Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Jaime Schechner April 2, 2021

Alzheimer's Disease Associated Genomic Variants Identified in White Populations are Present in African Americans at Risk for Alzheimer's Disease Due to Family History

By

Jaime Schechner

Whitney Wharton, PhD.
Adviser

Neuroscience and Behavioral Biology

Whitney Wharton, PhD.

Adviser

Leah Roesch, PhD.

Committee Member

Tracy Scott, PhD.

Committee Member

2021

Alzheimer's Disease Associated Genomic Variants Identified in White Populations are Present in African Americans at Risk for Alzheimer's Disease Due to Family History

By

Jaime Schechner

Whitney Wharton, PhD.

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2021

Abstract

Alzheimer's Disease Associated Genomic Variants Identified in White Populations are Present in African Americans at Risk for Alzheimer's Disease Due to Family History

By Jaime Schechner

Importance: The underlying genetic variants that contribute to the neurodevelopment of familial Alzheimer's Disease (AD) is poorly understood. Furthermore, most studies on the genetics of AD have been limited to White cohorts; therefore, exploring the prevalence of variants between White and African American populations can help determine if White derived SNPs are applicable to non-white cohorts. Further, comparing the prevalence of variants within cognitively normal White and African American individuals with familial history to those diagnosed with Alzheimer's has the potential to identify "at-risk" genes.

Objective: To investigate the prevalence of known and recently identified single nucleotide polymorphisms (SNPs) among African American and White individuals either diagnosed with AD or at risk via parental history.

Design/Setting: Observational and longitudinal multicenter study.

Participants: 160 cognitively normal and 55 Alzheimer's Disease diagnosed African American and White individuals with parental history of Alzheimer's disease.

Main Outcome and Measure: Genotyped whole blood samples

Results: 97.8% of the SNPs screened should be viable for screening in African American cohorts. Further, we report no significant results between the at-risk and Alzheimer's diagnosed cohorts, suggesting that the genetic markers of AD are complex and not one gene contributes to the neurodevelopment of AD.

Conclusion and Relevance: Alzheimer's-associated variants identified in White populations may be applicable to assess the development of familial Alzheimer's Disease within other populations such as African Americans, but all variants must be validated at a per SNP level. Furthermore, our results emphasize the complexity of Alzheimer's Disease and how its development and progression is dependent on genetic, environmental, and lifestyle factors.

Alzheimer's Disease Associated Genomic Variants Identified in White Populations are Present in African Americans at Risk for Alzheimer's Disease Due to Family History

By

Jaime Schechner

Whitney Wharton
Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2021

Acknowledgements

I would like to thank Dr. Whitney Wharton for all of her guidance and support throughout my time in her laboratory. I would also like to thank Dr. Dominika Swieboda for her explanations on PLINK and for her continuous assistance throughout my thesis process.

Table of Contents

Introduction 1
Methods 5
Study Sample: 5
Genotyping: 6
Data Analysis: 6
Results 9
Participant Demographics 9
SNP Summary 10
Race Association Analysis 10
Diagnosis Association Analysis 11
Discussion 13
References 19
Tables 30
Table 1: Patient Demographics 30
Table 2: SNP Summary For All Screened Variants. 31
Table 3: Race Association Analysis Results 34
Table 4: Race Genotyping Results 36
Table 5: Diagnosis Association Analysis Results 40
Table 6: Diagnosis Genotyping Results 42
Supplementary 45
Supplementary File One: PLINK Code 45

<u>Introduction</u>

Alzheimer's disease (AD) is a progressive neurodegenerative disease characterized by memory impairment and cognitive decline, which ultimately leads to deficits in executive functioning, alterations to personality and behavior, and death (DeTure and Dickson, 2019, Alz.org, 2020). Estimates suggest that more than 5 million people within the United States are currently diagnosed with AD and this number is projected to reach 13.8 million by 2050 (Alz.org, 2020). AD is among the top 10 global causes of death (World Health Organization, 2018) and the only disease within the top 10 without a cure or proven disease modifying therapy (Cummings and Fox, 2017). In order to progress the search for a cure, it is paramount that researchers understand the genetics and biological mechanisms that contribute to the neurodevelopment of AD.

Neurofibrillary tangles and beta-amyloid plaques are the pathological hallmarks for AD (National Institute on Aging, 2017). Neurofibrillary tangles, composed of abnormally phosphorylated tau proteins adhering to other tau proteins (Perl, 2010), and beta-amyloid plaques block the transport system of essential molecules within neurons leading to the neurodegeneration (Perl, 2010) and interfere with cell synaptic transmission (Robertson, 2018), respectively. In individuals diagnosed with AD, it is believed that amyloid and tau proteins begin to undetectably accumulate in the brain 10–20 years before the onset of clinical symptoms (Bateman et al., 2012, Beason-Held et al., 2013).

However, despite this understanding of AD pathology, identifying those at risk for AD has proven elusive. Indeed, researchers have spent significant resources investigating the genetics that contribute to the etiology of AD and identified three causative genes associated with mendelian forms of AD, contributing to less than 1% of AD cases (Bekris et al., 2010). These genes include: APP on chromosome 21, PSEN1 on chromosome 14, and PSEN2 on chromosome 1 (Loy et al.,

2014, and Berkis et al., 2010). APP has been mapped to chromosome 21 and accounts for previous observations that patients with Down syndrome (trisomy 21) develop significant amyloid plaques and other neuropathological features of AD when in their 40s (Berkis et al., 2010). Today, more than 32 different APP missense mutations have been identified and contribute to 15% of early-onset familial (EOF) AD (Bekris et al., 2010). Both Presenilin (PSEN) 1 and 2, located on chromosome 14 and 1 respectively, have been found to increase amyloid-beta production by increasing the amount of γ-secretase cleavage of APP. Previous research has also shown that PSEN1 missense mutations account for 18% to 50% of autosomal dominant EOFAD cases. (Bekris et al., 2010).

Apolipoprotein E (APOE) is another known genetic determinant of AD risk. Located on chromosome 19, APOE is a cholesterol carrier protein that facilitates lipid transport and injury repair within the brain (Liu et al., 2013). APOE has various polymorphic alleles that differ in AD risk. Specifically, carriers of the ε4 allele are at an increased risk of AD as the ε4 allele has been found to strongly affect the deposition of amyloid-beta plaques (Berkis et al., 2010). In contrast, carriers of the ε2 allele are at a decreased risk for AD (Liu et al., 2013).

Nonetheless, the genetics that further contribute to familial AD (FAD) or the sporadic development of AD remains elusive. A variety of genome wide association studies (GWAS) and genome wide significance studies (GWS) have been conducted trying to identify novel genes implicated in the neurodevelopment of AD (Xiao et al., 2020). These studies have identified over 75 AD risk associated loci (Bellenguez et al., 2021). However, studies limit evaluation to individuals diagnosed with AD thereby limiting power to identify genes associated with progression to AD by individuals at risk by virtue of family history. Studies comparing individuals

with normal cognition and family history of AD to individuals with AD and family history could therefore identify drivers of disease, protective elements, or interventions.

Furthermore, a majority of the GWAS and GWS studies performed to date evaluate predominantly White participants, or participants with European ancestry. As of 2017, there were only six US population-based publications and zero GWAS studies regarding AD incidence by race (Steenland et al., 2016). Further, the cohorts within the six studies had a greater representation of White participants compared to non-white, African Americans, and Hispanic participants. Focusing on White individuals to the exclusion of others may be a disservice to the greater community if the genetics that contribute to the neurodevelopment of AD differ by race or origin.

While race is a social construct that affects daily life including access to resources, healthcare, and both acute and chronic stress/anxiety, and in turn may impact biological health processes (American Psychological Association, 2017, Omi and Winant, 1994), there may be important interactions between race and genetic characteristics in disease susceptibility. Our group and others view race as a potential confound in the development of AD due to the impact race can have on genetic AD risk factors, co-morbid cerebrovascular and systemic diseases, socioeconomic status, and psychosocial inequity (Wharton et al., 2019). However, studies examining the racial differences in molecular biomarkers for AD have produced conflicting results (Morris, Schindler, and McCue, 2019). For example, studies have suggested an increased prevalence for AD in Black and African Americans compared to White individuals (Tang et al., 2001, Green et al., 2001 and Mayeda et al., 2016) while other publications have found no differences between race and risk for AD (Fillenbaum et al., 1998, and Fitzpatrick et al., 2004). Due to the paucity of genomic data in Black or African American individuals with AD, it is difficult to determine if there is a genetic foundation to the differences found in some groups, or if reported differences in AD are due to the

sociologically consequences of historic and current racial social determinates of health. Evaluating the genomic risk factors contributing to AD and determining if these are equally represented and function similarly in Black and African American populations, is therefore of paramount importance to developing cures and therapies which are equitable.

To address the literature gap regarding the genetics that contribute to FAD and the conflicting results discussing race and risk of AD, we investigate the prevalence of known and recently identified single nucleotide polymorphisms (SNPs) among two previous published cohorts (Kumar et al., 2020 and Weiner et al., 2012) of African American and White individuals either diagnosed with AD or at risk via parental history. The goals of this study are two-fold: (1) evaluating the prevalence of gene variations that may contribute to the neuropathology of AD in individuals with normal cognition and family history of AD versus individuals with AD and family history and (2) comparing how the prevalence of known risk factors differs between White and African American individuals with familial history of AD.

Methods

Study Sample:

Two distinct cohorts were utilized to examine how SNP prevalence and self-identified race implicate the neuropathology and development of AD. The first cohort consisted of eighty-two subjects enrolled in ASCEND (Association Between Cardiovascular Risk and Preclinical Alzheimer's Disease Pathology), a completed two-year observational research study (PI Wharton). ASCEND participants were recruited from the Emory Alzheimer's Disease Research Center (ADRC) and were eligible for enrollment if they were 45 years or older and had a parent with AD confirmed either by autopsy or defined by NINDS-ADRDA criteria. Under ASCEND IRB approved protocol, subjects received a blood draw, cognitive testing, medical and medication history, lumbar puncture, vascular ultrasound, and MRI at Baseline, Year 1, and Year 2 visits. Exclusion criteria included participant withdrawal (n=2), blood draw failure (n=2), and sample not meeting quality control standards for genotyping (n=23). Of the eighty-two ASCEND participants, fifty-seven participants' baseline blood samples were genotyped.

The second cohort consisted of 811 subjects enrolled in the ongoing Alzheimer's Disease Neuroimaging Initiative (ADNI) 1 study (PI Weiner). ADNI is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of AD. Subject enrollment, clinical, and subject characteristics for the ADNI-1 study are described elsewhere (Saykin et al. 2010). Data were downloaded from http://adni.loni.usc.edu. For this study, analyses were conducted under the participant's most recent diagnosis: normal cognition, mild cognitive impairment, or Alzheimer's Disease. Further analyses were conducted on a subset of ADNI participants (n= 266) who had indicated parental history of AD on their

Family History Questionnaire. This subset was then merged with the ASCEND cohort for a final analysis.

Genotyping:

ASCEND: Blood draws were completed by a member of the study team. Venous blood was collected into EDTA anticoagulated tubes and genomic DNA was obtained through standard protocol measures. Typically, 50 to 70 grams of DNA was isolated from 2 milliliters of whole blood per participant. Vials containing blood samples were labeled with subject ID numbers, name of the study, date collected, barcoded and temporarily stored at Executive Park 12. Every 6 months, samples were carried to a research freezer in a secure ADRC cryogenic storage facility located at the Whitehead Biomedical Research Building. Access to the freezer could only be accessed by authorized study personnel and the freezer was protected and equipped with a backup power and telephone alarm system. All blood samples collected from baseline visits (n=78) were sent to the Wingo Lab at Emory University for genotype sequencing. The Wingo lab completed the Genome Wide Association Study genotyping using an Axiom Genotyping Solution Array. Data was sent back to the Wharton Lab in PLINK data format.

ADNI: ADNI blood draw and genotyping protocols for the 811 ADNI-1 participants have previously been described (Saykin et al. 2010). Genotyped data was available for download in the PLINK data format. Genotyped data was available for download in the PLINK data format at http://adni.loni.usc.edu.

Data Analysis:

The primary study objective was to evaluate the potential relationship between self-identified race and the prevalence of SNPs known to influence risk or progression of AD. For the

ASCEND cohort, participant's self-reported race was stored in the secured REDcap online database. For the ADNI cohort, self-reported race was stored within the ADNIMERGE table, which is only authorized to download by approved researchers.

The researchers generated a list of 137 novel and known SNPs of interest implicated in the neuropathology of AD from previously published meta-analyses (Bellenguez et al. 2020 and Giri, Zhang, and Lü, 2016) and from publications addressing genes more prevalent in African genomes (Rotimi et al. 2017). For publications that listed only gene names, common SNP variants were recorded from SNPedia (SNPedia.com).

The researchers utilized PLINK version 1.90 to analyze the data. PLINK is an open-source whole genome association analysis toolset for performing computationally efficient large-scale analyses of genotype/phenotype data (https://zzz.bwh.harvard.edu/plink/). The PLINK pipeline for the ASCEND and merged ASCEND-ADNI cohorts is described in Supplementary File 1. The researchers extracted the SNPs of interest from the ASCEND and ADNI cohorts. The researchers then merged the ASCEND cohort and the subset of ADNI participants of all cognitive states (normal cognition, mild cognitive impairment, Alzheimer's Disease) but had parental history of AD together. A race cluster frequency analysis and association analysis was conducted on the merged cohorts. Minor allele frequency was analyzed in PLINK using chi squared tests, followed by controlling the false discovery rate using the Benjamini-Hochberg Procedure (FDR of 0.05) in GraphPad Prism.

The researchers then performed a diagnosis cluster frequency analysis and association analysis on the merged cohorts. However, the previously described ADNI participants were further stratified by diagnosis (normal cognition or AD) and then merged with the ASCEND participants. Minor allele frequency was analyzed in PLINK using chi squared tests, followed by controlling

the false discovery rate using the Benjamini-Hochberg Procedure (FDR of 0.05) in GraphPad Prism.

To analyze participant genotype, merged pedigree (.ped) and variant information (.map) files were imported to Excel. Counts of individuals stratified by race who were homozygous and heterozygous for disease associated-alleles, as well as homozygous for reference alleles, were generated. Sources used to identify risk-associated alleles can be found in Table 2. For the race analysis, counts of African American individuals (n=25) with at least one disease-associated allele or no disease-associated allele were compared to counts of White individuals (n=290) with at least one disease-associated allele or no disease-associated allele using Fisher's exact test in GraphPad Prism for each SNP, followed by controlling the false discovery rate using the Benjamini-Hochberg Procedure (FDR of 0.05). For the diagnosis analysis, counts of cognitively normal individuals with parental AD history (n=161) with at least one disease-associated allele or no disease-associated allele were compared to counts of AD diagnosed individuals with parental AD history (n=55) with at least one disease-associated allele or no disease-associated allele using Fisher's exact test in GraphPad Prism for each SNP, followed by controlling the false discovery rate using the Benjamini-Hochberg Procedure (FDR of 0.05). P-values, odd's ratios (OR), and confidence intervals are reported in both Table 4 and Table 6.

Results

Participant Demographics

Table 1 reports the demographic characteristics for the 57 ASCEND participants and the 269 ADNI participants. Results show ASCEND participants are middle aged (M=57.97 ± 9.95), mostly female (68.4%), of normal cognitive status (100%), are a mix of African American (36.8%) and White (57.9%) individuals and are carriers of the protective APOE2 allele (15.79%) and risk APOE4 allele (80.85%). ADNI participants are more elderly (M=70.82 ± 6.75), mostly female (53.2%), have varying cognitive states (normal cognition 38.23%, mild cognitive impairment 40.82%, Alzheimer's 19.85%), predominantly White (95.5%), and carriers of the APOE allele (10.82% APOE2 and 47.58% APOE4). All 326 participants were at risk for AD by having a parent diagnosed with the disease.

Table 1 also reports the demographics for the merged race analysis. Results show that the African American participants (n=25) were older (M=68.72 ±8.09), mostly female (80%), are primarily neurotypical (88% normal cognition, 8% mild cognitive impairment, and 4% Alzheimer's), and are more likely to be carriers of the protective APOE2 allele (8%) and risk APOE4 allele (56%). Our results also show White participants (n=290) to be older (M=68.34± 8.04), mostly female (53.79%), have varying cognitive states (46.215 normal cognition, 35.52% mild cognitive impairment, and 17.93% AD), and are carriers of the protective APOE2 allele (9.66%) and risk APOE4 allele (65.17%). African American and White cohorts differed significantly in biological sex representation (p=0.012) and in diagnosis (p=0.0003). All 315 participants were at risk for AD by having a parent diagnosed with the disease.

Table 1 also summarizes the demographics of the merged diagnosis analysis. Our results show that participants of normal cognitive status (n=160) are more elderly (M=68.34 ± 8.56),

mostly female (65%), are a mix of White (83.75%) and African American (13.75%) individuals and are carriers of the protective APOE2 allele (15%) and risk APOE4 allele (51.88%). Results also show participants diagnosed with AD (n=55) to be the most elderly (M=71.76± 6.43), primarily White (95.55%), and 100% carriers for the risk APOE4 allele. Cognitively normal and AD diagnosed cohorts differed significantly in age (p=0.002), biological sex (p=0.025), self-identified race (p=0.045), and in APOE carrier status (p=0.0002). All 215 participants were at risk for AD by having a parent diagnosed with the disease

SNP Summary

Table 2 shows a summary of all the SNPs investigated. From our original list of 137 new and novel SNPs known to be implicated in the progression of AD, 46 SNPs were sequenced between the two cohorts (33 in ASCEND, 26 in ADNI, 12 found in both). All variants are single nucleotide variations (SNVs) and the majority of the SNVs have the consequence of being an intron variant (n=32). Table 2 also identifies the chromosome, reference, alternate, and risk allele for each variant. The reference allele frequency within European, African, and African American populations is also reported to help assess the prevalence of the allele in these populations.

Race Association Analysis

Table 3 shows the results of the race association analysis. The major, minor, and disease associated allele is reported for each SNP. Of the 45 screened variants, we identified 16 SNPs with minor allele frequencies (MAF) which differed by race prior to correction for multiple hypotheses. Following correction, 7 remained significant specifically rs983392, rs1562990, rs2279590, rs7274581, rs11771145, rs3764650, and rs190982. Of these, only one variant, rs11771145, had disease-associated variant overrepresentation in Caucasians; all other disease-associated variants

were overrepresented in African Americans. Therefore, these data would suggest that while there seems to be a racial bias in the distribution of AD-associated SNPs, despite a majority of these variants being identified in White cohorts, a major subset are present within a cohort of African Americans at risk of AD by virtue of family history. Indeed, since only one SNP, rs11771145, was significantly over-represented in Caucasians, 97.8% of the SNPs we screen should be viable for screening in non-White cohorts, especially if multiple SNPs are used to evaluate risk.

Table 4 summarizes the findings of disease-associated genotype frequency analysis by race. The genotype, disease associated variant, and number of White and African American participants who possess each genotype is reported for each variant. Of the 45 screened variants, we identified six SNPs with disease-associated genotype frequencies which differed by race prior to correction for multiple hypotheses. Specifically, risk-associated genotypes in rs11771145, rs6656401, and rs769449 were found to be overrepresented in White participants and risk-associated genotypes in rs7274581, rs3764650, and rs190982 were found to be overrepresented in African American participants. However, following correction, no genotype frequency remained significant. This suggests that despite the fact that all SNPs analyzed were derived from a cohort with White and European ancestry, the majority of risk-associated genotypes are equally pervasive within African Americans at risk for AD. Therefore, these findings suggest that a major subset of risk-associated genotypes derived from White cohorts can be used to assess risk of AD in African Americans. Further, these data suggest that SNPs found in one population, after validation in other cohorts, may be useful in identifying the risk of AD in diverse populations.

Diagnosis Association Analysis

Table 5 shows the results of the diagnosis association analysis. The major, minor, and disease associated allele is reported for each SNP. Of the 26 screened variants, we identified one

SNP, rs7225151, with a MAF score which differed by diagnosis prior to correction for multiple hypotheses. Specifically, rs7225151's disease associated allele was found to be overrepresented in normally cognitive individuals. However, following correction, no SNPs remained significant. These data suggest that cognitively normal individuals who are at risk for AD and express the rs7225151 variant, may be less likely to progress to an AD diagnosis later in life. Further, these results suggest that the genetic markers of AD are complex, and not one gene contributes to the neurodevelopment of AD.

Table 6 summarizes the findings of disease-associated genotype frequency analysis by diagnosis. The genotype, disease associated variant, and number of cognitively normal and AD-diagnosed participants with who possess each genotype is reported for each variant. Of the 26 screened variants, we identified two SNPs with disease-associated genotype frequencies which differed by diagnosis prior to correction for multiple hypotheses. Specifically, rs2154481's disease-associated genotype was found to be overrepresented in individuals diagnosed with AD and rs7225151's disease-associated genotype was found to be overrepresented in cognitively normal participants. Following correction, SNP rs2154481 remained significant. These results suggest that rs2154481 may be a good variant to screen for in individuals at risk for AD as participants who expressed rs2154481's disease-associated genotype had progressed to an AD diagnosis. Further, these results indicate that rs7225151 may not be a good variant to screen for in the progression of AD, as the disease-associated genotype is overrepresented in individuals at risk by virtue of family history but not diagnosed with AD.

Discussion

Here, we conducted association and genotype analyses to (1.) investigate how the prevalence of known risk factors differ between White and African American individuals with a familial history of AD and (2.) explore how the frequency of different gene variations may contribute to the neuropathology of AD in individuals with normal cognition and family history of AD versus individuals with AD and family history. Within our limited study, our results showed that most AD-associated variants identified in White populations may be applicable to assess the development of familial AD within other populations such as African Americans, but all variants must be assessed at a per SNP level. Furthermore, we found no significant differences between neurotypical and AD diagnosed adults with parental history of AD. These results highlight the complexity of AD and emphasize how the development of AD is dependent on genetic, environmental, and lifestyle factors. To our knowledge, this is the first study to verify that novel variants implicated in the progression of AD within White and European cohorts are applicable to the development of AD within African American cohorts. Moreover, this is the only study to compare the prevalence of novel SNPs within cognitively normal U.S adults with parental history of AD to those with parental AD history who have progressed to an AD diagnosis.

Our results suggest that there may be ethnoracial differences in genetic risk factors associated with AD. Indeed, studies have investigated the differences in the prevalence and incidence of AD within the genetic and environmental domains between White and African American cohorts but have produced varied results. It is plausible that this variability results from differences in experimental methodology such as criteria for racial grouping, inclusion criteria, and diagnostic criteria (Chin et al., 2012). Further the increased prevalence of AD within African Americans may also result from the utilization of race-adjusted algorithms and diagnostic tools, and bias of AD diagnostic instruments and disease diagnosis by clinicians (Boyd et al., 2020, Vyas

et al., 2020). Various instruments for the diagnosis of AD, such as the Mini Mental Status Exam, have been found to have low specificity within African Americans, producing up to a 42% false-positive rate for cognitive impairment among African Americans compared to the 6% rate among Caucasians (Chin et al., 2011). Further, due to the overrepresentation of White participants within AD clinical studies, many researchers may assume the genomic landscape of White individuals with AD is applicable to minority populations. Our study overcame some of these potential barriers by having each participant indicate their self-identified race on a questionnaire. Moreover, ADNI AD diagnosis was determined by clinical evaluations and neuropsychological tests and further supported by AD biomarker assessments such as genetic testing, lumbar puncture, and MRI and PET scans (Weiner et al., 2012). Our results suggest that while there are genetic differences between White and African American populations, some White derived SNPs are applicable to African Americans. In order to create a more equitable research and healthcare environment, future studies must verify the applicability of all genetic biomarkers between different racial populations and work to fix the implicit biases found in these domains.

To validate our results, we compared SNP frequencies of our race association analysis to the overall frequency found within European and African American populations reported in previous publications (Sherry et al., 1999). For the purposes of this comparison, similar is defined as frequency within 10% of previously reported penetration. Out of the 45 SNPs screened within African American participants, 16 variants were similar in frequency to other publications and 29 SNPs were not. In addition, out of the 45 variants screened within White participants, 13 SNPs were similar in frequency to other publications and 32 were not. These data indicate that some of the analyzed at-risk variants are naturally prevalent within White and African American populations without a history of AD. However, the majority of the analyzed SNPs are more

prevalent within African American and White cohorts with parental history of AD. This suggests that these SNPs are a valid proxy and steppingstone in identifying the genetic markers that increase AD risk. Further, our results may indicate that the frequency of at-risk variants may differ by self-identified race supporting the idea that different races have different degrees of susceptibility to AD (Christensen et al., 2008).

Family history is the second greatest risk factor for AD following advanced age (Tanzi 2012). Twin and family studies estimate that up to 80% of AD cases involve the inheritance of genetic factors (Gatz et al. 2006). However, the development of AD has been associated with environmental and lifestyle factors such as diet, exposure to environmental toxins such as car exhaust and toxic heavy metals, and incidences of concussion (Grant, 2002, Killin et al., 2016, Ramos-Cejudo et al., 2018). Our inability to find an association between SNP prevalence between cognitively normal and AD diagnosed adults with parental history of AD suggests that familial AD results from an intricate pattern of inheritance in which genetic risk factors work in tandem with environmental factors and life exposure events to determine AD risk (Tanzi 2012).

Notable strengths of this study include the integration of multiple cohorts and the utilization of previously screened blood samples for the detection of novel AD variants. Furthermore, our study only included participants who had parental history of AD confirmed either by autopsy, defined by NINDS-ADRDA criteria, or family history questionnaire confirmed by parental medical records. By investigating familial AD, we can explore the genetic components that contribute to the progression of AD, identify potential detrimental variants, and utilize novel gene and cell targeting therapies to possibly prevent or delay AD development (Rasmussen and Langerman, 2019 and Lorea-Valencia et al., 2018). In addition, our study is the first to validate the utilization of novel White derived SNPs in other cohorts. Prior publications have assumed that

the variants identified in meta-GWAS are applicable to all populations (Cardon and Palmer, 2003 and Price et al., 2006). However, our findings evaluated each novel variant at the individual (by looking at genotype) and population (by looking at allele frequency) level, confirming the applicability of some White derived SNPs to diverse populations. A final strength of this study is our focus on African Americans. Within the United States healthcare and research domains, African Americans are an underrepresented minority group. As previously described, there is a variety of conflicting literature discussing the susceptibility of African Americans to AD, and these are compounded by few studies recruiting African American participants successfully (Barrett et al., 2017, and Ejigou et al., 2007). Our ability to recruit racial minorities into genetic research is a step in addressing conflicting publications and in improving patient care by increasing our understanding of genetic markers of disease in those who do not self-identify as White (Burke 2019, Knerr et al., 2017).

This study also has limitations that must be acknowledged. First, we recognize that the utilization of a small sample size (n=315) may influence the interpretation of results at the statistical level and result in overestimates in the magnitude of SNP prevalence associations (Hackshaw 2008). Further, the unbalanced race (CC=290, AA=25) and diagnosis (NC=160, AD=55) cohorts may lead to spurious associations and false-positive/negative statistical results (Leonard et al., 2018). Future studies should strive to expand the sample size and create balanced racial cohorts to further assess the applicability of White derived SNPs to diverse populations. We also acknowledge that SNPs implicated in the progression of AD within White and European cohorts will naturally be more prevalent within White subjects at risk for AD. However, due to the invention of novel genomic technology and demand for equity within the research domain, new GWAS/GWS studies are being conducted identifying AD variants within primarily African

American cohorts (Kunkle et al., 2020). Future research directions should account for AD SNPs implicated in both White and African American cohorts when assessing the prevalence of variants within diverse populations.

The utilization of microarrays for genotype sequencing is another limitation of this study. Microarrays require high quality control standards, are dependent on previously known genes, and tend to only contain common variants due the limited in the number of SNPs that they can sequence for (Baker 2013 and Jakis et al., 2015). Consequently, the blood of 23 participants was unable to be sequenced due sample to failure to meet quality control standards. Further, the two microarrays only sequenced 45 out of 137 novel SNPs; considering that the novel SNPs chosen were identified years after the original microarrays were run, 32% coverage was a pleasant surprise; this suggests that previously collected genetic data may constitute a data source for future studies even as understanding of disease evolves. Nonetheless, researchers should consider the utilization of Next Generation Sequencing technology to mitigate these limitations in future publications (Betram, 2016).

We also recognize that this study assumes that there is a biological and genetic foundation for race, which has recently been declared unfounded by the American Medical Association (O'Reilly 2020). As previously discussed, race is a social construct that affects daily life at the stress/anxiety level and in turn may impact biological health processes (American Psychological Association 2017). Race is a complex concept, and it is defined and operationalized differently in different disciplines. Increasingly, health researchers are calling for discussions about the nature of "race," the social construction of race, and the implications of these conceptualizations in health and medical research (Boyd et. al. 2020). The goal of this study was to assess the interactions between race and genetics in AD susceptibility and in order to investigate this relationship, we had

to assume that there is some genetic foundation for race. However, we acknowledge how this assumption is unsupported and may perpetuate the implicit biases found within the healthcare setting. Indeed, this study suggests that there is significant overlap in the genetic foundation of AD between people of different races. Future studies must account for ancestral history when assessing the interplay between genetic markers and disease progression.

In summary, our results showed that some White derived variants may be applicable to diverse populations, and that there are a variety of genetic and environmental factors contributing to the development of familial AD. However, our results stress the need for further research investigating the genetic factors that contribute to the progression and development of AD.

References

Alz.org (2020). 2020 Alzheimer's Disease Facts and Figures
American Psychological Association, APA Working Group on Stress and Health Disparities. (2017). Stress and health disparities: Contexts, mechanisms, and interventions among racial/ethnic minority and low-socioeconomic status populations. Retrieved from http://www.apa.org/pi/health-disparities/resources/stress-report.aspx
Antúnez, Carmen et al. "The membrane-spanning 4-domains, subfamily A (MS4A) gene cluster contains a common variant associated with Alzheimer's disease." Genome medicine vol. 3,5 33. 31 May. 2011, doi:10.1186/gm249
Arboleda-Velasquez JF, Lopera F, O'Hare M, Delgado-Tirado S, Marino C, Chmielewska N, Saez-Torres KL, Amarnani D, Schultz AP, Sperling RA, Leyton-Cifuentes D, Chen K, Baena A, Aguillon D, Rios-Romenets S, Giraldo M, Guzmán-Vélez E, Norton DJ, Pardilla-Delgado E, Artola A, Sanchez JS, Acosta-Uribe J, Lalli M, Kosik KS, Huentelman MJ, Zetterberg H, Blennow K, Reiman RA, Luo J, Chen Y, Thiyyagura P, Su Y, Jun GR, Naymik M, Gai X, Bootwalla M, Ji J, Shen L, Miller JB, Kim LA, Tariot PN, Johnson KA, Reiman EM, Quiroz YT. Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report. Nat Med. 2019 Nov;25(11):1680-1683. doi: 10.1038/s41591-019-0611-3. Epub 2019 Nov 4. PMID: 31686034; PMCID: PMC6898984.
Axiom™ Genotyping Solution. 2020. Axiom™ Genotyping Solution Data Analysis: *User Guide* Santa Clara, CA: Thermo Fisher Scientific Inc.
Baker, S. C. (2013, January 14). Next-Generation sequencing Vs. microarrays. Retrieved March 15, 2021, from https://www.genengnews.com/magazine/next-generation-sequencing-vs-microarrays-2/
Barrett NJ, Ingraham KL, Vann Hawkins T, Moorman PG. Engaging African Americans in Research: The Recruiter's Perspective. Ethn Dis. 2017;27(4):453-462. Published 2017 Dec 7. doi:10.18865/ed.27.4.453
Bateman RJ, Xiong C, Benzinger TL, Fagan AM, Goate A, Fox NC, Marcus DS, Cairns NJ, Xie X, Blazey TM, Holtzman DM, Santacruz A, Buckles V, Oliver A, Moulder K, Aisen PS, Ghetti B, Klunk WE, McDade E, Martins RN, Masters CL, Mayeux R, Ringman JM, Rossor MN, Schofield PR, Sperling RA, Salloway S, Morris JC; Dominantly Inherited Alzheimer Network. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med. 2012 Aug 30;367(9):795-804. doi: 10.1056/NEJMoa1202753. Epub 2012 Jul 11. Erratum in: N Engl J Med. 2012 Aug 23;367(8):780. PMID: 22784036; PMCID: PMC3474597.
Beason-Held, Lori L et al. "Changes in brain function occur years before the onset of cognitive impairment." The Journal of neuroscience : the official journal of the Society for Neuroscience vol. 33,46 (2013): 18008-14. doi:10.1523/JNEUROSCI.1402-13.2013
Bekris, L. M., Yu, C., Bird, T. D., & Tsuang, D. W. (2010). Genetics of Alzheimer Disease. Journal of Geriatric Psychiatry and Neurology, 23(4), 213-227. doi:10.1177/0891988710383571
Bellenguez, C., et al. (2020). "New insights on the genetic etiology of Alzheimer's and related dementia." medRxiv: 2020.2010.2001.20200659.

Bellenguez C, Küçükali F, Jansen I, Andrade V, Moreno-Grau S, Amin N, Naj AC, Grenier-Boley B, Campos-Martin R, Holmans PA, Boland A, Kleineidam L, Damotte V, van der Lee SJ, Kuulasmaa T, Yang Q, de Rojas I, Bis JC, Yaqub A, Prokic I, Costa MR, Chapuis J, Ahmad S, Giedraitis V, Boada M, Aarsland D, García-González P, Abdelnour C, Alarcón-Martín E, Alegret M, Alvarez I, Álvarez V, Armstrong NJ, Tsolaki A, Antúnez C, Appollonio I, Arcaro M, Archetti S, Pastor AA, Arosio B, Athanasiu L, Bailly H, Banaj N, Baquero M, Pastor AB, Benussi L, Berr C, Besse C, Bessi V, Binetti G, Bizzarro A, Alcolea D, Blesa R, Borroni B, Boschi S, Bossù P, Bråthen G, Bresner C, Brookes KJ, Brusco LI, Bûrger K, Bullido MJ, Burholt V, Bush WS, Calero M, Dufouil C, Carracedo A, Cecchetti R, Cervera-Carles L, Charbonnier C, Chillotti C, Brodaty H, Ciccone S, Claassen JAHR, Clark C, Conti E, Corma-Gómez A, Costantini E, Custodero C, Daian D, Dalmasso MC, Daniele A, Dardiotis E, Dartigues JF, de Deyn PP, de Paiva Lopes K, de Witte LD, Deb. New insights on the genetic etiology of Alzheimer's and related dementia. *medRxiv*. 2020 Dec 14.

Bellenguez, C., et al. (2020). "Large meta-analysis of genome-wide association studies expands knowledge of the genetic etiology of Alzheimer's disease and highlights potential translational opportunities." medRxiv: 2020.2010.2001.20200659.

Bertram L. Next Generation Sequencing in Alzheimer's Disease. Methods Mol Biol. 2016;1303:281-97. doi: 10.1007/978-1-4939-2627-5_17. PMID: 26235074.

Burke, C. (2019, June 18). The importance of increasing diversity in genetic research. Retrieved March 15, 2021, from https://www.biospace.com/article/the-importance-of-increasing-diversity-in-genetic-research/

Boyd, Rhea, Lindo, Edwin Lindo, Weeks, Lachelle, McLemore, Monica. (July 2, 2020). "On Racism: A New Standard for Publishing on Racial Health Inequities." *Health Affairs*. https://www.healthaffairs.org/do/10.1377/hblog20200630.939347/full/

Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet. 2003 Feb 15;361(9357):598-604. doi: 10.1016/S0140-6736(03)12520-2. PMID: 12598158.

Chin AL, Negash S, Hamilton R. Diversity and disparity in dementia: the impact of ethnoracial differences in Alzheimer disease. Alzheimer Dis Assoc Disord. 2011;25(3):187-195. doi:10.1097/WAD.0b013e318211c6c9

Christensen, K., Roberts, J., Royal, C. *et al.* Incorporating ethnicity into genetic risk assessment for Alzheimer disease: the REVEAL study experience. *Genet Med* **10,** 207–214 (2008). https://doi.org/10.1097/GIM.0b013e318164e4cf

Cummings, J., & Fox, N. (2017). Defining Disease Modifying Therapy for Alzheimer's Disease. The Journal of Prevention of Alzheimer's Disease, 4(2), 109-115. http://dx.doi.org/10.14283/jpad.2017

Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA, Thompson WK, Besser L, Kukull WA, Holland D, Chen CH, Brewer JB, Karow DS, Kauppi K, Witoelar A, Karch CM, Bonham LW, Yokoyama JS, Rosen HJ, Miller BL, Dillon WP, Wilson DM, Hess CP, Pericak-Vance M, Haines JL, Farrer LA, Mayeux R, Hardy J, Goate AM, Hyman BT, Schellenberg GD, McEvoy LK, Andreassen OA, Dale AM. Genetic assessment of age-associated Alzheimer

disease risk: Development and validation of a polygenic hazard score. PLoS Med. 2017 Mar 21;14(3):e1002258. doi: 10.1371/journal.pmed.1002258. Erratum in: PLoS Med. 2017 Mar 28;14 (3):e1002289. PMID: 28323831; PMCID: PMC5360219.

DeTure, M.A., Dickson, D.W. The neuropathological diagnosis of Alzheimer's disease. Mol Neurodegeneration 14, 32 (2019). https://doi.org/10.1186/s13024-019-0333-5

Ejiogu N, Norbeck JH, Mason MA, Cromwell BC, Zonderman AB, Evans MK. Recruitment and retention strategies for minority or poor clinical research participants: lessons from the Healthy Aging in Neighborhoods of Diversity across the Life Span study. Gerontologist. 2011;51 Suppl 1(Suppl 1):S33-S45. doi:10.1093/geront/gnr027

Fitzpatrick AL, Kuller LH, Ives DG, Lopez OL, Jagust W, Breitner JC, Jones B, Lyketsos C, Dulberg C. Incidence and prevalence of dementia in the Cardiovascular Health Study. J Am Geriatr Soc. 2004 Feb;52(2):195-204. doi: 10.1111/j.1532-5415.2004.52058.x. PMID: 14728627.

Gatz M, Reynolds CA, Fratiglioni L, et al. Role of Genes and Environments for Explaining Alzheimer Disease. Arch Gen Psychiatry. 2006;63(2):168–174. doi:10.1001/archpsyc.63.2.168

Gerda G. Fillenbaum, Albert Heyman, Marc S. Huber, Max A. Woodbury, Jack Leiss, Kenneth E. Schmader, Arline Bohannon, Barbara Trapp-Moen, The Prevalence and 3-Year Incidence of Dementia in Older Black and White Community Residents, Journal of Clinical Epidemiology, Volume 51, Issue 7, 1998, Pages 587-595, ISSN 0895-4356, https://doi.org/10.1016/S0895-4356(98)00024-9.

Giri, Mohan et al. "Genes associated with Alzheimer's disease: an overview and current status." Clinical interventions in aging vol. 11 665-81. 17 May. 2016, doi:10.2147/CIA.S105769

Grant, William B. et al. 'The Significance of Environmental Factors in the Etiology of Alzheimer's Disease'. 1 Jan. 2002 : 179 – 189.

Green RC, Cupples LA, Go R, Benke KS, Edeki T, Griffith PA, Williams M, Hipps Y, Graff-Radford N, Bachman D, Farrer LA; MIRAGE Study Group. Risk of dementia among white and African American relatives of patients with Alzheimer disease. JAMA. 2002 Jan 16;287(3):329-36. doi: 10.1001/jama.287.3.329. PMID: 11790212.

Hackshaw, A. (2008). "Small studies: strengths and limitations." European Respiratory Journal 32(5): 1141-1143.

Hamza TH, Zabetian CP, Tenesa A, Laederach A, Montimurro J, Yearout D, Kay DM, Doheny KF, Paschall J, Pugh E, Kusel VI, Collura R, Roberts J, Griffith A, Samii A, Scott WK, Nutt J, Factor SA, Payami H. Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. Nat Genet. 2010 Sep;42(9):781-5. doi: 10.1038/ng.642. Epub 2010 Aug 15. PMID: 20711177; PMCID: PMC2930111.

Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F,

Schürmann B, Heun R, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frölich L, Hampel H, Hüll M, Rujescu D, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Mühleisen TW, Nöthen MM, Moebus S, Jöckel KH, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009 Oct;41(10):1088-93. doi: 10.1038/ng.440. Epub 2009 Sep 6. Erratum in: Nat Genet. 2009 Oct;41(10):1156. Erratum in: Nat Genet. 2013 Jun;45(6):712. Haun, Reinhard [added]. PMID: 19734902; PMCID: PMC2845877.
Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, Abraham R, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Jones N, Stretton A, Thomas C, Richards A, Ivanov D, Widdowson C, Chapman J, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Beaumont H, Warden D, Wilcock G, Love S, Kehoe PG, Hooper NM, Vardy ER, Hardy J, Mead S, Fox NC, Rossor M, Collinge J, Maier W, Jessen F, Rüther E, Schürmann B, Heun R, Kölsch H, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frölich L, Hampel H, Gallacher J, Hüll M, Rujescu D, Giegling I, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Mühleisen TW, Nöthen MM, Moebus S, Jöckel KH, Klopp N, Wichmann HE, Pankratz VS, Sando SB, Aasly JO, Barcikowska M, Wszolek ZK, Dickson DW, Graff-Radford NR, Petersen RC; Alzheimer's Disease Neuroimaging Initiative, van Duijn CM, Breteler MM, Ikram MA, DeStefano AL, Fitzpatrick AL, Lopez O, Launer LJ, Seshadri S; CHARGE consortium, Berr C, Campion D, Epelbaum J, Dartigues JF, Tzourio C, Alpérovitch A, Lathrop M; EADI1 consortium, Feulner TM, Friedrich P, Riehle C, Krawczak M, Schreiber S, Mayhaus M, Nicolhaus S, Wagenpfeil S, Steinberg S, Stefansson H, Stefansson K, Snaedal J, Björnsson S, Jonsson PV, Chouraki V, Genier-Boley B, Hiltunen M, Soininen H, Combarros O, Zelenika D, Delepine M, Bullido MJ, Pasquier F, Mateo I, Frank-Garcia A, Porcellini E, Hanon O, Coto E, Alvarez V, Bosco P, Siciliano G, Mancuso M, Panza F, Solfrizzi V, Nacmias B, Sorbi S, Bossù P, Piccardi P, Arosio B, Annoni G, Seripa D, Pilotto A, Scarpini E, Galimberti D, Brice A, Hannequin D, Licastro F, Jones L, Holmans PA, Jonsson T, Riemenschneider M, Morgan K, Younkin SG, Owen MJ, O'Donovan M, Amouyel P, Williams J. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat Genet. 2011 May;43(5):429-35. doi: 10.1038/ng.803. Epub 2011 Apr 3. PMID: 21460840; PMCID: PMC3084173.

Jaksik, R., Iwanaszko, M., Rzeszowska-Wolny, J. et al. Microarray experiments and factors which affect their reliability. Biol Direct 10, 46 (2015). https://doi.org/10.1186/s13062-015-0077-2

Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, Sealock J, Karlsson IK, Hägg S, Athanasiu L, Voyle N, Proitsi P, Witoelar A, Stringer S, Aarsland D, Almdahl IS, Andersen F, Bergh S, Bettella F, Bjornsson S, Brækhus A, Bråthen G, de Leeuw C, Desikan RS, Djurovic S, Dumitrescu L, Fladby T, Hohman TJ, Jonsson PV, Kiddle SJ, Rongve A, Saltvedt I, Sando SB, Selbæk G, Shoai M, Skene NG, Snaedal J, Stordal E, Ulstein ID, Wang Y, White LR, Hardy J, Hjerling-Leffler J, Sullivan PF, van der Flier WM, Dobson R, Davis LK, Stefansson H, Stefansson K, Pedersen NL, Ripke S, Andreassen OA, Posthuma D. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet. 2019 Mar;51(3):404-413. doi: 10.1038/s41588-018-0311-9. Epub 2019 Jan 7. Erratum in: Nat Genet. 2020 Mar;52(3):354. PMID: 30617256; PMCID: PMC6836675.

Jun G, Ibrahim-Verbaas CA, Vronskaya M, Lambert JC, Chung J, Naj AC, Kunkle BW, Wang LS, Bis JC, Bellenguez C, Harold D, Lunetta KL, Destefano AL, Grenier-Boley B, Sims R, Beecham GW, Smith AV, Chouraki V, Hamilton-Nelson KL, Ikram MA, Fievet N, Denning N, Martin ER, Schmidt H, Kamatani Y, Dunstan ML, Valladares O, Laza AR, Zelenika D, Ramirez A, Foroud TM, Choi SH, Boland A, Becker T, Kukull WA, van der Lee SJ, Pasquier F, Cruchaga C, Beekly D, Fitzpatrick AL, Hanon O, Gill M, Barber R, Gudnason V, Campion D, Love S, Bennett DA, Amin N, Berr C, Tsolaki M, Buxbaum JD, Lopez OL, Deramecourt V, Fox NC, Cantwell LB, Tárraga L, Dufouil C, Hardy J, Crane PK, Eiriksdottir G, Hannequin D, Clarke R, Evans D, Mosley TH Jr, Letenneur L, Brayne C, Maier W, De Jager P, Emilsson V, Dartigues JF, Hampel H, Kamboh MI, de Bruijn RF, Tzourio C, Pastor P, Larson EB, Rotter JI, O'Donovan MC, Montine TJ, Nalls MA, Mead S, Reiman EM, Jonsson PV, Holmes C, St George-Hyslop PH, Boada M, Passmore P, Wendland JR, Schmidt R, Morgan K, Winslow AR, Powell JF, Carasquillo M, Younkin SG, Jakobsdóttir J, Kauwe JS, Wilhelmsen KC, Rujescu D, Nöthen MM, Hofman A, Jones L; IGAP Consortium, Haines JL, Psaty BM, Van Broeckhoven C, Holmans P, Launer LJ, Mayeux R, Lathrop M, Goate AM, Escott-Price V, Seshadri S, Pericak-Vance MA, Amouyel P, Williams J, van Duijn CM, Schellenberg GD, Farrer LA. A novel Alzheimer disease locus located near the gene encoding tau protein. Mol Psychiatry. 2016 Jan;21(1):108-17. doi: 10.1038/mp.2015.23. Epub 2015 Mar 17. PMID: 25778476; PMCID: PMC4573764.

Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E, Lyytikäinen LP, Kangas AJ, Soininen P, Würtz P, Silander K, Dick DM, Rose RJ, Savolainen MJ, Viikari J, Kähönen M, Lehtimäki T, Pietiläinen KH, Inouye M, McCarthy MI, Jula A, Eriksson J, Raitakari OT, Salomaa V, Kaprio J, Järvelin MR, Peltonen L, Perola M, Freimer NB, Ala-Korpela M, Palotie A, Ripatti S. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat Genet. 2012 Jan 29;44(3):269-76. doi: 10.1038/ng.1073. PMID: 22286219; PMCID: PMC3605033.

Killin LO, Starr JM, Shiue IJ, Russ TC. Environmental risk factors for dementia: a systematic review. BMC Geriatr. 2016;16(1):175. Published 2016 Oct 12. doi:10.1186/s12877-016-0342-y
Knerr S, Wayman D, Bonham VL. Inclusion of racial and ethnic minorities in genetic research: advance the spirit by changing the rules?. J Law Med Ethics. 2011;39(3):502-512. doi:10.1111/j.1748-720X.2011.00617
Kumar VV, Huang H, Zhao L, Verble DD, Nutaitis A, Tharwani SD, Brown AL, Zetterberg H, Hu W, Shin R, Kehoe PG, Quyyumi A, Nocera J, Kippels A, Wharton W. Baseline Results: The Association Between Cardiovascular Risk and Preclinical Alzheimer's Disease Pathology (ASCEND) Study. J Alzheimers Dis. 2020;75(1):109-117. doi: 10.3233/JAD-191103. PMID: 32280088.
Kunkle BW, Schmidt M, Klein HU, Naj AC, Hamilton-Nelson KL, Larson EB, Evans DA, De Jager PL, Crane PK, Buxbaum JD, Ertekin-Taner N, Barnes LL, Fallin MD, Manly JJ, Go RCP, Obisesan TO, Kamboh MI, Bennett DA, Hall KS, Goate AM, Foroud TM, Martin ER, Wang LS, Byrd GS, Farrer LA, Haines JL, Schellenberg GD, Mayeux R, Pericak-Vance MA, Reitz C; Writing Group for the Alzheimer's Disease Genetics Consortium (ADGC), Graff-Radford NR, Martinez I, Ayodele T, Logue MW, Cantwell LB, Jean-Francois M, Kuzma AB, Adams LD, Vance JM, Cuccaro ML, Chung J, Mez J, Lunetta KL, Jun GR, Lopez OL, Hendrie HC, Reiman EM, Kowall NW, Leverenz JB, Small SA, Levey AI, Golde TE, Saykin AJ, Starks TD, Albert MS, Hyman BT, Petersen RC, Sano M, Wisniewski T, Vassar R, Kaye JA, Henderson VW, DeCarli C, LaFerla FM, Brewer JB, Miller BL, Swerdlow RH, Van Eldik LJ, Paulson HL, Trojanowski JQ, Chui HC, Rosenberg RN, Craft S, Grabowski TJ, Asthana S, Morris JC, Strittmatter SM, Kukull WA. Novel Alzheimer Disease Risk Loci and Pathways in African American Individuals Using the African Genome Resources Panel: A Meta-analysis. JAMA Neurol. 2021 Jan 1;78(1):102-113. doi: 10.1001/jamaneurol.2020.3536. PMID: 33074286; PMCID: PMC7573798.
Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thorton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Morón FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fiévet N, Huentelman MW, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuiness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossù P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Deniz Naranjo MC, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F; European Alzheimer's Disease Initiative (EADI); Genetic and Environmental Risk in Alzheimer's Disease; Alzheimer's Disease Genetic Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannefelt L,

Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH Jr, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RF, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JS, Boerwinkle E, Riemenschneider M, Boada M, Hiltuenen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nöthen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013 Dec;45(12):1452-8. doi: 10.1038/ng.2802. Epub 2013 Oct 27. PMID: 24162737; PMCID: PMC3896259.

Leonard, H., et al. (2018). "Genetic variability and potential effects on clinical trial outcomes: perspectives in Parkinson's disease." bioRxiv: 427385.

Levesque G, Ikeda M, Liang Y, Chi H, Lin C, Holman K, Tsuda T, et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. Nature. 1995 Aug 31;376(6543):775-8. doi: 10.1038/376775a0. PMID: 7651536.

Liu C, Chyr J, Zhao W, Xu Y, Ji Z, Tan H, Soto C, Zhou X; Alzheimer's Disease Neuroimaging Initiative. Genome-Wide Association and Mechanistic Studies Indicate That Immune Response Contributes to Alzheimer's Disease Development. Front Genet. 2018 Sep 24;9:410. doi: 10.3389/fgene.2018.00410. PMID: 30319691; PMCID: PMC616600

Liu, C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: Risk, mechanisms and therapy. *Nature Reviews Neurology, 9*(2), 106-118. doi:10.1038/nrneurol.2012.263

Lo MT, Kauppi K, Fan CC, Sanyal N, Reas ET, Sundar VS, Lee WC, Desikan RS, McEvoy LK, Chen CH; Alzheimer's Disease Genetics Consortium. Identification of genetic heterogeneity of Alzheimer's disease across age. Neurobiol Aging. 2019 Dec;84:243.e1-243.e9. doi: 10.1016/j.neurobiolaging.2019.02.022. Epub 2019 Mar 12. PMID: 30979435; PMCID: PMC6783343.

Loera-Valencia R, Piras A, Ismail MAM, Manchanda S, Eyjolfsdottir H, Saido TC, Johansson J, Eriksdotter M, Winblad B, Nilsson P. Targeting Alzheimer's disease with gene and cell therapies. J Intern Med. 2018 Jul;284(1):2-36. doi: 10.1111/joim.12759. PMID: 29582495.

Loy, C. T., Schofield, P. R., Turner, A. M., & Kwok, J. B. (2014). Genetics of dementia. The Lancet, 383(9919), 828-840. doi:10.1016/s0140-6736(13)60630-3

Mayeda ER, Glymour MM, Quesenberry CP, Whitmer RA. Inequalities in dementia incidence between six racial and ethnic groups over 14 years. Alzheimers Dement. 2016 Mar;12(3):216-24. doi: 10.1016/j.jalz.2015.12.007. Epub 2016 Feb 11. PMID: 26874595; PMCID: PMC4969071.

Morris JC, Schindler SE, McCue LM, et al. Assessment of Racial Disparities in Biomarkers for Alzheimer Disease. JAMA Neurol. 2019;76(3):264–273. doi:10.1001/jamaneurol.2018.4249

Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J, Gallins PJ, Buxbaum JD, Jarvik GP, Crane PK, Larson EB, Bird TD, Boeve BF, Graff-Radford NR, De Jager PL, Evans D, Schneider JA, Carrasquillo MM, Ertekin-Taner N, Younkin SG, Cruchaga C, Kauwe JS, Nowotny P, Kramer P, Hardy J, Huentelman MJ, Myers AJ, Barmada MM, Demirci FY, Baldwin CT, Green RC, Rogaeva E, St George-Hyslop P, Arnold SE, Barber R, Beach T, Bigio EH, Bowen JD, Boxer A, Burke JR, Cairns NJ, Carlson CS, Carney RM, Carroll SL, Chui HC, Clark DG, Corneveaux J, Cotman CW, Cummings JL, DeCarli C, DeKosky ST, Diaz-Arrastia R, Dick M, Dickson DW, Ellis WG, Faber KM, Fallon KB, Farlow MR, Ferris S, Frosch MP, Galasko DR, Ganguli M, Gearing M, Geschwind DH, Ghetti B, Gilbert JR, Gilman S, Giordani B, Glass JD, Growdon JH, Hamilton RL, Harrell LE, Head E, Honig LS, Hulette CM, Hyman BT, Jicha GA, Jin LW, Johnson N, Karlawish J, Karydas A, Kaye JA, Kim R, Koo EH, Kowall NW, Lah JJ, Levey AI, Lieberman AP, Lopez OL, Mack WJ, Marson DC, Martiniuk F, Mash DC, Masliah E, McCormick WC, McCurry SM, McDavid AN, McKee AC, Mesulam M, Miller BL, Miller CA, Miller JW, Parisi JE, Perl DP, Peskind E, Petersen RC, Poon WW, Quinn JF, Rajbhandary RA, Raskind M, Reisberg B, Ringman JM, Roberson ED, Rosenberg RN, Sano M, Schneider LS, Seeley W, Shelanski ML, Slifer MA, Smith CD, Sonnen JA, Spina S, Stern RA, Tanzi RE, Trojanowski JQ, Troncoso JC, Van Deerlin VM, Vinters HV, Vonsattel JP, Weintraub S, Welsh-Bohmer KA, Williamson J, Woltjer RL, Cantwell LB, Dombroski BA, Beekly D, Lunetta KL, Martin ER, Kamboh MI, Saykin AJ, Reiman EM, Bennett DA, Morris JC, Montine TJ, Goate AM, Blacker D, Tsuang DW, Hakonarson H, Kukull WA, Foroud TM, Haines JL, Mayeux R, Pericak-Vance MA, Farrer LA, Schellenberg GD. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet. 2011 May;43(5):436-41. doi: 10.1038/ng.801. Epub 2011 Apr 3. PMID: 21460841; PMCID: PMC3090745.

Naj, A., Jun, G., Beecham, G. et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet 43, 436–441 (2011). https://doi.org/10.1038/ng.801

National Institute on Aging. (2017, May 16). What Happens to the Brain in Alzheimer's Disease? Retrieved October 08, 2020, from https://www.nia.nih.gov/health/what-happens-brain-alzheimers-disease

Omi, Michael, and Howard Winant. Racial Formation in the United States: From the 1960s to the 1990s. 2nd ed. New York: Routledge, 1994.

"On Racism: A New Standard For Publishing On Racial Health Inequities, " Health Affairs Blog, July 2, 2020.DOI: 10.1377/hblog20200630.939347

O'Reilly News Editor. (2020, November 16). AMA: Racism is a threat to public health. Retrieved March 15, 2021, from https://www.ama-assn.org/delivering-care/health-equity/ama-racism-threat-public-

health#:~:text=Recognizing%20race%20as%20social%20construct&text=Delegates%20adopted%20new%20policy%20to,education%2C%20research%20and%20clinical%20practice
Perl, D.P. (2010), Neuropathology of Alzheimer's Disease. Mt Sinai J Med, 77: 32-42. doi:10.1002/msj.20157
Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006 Aug;38(8):904-9. doi: 10.1038/ng1847. Epub 2006 Jul 23. PMID: 16862161.
Ramos-Cejudo J, Wisniewski T, Marmar C, et al. Traumatic Brain Injury and Alzheimer's Disease: The Cerebrovascular Link. EBioMedicine. 2018;28:21-30. doi:10.1016/j.ebiom.2018.01.021
Rasmussen, Jill, and Haya Langerman. "Alzheimer's Disease - Why We Need Early Diagnosis." Degenerative neurological and neuromuscular disease vol. 9 123-130. 24 Dec. 2019, doi:10.2147/DNND.S228939
Robertson, S. (2018, August 23). What are Amyloid Plaques? Retrieved October 08, 2020, from https://www.news-medical.net/health/What-are-Amyloid-Plaques.aspx
Rotimi, Charles N et al. "The genomic landscape of African populations in health and disease." Human molecular genetics vol. 26,R2 (2017): R225-R236. doi:10.1093/hmg/ddx253
Saykin, Andrew J et al. "Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans." Alzheimer's & dementia : the journal of the Alzheimer's Association vol. 6,3 (2010): 265-73. doi:10.1016/j.jalz.2010.03.013
Shen N, Chen B, Jiang Y, Feng R, Liao M, Zhang L, Li F, Ma G, Chen Z, Zhao B, Li K, Liu G. An Updated Analysis with 85,939 Samples Confirms the Association Between CR1 rs6656401 Polymorphism and Alzheimer's Disease. Mol Neurobiol. 2015;51(3):1017-23. doi: 10.1007/s12035-014-8761-2. Epub 2014 May 31. PMID: 24878768.
Sherry,S.T., Ward,M. and Sirotkin,K. (1999) dbSNP—Database for Single Nucleotide Polymorphisms and Other Classes of Minor Genetic Variation. *Genome Res.*, **9**, 677–679.
Steenland, Kyle et al. "A Meta-Analysis of Alzheimer's Disease Incidence and Prevalence Comparing African-Americans and Caucasians." Journal of Alzheimer's disease : JAD vol. 50,1 (2016): 71-6. doi:10.3233/JAD-150778
Tang MX, Cross P, Andrews H, Jacobs DM, Small S, Bell K, Merchant C, Lantigua R, Costa R, Stern Y, Mayeux R. Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. Neurology. 2001 Jan 9;56(1):49-56. doi: 10.1212/wnl.56.1.49. PMID: 11148235.
Tanzi RE. The genetics of Alzheimer disease. Cold Spring Harb Perspect Med. 2012;2(10):a006296. Published 2012 Oct 1. doi:10.1101/cshperspect.a006296
Tao QQ, Chen YC, Wu ZY. The role of CD2AP in the Pathogenesis of Alzheimer's Disease. Aging Dis. 2019;10(4):901-907. Published 2019 Aug 1. doi:10.14336/AD.2018.1025
Vuckovic D, Bao EL, Akbari P, Lareau CA, Mousas A, Jiang T, Chen MH, Raffield LM, Tardaguila M, Huffman JE, Ritchie SC, Megy K, Ponstingl H, Penkett CJ, Albers PK, Wigdor EM, Sakaue S, Moscati A, Manansala R, Lo KS, Qian H, Akiyama M, Bartz TM, Ben-Shlomo

Y, Beswick A, Bork-Jensen J, Bottinger EP, Brody JA, van Rooij FJA, Chitrala KN, Wilson PWF, Choquet H, Danesh J, Di Angelantonio E, Dimou N, Ding J, Elliott P, Esko T, Evans MK, Felix SB, Floyd JS, Broer L, Grarup N, Guo MH, Guo Q, Greinacher A, Haessler J, Hansen T, Howson JMM, Huang W, Jorgenson E, Kacprowski T, Kähönen M, Kamatani Y, Kanai M, Karthikeyan S, Koskeridis F, Lange LA, Lehtimäki T, Linneberg A, Liu Y, Lyytikäinen LP, Manichaikul A, Matsuda K, Mohlke KL, Mononen N, Murakami Y, Nadkarni GN, Nikus K, Pankratz N, Pedersen O, Preuss M, Psaty BM, Raitakari OT, Rich SS, Rodriguez BAT, Rosen JD, Rotter JI, Schubert P, Spracklen CN, Surendran P, Tang H, Tardif JC, Ghanbari M, Völker U, Völzke H, Watkins NA, Weiss S; VA Million Veteran Program, Cai N, Kundu K, Watt SB, Walter K, Zonderman AB, Cho K, Li Y, Loos RJF, Knight JC, Georges M, Stegle O, Evangelou E, Okada Y, Roberts DJ, Inouye M, Johnson AD, Auer PL, Astle WJ, Reiner AP, Butterworth AS, Ouwehand WH, Lettre G, Sankaran VG, Soranzo N. The Polygenic and Monogenic Basis of Blood Traits and Diseases. Cell. 2020 Sep 3;182(5):1214-1231.e11. doi: 10.1016/j.cell.2020.08.008. PMID: 32888494; PMCID: PMC7482360.

Vyas DA, Eisenstein LG, Jones DS. Hidden in Plain Sight - Reconsidering the Use of Race Correction in Clinical Algorithms. N Engl J Med. 2020 Aug 27;383(9):874-882. doi: 10.1056/NEJMms2004740. Epub 2020 Jun 17. PMID: 32853499.

Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, Harvey D, Jack CR, Jagust W, Liu E, Morris JC, Petersen RC, Saykin AJ, Schmidt ME, Shaw L, Siuciak JA, Soares H, Toga AW, Trojanowski JQ; Alzheimer's Disease Neuroimaging Initiative. The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. Alzheimers Dement. 2012 Feb;8(1 Suppl):S1-68. doi: 10.1016/j.jalz.2011.09.172. Epub 2011 Nov 2. PMID: 22047634; PMCID: PMC3329969

Wharton, Whitney et al. "Interleukin 9 alterations linked to alzheimer disease in african americans." Annals of neurology vol. 86,3 (2019): 407-418. doi:10.1002/ana.25543

Wightman, D. P., et al. (2020). "Largest GWAS (N=1,126,563) of Alzheimer's Disease Implicates Microglia and Immune Cells." medRxiv: 2020.2011.2020.20235275

World Health Organization. (2018, May 24). The top 10 causes of death. Retrieved October 08, 2020, from https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

Xiao, X., et al. (2020). "Association of Genes Involved in the Metabolic Pathways of Amyloid-β and Tau Proteins With Sporadic Late-Onset Alzheimer's Disease in the Southern Han Chinese Population." Frontiers in Aging Neuroscience 12(390).

Yuan Q, Chu C, Jia J. Association studies of 19 candidate SNPs with sporadic Alzheimer's disease in the North Chinese Han population. Neurol Sci. 2012 Oct;33(5):1021-8. doi: 10.1007/s10072-011-0881-0. Epub 2011 Dec 14. PMID: 22167654.

Zhou L, Li HY, Wang JH, et al. Correlation of gene polymorphisms of CD36 and ApoE with susceptibility of Alzheimer disease: A case-control study. Medicine (Baltimore). 2018;97(38):e12470. doi:10.1097/MD.0000000000012470

Zhuang, L., et al. (2019). "Genetic Variants of PICALM rs541458 Modulate Brain Spontaneous Activity in Older Adults With Amnestic Mild Cognitive Impairment." Frontiers in Neurology 10(494).

Tables

Table 1: Patient Demographics

NC: Normal Cognition; AD: Alzheimer’s Disease; Welch's t test was used to compare age; fisher's exact test was used to compare biological sex, parental AD diagnosis, and APOE genotype; and chi-squared test was used for race within diagnosis, and diagnosis within race.

			Merged for Race			Merged for Diagnosis		
Category	**ASCEND (n=57)**	**ADNI (n=269)**	**Black (n=25)**	**White (n=290)**	**P value**	**NC (n=160)**	**AD (n=55)**	**P value**
Age (years):	57.97 ± 9.95	70.82 ± 6.75	68.72± 8.09	68.86± 8.04	0.934	68.34 ± 8.56	71.76 ± 6.43	0.002
Biological Sex N (% female):	39 (68.4%)	116 (53.2%)	20 (80%)	156 (53.79%)	0.012	104 (65%)	26 (47.27%)	0.025
Self-Identified Race N (%):								0.045
Caucasian	33 (57.9%)	257 (95.5%)	0 (0%)	290 (100%)		134 (83.75%)	52 (95.55%)	
African American	21 (36.8%)	4 (1.56%)	25 (100%)	0 (0%)		22 (13.75%)	1 (1.82%)	
Other	3 (5.3%)	8 (2.94%)	0 (0%)	0 (0%)		4 (2.5%)	2 (3.64%)	
Diagnosis N (%):								
Normal Cognition	57 (100%)	103 (38.23%)	22 (88%)	134 (46.21%)	0.0003	160 (100%)	0 (0%)	
Mild Cognitive Impairment	0 (0%)	110 (40.82%)	2 (8%)	103 (35.52%)		0 (0%)	0 (0%)	
Alzheimer’s	0 (0%)	54 (19.85%)	1 (4%)	52 (17.93%)		0 (0%)	55 (100%)	
Parental AD Diagnosis N (%):	57 (100%)	269 (100%)	25 (100%)	290 (100%)	>0.999	160 (100%)	55 (100%)	>0.999
APOE N (%):					>0.999			0.0002
E2(+)	9 (15.79%)	29 (10.82%)	2 (8%)	28 (9.66%)		24 (15%)	1 (1.82%)	
E4(+)	38 (80.85%)	128 (47.58%)	14 (56%)	189 (65.17%)		83 (51.88%)	55 (100%)	

Table 2: SNP Summary For All Screened Variants.

Chr: chromosome; SNV: single nucleotide variation; Cons: consequence; MM: missense mutation; IV: intron variant; InV: intergenic variant; MV: missense variant; SV: Synonymous Various; RRV: Regulatory Region Variant; NCTV: Non-Coding Transcript Variant; 3PV: 3 Prime UTR Variant; SG: Stop Gained; Ref Allele: Reference allele as identified in dbSNP database. RAF: reference allele frequency; RAF in total, European, African, and African American (AA) populations determined from dbSNP from cohorts TopMed, gnomAD, gnomAD, and the PAGE study, respectively. Risk allele: variant associated with increased risk or progression of AD or the allele not associated with protection. protective allele: variant associated with reduced risk or progression of AD.

SNP	Chr	Gene	Type	Cons	Ref Allele	Alt Allele	RAF (total)	RAF (Euro)	RAF (African)	RAF (AA)	ASCEND	ADNI	Risk Allele	Prot. Allele	Source
rs28936379	1	PSEN2	SNV	MM	A	C, G, T	1.000	1.000	1.000	1.000	yes	no	G		Levesque et al. 1995
rs6656401	1	CR1	SNV	IV	A	G, T	0.179	0.190	0.032	0.033	yes	no	A		Shen et al. 2015
rs10933431	2	INPP5D	SNV	IV	G	C	0.372	0.236	0.628	0.622	yes	no	C	G	Jansen et al. 2019
rs35349669	2	INPP5D	SNV	IV	C	T	0.617	0.512	0.866	0.863	yes	no	T		Lambert et al. 2013
rs6733839	2	LOC105373605	SNV	IV	C	G, T	0.610	0.613	0.598	0.597	yes	no	T		Desikan et al. 2017
rs7561528	2	LOC105373605	SNV	IV	G	A	0.685	0.668	0.795	0.794	yes	yes	A		Lo et al. 2019
rs16824536	3	MME	SNV	IV	G	A	0.893	0.949	0.737	0.740	yes	no	G	A	Bellenguez et al. 2020
rs3822030	4	IDUA	SNV	IV	G	T	0.443	0.429	0.766	0.763	no	yes	T	G	Bellenguez et al. 2020; Vuckovic et al. 2020
rs6846529	4	CLNK	SNV	InV	C	G, T	0.280	0.275	0.389	0.390	no	yes	C		Bellenguez et al. 2020; Vuckovic et al. 2020
rs190982	5	MEF2C-AS1	SNV	IV	G	A, C	0.377	0.397	0.103	0.104	yes	yes	A		Lambert et al. 2013
rs871269	5	TNIP1	SNV	IV	C	A, T, G	0.659	0.673	0.593	0.595	yes	yes	C	T	Wightman et al. 2020
rs10947943	6	UNC5CL	SNV	IV	G	A	0.890	0.844	0.976	0.968	yes	no	G	A	Bellenguez et al. 2020
rs3129882	6	HLA-DRA	SNV	IV	G	A, C	0.443	0.435	0.431	0.432	yes	yes	G		Hamzaet al. 2010
rs75932628	6	TREM2/LOC105375056	SNV	MV/IV	C	A,T	0.998	0.998	1.000	1.000	yes	no	T		Bellenguez et al. 2020; Vuckovic et al. 2020
rs785129	6	HS3ST5/HDAC2-AS2	SNV	IV	T	C	0.343	0.340	0.233	0.234	no	yes	C		Bellenguez et al. 2020; Vuckovic et al. 2020
rs9296559	6	CD2AP	SNV	IV	T	C	0.740	0.735	0.797	0.795	no	yes	C		Tao et al. 2019

SNP	Chr	Gene	Type	Cons	Ref Allele	Alt Allele	RAF (total)	RAF (Euro)	RAF (African)	RAF (AA)	ASCEND	ADNI	Risk Allele	Prot. Allele	Source
rs10952097	7	ICA1	SNV	IV	T	A, C	0.231	0.099	0.553	0.532	no	yes	T	C	Bellenguez et al. 2020
rs11767557	7	EPHA-1-AS1	SNV	IV	T	C	0.806	0.803	0.845	0.845	yes	yes	T	C	Naj et al. 2011
rs11771145	7	EPHA1-AS1	SNV	IV	G	A, T	0.648	0.655	0.446	0.447	yes	yes	G	A	Lambert et al. 2013
rs3211956	7	CD36	SNV	IV	T	G	0.922	0.921	0.965	0.964	no	yes	T	G	Zhou et al. 2018
rs11136000	8	CLU	SNV	IV	T	A, C	0.392	0.397	0.549	0.551	yes	yes	C		Haroldet al. 2009
rs2279590	8	CLU	SNV	IV	T	A, C, G	0.386	0.409	0.105	0.108	yes	no	C		Jun et al. 2015
rs7068231	10	ANK3	SNV	InV	T	A, G	0.417	0.409	0.258	0.259	no	yes	T	G	Bellenguez et al. 2020; Vuckovic et al. 2020
rs1562990	11	MS4A	SNV	IV	C	A, T	0.409	0.413	0.203	0.206	yes	no	A	C	Antúnez et al. 2011
rs4938933	11	MS4A4A	SNV	IV	C	G, T	0.404	0.408	0.318	0.319	yes	yes	T	C	Hollingworth et al. 2011
rs541458	11	PICALM	SNV	RRV	C	T	0.315	0.318	0.205	0.195	yes	no	T	C	Zhuang et al. 2019
rs610932	11	MS4A6A	SNV	3PV	T	C, G	0.430	0.430	0.485	0.484	no	yes	C	A	Hollingworth et al. 2011
rs983392	11	MS4A2	SNV	InV	A	G	0.664	0.601	0.932	0.929	yes	no	A	G	Lambert et al. 2013
rs4985556	16	IL34	SNV	SG	C	A, T	0.894	0.888	0.976	0.975	yes	yes	A		Bellenguez et al. 2020; Vuckovic et al. 2020
rs56407236	16	FAM157C/LOC105376786	SNV	IV/MCTV	G	A	0.935	0.935	0.927	0.928	yes	no	G		Bellenguez et al. 2020; Vuckovic et al. 2020
rs889555	16	BCKDK	SNV	IV	C	T	0.708	0.720	0.604	0.603	no	yes	C	T	Bellenguez et al. 2020; Vuckovic et al. 2020
rs199515	17	LRRC37A2/WNT3	SNV	IV	G	A, C	0.171	0.186	0.116	0.117	yes	no	C		Bellenguez et al. 2020
rs4277405	17	ACE	SNV	InV	C	A, T, G	0.375	0.380	0.354	0.353	yes	no	T	C	Bellenguez et al. 2020; Vuckovic et al. 2020
rs7225151	17	SCIMP	SNV	IV	G	A	0.877	0.880	0.770	0.771	yes	yes	A		Bellenguez et al. 2020; Vuckovic et al. 2020
rs3752246	19	ABCA7	SNV	MV	G	C, T	0.158	0.171	0.046	0.046	yes	no	G		Naj et al. 2011
rs3764650	19	ABCA7	SNV	IV	T	G	0.897	0.906	0.740	0.741	yes	yes	G		Hollingworth et al. 2011
rs3826656	19	CD33	SNV	IV	G	A	0.255	0.240	0.185	0.186	yes	yes	G	A	Yuan et al. 2011

SNP	Chr	Gene	Type	Cons	Ref Allele	Alt Allele	RAF (total)	RAF (Euro)	RAF (African)	RAF (AA)	ASCEND	ADNI	Risk Allele	Prot. Allele	Source
rs429358	19	APOE	SNV	MV	T	C	0.926	0.932	0.871	0.872	yes	no	C	A	Arboleda-Velasquez et al. 2019
rs7412	19	APOE	SNV	MV	C	T	0.917	0.917	0.895	0.896	yes	no	C		Kettunen et al. 2012
rs769449	19	APOE	SNV	IV	G	A	0.901	0.898	0.976	0.975	yes	yes	A		Liu et al. 2018
rs9304690	19	SIGLEC11	SNV	SV	C	T	0.771	0.758	0.896	0.895	no	yes	T		Bellenguez et al. 2020; Vuckovic et al. 2020
rs16979934	20	CASS4	SNV	IV	T	C, G	0.949	0.947	0.944	0.944	no	yes	C		Hollingworth et al. 2011
rs6014724	20	CASS4	SNV	IV	A	C, G	0.896	0.904	0.897	0.897	no	yes	A	G,T	Jansenet al. 2020
rs7274581	20	CASS4	SNV	IV	T	C	0.906	0.914	0.797	0.797	yes	no	C		Desikan et al. 2017
rs2154481	21	APP	SNV	IV	C	A, T	0.504	0.477	0.947	0.945	no	yes	T		Bellenguez et al. 2020
rs2830489	21	ADAMTS1	SNV	InV	C	A, T	0.864	0.800	0.971	0.970	yes	no	C	T	Bellenguez et al. 2020

Table 3: Race Association Analysis Results

Disease associated alleles are bolded. MAF: minor allele frequency. UNADJ P: Asymptotic p-value calculated using basic allelic test chi-square. Final significance determined by controlling the false discovery rate using the Benjamini-Hochberg procedure.

SNP	Major Allele	Minor Allele	MAF Black (n=25)	MAF White (n=290)	UNADJ P	Rank	(i/m)Q	significant?
rs983392	**A**	G	0.071	0.453	**8.09E-06**	1	0.001	**yes**
rs1562990	**A**	C	0.119	0.394	**0.001**	2	0.002	**yes**
rs2279590	**C**	T	0.071	0.333	**0.001**	3	0.003	**yes**
rs7274581	T	**C**	0.381	0.106	**0.002**	4	0.004	**yes**
rs11771145	**G**	A	0.560	0.340	**0.003**	5	0.006	**yes**
rs3764650	T	**G**	0.229	0.095	**0.004**	6	0.007	**yes**
rs190982	**A**	G	0.160	0.361	**0.006**	7	0.008	**yes**
rs541458	**T**	C	0.119	0.333	**0.014**	8	0.009	no
rs6656401	G	**A**	0.048	0.250	**0.018**	9	0.010	no
rs2830489	**C**	T	0.071	0.273	**0.020**	10	0.011	no
rs56407236	**G**	A	0.119	0.016	**0.022**	11	0.012	no
rs35349669	C	**T**	0.191	0.422	**0.023**	12	0.013	no
rs7225151	G	**A**	0.260	0.148	**0.035**	13	0.014	no
rs769449	G	**A**	0.060	0.185	**0.036**	14	0.016	no
rs10933431	**C**	G	0.548	0.333	**0.038**	15	0.017	no
rs3826656	A	**G**	0.140	0.268	**0.055**	16	0.018	no
rs7068231	G	**T**	0.750	0.414	0.057			
rs6733839	C	**T**	0.571	0.394	0.067			
rs3752246	C	**G**	0.050	0.167	0.084			
rs610932	**C**	A	0.125	0.400	0.124			
rs16824536	**G**	A	0.238	0.141	0.154			

SNP	Major Allele	Minor Allele	MAF Black (n=25)	MAF White (n=290)	UNADJ P	Rank	(i/m)Q	significant?
rs7561528	G	**A**	0.420	0.322	0.162			
rs871269	**C**	T	0.420	0.331	0.209			
rs16979934	T	**C**	0.125	0.042	0.264			
rs9296559	T	**C**	0.125	0.265	0.334			
rs2154481	**T**	C	0.625	0.460	0.362			
rs6014724	**A**	G	0.000	0.084	0.385			
rs3211956	**T**	G	0.000	0.086	0.391			
rs11767557	**T**	C	0.140	0.189	0.396			
rs199515	**C**	G	0.095	0.152	0.409			
rs9304690	C	**T**	0.125	0.250	0.416			
rs75932628	C	**T**	0.000	0.015	0.430			
rs11136000	**C**	T	0.360	0.412	0.448			
rs429358	T	**C**	0.310	0.379	0.478			
rs7412	**C**	T	0.048	0.076	0.557			
rs6846529	T	**C**	0.375	0.281	0.558			
rs889555	**C**	T	0.375	0.281	0.578			
rs3822030	**T**	G	0.375	0.468	0.587			
rs785129	**C**	T	0.250	0.327	0.635			
rs4938933	**T**	C	0.400	0.368	0.653			
rs4277405	**T**	C	0.405	0.439	0.740			
rs10947943	**G**	A	0.048	0.061	0.772			
rs3129882	A	**G**	0.420	0.412	0.908			
rs4985556	C	**A**	0.080	0.078	0.959			
rs10952097	C	**T**	0.125	0.120	0.965			

Table 4: Race Genotyping Results

Counts of individuals with genotypes including at least one disease-associated allele (bolded) and genotypes without any disease-associated allele were analyzed with chi-square test. Final significance determined by controlling the false discovery rate using the Benjamini-Hochberg procedure.

SNP	Genotype	Black (n)	White (n)	p-value	(i/M)Q	significant?	OR	95% CI
rs11771145	**G G or G A**	16	253	**0.004**	0.001	no	0.253	0.1053 to 0.6062
	A A	9	36					
rs7274581	**C C or C T**	12	6	**0.007**	0.002	no	6.000	1.667 to 18.10
	T T	9	27					
rs6656401	**A A or A G**	2	12	**0.029**	0.003	no	0.175	0.03642 to 0.7974
	G G	19	20					
rs3764650	**G G or G T**	9	53	**0.033**	0.004	no	2.672	1.105 to 6.340
	T T	15	236					
rs190982	**A A or A G**	25	245	**0.033**	0.006	no	Infinity	1.250 to Infinity
	G G	0	44					
rs769449	**A A or A G**	3	91	**0.042**	0.007	no	0.297	0.09184 to 0.9895
	G G	22	198					
rs35349669	**T T or T C**	7	20	0.051			0.300	0.09478 to 0.9403
	C C	14	12					
rs7561528	**A A or A G**	18	152	0.093			2.301	0.9873 to 6.090
	G G	7	136					
rs3752246	**G G or G C**	2	10	0.105			0.256	0.05204 to 1.227
	C C	18	23					
rs7225151	**A A or A G**	11	82	0.113			1.983	0.9075 to 4.448
	G G	14	207					

SNP	Genotype	Black (n)	White (n)	p-value	(i/M)Q	significant?	OR	95% CI
rs6733839	**T T or T C**	18	21	0.120			3.429	0.9356 to 12.53
	C C	3	12					
rs3826656	**G G or A G**	7	131	0.140			0.469	0.1771 to 1.093
	A A	18	158					
rs983392	**A A or A G**	21	27	0.144			Infinity	0.9875 to Infinity
	G G	0	5					
rs1562990	**A A or A C**	21	29	0.148			Infinity	0.6184 to Infinity
	C C	0	4					
rs2830489	**C C or C T**	21	29	0.148			Infinity	0.6184 to Infinity
	T T	0	4					
rs541458	**T T or C T**	21	29	0.148			Infinity	0.6184 to Infinity
	C C	0	4					
rs871269	**C C or C T**	20	256	0.204			0.516	0.1810 to 1.331
	T T	5	33					
rs2279590	**C C or C T**	21	30	0.274			Infinity	0.5624 to Infinity
	T T	0	3					
rs16979934	**C C or C T**	1	20	0.288			3.933	0.2894 to 27.18
	T T	3	236					
rs7068231	**T T or T G**	4	168	0.303			Infinity	0.5073 to Infinity
	G G	0	88					
rs10933431	**C C or C G**	15	28	0.305			0.446	0.1223 to 1.571
	G G	6	5					
rs6846529	**C C or C T**	3	123	0.357			3.244	0.4767 to 42.41
	T T	1	133					
rs3129882	**G G or G A**	19	192	0.382			1.600	0.6525 to 3.974
	A A	6	97					

SNP	Genotype	Black (n)	White (n)	p-value	(i/M)Q	significant?	OR	95% CI
rs3822030	**T T or T G**	4	203	0.586			Infinity	0.2457 to Infinity
	G G	0	52					
rs11767557	**T T or C T**	24	278	0.606			0.863	0.1342 to 9.753
	C C	1	10					
rs9296559	**C C or C T**	1	124	0.623			0.355	0.02712 to 2.413
	T T	3	132					
rs9304690	**T T or C T**	1	111	0.636			0.432	0.03303 to 2.940
	C C	3	144					
rs11136000	**C C or C T**	21	249	0.764			0.843	0.2759 to 2.376
	T T	4	40					
rs4985556	**A A or A C**	3	45	0.779			0.733	0.2237 to 2.309
	C C	22	242					
rs10947943	**G G or A G**	21	33	>0.999			0.000	to Infinity
	A A	0	0					
rs10952097	**T T or T C**	1	52	>0.999			1.308	0.09893 to 8.908
	C C	3	204					
rs16824536	**G G or G A**	21	32	>0.999			0.300	to Infinity
	A A	0	0					
rs199515	**C C or C G**	21	32	>0.999			Infinity	0.07071 to Infinity
	G G	0	1					
rs2154481	**T T or C T**	3	197	>0.999			0.853	0.1252 to 11.26
	C C	1	56					
rs3211956	**T T or T G**	4	254	>0.999			Infinity	0.006587 to Infinity
	G G	0	2					
rs4277405	**T T or C T**	17	26	>0.999			1.144	0.2774 to 3.909
	C C	4	7					

SNP	Genotype	Black (n)	White (n)	p-value	(i/M)Q	significant?	OR	95% CI
rs429358	**C C or C T**	12	19	>0.999			0.983	0.3223 to 3.172
	T T	9	14					
rs4938933	**T T or T C**	22	250	>0.999			1.144	0.3586 to 3.767
	C C	3	39					
rs56407236	**G G or A G**	21	31	>0.999				
	A A	0	0					
rs6014724	**A A or A G**	4	254	>0.999			Infinity	0.001743 to Infinity
	G G	0	1					
rs610932	**C C or C A**	4	213	>0.999			Infinity	0.1925 to Infinity
	A A	0	43					
rs7412	**C C or C T**	21	33	>0.999			1.144	to Infinity
	T T	0	0					
rs75932628	**T T ot T C**	0	1	>0.999			0.000	0.000 to 14.14
	C C	21	32					
rs785129	**C C or C T**	4	233	>0.999			Infinity	0.09117 to Infinity
	T T	0	23					
rs889555	**C C or C T**	4	226	>0.999			Infinity	0.1159 to Infinity
	T T	0	28					

Table 5: Diagnosis Association Analysis Results

Disease associated alleles are bolded. MAF: minor allele frequency. UNADJ P: Asymptotic p-value calculated using basic allelic test chi-square. Final significance determined by controlling the false discovery rate using the Benjamini-Hochberg procedure.

SNP	Major Allele	Minor Allele	MAF No AD (n=161)	MAF AD (n=55)	UNADJ P	(i/m)Q	Significant?
rs7225151	G	**A**	0.178	0.077	**0.013**	0.002	no
rs11136000	**C**	T	0.369	0.462	0.092		no
rs6014724	**A**	G	0.068	0.125	0.093		no
rs769449	G	**A**	0.159	0.096	0.111		no
rs3764650	T	**G**	0.116	0.067	0.155		no
rs11767557	**T**	C	0.217	0.154	0.163		no
rs3211956	**T**	G	0.073	0.115	0.210		no
rs3129882	A	**G**	0.403	0.462	0.294		no
rs11771145	**G**	A	0.403	0.346	0.301		no
rs4985556	C	**A**	0.089	0.058	0.316		no
rs9296559	T	**C**	0.291	0.240	0.343		no
rs3826656	A	**G**	0.247	0.202	0.348		no
rs610932	**C**	A	0.432	0.385	0.424		no
rs7068231	G	**T**	0.422	0.462	0.511		no
rs871269	**C**	T	0.325	0.356	0.563		no
rs3822030	**T**	G	0.447	0.481	0.569		no
rs9304690	C	**T**	0.248	0.226	0.669		no
rs4938933	**T**	C	0.388	0.365	0.687		no
rs889555	**C**	T	0.287	0.308	0.708		no
rs6846529	T	**C**	0.296	0.279	0.752		no
rs190982	**A**	G	0.363	0.346	0.763		no
rs7561528	G	**A**	0.316	0.327	0.830		no

SNP	Major Allele	Minor Allele	MAF No AD (n=161)	MAF AD (n=55)	UNADJ P	(i/m)Q	Significant?
rs785129	**C**	T	0.320	0.327	0.908		no
rs16979934	T	**C**	0.049	0.048	0.986		no
rs2154481	**T**	C	0.451	0.451	1.000		no

Table 6: Diagnosis Genotyping Results

AD: Alzheimer's Disease, Counts of individuals with genotypes including at least one disease-associated allele (bolded) and genotypes without any disease-associated allele were analyzed with chi-square test. Final significance determined by controlling the false discovery rate using the Benjamini-Hochberg procedure.

SNP	Genotype	At Risk (n)	AD (n)	P Value	(i/m)Q	Significant?	OR	95% CL
rs2154481	**T T or C T**	78	46	**0.000**	0.002	**yes**	0.238	0.1113 to 0.5393
	C C	57	8					
rs7225151	**A A or A G**	54	10	**0.040**	0.004	no	2.292	1.111 to 4.656
	G G	106	45					
rs11767557	**T T or C T**	151	55	0.117			0.000	0.000 to 1.299
	C C	8	0					
rs610932	**C C or C A**	81	48	0.203			0.537	0.2112 to 1.287
	A A	22	7					
rs7068231	**T T or T G**	66	41	0.213			0.609	0.2965 to 1.244
	G G	37	14					
rs3764650	**G G or G T**	36	8	0.247			1.720	0.7347 to 3.730
	T T	123	47					
rs11136000	**C C or C T**	141	45	0.256			1.649	0.7202 to 3.886
	T T	19	10					
rs9296559	**C C or C T**	54	24	0.320			1.423	0.7281 to 2.687
	T T	49	31					
rs769449	**A A or A G**	43	11	0.370			1.470	0.7169 to 3.207
	G G	117	44					
rs4985556	**A A or A C**	27	6	0.387			1.683	0.6526 to 4.154
	C C	131	49					
rs3129882	**G G or G A**	108	41	0.398			0.709	0.3508 to 1.411
	A A	52	14					

SNP	Genotype	At Risk (n)	AD (n)	P Value	(i/m)Q	Significant?	OR	95% CL
rs6846529	**C C or C T**	53	24	0.405			1.369	0.7003 to 2.583
	T T	50	31					
rs3826656	**G G or A G**	67	20	0.526			1.261	0.6825 to 2.313
	A A	93	35					
rs11771145	**G G or G A**	133	48	0.528			0.718	0.2733 to 1.738
	A A	27	7					
rs9304690	**T T or C T**	45	21	0.612			1.219	0.6109 to 2.385
	C C	58	33					
rs871269	**C C or C T**	144	48	0.615			1.313	0.5133 to 3.392
	T T	16	7					
rs190982	**A A or A G**	135	48	0.667			0.788	0.2983 to 1.939
	G G	25	7					
rs16979934	**C C or C T**	8	5	0.769			0.842	0.2707 to 2.387
	T T	95	50					
rs785129	**C C or C T**	93	51	0.772			0.729	0.2424 to 2.337
	T T	10	4					
rs4938933	**T T or T C**	136	48	0.825			0.826	0.3159 to 2.055
	C C	24	7					
rs7561528	**A A or A G**	85	28	0.8758			1.093	0.6029 to 1.974
	G G	75	27					
rs10952097	**T T or T C**	20	11	>0.999			0.964	0.4175 to 2.087
	C C	83	44					
rs3211956	**T T or T G**	102	54	>0.999			1.889	0.09774 to 36.19
	G G	1	1					
rs3822030	**T T or T G**	45	40	>0.999			0.993	0.4455 to 2.165
	G G	17	15					

SNP	Genotype	At Risk (n)	AD (n)	P Value	(i/m)Q	Significant?	OR	95% CL
rs6014724	**A A or A G**	102	55	>0.999			0.000	0.000 to 16.85
	G G	1	0					
rs889555	**C C or C T**	91	50	>0.999			0.910	0.3318 to 2.649
	T T	10	5					

Supplementary

Supplementary File One: PLINK Code

Ascend

Files Used and Generated:

Data: ASCEND microarray results genotyped by the Wingo Lab as previously described in the Axiom™ Genotyping Solution Data Analysis User Guide

Mysnps.txt: List of SNPs of interest derived from Bellenguez et al. 2020, Giri, Zhang, Lü 2016, and Rotimi et al. 2016

ASCEND: Data containing ASCEND subjects who have SNPs of Interest

```
### down sample all SNPs assayed to include only SNPs of interest with ASCEND Cohort
./plink --file data --extract mysnps.txt --make-bed ASCEND
```

ADNI

Files Used and Generated:

Data: ADNI-1 microarray results produced by ADNI researchers

Mysnps.txt: List of SNPs of interest derived from Bellenguez et al. 2020, Giri, Zhang, Lü 2016, and Rotimi et al. 2016

ADNI: Data containing ADNI subjects who have SNPs of Interest

FHAD.txt: ADNI participants diagnosed with AD and had parental history

ADNIFH: Data containing ADNI subjects diagnosed with AD and have familial history of AD

```
### down sample all SNPs assayed to include only SNPs of interest with ADNI Cohort
./plink --file data --extract mysnps.txt --make-bed ADNI

### Extracted Subset of ADNI participants with AD and familial history of AD
./plink --bfile ADNI --keep FHAD.txt --make-bed --out ADNIFH
```

Merged ADNI-ASCEND

Files Used and generated:

ADNIFH: Data containing ADNI subjects diagnosed with AD and have familial history of AD

ASCEND: Data containing ASCEND subjects who have SNPs of Interest

Merged.missnp: List of 5 SNP variants that have more than two alleles

Out Source 1: Data that contain SNPS of interest excluding multi-allelic SNPS in ADNI Cohort

Out Source 2: Data that contain SNPS of interest excluding multi-allelic SNPS in ASCEND Cohort

MergedFiles: Merged ASCEND and ADNI participants

CombinedRace: List containing self-reported race of ADNI and ASCEND participants

Diagnosis: List containing cognitive status of ADNI and ASCEND participants

```
### ADNI with parental history without five multi-allelic SNPS
./plink --bfile ADNIFH --exclude merged.missnp --make-bed --out source1

#### ASCEND cohort without five multi-allelic SNPS
```

./plink --bfile ASCEND --exclude merge.missnp --make-bed --out source2

###Merged Cohorts

./plink --bfile out source1 --bmerge out source2 --make-bed --out mergedfiles

###Merged Cluster Race Frequency Analysis

./plink --bfile mergedfiles --freq --within CombinedRace.txt --out freq_stat

Race Association analysis within Merged ASCEND- ADNI cohort by race with Sidak's correction

./plink --bfile mergedfiles --pheno CombinedRace.txt --all-pheno --assoc --adjust --out AssocAnalysis

###Merged Diagnosis Frequency Analysis

./plink --bfile mergedfiles --freq --within diagnosis.txt --out freq_stat1

Diagnosis Association analysis within Merged ASCEND- ADNI cohort by race with Sidak's correction

./plink --bfile mergedfiles --pheno diagnosis.txt --all-pheno --assoc --adjust --out AssocAnalysis1