Examining the Association Between an Endolysosomal Polygenic Risk Score and Alzheimer's Disease

Samuel David Strohbehn

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2023

Committee:

Elizabeth Blue

Suman Jayadev

Program Authorized to Offer Degree:

Institute for Public Health Genetics

University of Washington

Abstract

Examining the Association Between an Endolysosomal Polygenic Risk Score and Alzheimer's Disease

Samuel David Strohbehn

Chair of the Supervisory Committee:

Elizabeth Blue

Department of Medicine

Alzheimer’s disease (AD) is a highly heritable but polygenic common neurodegenerative disorder. Recent discoveries of AD risk susceptibility loci by genome-wide association studies (GWAS) have small overall effects, and translation of these small effects into accurate personalized risk assessments is challenging. Pathway-specific polygenic risk scores (PRSs) to assess combined effects of multiple genetic variants within specific biological pathways can offer a more precise and accurate prediction of an individual's risk for that trait. We examine the association between an endolysosomal pathway-specific PRS (ePRS) and a general AD PRS with AD case-control status and AD neuropathology in a sample collected by the UW Alzheimer’s Disease Research Center. We conducted expression quantitative trait locus (eQTL)

analysis of genes contributing to our ePRS in bulk RNA sequencing from leptomeningeal tissue. ePRS was significantly associated with amyloid beta (Aβ)-related neuropathology and AD case-control status using strict cognitive testing criteria, but not significantly associated with tau-related neuropathology. A general AD PRS showed smaller effects on each AD phenotype than the ePRS. eQTL analysis of ePRS genes in leptomeningeal fibroblasts did not show an association between effect allele count and gene count beyond rs3764650 and *ABCA7,* although these results could change with a larger sample size. Together, our results suggest a pathway-specific PRS could assist in risk prediction and aid in interpretation of neuropathological changes that occur due to polygenic AD risk.

Acknowledgements

I would like to acknowledge the immense gratitude and extraordinary debt I owe to my committee chair, Dr. Elizabeth Blue, for her unending support, vision, and guidance throughout the past months. Without her, this would not have been possible.

I would like to express my sincere thanks towards my committee member, Dr. Suman Jayadev, and members of her research group, the University of Washington Alzheimer's Disease Research Center – notably Dr. Jessica Young, Dr. Corbin Johnson, and Dr. Katherine Prater. They generously provided me both with the data used in this study as well as valuable information, support, and expertise.

I am additionally grateful for the Institute for Public Health Genetics and the wonderful people that make up the program for their feedback and moral support that helped me along the way, including members of my cohort, Diane Xue, my course instructors, and my program supervisor, Annique Atwater.

Lastly, I would like to thank my friends and family for their consistent support throughout my program.

Table of Contents

Introduction

Alzheimer's disease (AD)

Alzheimer's disease (AD) is a common neurodegenerative disorder affecting the elderly, underlying 60-70% of dementia (Plassman *et al.*, 2007). AD is characterized by cognitive decline, behavioral issues, memory loss, problems with language, and mood swings (Tahami Monfared *et al.*, 2022). Autopsy of brains from patients with clinically diagnosed AD consistently shows decreased brain volume, particularly in the hippocampus, as well as abnormal clumping of two proteins: Amyloid beta (Aβ) peptides produced through the processing of transmembrane protein amyloid precursor protein (APP) by secretase enzymes accumulate into neuritic plaques (Chen *et al.*, 2017), and neurofibrillary tangles (NFTs) form through aggregation of hyperphosphorylated tau proteins within the brain (Joachim *et al.*, 1988). Onset of AD usually occurs > 65 years old, however 5% of cases have an early onset between ages 30 - 60 years (Mendez, 2019). While age-adjusted estimates of AD risk are equivalent between the sexes (Simon, 2022), prevalence of AD and related dementias at ages 65+ years varies by race and ethnicity in the United States (US): African American populations have the highest prevalence (13.8%), followed by Hispanics (12.2%), non-Hispanic Whites (NHW) (10.3%), Native American & Alaskan Natives (9.1%), and Asian & Pacific Islander populations (8.4%) (Lim *et al.*, 2023).

Public health burden of AD

AD represents a major public health burden. An estimated 50 million people are affected worldwide, with > 6 million cases in the US where it is the seventh leading cause of death (Breijyeh & Karaman, 2020; James *et al*, 2014). The estimated global financial cost of AD is $1

trillion annually, projected to reach $9.1 trillion by 2050 (Breijyeh & Karaman, 2020; Tahami Monfared, Byrnes, White *et al.*, 2022), as elder populations (ages 65+) are the fastest growing age group in the world (Jarzebski *et al.*, 2013). Despite its importance and recent advances in AD therapeutics targeting Aβ peptide aggregation (Varadharajan *et al.*, 2023; Wu *et al.*, 2023), there is still much to learn about the cause and treatment of AD (Lian *et al.,* 2023).

AD is genetically heterogeneous

Family history is a major risk factor for AD, and the disease is highly heritable (h^2 = 60–80%) (Gatz *et al.*, 2006). Rare pathogenic variants in *APP* (Goate *et al.*, 1991), *PSEN1* (Sherrington *et al.*, 1995) and *PSEN2* (Levy-Lehad *et al.*, 1995) are known to underlie autosomal dominant early-onset AD (EOAD; age < 65 years), but heritable mutations in these genes only explain ~5% of all cases of AD (Mendez, 2019). Additional risk alleles with more common allele frequencies and lower penetrance are also observed in these genes, exhibiting higher allele frequencies in AD cases than controls (Wang *et al.,* 2023; Benitez *et al.*, 2013; Cruchaga *et al.*, 2012).

APOE is the earliest locus associated with late-onset AD (LOAD; age-at-onset ≥ 65 years), a gene first associated with Aβ plaques diagnostic for AD in 1992 (Wisniewski and Frangione, 1992). Landmark studies in subsequent years linked the *APOE* ε4 allele to increased risk of AD (Strittmatter *et al.*, 1993; Corder *et al.*, 1993), and the *APOE* ε2 allele to a protective effect against AD (Corder *et al.*, 1994). *APOE* has three major isoforms (ε2/ε3/ε4) defined by two missense variants (p.Arg158Cys and p.Cys112Arg), each of which correspond to changes in protein structure and function (Mahley *et al.*, 2006). These isoforms are polymorphic, with a worldwide frequency of 8.4%, 77.9%, and 13.7%, respectively (Liu *et al.*, 2013). The effect of ε4 and ε2 alleles on AD risk has been replicated many times, confirming an increased risk for ε4

carriers that increases with two copies of the ε4 allele; and a protective effect in ε2 carriers (Bertram *et al.*, 2007; Farrer *et al.*, 1997; Choudhury *et al.*, 2021). However, the estimated effects of the ε4 and ε2 alleles on AD risk vary considerably by ancestry (Belloy *et al.*, 2023), explained in part by local ancestry effects (Blue *et al.*, 2019; Rajabli *et al.*, 2018).

The genetic basis of LOAD is still not fully understood, despite advances in sequencing methods and capability (Atri, 2019). Early AD genome-wide association studies (GWAS) had modest sample sizes and relied on genotype data from cohorts utilizing array-based imputed genotypes from reference populations (Grupe *et al.*, 2007; Bertram *et al.*, 2008; Harold *et al.*, 2009). Loci reaching genome-wide significance in these previous cohorts have only explained around 13% of SNP-based heritability (Baker *et al.*, 2023). More recent AD GWAS utilizing large samples of hundreds of thousands to >1 million participants that leverage biobank data and proxy AD cases – participants who reported at least one parent or sibling affected with AD, have brought the number of loci associated with LOAD up to 91 (Jansen *et al.*, 2019; de Rojas *et al.*, 2021; Bellenguez *et al.*, 2022; Kamboh, 2022). However, recent AD GWAS SNP discoveries tend to have low allele frequencies or effect size odds ratios (OR) between 0.90 and 1.10 (Bellenguez *et al.*, 2022; Kamboh, 2022; Zhang *et al.*, 2020; Andrews *et al.*, 2023), much more modest than *APOE* (OR = 3.32, p = 1.2e-881) (Kunkle *et al.*, 2019).

Pathway enrichment across AD loci

GWAS of complex traits like AD implicate hundreds to thousands of variants, but because each one only has a modest effect, there are limited conclusions that can be made from any individual finding (Andrews *et al.*, 2023). Different sets or combinations of risk alleles can result in the same phenotype (Silberstein *et al.*, 2021). Pathway analysis is a computational method that highlights biological pathways overrepresented in a set of input genes, aiding the

biological interpretation of statistical associations found in GWAS (Silberstein *et al.*, 2021). Pathway analysis incorporates prior biological knowledge about the input genes or loci including protein-protein interactions, metabolic pathways, and signaling pathways (White *et al.*, 2019). Pathway analysis of unique gene sets nominated by differing AD gene discovery strategies including AD GWAS, familial dementia genes, and novel AD risk variants discovered by whole-exome and whole-genome sequencing from the AD Sequencing Project (ADSP) (Beecham, *et al.*, 2017) highlighted shared pathways enriched across gene sets, while also demonstrating unique pathway enrichments from the ADSP that can assist in elucidating novel findings about the underlying biology of AD by providing additional target pathways for functional study and generating hypotheses about how such pathways might relate to AD pathogenesis (Xue *et al.*, 2021). Pathway analysis of genes implicated by AD GWAS have highlighted several key areas, such as immunity and inflammation, cholesterol metabolism, Aβ clearance, and endocytosis (Harrison *et al.*, 2020).

The endolysosomal pathway includes various membranous compartments within cells, such as early endosomes, recycling endosomes, late endosomes, and the lysosome. This system plays a pivotal role in the cyclic degradation and recycling of cellular components (Pillay *et al.*, 2002; Hu *et al.*, 2015). The endolysosomal system encapsulates several crucial biological factors associated with AD risk (Szabo *et al.*, 2022). Intracellular Aβ collects years before Aβ is deposited in the brain (Nixon, 2017). Proteins encoded by *APP* and *APOE*, two genes with large established effects on AD risk, are both altered by the endolysosomal pathway. Early endosomes have previously been shown to be the site of APP processing by cleaving enzyme BACE1, and the dissociation of the ApoE protein from the LDLR1 receptor (Vassar *et al.*, 1999). Importantly, the ApoE-LDLR1 receptor-complex influences the level of endocytosis and degradation of APP

(Ulery *et al.*, 2000; Kang *et al.*, 2000). This demonstrated connection between known AD genes and AD risk through the endolysosomal pathway motivates further analysis focused on this pathway to better understand AD risk and pathogenesis.

Genetic risk scores for AD

It has been challenging to translate the genes and variants implicated by AD GWAS into effective personalized risk assessment. Genetic risk scores (GRS) allow individual-level assessment of disease risk dependent on estimating cumulative risk of a given trait assessed in a well-matched and independent sample (Wray *et al.*, 2007), estimating the cumulative disease risk from independent risk loci. GRS are often extended to hundreds or thousands of loci with small effects on risk for complex, generating polygenic risk scores (PRS) to improve risk prediction (Igo *et al.*, 2019).

A GRS based upon 20 single nucleotide polymorphisms (SNPs) significantly associated with LOAD from the Lambert *et al.* (2013) GWAS was found to be significantly associated with risk of AD in the National Institute on Aging–Late-Onset Alzheimer's Disease Family Study (NIA-LOAD) (OR = 1.29; 95% CI: 1.21–1.37; $p < 0.001$) (Tosto *et al.* 2017). A subset of unrelated non-Hispanic white (NHW) samples within the NIA-LOAD study were represented in the original Lambert *et al.* (2013) GWAS. To replicate this result, they also found a significant association in an independent sample of Caribbean Hispanic families from the Estudio Familiar de Influencia Genetica en Alzheimer (EFIGA) cohort (OR = 1.73; 95% CI: 1.57–1.93; $p < 0.001$) (Tosto *et al.* 2017). A slightly different GRS based on 21 SNPs from Lambert *et al.* (2013) was tested using one sample per NIA-LOAD family as well as NHW AD cohorts including a sporadic early-onset AD (sEOAD) cohort – Knight-Alzheimer's Disease Research Center (Knight-ADRC) – and two sporadic late-onset AD (sLOAD) cohorts – the Dominantly Inherited

Alzheimer Network (DIAN), and the Alzheimer's Disease Neuroimaging Initiative (ADNI). They found significant association with sEOAD, (OR = 2.27; p = 1.29e-7), fLOAD (OR = 1.75; p = 1.12e-7) and sLOAD (OR = 1.40; p = 1.21e-3), revealing potential shared genetic architecture between the early- and late-onset forms of AD (Cruchaga *et al.*, 2018). Expanding the original 20 SNP GRS to include all (87,605) SNPs with $p < 0.5$ from Lambert *et al.* (2013) and ε4 and ε2 significantly improved prediction accuracy, as the area under the curve (AUC) increased from 0.715 to 0.745 (p = 1.3e-11) (Escott-Price *et al.*, 2015). However, PRS including SNPs with $p < 0.5$ from more recent GWAS including biobank data and proxy-AD cases (Jansen *et al.*, 2019) had significantly decreased prediction accuracy (AUC = 0.733; 95% CI: 0.706–0.758; p = 1.2e-45) compared to SNPs reaching an optimal significance threshold of $p < 1.69e-5$ (AUC = 0.827; 95% CI: 0.805–0.849; p = 9.71e-70). This result was replicated using three different AD GWAS with contrasting p-value thresholds (Zhang *et al.*, 2020). Conflicting results from Escott-Price *et al.* (2015) and Zhang *et al.* (2020) may be explained by the complex genetic structure for AD impacting multiple contributory biological pathways. However, including common low-effect loci below a certain significance threshold might not offer any additional risk prediction accuracy for PRSs.

Focusing on the combined effects of multiple genetic variants within specific biological pathways might better predict a specific AD-related biomarker or cognitive component. Pathway-specific PRSs have been previously shown to improve accuracy of risk prediction for Parkinson's disease, AD, coronary artery disease, and obesity (Choi *et al.*, 2023). For example, despite having fewer risk variants, a revised PRS including only genes with high correlations with known Aβ and glucose metabolism AD biomarkers outperformed a more typical PRS in AD prediction accuracy (Lee *et al.*, 2022). Small pathway-specific GRS built on GWAS SNPs

from Lambert *et al.* (2013) annotated as falling within the cholesterol metabolism (*PICALM, CLU, CR1,* and *APOE*) and Aβ (*ABCA7, CLU,* and *APOE*) pathways using Gene Ontology (GO; Ashburner *et al.*, 2000) showed significant correlation with several cognitive factors and cerebrospinal fluid (CSF) Aβ and had a stronger effect size than the overall PRS. However, these PRSs were not found to be better predictors of these outcomes than an *APOE*-focused risk score measuring ε2, ε3, and ε4 alleles alone (Darst *et al.*, 2017). Because both GRS included *APOE,* it is not clear what value the other pathway-specific SNPs bring in additional risk prediction.

Studies of AD PRS which did not include *APOE* found a stronger PRS effect in ε4 carriers than in non-ε4 carriers, suggestive of an effect of *APOE* on polygenic AD risk (Riaz *et al.*, 2021; Stocker *et al.*, 2023). Additional evidence for an *APOE* effect on AD GRS performance was found by Xu *et al.* (2023), who compared six pathway-specific GRS built on the AD GWAS by Kunkle *et al.* (2019). None of the pathway-specific GRSs outperformed an *APOE* GRS regarding prediction of cognition or CSF Aβ and tau, however the Aβ (*CLU, SORL1, ABCA7, PICALM, ADAM10*) and cholesterol metabolism (*CLU, SORL1, ABCA7*) GRS could predict changes in Aβ nearly 15 years earlier than the overall GRS, and each GRS evaluated showed an interaction with *APOE* and age. This suggests that pathway-specific GRS are useful for better understanding AD.

Thesis Statement

We propose a pathway-specific GRS focused on the endolysosomal pathway (ePRS), given its prominent role in protein trafficking and degradation of AD-related proteins (Nixon, 2017; Ulery *et al.*, 2000; Kang *et al.*, 2000). We contrast the association between AD-related phenotypes and a general GRS versus our ePRS based on the subset of SNPs implicating genes in the endolysosomal pathway. Given the endolysosomal pathway is crucial to the processing

and degradation of proteins such as Aβ and tau, we explore whether our GRS values are correlated with AD neuropathology and contrast their performance. This association could provide insight into both the link between genetic predisposition to endolysosomal dysfunction and the development of AD-related neuropathological changes, as well as differentiation between endolysosomal dysfunction and overall risk of developing AD.

Our ePRS definition includes more SNPs than previous endolysosomal-related PRS and excludes *APOE* to prevent overpowering of small effects from other SNPs. Previous endolysosomal-related GRS include *APOE*, causing them to resemble *APOE* in risk prediction (Darst *et al.*, 2017, Xu *et al.*, 2023). Given the evidence of potential interaction of *APOE* with AD PRS (Xu *et al.,* 2023, Riaz *et al.*, 2021; Stocker *et al.*, 2023), we examine potential interaction between each risk score and *APOE*, and include an interaction effect in testing where indicated. Additionally, given the strong effect of *APOE* on AD phenotypes, our tests adjust for ε4 carrier status to avoid confounding by confining comparisons to those with the same ε4 carrier status. After controlling for *APOE* status, we hypothesize the odds of AD to be significantly higher in those with high ePRS and general GRS than those with low and intermediate scores.

Our GRS are built on GWAS signals driven by noncoding or intergenic SNPs. We intend to explore the gene regulatory consequences of the ePRS with pilot analysis of transcriptomes from participants representing the tails of the ePRS distribution, where we would expect to see differences in endolysosomal pathway gene expression between high versus low ePRS. As a first step, we test differences in expression of ePRS genes in leptomeningeal fibroblasts among samples with differing effect SNP counts. We use this cell type because its endolysosomal phenotype is similar to that observed in neurons, but samples are much easier to collect and maintain. A lack of significant results does not conclude that the tag SNP-target gene link is not

robust, as SNP effects on gene expression can be cell-type specific. We contrast the expression of genes linked to our ePRS in leptomeningeal fibroblast tissue in our sample to GTEx cultured skin fibroblasts that were not ascertained for AD to provide context through a comparison tissue type, and inform future work comparing expression quantitative trait loci (eQTLs) by cell type.

Methods

Sample Ascertainment

Autopsied brain samples from 247 donors were collected by the University of Washington (UW) Alzheimer's Disease Research Center (ADRC) and Seattle Longitudinal Study (SLS). Most donors participated in the Adult Changes in Thought Study (ACT; n = 160), a longitudinal prospective cohort study started in 1994 by the Health Research Institute of Kaiser Permanente Washington in which cognitively normal participants aged 65+ years are randomly selected from members of Group Health in Seattle, Washington (ACT, 2023). Incident cases of dementia are identified via cognitive testing every two years; all participants who scored cognitively normal initially who then show reductions in scores below a threshold are then evaluated with comprehensive neurological and neuropsychological assessment. Sixty-six samples were collected by the UW ADRC clinical core, part of a network funded by the National Institute on Aging that studies the early transition from normal function to mild cognitive impairment and early stages of dementia (UW-ADRC, 2023). Eighteen samples were ascertained by the SLS, a UW Medicine Department of Psychiatry and Behavioral Sciences longitudinal study that began in 1956 that examines cognitive and psychosocial changes in adulthood through study of multiple both cohorts (SLS, 2023). Two additional samples came from the Pacific Udall Center, a research center that conducts studies relating to causes of Parkinson's disease and

subsequent treatment of patients with Parkinson's disease and related neurodegenerative disorders (Pacific Udall Center, 2023), and the final sample represents a donor to the Seattle Alzheimer's Disease Brain Cell Atlas Consortium (SEA-AD, 2023).

Phenotyping

The UW-ADRC provided a consensus clinical diagnosis for each sample based on neuropathology data, imaging, and medical history (UW-ADRC, 2013). Each sample was labeled with one or more clinical diagnoses including AD, possible/probable AD, Parkinson's disease, Lewy body disease, vascular dementia, Huntington's disease, frontotemporal lobar degeneration, multiple system atrophy, control, or other. We define AD cases as those samples with a consensus clinical diagnosis of AD or possible/probable AD, including those for which AD was one of multiple diagnoses. Controls included all samples with a consensus clinical diagnosis of "control." All samples with exclusively non-AD dementia diagnoses were excluded from our analyses.

Neuropathology data for each sample was provided by the UW Precision Neuropathology Core within the UW-ADRC (UW-ADRC, 2023). Neuropathology scores were provided following the protocol from the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (Mirra *et al.*, 1991). CERAD scores evaluate neuritic plaques in the neocortex, hippocampus, and amygdala, where severity of the pathology is graded on a four-point scale from 0-3, representing no AD, possible AD, probable AD, and definite AD, respectively. Aβ pathology was scored from 0 to 5 using Thal amyloid phasing criteria, defined by progressive deposition of immunopositive Aβ in five locations: none = 0, the neocortex = 1, allocortex / limbic system = 2, diencephalon/basal ganglia = 3, brainstem/midbrain = 4, and cerebellum = 5 (Thal *et al.*, 2002) and cerebral amyloid angiopathy (CAA), defined by Aβ accumulation in small

to medium-sized cerebral blood vessels, classified as definite = 3, probable = 2, possible = 1, and none = 0 (Biffi & Greenberg, 2011). Braak staging was performed for each sample, in which progression of NFTs within an autopsied brain is assessed (Braak & Braak, 1991), beginning in the transentorhinal cortex (stage I), then progressing to the hippocampus (stage II), temporal cortex (stage III), other regions of cerebral cortex (stage IV), visual association cortex (stage V), and primary visual cortex (stage VI). The UW-ADRC provides an AD-related neuropathologic change (ADNC) score which combines CERAD, Thal, and Braak scores to define overall neuropathological change as none = 0, low = 1, intermediate = 2, or high = 3, for which intermediate or high is considered sufficient as explanatory of one's dementia (Montine *et al.*, 2012).

Selecting variants for genetic risk scores

We created a uniform score to measure the evidence connecting 29 SNPs previously associated with AD risk to its published candidate gene. We focus our review on AD GWAS with SNPs of higher frequency and stronger effects (Lambert *et al.* 2013, Sleegers *et al.*, 2015, Cauwenberghe *et al.*, 2016) rather than common, low-effect SNPs discovered by more recent GWAS (Jansen *et al.*, 2019; de Rojas *et al.*, 2021; Wightman *et al.*, 2021; Bellenguez *et al.*, 2022) given previous evidence that they do not significantly improve PRS performance. Each SNP was annotated to recognize potential effects on coding sequence and gene regulation. We summarized this evidence linking SNPs to their gene target with the score *S* where, for each gene, *a* is equal to 1 if ≥ 1 SNP on the same haplotype as the GWAS SNP is predicted to alter the candidate protein structure (ex., missense) and 0 otherwise; N_{SNPs} is equal to the number of haplotype SNPs; N_{eQTLs} is equal to the number of eQTLs; $N_{regulatory}$ is equal to the number of haplotype SNPs falling within promoter, enhancer, or DNase I hypersensitivity (DHS) regions

according to HaploReg v4.0 (Ward & Kellis, 2012); and N_{bonus} is equal to the number of haplotype SNPs falling within a promoter, enhancer, or transcription factor binding site (TFBS) according to Variant Effect Predictor v91 (VEP; Mclaren *et al.*, 2016).

$$S = 5\alpha + \frac{2N_{eQTL}}{N_{SNPs}} + \frac{N_{Regulatory}}{N_{SNPs}} + \frac{N_{Bonus}}{N_{SNPs}}$$

Each gene with $S > 1$ was assigned to three distinct AD-related biological pathways based on synthesis of GO terms provided by the Monarch Initiative (Mungall *et al.*, 2017): "endolysosomal" genes were represented in pathways related to endocytosis and cholesterol metabolism.

Genotyping

DNA was extracted from cerebellar tissue of 324 samples provided by the UW Precision Neuropathology Core. Samples with consensus clinical AD diagnoses, AD-neuropathology scores, *APOE* genotype, sex, and age at onset of cognitive symptoms (AAO) for cases or age at last cognitive evaluation (AAE) for controls were included in analysis. Early in the study, the UW Center for Precision Diagnostics performed targeted sequencing to genotype the 24 SNPs from previously published AD GWAS (Lambert *et al.*, 2013; Sleegers *et al.*, 2015; Cauwenberghe *et al.*, 2016) (n = 229). Later, the Northwest Genome Center genotyped samples using the Illumina Infinium Global Diversity Array 8v1.0 (n = 95) and provided the data in PLINK file format (Purcell *et al.*, 2009). Three of the Lambert *et al.* (2013) SNPs were not on the array (rs670139 near *MS4A4E*, rs4147929 near *ABCA7*, and rs8093731 near *DSG2*); we therefore identified proxy-SNPs in the array in linkage disequilibrium (LD; $r^2 > 0.8$) in the 1000 Genomes European reference (The 1000 Genomes Project Consortium, 2015) using HaploReg

and LDproxy (Ward & Kellis, 2012; Machiela & Chanock, 2015). Across the sample, the genotyping rate was high: 1.0 for 23/34 SNPs, while rs10838725 had a genotyping rate of 0.8947; the 10 samples with missing genotypes were excluded from analysis.

Genetic Risk Scores

We define GRS using the explained variance weighted GRS (EV-GRS) approach (Che & Motsinger-Reif, 2012). Bayesian modeling allows prior knowledge to make refined iterative predictions in new data. In the case of the EV-GRS, this prior knowledge includes allele frequencies (AFs) and allele-specific ORs and is used to model variance in the trait explained by these variants and subsequent disease risk. This additional information improved the predictive ability of EV-GRS compared to other methods (Che & Motsinger-Reif, 2012). Each SNP contributing to the GRS is weighted by w_{EVi}, calculated with the following formula, where for each SNP *i*: OR = odds ratio in an appropriate reference population and MAF is the minor allele frequency in the appropriate reference population.

$$w_{EV_i} = log(OR_i)\sqrt{2MAF_i(1 - MAF_i)}$$

For each sample, the EV-GRS can then be calculated using the following formula where, for each SNP *i*: *I* is the total number of SNPs in the score, w_{EVi} is the SNP weight as defined above, and *G* is the number of risk alleles observed, defined as the allele associated with a higher risk of a given trait:

$$EV\text{-}GRS = \sum_{i=1}^{I} w_{ev_i} G_i \left(\sum_{i=1}^{I} 1\right) / \left(\sum_{i=1}^{I} w_{ev_i}\right)$$

Measuring Gene Expression

The UW ADRC provided expression data from individuals who were *APOE* ε3 homozygotes ascertained from the high (n = 13) and low (n = 13) tails of the ePRS distribution, based on sample availability over time. Expression data was collected from bulk RNA sequencing (RNAseq) of leptomeningeal fibroblasts derived from cultured primary tissue taken at autopsy. RNAseq was performed by Genewiz from Azenta Life Sciences. The generated dataset consisted of raw counts of each genomic feature present in the dataset. Python packages Pandas and NumPy were used to analyze raw counts data of all reads (McKinney *et al.*, 2010; Harris *et al.*, 2020).

RNA sample and sequencing quality was measured using several metrics: *Assigned*, the number of reads mapped to a feature; *Unassigned Unmapped*, the number of reads reported as unmapped in SAM/BAM input; *Unassigned MultiMapping*, the number of reads that mapped to more than one location in the genome; *Unassigned No Features*, the number of reads not overlapping with any features included in annotation; *Unassigned Ambiguity*, the number of reads overlapping with two or more annotation features; *Imputed RIN*, the sample's RNA integrity score; *PMI Hours*, the Post-Mortem Interval (PMI) defined as the number of hours between death and sample collection; *megabase yield*, the number of Megabases sequenced; *Mean Quality Score*, $-10*\log_{10}$(estimated probability of base call being wrong); and *Q30*, the percentage of bases with a Quality Score of 30 or higher (Liao Y, Smyth GK, Shi W, 2019).

Statistical analyses

All statistical analyses were performed using R Statistical Software v4.1.2 (R Core Team, 2021). All figures were created using R package ggplot2 (Wickham, 2016). Significance was defined at the $p < 0.05$ level for each statistical test.

GRS contrasts

Differences in mean GRS between AD cases and controls and between genotyping platforms were tested using Welch Two Sample t-test (Welch, 1947). Given the common frequency of the ε4 allele and its large impact on AD risk, we performed analyses either stratified by or adjusting for ε4 carrier status. Our relatively modest sample size prohibits further stratification by ε2/ε3/ε4 genotypes. Stratified analyses allow those with and without ε4 to have different distributions of case/control status across risk scores. We created scatterplots of each AD-related phenotype and each GRS, stratified by ε4 carrier status, and estimated the correlation between each GRS and the outcome. The association between each GRS and AD case/control status was evaluated by logistic regression, adjusting for age, sex, and ε4 carrier status given their large effects on AD risk (Bellou *et al.*, 2020; Buckley *et al.*, 2023) and AD-related neuropathology (Walker *et al.*, 2022; Oveisgharan *et al.*, 2018; Luchsinger *et al.*, 2022). Similarly, we tested the association between each neuropathology score and each GRS using linear regression, adjusted for age, sex, and ε4 carrier status. If evidence of interaction between ε4 carrier status was observed in correlation plots, we also tested for an interaction between each GRS and ε4 carrier status.

Gene expression analyses

We first $\log_2$ transformed gene counts to avoid confounding by sequencing depth and compress variance between samples. We performed principal component analysis (PCA) of these counts to define the main drivers of variance in the dataset. Pearson correlation coefficients were calculated between principal components (PCs) that captured the majority of the variance within the dataset and covariates describing each sample and the respective data generation process.

Analysis of variance testing (ANOVA) was performed between the number of effect alleles and gene counts for each of the primary genes linked to the SNPs included in our GRS, adjusting for age, PC1, mean sequencing quality score, and PMI. We adjusted for PC1 as it was significantly correlated with multiple quality variables (*Assigned*, *Unassigned No Features*, U*nassigned Ambiguity*, *Read Number*, and *Megabase Yield)*, and adjusted for mean sequencing quality score and PMI as these technical variables were significantly correlated with PC2 and PC3, respectively. Sex was not included as a covariate as the sex ratio was balanced in the high and low ePRS groups. *APOE* ε4 carrier status was not included in this test, as each sample was homozygous for the ε3 allele. Box plots were created for each of these contrasts stratified by allele count to aid interpretation.

Regression analysis between $\log_2$ normalized gene counts and tag SNP risk allele count was performed, adjusted for AAO or AAE, PC1, mean sequencing quality score, and PMI. Regression coefficients for tag SNP risk allele count were compared to the normalized effect size (NES) in GTEx cultured fibroblasts, representing the slope of the linear regression as the effect of the alternative allele (ALT) relative to the reference allele (REF) to measure the eQTL effect of each SNP.

Results

Sample description

Of the 324 samples with DNA, 247 had neuropathology data. This sample was restricted to only include clinical consensus AD cases and controls, resulting in 188 samples used in our study (**Table 1**). Our sample was fairly balanced with respect to AD status, AAO vs. AAE, and reported race, but overrepresents females and ε4, as is typical of AD studies. Nearly 3/4 of our

AD cases are female, and 44% carry at least one ε4 allele. We see similar enrichment of the protective ε2 allele in controls (14%) vs. AD cases (10%). Surprisingly, nearly one in five of our cases had EOAD (n = 17/88, 19%).

The distribution of AD-related neuropathy in our sample is summarized in **Table 2**. With the exception of CAA, most clinically-diagnosed AD cases had high AD-related neuropathology measures: 97% of AD cases had Thal stage of 3 or higher; 95% had Braak stage of 4 or higher; 88% had a CERAD evaluation of probable or definite AD; and 97% had overall ADNC score of intermediate or high. Among clinically-diagnosed controls, there was less correlation: around half of controls had expected low AD-related neuropathology measures, but the other half showed high AD-related neuropathology measures, defined as Thal stage ≥ 3, CAA score ≥ 2, Braak stage ≥ 4, CERAD score ≥ 2, and ADNC score ≥ 2.

Deriving Genetic Risk Scores

We contrast our general AD GRS with our ePRS to address the hypothesis that focusing on a specific pathway will improve the predictive value of a GRS. Our general AD GRS was created containing SNPs reaching genome-wide significance after stages 1 and 2 in Lambert *et al.* (2013), summarized in **Table 3** and **Supplemental Table 1**. Three of these SNPs were not included on the genotyping array used for 95 samples and were replaced by proxy SNPs, modifying their GRS weight by multiplying it by the r^2 between the desired and proxySNP: rs670139/rs600550, $r^2 = 0.87$; rs8093731/rs139544178, $r^2 = 1.0$; rs4147929/ rs3752246, $r^2 = 0.88$ (Chou *et al.*, 2016). Our annotation-based scores linking GWAS SNPs to target genes resulted in 14 SNPs linked to 12 genes within an endolysosomal (endocytosis & cholesterol metabolism) pathway (**Supplemental Table 2**). We calculated each sample's ePRS using the endolysososmal pathway SNPs as specified in **Table 4**. The distribution of both the ePRS and AD GRS did not

differ significantly between groups defined by AD status, ε4 carrier status, or genotyping strategy ($p < 0.05$; **Figure 1, Supplementary Figure 1**). Not surprising given they are based on shared SNPs, the ePRS and general AD GRS were highly correlated ($r^2 = 0.84$, $p = 2e\text{-}16$).

Association between GRS and AD-related phenotypes

Stratified correlation plots between each GRS and AD-related outcome provide a broad summary of their relationship to disease progression and neuropathology (**Figure 2**). These stratified correlation plots provide evidence of potential interaction between ε4 carrier status and consensus clinical AD diagnosis, CERAD, and CAA, while presence of the ε4 allele shifts the sample toward higher ADNC score, Thal stage, and less so for Braak stage. ePRS and the general AD GRS were highly correlated in our sample ($r^2 = 0.84$, $p = 2e\text{-}16$). The evidence of association (both effect size and p-value) was stronger for the ePRS than the general AD GRS for all associations tested. Neuropathology measures focused on Aβ have nominally significant ($p < 0.05$) evidence of association with ePRS and AD GRS, while the association with ADNC is suggestive ($p < 0.1$) and the association with CAA is suggestive only in the interaction model for the ePRS **(Table 5).** AAO/AAE was significantly negatively associated with every AD-related phenotype other than CAA, sex was not significantly associated with any AD-related phenotype, while ε4 carrier status was significantly positively associated with every AD-related outcome (**Table 5**).

A portion of controls had their last cognitive exam [Mini-Mental State Examination (MMSE), Cognitive Abilities Screening Instrument (CASI), or Montreal Cognitive Assessment (MoCA)] more than 2 years prior to their death. We repeated our correlation and association analyses restricted to only controls with cognitive tests within the last two years, in line with the cognitive testing timeline for ACT. This dropped the number of controls from 100 to 52 and

reduced the number of controls with probable (27 to 17) or definite (20 to 5) AD as defined by CERAD. Despite a reduction in large reduction in sample size, the association between ePRS and AD case-control status strengthened dramatically, particularly among non-ε4 carriers, while the association between our general AD GRS and AD status was minimally impacted (**Figure 2**).

Relationship between ePRS and Gene Expression

Bulk RNAseq data was collected from cultured leptomeningeal fibroblasts representing 26 donors with the 13 highest and 13 lowest ePRS. These high and low ePRS groups were balanced with respect to age, cognitive status, and sex (**Table 6**). Among high ePRS samples, the ePRS ranged from 14.56 to 19.59, and among low ePRS samples, ePRS ranged from 3.12 to 9.38 (**Supplementary Figure 2**). There were no significant differences between high and low ePRS groups in RNA, sequencing, and sample quality metrics (**Supplementary Table 3**).

We found that PC1, PC2, and PC3 explained 26.5%, 10.7% and 7.4% of total variance in the RNAseq data, while all other PCs explained < 5% of the variance. PC1 was significantly negatively correlated with technical sequencing quality variables *Assigned*, *Unassigned No Features*, *Unassigned Ambiguity*, *Read Number*, and *Yield megabases,* while mean RNA quality score and PMI hours were highly correlated with PC2 and PC3, respectively High versus low ePRS status was not shown to drive significant variance in any PCs.

We evaluated the relationship between risk allele count at each ePRS SNP and its target gene's $\log_2$-transformed gene count **(Table 7)** using box plots and ANOVA adjusting for PC1, RIN, and PMI. We did not do ANOVA testing for *HLA-DRB5*, *EPHA1*, *MS4A4E*, *SORL1*, and *CD33* due to a low overall read count for these genes in this tissue type. Among ePRS SNPs, only rs3764650 was associated with the expression of its target gene, *ABCA7* ($F = 17.14$, $p = 6.1e\text{-}4$) (**Figure 3**).

Limited evidence of association between ePRS SNPs and target gene expression could be explained by poor statistical power due to low levels of gene expression in this cell type. We explored this possibility by comparing our observed transcript-per-million (TPM) counts with those for fibroblasts in the Genotype-Tissue Expression Project (GTEx) data (**Table 8**). In most ePRS genes, our leptomeningeal RNAseq TPM was higher than the cultured fibroblast TPM from GTEx, while *CLU* TPM was two orders of magnitude higher in our leptomeningeal fibroblasts than the GTEx cultured fibroblasts. These results suggest that leptomeningeal fibroblasts are not directly comparable to skin fibroblasts, as the expression of *BIN1, MEF2C, CLU,* and *ABCA7* were strikingly different between the tissue types.

Discussion

In our sample, a relatively high proportion of controls meeting clinical consensus criteria had high AD-related neuropathology scores (**Table 2**): 59% of controls had a Thal stage ≥ 3, 33% had intermediate or high CAA, 76% had a Braak stage ≥ 4, 47% had a CERAD score ≥ 2, and 58% had an ADNC score ≥ 2. While cognitively intact individuals with high neuropathological burden are not unique (Pike *et al.*, 2007), our sample shows a larger percentage of unaffected controls with high AD-related neuropathology than has previously been published (Knopman *et al.*, 2003). Consensus clinical diagnosis made based on cognitive, physical, and neurological evaluations while participants are still alive do not necessarily reflect the cognitive status of the participant at death depending on the time elapsed since cognitive testing. Long intervals between clinical evaluation and death increase the opportunity for controls to develop AD prior to death, which could explain a portion of the high neuropathological burden in half of our sample controls. Our follow-up testing restricted to

controls with interval < 2 years resembles cognitive testing intervals for the ACT, for which participants are screened regularly to maintain up-to-date cognitive evaluations (ACT, 2023). Restricting controls to those with a recent cognitive exam decreased our sample size from 100 to 52 but strengthened the association between AD status and both ePRS (OR = 1.07, $p = 0.23$ vs. OR = 1.21, $p = 0.039$) and general AD GRS (OR = 1.01, $p = 0.43$ vs. OR = 1.03, $p = 0.27$). This suggests that a portion of controls converted to cases between the time of their consensus clinical diagnosis and death.

Across measures of AD neuropathology, ePRS showed the strongest association with measures of Aβ and amyloid plaques. Significant associations were observed in CERAD (β = 0.085, $p = 4.1e-3$) and Thal (β = 0.076, $p = 0.026$), which were followed by nearly significant association with ADNC (β = 0.042, $p = 0.061$) and weak associations with CAA (β = 0.041, $p = 0.16$) and Braak (β = 0.036, $p = .20$). Increases in both ePRS and general AD GRS among non-ε4 carriers corresponded to increases in CERAD score; among ε4 carriers, there was no observed positive correlation. As ADNC incorporates Thal, CERAD, and Braak scores, positive associations between GRS and ADNC were likely driven by Thal and CERAD, as both of our GRS were only weakly associated with Braak. Taken together, these results suggest that our GRSs, and specifically the ePRS, capture information about endolysosomal function tied to Aβ degradation and clearance necessary to prevent AD neuropathology. This is supported by previous studies showing that Aβ peptide degradation occurs in lysosomes, and that progressive dysfunction of endososomes is correlated with Aβ accumulation in AD and AD models (Nixon, 2013; Nixon, 2007, Yu *et al.*, 2005, Takahashi *et al.*, 2002). Given previous literature demonstrating that tau accumulates in lysosomes, and that increases in tau pathology can be observed with dysregulated endolysosome function (Distl *et al.*, 2003; Bi *et al.*, 2007; Liao *et al.*,

2007; Yan & Zheng, 2021), insignificant association between both GRS and Braak is surprising, although these association signals may strengthen with a larger sample size.

Associations were stronger with ePRS than general AD GRS across all tests performed: Thal stage (β = 0.077 vs 0.054), CERAD (β = 0.085 vs 0.052), CAA (β = 0.041 vs 0.029), ADNC (β = 0.042 vs 0.024), and Braak stage (β = 0.036 vs 0.018). Our results suggest that AD GWAS SNPs with modest effect sizes can offer useful insights into AD risk beyond ε4 carrier status when compiled into a risk score informed by biological pathways. Additionally, these results suggest that a pathway-specific GRS capturing the endolysosomal pathway might be more suited to predicting Aβ-related neuropathology than an overall GRS or previous pathways focusing on cholesterol or Aβ clearance alone (Darst *et al.*, 2017; Xu *et al.*, 2023).

Both the ePRS and general AD GRS showed evidence of an interaction effect with ε4 carrier status on AD status, CAA, and CERAD in stratified correlation plots (**Figure 2**). While the interaction terms were not found to be significant, they strengthened each association between GRS and AD-related outcomes with evidence for interaction with ε4 (**Table 5**). Among non-ε4 carriers, those with higher GRS were more likely to be affected by AD or exhibit Aβ-related neuropathy, whereas this association was attenuated or inverted among ε4 carriers, of whom only 33% were controls. In our modest sample, this suggests that the presence of an ε4 allele might overpower a potential effect from high ePRS and general AD GRS. The effect of presence of an ε4 allele on neuritic plaque density measured by CERAD overwhelmed the GRS – only 20% of samples with an ε4 allele had a CERAD score of 0 or 1, which limited potential association that can be observed in this subset of samples. CAA showed a similar effect: among non-ε4 carriers, those with higher GRS were more likely to have high CAA, whereas among ε4 carriers, those with higher ePRS and general AD GRS were not more likely to have high CAA

score, as only 36% of ε4 carriers having a CAA score of 0 or 1. This contrasts with previous findings that general AD PRS show stronger associations in ε4 carriers than in non-ε4 carriers (Riaz *et al.*, 2021; Stocker *et al.*, 2023). Associations with Thal, ADNC, and Braak did not seem to change based on ε4 carrier status. Future studies are needed to investigate the relationship between polygenic risk of AD and a potential interaction with *APOE* genotype in AD-related phenotypes.

Most SNPs contributing to the ePRS are non-coding, thus we expect that they influence AD risk through gene regulation (Zhang & Lupski, 2015). Several endolysosomal genes including *MEF2C*, *CD2AP*, *EPHA1*, *MS4A4E*, *SORL1*, and *ABCA7* showed observable differences in mean gene count by number of effect alleles in our leptomeningeal RNAseq data, but only *ABCA7* showed significant differences. Our results also suggest that cultured skin fibroblasts do not offer an effective proxy for leptomeningeal fibroblasts when examining eQTLs, which is supported by previous literature showing that eQTLs are cell-type specific (Donovan *et al.*, 2020). While GTEx cultured fibroblasts showed that ePRS SNPs rs4147929, rs9271192, rs10948363, rs28834970, rs9331896, and rs10792832 were significant eQTLs for the target gene we identified, this was not observed in our leptomeningeal fibroblasts. Further comparisons should be made with a larger sample size, given the high variance of gene counts in expression datasets. In future RNAseq analyses, it would be beneficial to analyze the ePRS using single cell RNAseq in microglia, neurons, and astrocytes involved in cellular waste removal of protein aggregates observed in AD cases (Colonna & Butovsky, 2021). Pathway analyses of genes differentially expressed in cells from those with high vs low ePRS would allow us to examine shared mechanisms potentially influencing the relationship between the endolysososmal pathway and AD pathogenesis (Thomas *et al.*, 2021).

The current study has limitations. The accuracy of our results is based on the classification of genes and pathways at the time of the ePRS creation. Recent non-proxy AD GWAS (Kunkle *et al.*, 2019) identified more genome-wide significant associations with AD that are not represented in our ePRS or general AD GRS, given the date of targeted genotyping of our samples. Second, statistical testing only considered the effect of *APOE* through ε4 carrier status, while other *APOE* variants are associated with AD risk (Belloy *et al.*, 2023); grouping *APOE* genotypes into just two groups obscured potentially useful information. Third, the SNP array used to genotype 95 samples did not contain all SNPs used in the ePRS and general AD GRS, and proxy SNP genotypes were not guaranteed to be accurate. Their SNP weights did not reflect the ORs listed in GWAS, which might have caused incorrect ePRS calculations. Lastly, the sample size of our study, both in AD-phenotype and expression analyses, is relatively small and of NHW ethnicity, which limits the generalizability of our results, especially without a replication analysis. The present findings and limitations motivate future research examining the association between ePRS and AD-related outcomes with a larger sample size, as well as comparisons to other pathway-specific GRS, an overall GRS, and to an updated ePRS including small-effect SNPs identified by recent larger-scale GWAS. A larger study would improve the power to detect small differences or correlations in association tests, and potentially increase the external validity of the results. Additionally, sample genotyping and ePRS scoring done by whole-genome sequencing or genome imputation to evaluate evidence for association at genetic markers not directly sequenced, would allow for more confidence in SNP genotyping than array data using proxy SNPs provides (Li *et al.*, 2010). Further, correlations between ePRS and laboratory measures of endolysosome function would allow for investigation to further validate

the weightings assigned to each SNP and the ePRS score as a representation of one's endolysosomal function.

In conclusion, our GRSs are significantly associated with AD status when controls were restricted to recent cognitive testing, and with Aβ-related neuropathology. Our GRS demonstrated evidence of interaction with ε4 carrier status for AD status and Aβ-related neuropathology measures CERAD and CAA. While ε4 carrier status showed higher association with consensus clinical AD diagnosis than either GRS, this evidence of interaction suggests that our GRSs contain useful information regarding risk of AD phenotypes. The ePRS showed stronger associations with AD status and AD-related neuropathology than an overall AD GRS that included 10 additional SNPs not linked to genes in the endolysosomal pathway. If further studies show similar associations with a larger and more diverse sample, the ePRS could be validated as an effective tool in identifying elevated genetic risk of AD and age-related neuropathology like accumulation of Aβ for each *APOE* genotype (Kang *et al.*, 2000). The additional risk prediction that our pathway-specific GRS provides could be useful, especially among non-ε4 carriers, and might allow a greater understanding of neuropathological changes that predicate AD. More accurate prediction of AD precursors could be beneficial for early intervention to slow the progressive neurodegenerative process in AD (Viola & Klein, 2015).

References

"Adult Changes in Thought (ACT) Study." *Kaiser Permanente Health Research Institute* (2023).

Altmann, A *et al*. "A comprehensive analysis of methods for assessing polygenic burden on Alzheimer's disease pathology and risk beyond APOE." *Brain communications* vol. 2,1 (2020): fcz047. doi:10.1093/braincomms/fcz047

Andrews, S J *et al*. "The complex genetic architecture of Alzheimer's disease: novel insights and future directions." *EBioMedicine* vol. 90 (2023): 104511. doi:10.1016/j.ebiom.2023.104511

Arbo, B D *et al*. "Resveratrol Derivatives as Potential Treatments for Alzheimer's and Parkinson's Disease." *Frontiers in aging neuroscience* vol. 12 103 (2020). doi:10.3389/fnagi.2020.00103

Ashburner, M *et al*. "Gene ontology: tool for the unification of biology." *Nat Genet* vol 25 1 (2000): 25-9. doi:10.5281/zenodo.7942786

Atri, A. "The Alzheimer's Disease Clinical Spectrum." *Medical Clinics of North America* vol 103 2, 1 Mar. 2019.

Baker, E *et al*. "What does heritability of Alzheimer's disease represent?." *PloS one* vol. 18,4 e0281440. 28 Apr. 2023, doi:10.1371/journal.pone.0281440.

Bandres-Ciga, S *et al*. "Large-scale pathway specific polygenic risk and transcriptomic community network analysis identifies novel functional pathways in Parkinson disease." *Acta neuropathologica* vol. 140,3 (2020): 341-358. doi:10.1007/s00401-020-02181-3

Beecham, G W *et al*. "The Alzheimer's Disease Sequencing Project: Study design and sample selection." *Neurology. Genetics* vol. 3,5 e194. 13 Oct. 2017, doi:10.1212/NXG.0000000000000194

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

Bellou, E *et al*. "Age-dependent effect of APOE and polygenic component on Alzheimer's disease." *Neurobiology of aging* vol. 93 (2020): 69-77. doi:10.1016/j.neurobiolaging.2020.04.024

Belloy, M E *et al*. "APOE Genotype and Alzheimer Disease Risk Across Age, Sex, and Population Ancestry." *JAMA neurology*, e233599. 6 Nov. 2023, doi:10.1001/jamaneurol.2023.3599

Benitez, B A *et al*. "The PSEN1, p.E318G variant increases the risk of Alzheimer's disease in APOE-ε4 carriers." *PLoS genetics* vol. 9,8 (2013): e1003685. doi:10.1371/journal.pgen.1003685

Bennett, D A *et al*. "Neuropathology of older persons without cognitive impairment from two community-based studies." *Neurology* vol. 66,12 (2006): 1837-44. doi:10.1212/01.wnl.0000219668.47116.e6

Bertram, L *et al*. "Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database." *Nature genetics* vol. 39,1 (2007): 17-23. doi:10.1038/ng1934

Bertram, L *et al*. "Genome-wide association analysis reveals putative Alzheimer's disease susceptibility loci in addition to APOE." *American journal of human genetics* vol. 83,5 (2008): 623-32. doi:10.1016/j.ajhg.2008.10.008

Bi, X, and Liao, G. "Autophagic-lysosomal dysfunction and neurodegeneration in Niemann-Pick Type C mice: lipid starvation or indigestion?." *Autophagy* vol. 3,6 (2007): 646-8. doi:10.4161/auto.5074

Biffi, A, and Greenberg, S M. “Cerebral amyloid angiopathy: a systematic review.” *Journal of clinical neurology (Seoul, Korea)* vol. 7,1 (2011): 1-9. doi:10.3988/jcn.2011.7.1.1

Braak, H, and Braak, E. “Neuropathological stageing of Alzheimer-related changes.” *Acta neuropathologica* vol. 82,4 (1991): 239-59. doi:10.1007/BF00308809

Breijyeh, Z, and Karaman, R. “Comprehensive Review on Alzheimer’s Disease: Causes and Treatment.” *Molecules*, vol. 25 24, (2020). doi:10.3390/molecules25245789

Bright, N A *et al*. “Endolysosomes Are the Principal Intracellular Sites of Acid Hydrolase Activity.” *Current biology : CB* vol. 26,17 (2016): 2233-45. doi:10.1016/j.cub.2016.06.046

Blanchard, J W, Akay, L A, Davila-Velderrain, J *et al*. “APOE4 impairs myelination via cholesterol dysregulation in oligodendrocytes.” *Nature* vol. 611 (2022): 769–779. doi:10.1038/s41586-022-05439-w

Blue, E E *et al*. “Local ancestry at APOE modifies Alzheimer's disease risk in Caribbean Hispanics.” *Alzheimer's & dementia : the journal of the Alzheimer's Association* vol. 15,12 (2019): 1524-1532. doi:10.1016/j.jalz.2019.07.016

Buckley, R F *et al*. “A Call to Action to Address Sex Differences in Alzheimer Disease Clinical Trials.” *JAMA neurology* vol. 80,8 (2023): 769-770. doi:10.1001/jamaneurol.2023.1059

Buenrostro, J D *et al*. “ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide.” *Current protocols in molecular biology* vol. 109 21.29.1-21.29.9. 5 Jan. 2015, doi:10.1002/0471142727.mb2129s109

Buchman, A S *et al*. “Association of brain pathology with the progression of frailty in older adults.” *Neurology* vol. 80,22 (2013): 2055-61. doi:10.1212/WNL.0b013e318294b462

Che, R, and Motsinger-Reif, A A. “A new explained-variance based genetic risk score for predictive modeling of disease risk.” *Statistical applications in genetics and molecular biology* vol. 11,4 Article 15. 25 Sep. 2012, doi:10.1515/1544-6115.1796

Chen, G *et al*. “Amyloid beta: structure, biology and structure-based therapeutic development.” *Acta pharmacologica Sinica* vol. 38,9 (2017): 1205-1235. doi:10.1038/aps.2017.28

Chen, S, Francioli, L C, *et al*. “A genome-wide mutational constraint map quantified from variation in 76,156 human genomes.” *bioRxiv* 2022.03.20.485034 (2022). doi:10.1101/2022.03.20.485034

Choi, S W *et al*. “PRSet: Pathway-based polygenic risk score analyses and software.” *PLoS genetics* vol. 19,2 e1010624. 7 Feb. 2023, doi:10.1371/journal.pgen.1010624

Choudhury, P *et al*. “APOE ε4 Allele Testing and Risk of Alzheimer Disease.” *JAMA* vol. 325,5 (2021): 484-485. doi:10.1001/jama.2020.15085

Colonna, M, and Butovsky, O. “Microglia Function in the Central Nervous System During Health and Neurodegeneration.” *Annual review of immunology* vol. 35 (2017): 441-468. doi:10.1146/annurev-immunol-051116-052358

Corder, E H *et al*. “Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families.” *Science (New York, N.Y.)* vol. 261,5123 (1993): 921-3. doi:10.1126/science.8346443

Corder, E, Saunders, A, Risch, N *et al*. “Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease.” *Nat Genet* 7, 180–184 (1994). doi:10.1038/ng0694-180

Cruchaga, C *et al*. “Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families.” *PloS one* vol. 7,2 (2012): e31039. doi:10.1371/journal.pone.0031039

Cruchaga, C *et al*. “Polygenic risk score of sporadic late-onset Alzheimer's disease reveals a shared architecture with the familial and early-onset forms.” *Alzheimer's & dementia : the journal of the Alzheimer's Association* vol. 14,2 (2018): 205-214. doi:10.1016/j.jalz.2017.08.013

Darst, B F *et al*. “Pathway-Specific Polygenic Risk Scores as Predictors of Amyloid-β Deposition and Cognitive Function in a Sample at Increased Risk for Alzheimer's Disease.” *Journal of Alzheimer's disease : JAD* vol. 55,2 (2017): 473-484. doi:10.3233/JAD-160195

Distl, R *et al*. “Cholesterol storage and tau pathology in Niemann-Pick type C disease in the brain.” *The Journal of pathology* vol. 200,1 (2003): 104-11. doi:10.1002/path.1320

de Rojas, I, Moreno-Grau, S, Tesi, N *et al*. “Common variants in Alzheimer’s disease and risk stratification by polygenic risk scores.” *Nat Commun* vol. 12, 3417 (2021). doi:10.1038/s41467-021-22491-8

Donovan, M K R *et al*. “Cellular deconvolution of GTEx tissues powers discovery of disease and cell-type associated regulatory variants.” *Nature Commun* vol. 11,1 955 (2020). doi:10.1038/s41467-020-14561-0

Escott-Price, V *et al*. “Common polygenic variation enhances risk prediction for Alzheimer's disease.” *Brain : a journal of neurology* vol. 138, 12 (2015): 3673-84. doi:10.1093/brain/awv268

Escott-Price, V *et al*. “Polygenic score prediction captures nearly all common genetic risk for Alzheimer's disease.” *Neurobiology of aging* vol. 49 (2017): 214.e7-214.e11. doi:10.1016/j.neurobiolaging.2016.07.018

Farrer, L A *et al*. “Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium.” *JAMA* vol. 278,16 (1997): 1349-56.

Finkemeier I, König A C, Heard W *et al*. “Transcriptomic analysis of the role of carboxylic acids in metabolite signaling in Arabidopsis leaves.” *Plant Physiol*. 2013 May;162(1):239-53. doi: 10.1104/pp.113.214114. Epub 2013 Mar 13. PMID: 23487434; PMCID: PMC3641205.

Gatz, M *et al*. “Role of genes and environments for explaining Alzheimer disease.” *Archives of general psychiatry* vol. 63,2 (2006): 168-74. doi:10.1001/archpsyc.63.2.168

Goate, A *et al*. “Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease.” *Nature* vol. 349,6311 (1991): 704-6. doi:10.1038/349704a0

Grupe, A *et al*. “Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants.” *Human molecular genetics* vol. 16,8 (2007): 865-73. doi:10.1093/hmg/ddm031

GTEx Consortium. “The Genotype-Tissue Expression (GTEx) project.” *Nature genetics* vol. 45,6 (2013): 580-5. doi:10.1038/ng.2653

Halcrow, P *et al*. “Importance of measuring endolysosome, cytosolic, and extracellular pH in understanding the pathogenesis of and possible treatments for glioblastoma multiforme.” *Cancer Reports* vol. 2,6 (2019). doi:10.1002/cnr2.1193

Harold, D *et al*. “Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease.” *Nature genetics* vol. 41,10 (2009): 1088-93. doi:10.1038/ng.440

Harris, C R, Millman, K J, van der Walt, S J *et al*. “Array programming with NumPy.” *Nature* vol. 585 (2020): 357–362. doi:10.1038/s41586-020-2649-2.

Harrison, J R *et al*. “From Polygenic Scores to Precision Medicine in Alzheimer's Disease: A Systematic Review.” *Journal of Alzheimer's disease : JAD* vol. 74,4 (2020): 1271-1283. doi:10.3233/JAD-191233

Hippius, H, and Neundörfer, G. “The discovery of Alzheimer's disease.” *Dialogues in clinical neuroscience* vol. 5,1 (2003): 101-8. doi:10.31887/DCNS.2003.5.1/hhippius

Hohman, T J *et al*. “Sex-Specific Association of Apolipoprotein E With Cerebrospinal Fluid Levels of Tau.” *JAMA neurology* vol. 75,8 (2018): 989-998. doi:10.1001/jamaneurol.2018.0821

Hollingworth, P *et al*. “Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease.” *Nature genetics* vol. 43,5 (2011): 429-35. doi:10.1038/ng.803

Hu, Y B, Dammer, E, Ren, R J *et al.* “The endosomal-lysosomal system: from acidification and cargo sorting to neurodegeneration.” *Transl Neurodegener* 4, 18 (2015). doi:10.1186/s40035-015-0041-1

Hyman, B T *et al*. “National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease.” *Alzheimer's & dementia : the journal of the Alzheimer's Association* vol. 8,1 (2012): 1-13. doi:10.1016/j.jalz.2011.10.007

Igo, R P Jr *et al*. “Genetic Risk Scores.” *Current protocols in human genetics* vol. 104,1 (2019): e95. doi:10.1002/cphg.95

International Schizophrenia Consortium. “Common polygenic variation contributes to risk of schizophrenia and bipolar disorder.” *Nature* 460, 748–752 (2009). doi:10.1038/nature08185

James, B D *et al*. “Contribution of Alzheimer disease to mortality in the United States.” *Neurology* vol. 82,12 (2014): 1045-50. doi:10.1212/WNL.0000000000000240

Jansen, I E *et al*. “Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk.” *Nature genetics* vol. 51,3 (2019): 404-413. doi:10.1038/s41588-018-0311-9

Jarzebski, M P, Elmqvist, T, Gasparatos, A *et al*. “Ageing and population shrinking: implications for sustainability in the urban century.” *npj Urban Sustain* 1, 17 (2021). doi:10.1038/s42949-021-00023-z

Joachim, C L *et al*. “Clinically diagnosed Alzheimer's disease: autopsy results in 150 cases.” *Annals of neurology* vol. 24,1 (1988): 50-6. doi:10.1002/ana.410240110

Kahan, B C, Jairath, V, Doré, C J *et al*. “The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies.” *Trials* 15, 139 (2014). doi:10.1186/1745-6215-15-139

Kamboh, M I. “Genomics and Functional Genomics of Alzheimer's Disease.” *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics* vol. 19,1 (2022): 152-172. doi:10.1007/s13311-021-01152-0

Kang, D E *et al*. “Modulation of amyloid beta-protein clearance and Alzheimer's disease susceptibility by the LDL receptor-related protein pathway.” *The Journal of clinical investigation* vol. 106,9 (2000): 1159-66. doi:10.1172/JCI11013

Karczewski, K J, Francioli, L C, Tiao, G *et al*. “The mutational constraint spectrum quantified from variation in 141,456 humans.” *Nature* vol. 581 (2020): 434–443. doi:/10.1038/s41586-020-2308-7

Knopman, D S *et al*. “Neuropathology of cognitively normal elderly.” *Journal of neuropathology and experimental neurology* vol. 62,11 (2003): 1087-95. doi:10.1093/jnen/62.11.1087

Kunkle, B W, Grenier-Boley, B, Sims, R *et al*. “Genetic meta-analysis of diagnosed Alzheimer’s disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing.” Nat Genet 51, 414–430 (2019). doi:10.1038/s41588-019-0358-2

Mi, H *et al*. “PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees.” *Nucleic acids research* vol. 41, Database issue (2013): D377-86. doi:10.1093/nar/gks1118

Nuriel T E *et al*. “The Endosomal-Lysosomal Pathway Is Dysregulated by APOE4 Expression in Vivo.” *Front Neurosci.* vol. 12, 11 (2017): 702. doi:10.3389/fnins.2017.00702.

Lambert, J C, Ibrahim-Verbaas, C, Harold, D *et al*. “Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease.” *Nat Genet* vol 45, (2013): 1452–1458. doi:10.1038/ng.2802

Lee, G *et al*. “Developing an oligogenic risk score for Alzheimer’s disease (adORS).” *Alzheimer's Dement*. vol. 18 (2022): e066011. doi:10.1002/alz.066011

Levy-Lahad, E *et al*. “Candidate gene for the chromosome 1 familial Alzheimer's disease locus.” *Science (New York, N.Y.)* vol. 269,5226 (1995): 973-7. doi:10.1126/science.7638622

Li, Y *et al*. “Genotype imputation.” *Annual review of genomics and human genetics* vol. 10 (2010): 387-406. doi:10.1146/annurev.genom.9.081307.164242

Li, H *et al*. “The Sequence Alignment/Map format and SAMtools.” *Bioinformatics (Oxford, England)* vol. 25,16 (2009): 2078-9. doi:10.1093/bioinformatics/btp352

Lian, Y *et al*. “Clarity on the blazing trail: clearing the way for amyloid-removing therapies for Alzheimer's disease.” *Molecular psychiatry*, 10.1038/s41380-023-02324-4. 24 Nov. 2023, doi:10.1038/s41380-023-02324-4

Liao, G *et al*. “Cholesterol accumulation is associated with lysosomal dysfunction and autophagic stress in Npc1 -/- mouse brain.” *The American journal of pathology* vol. 171,3 (2007): 962-75. doi:10.2353/ajpath.2007.070052

Liao, Y *et al*. “featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.” *Bioinformatics (Oxford, England)* vol. 30,7 (2014): 923-30. doi:10.1093/bioinformatics/btt656

Lim, A C *et al*. “Quantification of race/ethnicity representation in Alzheimer's disease neuroimaging research in the USA: a systematic review.” *Communications medicine* vol. 3,1 101. 25 Jul. 2023, doi:10.1038/s43856-023-00333-6

Liu, C *et al*. “Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy.” *Nature Reviews Neurology* vol. 9,2 (2013): 106-18. doi:10.1038/nrneurol.2012.263

Love M I, Huber W, Anders S (2014). “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.” *Genome Biology*, 15, 550. doi:10.1186/s13059-014-0550-8)

Machiela, M J, and Chanock, S J. “LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants.” *Bioinformatics (Oxford, England)* vol. 31,21 (2015): 3555-7. doi:10.1093/bioinformatics/btv402

Mahley, R W *et al*. “Apolipoprotein E4: a causative factor and therapeutic target in neuropathology, including Alzheimer's disease.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 103,15 (2006): 5644-51. doi:10.1073/pnas.0600549103

Mahley, R W. “Apolipoprotein E: cholesterol transport protein with expanding role in cell biology.” *Science (New York, N.Y.)* vol. 240,4852 (1988): 622-30. doi:10.1126/science.3283935

Martin, R, Splitt, M, Genevieve, D *et al*. “De novo variants in *CNOT3* cause a variable neurodevelopmental disorder.” *Eur J Hum Genet* 27, 1677–1682 (2019). doi:10.1038/s41431-019-0413-6

McLaren, W *et al*. “The Ensembl Variant Effect Predictor.” *Genome biology* vol. 17,1 122 (2016): doi:10.1186/s13059-016-0974-4

Mendez, M F. “Early-onset Alzheimer Disease and Its Variants.” *Continuum (Minneapolis, Minn.)* vol. 25,1 (2019): 34-51. doi:10.1212/CON.0000000000000687

McKinney, W *et al*. “Data structures for statistical computing in python.” *Proceedings of the 9th Python in Science Conference* vol. 445 (2010): 51–56.

Mirra, S S *et al*. “The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease.” *Neurology* vol. 41,4 (1991): 479-86. doi:10.1212/wnl.41.4.479

Molania, R, Foroutan, M, Gagnon-Bartsch, J A *et al*. “Removing unwanted variation from large-scale RNA sequencing data with PRPS.” *Nat Biotechnol* vol. 41 (2023): 82–95. doi:10.1038/s41587-022-01440-w

Montine, T J *et al*. “Adult Changes in Thought study: dementia is an individually varying convergent syndrome with prevalent clinically silent diseases that may be modified by some commonly used therapeutics.” *Current Alzheimer research* vol. 9,6 (2012): 718-23. doi:10.2174/156720512801322555

Mungall, C J *et al*. “The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species.” *Nucleic acids research* vol. 45,D1 (2017): D712-D722. doi:10.1093/nar/gkw1128

Naj, A C *et al*. “Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease.” *Nature genetics* vol. 43,5 (2011): 436-41. doi:10.1038/ng.801

Nelson, P T, Kukull, W A, Frosch, M P. “Thinking outside the box: Alzheimer-type neuropathology that does not map directly onto current consensus recommendations.” *J Neuropathol Exp Neurol* 69,5 (2010): 449–454. doi:10.1097/NEN.0b013e3181d8db07

Nickerson, D A *et al*. “Sequence diversity and large-scale typing of SNPs in the human apolipoprotein E gene.” *Genome research* vol. 10,10 (2000): 1532-45. doi:10.1101/gr.146900

Nixon, R A. “Autophagy, amyloidogenesis and Alzheimer disease.” *Journal of cell science* vol. 120,Pt 23 (2007): 4081-91. doi:10.1242/jcs.019265

Nixon, R A. “The role of autophagy in neurodegenerative disease.” *Nature medicine* vol. 19,8 (2013): 983-97. doi:10.1038/nm.3232

Nixon, R A. “Amyloid precursor protein and endosomal-lysosomal dysfunction in Alzheimer's disease: inseparable partners in a multifactorial disease.” *FASEB journal : official*

publication of the Federation of American Societies for Experimental Biology vol. 31,7 (2017): 2729-2743. doi:10.1096/fj.201700359

Oveisgharan, S *et al*. "Sex differences in Alzheimer's disease and common neuropathologies of aging." *Acta neuropathologica* vol. 136,6 (2018): 887-900. doi:10.1007/s00401-018-1920-1

"Pacific Udall Center." *Stanford University* (2023).

Pagès H, Carlson M, Falcon S, Li N. *AnnotationDbi: Manipulation of SQLite-based annotations in Bioconductor*. R package version 1.62.2.

Pike, K E *et al*. "Beta-amyloid imaging and memory in non-demented individuals: evidence for preclinical Alzheimer's disease." *Brain : a journal of neurology* vol. 130,Pt 11 (2007): 2837-44. doi:10.1093/brain/awm238

Pillay, C S *et al*. "Endolysosomal proteolysis and its regulation." *The Biochemical journal* vol. 363,Pt 3 (2002): 417-29. doi:10.1042/0264-6021:3630417

Plassman, B L *et al*. "Prevalence of dementia in the United States: the aging, demographics, and memory study." *Neuroepidemiology* vol. 29,1-2 (2007): 125-32. doi:10.1159/000109998

Pruim, R J *et al*. "LocusZoom: regional visualization of genome-wide association scan results." *Bioinformatics (Oxford, England)* vol. 26,18 (2010): 2336-7. doi:10.1093/bioinformatics/btq419

Purcell, S *et al*. "PLINK: a tool set for whole-genome association and population-based linkage analyses." *American journal of human genetics* vol. 81,3 (2007): 559-75. doi:10.1086/519795

Qian, J *et al*. "Association of APOE Genotype With Heterogeneity of Cognitive Decline Rate in Alzheimer Disease." *Neurology* vol. 96,19 (2021): e2414-e2428. doi:10.1212/WNL.0000000000011883

R Core Team. "R: A language and environment for statistical computing." R *Foundation for Statistical Computing* (2021).

Rajabli, F *et al*. "Ancestral origin of ApoE ε4 Alzheimer disease risk in Puerto Rican and African American populations." *PLoS genetics* vol. 14,12 (2018): e1007791. doi:10.1371/journal.pgen.1007791

Riaz, M *et al*. "Effect of APOE and a polygenic risk score on incident dementia and cognitive decline in a healthy older population." *Aging cell* vol. 20,6 (2021): e13384. doi:10.1111/acel.13384

Ruffier, M *et al*. "Ensembl core software resources: storage and programmatic access for DNA sequence and genome annotation." *Database: the journal of biological databases and curation* vol. 2017,1 (2017). doi:10.1093/database/bax020

"Seattle Alzheimer's Disease Brain Cell Atlas Consortium (SEA-AD)." *Allen Brain Map, The Allen Institute* (2023).

"Seattle Longitudinal Study (SLS)." *University of Washington Medicine Department of Psychiatry and Behavioral Sciences* (2023).

Seshadri S, Fitzpatrick A L, Ikram M A, *et al*. "Genome-wide Analysis of Genetic Loci Associated With Alzheimer Disease." *JAMA*. 2010;303(18):1832–1840. doi:10.1001/jama.2010.574

Silberstein, M *et al*. “Pathway analysis for genome-wide genetic variation data: Analytic principles, latest developments, and new opportunities.” *Journal of genetics and genomics* vol. 48,3 (2021): 173-183. doi:10.1016/j.jgg.2021.01.007

Sherrington, R *et al*. “Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease.” *Nature* vol. 375,6534 (1995): 754-60. doi:10.1038/375754a0

Sleegers, K *et al*. “A 22-single nucleotide polymorphism Alzheimer's disease risk score correlates with family history, onset age, and cerebrospinal fluid Aβ42.” *Alzheimer's & dementia : the journal of the Alzheimer's Association* vol. 11,12 (2015): 1452-1460. doi:10.1016/j.jalz.2015.02.013

Stocker, H, Trares, K, Beyer, L *et al*. “Alzheimer’s polygenic risk scores, APOE, Alzheimer’s disease risk, and dementia-related blood biomarker levels in a population-based cohort study followed over 17 years.” *Alz Res Therapy* 15, 129 (2023). doi:10.1186/s13195-023-01277-8

Strittmatter, W J *et al*. “Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 90,5 (1993): 1977-81. doi:10.1073/pnas.90.5.1977

Szabo, M P, Mishra, S, Knupp, A, Young, J. “The role of Alzheimer's disease risk genes in endolysosomal pathways.” *Neurobiology of Disease* vol. 162 (2022). doi:10.1016/j.nbd.2021.105576.

Tahami Monfared, A A, Byrnes, M J, White, L A *et al.* “The Humanistic and Economic Burden of Alzheimer's Disease.” *Neurol Ther* 11, (2022): 525–551. doi:10.1007/s40120-022-00335-x

Tahami Monfared, A A *et al*. “Alzheimer's Disease: Epidemiology and Clinical Progression.” *Neurology and therapy* vol. 11,2 (2022): 553-569. doi:10.1007/s40120-022-00338-8

Takahashi, R H *et al*. “Intraneuronal Alzheimer abeta42 accumulates in multivesicular bodies and is associated with synaptic pathology.” *The American journal of pathology* vol. 161,5 (2002): 1869-79. doi:10.1016/s0002-9440(10)64463-x

Tanzi, R E. “The genetics of Alzheimer disease.” *Cold Spring Harbor perspectives in medicine* vol. 2,10 a006296. 1 Oct. 2012, doi:10.1101/cshperspect.a006296

Thal, D R *et al*. “Phases of A beta-deposition in the human brain and its relevance for the development of AD.” *Neurology* vol. 58,12 (2002): 1791-800. doi:10.1212/wnl.58.12.1791

“The 1000 Genomes Project Consortium, a global reference for human genetic variation.” *Nature* vol. 526 (2015): 68-74. doi:10.1038/nature15393

Thomas, P D *et al*. “PANTHER: Making genome-scale phylogenetics accessible to all” *Protein Science* (2021). doi:10.1002/pro.4218

Todd, S *et al*. “Survival in dementia and predictors of mortality: a review.” *International journal of geriatric psychiatry* vol. 28,11 (2013): 1109-24. doi:10.1002/gps.3946

Tosto, G *et al*. “Polygenic risk scores in familial Alzheimer disease.” *Neurology* vol. 88,12 (2017): 1180-1186. doi:10.1212/WNL.0000000000003734

Ulery, P G *et al*. “Modulation of beta-amyloid precursor protein processing by the low density lipoprotein receptor-related protein (LRP). Evidence that LRP contributes to the pathogenesis of Alzheimer's disease.” *The Journal of biological chemistry* vol. 275,10 (2000): 7410-5. doi:10.1074/jbc.275.10.7410

"University of Washington Alzheimer's Disease Research Center (UW-ADRC)." *UW Memory and Brain Wellness Center* (2023).

Van Acker, Z P, Bretou, M. & Annaert, W. "Endo-lysosomal dysregulations and late-onset Alzheimer's disease: impact of genetic risk factors." *Mol Neurodegeneration* 14, 20 (2019). doi:/10.1186/s13024-019-0323-7

Van Cauwenberghe, C *et al*. "The genetic landscape of Alzheimer disease: clinical implications and perspectives." *Genetics in medicine : official journal of the American College of Medical Genetics* vol. 18,5 (2016): 421-30. doi:10.1038/gim.2015.117

van Weering, J R T, and Scheper, W. "Endolysosome and Autolysosome Dysfunction in Alzheimer's Disease: Where Intracellular and Extracellular Meet." *CNS drugs* vol. 33,7 (2019): 639-648. doi:10.1007/s40263-019-00643-1

Varadharajan, A *et al*. "Guidelines for pharmacotherapy in Alzheimer's disease - A primer on FDA-approved drugs." *Journal of neurosciences in rural practice* vol. 14,4 (2023): 566-573. doi:10.25259/JNRP_356_2023

Vassar, R *et al*. "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE." *Science (New York, N.Y.)* vol. 286,5440 (1999): 735-41. doi:10.1126/science.286.5440.735

Viola, K L, and Klein, W L. "Amyloid β oligomers in Alzheimer's disease pathogenesis, treatment, and diagnosis." *Acta neuropathologica* vol. 129,2 (2015): 183-206. doi:10.1007/s00401-015-1386-3

Walker, J M *et al*. "Age-related progression of Alzheimer disease neuropathologic change (ADNC) in cognitively normal individuals: Toward a definition of resilience and resistance." *Alzheimer's and Dementia* vol. 17,3 (2022): doi:10.1002/alz.054305

Wall, L, *et al*. “The Perl programming language.” *Prentice Hall Software Series* (1994).

Wang, D *et al*. “Frequency of Variants in Mendelian Alzheimer's Disease Genes within the Alzheimer's Disease Sequencing Project (ADSP).” *medRxiv : the preprint server for health sciences* 25 Oct. 2023, doi:10.1101/2023.10.24.23297227. Preprint.

Ward, L, and Kellis, M. “HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants.” *Nucleic Acids Research* vol 40, 1 (2012): 930-934, doi:10.1093/nar/gkr917

Welch, B L. “The generalisation of student's problems when several different population variances are involved.” *Biometrika* vol. 34,1-2 (1947): 28-35. doi:10.1093/biomet/34.1-2.28

Wickham, H. “ggplot2: Elegant Graphics for Data Analysis.” *Springer-Verlag*, New York (2016). ISBN 978-3-319-24277-4

Wisniewski, T, and Frangione, B. “Apolipoprotein E: a pathological chaperone protein in patients with cerebral and systemic amyloid.” *Neuroscience letters* vol. 135,2 (1992): 235-8. doi:10.1016/0304-3940(92)90444-c

Wharton, S B, Wang, D, Parikh, C *et al*. “Epidemiological pathology of Aβ deposition in the ageing brain in CFAS: addition of multiple Aβ-derived measures does not improve dementia assessment using logistic regression and machine learning approaches.” *acta neuropathol commun* 7, 198 (2019). doi:10.1186/s40478-019-0858-4

White, M J *et al*. “Strategies for Pathway Analysis Using GWAS and WGS Data.” *Current protocols in human genetics* vol. 100,1 (2019): e79. doi:10.1002/cphg.79

Wong, E D *et al*. “The YeastGenome app: the Saccharomyces Genome Database at your fingertips.” *Database : the journal of biological databases and curation* vol. 2013 bat004. 8 Feb. 2013, doi:10.1093/database/bat004

Wray, N R *et al*. “Prediction of individual genetic risk to disease from genome-wide association studies.” *Genome research* vol. 17,10 (2007): 1520-8. doi:10.1101/gr.6665407

Wu, W *et al*. “The FDA-approved anti-amyloid-β monoclonal antibodies for the treatment of Alzheimer's disease: a systematic review and meta-analysis of randomized controlled trials.” *European journal of medical research* vol. 28,1 544. 28 Nov. 2023, doi:10.1186/s40001-023-01512-w

Xu, Y *et al*. “Effect of Pathway-Specific Polygenic Risk Scores for Alzheimer's Disease (AD) on Rate of Change in Cognitive Function and AD-Related Biomarkers Among Asymptomatic Individuals.” *Journal of Alzheimer's disease : JAD* vol. 94,4 (2023): 1587-1605. doi:10.3233/JAD-230097

Xue, D *et al*. “Large-scale sequencing studies expand the known genetic architecture of Alzheimer's disease.” *Alzheimer's & dementia (Amsterdam, Netherlands)* vol. 13,1 e12255. 31 Dec. 2021, doi:10.1002/dad2.12255

Yan, M, & Zheng, T. “Role of the endolysosomal pathway and exosome release in tau propagation.” *Neurochemistry international* vol. 145 (2021): 104988. doi:10.1016/j.neuint.2021.104988

Yu, W H *et al*. “Macroautophagy--a novel Beta-amyloid peptide-generating pathway activated in Alzheimer's disease.” *The Journal of cell biology* vol. 171,1 (2005): 87-98. doi:10.1083/jcb.200505082

Zhang, F, & Lupski, J. “Non-coding genetic variants in human disease.” *Human molecular genetics* vol. 24,R1 (2015): R102-10. doi:10.1093/hmg/ddv259

Zhang, Q, Sidorenko, J, Couvy-Duchesne, B *et al.* “Risk prediction of late-onset Alzheimer’s disease implies an oligogenic architecture.” *Nat Commun* 11, 4799 (2020). doi:10.1038/s41467-020-18534-1

Zhu, A, Ibrahim, J G, Love, M I. “Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences.” *Bioinformatics* vol 35 12 (2019): 2084–2092. doi:10.1093/bioinformatics/bty895

Tables

Table 1. Descriptive Statistics for baseline variables of all samples by case-control status.

*Refers to age of onset of cognitive symptoms (AAO) for cases and age at last cognitive exam (AAE) for controls.

N (%)	AD Cases (n = 88)	Controls (n =100)	Total (n = 188)
Sex			
Male	23 (26.1%)	41 (41.0%)	64 (34.0%)
Female	65 (73.9%)	59 (59.0%)	124 (66.0%)
Reported race			
Non-Hispanic White	76 (86.4%)	91 (91.0%)	167 (88.8%)
Asian	2 (2.3%)	1 (1.0%)	3 (1.6%)
Black/African American	0 (0.0%)	2 (2.0%)	2 (1.1%)
American/Alaskan Native	1 (1.1%)	0 (0.0%)	1 (0.5%)
Unknown/Unreported	6 (6.8%)	4 (4.0%)	10 (5.3%)
Age*			
Mean (SD)	78.1 (14.6)	87.3 (7.7)	82.6 (12.5)
Median [Range]	83 [48, 100]	88 [68, 101]	85 [48, 101]
***APOE* Genotype**			
ε2ε2	0 (0.0%)	0 (0.0%)	0 (0.0%)
ε2ε3	7 (8.0%)	13 (13.0%)	20 (10.6%)
ε3ε3	42 (47.7%)	60 (60.0%)	102 (54.3%)
ε2ε4	2 (2.3%)	1 (1.0%)	3 (1.6%)
ε3ε4	31 (35.2%)	25 (25.0%)	56 (29.8%)
ε4ε4	6 (6.8%)	1 (1.0%)	7 (3.7%)

Table 2. Neuropathology by clinical consensus AD status.

	AD Cases (n = 88)	Controls (n = 100)	Total (n = 188)
Thal Stage			
0 - None	1 (1.1%)	15 (15.0%)	16 (8.5%)
1 - Neocortex	1 (1.1%)	10 (10.0%)	11 (5.9%)
2 - Allocortex/limbic	1 (1.1%)	16 (16.0%)	17 (5.9%)
3 - Diencephalon/Basal Ganglia	6 (6.8%)	16 (16.0%)	22 (11.7%)
4 - Brainstem/midbrain	33 (37.5%)	34 (34.0%)	67 (35.6%)
5 - Cerebellum	46 (52.3%)	9 (9.0%)	55 (29.3%)
Cerebral Amyloid Angiopathy Score			
0 - None	28 (31.8%)	39 (39.0%)	67 (35.6%)
1 - Low	18 (20.5%)	28 (28.0%)	46 (24.5%)
2 - Intermediate	30 (34.1%)	29 (29.0%)	59 (31.4%)
3 - High	12 (13.6%)	4 (4.0%)	16 (8.0%)
Braak Stage			
0 - None	0 (0.0%)	1 (1.0%)	1 (0.5%)
1 - Moderate entorhinal	0 (0.0%)	4 (4.0%)	4 (2.1%)
2 - Greater entorhinal	0 (0.0%)	3 (3.0%)	3 (1.6%)
3 - Moderate limbic	2 (2.3%)	16 (16.0%	18 (9.6%)
4 - Greater limbic	13 (14.8%)	36 (36.0%)	49 (26.0%)
5 - Moderate neocortical	26 (29.5%)	36 (36.0%)	62 (33.0%)
6 - Severe neocortical	47 (54.5%)	4 (4.0%)	51 (27.1%)
CERAD Score			
0 - No AD	2 (2.3%)	26 (26.0%)	28 (14.9%)
1 - Possible AD	8 (9.1%)	27 (27.0%)	35 (18.6%)
2 - Probable AD	30 (34.1%)	27 (27.0%)	57 (30.3%)
3 - Definite AD	48 (53.4%)	20 (20.0%)	68 (36.2%)
Overall AD Neuropathologic Change			
0 - None	1 (1.1%)	17 (17.0%)	16 (8.5%)
1 - Low	2 (2.3%)	25 (25.0%)	27 (14.4%)
2 - Intermediate	17 (19.3%)	30 (30.0%)	47 (25.0%)
3 - High	68 (77.3%)	28 (28.0%)	96 (51.1%)

Table 3. GWAS SNPs used to calculate General AD GRS and ePRS. "SNP" refers to each variant position by chromosome:position on GRCh37:reference allele:alternate allele, rsID is the marker identifier, "OR" is the overall odds ratio from Lambert *et al.* (2013), "EAF" is the effect allele frequency as defined by the allele frequencies in the 1000 Genomes European reference (1000 Genomes Project Consortium, 2015). * indicates a SNP for which a proxy SNP was used in place of a lead SNP in samples genotyped using an array. SNP weights for these proxy SNPs were downweighted by the linkage disequilibrium r^2 value with the lead SNP.

SNP	rsID	Gene	Effect Allele	OR [95% CI]	EAF	PRS weight
1:207692049:A:G	rs6656401	*CR1*	A	1.18	.20	.04039
2:127892810:C:T	rs6733839	*BIN1*	T	1.22	.39	.05962
2:127894615:A:G	rs744373	*BIN1*	G	1.24	.29	.06001
2:234068476:C:T	rs35349669	*INPP5D*	T	1.08	.50	.02363
5:88223420:A:G	rs190982	*MEF2C*	A	1.08	.62	.02158
6:32578530:C:A	rs9271192	*HLA-DRB5*	C	1.11	.28	.02865
6:47487762:A:G	rs10948363	*CD2AP*	G	1.10	.27	.02590
7:37841534:A:G	rs2718058	*NME8*	A	1.08	.64	.02136
7:100004446:C:T	rs1476679	*ZCWPW1*	T	1.10	.70	.02645
7:143110762:G:A	rs11771145	*EPHA1*	G	1.11	.64	.03107
8:27195121:T:C	rs28834970	*PTK2B*	C	1.10	.33	.02763
8:27467686:C:T	rs9331896	*CLU*	T	1.16	.60	.04534
11:47557871:T:C	rs10838725	*CELF1*	C	1.08	.30	.02162
11:59923508:A:G	rs983392	*MS4A6A*	A	1.11	.60	.03169
11:59971795:G:T	rs670139*	*MS4A4E*	T	1.08	.43	.02338
11:85867875:A:G	rs10792832	*PICALM*	A	1.15	.65	.04089
11:121435587:T:C	rs11218343	*SORL1*	T	1.30	.96	.03105
14:53400629:T:C	rs17125944	*FERMT2*	C	1.14	.095	.02361
14:92926952:G:T	rs10498633	*SLC24A4*	G	1.10	.77	.02420
18:29088958:C:T	rs8093731*	*DSG2*	C	1.37	.99	.01952
19:1046520:T:G	rs3764650	*ABCA7*	G	1.37	.084	.05366
19:1063443:A:G	rs4147929*	*ABCA7*	A	1.15	.83	.03232
19:51727962:C:A	rs3865444	*CD33*	C	1.06	.68	.01775
20:55018260:T:C	rs7274581	*CASS4*	T	1.14	.91	.02223

Table 4. Measuring evidence connecting each GWAS SNP to its published candidate.
Evidence linking SNPs to their gene target. A = 1 if ≥ 1 haplotype SNP is predicted to alter the candidate protein structure and 0 otherwise; N_{SNPs} is equal to the number of haplotype SNPs; N_{eQTLs} = number of eQTLs; $N_{regulatory}$ = number of SNPs falling within promoter, enhancer, or DNase I hypersensitivity (DHS) regions according to HaploReg v4.0 (Ward & Kellis, 2012); N_{bonus} = number of regulatory SNPs within a promoter, enhancer, or transcription factor binding site (TFBS) according to Variant Effect Predictor v91 (VEP; McLaren *et al.*, 2016).

Gene	SNP(s)	Endolysosomal Pathway	N_{SNPs}	A	N_{eQTLs}	$N_{regulatory}$	N_{bonus}	Score
CLU	rs1136000	Yes	11	1	10	11	3	**8.09**
HLA-DRB5	rs9271192	Yes	42	1	42	42	0	**8.00**
MS4A4E	rs670139	Yes	33	1	15	30	3	**6.91**
CR1	rs6656401	No	22	1	10	18	1	**6.78**
ABCA7	rs3764650, rs4147929	Yes	9	1	2	9	1	**6.56**
CD33	rs3865444	No	8	1	2	8	0	**6.50**
DSG2	rs8093731	No	15	1	0	15	2	**6.13**
BIN1	rs6733839, rs744373	Yes	4	0	4	4	0	**3.00**
CASS4	rs7274581	No	13	0	13	12	1	**3.00**
CELF1	rs10838725	No	8	0	8	8	0	**3.00**
EPHA1	rs11771145	Yes	1	0	1	1	0	**3.00**
INPP5D	rs35349669	No	54	0	54	54	0	**3.00**
MEF2C	rs190982	Yes	2	0	2	2	0	**3.00**
SORL1	rs11218343	Yes	1	0	1	1	0	**3.00**
CD2AP	rs10948363	Yes	83	0	81	76	2	**2.89**
MS4A6A	rs983392	No	108	0	91	91	0	**2.53**
ZCWPW1	rs1476679	No	4	0	3	4	0	**2.50**
SLC24A2	rs10498633	Yes	3	0	2	3	0	**2.33**
PTK2B	rs28834970	Yes	6	0	2	6	0	**1.67**
PICALM	rs10792832	Yes	4	0	1	4	0	**1.50**
FERMT2	rs1712944	No	10	0	0	10	0	**1.00**
NME8	rs2718058	No	1	0	0	0	0	**0.00**

Table 5. Association tests between GRS and AD-related phenotypes and covariates. Bold font indicates significant association signals at $p < 0.05$, while bold italic font indicates suggestive signals with $p < 0.10$. AD* indicates testing done on AD case control status restricted to controls with cognitive testing within 2 years of death. ‡ indicates age of onset of cognitive symptoms (AAO) for case sand age at last cognitive evaluation (AAE) for controls. § indicates regression testing coding 1 = male, 0 = female. OR indicates odds ratios, while β is the regression coefficient for the covariate.

Covariates									
	age‡			**sex§**			**ε4 carrier status**		
	effect size	95% CI	p-value	effect size	95% CI	p-value	effect size	95% CI	p-value
AD status	OR = .93	[.90, .96]	**1.3e-4**	OR = .54	[.26, 1.06]	***.08***	OR = 2.65	[1.33, 6.06]	**4.9e-3**
Thal	β = -.035	[-.053, -.017]	**1.4e-4**	β = -.11	[-0.58, .34]	.63	β = .90	[.43, 1.36]	**2.0e-4**
CERAD	β = -.024	[- .036, -.012]	**1.7e-4**	β = -.15	[0.46, .16]	.35	β = 1.23	[.013, 2.48]	**.048**
CAA	β = -7.8e-3	[-.020, 4.3e-3]	.20	β = .14	[-.17, .44]	.38	β = 1.59	[.39, 2.78]	**9.4e-3**
ADNC	β = -.017	[-.029, -5.4e-3]	**4.6e-3**	β = -.13	[-.44, .18]	.42	β = .51	[.20, .82]	**1.2e-3**
Braak	β = -.021	[-0.036, 6.7e-3]	**4.4e-3**	β = -.18	[-.56, .18]	.32	β = .50	[.12, .87]	**9.9-e3**
General AD GRS									
	APOE adjusted			*APOE* interaction					
	effect size	95% CI	p-value	effect size	95% CI	p-value	effect	95% CI	p-value
AD status	OR = .99	[.92, 1.08]	.87	OR = 1.01	[.92, 1.29]	.43	OR = .99	[.83, 1.17]	.87
AD*	OR = 1.01	[.92, 1.12]	.80	OR = 1.04	[.92, 1.18]	.27	OR = .93	[.75, 1.16]	.54
Thal	β = .054	[1.3e-3, .11]	**.045**	-	-	-	-	-	-
CERAD	β = .037	[9.5e-4, .074]	**.044**	β = .052	[-7.1e-3, .096]	**.023**	β = -.044	[-.12,.035]	.27
CAA	β = 7.0e-3	[-.028, .042]	.70	β = .029	[-0.014, 0.072]	0.19	β = -.066	[-0.14, 9.5e-3]	***.082***
ADNC	β = .024	[-.011, .059]	.18	-	-	-	-	-	-
Braak	β = .018	[-.025, .061]	.41	-	-	-	-	-	-
ePRS									
	APOE adjusted			*APOE* interaction					
	effect size	95% CI	p-value	effect size	95% CI	p-value	effect	95% CI	p-value
AD status	OR = 1.02	[.92, 1.13]	.69	OR = 1.07	[.92, 1.17]	0.23	OR = .96	[.77, 1.20]	.35
AD*	OR = 1.11	[.98, 1.27]	.12	OR = 1.21	[1.03, 1.43]	**.039**	OR = .87	[.63, 1.07]	.15
Thal	β = .076	[9.1e-3, .14]	**.026**	-	-	-	-	-	-
CERAD	β = .059	[.014, .11]	**.011**	β = .085	[.027, .14]	**4.1e-3**	β = -.070	[-.17, .032]	.15
CAA	β = .011	[-.034, .056]	.64	β = .041	[-.015, .097]	.16	β = -.084	[-.17, .011]	***.082***
ADNC	β = .042	[-1.0e-3, .083]	***.061***	-	-	-	-	-	-
Braak	β = .036	[-.019, .091]	.20	-	-	-	-	-	-

Table 6. Descriptive Statistics of Samples in RNA Expression Dataset

	High ePRS (n = 13)	Low ePRS (n = 13)	Total (n = 26)
ePRS			
Mean (SD)	16.42 (1.79)	7.90 (1.60)	12.16 (4.65)
Median [Range]	15.67 [14.56, 19.59]	8.21 [3.12, 9.38]	11.97 [3.12, 19.59]
Sex			
Male	5 (38%)	5 (38%)	10 (38%)
Female	8 (62%)	8 (62%)	16 (62%)
Age			
Mean (SD)	86.3 (13.2)	90.5 (9.8)	88.4 (11.5)
Median [Range]	89 [60, 103]	93 [68, 100]	91 [60, 103]
Cognitive Status			
Dementia	6 (46%)	6 (46%)	12 (46%)
No dementia	6 (46%)	7 (54%)	13 (50%)
Other	1 (8%)	0 (0%)	1 (4%)

Table 7. ANOVA testing log_2 transformed ePRS gene counts by tag SNP count. *Adjusted for PC1, RIN, & PMI. Bold font indicates significant association signals at $p < 0.05$.

Tag SNP	Target Gene	ANOVA F-statistic*	p-value
rs6733839	*BIN1*	1.37	.25
rs744373	*BIN1*	1.37	.25
rs190982	*MEF2C*	.010	.92
rs10948363	*CD2AP*	1.19	.29
rs28834970	*PTK2B*	.12	.74
rs9331896	*CLU*	.10	.75
rs10792832	*PICALM*	.21	.65
rs3764650	*ABCA7*	17.14	**6.1e-4**
rs4147929	*ABCA7*	1.58	.22

Table 8. Expression of ePRS genes across fibroblast types, measured by TPM.

Gene	Leptomeningeal TPM (n = 26)	Cultured Fibroblast TPM (n = 504)
BIN1	127.88	45.69
MEF2C	67.23	.88
HLA-DRB5	.12	.23
CD2AP	58.70	15.11
EPHA1	.77	.91
PTK2B	21.48	16.95
CLU	7042.50	13.02
MS4A4E	.03	.00
PICALM	321.32	133.9
SORL1	.81	.07
ABCA7	48.73	2.94
CD33	.03	.02

Figures

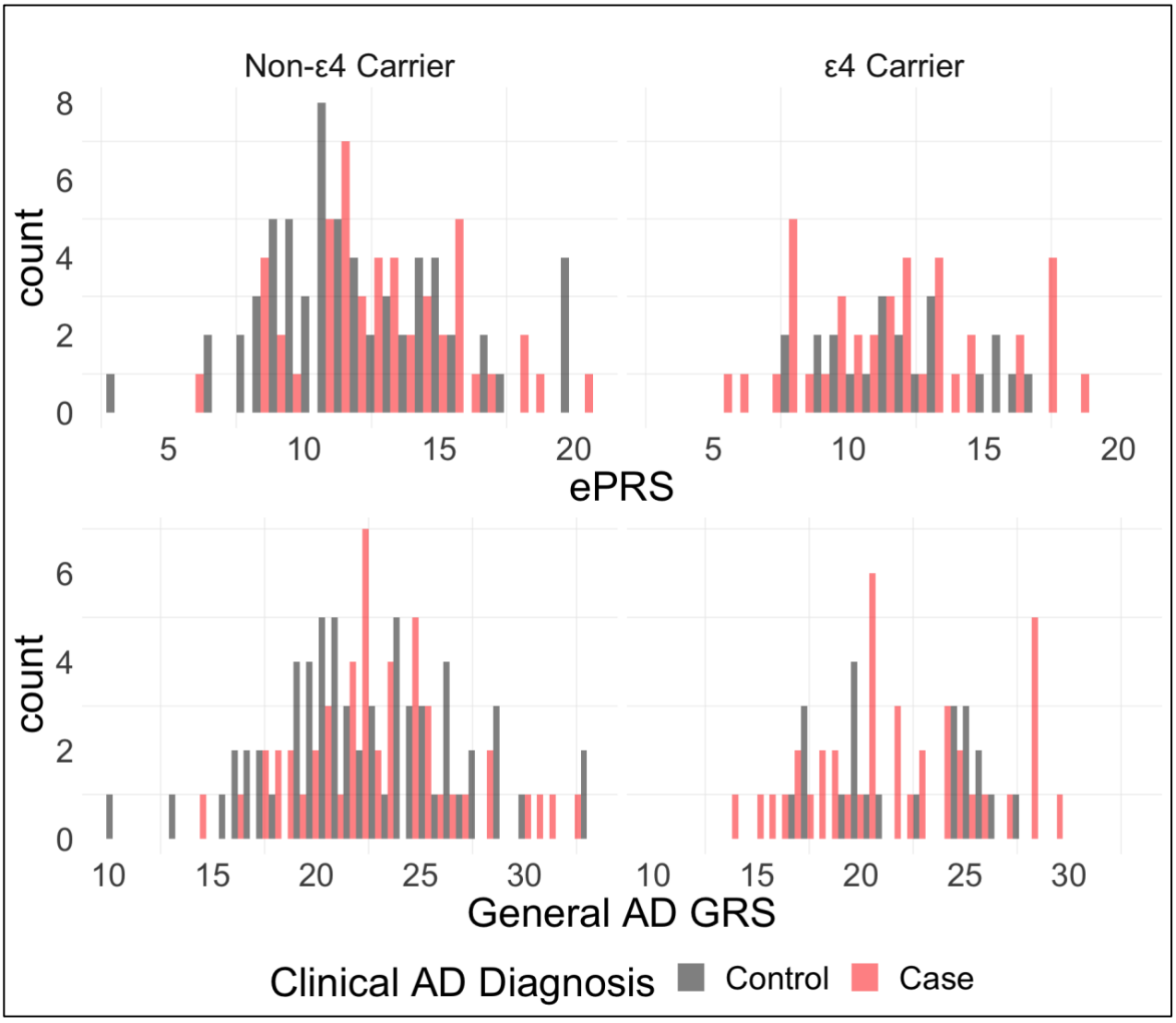


Figure 1. ePRS and General AD GRS score by AD case-control status and *APOE* ε4 carrier status.

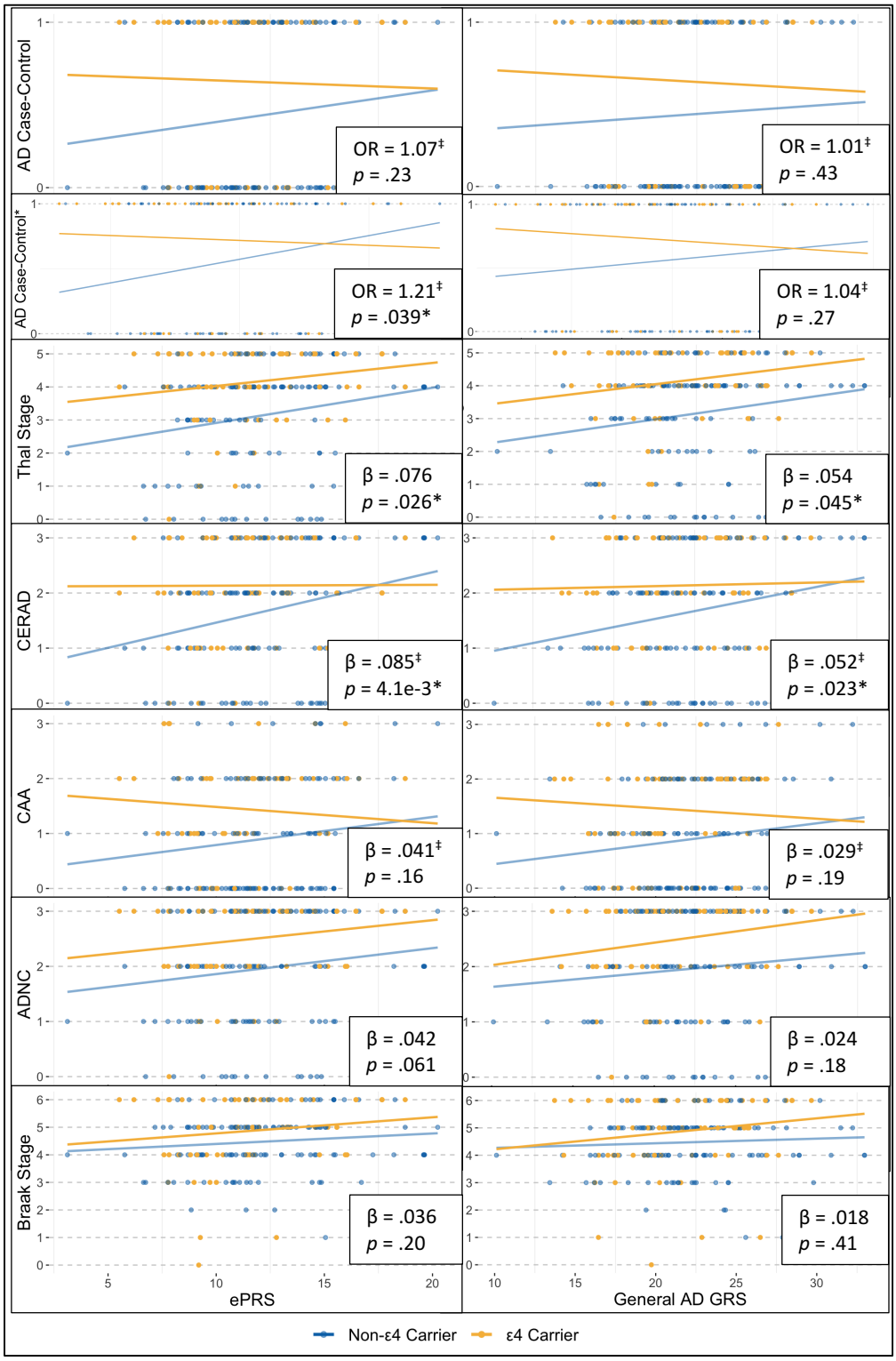


Figure 2. GRS-AD phenotype correlations stratified by ε4 carrier status. ‡ indicates inclusion of an interaction term between *APOE* ε4 Carrier Status and risk score in regression model. AD Case-Control* indicates regression with controls with cognitive testing within 2 years of death. * Indicates a significant correlation. Case-control status: 1 indicates AD Case, 0 = Control.

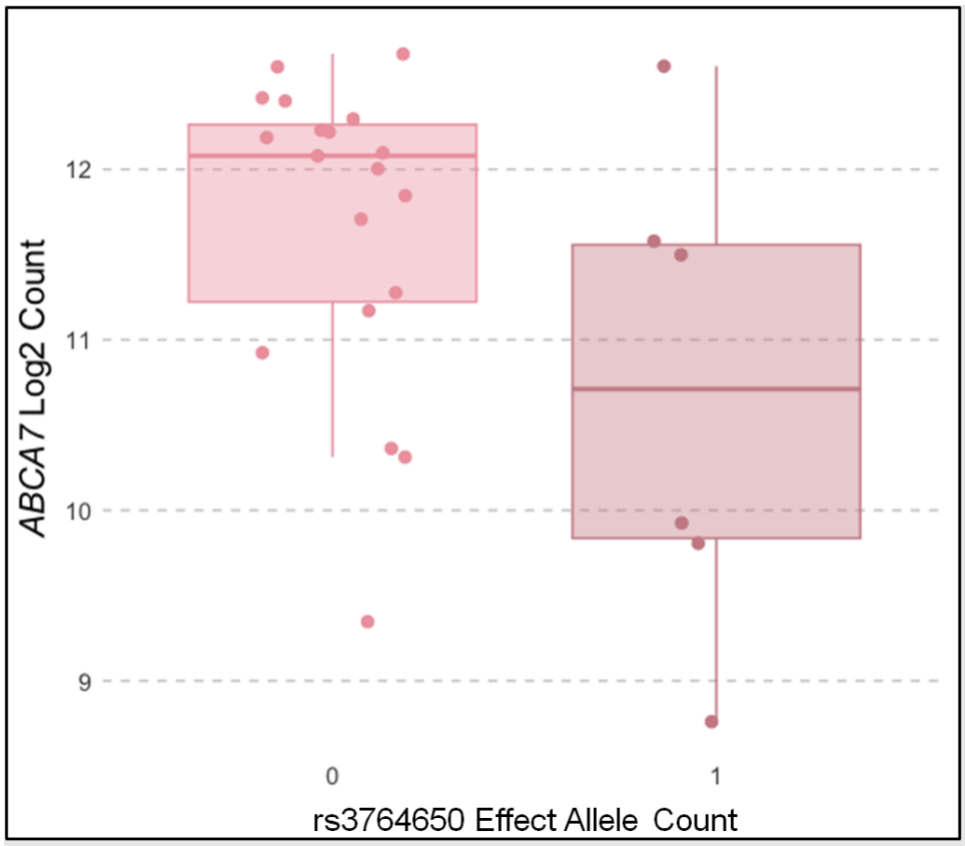


Figure 3. *ABCA7* Log2 counts by rs3764650 count.

Supplementary Tables

Supplementary Table 1. PRS weights used to calculate ePRS and general AD GRS. Bp37: SNP position on GRCh37. OR: Odds ratio. EAF: the effect allele frequencies in the 1000 Genomes European reference (1000 Genomes Project Consortium, 2015). leadSNP: specifies which SNP a proxy SNP represents. r^2: linkage disequilibrium between lead SNP and proxy SNP. Weight: final SNP weight used in GRS calculations.

chr	bp37	rsID	gene	ref	alt	effect	OR	EAF	leadSNP	r^2	weight
1	207692049	rs6656401	*CR1*	A	G	A	1.18	0.20	.	.	0.040393307
2	127892810	rs6733839	*BIN1*	C	T	T	1.22	0.39	.	.	0.059626768
2	127894615	rs744373	*BIN1*	A	G	G	1.24	0.30	.	.	0.060005119
2	234068476	rs35349669	*INPP5D*	C	T	T	1.08	0.50	.	.	0.023634043
5	88223420	rs190982	*MEF2C*	G	A	A	1.075	0.62	.	.	0.02158055
6	32578530	rs9271192	*HLA-DRB5*	C	A	C	1.11	0.28	.	.	0.028652153
6	47487762	rs10948363	*CD2AP*	A	G	G	1.1	0.27	.	.	0.025899889
7	37841534	rs2718058	*NME8*	A	G	A	1.075	0.64	.	.	0.021362993
7	100004446	rs1476679	*ZCWPW1*	C	T	T	1.099	0.70	.	.	0.026454806
7	143110762	rs11771145	*EPHA1*	G	A	G	1.111	0.64	.	.	0.031072513
8	27195121	rs28834970	*PTK2B*	T	C	C	1.10	0.36	.	.	0.027629424
8	27467686	rs9331896	*CLU*	C	T	T	1.163	0.60	.	.	0.045342593
11	47557871	rs10838725	*CELF1*	T	C	C	1.08	0.30	.	.	0.021623737
11	59923508	rs983392	*MS4A6A*	A	G	A	1.111	0.60	.	.	0.031688442
11	59971795	rs670139	*MS4A4E*	G	T	T	1.08	0.43	.	.	0.023383003
11	59997666	rs600550	*MS4A4E*	C	T	T	.	0.43	rs670139	0.87	0.020343212
11	85867875	rs10792832	*PICALM*	A	G	G	1.149	0.65	.	.	0.040886814
11	121435587	rs11218343	*SORL1*	T	C	T	1.299	0.96	.	.	0.031050318
14	53400629	rs17125944	*FERMT2*	T	C	C	1.14	0.095	.	.	0.023608874
14	92926952	rs10498633	*SLC24A4*	G	T	G	1.099	0.77	.	.	0.024199565
18	29088958	rs8093731	*DSG2*	C	T	C	1.370	0.99	.	.	0.019524944
18	29143724	rs139544178	*DSG2*	G	A	G	.	0.98	rs8093731	1.0	0.019524944
19	1046520	rs3764650	*ABCA7*	T	G	G	1.37	0.084	.	.	0.053659599
19	1063443	rs4147929	*ABCA7*	A	G	A	1.15	0.171	.	.	0.03231942
19	1056492	rs3752246	*ABCA7*	G	C	G	.	0.17	rs4147929	0.88	0.02844109
19	51727962	rs3865444	*CD33*	C	A	C	1.064	0.69	.	.	0.017745001
20	55018260	rs7274581	*CASS4*	T	C	T	1.14	0.91	.	.	0.02222644

Supplementary Table 2. Gene ontology terms to define pathway assignments of PRS genes based on Monarch Query by Elizabeth Blue, Ph.D. (Mungall *et al.*, 2016). Chr:chromosome, rsID(s): marker identifier.

chr	rsID(s)	gene	Pathway	Gene Ontology Terms
1	rs6656401	*CR1*	Microglial	complement activation, classical pathway, positive regulation of serine-type endopeptidase activity, ficolin-1-rich granule membrane, negative regulation of serine-type endopeptidase activity, regulation of regulatory T cell differentiation, complement component C4b binding, complement receptor mediated signaling pathway, secretory granule membrane, complement component C4b receptor activity, complement component C3b binding, viral entry into host cell, negative regulation of complement activation, classical pathway, regulation of complement activation, virus receptor activity, cell surface, neutrophil degranulation, plasma membrane, innate immune response, extracellular exosome, negative regulation of complement activation, alternative pathway, integral component of plasma membrane, complement component C3b receptor activity
2	rs6733839, rs744373	*BIN1*	Endolysosomal	regulation of neuron differentiation, muscle cell differentiation, identical protein binding, T-tubule, nuclear envelope, membrane, actin filament binding, lipid tube assembly, RNA polymerase binding, endocytosis, cytoplasm, Z disc, actin cytoskeleton, axon, nucleus localization, lipid tube, cytosol, nucleus organization, cell proliferation, positive regulation of apoptotic process, regulation of cell cycle arrest, positive regulation of astrocyte differentiation, I band, membrane organization, regulation of endocytosis, node of Ranvier, protein binding, viral process, axon initial segment, tau protein binding
2	rs35349669	*INPP5D*	Microglial	negative regulation of neutrophil differentiation, inositol phosphate metabolic process, cytosol, apoptotic process, leukocyte migration, T cell receptor signaling pathway, inositol-1,3,4,5-tetrakisphosphate 5-phosphatase activity, membrane raft, determination of adult lifespan, inositol-polyphosphate 5-phosphatase activity, protein binding, positive regulation of B cell differentiation, phosphatidylinositol biosynthetic process, negative regulation of osteoclast differentiation, positive regulation of apoptotic process, phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase activity, negative regulation of monocyte differentiation, plasma membrane, negative regulation of immune response, immunoglobulin mediated immune response, phosphatidylinositol-3,4,5-trisphosphate 5-phosphatase activity, intracellular signal transduction, negative regulation of bone resorption, phosphate-containing compound metabolic process, phosphatidylinositol dephosphorylation, SH3 domain binding, negative regulation of B cell proliferation, negative regulation of interleukin-6 biosynthetic process, positive regulation of erythrocyte differentiation, negative regulation of signal transduction, signal transduction, cytoskeleton, PTB domain binding

5	rs190982	*MEF2C*	Endolysosomal	RNA polymerase II proximal promoter sequence-specific DNA binding, renal tubule morphogenesis, outflow tract morphogenesis, chondrocyte differentiation, positive regulation of skeletal muscle tissue development, negative regulation of gene expression, regulation of megakaryocyte differentiation, transcriptional activator activity, RNA polymerase II distal enhancer sequence-specific DNA binding, nuclear speck, regulation of AMPA receptor activity, osteoblast differentiation, regulation of neurotransmitter secretion, learning or memory, smooth muscle cell differentiation, positive regulation of B cell proliferation, chromatin binding, apoptotic process, embryonic viscerocranium morphogenesis, heart looping, RNA polymerase II transcription factor activity, sequence-specific DNA binding, positive regulation of skeletal muscle cell differentiation, activating transcription factor binding, histone deacetylase binding, positive regulation of myoblast differentiation, positive regulation of cardiac muscle cell differentiation, cellular response to calcium ion, cytoplasm, neuron differentiation, primary heart field specification, positive regulation of behavioral fear response, nephron tubule epithelial cell differentiation, RNA polymerase II distal enhancer sequence-specific DNA binding, postsynapse, humoral immune response, blood vessel development, B cell receptor signaling pathway, positive regulation of macrophage apoptotic process, blood vessel remodeling, cartilage morphogenesis, cellular response to parathyroid hormone stimulus, heart development, transcription regulatory region DNA binding, B cell proliferation, regulation of synapse assembly, positive regulation of gene expression, regulation of synaptic transmission, glutamatergic, nervous system development, melanocyte differentiation, positive regulation of osteoblast differentiation, nucleus, positive regulation of bone mineralization, negative regulation of epithelial cell proliferation, excitatory postsynaptic potential, HMG box domain binding, muscle organ development, cell morphogenesis involved in neuron differentiation, regulation of synaptic activity, muscle cell fate determination, protein complex, negative regulation of transcription by RNA polymerase II, monocyte differentiation, glomerulus morphogenesis, negative regulation of neuron apoptotic process, skeletal muscle cell differentiation, cardiac ventricle formation, positive regulation of neuron differentiation, regulation of germinal center formation, positive regulation of cardiac muscle cell proliferation, cellular response to fluid shear stress, epithelial cell proliferation involved in renal tubule morphogenesis, sarcoplasm, cardiac muscle hypertrophy in response to stress, regulation of NMDA receptor activity, positive regulation of transcription, DNA-templated, regulation of neuron apoptotic process, platelet formation, transcription factor activity, RNA polymerase II core promoter sequence-specific DNA binding, ventricular cardiac muscle cell differentiation, neuron migration, germinal center formation, secondary heart field specification, core promoter sequence-specific DNA binding, skeletal muscle tissue development, transcription from RNA polymerase II promoter, regulation of transcription, DNA-templated, intracellular membrane-bounded organelle, protein binding, myotube differentiation, nucleoplasm, endochondral ossification, neuron development, RNA polymerase

				II regulatory region sequence-specific DNA binding, cellular response to drug, transcriptional activator activity, RNA polymerase II proximal promoter sequence-specific DNA binding, positive regulation of transcription from RNA polymerase II promoter, regulation of sarcomere organization, B cell homeostasis, protein heterodimerization activity, sinoatrial valve morphogenesis, regulation of dendritic spine development, positive regulation of muscle cell differentiation, cellular response to transforming growth factor beta stimulus, positive regulation of protein homodimerization activity, regulation of synaptic plasticity, positive regulation of cell proliferation in bone marrow, positive regulation of alkaline phosphatase activity, MAPK cascade, cellular response to trichostatin A, AT DNA binding, negative regulation of ossification, cellular response to lipopolysaccharide, roof of mouth development, neural crest cell differentiation, DNA binding transcription factor activity
6	rs9271192	*HLA-DRB5*	Microglial	membrane, Golgi membrane, antigen processing and presentation of exogenous peptide antigen via MHC class II, trans-Golgi network membrane, T cell costimulation, MHC class II protein complex, extracellular exosome, late endosome membrane, endocytic vesicle membrane, cell surface, ER to Golgi transport vesicle membrane, polysaccharide assembly with MHC class II protein complex, MHC class II protein complex binding, integral component of lumenal side of endoplasmic reticulum membrane, transport vesicle membrane, lysosomal membrane, clathrin-coated endocytic vesicle membrane, interferon-gamma-mediated signaling pathway, peptide antigen binding, plasma membrane, T cell receptor signaling pathway
6	rs10948363	*CD2AP*	Endolysosomal	vascular endothelial growth factor receptor binding, cell division, negative regulation of transforming growth factor beta1 production, perinuclear region of cytoplasm, actin cytoskeleton, cell-cell adhesion, beta-catenin binding, positive regulation of protein localization to nucleus, endocytic vesicle, actin filament organization, protein complex assembly, signal transduction, ruffle, filamentous actin, substrate-dependent cell migration, cell extension, plasma membrane, SH3 domain binding, cytoplasm, extracellular exosome, cell cycle, cadherin binding, vesicle organization, regulation of actin cytoskeleton reorganization, regulation of receptor-mediated endocytosis, structural constituent of cytoskeleton, proteasome-mediated ubiquitin-dependent protein catabolic process, cell-cell junction, protein binding, negative regulation of small GTPase mediated signal transduction, protein C-terminus binding
7	rs2718058	*NME8*	Neuronal	sperm cytoplasmic droplet, nucleoside diphosphate kinase activity, GTP biosynthetic process, outer dynein arm, flagellated sperm motility, multicellular organism development, sperm principal piece, cellular response to reactive oxygen species, nucleoside diphosphate phosphorylation, CTP biosynthetic process, UTP biosynthetic process, cilium assembly, microtubule binding, spermatogenesis, cell redox homeostasis, cell differentiation
7	rs1476679	*ZCWPW1*	--	zinc ion binding

7	rs11771145	*EPHA1*	Endolysosomal, Microglial	regulation of GTPase activity, negative regulation of cell migration, angiogenesis, ATP binding, positive regulation of cell-matrix adhesion, somatic stem cell population maintenance, peptidyl-tyrosine phosphorylation, protein autophosphorylation, cell surface receptor signaling pathway, integral component of plasma membrane, protein kinase binding, positive regulation of cell migration, activation of GTPase activity, ephrin receptor signaling pathway, negative regulation of protein kinase activity, positive regulation of cell proliferation, positive regulation of angiogenesis, plasma membrane, positive regulation of stress fiber assembly, transmembrane-ephrin receptor activity, substrate adhesion-dependent cell spreading, protein kinase activity
8	rs28834970	*PTK2B*	Endolysosomal	epidermal growth factor receptor signaling pathway, bone resorption, angiogenesis, MAPK cascade, lamellipodium, blood vessel endothelial cell migration, positive regulation of cell migration, positive regulation of ERK1 and ERK2 cascade, positive regulation of JUN kinase activity, apoptotic process, protein phosphorylation, nucleus, protein complex assembly, regulation of release of sequestered calcium ion into cytosol, positive regulation of neuron projection development, positive regulation of angiogenesis, protein tyrosine kinase activity, cytoskeleton, interleukin-7-mediated signaling pathway, regulation of NMDA receptor activity, signaling receptor binding, long-term synaptic potentiation, protein autophosphorylation, positive regulation of nitric-oxide synthase activity, vascular endothelial growth factor receptor signaling pathway, ATP binding, peptidyl-tyrosine autophosphorylation, signal transducer activity, growth cone, positive regulation of peptidyl-tyrosine phosphorylation, positive regulation of excitatory postsynaptic potential, activation of GTPase activity, peptidyl-tyrosine phosphorylation, positive regulation of ubiquitin-dependent protein catabolic process, negative regulation of bone mineralization, protein binding, glial cell proliferation, cellular response to retinoic acid, response to hypoxia, sprouting angiogenesis, positive regulation of phosphatidylinositol 3-kinase activity, innate immune response, negative regulation of myeloid cell differentiation, cytosol, ionotropic glutamate receptor signaling pathway, oocyte maturation, stress fiber assembly, extrinsic component of cytoplasmic side of plasma membrane, response to lithium ion, dendrite, integrin-mediated signaling pathway, regulation of establishment of cell polarity, positive regulation of endothelial cell migration, regulation of cell adhesion, positive regulation of nitric oxide biosynthetic process, cell cortex, response to osmotic stress, response to ethanol, response to hormone, signal complex assembly, response to calcium ion, non-membrane spanning protein tyrosine kinase activity, positive regulation of translation, cytoplasm, response to immobilization stress, marginal zone B cell differentiation, negative regulation of cell proliferation, positive regulation of cell growth, response to hydrogen peroxide, response to mechanical stimulus, regulation of macrophage chemotaxis, positive regulation of cytosolic calcium ion concentration, postsynaptic density, cellular response to fluid shear stress, signal transduction, positive regulation of cell-matrix adhesion, cell body,

				response to cocaine, focal adhesion, regulation of cGMP biosynthetic process, cell surface receptor signaling pathway, regulation of cell shape, positive regulation of protein kinase activity, negative regulation of apoptotic process, response to glucose, apical dendrite, neuron projection development, 3-phosphoinositide-dependent protein kinase binding, membrane raft, NMDA selective glutamate receptor complex, regulation of synaptic plasticity, activation of Janus kinase activity, regulation of calcium-mediated signaling, response to stress, dendritic spine, positive regulation of synaptic transmission, glutamatergic, positive regulation of cell proliferation, long term synaptic depression, regulation of ubiquitin-dependent protein catabolic process, axon, positive regulation of DNA biosynthetic process, regulation of cGMP-mediated signaling, negative regulation of muscle cell apoptotic process, regulation of inositol trisphosphate biosynthetic process, cellular defense response, positive regulation of B cell chemotaxis, regulation of actin cytoskeleton reorganization, negative regulation of potassium ion transport, tumor necrosis factor-mediated signaling pathway, focal adhesion assembly, chemokine-mediated signaling pathway, adaptive immune response, neuronal cell body, calmodulin-dependent protein kinase activity, ubiquitin protein ligase binding, positive regulation of actin filament polymerization, negative regulation of neuron apoptotic process, response to cAMP, positive regulation of JNK cascade, perinuclear region of cytoplasm
8	rs9331896	*CLU*	Endolysosomal, Microglial	response to misfolded protein, ubiquitin protein ligase binding, response to virus, positive regulation of tau-protein kinase activity, innate immune response, platelet alpha granule lumen, protein binding, protein stabilization, positive regulation of ubiquitin-dependent protein catabolic process, positive regulation of amyloid fibril formation, misfolded protein binding, negative regulation of cellular response to tunicamycin, central nervous system myelin maintenance, cytoplasm, apical dendrite, protein complex, regulation of amyloid-beta clearance, chaperone-mediated protein transport involved in chaperone-mediated autophagy, positive regulation of tumor necrosis factor production, nucleus, microglial cell activation, extracellular matrix, mitochondrion, cell surface, blood microparticle, microglial cell proliferation, positive regulation of proteasomal ubiquitin-dependent protein catabolic process, cell morphogenesis, protein targeting to lysosome involved in chaperone-mediated autophagy, negative regulation of response to endoplasmic reticulum stress, chromaffin granule, positive regulation of gene expression, extracellular exosome, positive regulation of receptor-mediated endocytosis, cell periphery, regulation of neuron death, antimicrobial humoral response, negative regulation of amyloid-beta formation, amyloid-beta binding, spherical high-density lipoprotein particle, reverse cholesterol transport, chaperone binding, negative regulation of protein homooligomerization, complement activation, classical pathway, chaperone-mediated protein folding, extracellular space, negative regulation of cellular response to thapsigargin, platelet degranulation, positive regulation of protein homooligomerization, mitochondrial membrane, positive regulation of amyloid-beta

				formation, positive regulation of NF-kappaB transcription factor activity, perinuclear region of cytoplasm, lipid metabolic process, release of cytochrome c from mitochondria, regulation of complement activation, Golgi apparatus, complement activation, protein import, chaperone-mediated protein complex assembly, positive regulation of neurofibrillary tangle assembly, endoplasmic reticulum, negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage, extracellular region, low-density lipoprotein particle receptor binding, positive regulation of neuron death, negative regulation of amyloid fibril formation, positive regulation of nitric oxide biosynthetic process, intracellular, negative regulation of cell death, cytosol, regulation of neuronal signal transduction, neurofibrillary tangle, negative regulation of release of cytochrome c from mitochondria
11	rs10838725	*CELF1*	--	membrane, intracellular ribonucleoprotein complex, translation repressor activity, mRNA regulatory element binding, negative regulation of translation, regulation of RNA splicing, germ cell development, mRNA processing, nucleoplasm, RNA binding, cytoplasm, protein binding, pre-mRNA binding, mRNA splice site selection, BRE binding, nucleus, mRNA binding, RNA interference, embryo development
11	rs983392	*MS4A6A*	Microglial	integral component of membrane
11	rs670139	*MS4A4E*	Microglial	integral component of membrane
11	rs10792832	*PICALM*	Endolysosomal	protein complex assembly, cell surface, dendrite morphogenesis, cadherin binding, negative regulation of metalloendopeptidase activity involved in amyloid precursor protein catabolic process, perinuclear region of cytoplasm, protein binding, regulation of aspartic-type endopeptidase activity involved in amyloid precursor protein catabolic process, AP-2 adaptor complex, axonogenesis, regulation of vesicle size, Golgi apparatus, phosphatidylinositol-4,5-bisphosphate binding, endosomal transport, cell proliferation, membrane, endosome to plasma membrane transport vesicle, receptor internalization, neuronal cell body, presynaptic membrane, negative regulation of protein localization to cell surface, vesicle-mediated transport, clathrin-coated pit, negative regulation of protein localization to plasma membrane, clathrin-dependent endocytosis, positive regulation of aspartic-type endopeptidase activity involved in amyloid precursor protein catabolic process, clathrin coat of coated pit, positive regulation of transcription, DNA-templated, positive regulation of GTPase activity, modulation of age-related behavioral decline, clathrin-coated endocytic vesicle, clathrin coat assembly, Rab GTPase binding, nucleus, regulation of endocytosis, low-density lipoprotein particle receptor binding, negative regulation of gene expression, intracellular membrane-bounded organelle, positive regulation of amyloid-beta formation, synaptic vesicle maturation, positive regulation of clathrin-dependent endocytosis, clathrin-coated vesicle, postsynaptic membrane, neurofibrillary tangle, cargo loading into vesicle, cytosol, clathrin binding, vesicle organization, early endosome, membrane bending, transcytosis, clathrin heavy

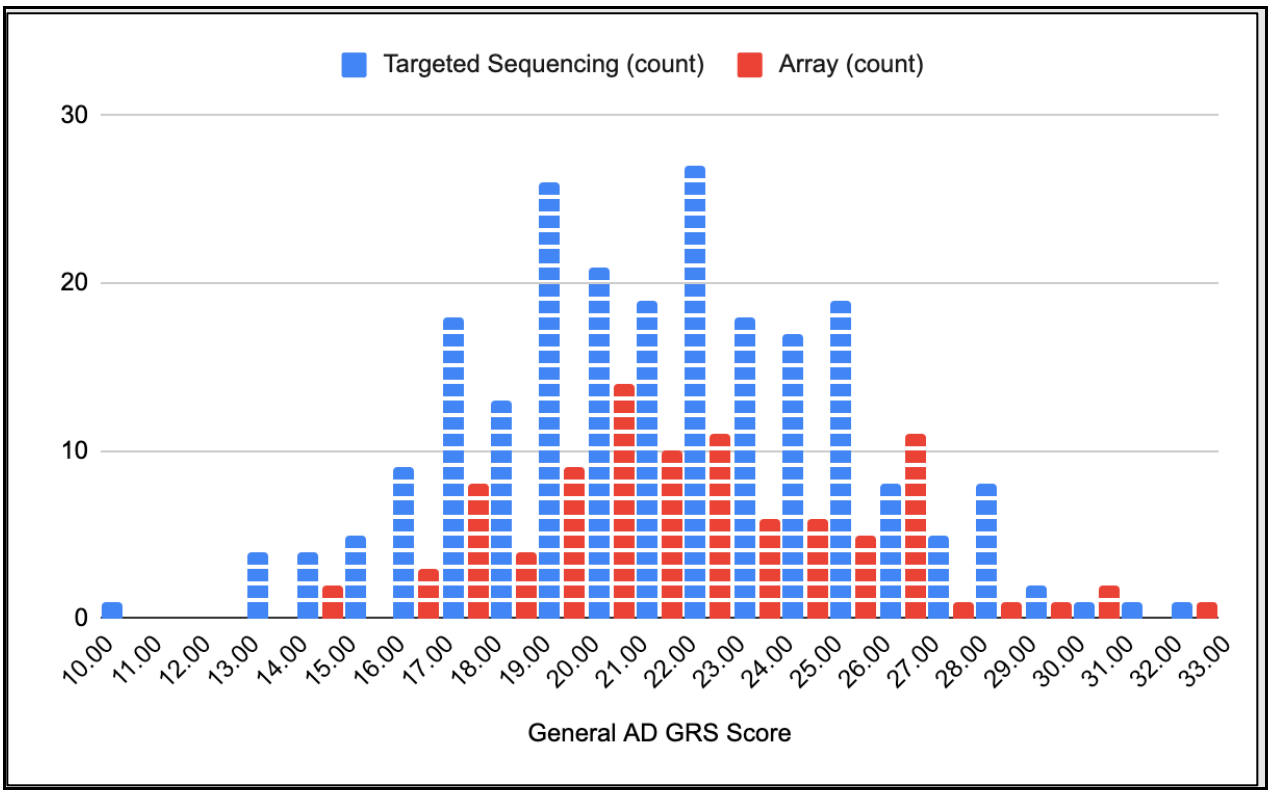
				chain binding, learning or memory, iron ion import across plasma membrane, intrinsic component of membrane, regulation of protein localization, hemopoiesis, negative regulation of receptor-mediated endocytosis, iron ion homeostasis, 1-phosphatidylinositol binding, positive regulation of amyloid-beta clearance, positive regulation of neuron death, vesicle, receptor-mediated endocytosis, clathrin adaptor activity, membrane organization
11	rs11218343	*SORL1*	Endolysosomal	positive regulation of protein catabolic process, lipid transport, ADP-ribosylation factor binding, negative regulation of aspartic-type endopeptidase activity involved in amyloid precursor protein catabolic process, membrane, endosome membrane, negative regulation of metalloendopeptidase activity involved in amyloid precursor protein catabolic process, positive regulation of ER to Golgi vesicle-mediated transport, cellular protein metabolic process, integral component of plasma membrane, positive regulation of protein localization to early endosome, negative regulation of protein binding, recycling endosome, negative regulation of tau-protein kinase activity, positive regulation of endocytic recycling, endoplasmic reticulum, positive regulation of early endosome to recycling endosome transport, Golgi cisterna, extracellular space, trans-Golgi network, signal transduction, protein targeting to lysosome, amyloid-beta binding, regulation of smooth muscle cell migration, negative regulation of neurogenesis, transmembrane signaling receptor activity, low-density lipoprotein particle receptor activity, negative regulation of protein oligomerization, negative regulation of MAP kinase activity, nuclear envelope lumen, extracellular exosome, positive regulation of protein exit from endoplasmic reticulum, negative regulation of neuron death, protein binding, negative regulation of neurofibrillary tangle assembly, protein retention in Golgi apparatus, Golgi membrane, early endosome, cholesterol metabolic process, endosome, Golgi apparatus, low-density lipoprotein particle, negative regulation of amyloid-beta formation, positive regulation of choline O-acetyltransferase activity, protein maturation, post-Golgi vesicle-mediated transport, receptor-mediated endocytosis, low-density lipoprotein particle binding, protein targeting
14	rs17125944	*FERMT2*	Neuronal	substrate adhesion-dependent cell spreading, nucleus, cell junction assembly, transforming growth factor beta receptor signaling pathway, focal adhesion, cell cortex, nucleoplasm, focal adhesion assembly, lamellipodium membrane, extrinsic component of cytoplasmic side of plasma membrane, regulation of cell shape, stress fiber, integrin-mediated signaling pathway, cytoplasm, cell-matrix adhesion, cytosol, I band, filamentous actin, protein localization to membrane, phosphatidylinositol-3,4,5-trisphosphate binding, Wnt signaling pathway, cell surface, integrin activation, protein binding

14	rs10498633	*SLC24A4*	Endolysosomal	memory, integral component of plasma membrane, visual perception, potassium ion transport, protein dimerization activity, calcium ion binding, sodium ion transmembrane transport, calcium channel activity, cadmium ion binding, sodium ion binding, calcium, potassium:sodium antiporter activity, potassium ion binding, long-term synaptic potentiation, nickel cation binding, manganese ion binding, symporter activity, calcium ion transmembrane transport, response to stimulus, plasma membrane, learning, ion transport, ion transmembrane transport, long term synaptic depression, cellular calcium ion homeostasis
18	rs8093731	*DSG2*	Neuronal	cell-cell junction, intracellular membrane-bounded organelle, cell adhesion, cell surface, plasma membrane, bundle of His cell-Purkinje myocyte adhesion involved in cell communication, calcium ion binding, Purkinje myocyte development, desmosome organization, response to progesterone, cornified envelope, desmosome, intercalated disc, keratinization, regulation of heart rate by cardiac conduction, apical plasma membrane, integral component of membrane, cell adhesion molecule binding, cell junction, cornification, maternal process involved in female pregnancy, extracellular exosome, regulation of ventricular cardiac muscle cell action potential, lateral plasma membrane, homophilic cell adhesion via plasma membrane adhesion molecules, cell adhesive protein binding involved in bundle of His cell-Purkinje myocyte communication
19	rs3764650, rs4147929	*ABCA7*	Endolysosomal	transmembrane transport, phospholipid translocation, transporter activity, plasma membrane, positive regulation of engulfment of apoptotic cell, apolipoprotein A-I-mediated signaling pathway, early endosome membrane, peptide cross-linking, positive regulation of amyloid-beta clearance, apolipoprotein A-I receptor activity, phagocytosis, ATP binding, cell junction, positive regulation of phospholipid efflux, positive regulation of ERK1 and ERK2 cascade, negative regulation of amyloid-beta formation, protein localization to nucleus, ruffle membrane, phagocytic cup, positive regulation of phagocytosis, positive regulation of cholesterol efflux, phosphatidylcholine-translocating ATPase activity, phospholipid efflux, intracellular membrane-bounded organelle, ATPase activity, coupled to transmembrane movement of substances, memory, phosphatidylserine-translocating ATPase activity, cholesterol efflux, Golgi apparatus, ATP-binding cassette (ABC) transporter complex, ATPase activity, Golgi membrane, negative regulation of amyloid precursor protein biosynthetic process, integral component of membrane, cell surface, high-density lipoprotein particle assembly
19	rs3865444	*CD33*	Endolysosomal, Microglial	negative regulation of cell proliferation, signal transduction, neutrophil degranulation, carbohydrate binding, receptor activity, tertiary granule membrane, regulation of immune response, extracellular exosome, protein binding, integral component of plasma membrane, specific granule membrane, cell-cell signaling, cell adhesion, nucleus, external side of plasma membrane, plasma membrane
20	rs7274581	*CASS4*	Neuronal	cytoplasm, cell adhesion, focal adhesion, cytoskeleton

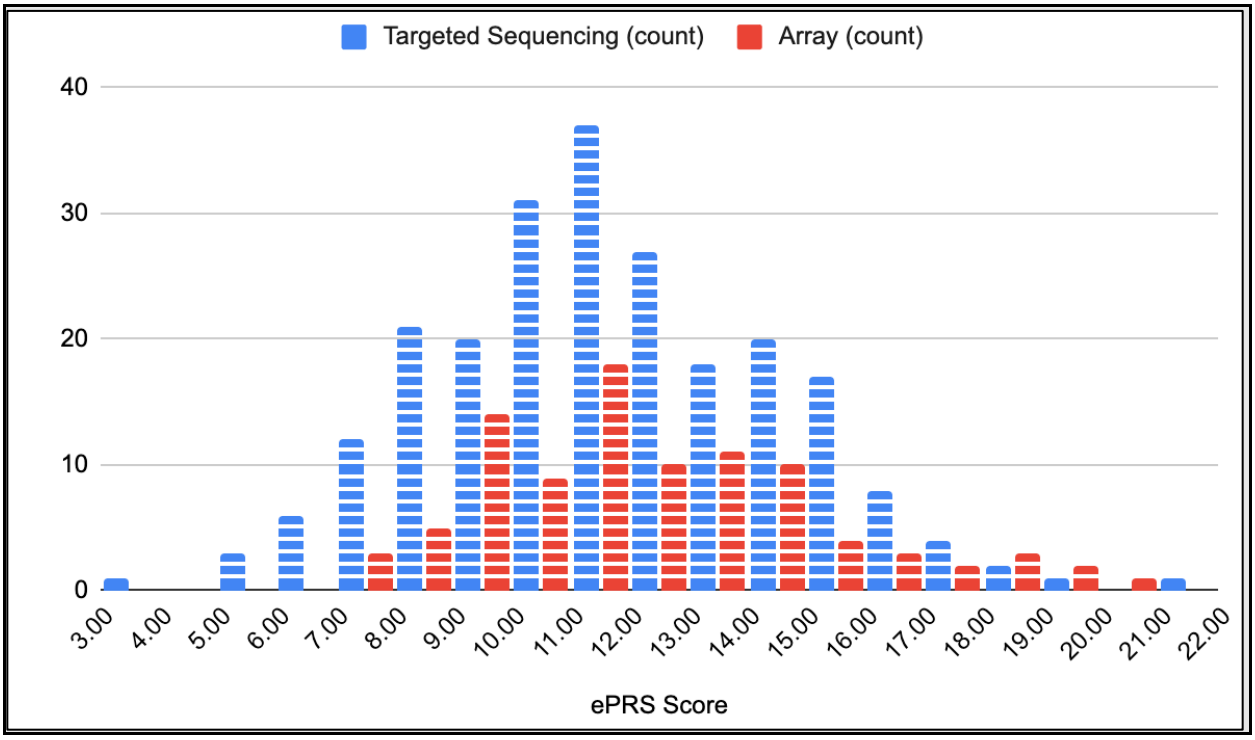
Supplementary Table 3. Mean and standard deviation FeatureCounts Read Numbers, RNA Quality, and Sample Collection Metrics for bulk RNA sequencing data. *Assigned*: reads mapped to a feature; *Unassigned Unmapped*: reads are reported as unmapped in SAM/BAM input; *Unassigned MultiMapping:* reads that map to more than one location in the genome; *Unassigned No Features:* not overlapping with any features included in the annotation; *Unassigned Ambiguity*: overlapping with two or more features; *Imputed RIN*: *RNA Integrity Number*; *PMI Hours*: Post-Mortem Interval hours prior to sample collection; Yield mBases: Megabases sequenced; *Mean Quality Score*: Q = -10log10(e); e = estimated probability of base call being wrong; *%Q30*: Percentage of bases with a Phred Quality Score of 30 or higher.

	High ePRS (N = 13)	Low ePRS (N = 13)	Total (N = 26)
Read Number	68,166,285 (7,263,497)	71,383,453 (7,445,327)	69,774,869 (7,390,728)
Assigned	104,692,574 (10,502,963)	108,782,944 (10,741,384)	106,737,759 (10,615,135)
Unassigned Unmapped	9,186,336 (1,217,219)	8,899,810 (949,603)	9,043,073 (1,079,519)
Unassigned MultiMapping	24,181,609 (5,524,832)	33,386,600 (16,941,992)	28,784,104 (13,208,202)
Unassigned No Features	5,179,143 (1,943,783)	5,218,008 (1,954,327)	5,198,575 (1,909,786)
Unassigned Ambiguity	11,188,047 (1,446,980)	12,393,494 (1,618,488)	11,790,770 (1,624,859)
Yield mBases	20,450 (2,179)	21,415 (2,234)	20,932 (2,217)
Mean Quality Score	35.4 (0.1)	35.5 (0.1)	35.4 (0.1)
%Q30	90.9 (0.7)	91.6 (0.4)	91.3 (0.7)
Imputed RIN	9.5 (0.4)	9.4 (0.7)	9.4 (0.6)
PMI Hours	9 (4.9)	6.1 (3.6)	7.5 (4.4)

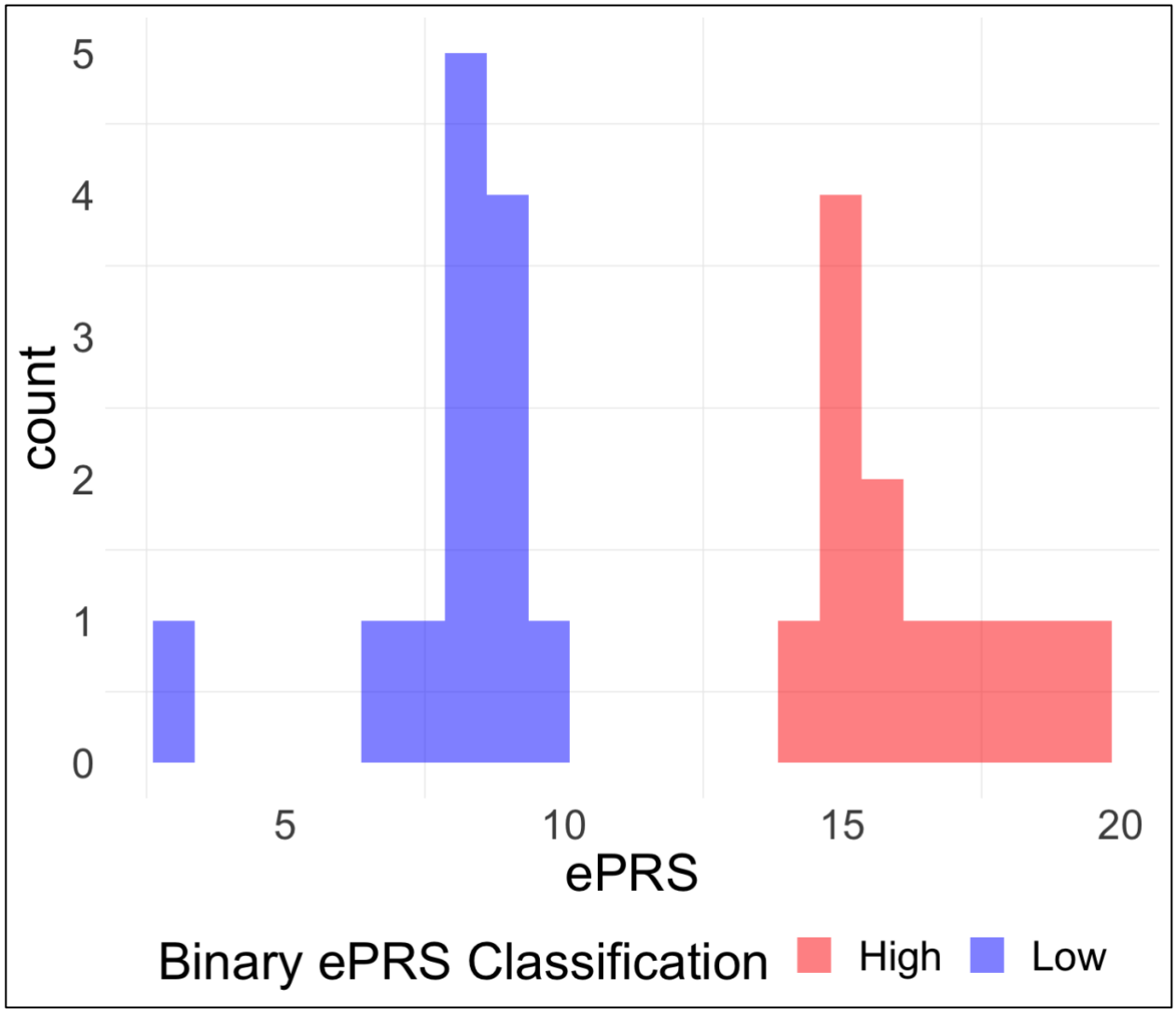
Supplementary Figures



Supplementary Figure 1a. General AD GRS by Genotyping Type.



Supplementary Figure 1b. ePRS by Genotyping Type.



Supplementary Figure 2. ePRS score among samples used in Bulk RNAseq.