, Thomas MA. Integrating the Alzheimer's Disease Proteome and Transcriptome: A Comprehensive Network Model of a Complex Disease. OMICS. 2012 -1;16(1-2):37-49.

180. Herold C, Hooli BV, Mullin K, Liu T, Roehr JT, Mattheisen M, et al. Family-based association analyses of imputed genotypes reveal genome-wide significant association of Alzheimer's disease with OSBPL6, PTPRG, and PDCL3. Molecular Psychiatry. 2016 11;21(11):1608-1612.

181. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer’s disease risk. Nature Genetics. 2019 -3;51(3):404-413.

182. Johnson JL, Chambers E, Jayasundera K. Application of a Bioinformatics-Based Approach to Identify Novel Putative in vivo BACE1 Substrates. Biomedical Engineering and Computational Biology 2013 -2-03;5:1-15.

183. Vélez JI, Lopera F, Sepulveda-Falla D, Patel HR, Johar AS, Chuah A, et al. APOE*E2 allele delays age of onset in PSEN1 E280A Alzheimer’s disease. Molecular Psychiatry. 2016 -07;21(7):916-924.

184. Lim B, Tsolaki M, Batruch I, Anastasiou A, Frontistis A, Prassas I, et al. Putative autoantibodies in the cerebrospinal fluid of Alzheimer’s disease patients. F1000Research. 2019 -11-11;8.

185. Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nature Genetics. 2014 -09;46(9):989-993.

186. Song F, Poljak A, Kochan NA, Raftery M, Brodaty H, Smythe GA, et al. Plasma protein profiling of Mild Cognitive Impairment and Alzheimer’s disease using iTRAQ quantitative proteomics. Proteome Science. 2014 -1-17;12:5.

187. Wand GS, May C, May V, Whitehouse PJ, Rapoport SI, Eipper BA. Alzheimer's disease: low levels of peptide alpha-amidation activity in brain and CSF. Neurology. 1987 Jun;37(6):1057-1061.

188. Liang X, Slifer M, Martin ER, Schnetz-Boutaud N, Bartlett J, Anderson B, et al. Genomic convergence to identify candidate genes for Alzheimer disease on chromosome 10. Human Mutation. 2009 -3;30(3):463-471.

189. C P, H Y, B H, S W, M C, Ds C, et al. Identification and Analysis of Alzheimer's Candidate Genes by an Amplitude Deviation Algorithm. Journal of Alzheimer’s Disease & Parkinsonism. 2019 /02/02;9(1).

190. 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;2011:729478.

191. Vardarajan BN, Logue M, Cupples A, Lunetta K, Jun G, Buros J, et al. Candidate gene study in the endosome-to-Golgi retrieval pathway reveals association of retromer genes with Alzheimer's disease. Alzheimer's & Dementia. 2010 /07/01;6(4):S194.

192. Kong W, Mou X, Liu Q, Chen Z, Vanderburg CR, Rogers JT, et al. Independent component analysis of Alzheimer's DNA microarray gene expression data. Molecular Neurodegeneration. 2009 Jan 28,;4:5.

193. Hsu M, Dedhia M, Crusio WE, Delprato A. Sex differences in gene expression patterns associated with the APOE4 allele. F1000Research. 2019 -7-23;8.

194. Rademakers R, Cruts M, Sleegers K, Dermaut B, Theuns J, Aulchenko Y, et al. Linkage and Association Studies Identify a Novel Locus for Alzheimer Disease at 7q36 in a Dutch Population-Based Sample. The American Journal of Human Genetics. 2005 October 1;77(4):643-652.

195. Zhao B, Luo T, Li T, Li Y, Zhang J, Shan Y, et al. GWAS of 19,629 individuals identifies novel genetic variants for regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. bioRxiv. 2019:586339.

196. Lutz MW, Sprague D, Barrera J, Chiba-Falek O. Shared genetic etiology underlying Alzheimer’s disease and major depressive disorder. Translational Psychiatry. 2020 -3-9;10.

197. Bik-Multanowski M, Dobosz A. Detection of high expression of complex I mitochondrial genes can indicate low risk of Alzheimer' s disease. International Journal of Clinical and Experimental Pathology; 2017.

198. Cacabelos R. Pharmacogenomics of Alzheimer's and Parkinson's diseases. Neuroscience Letters. 2020 May 01,;726:133807.

199. Pino-Lagos K, Quezada S, Catalán DF, Aguillon JC. Searching for Immune Tolerance Manipulating New Molecules and Exploiting New Concepts on Lymphocyte Biology. Frontiers Media SA; 2016.

200. Blood (human) - MACS Handbook - Miltenyi Biotec - USA [Internet].; 2020. Available from: https://www.miltenyibiotec.com/US-en/resources/macs-handbook/human-cells-and-organs/human-cell-sources/blood-human.html#gref.

201. Frequencies of Cell Types in Human Peripheral Blood [Internet].; 2019. Available from: https://www.stemcell.com/media/files/wallchart/WA10006-Frequencies_Cell_Types_Human_Peripheral_Blood.pdf.

202. von Bartheld CS, Bahney J, Herculano-Houzel S. The Search for True Numbers of Neurons and Glial Cells in the Human Brain: A Review of 150 Years of Cell Counting. The Journal of Comparative Neurology. 2016 -12-15;524(18): 3865-3895.

203. Dinkins MB, Dasgupta S, Wang G, Zhu G, He Q, Kong JN, et al. The 5XFAD Mouse Model of Alzheimer's Disease Exhibits an Age-Dependent Increase in Anti-Ceramide IgG and Exogenous Administration of Ceramide Further Increases Anti-Ceramide Titers and Amyloid Plaque Burden. Journal of Alzheimer's Disease. 2015 -5-7;46(1):55-61.

204. Abou-Raya A, Abou-Raya S. Inflammation: A pivotal link between autoimmune diseases and atherosclerosis. Autoimmunity Reviews. 2006 May 1,; 5(5):331-337.

205. Ivashkiv LB, Donlin LT. Regulation of type I interferon responses. Nature Reviews. Immunology 2014 -1;14(1):36-49.

206. Patel T, Brookes KJ, Turton J, Chaudhury S, Guetta-Baranes T, Guerreiro R, et al. Whole-exome sequencing of the BDR cohort: evidence to support the role of the PILRA gene in Alzheimer's disease. Neuropathology and Applied Neurobiology. 2017 Nov 27.

207. Logue MW, Schu M, Vardarajan BN, Farrell J, Lunetta KL, Jun G, et al. Search for age-related macular degeneration risk variants in Alzheimer disease genes and pathways. Neurobiology of Aging. 2014 Jun;35(6):1510.e7-18.

208. Se S, M L, Rv S, C W, Vn L, J X, et al. Candidate-based screening via gene modulation in human neurons and astrocytes implicates FERMT2 in Aβ and TAU proteostasis. Human Molecular Genetics. 2019 /03/01;28(5):718-735.

209. de Kreuk B, Schaefer A, Anthony EC, Tol S, Fernandez-Borja M, Geerts D, et al. The Human Minor Histocompatibility Antigen1 Is a RhoGAP. PLoS One. 2013 -9-23;8(9).

210. Katsumata Y, Nelson PT, Estus S, Fardo DW. Translating Alzheimer's disease–associated polymorphisms into functional candidates: a survey of IGAP genes and SNPs. Neurobiology of Aging. 2019 February 1,;74:135-146.

211. Del Villar K, Miller CA. Down-regulation of DENN/MADD, a TNF receptor binding protein, correlates with neuronal cell death in Alzheimer's disease brain and hippocampal neurons. Proceedings of the National Academy of Sciences of the United States of America. 2004 Mar 23,;101(12):4210-4215.

212. Dourlen P, Fernandez-Gomez FJ, Dupont C, Grenier-Boley B, Bellenguez C, Obriot H, et al. Functional screening of Alzheimer risk loci identifies PTK2B as an in vivo modulator and early marker of Tau pathology. Molecular Psychiatry. 2017 -06;22(6):874-883.

213. Allen M, Burgess JD, Ballard T, Serie D, Wang X, Younkin CS, et al. Gene expression, methylation and neuropathology correlations at progressive supranuclear palsy risk loci. Acta Neuropathologica. 2016 -8;132(2):197-211.

214. Shigemoto-Mogami Y, Hoshikawa K, Sato K. Activated Microglia Disrupt the Blood-Brain Barrier and Induce Chemokines and Cytokines in a Rat in vitro Model. Frontiers in Cell Neuroscience. 2018;12.

215. Sakae N, Liu C, Shinohara M, Frisch-Daiello J, Ma L, Yamazaki Y, et al. ABCA7 Deficiency Accelerates Amyloid-β Generation and Alzheimer's Neuronal Pathology. The Journal of Neuroscience. 2016 -03-30 00:00:00;36(13):3848-3859.

216. Li H, Karl T, Garner B. Understanding the function of ABCA7 in Alzheimer's disease. Biochemical Society Transactions. 2015 Oct;43(5):920-923.

217. Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M, Myers AJ, et al. Genome-wide association meta-analysis of neuropathologic features of

Alzheimer's disease and related dementias. PLoS Genetics. 2014 Sep;10(9): e1004606.

218. Narain Y, Yip A, Murphy T, Brayne C, Easton D, Evans JG, et al. The ACE gene and Alzheimer's disease susceptibility. Journal of Medical Genetics. 2000 Sep;37(9):695-697.

219. Sogorb-Esteve A, García-Ayllón M, Gobom J, Alom J, Zetterberg H, Blennow K, et al. Levels of ADAM10 are reduced in Alzheimer's disease CSF. Journal of Neuroinflammation. 2018 Jul 25,;15(1):213.

220. Yuan X, Sun S, Tan C, Yu J, Tan L. The Role of ADAM10 in Alzheimer's Disease. Journal of Alzheimer’s Disease. 2017;58(2):303-322.

221. Gurses MS, Ural MN, Gulec MA, Akyol O, Akyol S. Pathophysiological Function of ADAMTS Enzymes on Molecular Mechanism of Alzheimer's Disease. Aging and Disease. 2016 Aug;7(4):479-490.

222. Cheng Y, Huang Y, Liu H. Effect of Apolipoprotein E ε4 Carrier Status on Cognitive Response to Acetylcholinesterase Inhibitors in Patients with Alzheimer's Disease: A Systematic Review and Meta-Analysis. Dementia and Geriatric Cognitive Disorders. 2018;45(5-6):335-352.

223. Tulloch J, Leong L, Thomson Z, Chen S, Lee E, Keene CD, et al. Glia-specific APOE epigenetic changes in the Alzheimer's disease brain. Brain Research. 2018 Aug 03,.

224. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991 Feb 21,;349(6311):704-706.

225. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA: The Journal of the American Medical Association. 2010 May 12,; 303(18):1832-1840.

226. Rosenthal SL, Barmada MM, Wang X, Demirci FY, Kamboh MI. Connecting the dots: potential of data integration to identify regulatory SNPs in late-onset Alzheimer's disease GWAS findings. PLoS ONE. 2014;9(4):e95152.

227. Rehker J, Rodhe J, Nesbitt RR, Boyle EA, Martin BK, Lord J, et al. Caspase-8, association with Alzheimer's Disease and functional analysis of rare variants. PLoS ONE. 2017;12(10):e0185777.

228. Bannwarth S, Ait-El-Mkadem S, Chaussenot A, Genin EC, Lacas-Gervais S, Fragaki K, et al. A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. Brain. 2014 Aug; 137(Pt 8):2329-2345.

229. Skibinski G, Parkinson NJ, Brown JM, Chakrabarti L, Lloyd SL, Hummerich H, et al. Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. Nature Genetics. 2005 Aug;37(8):806-808.

230. van der Zee J, Urwin H, Engelborghs S, Bruyland M, Vandenberghe R, Dermaut B, et al. CHMP2B C-truncating mutations in frontotemporal lobar degeneration are associated with an aberrant endosomal phenotype in vitro. Human Molecular Genetics. 2008 Jan 15,;17(2):313-322.

231. Mez J, Chung J, Jun G, Kriegel J, Bourlas AP, Sherva R, et al. Two novel loci, COBL and SLC10A2, for Alzheimer's disease in African Americans. Alzheimer’s & Dementia. 2017 Feb;13(2):119-129.

232. Rademakers R, Baker M, Nicholson AM, Rutherford NJ, Finch N, Soto-Ortolaza A, et al. Mutations in the colony stimulating factor 1 receptor (CSF1R) gene cause hereditary diffuse leukoencephalopathy with spheroids. Nature Genetics. 2011 Dec 25,;44(2):200-205.

233. Chung J, Zhang X, Allen M, Wang X, Ma Y, Beecham G, et al. Genome-wide pleiotropy analysis of neuropathological traits related to Alzheimer's disease. Alzheimer’s Research & Therapy. 2018 02 20,;10(1):22.

234. van der Knaap, Marjo S., Leegwater PAJ, Könst AAM, Visser A, Naidu S, Oudejans CBM, et al. Mutations in each of the five subunits of translation initiation factor eIF2B can cause leukoencephalopathy with vanishing white matter. Annals of Neurology. 2002 Feb;51(2):264-270.

235. Matsukawa T, Wang X, Liu R, Wortham NC, Onuki Y, Kubota A, et al. Adult-onset leukoencephalopathies with vanishing white matter with novel missense mutations in EIF2B2, EIF2B3, and EIF2B5. Neurogenetics. 2011 Aug;12(3):259-261.

236. Hollingworth P, Harold D, Sims R, Gerrish A, Lambert J, Carrasquillo MM, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nature Genetics. 2011 May;43(5):429-435.

237. Tosto G, Fu H, Vardarajan BN, Lee JH, Cheng R, Reyes-Dumeyer D, et al. F-box/LRR-repeat protein 7 is genetically associated with Alzheimer's disease. Annals of Clinical and Translational Neurology. 2015 Aug;2(8):810-820.

238. Lambert J-, Grenier-Boley B, Harold D, Zelenika D, Chouraki V, Kamatani Y, et al. Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer's disease. Molecular Psychiatry. 2013 Apr;18(4):461-470.

239. Yan J, Deng H-, Siddique N, Fecto F, Chen W, Yang Y, et al. Frameshift and novel mutations in FUS in familial amyotrophic lateral sclerosis and ALS/dementia. Neurology. 2010 Aug 31,;75(9):807-814.

240. Rademakers R, Eriksen JL, Baker M, Robinson T, Ahmed Z, Lincoln SJ, et al. Common variation in the miR-659 binding-site of GRN is a major risk factor for TDP43-positive frontotemporal dementia. Human Molecular Genetics. 2008 Dec 01,;17(23):3631-3642.

241. Lee JH, Kahn A, Cheng R, Reitz C, Vardarajan B, Lantigua R, et al. Disease-related mutations among Caribbean Hispanics with familial dementia. Molecular Genetics & Genomic Medicine. 2014 -9;2(5):430-437.

242. Yu L, Chibnik LB, Srivastava GP, Pochet N, Yang J, Xu J, et al. Association of Brain DNA methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 with pathological diagnosis of Alzheimer disease. JAMA Neurology. 2015 Jan; 72(1):15-24.

243. Escott-Price V, Bellenguez C, Wang L, Choi S, Harold D, Jones L, et al. Gene-wide analysis detects two new susceptibility genes for Alzheimer's disease. PLoS ONE. 2014;9(6):e94661.

244. Vidal R, Frangione B, Rostagno A, Mead S, Révész T, Plant G, et al. A stop-codon mutation in the BRI gene associated with familial British dementia. Nature. 1999 Jun 24,;399(6738):776-781.

245. Ghiso JA, Holton J, Miravalle L, Calero M, Lashley T, Vidal R, et al. Systemic amyloid deposits in familial British dementia. Journal of Biological Chemistry. 2001 Nov 23,;276(47):43909-43914.

246. Vidal R, Revesz T, Rostagno A, Kim E, Holton JL, Bek T, et al. A decamer duplication in the 3' region of the BRI gene originates an amyloid peptide that is associated with dementia in a Danish kindred. Proceedings of the National Academy of Sciences of the United States of America. 2000 Apr 25,;97(9):4920-4925.

247. Schuster J, Sundblom J, Thuresson A, Hassin-Baer S, Klopstock T, Dichgans M, et al. Genomic duplications mediate overexpression of lamin B1 in adult-onset autosomal dominant leukodystrophy (ADLD) with autonomic symptoms. Neurogenetics. 2011 Feb;12(1):65-72.

248. Hoekstra EJ, Mesman S, de Munnik WA, Smidt MP. LMX1B is part of a transcriptional complex with PSPC1 and PSF. PLoS ONE. 2013;8(1):e53122.

249. Chung J, Wang X, Maruyama T, Ma Y, Zhang X, Mez J, et al. Genome-wide association study of Alzheimer's disease endophenotypes at prediagnosis stages. Alzheimer’s & Dementia. 2018 May;14(5):623-633.

250. McKenzie AT, Moyon S, Wang M, Katsyv I, Song W, Zhou X, et al. Multiscale network modeling of oligodendrocytes reveals molecular components of myelin dysregulation in Alzheimer's disease. Molecular Neurodegeneration. 2017 11 06,;12(1):82.

251. Sil S, Periyasamy P, Thangaraj A, Chivero ET, Buch S. PDGF/PDGFR axis in the neural systems. Molecular Aspects of Medicine. 2018 Aug;62:63-74.

252. Cruchaga C, Karch CM, Jin SC, Benitez BA, Cai Y, Guerreiro R, et al. Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. Nature. 2014 Jan 23,;505(7484):550-554.

253. Jun G, Asai H, Zeldich E, Drapeau E, Chen C, Chung J, et al. PLXNA4 is associated with Alzheimer disease and modulates tau phosphorylation. Annals of Neurology. 2014 Sep;76(3):379-392.

254. Gyungah R.Jun, JaeyoonChung, Kathryn L.Lunetta, Jonathan L.Haines, Margaret A.Pericak-Vance, RichardMayeux, et al. Further Stratification of APOE E4-Negative Subjects Identifies Novel Genes for Alzheimer's Disease. Alzheimer's & Dementia. 2016;12(7).

255. Mead S, Mahal SP, Beck J, Campbell T, Farrall M, Fisher E, et al. Sporadic--but not variant--Creutzfeldt-Jakob disease is associated with polymorphisms upstream of PRNP exon 1. American Journal of Human Genetics. 2001 Dec; 69(6):1225-1235.

256. Delabio R, Rasmussen L, Mizumoto I, Viani G, Chen E, Villares J, et al. PSEN1 and PSEN2 gene expression in Alzheimer's disease brain: a new approach. Journal of Alzheimer’s Disease. 2014;42(3):757-760.

257. Xu W, Fang F, Ding J, Wu C. Dysregulation of Rab5-mediated endocytic pathways in Alzheimer's disease. Traffic. 2018 Apr;19(4):253-262.

258. Reitz C, Tokuhiro S, Clark LN, Conrad C, Vonsattel J, Hazrati L, et al. SORCS1 alters amyloid precursor protein processing and variants may increase Alzheimer's disease risk. Annals of Neurology. 2011 Jan;69(1):47-64.

259. Reitz C, Tosto G, Vardarajan B, Rogaeva E, Ghani M, Rogers RS, et al. Independent and epistatic effects of variants in VPS10-d receptors on Alzheimer disease risk and processing of the amyloid precursor protein (APP). Translational Psychiatry. 2013 May 14,;3:e256.

260. Kovacs GG, Murrell JR, Horvath S, Haraszti L, Majtenyi K, Molnar MJ, et al. TARDBP variation associated with frontotemporal dementia, supranuclear gaze palsy, and chorea. Movement Disorders . 2009 Sep 15,;24(12):1843-1847.

261. Correction: Rare Functional Variant in TM2D3 is Associated with Late-Onset Alzheimer's Disease. PLoS Genetics. 2016 Nov;12(11):e1006456.

262. Ruiz A, Heilmann S, Becker T, Hernández I, Wagner H, Thelen M, et al. Follow-up of loci from the International Genomics of Alzheimer's Disease Project identifies TRIP4 as a novel susceptibility gene. Translational Psychiatry. 2014 Feb 04,;4:e358.

263. Mehta SG, Khare M, Ramani R, Watts GDJ, Simon M, Osann KE, et al. Genotype-phenotype studies of VCP-associated inclusion body myopathy with Paget disease of bone and/or frontotemporal dementia. Clinical Genetics. 2013 May;83(5):422-431.

264. Levy-Lahad E, Wijsman EM, Nemens E, Anderson L, Goddard KA, Weber JL, et al. A familial Alzheimer's disease locus on chromosome 1. Science. 1995 Aug 18,;269(5226):970-973.

265. Dai M, Zheng H, Zeng L, Zhang Y. The genes associated with early-onset Alzheimer’s disease. Oncotarget. 2017 -12-15;9(19):15132-15143.

266. Ma Y, Jun GR, Zhang X, Chung J, Naj AC, Chen Y, et al. Analysis of Whole-Exome Sequencing Data for Alzheimer Disease Stratified by APOE Genotype. JAMA Neurology. 2019 Jun 10,.

267. Deming Y, Filipello F, Cignarella F, Cantoni C, Hsu S, Mikesell R, et al. The MS4A gene cluster is a key modulator of soluble TREM2 and Alzheimer's disease risk. Science Translational Medicine. 2019 08 14,;11(505).

268. Kumar-Singh S, Theuns J, Van Broeck B, Pirici D, Vennekens K, Corsmit E, et al. Mean age-of-onset of familial alzheimer disease caused by presenilin

mutations correlates with both increased Abeta42 and decreased Abeta40. Human Mutation. 2006 Jul;27(7):686-695.

269. Jakobsdottir J, van der Lee, Sven J., Bis JC, Chouraki V, Li-Kroeger D, Yamamoto S, et al. Rare Functional Variant in TM2D3 is Associated with Late-Onset Alzheimer's Disease. PLoS Genetics. 2016 Oct;12(10):e1006327.

270. Meng Y, Baldwin CT, Bowirrat A, Waraska K, Inzelberg R, Friedland RP, et al. Association of Polymorphisms in the Angiotensin-Converting Enzyme Gene with Alzheimer Disease in an Israeli Arab Community. American Journal of Human Genetics. 2006 -5;78(5):871-877.

271. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993 Aug 13,;261(5123):921-923.

272. Pontén F, Jirström K, Uhlen M. The Human Protein Atlas--a tool for pathology; The human brain in cell types. The Journal of Pathology. 2008 Dec; 216(4):387-393.

CURRICULUM VITAE