Supplemental Online Content

Le Guen Y, Raulin A-C, Logue MW, et al. Association of African ancestry–specific *APOE* missense variant R145C with risk of Alzheimer disease. *JAMA*. Published February 21, 2023. doi:10.1001/jama.2023.0268

eAppendix 1. ADSP Cohort Descriptions

eAppendix 2. Supplemental Methods

eAppendix 3. Additional Acknowledgments

eFigure 1. Local Ancestry Inference at rs769455 in Admixed American Participants in gnomAD

eFigure 2. Flowchart Describing the *APOE* Alleles and R145C Genotyping Among African and Admixed-African Individuals Included in the Discovery Analysis

eFigure 3. Principal Ancestry Components Computed on the Merge of the ADSP (Discovery) African Ancestry Participants and African Participants in 1000 Genomes Project and Human Genome Diversity Project

eFigure 4. Principal Ancestry Components Computed on the Merge of the ADGC (Replication) African Ancestry Participants and African Participants in 1000 Genomes Project and Human Genome Diversity Project

eFigure 5. Age-at-AD-Onset Distribution Across *APOE* Genotype Groups in the ADSP Discovery and ADGC Replication in Function of R145C Allele

eFigure 6. R145C Association With Alzheimer Disease Risk and Age-at-AD-Onset

eFigure 7. *APOE* ε3[R145C]/ε4 Individuals Have an Alzheimer Disease Risk Comparable to *APOE* ε4/ε4 Individuals Regardless of the African Ancestry Cutoff

eFigure 8. ApoE3-R145C Shows Significantly Reduced Heparin Binding Compared to apoE3

eTable 1. Missense Mutations on *APOE* Canonical Transcript

eTable 2. Cohorts Overview

eTable 3. Overview of Alzheimer's Disease Sequencing Project Studies

(continued on next page)

eTable 4. Pathogenic Variants Identified on *APP*, *PSEN1*, *PSEN2*, *MAPT* in ADSP WES and WGS

eTable 5. Demographics of All the Queried Cohorts

eTable 6. R145C Per Cohort, Diagnosis and APOE Genotypes

eTable 7. *APOE* Variants Genotype Call in the Data Released by NIAGADS

eTable 8. Demographics Per Cohort After Ancestry Selection, Quality Control and Duplicates Removal

eTable 9. *APOE* R145C (rs769455) Allelic Breakdown by *APOE* Genotype

eTable 10. African Ancestry Cutoff Sensitivity Analyses

eTable 11. African Ancestry Cutoff Sensitivity Analyses Restricted to Individuals With African Local Ancestry at Both of Their *APOE* Haplotypes

eTable 12. Mega-analysis of the Discovery and Replication Samples

eTable 13. African Ancestry Cutoff Sensitivity Analyses Only Including Individuals Directly Genotyped in the ADSP WES Data Distributed by NIAGADS

eTable 14. Primary and Secondary Analyses Considering Standard Model (Non-stratified by *APOE* Genotype) and Adjusting for ε2 and ε4 Dosages

eTable 15. Interaction Between R145C*ε4 for Association With Alzheimer Disease Risk, Considering Standard Model (Non-stratified by APOE Genotype) and Adjusting for ε2 and ε4 Dosages

eTable 16. Risk of Alzheimer Disease by *APOE* Genotype Including *APOE* ε3[R145C] Subtypes

eTable 17. Association of R145C With CDR-SOB Changes With Age in ε3/ε4 Stratified Analysis

eTable 18. Association of R145C*ε4 With CDR-SOB Changes With Age in Unstratified Analysis

eReferences

This supplemental material has been provided by the authors to give readers additional information about their work.

eAppendix 1. ADSP Cohort Descriptions

Adult Changes in Thought (ACT)

The Adult Changes in Thought (ACT) study is a longitudinal prospective cohort study that began in 1994. Participants are randomly selected from Seattle area members of Group Health aged 65 years or older. Participants are cognitively intact at study enrollment, defined operationally as a CASI (Cognitive Abilities Screening Instrument) score of >85 or consensus diagnosis of "not demented" following comprehensive neurological and neuropsychological assessment. Incident cases of dementia are identified with the same 2-stage sampling scheme, where all participants with CASI scores <86 are evaluated with a comprehensive neurological and neuropsychological assessment, and all results are considered at a consensus conference. Autopsy is also available for this collection; autopsy consent rates are about 25% of the cohort. Genomic DNA from blood and/or brain tissue is available from this collection. This study includes 4690 subjects across three sub-cohorts.

Website:
https://www.maelstrom-research.org/mica/individual-study/act

Alzheimer's Disease Neuroimaging Initiative (ADNI)

ADNI is a global research study that actively supports the investigation and development of treatments that slow or stop the progression of AD. In this multisite longitudinal study, researchers at 63 sites in the US and Canada track the progression of AD in the human brain with clinical, imaging, genetic and biospecimen biomarkers through the process of normal aging, early mild cognitive impairment (EMCI), and late mild cognitive impairment (LMCI) to dementia or AD. Participants undergo a series of initial tests that are repeated at intervals over subsequent years, including a clinical evaluation, neuropsychological tests, genetic testing, lumbar puncture, and MRI and PET scans. The overall goal of ADNI is to validate biomarkers for use in Alzheimer' s disease clinical treatment trials. 1338 cases and 483 controls are included in this study across 4 phases.

Atherosclerosis Risk in Communities (ARIC)

The Atherosclerosis Risk in Communities Study (ARIC), sponsored by the National Heart, Lung, and Blood Institute (NHLBI) is a prospective epidemiologic study conducted in four U.S. communities. ARIC is designed to investigate the causes of atherosclerosis and its clinical outcomes, and variation in cardiovascular risk factors, medical care, and disease by race, gender, location, and date. To date, the ARIC study has published over 2,000 articles in peer-reviewed journals. ARIC includes two parts: the Cohort Component and the Community Surveillance Component. The Cohort Component began in 1987, and each ARIC field center randomly selected and recruited a cohort sample of approximately 4,000 individuals aged 45-64 from a defined population in their community, to receive extensive examinations, including medical, social, and demographic data. Follow-up also occurs semi-annually, by telephone, to maintain

contact and to assess health status of the cohort. In the Community Surveillance Component, the four communities are investigated to determine the long term trends in hospitalized myocardial infarction (MI) and coronary heart disease (CHD) deaths in approximately 470,000 men and women aged 35-84 years.
Website:
https://sites.cscc.unc.edu/aric/desc_pub

Cache County Study (CCS)

The Cache County Study on Memory Health and Aging (CCS) was initiated in 1994 to investigate the association of APOE genotype and environmental exposures on cognitive function and dementia. This cohort of 5092 Cache County, Utah, residents (90% of those aged 65 years or older in 1994), has been followed continuously for over 15 years, with four triennial waves of data collection and additional clinical assessments for those at high-risk for dementia. DNA samples were obtained from 97.6% of participants. The Cache County population is exceptionally long-lived and ranked number one in life expectancy among all counties in the 1990 US Census. All but one of the members of the CCS have been linked to the UPDB and their extended genealogies are known. This population was the source of most of the Centre d'Etude du Polymorphisme Humain (CEPH) families that have been used to represent Caucasians in many genetic studies worldwide, including the HapMap project. Recent analyses confirm that these data are representative of the general European-American population. For this study, we needed both AD cases and resilient individuals identified in the same pedigrees." (Ridge, P.G., Karch, C.M., Hsu, S. et al. Linkage, whole genome sequence, and biological data implicate variants in RAB10 in Alzheimer's disease resilience. Genome Med 9, 100 (2017). https://doi.org/10.1186/s13073-017-0486-1)

Cardiovascular Health Study (CHS)

The Cardiovascular Health Study (CHS) is an NHLBI-funded observational study of risk factors for cardiovascular disease in adults 65 years or older (n=5888, including a secondary cohort of predominately African-American subjects (n=687)). Starting in 1989, and continuing through 1999, participants underwent annual extensive clinical examinations. Measurements included traditional risk factors such as blood pressure and lipids as well as measures of subclinical disease, including echocardiography of the heart, carotid ultrasound, and cranial magnetic-resonance imaging (MRI). Examination also included cognitive measures. At six month intervals between clinic visits, and once clinic visits ended, participants were contacted by phone to ascertain hospitalizations and health status. The main outcomes are coronary heart disease (CHD), angina, heart failure (HF), stroke, transient ischemic attack (TIA), claudication, and mortality. Participants continue to be contacted by phone every 6 months.
Website:
https://chs-nhlbi.org/

Case Western Reserve University (CWRU) Autopsy

The Case Western Reserve University (CWRU) Autopsy Cohort was an Alzheimer's Disease Research Center (ADRC) clinic-based sample. Subjects included in the Case Western Reserve Autopsy cohort include individuals who were participants in the Brain Health and Memory Center Research Brain Donation Protocol at University Hospitals Cleveland Medical Center in Cleveland, OH. The target population is patients who have degenerative disorders of the central nervous system. Recruitment of potential subjects is primarily based on referrals from health care providers. In addition, individuals who learn of the program through other means may participate. This is an autopsy-based study designed to analyze post-mortem brain tissue. Medical records are also obtained as available to help correlate brain behavior relationships.

Participants were classified into clinical categories (Alzheimer's disease, Controls, and other dementia (ADRD) based on description of the brain gross examination and neuropath microscopic description. Cases consisted of individuals who were diagnosed as AD. Controls consisted of individuals who had insufficient findings of neurofibrillary tangles or BRAAK stage I/II or essentially normal brain for age. Other (ADRD) consisted of individuals with diagnosis not being AD or Control.

Case Western Reserve University (CWRU) Rapid Decline

The Case Western Reserve University (CWRU) Rapid Decline Cohort (Cohen et al. 2015) represents patients with rapid decline. Subjects included in the Case Western Reserve Rapid Decline dataset include individuals who were initially suspected of having Creutzfelt-Jakob disease (CJD). The Brains of these individuals were obtained by the National Prion Disease Pathology Surveillance Center (NPDPSC) for testing and confirmation of CJD. Samples are received from across the United States. After testing, these samples were determined not to have CJD, but to have pathology consistent with Alzheimer disease. The usual progression of CJD is quite rapid, so that the vast majority of these individuals progressed from onset to death in less than 3 years. This is an autopsy-based study designed to analyze post-mortem brain tissue. Medical records are obtained when available to help correlate brain behavior relationships.

Participants were classified as either Alzheimer's disease or another dementia (ADRD; Lewy body Dementia, Frontotemporal dementia) based on description of the brain gross examination and neuropath microscopic description.

Chicago Health and Aging Project (CHAP)

The Chicago Healthy Aging Project (CHAP) is a longitudinal population study of an urban general population sample (n= 10,000+) lasting from 1993 to 2012 of common chronic health problems of older persons, especially of risk factors for incident Alzheimer's disease, based in three neighborhoods on the south side of Chicago. An initial enrollment was supplemented by enrollment of successive age cohorts of community residents as they attained the age of 65. After the enrollment period, the CHAP pursued a complex strategy for follow-up evaluations, interviewing all participants about every

three years and conducting in-depth clinical evaluations among a stratified random sample of participants at each of these cycles. Cognition was assessed for all CHAP subjects and the presence of Alzheimer's disease (AD) was assessed for those in the Clinical Evaluation sample.
Clinical evaluation included a neuropsychological battery, structured neurological examination and medical history. For persons in whom there was evidence of dementia and uncertainty as to whether a stroke had occurred or its relation to detention, limited diagnostic use of brain magnetic resonance imaging (MRI) occurred. Diagnosis of dementia required loss of cognitive function by the neurologist's assessment and impairment in two or more functions on cognitive performance tests. The diagnosis of Alzheimer's disease was by criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) for probable Alzheimer's disease.
Website:
https://www.maelstrom-research.org/mica/individual-study/chap

Corticobasal Degeneration (CBD)

Patients with a clinical Parkinsonism in life and neuropathological confirmation of Corticobasal Degeneration (CBD) were identified from brain banks, research hospitals and neuropathologists. The top three contributing sites were the Mayo Clinic, University College London, and the University of Pennsylvania and additional small numbers of cases were obtained from various other institutions across the US and some international collaborators. The neuropathological diagnosis was made according to NINDS neuropathologic diagnostic criteria. DNA was extracted from brain tissue from patients who had consented for brain donation. DNA samples and/or tissue were sent to the University of Pennsylvania for preparation for genotyping.

Cuban American Alzheimer's Disease Initiative (CuAADI)

The Cuban American Alzheimer's Disease Initiative (CuAADI) is a convenience sample ascertained through community outreach to Alzheimer's and adult day care centers in Southern Florida, lay conferences, and Neurology and Memory Disorders clinics. Eligibility was based on self-reported Cuban heritage. The participants were recruited in South Florida which is home to the largest number of Cubans in the US. Most of our participants have been living in South Florida since they moved from Cuba in the last 20 to 45 years.
Participants were ascertained and evaluated through community centers, referrals from University of Miami memory clinics and adult day care centers. Some participants were evaluated in their homes. All participants greater than 60 years of age underwent a standard clinical evaluation consisting of a medical and family history interview, neuropsychological testing, behavioral and emotional assessments, and functional measures. Venous blood samples (or saliva samples when needed) were collected on all

participants. All assessments were conducted in the preferred language of the participant or knowledgeable informant.

Erasmus Rucphen Family (ERF)

A family-based cohort study that is embedded in the Genetic Research in Isolated Populations (GRIP) program in the South West of the Netherlands. The aim of this program was to identify genetic risk factors in the development of complex disorders. For the Erasmus Rucphen Family Study (ERF), 22 families that had at least five children baptized in the community church between 1850-1900 were identified with the help of genealogical records. All living descendants of these couples and their spouses were invited to take part in the study. Data collection started in June 2002 and was finished in February 2005 (n=2065).

Website:

https://www.neurodegenerationresearch.eu/cohort/erasmus-rucphen-family-study/

Estudio Familiar de Influencia Genetica en Alzheimer (EFIGA)

Estudio Familiar de Influencia Genetica en Alzheimer (EFIGA) included 683 at-risk family members from 242 AD-affected families of Caribbean Hispanic descent. These families have 2 or more individuals affected with Alzheimer's disease. A system of recruitment was also set up in the Dominican Republic with the help of several local physicians, including the president of the Dominican Society of Geriatrics and Gerontology. All affected and unaffected family members are evaluated in person both in the Dominican Republic and New York. A case was defined as any individual meeting NINCDS-ADRDA criteria for probable or possible LOAD. The Clinical Dementia Rating was used to rate the severity of dementia. Brain imaging and other laboratory study results were reviewed, when available, to ensure full implementation of the NINCDS-ADRDA criteria. Once patients with LOAD were identified, their illnesses were documented with standardized neurological and neuropsychological evaluations. Structured family history interviews were then conducted with available family members to determine whether patients had living siblings or relatives with the disease. Medical and neurological examinations were completed for all family members. Brains of participants with dementia and history of stroke were administered magnetic resonance imaging scans to exclude patients with comorbid cerebrovascular disease. DNA samples and cell lines are stored for all participating individuals.

The goal of this study is to root out genetic variants that increase late onset Alzheimer disease risk in this ethnic group. This study was initiated in 1998 and recruited subjects from the Taub Institute for Research on Alzheimer's Disease and the Aging Brain in New York as well as from clinics in the Dominican Republic.

Website:

http://www.cumc.columbia.edu/adrc/investigators

Framingham Heart Study (FHS)

The Framingham Heart Study, under the direction of the National Heart, Lung and Blood Institute (NHLBI), formerly known as the National Heart Institute, has been committed to identifying the common factors or characteristics that contribute to cardiovascular disease (CVD) since its beginning in 1948. FHS has followed CVD development over a long period of time in three generations of participants. The Study began in 1948 by recruiting an Original Cohort of 5,209 men and women between the ages of 30 and 62 from the town of Framingham, Massachusetts, who had not yet developed overt symptoms of cardiovascular disease or suffered a heart attack or stroke. Since that time the Study has added an Offspring Cohort in 1971, the Omni Cohort in 1994, a Third Generation Cohort in 2002, a New Offspring Spouse Cohort in 2003, and a Second Generation Omni Cohort in 2003.

Data collected over the course of FHS have included those derived from physical examinations, lifestyle interviews, detailed medical histories, and laboratory testing. DNA has been collected from blood samples for the Original, Offspring, and Third Generation Cohorts. Available phenotype information includes quantitative measures of the major CVD risk factors such as systolic blood pressure, total and HDL cholesterol, fasting glucose, and cigarette use, as well as anthropomorphic measures such as body mass index, biomarkers such as fibrinogen and CRP, and electrocardiography measures such as the QT interval.

Website:

https://www.framinghamheartstudy.org/

Genetic and Environmental Risk Factors for Alzheimer's Disease Among African Americans (GenerAAtions)

Participants of the GenerAAtions Study were identified through the electronic claims database of the Henry Ford Health System. Community-dwelling African Americans aged 65 and older who had at least one encounter with the Henry Ford Health System in the three years prior to their recruitment and who had an available proxy informant were eligible for this study. Cases met NINCDS-ADRDA criteria for possible or probable AD, determined in a consensus conference which included a behavioral neurologist, psychiatrist, neuropsychologist, and a behavioral neurology nurse practitioner. Phenotypic data were available for 242 AD cases and 204 cognitively normal controls.

Genetic Differences (GenDiff)

Genetic Differences (GenDiff) was an epidemiologic case control study. Cases (235) were newly recognized (e.g., "incident") "Probable AD" (McKhann criteria NINCDS-ADRDA). Subjects were discovered and diagnosed by the Alzheimer's Disease Patient Registry from a community based HMO. Controls were selected at random from the same HMO, without cognitive impairment, and frequency matched for age and sex. The cohort was derived from the same population as the Adult Changes in Thought (ACT)

study. They were also cognitively screened and followed. Blood samples were obtained on most that entered GenDiff with consent for any future genetic analyses.
Website:
http://grantome.com/grant/NIH/R01-AG007584-10S1

Hillblom Aging Network (HAN)

Participants were enrolled in the Hillblom Aging Network at the University of California, San Francisco (UCSF) Memory and Aging Center. All participants underwent comprehensive neurobehavioral evaluations and met the following inclusionary criteria at baseline: 1) clinically normal based on consensus conference with a neurologist and board-certified neuropsychologist; 2) no history of neurological disorder known to impact cognition (e.g., epilepsy, stroke); and 3) functionally intact as defined by an informant-obtained CDR global score of 0 (Morris, 1993). More specifically, the determination of clinically normal by consensus conference involved ruling out the presence of mild cognitive impairment, dementia, or any other neurological condition resulting in cognitive, behavioral, motor, or functional decline (e.g., Parkinson's disease), according to widely used diagnostic criteria (e.g., Albert et al., 2011; Armstrong et al., 2013; Gorno-Tempini et al., 2011; Höglinger et al., 2017; McKeith et al., 2017; McKhann et al., 2011; Postuma et al., 2015; Rascovsky et al., 2011). Three main sources of information were considered by the neurologist and neuropsychologist during the diagnostic conference. First, participants underwent a thorough evaluation with the neurologist that involved a comprehensive neurological examination, clinical interview, and review of systems. Second, neuroimaging (structural MRI) was reviewed to screen out gross brain pathology with potential to negatively impact cognition (e.g., tumor). Third, participants completed a battery of neuropsychological tests to objectively assess major domains of cognitive function, including attention, executive functioning, memory, language, and visuospatial skills. Cognitive impairment was defined by the presence of subjective cognitive decline, as reported by the participant or informant, together with objective performance on neuropsychological testing that was below expectation given the participant's age and level of premorbid functioning (Albert et al., 2011). In making the determination of clinically normal, emphasis was placed on ruling out any declines in the participant's ability to perform everyday tasks due to cognitive changes.
University of California San Francisco Alzheimer's Disease Research Center under grant P30AG062422;
Larry L. Hillblom Network under Grant 2014-A-004-NET;
R01AG032289 (PI: JK);
R01AG048234 (PI: JK)
Espeland MA, Yassine H, Hayden KD, Hugenschmidt C, Bennett WL, Chao A, Neiberg R, Kahn SE, Luchsinger JA; Action for Health in Diabetes (Look AHEAD) Research Group. Sex-related differences in cognitive trajectories in older individuals with type 2 diabetes and overweight or obesity. Alzheimers Dement (N Y). 2021 Apr 9;7(1):e12160. doi: 10.1002/trc2.12160. PMID: 33860069; PMCID: PMC8033410.

Casaletto KB, Elahi FM, Staffaroni AM, Walters S, Contreras WR, Wolf A, Dubal D, Miller B, Yaffe K, Kramer JH. Cognitive aging is not created equally: differentiating unique cognitive phenotypes in “normal” adults. Neurobiol Aging. 2019 May;77:13-19. doi: 10.1016/j.neurobiolaging.2019.01.007. Epub 2019 Jan 24. PMID: 30772736; PMCID: PMC6486874.
Staffaroni AM, Brown JA, Casaletto KB, Elahi FM, Deng J, Neuhaus J, Cobigo Y, Mumford PS, Walters S, Saloner R, Karydas A, Coppola G, Rosen HJ, Miller BL, Seeley WW, Kramer JH. The Longitudinal Trajectory of Default Mode Network Connectivity in Healthy Older Adults Varies As a Function of Age and Is Associated with Changes in Episodic Memory and Processing Speed. J Neurosci. 2018 Mar 14;38(11):2809-2817. doi: 10.1523/JNEUROSCI.3067-17.2018. Epub 2018 Feb 13. PMID: 29440553; PMCID: PMC5852659.
Yokoyama JS, Sturm VE, Bonham LW, Klein E, Arfanakis K, Yu L, Coppola G, Kramer JH, Bennett DA, Miller BL, Dubal DB. Variation in longevity gene KLOTHO is associated with greater cortical volumes. Ann Clin Transl Neurol. 2015 Mar;2(3):215-30. doi: 10.1002/acn3.161. Epub 2015 Jan 26. PMID: 25815349; PMCID: PMC4369272.

Indianapolis-Ibadan (IIAA/IIBD)

The Indianapolis-Ibadan Dementia Project, established in 1991, is a longitudinal prospective population-based comparative epidemiological study of the prevalence and incidence rates and risk factors for Alzheimer’s disease and other age associated dementias. Enrollment of community-dwelling elderly (age>65 years) African Americans living in Indianapolis and Yoruba living in Ibadan, Nigeria employed the same research design, methods, and investigators (see study description at https://iidpportal.medicine.iu.edu/). The first enrollment wave began in 1992 and participants were followed every 2 to 3 years.

Participants were ascertained and evaluated through community centers, clinical and hospital settings as well in community centers and at home. All participants greater than 65 years of age who agreed to participate were screened using measures. Those who failed the screen underwent a more comprehensive clinical evaluation. Medical and family history interview, neuropsychological testing, behavioral and emotional assessments, and functional measures, including collateral informant report, were available for all most participants. Venous blood samples (were collected on all participants. All assessments were conducted in the preferred language of the participant or knowledgeable informant. Finally, all participants were adjudicated by a clinical consensus panel and were classified according to various criteria in place at the time of the clinical data collection Details on diagnosis criteria and process were described in Hendrie et al JAMA 2001.

Website:
https://iidpportal.medicine.iu.edu/

Knight Alzheimer's Disease Research Center (KGAD)

The search for novel risk factors for Alzheimer disease relies on access to accurate and deeply phenotyped datasets. The Memory and Aging Project at the Knight-ADRC (Knight ADRC-MAP) collects plasma, CSF, fibroblast, neuroimaging clinical and cognition data longitudinally and autopsied brain samples. We are using multi-tissue (brain, CSF and plasma) multi-omic data (genetics, epigenomics, transcriptomics, proteomics and metabolomics) to identify novel risk and protective variants, create new prediction models and identify drug targets. The study cohort includes MAP participants from the Knight-ADRC at Washington University in St. Louis (MO). MAP participants have to be at least 65 years old and have no memory problems or mild dementia at the time of enrollment. There is no age at onset criteria for this cohort. Cases had to have a CDR >=0.5 whereas controls had to have a CDR=0 at last assessment. AD definition is based on a combination of both clinical and pathological information if available. Pathologic diagnosis will overrule clinical diagnosis. Participants are Non-Hispanic white from North America (95%) and African American (5%). Autopsy information was provided if available, but it is not a requirement for enrollment.

Website:

https://knightadrc.wustl.edu/

LonGenity

The LonGenity study at Albert Einstein College of Medicine has been recruiting community dwelling Ashkenazi Jewish seniors aged 65 or older in the United States since 2008. Offspring of Parents with Exceptional Longevity (OPEL), defined by having at least one parent who lived to age 95 or older and Offspring of Parents with Usual Survival (OPUS), defined by having neither parent survived to age 95 are being followed annually in this longitudinal study. The goal of this study is to search for longevity genes that may act to slow the aging process and/or protect from age-related diseases. Participants undergo comprehensive cognitive testing, physical performance assessments, and complete medical and family history questionnaires at annual visits. Blood samples are collected biennially and are used for DNA analysis. Participants selected for this sub-study were either (1) age ³70, carriers of APOe4/e4 genotype, and exhibited normal cognitive function or (2) were age ³80, carriers of APOe3/e4 genotype, and exhibited normal cognitive function. Cognitive function was evaluated annually with comprehensive neurocognitive test battery.

Funding:
Grants from the National Institutes of Health R01AG042188, R01AG044829, R01AG046949, R01AG057909, R01AG061155, P30AG038072, the Einstein-Paul Glenn Foundation for Medical Research Center for the Biology of Human Aging.

Longevity Genes Project (LGP)

The Longevity Genes Project (LGP), established in 1998 at Albert Einstein College of Medicine, recruits Ashkenazi Jewish centenarians (age 95 and older) who are in general

good health at age 95, offspring of centenarians, and spouses of offspring in the Eastern United States. The goal of this cross-sectional study is to identify longevity genes that help to slow the aging process and/or protect from age-related diseases. Participants undergo a physical examination (including physical measurements and mini mental state examination (MMSE)), complete a series of questionnaires (including medical and family history, physical activity, etc.), as well as a blood draw or cheek swab collection for DNA analysis. Participants selected for this sub-study were living in the community and were either (1) age ³70, carriers of APOe4/e4 genotype, and exhibited normal cognitive function or (2) were age ³80, carriers of APOe3/e4 genotype, and exhibited normal cognitive function. For individuals age 95 and older, normal cognitive function was defined as full MMSE score >22 or blind MMSE score ³16. Fo individuals age <95, normal cognition was defined as MMSE >25.

Funding:
Grants from the National Institutes of Health R01AG042188, R01AG044829, R01AG046949, R01AG057909, R01AG061155, P30AG038072, the Einstein-Paul Glenn Foundation for Medical Research Center for the Biology of Human Aging.

Mayo Clinic (MAYO)

All 248 cases and 98 controls consisted of Caucasian subjects from the United States ascertained at the Mayo Clinic. All subjects were diagnosed by a neurologist at the Mayo Clinic in Jacksonville, Florida or Rochester, Minnesota. The neurologist confirmed a Clinical Dementia Rating score of 0 for all controls; cases had diagnoses of possible or probable AD made according to NINCDS-ADRDA criteria. Autopsy-confirmed samples came from the brain bank at the Mayo Clinic in Jacksonville, FL and were evaluated by a single neuropathologist. In clinically-identified cases, the diagnosis of definite AD was made according to NINCDS-ADRDA criteria.

Mexican Health and Aging Study (MHAS)

The Mexican Health and Aging Study (MHAS) is a national longitudinal study of adults 50 years and older in Mexico.

The baseline survey, with national and urban/rural representation of adults born in 1951 or earlier, was conducted in 2001 with follow-up interviews in 2003, 2012, 2015, and 2018. A new sample of adults born between 1952-1962 was added in 2012. Similarly, in 2018 a new cohort of adults born between 1963 and 1968 was added to refresh the sample.

The study is a collaborative effort among researchers from the University of Texas Medical Branch (UTMB), the Instituto Nacional de Estadística y Geografía (INEGI, Mexico), the University of Wisconsin, the Instituto Nacional de Geriatría (INGER, Mexico), the Instituto Nacional de Salud Pública (INSP, Mexico), and University of California Los Angeles (UCLA). The MHAS is partly supported by the National Institutes of Health/National Institute on Aging (R01AG018016, R Wong, PI) in the United States and the Instituto Nacional de Estadística y Geografía (INEGI) in Mexico.

Cohort description taken from the MHAS website, July 29, 2022.
Website:
http://www.mhasweb.org/index.aspx

Minority Aging Research Study (MARS)

The Minority Aging Research Study (MARS) is a longitudinal, epidemiological cohort study of decline in cognitive function and risk of Alzheimer's disease in older African-Americans. MARS began in 2004 and over 600 persons have enrolled, 560 of which are currently alive. The study enrolls African-American men and women over age 65 that haven't been diagnosed with dementia. After consent is obtained, a baseline evaluation is scheduled and performed in the participant's home. At the baseline visit, a uniform, structured clinical evaluation is completed consisting of an interview to ascertain a variety of lifestyle and experiential risk factors, a neurological examination, a blood draw, and a comprehensive neuropsychological battery of 23 cognitive tests. The clinical evaluation is repeated on an annual basis. The autopsy program is offered to all subjects as well but not required. Participants also have yearly blood draws which result in the storage of serum, plasma and cells.
Website:
https://www.rushu.rush.edu/research/departmental-research/minority-aging-research-study

Mount Sinai Brain Bank (MSBB)

Human brains were accessed from the Mount Sinai/JJ Peters VA Medical Center Brain Bank (MSBB–Mount Sinai NIH Neurobiobank) cohort, which holds over 2,040 well-characterized brains. This cohort was assembled after applying stringent inclusion/exclusion criteria and represents the full spectrum of cognitive and neuropathological disease severity in the absence of discernable non-AD neuropathology. For each sample, neuropathological assessment was performed according to the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) protocol and included assessment by hematoxylin and eosin, modified Bielschowski, modified thioflavin S, and anti-β amyloid (4G8), anti-tau (AD2) and anti-ubiquitin. A Braak AD-staging score for progression of neurofibrillary neuropathology was assigned to each case. Quantitative data regarding the mean of the density of neuritic plaques in the middle frontal gyrus, orbital frontal cortex, superior temporal gyrus, inferior parietal cortex and calcarine cortex were also collected. Clinical dementia rating scale (CDR) was conducted for assessment of dementia and cognitive status. (Wang, M., Beckmann, N., Roussos, P. et al. The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease. Sci Data 5, 180185 (2018). https://doi.org/10.1038/sdata.2018.185)

Multi-Institutional Research in Alzheimer's Genetic Epidemiology (MIRAGE)

The Multi-Institutional Research in Alzheimer's Genetic Epidemiology (MIRAGE) Study is a family study funded by the NIA that began in 1991. The goal of MIRAGE is to identify genetic and non-genetic risk factors for Alzheimer's disease. Approximately 2,500 members of 1,000 Caucasian and African American families were recruited and blood was collected for DNA and cell lines. These families included both subjects who are cognitively normal and others meeting NINCDS/ADRDA criteria for probable or definite AD. Subjects were ascertained at 17 centers in the US, Canada, Germany and Greece.

Website:

https://clinicaltrials.gov/ct2/show/NCT00239759

National Cell Repository for Alzheimer's Disease Family (NCRAD Family)

The National Cell Repository for Alzheimer's Disease (NCRAD) family cohort was started in 1990 and consists of families with two or more members with early or late onset AD and related dementias. This collection maintains DNA and cell lines on affected family members and unaffected relatives (typically over age 60). These families are not evaluated in person and all clinical information is obtained through medical record review. Therefore, data is limited to the following: family history; demographic data; medical records on the evaluation; diagnosis and treatment of symptomatic subjects; telephone cognitive battery; neuropathological findings when available. This is a longitudinal study with autopsy available to all participants. Genomic DNA, Cell Line DNA, Lymphoblastoid Cell Lines (LCLs), and PBMCs are available.

Website:

https://ncrad.iu.edu/accessing_data.html

National Institute of Mental Health (NIMH)

Case studies generated for members of this cohort include clinical data regarding onset and progression of cognitive problems, medical history and medications, family history of memory problems, neurological history and examination results, psychiatric history and examination results, neuropsychological examination results, and diagnostic imaging and laboratory results. In addition to demographic data, a clinical history, and a neurological history and examination, virtually all of the case summaries included a computed tomographic or magnetic resonance imaging scan, laboratory studies, at least a brief description of behavioral symptoms and/or signs, and a Mini-Mental State Examination and/or more detailed neuropsychological testing.

Diagnosticians followed the NINCDS-ADRDA criteria for the diagnosis of probable and possible AD, with two further specifications. First, a gradual progressive language deficit as the initial symptom would warrant a diagnosis of possible, not probable, AD. This change was instituted based on recent autopsy findings that such patients sometimes do not have AD. Second, subjects with a prominent early behavioral disturbance should

have a diagnosis of possible AD based on evidence that such patients do not invariably have AD.
In phase 1 of the study, at least two clinicians at each of the two non-originating sites rated each patient as having probable AD, possible AD, or non-AD using the modified NINCDS-ADRDA criteria described above. All diagnosticians were "blinded" to the autopsy diagnosis and were not told which of the other two sites had provided the case. Where internal disagreements occurred, the clinicians at each site discussed the cases informally and agreed on a clinical diagnosis, which was tabulated as a pre-consensus diagnosis. In phase 2 of the study, consensus conferences were conducted for each case on which the non-originating sites disagreed, and post-consensus diagnoses (including the cases with agreement) were tabulated.
Website:
https://www.nimhgenetics.org/

National Institute on Aging Late Onset of Alzheimer's Disease Family (NIA-LOAD)

The LOAD collection is a longitudinal, multi-center late onset AD sibling genetics initiative. This NIA-funded study began in 2002 and maintains DNA and cell lines on families with 2 or more siblings with AD (at least one probable or confirmed AD) (n=5291). A third family member who is either affected or unaffected is also required. These individuals are evaluated in person and/or over the phone. A minimum dataset is collected for each person in the family. Definite AD is defined by established neuropathological criteria (and confirmed by autopsy). Probable or possible AD is defined according to NINCDS-ADRDA criteria. Autopsy is offered to all subjects.
Website:
https://www.alz.washington.edu/WEB/researcher_home.html

NIA Alzheimer Disease Research Centers (ADRC)

The NIA ADRC cohort included subjects ascertained and evaluated by the clinical and neuropathology cores of the 32 NIA-funded ADRCs. Data collection is coordinated by the National Alzheimer's Coordinating Center (NACC). NACC coordinates collection of phenotype data from the 32 ADRCs, cleans all data, coordinates implementation of definitions of AD cases and controls, and coordinates collection of samples. The ADRC cohort consists of autopsy-confirmed and clinically-confirmed AD cases, and cognitively normal elders (CNEs) with complete neuropathology data who were older than 60 years at age of death, and living CNEs evaluated using the Uniform dataset (UDS) protocol who were documented to not have mild cognitive impairment (MCI) and were between 60 and 100 years of age at assessment.
Based on the data collected by NACC, the ADGC Neuropathology Core Leaders Subcommittee derived inclusion and exclusion criteria for AD and control samples. All autopsied subjects were age ≥ 60 years at death. AD cases were demented according to DSM-IV criteria or Clinical Dementia Rating (CDR) ≥1. Neuropathologic stratification of

cases followed NIA/Reagan criteria explicitly, or used a similar approach when NIA/Reagan criteria were coded as not done, missing, or unknown. Cases were intermediate or high likelihood by NIA/Reagan criteria with moderate to frequent amyloid plaques and neurofibrillary tangle (NFT) Braak stage of III-VI. Persons with Down syndrome, non-AD tauopathies and synucleinopathies were excluded. All autopsied controls had a clinical evaluation within two years of death.
Controls did not meet DSM-IV criteria for dementia, did not have a diagnosis of mild cognitive impairment (MCI), and had a CDR of 0, if performed. Controls did not meet or were low-likelihood AD by NIA/Reagan criteria, had sparse or no amyloid plaques, and a Braak NFT stage of 0 – II. ADRCs sent frozen tissue from autopsied subjects and DNA samples from some autopsied subjects and from living subjects to the ADRCs to the National Cell Repository for Alzheimer's Disease (NCRAD). DNA was prepared by NCRAD for genotyping and sent to the genotyping site at Children's Hospital of Philadelphia. ADRC samples were genotyped and analyzed in separate batches.

Northern Manhattan Study (NOMAS)

The Northern Manhattan Study (NOMAS) is a research study of stroke and stroke risk factors among the multi-ethnic community of Northern Manhattan, New York.
The study is a collaboration between the Department of Neurology at University of Miami and the Neurological Institute at Columbia University.
NOMAS is a NINDS-funded study of the population of Washington Heights in Northern Manhattan. The ongoing study, which began in 1990, is now a collaboration between the Department of Neurology at University of Miami and the Neurological Institute at Columbia University. The study's interdisciplinary team of doctors and researchers has enrolled over 4,400 people from the community, some of whom have suffered a stroke or related neurological diseases (Sacco et al. 2004).
Website:
https://northernmanhattanstudy.org/

Progressive Supranuclear Palsy (PSP)

Patients with a clinical Parkinsonism in life and neuropathological confirmation of Progressive Supranuclear Palsy (PSP) were identified from brain banks, research hospitals and neuropathologists. The top three contributing sites were the Mayo Clinic, Harvard Brain Tissue Resource Center at McClean Hospital, and the University of Pennsylvania and additional small numbers of cases were obtained from various other institutions across the US. The neuropathological diagnosis was made according to NINDS neuropathologic diagnostic criteria. DNA was extracted from brain tissue from patients who had consented for brain donation. DNA samples and/or tissue were sent to the University of Pennsylvania for preparation for genotyping.

Progressive Supranuclear Palsy at the University of California, Los Angeles (PSP UCLA)

This study includes subjects with Progressive Supranuclear Palsy (PSP). Subjects were enrolled in the davenutide PSP Phase Trail 2/3. Additional subjects were clinically diagnosed PSP at the University of California, San Francisco Memory and Aging Center. Participants met the modified Neuroprotection and Natural History in Parkinson Plus Syndrome study criteria for PSP.

See https://clinicaltrials.gov/ct2/show/NCT01110720 for comprehensive inclusion and exclusion criteria.

Boxer AL, Lang AE, Grossman M, et al. Davunetide in patients with progressive supranuclear palsy: a randomised, double-blind, placebo-controlled phase 2/3 trial. Lancet Neurol. 2014;13(7):676-685. doi:10.1016/S1474-4422(14)70088-2. PMID: 24873720

Puerto Rican 10/66 Study (PR1066)

The Puerto Rican 10/66 Study (PR1066) (https://www.alz.co.uk/1066/), is an Alzheimer's Disease International study of dementia in Puerto Rico that began in 2007 (Dr. Ivonne Jimenez-Velazquez, PI). Individuals were recruited as part of the the 10/66 Population Based Study of Dementia using standard protocols. As part of this study, sociodemographic information and detailed clinical history of memory decline were collected. In addition, the Clinical Dementia Rating scale (CDR), and Community Screening Interview for Dementia, and Petersen ADL criteria were collected for all individuals. Neurocognitive testing (CERAD battery) was available for some participants. Participants were adjudicated for dementia. Lastly, the total number of samples that were whole genome sequenced is 1,565 with the breakdown of 140 Cases; 1,245 Controls and 180 MCI's.

Prince M, Ferri CP, Acosta D, et al. The protocols for the 10/66 dementia research group population- based research programme. *BMC Public Health*. 2007;7(1):165.

Puerto Rican Alzheimer's Disease Initiative (PRADI)

The **Puerto Rican Alzheimer's Disease Initiative (PRADI)** is a NIH NIA study of late-onset Alzheimer disease focused on the Caribbean-Hispanic Puerto Rican population. Participants were ascertained for an AD/ADRD Memory Study, including healthy controls. Eligibility was based on self-reported Puerto Rican heritage. Most participants were recruited in the Island of Puerto Rico with a small fraction being ascertained in South Florida, New York, and Connecticut.

Participants were ascertained and evaluated through community centers, private memory clinics and adult day care centers. Some participants were evaluated in their homes. All participants greater than 60 years of age underwent a standard clinical evaluation consisting of a medical and family history interview, neuropsychological testing, behavioral and emotional assessments, and functional measures. Venous blood

samples (or saliva samples when needed) were collected on all participants. All assessments were conducted in the preferred language of the participant or knowledgeable informant.

Religious Orders Study/Memory and Aging Project (ROSMAP)

The Religious Orders Study (ROS) is a longitudinal, epidemiologic clinical-pathological study of memory, motor, and functional problems in older Catholic nuns, priests, and brothers aged 65 years and older from across the United States. Participants without known dementia agree to medical and psychological evaluation each year and brain donation after death. Since 1994, approximately 1,200 older persons have been enrolled and 580 are currently alive. Participants also have yearly blood draws which result in the storage of serum, plasma and cells.

The Memory and Aging Project (MAP) is a longitudinal, epidemiologic clinical-pathologic study of dementia and other chronic diseases of aging. Older persons are recruited from about 40 continuous care retirement communities and senior subsidized housing facilities around the Chicago metropolitan area. Participants without known dementia agree to annual detailed clinical evaluation and donation of brain, spinal cord and muscle after death. MAP began in 1997 and over 1,600 older adults have enrolled. Approximately 1,000 participants are currently alive. Participants also have yearly blood draws which result in the storage of serum, plasma and cells.

Clinical evaluation, self-report, and medication inspection are used to document medical conditions. The diagnostic process is the same for ROS and MAP. Briefly, a decision tree designed to mimic expert clinical judgment was implemented by computer to inform several clinical diagnoses, including dementia and AD. It combines data reduction techniques for the cognitive performance testing with a series of discrete clinical judgments made in series by a neuropsychologist and a clinician. Presumptive diagnoses of dementia and AD are calculated that conform to accepted clinical criteria. The clinician is asked to agree or disagree with the decisions. An algorithm uses these decisions to provide diagnoses of MCI and amnestic MCI. Persons with MCI are judged to have cognitive impairment by neuropsychologic testing without a diagnosis of dementia by the clinician. Persons without dementia or MCI are categorized as having no cognitive impairment (NCI).

Subjects are also evaluated neurologically every year, and, at the time of death, a review of all ante-mortem data leads to a final clinical diagnosis for each participant: each individual receives a diagnosis of syndromic Alzheimer's disease (AD), of mild cognitive impairment (MCI), or of no cognitive impairment (NCI). After the autopsy is concluded, a spectrum of neuropathologic diagnoses are obtained, such as a pathologic diagnosis of AD as defined using the modified NIA Reagan criteria based on a modified Bielschowsky silver stain to visualize amyloid plaques and neurofibrillary tangles.

Website:
https://www.rushu.rush.edu/research/departmental-research/rush-alzheimers-disease-center/radc-research/epidemologic-research

Research in African-American Alzheimer's Disease Initiative (REAAADI)

The **Research in African-American Alzheimer's Disease Initiative (REAAADI)** is a NIH NIA study focused on identifying genetic factors for Alzheimer disease within the African-American population in order to detect new targets for drug development and improve accessibility to Alzheimer's disease education within the community. Participants were ascertained for multiple studies of AD/ADRD over the past 20 years, including healthy controls. Eligibility across studies was based on self-reported African American heritage. While participants were enrolled initially as part of larger studies of AD/ADRD, the first formal study focusing exclusively on African Americans began in in 2007 (Genetic Epidemiology of Alzheimer's Disease in African Americans; AG028786). Ascertainment has continued since 2007 as part of multiple studies of AD/ADRD including the REAAADI Study (AG052410). Since 2007, participants have been ascertained via academic centers in North Carolina (Duke University, NC &T, Wake Forest University), Florida (University of Miami), New York (Columbia University), Ohio (Case Western University) and Tennessee (Vanderbilt University).
Participants were ascertained and evaluated through community centers, private memory clinics and adult day care centers. Some participants were evaluated in their homes. All participants greater than 60 years of age underwent a standard clinical evaluation consisting of a medical and family history interview, neuropsychological testing, behavioral and emotional assessments, and functional measures. Venous blood samples (or saliva samples when needed) were collected on all participants. All assessments were conducted in the preferred language of the participant or knowledgeable informant. Note that the assessment protocols have changed over the years, but a core group of clinical measures (including neuropsychological tests) are available for all participants.

The Rotterdam Elderly Study is a prospective cohort study in the Ommoord district in the city of Rotterdam, the Netherlands [Hofman et al., 1991]. Following the pilot in 1989, recruitment started in January 1990. The main objectives of the Rotterdam Study were to investigate the risk factors of cardiovascular, neurological, ophthalmological and endocrine diseases in the elderly. Up to 2008, approximately 15,000 subjects aged 45 years or over have been recruited. Participants were interviewed at home and went through an extensive set of examinations, bone mineral densitometry, including sample collections for in-depth molecular and genetic analyses. Examinations were repeated every 3-4 years in potentially changing characteristics. Participants were followed for the most common diseases in the elderly, including coronary heart disease, heart failure and stroke, Parkinson's disease, Alzheimer's disease and other dementias, depression and anxiety disorders, macular degeneration and glaucoma, diabetes mellitus and osteoporosis.
In the baseline and follow-up examinations participants undergo an initial screen for dementia with the Mini Mental State Examination (MMSE) and the Geriatric Mental Schedule (GMS), followed by an examination and informant interview with the Cambridge Examination for Mental Disorders of the Elderly (CAMDEX) in screen positives (MMSE <26 or GMS >0), and subsequent neurological, neuropsychological and

neuroimaging examinations. Of subjects who cannot be reexamined in person, information is obtained from the GPs and the regional institute for outpatient mental health care. A consensus panel makes the final diagnoses in accordance with standard criteria (DSM-III-R criteria; NINCDS-ADRDA; NINDS-AIREN).
Website:
http://www.erasmus-epidemiology.nl/research/ergo.htm

Stanford Extreme Phenotypes in AD (StEP AD)

This cohort contains samples from Brains for Dementia Research (BDR) Genetics project (Dr. Kevin Morgan, PI and Directory of the Alzheimer's Research UK (ARUK) DNA Consortium and BDR).
The BDR cohort and program was for planned brain donation across five UK brain banks and one donation point, with standardized operating procedures, following longitudinal clinical and psychometric assessments for people with no cognitive impairment as well as those with dementia. See the Francis et al. publication for a detailed description about the BDR cohort set-up, clinical data, and psychometric assessment measures collected.
Francis PT, Costello H, Hayes GM. Brains for Dementia Research: Evolution in a Longitudinal Brain Donation Cohort to Maximize Current and Future Value. *J Alzheimers Dis*. 2018;66(4):1635-1644. doi:10.3233/JAD-180699

Texas Alzheimer's Research and Care Consortium (TARCC)

Data from the Texas Alzheimer's Research and Care Consortium (TARCC) includes cases enrolled at several major medical research institutions (as of 2013 this included Baylor College of Medicine, Texas Tech University Health Sciences Center, University of North Texas Health Science Center, The University of Texas Health Sciences Center at San Antonio, The University of Texas Southwestern Medical Center, and Texas A & M Health Science Center).
Individuals must be at least 55 years of age with a diagnosis of probable AD or normal cognition based on a Clinical Dementia Rating Global Score of 0. Clinical, neurological, and neuropsychological examinations performed at each site follow the TARCC research protocol that has been adopted from the standard clinical work-up for dementia. All subjects are examined at baseline and at each annual follow-up visit.
Information is obtained from the clinical and neurological examination on age at onset of symptoms (if AD patient), family history of dementia in first degree relatives, cardiovascular disease and cardiovascular disease risk factors. Subjects also undergo a battery of neuropsychological tests as part of the TARC research protocol, with all information reviewed by a consensus panel made up of at least a physician, neuropsychologist, and research coordinator at each site to assign the final clinical diagnosis according to NINCDS-ADRDA criteria.
Website:
http://www.txalzresearch.org/

University of Miami (MIA)

Each affected individual met NINCDS-ADRDA criteria for probable or definite AD with age at onset greater than 60 years, as determined from specific probe questions within the clinical history provided by a reliable family informant or from documentation of significant cognitive impairment in the medical record. Cognitively healthy controls were unrelated individuals from the same catchment areas and frequency matched by age and gender, and had a documented MMSE or 3MS score in the normal range.
Samples sequenced in ADSP-FUS1 include participants from the John P. Hussman Institute for Human Genomics (HIHG) Brain Bank. The HIHG Brain, Bank autopsy individuals followed the typical clinical course of disease. This sample is a clinic and community outreach sample ascertained in North and South Carolina and Virginia.
Website:
http://hihg.med.miami.edu/alzheimers

University of Miami Brain Bank (MBB)

The University of Miami Brain Bank Cohort was ascertained through self-referred cases and controls to the University of Miami Brain Endowment Bank, a National Institutes of Health (NIH) NeuroBioBank, one of six designated brain and tissue biorepositories in the nation. Medical records are available on all cases and controls, all of whom were tested before death for cognitive function.
Subjects included in the University of Miami Brain Endowment Bank cohort include individuals who were donors to the University of Miami Brain Endowment Bank from 1986 to 2020. Donors were either individuals with history of cognitive impairment and neuropathological changes consistent with Alzheimer disease, or cognitively unaffected individuals without such neuropathological changes. Participants were recruited from community organizations, organ/tissue donation program registries, or self-referral. This is an autopsy-based study designed to analyze post-mortem brain tissue. Medical records are also obtained as available to help correlate brain behavior relationships.
Website:
https://med.miami.edu/programs/brain-endowment-bank

University of Pittsburg (PITT)

Study participants were enrolled at the University of Pittsburgh Alzheimer's Disease Research Center (ADRC), all of whom met the National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association clinical criteria for probable AD. Each participant had undergone an extensive neuropsychiatric evaluation, which has been described in detail previously. Briefly, this involved a physical examination, neurological examination, semistructured psychiatric interview, and neuropsychological assessment. All patient records were reviewed at a multidisciplinary clinical consensus conference for assignment of a diagnosis. The inclusion/exclusion criteria have been published previously, but it is important to note that one of the criteria was that all the patients

had to have a reliable caregiver that could provide detailed information about the patients' clinical symptoms and their activities of daily living (ADLs). The caregiver was interviewed in person once a year and by phone every 6 months. In case of loss of a caregiver, the data were censored as the last date that reliable information was received from the caregiver.
Lopez OL, Becker JT, Saxton J, Sweet RA, Klunk W, DeKosky ST. Alteration of a clinically meaningful outcome in the natural history of Alzheimer's disease by cholinesterase inhibition. J Am Geriatr Soc. 2005 Jan;53(1):83-7. doi: 10.1111/j.1532-5415.2005.53015.x. PMID: 15667381.

University of Toronto

This study was carried out under the direction of Dr. Peter St George-Hyslop and Dr Rogaeva at the Tanz Centre for Research in Neurodegenerative Diseases (CRND), University of Toronto. In order to explore the potential role of hereditary factors, a registry at the Tanz CRND with a coordinator was established for families in which two or more individuals have the suspected diagnosis of Alzheimer's Disease or other forms of dementia. Eligibility requirements were disseminated through the CRND website. Individuals who contacted the Registry received a family history questionnaire in the mail. For sufficiently informative families, researchers sought consent to obtain blood samples from available family members required for genetic research. Collaborating neurologists and genetic counselors also referred families that fit the requirements. All collected individuals were given numeric IDs to protect their privacy. Families were eligible if: two or more members, living or decease, were affected by Alzheimer's Disease, Parkinson's Disease, amyotrophic lateral sclerosis, frontal temporal dementia or Creutzfeldt-Jacob Disease. Families were ineligible if: there were fewer than 2 affected members or family history was unavailable to document.
Website:
http://www.tanz.med.utoronto.ca/familial-alzheimer%E2%80%99s-disease-registry

University of Washington Families (RAS)

131 families with LOAD (751 individuals) were ascertained and evaluated through the University of Washington Alzheimer Disease Research Center. Clinical and neuropathological assessments of cases and controls, including blood sampling, medical record reviews, brain autopsies, and genetic analyses were performed under protocols approved by the institutional review boards of the University of Washington and the Seattle Veterans Affairs Puget Sound Health Care System.

Vanderbilt University (VAN)

The UM/VU dataset contains 1,186 cases and 1,135 CNEs (new and previously published) ascertained at the University of Miami and Vanderbilt University, including 409 autopsy-confirmed cases and 136 controls. An additional 16 cases were included

and 34 controls excluded from the data analyzed in the prior study. Each affected individual met NINCDS-ADRDA criteria for probably or definite AD with age at onset greater than 60 years as determined from specific probe questions within the clinical history provided by a reliable family informant or from documentation of significant cognitive impairment in the medical record. Cognitively healthy controls were unrelated individuals from the same catchment areas and frequency matched by age and gender, and had a documented MMSE or 3MS score in the normal range. Cases and controls had similar demographics: both had ages-at-onset/ages-at-exam of 74 (± 8 standard deviations), and cases were 63% female, and controls were 61% female.

Washington Heights and Inwood Community Aging project (WHICAP)

Since inception of the study in 1992, over 6,000 participants have enrolled in the Washington Heights and Inwood Community Aging project (WHICAP). The cohort participants were nondemented initially, 65 years of age or older, and comprised of non-Hispanic whites (32%), African Americans (28%), and Caribbean Hispanics from the Dominican Republic (44%). During each assessment, participants received a neuropsychological test battery, medical interview, and were re-consented for sharing of genetic information and autopsy. A consensus diagnosis was derived for each participant by experienced clinicians based on NINCDS-ADRDA criteria for possible, probable, or definite AD, or moderate or high likelihood of neuropathological criteria of AD. Every individual with whole-exome sequencing has at least a baseline and one follow-up assessment and examination, and for those who have died, the presence or absence of dementia was determined using a brief, validated telephone interview with participant informants: The Dementia Questionnaire (DQ) and the Telephone Interview of Cognitive Status (TICS).

Over the length of the project, WHICAP have identified environmental, health-related and genetic risk factors of disease and predictors of disease progression by collecting longitudinal data on cognitive performance, emotional health, independence in daily activities, blood pressure, anthropometric measures, cardiovascular status and selected biomarkers in this elderly, multi-ethnic cohort. Biomarker studies include lipids, amyloid peptides, sex hormones, homocysteine, insulin and C-reactive protein (CRP), and MRI in these elderly participants. WHICAP have reported that the rates of disease and the frequency of disease risk factors vary across ethnic groups, and have identified one of the largest, multi-ethnic groups of incident LOAD cases.

Website:
https://www.maelstrom-research.org/mica/individual-study/whicap

Wisconsin Registry for Alzheimer's Prevention (WRAP)

The Wisconsin Registry for Alzheimer's Prevention (WRAP) is an ongoing longitudinal observational cohort study of individuals age 40-65 at baseline who do not have dementia. Since 2001, WRAP has enrolled more than 1,700 individuals, 73% of whom had a parental history of probable Alzheimer's disease (AD) dementia. Participants return

for a second visit approximately 4 years after baseline, and subsequent visits occur every 2 years. At each visit, a cognitive test battery is administered, self-reported medical and lifestyle histories (e.g., diet, physical and cognitive activity, sleep, and mood) are assessed via questionnaire, and blood is drawn for laboratory tests, metabolomics, and genomics. A subset of participants have also undergone molecular imaging, structural imaging, and cerebrospinal fluid collection for biomarker measurement.
GeneRations Of WRAP (GROW) is an ancillary study to the Wisconsin Registry for Alzheimer's Prevention (WRAP), a longitudinal cohort study enriched for individuals with a parental history of probable Alzheimer's disease (AD) dementia. GROW is a family study that recruited extended family members of existing WRAP participants. Genomic data from family members with probable AD dementia were collected via banked brains or saliva. Individuals free of dementia and within the WRAP eligibility age of 40-65 were enrolled in WRAP. At each study visit, a cognitive test battery is administered, self-reported medical and lifestyle histories (e.g., diet, physical and cognitive activity, sleep, and mood) are assessed via questionnaire, and blood is drawn for laboratory tests, metabolomics, and genomics. A subset of participants have also undergone molecular imaging, structural imaging, and cerebrospinal fluid collection for biomarker measurement.
Website:
https://wrap.wisc.edu/

eAppendix 2. Supplemental Methods

Quality control procedures

In each cohort-platform, variants were excluded based on genotyping rate (< 95%), MAF < 1%, and Hardy-Weinberg equilibrium in controls ($P < 10^{-6}$) using PLINK v1.9[1]. gnomAD[2] database-derived information was used to filter out SNPs that met one of the following exclusion criteria[3,4]: (i) located in a low complexity region, (ii) located within common structural variants (MAF > 1%), (iii) multiallelic SNPs with MAF > 1% for at least two alternate alleles, (iv) located within a common insertion/deletion, (v) having any flag different than PASS in gnomADv.3, and (vi) having potential probe polymorphisms. The latter are defined as SNPs for which the probe may have variable affinity due to the presence of other SNP(s) within 20 bp and with MAF > 1%. Individuals with more than 5% genotype missingness were excluded. Duplicate individuals were identified with KING[5] and their clinical, diagnostic and pathological data (including age-at-onset of cognitive symptoms, age-at-examination for clinical diagnosis, age-at-last exam, age-at-death), as well as sex, race, and *APOE* genotype were cross-referenced across cohorts. Duplicate entries with irreconcilable phenotype or discordant sex were flagged for exclusion. For individuals with duplicated genotype in sequencing and imputed data, the sequencing entry was used in the discovery set and the imputed entry was not included in the replication set. As some cohorts contributed to both the sequencing and genotyping platforms, some individuals in the discovery were related to individuals in the replication. Mega-analyses using linear mixed models that account for relatedness were run as sensitivity analyses (see Statistical analysis section).

Quality control for the MVP genotype data has been previously described elsewhere[6].

APOE genotype ascertainment

We directed specific attention to the genotyping of the SNPs determining the main *APOE* genotype (rs429358 and rs7412), rs769455-T (*APOE*[R145C]), and rs376170967-A

(*APOE*[R150H]). Note that Arg145Cys (R145C) is also sometimes referred to as Arg163Cys when the first 18 codons of *APOE* encoding a signal peptide are included; and respectively Arg150Cys is also referred to as Arg168Cys. The rs429358 and rs7412 alternate (minor) alleles were respectively used to determine the number of *APOE ε4* and *APOE ε2* alleles. The *ε3* allele was determined by a reference (major) allele call at both rs429359 and rs7412. Determining the *ε2/ε4* genotype would in theory require phase information (i.e., knowing whether alleles are observed on the same chromosome copy) to distinguish it from *ε1/ε3,* but since the *ε1* allele is extremely rare[7], we followed common practice to assign the *ε2/ε4* genotype. We considered the two triplets of variants separately (rs429358, rs7412, rs769455-T) and (rs429358, rs7412, rs376170967-A). Hereafter, we described the rs769455-T triplet genotyping and the same protocol was applied for the rs376170967-A triplet genotyping.

In the stage I discovery dataset, composed of WGS and WES data, the three variants were directly called in the sequencing data distributed by NIAGADS (**eTable 7).** In ADSP WGS, the three variants showed low genotype missingness across subjects (2-3%). We therefore decided to simply exclude subjects in ADSP WGS with missing genotypes for any of the three variants. In ADSP WES, there was a high genotype missingness at rs7412 (32.5%). This resulted from a low read depth and genotype quality in some of the different WES capture kits that were used in the ADSP WES project[8]. We therefore sought to re-call all three variants in order to fill out missing information where possible. We inferred the variants' genotype if one of the following two conditions was met: (i) read depth (DP) and genotype quality (GQ) were, respectively, greater than or equal to 6 and 20, observing at least 20% alternate allele reads to call a heterozygote (e.g. *ε2/ε3*); or (ii) considering only DP greater than or equal to 4 with, either, all calls corresponding to one allele to call a homozygote (e.g. *ε4/ε4*), or, each allele called on at least 40% of the reads, to call a heterozygote. Importantly, as a quality check, using these thresholds, we did not observe any discordance in the inferred *APOE* genotype across 3499 duplicates between WGS and WES. After this first round of *APOE* genotype ascertainment, some individuals still had either the rs7412 or rs429358 genotype missing (i.e., only one of the

two variants could be called using the above criteria), making it impossible to infer their *APOE* genotype from WES data alone. 110 individuals (3.4%) included in the discovery analysis were in this situation (**eFigure 2**). In order to determine the main *APOE* genotype in these 110 individuals, we used the following approach. Many of these individuals had a reported *APOE* genotype in their demographics that could be used to complete the missing information in a second additional round of *APOE* genotype ascertainment. This approach was preferred over relying solely on the *APOE* genotype in the demographics, since the genotype calls on the WES data are expected to provide higher accuracy compared to other commonly used *APOE* genotyping methods[9]. Here, it is worth noting that the standard read length of the WES data (100 to 150 bp) entails that rs769455 will often be on the same read as rs7412 (39 bp apart) and in rarer instances on the same read as rs429358 (99 bp apart). Thus, for this second ascertainment round, to avoid observing only one of the chromosome 19 pair in the sequencing reads at rs769455, we required rs769455 and other available APOE variants to meet criteria (i) from above. To illustrate, consider the example where an individual in the sequencing data was homozygous for the reference allele at rs429358 but had a missing genotype at rs7412. In this case, from the WES information, we know that this individual is not carrying an *ε4* allele, but we cannot determine the presence or absence of an *ε2* allele. We then turned to the information from the *APOE* genotype reported in the demographics to infer the most likely *APOE* genotype. In our example, if the individual has a reported *APOE* genotype *ε3/ε3, ε2/ε3,* or *ε2/ε2* then the information in the WES data is deemed concordant with the reported *APOE* and we used the reported *APOE*. However, if the reported *APOE* genotype was *ε4/ε4, ε3/ε4* then we would correct this genotype to *ε3/ε3*, based on the WES information that clearly indicated there were no *ε4* reads. This can be generalized as: (a) for individuals with high DP at rs429358 and for whom resolving the *APOE* genotype simply required changing *ε3* to *ε4* or vice-versa, then the information from the WES at rs429358 was used, (b) similarly for individuals with high DP at rs7412 and for whom resolving *APOE* genotype simply required changing *ε3* to *ε2* or vice-versa the information from the WES at rs7412 was used. Finally,

subjects who had no call at rs7412 and rs429358 in the sequencing data were excluded. **eFigure 2** provides a graphical representation of this workflow and provides the exact number of individuals in each category. Additionally, sensitivity analyses, in which we restricted the analysis to ADSP WES individuals who had direct calls at all 3 *APOE* missense variants, emphasize that the significance of our results and effect sizes remain unchanged (**eTable 13**).

In the stage II replication dataset, which did not contain individuals with next generation sequencing data, the main *APOE* genotype was prioritized in the following order: (i) directly obtained from the microarray genotyping, (ii) provided with the primary study demographics (genotyping methods described elsewhere[10,11]), (iii) imputed with high confidence ($r^2 > 0.8$). rs769455_T was directly genotyped on the Exome Chip microarrays and was imputed in other datasets (cf Imputation section in the main text).

In stage III replication (Million Veteran Project), *APOE* genotypes were calculated from the "best guess" imputed genotypes for rs429358 and rs7412, both of which were well imputed ($r^2>.8$).

APOE haplotype local ancestry estimation

To estimate the local ancestry of the *APOE* haplotype we considered a region encompassing the *APOE* gene with a 200kb-flank upstream and downstream (coordinates in build hg38 chr19:44705791-45109393). We phased separately the whole sample of ADSP WES and ADSP WGS using *Eagle* v2.4.1[12], without using an external reference panel. Publicly available sequencing reference panels are much smaller than these two datasets and the *Eagle*'s documentation suggests that using a reference panel in this scenario is unlikely to significantly increase phasing accuracy. To estimate local ancestry we used *RFMix* v.2[13] with the 893 AFR individuals and 633 EUR individuals from the expanded high-coverage (30x) whole-genome sequencing from the 1000 Genome Project data[14]. In sensitivity analyses, we re-analyzed the discovery sample data solely including individuals with AFR local ancestry at both *APOE* haplotypes.

Quantifying the diversity of African subpopulations

To quantify the diversity in terms of African subpopulations in our ADSP discovery and ADGC replication, we merged individuals with at least 75% African global ancestry with the African participants from the 1000 Genomes Project[14] and Human Genome Diversity Project[15] callsets available through gnomAD(v3.1.2)[2]. EIGENSTRAT (v7.2.1)[16] *smartpca* method was used to compute ten principal components accounting for different genetic ancestries across these African populations. The ADSP discovery and ADGC replication were merged separately with the reference panels and two sets of principal components were obtained (**eFigures 3-4**). After principal components outliers removal performed by EIGENSTRAT on each merge, the 1000 Genomes Project was composed of the following African groups: Mandinka in Western Division – Gambia (178), Yoruba in Ibadan – Nigera (174), Esan in Nigeria (149), African Carribean in Barbados (116), Luhya in Webuye – Kenya (99), Mende in Sierra Leone (99), and African ancestry in Southwestern United-States (74). The Human Genome Diversity Project was composed of the following African groups: Mandinka in Senegal (20), Yoruba in Nigeria (18), Biaka in Central African Republic (15), Bantu in Kenya (10), Bantu in South Africa (4), Mbuti in Democratic Republic of Congo (2), and San in Namibia (1).

ApoE3 and apoE3-R145C production and purification

Human apoE*3* plasmid was a gift from M. J. LaDu at University of Illinois, Chicago. Human apoE*3*-R145C plasmid was generated by site-directed mutagenesis.DNA sequences were confirmed by Sanger sequencing.

ApoE3 and apoE3-R145C were recombinantly produced in HEK293 cells via transient transfection. Briefly, HEK293 cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) containing glucose (4.5 g/L), L-glutamine, 1% penicillin/ streptomycin, 1% amphotericin-B and supplemented with 10% FBS. *APOE* plasmids were transfected into HEK293 cells using PEI Max 40,000 MW (Polysciences, Inc).

Medium was changed to serum-free medium 24 h after transfection. Conditioned medium was collected after 3 days, and apoE protein was purified using heparin affinity columns with FPLC (Cytiva Akta Pure). Purified apoE was then buffer exchanged into PBS for downstream experiments. Protein concentrations were determined by nanodrop.

Heparin binding affinity

The heparin binding affinity of apoE3-R145C was compared to that of apoE3 using 1 ml HiTrap heparin columns (Cytiva) equilibrated in 20 mM Tris-HCl, pH 7.4. Recombinantly produced apoE3 and apoE3-R145C were diluted to 50 μg/mL in 20 mM phosphate buffer, pH 7.4 and 1 mL was injected onto the column, followed by a wash with 10 column volumes of 20 mM Tris-HCl, pH 7.4. Bound proteins were then eluted with a linear NaCl gradient (0-0.8 M). ApoE in each of the 0.5 mL fractions was detected by enzyme-linked immunosorbent assay (ELISA; capture antibody: AB947 from EMD Millipore; detection antibody: K74180B from Meridian Life Science). Heparin affinity chromatography.

Data availability

Data used in preparation of this manuscript can be obtained upon application at:

- dbGaP (https://www.ncbi.nlm.nih.gov/gap/advanced_search/)
- NIAGADS and NIAGADS DSS (https://www.niagads.org/)
- LONI (https://ida.loni.usc.edu/)
- Synapse (https://adknowledgeportal.synapse.org/)
- RADC Rush (https://www.radc.rush.edu/)
- NACC (https://naccdata.org/)

eTables 3 and 4 provide the details of repositories and accession number per cohort-platform group.

eAppendix 3. Additional Acknowledgments

Data for this study were prepared, archived, and distributed by the National Institute on Aging Alzheimer's Disease Data Storage Site (NIAGADS) at the University of Pennsylvania (U24-AG041689), funded by the National Institute on Aging.

Acknowledgments for the use of ADSP WES and WGS data

The Alzheimer's Disease Sequencing Project (ADSP) is comprised of two Alzheimer's Disease (AD) genetics consortia and three National Human Genome Research Institute (NHGRI) funded Large Scale Sequencing and Analysis Centers (LSAC). The two AD genetics consortia are the Alzheimer's Disease Genetics Consortium (ADGC) funded by NIA (U01 AG032984), and the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) funded by NIA (R01 AG033193), the National Heart, Lung, and Blood Institute (NHLBI), other National Institute of Health (NIH) institutes and other foreign governmental and non-governmental organizations. The Discovery Phase analysis of sequence data is supported through UF1AG047133 (to Drs. Schellenberg, Farrer, Pericak-Vance, Mayeux, and Haines); U01AG049505 to Dr. Seshadri; U01AG049506 to Dr. Boerwinkle; U01AG049507 to Dr. Wijsman; and U01AG049508 to Dr. Goate and the Discovery Extension Phase analysis is supported through U01AG052411 to Dr. Goate, U01AG052410 to Dr. Pericak-Vance and U01 AG052409 to Drs. Seshadri and Fornage.

Sequencing for the Follow Up Study (FUS) is supported through U01AG057659 (to Drs. PericakVance, Mayeux, and Vardarajan) and U01AG062943 (to Drs. Pericak-Vance and Mayeux). Data generation and harmonization in the Follow-up Phase is supported by U54AG052427 (to Drs. Schellenberg and Wang). The FUS Phase analysis of sequence data is supported through U01AG058589 (to Drs. Destefano, Boerwinkle, De Jager, Fornage, Seshadri, and Wijsman), U01AG058654 (to Drs. Haines, Bush, Farrer, Martin, and Pericak-Vance), U01AG058635 (to Dr. Goate), RF1AG058066 (to Drs. Haines, Pericak-Vance, and Scott), RF1AG057519 (to Drs. Farrer and Jun), R01AG048927 (to Dr. Farrer), and RF1AG054074 (to Drs. Pericak-Vance and Beecham).

The ADGC cohorts include: Adult Changes in Thought (ACT) (UO1 AG006781, UO1 HG004610, UO1 HG006375, U01 HG008657), the Alzheimer's Disease Centers (ADC) (P30 AG019610, P30 AG013846, P50 AG008702, P50 AG025688, P50 AG047266, P30 AG010133, P50 AG005146, P50 AG005134, P50 AG016574, P50 AG005138, P30 AG008051, P30 AG013854, P30 AG008017, P30 AG010161, P50 AG047366, P30 AG010129, P50 AG016573, P50 AG016570, P50 AG005131, P50 AG023501, P30 AG035982, P30 AG028383, P30 AG010124, P50 AG005133, P50 AG005142, P30 AG012300, P50 AG005136, P50 AG033514, P50 AG005681, and P50 AG047270), the Chicago Health and Aging Project (CHAP) (R01 AG11101, RC4 AG039085, K23 AG030944), Indianapolis Ibadan (R01 AG009956, P30 AG010133), the Memory and Aging Project (MAP) (R01 AG17917), Mayo

Clinic (MAYO) (R01 AG032990, U01 AG046139, R01 NS080820, RF1 AG051504, P50 AG016574), Mayo Parkinson's Disease controls (NS039764, NS071674, 5RC2HG005605), University of Miami (R01 AG027944, R01 AG028786, R01 AG019085, IIRG09133827, A2011048), the Multi-Institutional Research in Alzheimer's Genetic Epidemiology Study (MIRAGE) (R01 AG09029, R01 AG025259), the National Cell Repository for Alzheimer's Disease (NCRAD) (U24 AG21886), the National Institute on Aging Late Onset Alzheimer's Disease Family Study (NIA- LOAD) (R01 AG041797), the Religious Orders Study (ROS) (P30 AG10161, R01 AG15819), the Texas Alzheimer's Research and Care Consortium (TARCC) (funded by the Darrell K Royal Texas Alzheimer's Initiative), Vanderbilt University/Case Western Reserve University (VAN/CWRU) (R01 AG019757, R01 AG021547, R01 AG027944, R01 AG028786, P01 NS026630, and Alzheimer's Association), the Washington Heights-Inwood Columbia Aging Project (WHICAP) (RF1 AG054023), the University of Washington Families (VA Research Merit Grant, NIA: P50AG005136, R01AG041797, NINDS: R01NS069719), the Columbia University HispanicEstudio Familiar de Influencia Genetica de Alzheimer (EFIGA) (RF1 AG015473), the University of Toronto (UT) (funded by Wellcome Trust, Medical Research Council, Canadian Institutes of Health Research), and Genetic Differences (GD) (R01 AG007584). The CHARGE cohorts are supported in part by National Heart, Lung, and Blood Institute (NHLBI) infrastructure grant HL105756 (Psaty), RC2HL102419 (Boerwinkle) and the neurology working group is supported by the National Institute on Aging (NIA) R01 grant AG033193.

The CHARGE cohorts participating in the ADSP include the following: Austrian Stroke Prevention Study (ASPS), ASPS-Family study, and the Prospective Dementia Registry-Austria (ASPS/PRODEM-Aus), the Atherosclerosis Risk in Communities (ARIC) Study, the Cardiovascular Health Study (CHS), the Erasmus Rucphen Family Study (ERF), the Framingham Heart Study (FHS), and the Rotterdam Study (RS). ASPS is funded by the Austrian Science Fond (FWF) grant number P20545-P05 and P13180 and the Medical University of Graz. The ASPS-Fam is funded by the Austrian Science Fund (FWF) project I904),the EU Joint Programme - Neurodegenerative Disease Research (JPND) in frame of the BRIDGET project (Austria, Ministry of Science) and the Medical University of Graz and the Steiermärkische Krankenanstalten Gesellschaft. PRODEM-Austria is supported by the Austrian Research Promotion agency (FFG) (Project No. 827462) and by the Austrian National Bank (Anniversary Fund, project 15435. ARIC research is carried out as a collaborative study supported by NHLBI contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C). Neurocognitive data in ARIC is collected by U01 2U01HL096812, 2U01HL096814, 2U01HL096899, 2U01HL096902, 2U01HL096917 from the NIH (NHLBI, NINDS, NIA and NIDCD), and with previous brain MRI examinations funded by R01-HL70825 from the NHLBI. CHS research was supported by contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants U01HL080295 and U01HL130114 from the NHLBI with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional

support was provided by R01AG023629, R01AG15928, and R01AG20098 from the NIA. FHS research is supported by NHLBI contracts N01-HC-25195 and HHSN268201500001I. This study was also supported by additional grants from the NIA (R01s AG054076, AG049607 and AG033040 and NINDS (R01 NS017950). The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4- 2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002- 01254). High-throughput analysis of the ERF data was supported by a joint grant from the Netherlands Organization for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the municipality of Rotterdam. Genetic data sets are also supported by the Netherlands Organization of Scientific Research NWO Investments (175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), and the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project 050-060-810. All studies are grateful to their participants, faculty and staff. The content of these manuscripts is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the U.S. Department of Health and Human Services.

The FUS cohorts include: the Alzheimer’s Disease Centers (ADC) (P30 AG019610, P30 AG013846, P50 AG008702, P50 AG025688, P50 AG047266, P30 AG010133, P50 AG005146, P50 AG005134, P50 AG016574, P50 AG005138, P30 AG008051, P30 AG013854, P30 AG008017, P30 AG010161, P50 AG047366, P30 AG010129, P50 AG016573, P50 AG016570, P50 AG005131, P50 AG023501, P30 AG035982, P30 AG028383, P30 AG010124, P50 AG005133, P50 AG005142, P30 AG012300, P50 AG005136, P50 AG033514, P50 AG005681, and P50 AG047270), Alzheimer's Disease Neuroimaging Initiative (ADNI) (U19AG024904), Amish Protective Variant Study (RF1AG058066), Cache County Study (R01AG11380, R01AG031272, R01AG21136, RF1AG054052), Case Western Reserve University Brain Bank (CWRUBB) (P50AG008012), Case Western Reserve University Rapid Decline (CWRURD) (RF1AG058267, NU38CK000480), CubanAmerican Alzheimer's Disease Initiative (CuAADI) (3U01AG052410), Estudio Familiar de Influencia Genetica en Alzheimer (EFIGA) (5R37AG015473, RF1AG015473, R56AG051876), Genetic and Environmental Risk Factors for Alzheimer Disease Among African Americans Study (GenerAAtions) (2R01AG09029, R01AG025259, 2R01AG048927), Gwangju Alzheimer and Related Dementias Study (GARD) (U01AG062602), Hussman Institute for Human Genomics Brain Bank (HIHGBB) (R01AG027944, Alzheimer's

Association "Identification of Rare Variants in Alzheimer Disease"), Ibadan Study of Aging (IBADAN) (5R01AG009956), Mexican Health and Aging Study (MHAS) (R01AG018016), Multi-Institutional Research in Alzheimer's Genetic Epidemiology (MIRAGE) (2R01AG09029, R01AG025259, 2R01AG048927), Northern Manhattan Study (NOMAS) (R01NS29993), Peru Alzheimer's Disease Initiative (PeADI) (RF1AG054074), Puerto Rican 1066 (PR1066) (Wellcome Trust (GR066133/GR080002), European Research Council (340755)), Puerto Rican Alzheimer Disease Initiative (PRADI) (RF1AG054074), Reasons for Geographic and Racial Differences in Stroke (REGARDS) (U01NS041588), Research in African American Alzheimer Disease Initiative (REAAADI) (U01AG052410), Rush Alzheimer's Disease Center (ROSMAP) (P30AG10161, R01AG15819, R01AG17919), University of Miami Brain Endowment Bank (MBB), and University of Miami/Case Western/North Carolina A&T African American (UM/CASE/NCAT) (U01AG052410, R01AG028786).

The four LSACs are: the Human Genome Sequencing Center at the Baylor College of Medicine (U54 HG003273), the Broad Institute Genome Center (U54HG003067), The American Genome Center at the Uniformed Services University of the Health Sciences (U01AG057659), and the Washington University Genome Institute (U54HG003079).

Biological samples and associated phenotypic data used in primary data analyses were stored at Study Investigators institutions, and at the National Cell Repository for Alzheimer’s Disease (NCRAD, U24AG021886) at Indiana University funded by NIA. Associated Phenotypic Data used in primary and secondary data analyses were provided by Study Investigators, the NIA funded Alzheimer’s Disease Centers (ADCs), and the National Alzheimer’s Coordinating Center (NACC, U01AG016976) and the National Institute on Aging Genetics of Alzheimer’s Disease Data Storage Site (NIAGADS, U24AG041689) at the University of Pennsylvania, funded by NIA This research was supported in part by the Intramural Research Program of the National Institutes of health, National Library of Medicine. Contributors to the Genetic Analysis Data included Study Investigators on projects that were individually funded by NIA, and other NIH institutes, and by private U.S. organizations, or foreign governmental or nongovernmental organizations.

An up to date acknowledgment statement can be found on the ADSP site: https://www.niagads.org/adsp/content/acknowledgement-statement.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery

Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Additional information to include in an acknowledgment statement can be found on the LONI site: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Data_Use_Agreement.pdf.

The Alzheimer's Disease Genetics Consortium (ADGC) supported sample preparation, whole exome sequencing and data processing through NIA grant U01AG032984. Sequencing data generation and harmonization is supported by the Genome Center for Alzheimer's Disease, U54AG052427, and data sharing is supported by NIAGADS, U24AG041689. Samples from the National Centralized Repository for Alzheimer's Disease and Related Dementias (NCRAD), which receives government support under a cooperative agreement grant (U24 AG021886) awarded by the National Institute on Aging (NIA), were used in this study. We thank contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible. NIH grants supported enrollment and data collection for the individual studies including: GenerAAtions R01AG20688 (PI M. Daniele Fallin, PhD); Miami/Duke R01 AG027944, R01 AG028786 (PI Margaret A. Pericak-Vance, PhD); NC A&T P20 MD000546, R01 AG28786-01A1 (PI Goldie S. Byrd, PhD); Case Western (PI Jonathan L. Haines, PhD); MIRAGE R01 AG009029 (PI Lindsay A. Farrer, PhD); ROS P30AG10161, R01AG15819, R01AG30146, TGen (PI David A. Bennett, MD); MAP R01AG17917, R01AG15819, TGen (PI David A. Bennett, MD). The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P30 AG062428-01 (PI James Leverenz, MD) P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P30 AG062421-01 (PI Bradley Hyman, MD, PhD), P30 AG062422-01 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI Robert Vassar, PhD), P30 AG008017 (PI Jeffrey Kaye, MD),

P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P30 AG062429-01(PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P30 AG062715-01 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD).

This work was supported by grants from the National Institutes of Health (R01AG044546, P01AG003991, RF1AG053303, R01AG058501, U01AG058922, RF1AG058501 and R01AG057777). The recruitment and clinical characterization of research participants at Washington University were supported by NIH P50 AG05681, P01 AG03991, and P01 AG026276. This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, and the Departments of Neurology and Psychiatry at Washington University School of Medicine.

We thank the contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible. Members of the National Institute on Aging Late-Onset Alzheimer Disease/National Cell Repository for Alzheimer Disease (NIA-LOAD NCRAD) Family Study Group include the following: Richard Mayeux, MD, MSc; Martin Farlow, MD; Tatiana Foroud, PhD; Kelley Faber, MS; Bradley F. Boeve, MD; Neill R. Graff-Radford, MD; David A. Bennett, MD; Robert A. Sweet, MD; Roger Rosenberg, MD; Thomas D. Bird, MD; Carlos Cruchaga, PhD; and Jeremy M. Silverman, PhD.

This work was partially supported by grant funding from NIH R01 AG039700 and NIH P50 AG005136. Subjects and samples used here were originally collected with grant funding from NIH U24 AG026395, U24 AG021886, P50 AG008702, P01 AG007232, R37 AG015473, P30 AG028377, P50 AG05128, P50 AG16574, P30 AG010133, P50 AG005681, P01 AG003991, U01MH046281, U01 MH046290 and U01 MH046373. The funders had no role in study design, analysis or preparation of the manuscript. The authors declare no competing interests.

This work was supported by the National Institutes of Health (R01 AG027944, R01 AG028786 to MAPV, R01 AG019085 to JLH, P20 MD000546); a joint grant from the Alzheimer's Association (SG-14-312644) and the Fidelity Biosciences Research Initiative to MAPV; the BrightFocus Foundation (A2011048 to MAPV). NIA-LOAD Family-Based Study supported the collection of samples used in this study through NIH grants U24 AG026395 and R01 AG041797 and the MIRAGE cohort was supported through the NIH grants R01 AG025259 and R01 AG048927. We thank contributors,

including the Alzheimer's disease Centers who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible. Study design: HNC, BWK, JLH, MAPV; Sample collection: MLC, JMV, RMC, LAF, JLH, MAPV; Whole exome sequencing and Sanger sequencing: SR, PLW; Sequencing data analysis: HNC, BWK, KLHN, SR, MAK, JRG, ERM, GWB, MAPV; Statistical analysis: BWK, KLHN, JMJ, MAPV; Preparation of manuscript: HNC, BWK. The authors jointly discussed the experimental results throughout the duration of the study. All authors read and approved the final manuscript.

Data collection and sharing for this project was supported by the Washington Heights-Inwood Columbia Aging Project (WHICAP, PO1AG07232, R01AG037212, RF1AG054023) funded by the National Institute on Aging (NIA) and by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1TR001873. This manuscript has been reviewed by WHICAP investigators for scientific content and consistency of data interpreTation with previous WHICAP Study publications. We acknowledge the WHICAP study participants and the WHICAP research and support staff for their contributions to this study.

This work was supported by grants from the National Institutes of Health (R01AG044546, P01AG003991, RF1AG053303, R01AG058501, U01AG058922, RF1AG058501 and R01AG057777). The recruitment and clinical characterization of research participants at Washington University were supported by NIH P50 AG05681, P01 AG03991, and P01 AG026276. This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, and the Departments of Neurology and Psychiatry at Washington University School of Medicine.

We thank the contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible. Members of the National Institute on Aging Late-Onset Alzheimer Disease/National Cell Repository for Alzheimer Disease (NIA-LOAD NCRAD) Family Study Group include the following: Richard Mayeux, MD, MSc; Martin Farlow, MD; Tatiana Foroud, PhD; Kelley Faber, MS; Bradley F. Boeve, MD; Neill R. Graff-Radford, MD; David A. Bennett, MD; Robert A. Sweet, MD; Roger Rosenberg, MD; Thomas D. Bird, MD; Carlos Cruchaga, PhD; and Jeremy M. Silverman, PhD.

This work was supported by grants from the National Institutes of Health (R01AG044546, P01AG003991, RF1AG053303, R01AG058501, U01AG058922, RF1AG058501 and R01AG057777). The recruitment and clinical characterization of research participants at Washington University were supported by NIH P50 AG05681, P01 AG03991, and P01 AG026276. This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, and the Departments of Neurology and Psychiatry at Washington University School of Medicine.

We thank the contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible. Members of the National Institute on Aging Late-Onset Alzheimer Disease/National Cell Repository for Alzheimer Disease (NIA-LOAD NCRAD) Family Study Group include the following: Richard Mayeux, MD, MSc; Martin Farlow, MD; Tatiana Foroud, PhD; Kelley Faber, MS; Bradley F. Boeve, MD; Neill R. Graff-Radford, MD; David A. Bennett, MD; Robert A. Sweet, MD; Roger Rosenberg, MD; Thomas D. Bird, MD; Carlos Cruchaga, PhD; and Jeremy M. Silverman, PhD.

Mayo RNAseq Study- Study data were provided by the following sources: The Mayo Clinic Alzheimer's Disease Genetic Studies, led by Dr. Nilufer Ertekin-Taner and Dr. Steven G. Younkin, Mayo Clinic, Jacksonville, FL using samples from the Mayo Clinic Study of Aging, the Mayo Clinic Alzheimer's Disease Research Center, and the Mayo Clinic Brain Bank. Data collection was supported through funding by NIA grants P50 AG016574, R01 AG032990, U01 AG046139, R01 AG018023, U01 AG006576, U01 AG006786, R01 AG025711, R01 AG017216, R01 AG003949, NINDS grant R01 NS080820, CurePSP Foundation, and support from Mayo Foundation. Study data includes samples collected through the Sun Health Research Institute Brain and Body Donation Program of Sun City, Arizona. The Brain and Body Donation Program is supported by the National Institute of Neurological Disorders and Stroke (U24 NS072026 National Brain and Tissue Resource for Parkinson's Disease and Related Disorders), the National Institute on Aging (P30 AG19610 Arizona Alzheimer's Disease Core Center), the Arizona Department of Health Services (contract 211002, Arizona Alzheimer's Research Center), the Arizona Biomedical Research Commission (contracts 4001, 0011, 05-901 and 1001 to the Arizona Parkinson's Disease Consortium) and the Michael J. Fox Foundation for Parkinson's Research

ROSMAP- We are grateful to the participants in the Religious Order Study, the Memory and Aging Project. This work is supported by the US National Institutes of Health [U01 AG046152, R01 AG043617, R01 AG042210, R01 AG036042, R01 AG036836, R01 AG032990, R01 AG18023, RC2 AG036547, P50 AG016574, U01 ES017155, KL2 RR024151, K25 AG041906-01, R01 AG30146, P30 AG10161, R01 AG17917, R01 AG15819, K08 AG034290, P30 AG10161 and R01 AG11101.

Mount Sinai Brain Bank (MSBB)- This work was supported by the grants R01AG046170, RF1AG054014, RF1AG057440 and R01AG057907 from the NIH/National Institute on Aging (NIA). R01AG046170 is a component of the AMP-AD Target Discovery and Preclinical Validation Project. Brain tissue collection and characterization was supported by NIH HHSN271201300031C.

This study was supported by the National Institute on Aging (NIA) grants AG030653, AG041718, AG064877 and P30-AG066468.

We would like to thank study participants, their families, and the sample collectors for their invaluable contributions. This research was supported in part by the National Institute on Aging grant U01AG049508 (PI Alison M. Goate). This research was supported in part by Genentech, Inc. (PI Alison M. Goate, Robert R. Graham).

The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by these NIA-funded ADCs: P30 AG013846 (PI Neil Kowall, MD), P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P30 AG013854 (PI M. Marsel Mesulam, MD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG005131 (PI Douglas Galasko, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG005681 (PI John Morris, MD), P30 AG028377 (Kathleen Welsh-Bohmer, PhD), and P50 AG008671 (PI Henry Paulson, MD, PhD).

Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), which receives government support under a cooperative agreement grant (U24 AG21886) awarded by the National Institute on Aging (NIA), were used in this study. We thank contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible.

The Alzheimer's Disease Genetics Consortium supported the collection of samples used in this study through National Institute on Aging (NIA) grants U01AG032984 and RC2AG036528.

We acknowledge the generous contributions of the Cache County Memory Study participants. Sequencing for this study was funded by RF1AG054052 (PI: John S.K. Kauwe)

Acknowledgments for the use of GWAS data distributed by NIAGADS

The NIA Genetics of Alzheimer's Disease Data Storage Site (NIAGADS) is supported by a collaborative agreement from the National Institute on Aging, U24AG041689.

NG00047: The NIA supported this work through grants U01-AG032984, RC2-AG036528, U01-AG016976 (Dr Kukull); U24 AG026395, U24 AG026390, R01AG037212, R37 AG015473 (Dr Mayeux); K23AG034550 (Dr Reitz); U24-AG021886 (Dr Foroud); R01AG009956, RC2 AG036650 (Dr Hall); UO1 AG06781, UO1 HG004610 (Dr Larson); R01 AG009029 (Dr Farrer); 5R01AG20688 (Dr

Fallin); P50 AG005133, AG030653 (Dr Kamboh); R01 AG019085 (Dr Haines); R01 AG1101, R01 AG030146, RC2 AG036650 (Dr Evans); P30AG10161, R01AG15819, R01AG30146, R01AG17917, R01AG15819 (Dr Bennett); R01AG028786 (Dr Manly); R01AG22018, P30AG10161 (Dr Barnes); P50AG16574 (Dr Ertekin-Taner, Dr Graff-Radford), R01 AG032990 (Dr Ertekin-Taner), KL2 RR024151 (Dr Ertekin-Taner); R01 AG027944, R01 AG028786 (Dr Pericak-Vance); P20 MD000546, R01 AG28786-01A1 (Dr Byrd); AG005138 (Dr Buxbaum); P50 AG05681, P01 AG03991, P01 AG026276 (Dr Goate); and P30AG019610, P30AG13846, U01-AG10483, R01CA129769, R01MH080295, R01AG017173, R01AG025259, R01AG33193, P50AG008702, P30AG028377, AG05128, AG025688, P30AG10133, P50AG005146, P50AG005134, P01AG002219, P30AG08051, MO1RR00096, UL1RR029893, P30AG013854, P30AG008017, R01AG026916, R01AG019085, P50AG016582, UL1RR02777, R01AG031581, P30AG010129, P50AG016573, P50AG016575, P50AG016576, P50AG016577, P50AG016570, P50AG005131, P50AG023501, P50AG019724, P30AG028383, P50AG008671, P30AG010124, P50AG005142, P30AG012300, AG010491, AG027944, AG021547, AG019757, P50AG005136 (Alzheimer Disease GeneticsConsortium [ADGC]). We thank Creighton Phelps, Stephen Synder, and Marilyn Miller from the NIA, who are ex-officio members of the ADGC. Support was also provided by the Alzheimer's Association (IIRG-08-89720 [Dr Farrer] and IIRG-05-14147 [Dr Pericak-Vance]), National Institute of Neurological Disorders and Stroke grant NS39764, National Institute of Mental Health grant MH60451, GlaxoSmithKline, and the Office of Research and Development, Biomedical Laboratory Research Program, US Department of Veterans Affairs Administration. For the ADGC, biological samples and associated phenotypic data used in primary data analyses were stored at principal investigators' institutions and at the National Cell Repository for Alzheimer's Disease (NCRAD) at Indiana University, funded by the NIA. Associated phenotypic data used in secondary data analyses were stored at the National Alzheimer's Coordinating Center and at the NIA Alzheimer's Disease Data Storage Site at the University of Pennsylvania, funded by the NIA. Contributors to the genetic analysis data included principal investigators on projects ndividually funded by the NIA, other NIH institutes, or private entities.

Acknowledgments for the use of MARS and LATC

We thank all Minority Aging Research Study and Latino Core participants and the Rush Alzheimer’s Disease Center staff. This database was funded by the NIH/NIA grants R01AG22018 (MARS) and P30AG 072975 (ADC).

Acknowledgments for other GWAS and phenotype data

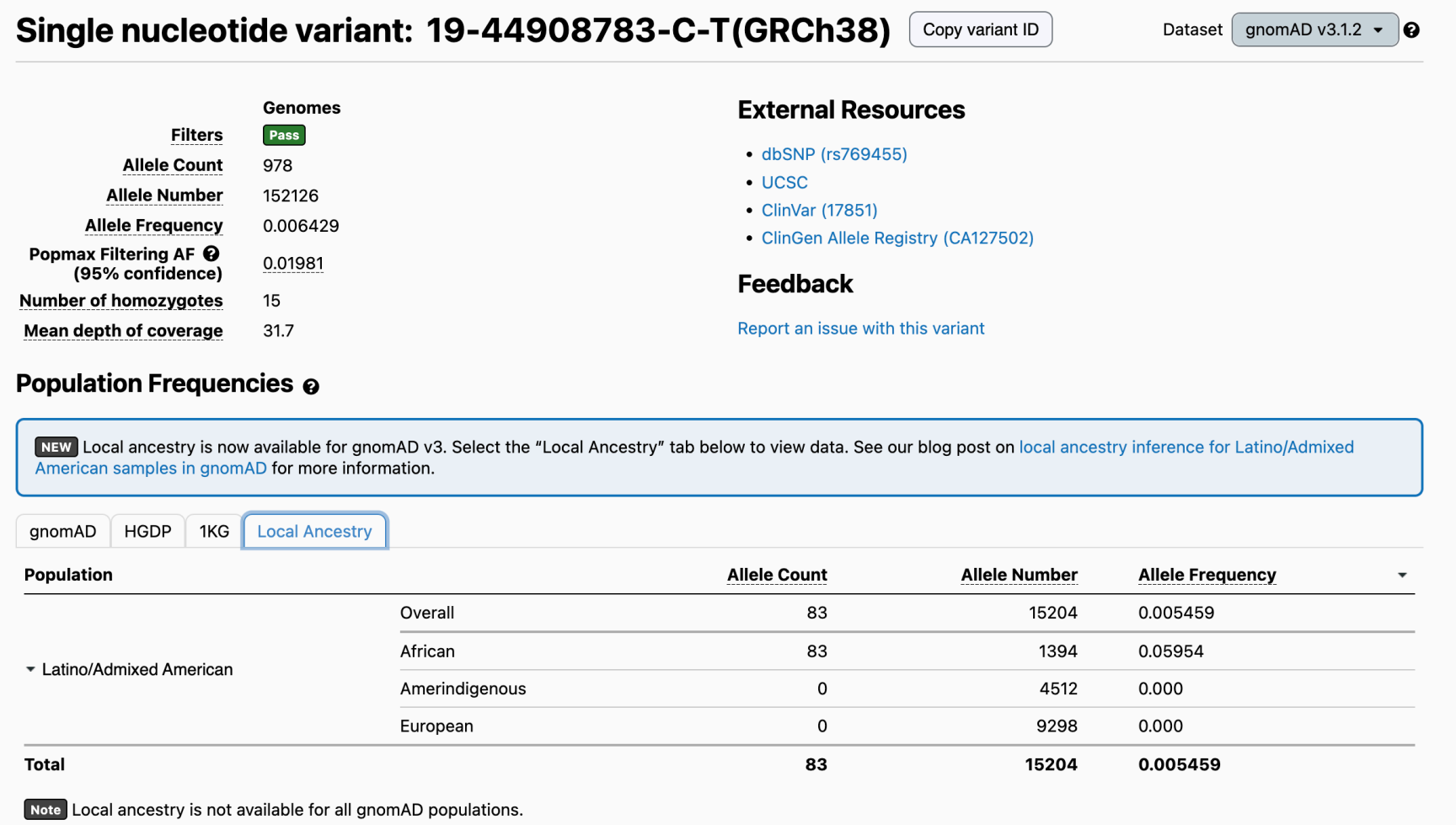
The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P30 AG062428-01 (PI James Leverenz, MD) P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin,

PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P30 AG062421-01 (PI Bradley Hyman, MD, PhD), P30 AG062422-01 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI Robert Vassar, PhD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P30 AG062429-01(PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P30 AG062715-01 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD).

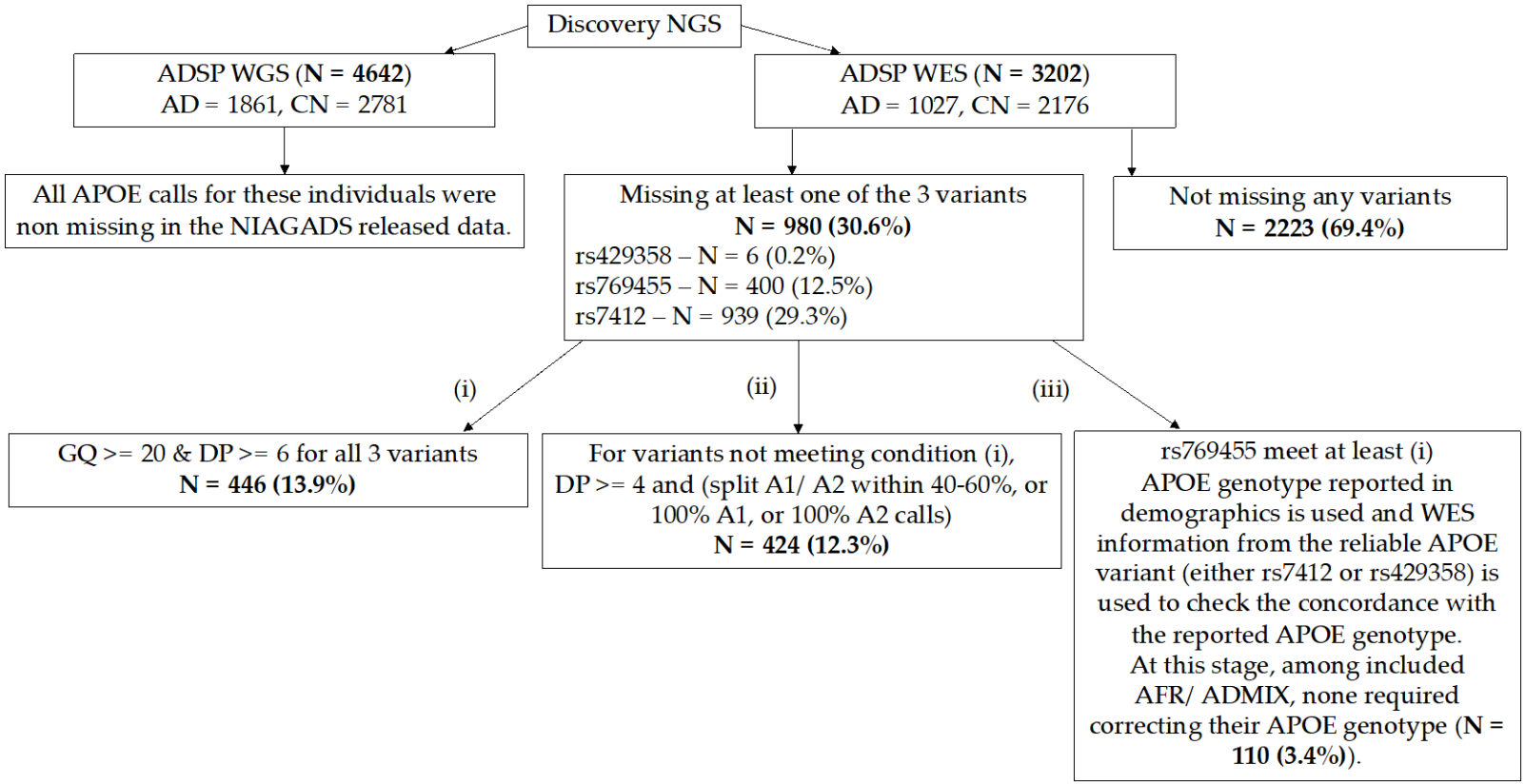
The genotypic and associated phenotypic data used in the study "Multi-Site Collaborative Study for Genotype-Phenotype Associations in Alzheimer's Disease (GenADA)" were provided by the GlaxoSmithKline, R&D Limited.

ROSMAP study data were provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. Data collection was supported through funding by NIA grants P30AG10161, R01AG15819, R01AG17917, R01AG30146, R01AG36836, U01AG32984, U01AG46152, the Illinois Department of Public Health, and the Translational Genomics Research Institute.

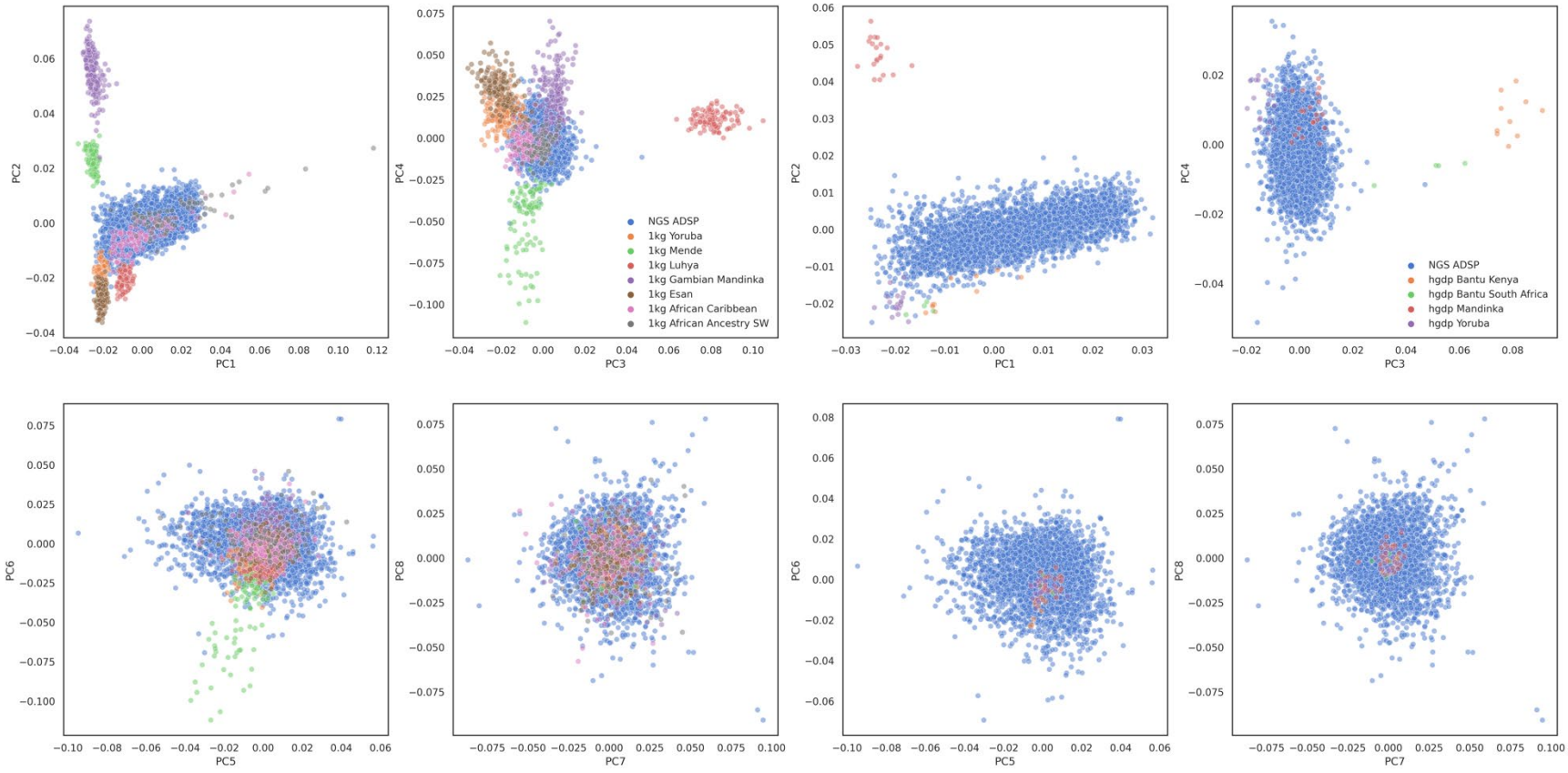
The AddNeuroMed data are from a public-private partnership supported by EFPIA companies and SMEs as part of InnoMed (Innovative Medicines in Europe), an Integrated Project funded by the European Union of the Sixth Framework program priority FP6-2004-LIFESCIHEALTH-5. Clinical leads responsible for data collection are Iwona Kłoszewska (Lodz), Simon Lovestone (London), Patrizia Mecocci (Perugia), Hilkka Soininen (Kuopio), Magda Tsolaki (Thessaloniki), and Bruno Vellas (Toulouse), imaging leads are Andy Simmons (London), Lars-Olad Wahlund (Stockholm) and Christian Spenger (Zurich) and bioinformatics leads are Richard Dobson (London) and Stephen Newhouse (London).



eFigure 1. Local Ancestry Inference at rs769455 in Latino/Admixed American Participants in gnomAD v3.1.2

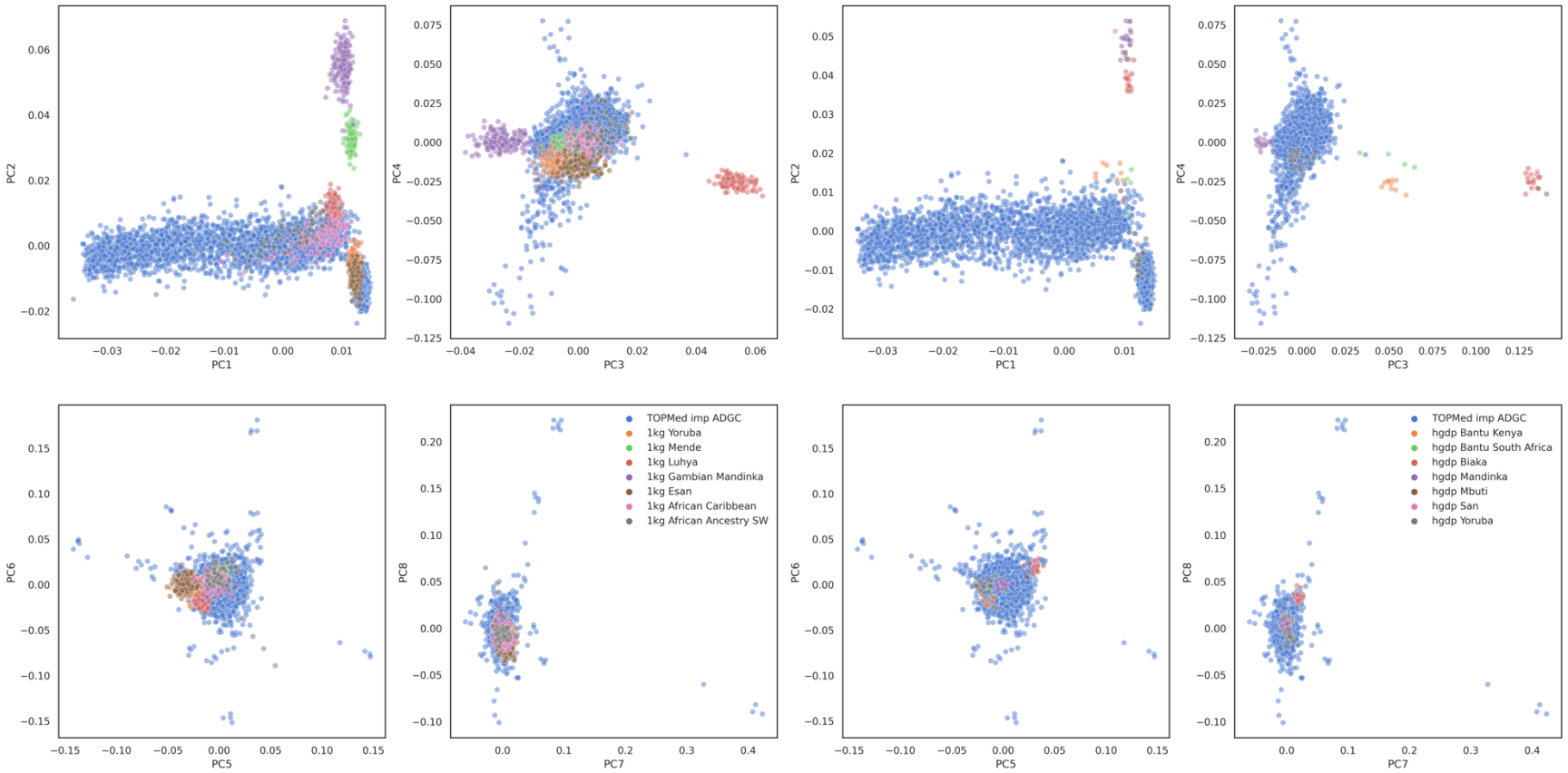


eFigure 2. Flowchart Describing the *APOE* Alleles and R145C Genotyping Among African and Admixed-African Individuals Included in the Discovery Analysis



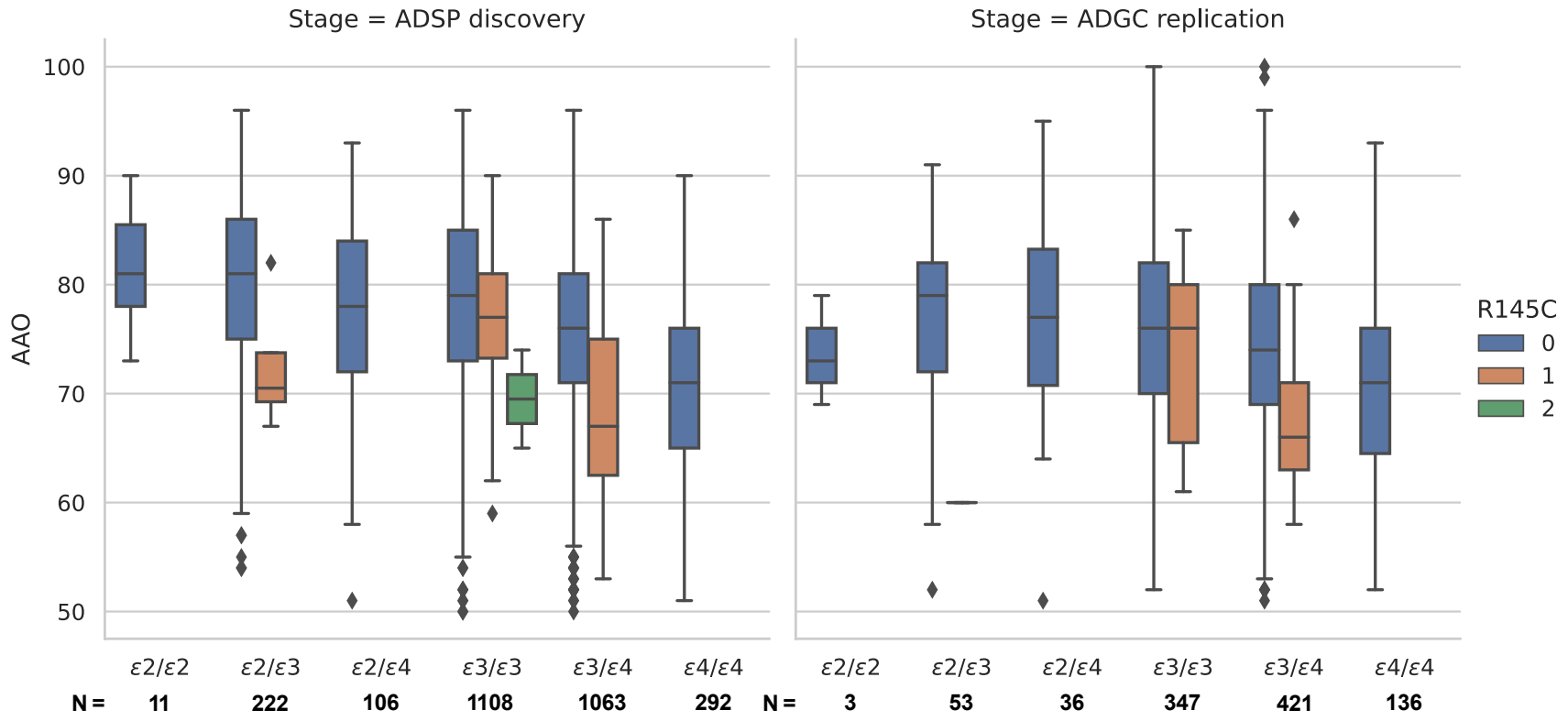
eFigure 3. Principal Ancestry Components Computed on the Merge of the ADSP (Discovery) African Ancestry Participants (AFR% > 75%) and African Participants in 1000 Genones Project (1kg) and Human Genome Diversity Project (hgdp)

Left panels present these principal components with the 1kg reference populations and right panels with the hgdp.



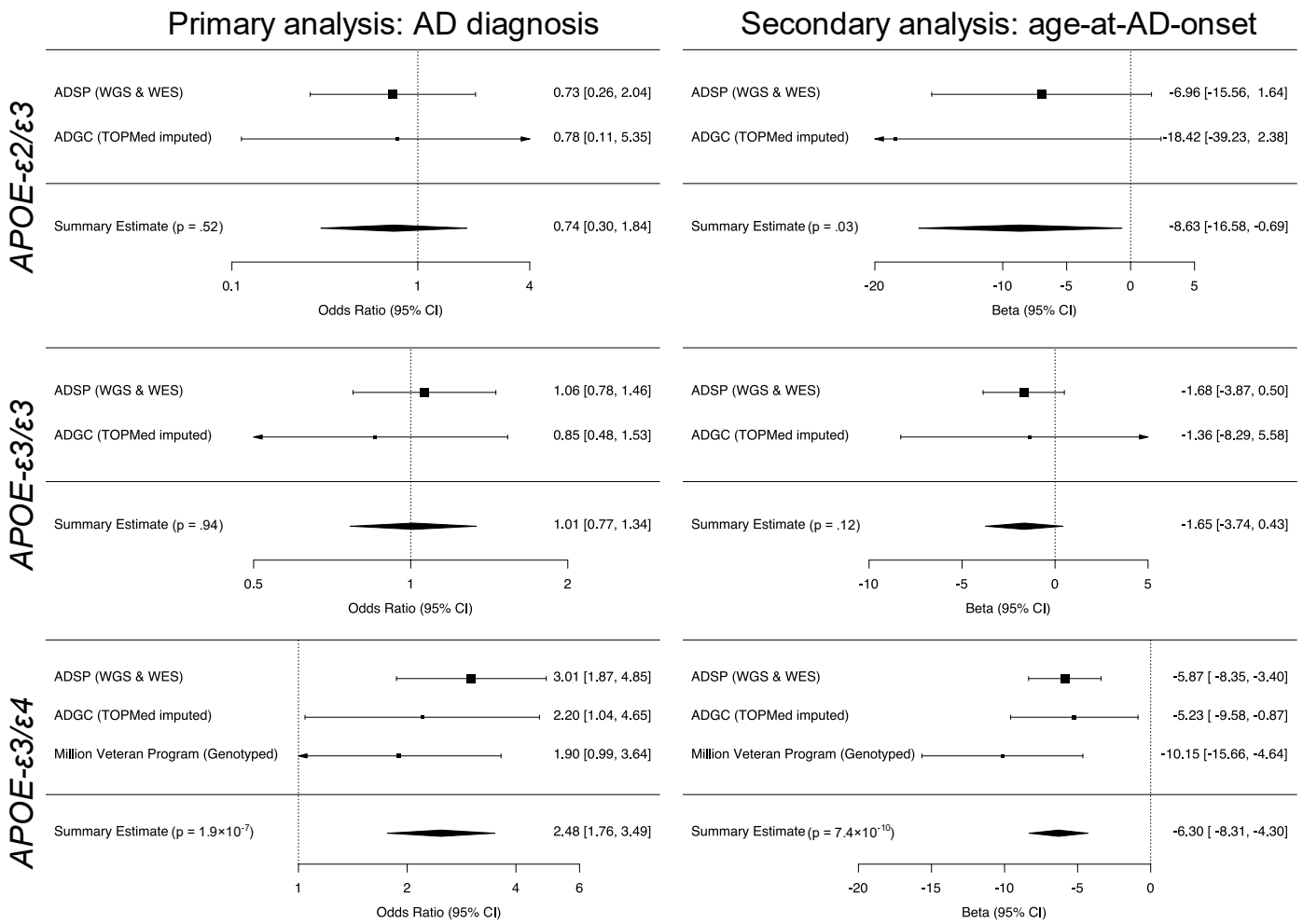
eFigure 4. Principal Ancestry Components Computed on the Merge of the ADGC (Replication) African Ancestry Participants (AFR% > 75%) and African Participants in 1000 Genones Project and Human Genome Diversity Project

Left panels present these principal components with the 1kg reference populations and right panels with the hgdp.



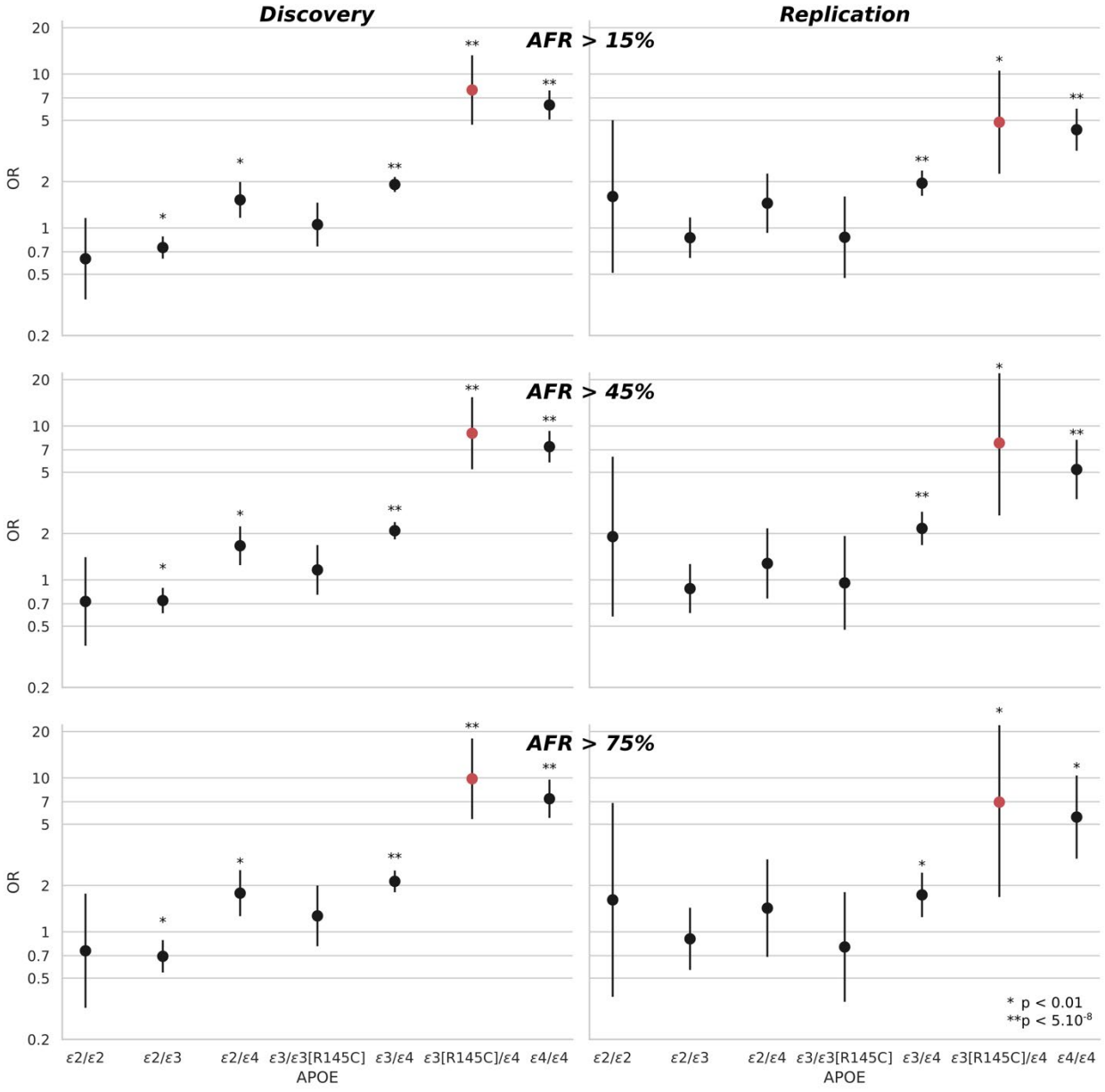
eFigure 5. Age-at-AD-Onset (AAO) Distribution Across *APOE* Genotype Groups in the ADSP Discovery and ADGC Replication in Function of R145C Allele

The box shows the quartiles of the dataset while the whiskers extend to show the rest of the distribution, except for points that are determined to be "outliers" using a method that is a function of the inter-quartile range (see *seaborn* Python package documentation for more details).



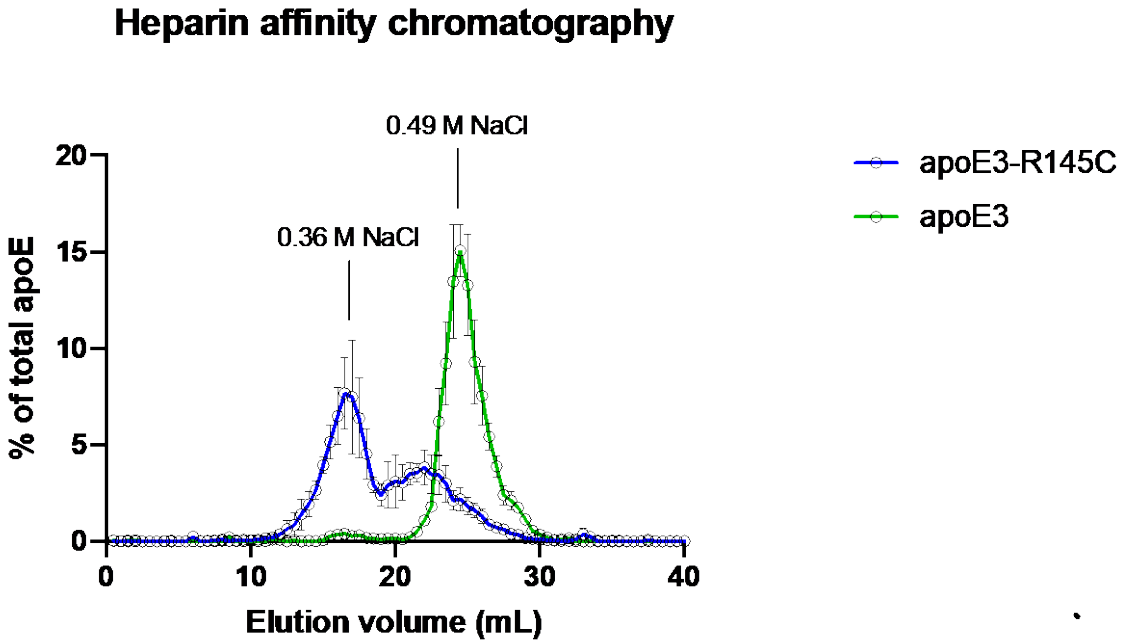
eFigure 6. R145C Association With Alzheimer Disease (AD) Risk and Age-at-AD-Onset

Forest plots equivalent to Table 3.



eFigure 7. *APOE ε3*[R145C]/*ε4* Individuals Have an AD Risk Comparable to *APOE ε4/ε4* Individuals Regardless of the African Ancestry Cutoff

Sensitivity analysis for the AD risk odds ratio (OR) per *APOE* group in the discovery and replication for several African ancestry cutoffs (15%, 45%, and 75%). AD risk per *APOE* group assessed compared to the *APOE ε3/ε3* reference group (i.e., $OR_{APOE\varepsilon3/\varepsilon3} = 1$) in our discovery sample (left column) composed of next generation sequencing data from the ADSP dataset, and in our replication (right column) composed of microarray data imputed on the TOPMed reference panel.



eFigure 8. ApoE3-R145C Shows Significantly Reduced Heparin Binding Compared to apoE3

Elution profile of apoE3 and apoE3-R145C from heparin affinity chromatography. ApoE was eluted from the column with a linear gradient of NaCl (0 to 0.8 M); apoE in each fraction was quantified by ELISA and expressed as a percentage of total apoE. Error bars represent SEM, N=3. ApoE3-R145C was primarily eluted at a significantly lower salt concentration than apoE3 (0.36 M NaCl vs 0.49 M NaCl; paired t-test: p=0.0258).

eTable 1. Missense Variants on *APOE* Canonical Transcript Reported in gnomADv.3.1

Pos: position on chromosome 19 in build hg38, AF: alternate allele frequency, AC: allele count, AN: allele number, Hom: Homozygote count. Only the first ten missense variants in term of allele frequency in gnomAD are reported here (AC > 20).

				HGVs		Overall gnomAD.v3.1				African/African-American				Latino/Admixed American				European (non-Finnish)			
Pos (hg38)	rsIDs	Ref	Alt	new	standard	AC	AN	AF	Hom	AC	AN	AF	Hom	AC	AN	AF	Hom	AC	AN	AF	Hom
44908684	rs429358	T	C	p.Cys130Arg	p.Cys112Arg	23875	151972	0.157101	1998	8917	41390	0.215439	950	1718	15280	0.112435	90	9371	67924	0.137963	676
44908822	rs7412	C	T	p.Arg176Cys	p.Arg158Cys	11838	152004	0.07788	546	4335	41402	0.104705	247	640	15262	0.041934	17	5401	67956	0.079478	224
44908783	rs769455	C	T	p.Arg163Cys	p.Arg145Cys	978	152126	0.006429	15	869	41444	0.020968	15	83	15274	0.005434	0	5	67996	7.35E-05	0
44907853	rs769452	T	C	p.Leu46Pro	p.Leu28Pro	293	152188	0.001925	0	11	41454	0.000265	0	14	15282	0.000916	0	168	68034	0.002469	0
44909057	rs199768005	T	A	p.Val254Glu	p.Val236Glu	72	152080	0.000473	0	16	41424	0.000386	0	0	15272	0	0	52	67984	0.000765	0
44907807	rs201672011	G	A	p.Glu31Lys	p.Glu13Lys	72	152186	0.000473	0	10	41458	0.000241	0	48	15282	0.003141	0	5	68024	7.35E-05	0
44908799	rs376170967	G	A	p.Arg168His	p.Arg150His	62	152122	0.000408	0	58	41452	0.001399	0	3	15262	0.000197	0	1	67994	1.47E-05	0
44909101	rs267606661	C	G	p.Arg269Gly	p.Arg251Gly	46	152200	0.000302	0	5	41450	0.000121	0	7	15280	0.000458	0	33	68022	0.000485	0
44908730	rs267606664	G	A	p.Gly145Asp	p.Gly127Asp	22	152152	0.000145	0	3	41458	7.24E-05	0	1	15284	6.54E-05	0	17	67992	0.00025	0
44908915	rs749750245	C	T	p.Arg207Cys	p.Arg189Cys	21	151884	0.000138	1	0	41406	0	0	21	15254	0.001377	1	0	67920	0	0

eTable 2. Queried Cohort Overview to Identify Admixed and African Ancestry Individuals

Cohort/Project	Genotyping Platform	Cohort-Platform ID	Sample (N)	Data Repository and Access ID
ADSP WES	Whole Exome Sequencing	ADSP_WES	20503	NIAGADS DSS (NG00067.v3) / NACC
ADSP WGS	Whole Genome Sequencing	ADSP_WGS	16906	NIAGADS DSS (NG00067.v5) / NACC
ACT	Illumina Human 660W-Quad	ACT	2790	NIAGADS (NG00034) / dbGaP (phs000234)
ADC1	Illumina Human 660W-Quad	ADC1	2731	NIAGADS (NG00022) / NACC
ADC2	Illumina Human 660W-Quad	ADC2	928	NIAGADS (NG00023) / NACC
ADC3	Illumina Human OmniExpress	ADC3	1526	NIAGADS (NG00024) / NACC
ADC4	Illumina Human OmniExpress	ADC4	1054	NIAGADS (NG00068) / NACC
ADC5	Illumina Human OmniExpress	ADC5	1224	NIAGADS (NG00069) / NACC
ADC6	Illumina Human OmniExpress	ADC6	1333	NIAGADS (NG00070) / NACC
ADC7	Illumina Infinium Human OmniExpressExome	ADC7	1462	NIAGADS (NG00071) / NACC
ADDNEUROMED	Illumina Human 610-Quad	ADM_Q	315	Synapse AddNeuroMed (syn4907804)
	Illumina Human OmniExpress	ADM_O	329	Synapse AddNeuroMed (syn4907804)
ADGC-ExomeChip	Illumina HumanExome BeadChip v1.0 at CHOP	CHOP	5180	NIAGADS (NG00081) / NACC
	Illumina HumanExome BeadChip v1.0 at Miami	MIA	1923	NIAGADS (NG00080) / NACC
	Illumina HumanExome BeadChip v1.0 at Northshore	NS	5998	NIAGADS (NG00079) / NACC
	Illumina HumanExome BeadChip v1.0 at WashU	WU	868	NIAGADS (NG00085) / NACC
ADNI	Illumina Human 610-Quad	ADNI_Q	757	LONI ADNI
	Illumina Human OmniExpress	ADNI_OE	361	LONI ADNI
	Illumina Omni 2.5	ADNI_O25	812	LONI ADNI
	Illumina Human OmniExpress	ADNI_DOD	204	LONI ADNIDOD
ADNI3	Illumina Global Screening Array (GSA)	ADNI3	327	LONI ADNI
IIDP African Americans	Illumina Human 1M-Duo	IIDP_AA	1175	NIAGADS (NG00047)
IIDP Yorubans	Illumina Human 1M-Duo	IIDP_YOR	1264	NIAGADS (NG00047) / cf. gaaindata.org/partner/IIDP
CIDR	Illumina Human Omni1-Quad	CIDR	3101	NIAGADS (NG00015) / dbGAP (phs000160)
GenADA	Affymetrix 500K	GSK	1571	dbGaP (phs000219)
LATC	Illumina Multi-Ethnic – BU	LATC	63	RADC Rush / Latino CORE Study
NIA-LOAD	Illumina Human 610-Quad	LOAD	5220	NIAGADS (NG00020)

MARS	Illumina Multi-Ethnic – BU	MARS	708	RADC Rush / Minority Aging Research Study
MAYO	Illumina Human Hap300	MAYO_1	2099	Synapse AMP-AD (syn5591675)
MAYO2	Illumina Omni 2.5	MAYO_2	314	Synapse AMP-AD (syn5550404)
MIRAGE	Illumina Human CNV370-Duo	MIRAGE_370	397	NIAGADS (NG00031)
	Illumina Human 610-Quad	MIRAGE_610	1105	NIAGADS (NG00031)
MTC	Illumina Human OmniExpress	MTC	542	NIAGADS (NG00096)
OHSU	Illumina Human CNV370-Duo	OHSU	647	NIAGADS (NG00017)
ROSMAP	Affymetrix GeneChip 6.0 - Broad Institute	ROSMAP_1B	1126	RADC Rush / Synapse AMP-AD (syn3219045)
	Affymetrix GeneChip 6.0 - TGen	ROSMAP_1T	582	RADC Rush / Synapse AMP-AD (syn3219045)
	Illumina Human OmniExpress 12 - Chop	ROSMAP_2C	382	RADC Rush / Synapse AMP-AD (syn7824841)
	Illumina Multi-Ethnic - BU	ROSMAP_3BU	494	RADC Rush
TARCC	Affymetrix 6.0	TARCC	2718	NIAGADS (NG00097) / TARCC study
TGEN2	Affymetrix 6.0	TGEN	1599	NIAGADS (NG00028)
UPITT	Illumina Human Omni1-Quad	UPITT	2440	NIAGADS (NG00026)
UM-VU-MSSM	Illumina Human 1M-Duo, Illumina 1M	UVM_A	1153	NIAGADS (NG00042)
	Affymetrix 6.0	UVM_B	864	NIAGADS (NG00042)
	Illumina Human 550K. Illumina Human 610-Quad	UVM_C	445	NIAGADS (NG00042)
WASHU	Illumina Human 610-Quad	WASHU_1	670	NIAGADS (NG00030)
WASHU2	Illumina Human OmniExpress	WASHU_2	235	NIAGADS (NG00087)
WHICAP	Illumina Human OmniExpress	WHICAP	647	NIAGADS (NG00093)

eTable 3. Overview of Alzheimer's Disease Sequencing Project (ADSP) Studies With Whole-Exome Sequencing (WES) and/or Whole-Genome Sequencing (WGS) Available at NIAGADS DSS (NG00067)

Study	Accession Number	Related Datasets
Accelerating Medicines Partnership- Alzheimer's Disease (AMP-AD)	sa000011	NG00067 – ADSP Umbrella
Cache County Study	sa000014	NG00067 – ADSP Umbrella
University of Pittsburgh- Kamboh WGS	sa000012	NG00067 – ADSP Umbrella
CurePSP and Tau Consortium PSP WGS	sa000016	NG00067 – ADSP Umbrella
NIH, CurePSP and Tau Consortium PSP WGS	sa000015	NG00067 – ADSP Umbrella
UCLA Progressive Supranuclear Palsy	sa000017	NG00067 – ADSP Umbrella
NACC Genentech WGS	sa000013	NG00067 – ADSP Umbrella
Alzheimer's Disease Sequencing Project (ADSP)	sa000001	NG00067 – ADSP Umbrella
Alzheimer's Disease Neuroimaging Initiative (ADNI)	sa000002	NG00067 – ADSP Umbrella
Alzheimer's Disease Genetics Consortium: African Americans (ADGC AA)	sa000003	NG00067 – ADSP Umbrella
The Familial Alzheimer Sequencing (FASe) project	sa000004	NG00067 – ADSP Umbrella
Brkanac – Family-based genome scan for AAO of LOAD	sa000005	NG00067 – ADSP Umbrella
HIHG Miami Families with AD	sa000006	NG00067 – ADSP Umbrella
Washington Heights/Inwood Columbia Aging Project (WHICAP)	sa000007	NG00067 – ADSP Umbrella
Charles F. and Joanne Knight Alzheimer's Disease Research Center (Knight ADRC)	sa000008	NG00067 – ADSP Umbrella
Corticobasal degeneration Study (CBD)	sa000009	NG00067 – ADSP Umbrella
Progressive Supranuclear Palsy Study (PSP)	sa000010	NG00067 – ADSP Umbrella

eTable 4. Pathogenic Variants Identified on *APP, PSEN1, PSEN2, MAPT* in ADSP WES and WGS

Briefly all variants within these genes were extracted and annotated with VEP (Ensembl Variant Effect Predictor) using the ClinVar plugins. Variants with ClinVar "pathogenic" status linked to a neurodegenerative disease in these genes were selected and their carriers' details were annotated. Evidence of pathogenic status was also checked in the Alzforum mutations library. Demographics are reported as follows AgeSex*APOE* (eg. 40M33 means 40 years old male ε3ε3), -9 represents missing information. These individuals were excluded from statistical analyses.

ADSP	Gene	HGVSp	rsid	Consequence	CADD	AD demographics	CN demographics	Other diagnoses demographics	Link Alzforum
WES	*APP*	p.I716T	rs63750851	missense	27.8	40M33			https://www.alzforum.org/mutations/app-i716t
WES	*APP*	p.I716V	rs63750399	missense	25.9	54F24		-9M24	https://www.alzforum.org/mutations/app-i716v
WES	*APP*	p.V717F	rs63750264	missense	29.7	57F33			https://www.alzforum.org/mutations/app-v717f-indiana
WGS	*APP*	p.V717G	rs63749964	missense	28.7	61F33			https://www.alzforum.org/mutations/app-v717g
WGS	*MAPT*	p.R741W	rs63750424	missense	29.8	50M33,49F33		61M23	https://www.alzforum.org/mutations/mapt-r406w
WES	*MAPT*	p.R741W	rs63750424	missense	29.8	67M33,62F33,62F33,61M33	41M22	-9M-9	https://www.alzforum.org/mutations/mapt-r406w
WGS	*PSEN1*	p.A431E	rs63750083	missense	27.3	42M33		-9F34,-9M33	https://www.alzforum.org/mutations/psen1-a431e
WGS	*PSEN1*	p.A79V	rs63749824	missense	27.1	56F34,51M34,56M34,66M34	73M34		https://www.alzforum.org/mutations/psen1-a79v
WES	*PSEN1*	p.A79V	rs63749824	missense	27.1	74M34,63M33,68F33,68M33, 69M34,64F33,67M34			https://www.alzforum.org/mutations/psen1-a79v
WGS	*PSEN1*	p.C410Y	rs661	missense	33	54M33,-9M33			https://www.alzforum.org/mutations/psen1-c410y
WGS	*PSEN1*	p.G206A	rs63750082	missense	27.2	56F33,50F33,64M33,43M34, 54M33,72F44,64M23,49F34, 63M33,61F33,55F33,58M34, 59M33,63F33,70F24,57M23	74F33	-9F33,-9M34, -9F33	https://www.alzforum.org/mutations/psen1-g206a
WES	*PSEN1*	p.G206A	rs63750082	missense	27.2	76F23,74F34,74F33,63M33, 65F33		-9F33	https://www.alzforum.org/mutations/psen1-g206a
WGS	*PSEN1*	p.H163R	rs63750590	missense	23.9	49M33			https://www.alzforum.org/mutations/psen1-h163r
WGS	*PSEN1*	p.M139V	rs63751037	missense	23.4	41M33,48F33			https://www.alzforum.org/mutations/psen1-m139v
WGS	*PSEN1*	p.P264L	rs63750301	missense	32	41M33		-9M33	https://www.alzforum.org/mutations/psen1-p264l
WGS	*PSEN1*	p.R269H	rs63750900	missense	29.3	54F34,60M33,61F34,57M33			https://www.alzforum.org/mutations/psen1-r269h
WGS	*PSEN1*	p.Y115C	rs63750450	missense	27.8			-9F33	https://www.alzforum.org/mutations/psen1-y115c
WGS	*PSEN2*	p.N141I	rs63750215	missense	25.3	47F34,59F34			https://www.alzforum.org/mutations/psen2-n141i

eTable 5. Demographics of All the Queried Cohorts

AFR: African, AMR: American (central and south; admixed), EAS: East Asian, SAS South Asian, EUR: European, otherwise ADMIX: admixed of these super ancestry categories.

			Ancestry						Diagnosis		Sex - Females		Age	
	Cohort	N total	AFR N	ADMIX N	AMR N	EAS N	SAS N	EUR N	CN N	AD N	CN N(%)	AD N(%)	CN μ(σ)	AD μ(σ)
Discovery	ADSP WES	20503	3171	3174	125	7	2	14024	9617	8723	6101(63.4)	5394(61.8)	82.0(8.5)	75.7(8.8)
	ADSP WGS	16906	2240	4012	58	68	19	10509	6717	6434	4510(67.1)	3896(60.6)	78.2(8.5)	74.1(10.5)
Replication	ACT	2790	70	64	7	73	0	2576	1833	713	1000(54.6)	462(64.8)	82.9(6.5)	82.1(6.6)
	ADC1	2731	92	58	47	20	0	2514	603	1946	354(58.7)	1039(53.4)	79.8(10.8)	70.7(9.5)
	ADC2	928	0	2	0	0	0	926	124	707	87(70.2)	366(51.8)	80.1(9.2)	72.9(7.1)
	ADC3	1526	0	5	0	0	0	1521	482	858	305(63.3)	468(54.5)	79.6(9.6)	72.5(10.3)
	ADC4	1054	6	10	1	0	0	1037	420	452	257(61.2)	237(52.4)	79.2(8.7)	72.6(9.0)
	ADC5	1224	0	1	0	0	0	1223	579	415	376(64.9)	226(54.5)	82.0(8.9)	74.1(8.7)
	ADC6	1333	0	2	0	0	0	1331	352	567	238(67.6)	304(53.6)	80.1(8.9)	66.9(12.0)
	ADC7	1462	0	4	0	0	0	1458	763	536	493(64.6)	281(52.4)	78.0(7.9)	72.8(7.7)
	ADDNEURO	644	0	2	0	0	0	642	186	256	105(56.5)	164(64.1)	76.4(6.6)	73.0(6.7)
	ADGC-ExomeChip	13969	55	197	12	32	0	13673	5250	7830	3136(59.7)	4585(58.6)	79.6(9.0)	73.0(9.1)
	ADNI	2134	63	69	21	30	5	1945	606	761	260(42.9)	330(43.4)	78.5(7.8)	74.1(7.4)
	ADNI3	327	4	12	1	4	0	306	228	24	142(62.3)	10(41.7)	72.5(6.1)	72.7(9.5)
	CIDR	3101	93	2780	70	0	0	158	1505	1530	1033(68.6)	986(64.4)	74.5(9.4)	75.5(9.6)
	GSK	1571	0	1	1	0	0	1569	773	798	497(64.3)	459(57.5)	73.4(7.9)	72.5(8.6)
	IIDP AA	1175	815	359	0	0	0	1	1001	172	663(66.2)	107(62.2)	83.3(5.3)	83.6(6.7)
	IIDP YOR	1264	1253	10	0	0	0	1	1145	104	732(63.9)	79(76.0)	82.6(5.9)	77.9(7.2)
	LATC	63	13	23	24	0	0	0	15	2	15(100.0)	2(100.0)	77.4(5.4)	78.0(0.0)
	MARS	708	423	275	1	0	0	1	463	79	392(84.7)	54(68.4)	79.6(6.1)	77.3(7.1)
	MAYO	2413	7	24	2	4	0	2335	1225	948	642(52.4)	546(57.6)	75.5(6.5)	74.0(6.0)
	MIRAGE	1502	1	28	2	0	0	1471	738	601	436(59.1)	366(60.9)	72.1(7.3)	68.8(8.6)
	MTC	542	5	29	12	0	0	496	202	272	130(64.4)	157(57.7)	71.7(8.9)	72.6(9.3)
	NIA-LOAD	5220	112	642	13	8	0	4445	2091	2351	1278(61.1)	1546(65.8)	70.6(12.6)	73.6(7.8)
	OHSU	647	3	2	0	1	0	635	379	201	205(54.1)	127(63.2)	85.7(7.5)	85.0(6.9)
	ROSMAP	2584	13	50	28	9	0	2451	1102	951	795(72.1)	690(72.6)	85.4(7.4)	84.1(6.5)
	TARCC	2718	75	218	821	7	2	1557	1124	908	788(70.1)	502(55.3)	70.1(9.8)	70.1(8.9)
	TGEN2	1599	0	9	1	0	1	1512	573	1005	255(44.5)	640(63.7)	80.8(8.7)	72.8(8.0)
	UM-VU-MSSM	2462	5	16	0	0	0	2441	1195	1206	724(60.6)	778(64.5)	74.1(8.2)	74.2(7.9)
	UPITT	2440	7	8	1	0	0	2355	896	1406	563(62.8)	908(64.6)	75.6(6.2)	73.2(6.6)
	WASHU	670	0	0	0	0	0	670	202	429	125(61.9)	239(55.7)	77.9(8.7)	74.0(9.6)
	WASHU2	235	10	1	0	0	0	224	116	68	65(56.0)	38(55.9)	73.7(8.6)	74.0(8.1)
	WHICAP	647	0	7	0	0	0	640	554	85	335(60.5)	60(70.6)	82.7(6.7)	84.1(7.5)

eTable 6. R145C Per Cohort, Diagnosis and *APOE* Genotypes

DX: diagnosis, N: number of individuals, n: number of R145C allele, Rsq: imputation quality, MAF: R145C minor allele frequency within the considered subset, Ntot: number of individuals within the considered *APOE* genotype.

				R145C			*APOE* ε2ε3			*APOE* ε3ε3			*APOE* ε3ε4		
	Cohort	DX	N	n	Rsq	MAF	Ntot	n	MAF	Ntot	n	MAF	Ntot	n	MAF
Discovery	ADSP WES	AD	1027	40	Sequenced	0.019	88	26	0	415	26	0.031	386	14	0.018
		CN	2176	70		0.016	308	55	0.01	1124	55	0.024	588	9	0.008
	ADSP WGS	AD	1861	77	Sequenced	0.021	137	35	0.015	730	35	0.024	707	38	0.027
		CN	2781	102		0.018	401	80	0.015	1498	80	0.027	700	10	0.007
Replication	ACT	AD	8	0	0.95	0	2	0	0	2	0	0	2	0	0
		CN	13	2		0.077	3	2	0	6	2	0.167	4	0	0
	ADC1	AD	10	0	0.85	0	1	0	0	2	0	0	2	0	0
		CN	11	0		0	5	0	0	0	0	0	6	0	0
	ADGC-ExomeChip	AD	24	2	Genotyped	0.042	1	1	0	9	1	0.056	11	1	0.045
		CN	44	2		0.023	9	2	0	21	2	0.048	13	0	0
	ADNI	AD	6	1	1	0.083	0	1	0	2	1	0.25	3	0	0
		CN	1	0		0	0	0	0	0	0	0	0	0	0
	ADNI3	AD	0	0	1	0	0	0	0	0	0	0	0	0	0
		CN	5	1		0.1	0	1	0	2	1	0.25	2	0	0
	CIDR	AD	651	18	0.96	0.014	35	5	0	238	5	0.011	275	13	0.024
		CN	406	9		0.011	31	7	0	218	7	0.016	120	2	0.008
	IIDP AA	AD	41	1	0.98	0.012	5	1	0	17	1	0.029	14	0	0
		CN	642	23		0.018	116	15	0.013	301	15	0.025	170	5	0.015
	IIDP YOR	AD	98	5	0.99	0.026	10	3	0	34	3	0.044	39	2	0.026
		CN	1073	54		0.025	144	40	0.024	502	40	0.04	319	7	0.011
	LATC	AD	0	0	0.95	0	0	0	0	0	0	0	0	0	0
		CN	9	1		0.056	2	0	0	1	0	0	5	1	0.1
	MARS	AD	44	0	0.95	0	7	0	0	20	0	0	11	0	0
		CN	368	22		0.03	56	16	0.009	181	16	0.044	91	5	0.027
	NIA-LOAD	AD	275	9	0.96	0.016	14	2	0.036	91	2	0.011	111	6	0.027
		CN	137	0		0	8	0	0	67	0	0	46	0	0
	ROSMAP	AD	11	1	0.97	0.045	0	1	0	8	1	0.062	3	0	0
		CN	15	1		0.033	1	1	0	10	1	0.05	3	0	0
	TARCC	AD	32	1	0.99	0.016	1	0	0	7	0	0	19	1	0.026
		CN	20	3		0.075	2	2	0	9	2	0.111	7	1	0.071
	UM-VU-MSSM	AD	1	0	0.94	0	0	0	0	0	0	0	1	0	0
		CN	0	0		0	0	0	0	0	0	0	0	0	0

eTable 7. *APOE* Variants Genotype Call in the Data Released by NIAGADS

First set of columns corresponds to number of individuals missing either of the three SNPs, the second set of columns to individuals missing any pair of three SNPs, and third set corresponds to individuals missing all three SNPs.

Cohort N	Missing genotypes – N (%)						
	rs429358	rs769455	rs7412	rs429358 & rs769455	rs429358 & rs7412	rs769455 & rs7412	rs429358 & rs769455 & rs7412
ADSP WES 20,504	760 (3.7%)	3,730 (18.2%)	6,661 (32.5%)	667 (3.3%)	701 (3.4%)	3,607 (17.6%)	666 (3.2%)
ADSP WGS 16,906	330 (2.0%)	499 (3.0%)	481 (2.8%)	278 (1.6%)	263 (1.6%)	442 (2.6%)	256 (1.5%)

eTable 8. Demographics Per Cohort After Ancestry Selection, Quality Control and Duplicates Removal

Dx: diagnosis, AAD: age-at death, AAL: age-at-last-exam, AAE: age-at-exam (and Dx), AAO: age-at-onset.

				Sex	Age	Age Type			
	Cohort	Dx	N	Females (%)	μ(σ)	AAD μ(σ) [%]	AAL μ(σ) [%]	AAE μ(σ) [%]	AAO μ(σ) [%]
Discovery	ADSP WES	AD	1027	69.0%	78.4(8.7)	75.0(-)[0.1%]	-	79.9(6.9)[4.3%]	78.3(8.8)[94.1%]
		CN	2176	69.3%	78.3(7.8)	80.9(8.4)[1.1%]	78.3(7.8)[97.7%]	-	-
	ADSP WGS	AD	1861	68.5%	75.3(8.9)	-	-	74.7(7.5)[0.8%]	75.3(8.9)[98.7%]
		CN	2781	74.1%	75.4(8.7)	82.3(8.8)[3.9%]	75.1(8.6)[92.4%]	-	-
Replication	ACT	AD	8	75.0%	76.1(4.6)	74.2(3.3)[62.5%]	-	-	81.0(4.2)[25.0%]
		CN	13	53.8%	80.4(5.5)	81.1(5.8)[61.5%]	79.2(5.4)[38.5%]	-	-
	ADC1	AD	10	80.0%	71.3(8.1)	-	-	71.0(-)[10.0%]	71.3(8.6)[90.0%]
		CN	11	72.7%	73.8(10.7)	84.0(-)[9.1%]	72.8(10.7)[90.9%]	-	-
	ADGC-ExomeChip	AD	24	79.2%	75.0(7.9)	-	-	-	75.0(7.9)[100.0%]
		CN	44	63.6%	80.3(6.7)	80.6(6.7)[54.5%]	79.9(6.8)[45.5%]	-	-
	ADNI	AD	6	83.3%	73.7(8.9)	-	-	75.4(8.8)[83.3%]	65.0(-)[16.7%]
		CN	1	100%	81.0(-)	-	81.0(-)[100.0%]	-	-
	ADNI3	AD	0	-	-	-	-	-	-
		CN	5	100%	68.6(6.9)	-	68.6(6.9)[100.0%]	-	-
	CIDR	AD	651	65.6%	74.0(10.2)	-	-	-	74.0(10.2)[100.0%]
		CN	406	71.2%	67.4(8.5)	-	67.4(8.5)[100.0%]	-	-
	IIDP AA	AD	41	75.6%	87.6(8.8)	-	-	87.6(8.8)[100.0%]	-
		CN	642	64.0%	82.9(5.5)	-	82.9(5.5)[100.0%]	-	-
	IIDP YOR	AD	98	74.5%	78.0(6.8)	-	-	78.0(6.8)[100.0%]	-
		CN	1073	64.0%	82.6(5.9)	-	82.6(5.9)[100.0%]	-	-
	LATC	AD	0	-	-	-	-	-	-
		CN	9	100%	73.9(3.3)	-	73.9(3.3)[100.0%]	-	-
	MARS	AD	44	68.2%	77.4(6.8)	-	-	77.4(6.8)[100.0%]	-
		CN	368	84.8%	79.3(6.0)	80.0(7.7)[17.4%]	79.2(5.6)[82.6%]	-	-
	NIA-LOAD	AD	275	70.9%	73.9(8.8)	-	-	-	73.9(8.8)[100.0%]
		CN	137	68.6%	64.3(10.7)	78.0(7.9)[3.6%]	63.7(10.5)[96.4%]	-	-
	ROSMAP	AD	11	90.9%	77.5(4.9)	-	-	77.1(4.9)[90.9%]	82.0(-)[9.1%]
		CN	15	93.3%	79.3(8.4)	79.8(6.7)[26.7%]	79.2(9.2)[73.3%]	-	-
	TARCC	AD	32	81.2%	69.9(8.3)	-	-	-	69.9(8.3)[100.0%]
		CN	20	85.0%	66.3(10.1)	-	66.3(10.1)[100.0%]	-	-
	UM-VU-MSSM	AD	1	100%	73.0(-)	-	-	-	73.0(-)[100.0%]
		CN	0	-	-	-	-	-	-

eTable 9. *APOE* R145C (rs769455) Allelic Breakdown by *APOE* Genotype

Rs769455 alternate allele (T) is not observed in *APOE ε2/ε2, ε2/ε4, ε4/ε4*, and is only present in the homozygous state in *APOE ε3/ε3*, supporting the finding in sequencing databases that the alternate allele is always found in phase with *APOE ε3*. Note that rs769455 is located between rs7412 (99 bp apart) and rs429358 (39 bp apart) which define the *APOE* allele genotype. CN: cognitively normal, AD: Alzheimer's disease, N: number of individuals, ADSP: Alzheimer's Disease Sequencing Project, ADGC: Alzheimer's Disease Genetic Consortium, MVP: Million Veteran Program.

			APOE ε2/ε2		*APOE ε2/ε3*		*APOE ε3/ε3*		*APOE ε2/ε4*		*APOE ε3/ε4*		*APOE ε4/ε4*	
Sample	rs769455	N total	CN	AD	CN	AD	CN	AD	CN	AD	CN	AD	CN	AD
ADSP Discovery (Stage I)	C C	7561	41	11	691	221	2490	1086	179	111	1269	1041	118	303
	C T	279	0	0	18	4	129	57	0	0	19	52	0	0
	T T	5	0	0	0	0	3	2	0	0	0	0	0	0
ADGC Replication (Stage II)	C C	3793	26	5	366	75	1236	416	120	42	765	468	117	157
	C T	148	0	0	11	1	78	14	0	0	21	23	0	0
	T T	4	0	0	0	0	4	0	0	0	0	0	0	0
MVP Replication (Stage III)	C C	19283	231	1	2827	60	8775	228	803	37	5265	278	679	99
	C T	820	0	0	107	0	534	19	0	0	149	11	0	0
	T T	8	0	0	0	0	8	0	0	0	0	0	0	0

eTable 10. African Ancestry Cutoff Sensitivity Analyses

Three representative thresholds are shown: 15%, 45%, and 75% corresponding respectively to the threshold use in the main analysis (15%), a first generation admix individual (45%), and the traditional cut off for super ancestry assignment in SNPWeight (75%). Overall, the results are very similar for the three cutoffs and the main finding in the *APOE* ε3/ε4 remains unchanged. Some analyses are more significant or show larger effect size at 45% or 75% cutoffs or other intermediate values (data not shown) than at 15% cutoff.

Sample		AFR (≥%)	AD Case-Control Regression N	MAC	OR [95% CI]	P	AD Age-at-onset Regression N	MAC	β [95% CI]	P
Discovery	*APOE* ε2ε3	15	934	22	0.73 [0.26; 2.04]	0.55	222	4	-6.96 [-15.56; 1.64]	0.11
		45	738	20	0.9 [0.3; 2.66]	0.84	163	4	-6.84 [-15.17; 1.5]	0.11
		75	465	12	2.2 [0.53; 9.13]	0.28	95	4	-6.68 [-15.43; 2.06]	0.13
	APOE ε3ε3	15	3767	196	1.06 [0.78; 1.46]	0.71	1108	58	-1.68 [-3.87; 0.5]	0.13
		45	2449	149	1.12 [0.78; 1.61]	0.54	676	44	-0.01 [-2.63; 2.61]	0.99
		75	1487	102	1.19 [0.77; 1.84]	0.43	400	30	0.38 [-2.84; 3.6]	0.82
	APOE ε3ε4	15	2381	71	3.01 [1.87; 4.85]	**6.0E-06**	1063	51	-5.87 [-8.35; -3.4]	**3.4E-06**
		45	1845	66	3.17 [1.93; 5.2]	**4.7E-06**	833	48	-5.48 [-8.04; -2.92]	**2.7E-05**
		75	1227	53	3.4 [1.95; 5.9]	**1.5E-05**	552	39	-4.98 [-7.76; -2.19]	**4.6E-04**
Replication	*APOE* ε2ε3	15	453	12	0.78 [0.11; 5.35]	0.8	53	1	-18.42 [-39.23; 2.38]	0.08
		45	395	12	0.79 [0.11; 5.5]	0.81	23	1	-24.18 [-46.48; -1.89]	0.03
		75	304	9	0.32 [0.02; 4.92]	0.41	-	-	-	-
	APOE ε3ε3	15	1748	100	0.85 [0.48; 1.53]	0.6	347	8	-1.36 [-8.29; 5.58]	0.7
		45	1291	91	0.93 [0.47; 1.81]	0.83	108	4	-7.36 [-17.38; 2.66]	0.15
		75	923	77	0.78 [0.36; 1.69]	0.53	19	2	-12.1 [-28.68; 4.47]	0.15
	APOE ε3ε4	15	1277	44	2.2 [1.04; 4.65]	**0.04**	421	21	-5.23 [-9.58; -0.87]	**0.02**
		45	900	31	2.48 [0.95; 6.5]	**0.06**	170	10	-5.5 [-11.34; 0.33]	**0.06**
		75	632	22	2.53 [0.75; 8.5]	**0.13**	45	4	-6.62 [-15.89; 2.66]	**0.16**
Meta-analysis	*APOE* ε2ε3	15	1387	34	0.74 [0.3; 1.84]	0.52	275	5	-8.63 [-16.58; -0.69]	0.03
		45	1133	32	0.87 [0.34; 2.24]	0.77	186	5	-8.96 [-16.77; -1.15]	0.02
		75	769	21	1.45 [0.41; 5.15]	0.56	95	4	-6.68 [-15.43; 2.06]	0.13
	APOE ε3ε3	15	5515	296	1.01 [0.77; 1.34]	0.94	1455	66	-1.65 [-3.74; 0.43]	0.12
		45	3740	240	1.07 [0.78; 1.47]	0.66	784	48	-0.48 [-3.02; 2.05]	0.71
		75	2410	179	1.08 [0.74; 1.57]	0.7	419	32	-0.07 [-3.23; 3.09]	0.96
	APOE ε3ε4	15	3658	115	2.75 [1.84; 4.11]	**8.3E-07**	1484	72	-5.72 [-7.87; -3.56]	**2.0E-07**
		45	2745	97	3.01 [1.94; 4.68]	**8.9E-07**	1003	58	-5.49 [-7.83; -3.14]	**4.4E-06**
		75	1859	75	3.23 [1.95; 5.34]	**5.0E-06**	597	43	-5.11 [-7.78; -2.45]	**1.7E-04**

eTable 11. African Ancestry Cutoff Sensitivity Analyses Restricted to Individuals With African Local Ancestry at Both of Their *APOE* Haplotypes

		AD Case-Control Regression				AD Age-at-onset Regression			
Sample	AFR (≥%)	N	MAC	OR [95% CI]	P	N	MAC	β [95% CI]	P
APOE ε3ε4	15			3.10				-5.11	
		978	44	[1.68; 5.7]	2.8E-04	434	32	[-8.19; -2.03]	1.1E-03
	45			3.16				-5.10	
		917	44	[1.72; 5.82]	2.2E-04	409	32	[-8.24; -1.96]	1.4E-03
	75			3.1				-4.78	
		719	36	[1.57; 6.13]	1.1E-03	321	26	[-8.25; -1.31]	7.0E-03

eTable 12. Mega-analysis of the Discovery and Replication Samples

Sample		AFR (≥%)	AD Case-Control Regression N	MAC	OR [95% CI]	P	AD Age-at-onset Regression N	MAC	β [95% CI]	P
Mega-analysis	*APOE* ε2ε3	15	1387	34	0.76 [0.31; 1.87]	0.55	275	5	-9.07 [-17.08; -1.06]	0.03
		45	1133	32	0.87 [0.34; 2.2]	0.76	186	5	-8.97 [-16.69; -1.26]	0.02
		75	769	21	1.51 [0.43; 5.26]	0.52	100	4	-7.21 [-16.02; 1.6]	0.11
	APOE ε3ε3	15	5515	296	0.98 [0.75; 1.28]	0.86	1455	66	-1.36 [-3.5; 0.77]	0.21
		45	3740	240	1.05 [0.77; 1.44]	0.74	784	48	-0.32 [-2.88; 2.24]	0.81
		75	2410	179	1.04 [0.72; 1.49]	0.84	419	32	-0.06 [-3.2; 3.09]	0.97
	APOE ε3ε4	15	**3658**	**115**	**2.93 [1.99; 4.31]**	**4.8E-08**	**1484**	**72**	**-5.86 [-8.05; -3.66]**	**1.7E-07**
		45	**2745**	**97**	**3.02 [1.97; 4.61]**	**3.6E-07**	**1003**	**58**	**-5.7 [-8.06; -3.33]**	**2.4E-06**
		75	**1859**	**75**	**3.25 [1.99; 5.3]**	**2.4E-06**	**597**	**43**	**-5.35 [-8.01; -2.69]**	**8.0E-05**

eTable 13. African Ancestry Cutoff Sensitivity Analyses Only Including Individuals Directly Genotyped in the ADSP WES Data Distributed by NIAGADS

Note that compared to **eTable 10**, solely the discovery sample changed by removing the 980 individuals who had any of the 3 *APOE* missenses variants missing in the NIAGADS call of ADSP WES. (**eFigure 2**). As in **eTable 10,** three representative thresholds are shown: 15%, 45%, and 75% corresponding respectively to the threshold use in the main analysis (15%), a first generation admix individual (45%), and the traditional cut off for super ancestry assignment in SNPWeight (75%). Overall, the results are very similar for the three cutoffs and the main finding in ε3/ε4 remains unchanged. Some analyses are more significant or show larger effect size at 45% or 75% cutoffs or other intermediate values (data not shown) than at 15% cutoff.

Sample		AFR (≥%)	AD Case-Control Regression N	MAC	OR [95% CI]	P	AD Age-at-onset Regression N	MAC	β [95% CI]	P
	APOE ε2ε3	15	812	20	0.84 [0.28; 2.46]	0.75	196	4	-6.35 [-14.72; 2.01]	0.14
		45	651	18	1.02 [0.33; 3.22]	0.97	147	4	-6.26 [-14.54; 2.01]	0.14
		75	415	12	2.17 [0.52; 9.06]	0.29	87	4	-6.31 [-15.03; 2.41]	0.16
	APOE ε3ε3	15	3247	175	1.05 [0.75; 1.46]	0.79	983	53	-1.56 [-3.87; 0.75]	0.19
		45	2133	132	1.19 [0.81; 1.73]	0.38	616	42	-0.31 [-2.97; 2.34]	0.82
		75	1328	92	1.26 [0.8; 1.99]	0.31	374	29	0.28 [-2.99; 3.54]	0.87
	APOE ε3ε4	15	2103	65	3.1 [1.88; 5.1]	**9.2E-06**	989	49	-5.65 [-8.2; -3.09]	**1.4E-05**
		45	1663	63	3.23 [1.95; 5.36]	**5.4E-06**	798	48	-5.24 [-7.8; -2.68]	**6.1E-05**
		75	1128	52	3.17 [1.82; 5.55]	**5.0E-05**	540	39	-4.82 [-7.6; -2.04]	**6.9E-04**
Replication	*APOE* ε2ε3	15	453	12	0.78 [0.11; 5.35]	0.8	53	1	-18.42 [-39.23; 2.38]	0.08
		45	395	12	0.79 [0.11; 5.5]	0.81	23	1	-24.18 [-46.48; -1.89]	0.03
		75	304	9	0.32 [0.02; 4.92]	0.41	-	-	-	-
	APOE ε3ε3	15	1748	100	0.85 [0.48; 1.53]	0.6	347	8	-1.36 [-8.29; 5.58]	0.7
		45	1291	91	0.93 [0.47; 1.81]	0.83	108	4	-7.36 [-17.38; 2.66]	0.15
		75	923	77	0.78 [0.36; 1.69]	0.53	19	2	-12.1 [-28.68; 4.47]	0.15
	APOE ε3ε4	15	1277	44	2.2 [1.04; 4.65]	**0.04**	421	21	-5.23 [-9.58; -0.87]	**0.02**
		45	900	31	2.48 [0.95; 6.5]	**0.06**	170	10	-5.5 [-11.34; 0.33]	**0.06**
		75	632	22	2.53 [0.75; 8.5]	**0.13**	45	4	-6.62 [-15.89; 2.66]	**0.16**
Meta-analysis	*APOE* ε2ε3	15	1265	32	0.82 [0.32; 2.11]	0.68	249	5	-8.03 [-15.79; -0.27]	0.04
		45	1046	30	0.96 [0.36; 2.57]	0.93	170	5	-8.43 [-16.19; -0.67]	0.03
		75	719	21	1.44 [0.41; 5.11]	0.57	87	4	-6.31 [-15.03; 2.41]	0.16
	APOE ε3ε3	15	4995	275	1.0 [0.75; 1.33]	0.98	1330	61	-1.54 [-3.73; 0.65]	0.17
		45	3424	223	1.12 [0.8; 1.55]	0.51	724	46	-0.78 [-3.35; 1.79]	0.55
		75	2251	169	1.12 [0.75; 1.65]	0.58	393	31	-0.19 [-3.4; 3.02]	0.91
	APOE ε3ε4	15	3380	109	2.79 [1.84; 4.22]	**1.3E-06**	1410	70	-5.54 [-7.74; -3.34]	**8.2E-07**
		45	2563	94	3.05 [1.95; 4.78]	**1.0E-06**	968	58	-5.28 [-7.63; -2.94]	**1.0E-05**
		75	1760	74	3.05 [1.84; 5.06]	**1.6E-05**	585	43	-4.96 [-7.63; -2.3]	**2.6E-04**

eTable 14. Primary and Secondary Analyses Considering a Standard Model (Non-stratified by *APOE* Genotype) and Adjusting for ε2 and ε4 Dosages

		AD Case-Control Regression				AD Age-at-onset Regression			
Sample	AFR (≥%)	N	MAC	OR [95% CI]	P	N	MAC	β [95% CI]	P
Discovery	15	7845	289	1.40 [1.08; 1.8]	0.01	2802	113	-3.55 [-5.15; -1.95]	1.4E-05
	45	5684	235	1.56 [1.18; 2.07]	2.0E-03	2030	96	-3.01 [-4.8; -1.21]	1.0E-03
	75	3633	167	1.8 [1.29; 2.51]	5.1E-04	1298	73	-2.81 [-4.82; -0.79]	6.3E-03
Replication	15	3945	156	1.23 [0.78; 1.94]	0.38	996	30	-4.42 [-7.98; -0.86]	0.01
	45	2928	134	1.28 [0.75; 2.17]	0.37	378	15	-6.80 [-11.76; -1.83]	7.3E-03
	75	2086	108	1.03 [0.55; 1.94]	0.93	90	6	-8.54 [-15.72; -1.35]	0.02
Meta-analysis	15	11790	445	1.36 [1.08; 1.69]	7.5E-03	3798	143	-3.70 [-5.16; -2.24]	7.0E-07
	45	8612	369	1.49 [1.16; 1.91]	1.6E-03	2408	111	-3.44 [-5.13; -1.76]	6.3E-05
	75	5719	275	1.6 [1.19; 2.14]	1.8E-03	1388	79	-3.23 [-5.17; -1.28]	1.1E-03

eTable 15. Interaction Between R145C*ε4 for Association With AD Risk, Considering a Standard Model (Non-stratified by *APOE* Genotype) and Adjusting for ε2 and ε4 Dosages

Sample	AFR (≥%)	AD Case-Control Regression			
		N	MAC	OR – R145C*ε4 [95% CI]	P
Discovery	15	7845	289	2.73 [1.57; 4.75]	3.6E-04
	45	5684	235	2.58 [1.43; 4.67]	1.7E-03
	75	3633	167	2.38 [1.20; 4.70]	0.01
Replication	15	3945	156	2.51 [1.08; 5.85]	0.03
	45	2928	134	2.71 [0.91; 8.05]	0.07
	75	2086	108	3.16 [0.84; 11.87]	0.09
Meta-analysis	15	11790	445	2.66 [1.68; 4.23]	3.4E-05
	45	8612	369	2.61 [1.55; 4.40]	3.1E-04
	75	5719	275	2.52 [1.38; 4.63]	2.8E-03
MVP	-	20111	828	0.99 [0.40; 2.48]	0.98

etable 16. Risk of Alzheimer disease by *APOE* Genotype Including *APOE* ε3[R145C] Subtypes

This table corresponds to values reported in **Figure 2**.

Stage I - Discovery				Stage II - Replication			
APOE	**N**	**OR [95% CI]**	**P**	***APOE***	**N**	**OR [95% CI]**	**P**
ε2/ε2	52	0.63 [0.34; 1.16]	0.14	ε2/ε2	31	1.6 [0.51; 5.02]	0.42
ε2/ε3	912	0.75 [0.63; 0.88]	5.3E-04	ε2/ε3	441	0.86 [0.64; 1.17]	0.34
ε2/ε4	290	1.52 [1.16; 1.99]	2.1E-03	ε2/ε4	162	1.45 [0.93; 2.26]	0.1
ε3/ε3	3576	1.00	/	ε3/ε3	1652	1.00	/
ε3/ε3[R145C]	2310	1.92 [1.71; 2.14]	8.1E-30	ε3/ε3[R145C]	1233	1.95 [1.62; 2.36]	3.6E-12
ε3/ε4	71	7.88 [4.69; 13.23]	6.1E-15	ε3/ε4	44	4.87 [2.25; 10.53]	5.8E-05
ε3[R145C]/ε4	191	1.05 [0.76; 1.46]	0.76	ε3[R145C]/ε4	96	0.87 [0.47; 1.6]	0.66
ε4/ε4	421	6.3 [5.07; 7.83]	8.5E-62	ε4/ε4	274	4.35 [3.18; 5.96]	5.4E-20

eTable 17. Association of R145C With CDR-SB Change Over Time in ε3/ε4 Stratified Analysis

	Estimate	Std.Error	t-value	p-val	-log$_{10}$(p-val)
(Intercept)	-0.40	0.89	-0.45	0.65	0.18
CDR-SB at baseline	1.15	0.02	47.35	0	488.60
Age at baseline	-0.38	0.02	-20.91	4.31E-97	96.37
Age	0.38	0.02	22.94	1.68E-116	115.77
R145C	-11.30	3.40	-3.32	8.87E-04	3.05
Sex (female)	-0.29	0.18	-1.59	0.11	0.95
Years of education	0.05	0.02	2.33	0.02	1.70
Age*R145C	0.19	0.05	3.87	1.11E-04	3.95

eTable 18. Association of R145C*ε4 With CDR-SB Change Over Time in Unstratified Analysis

	Estimate	Std.Error	t-value	p-val	-$\log_{10}$(p-val)
(Intercept)	3.02	0.48	6.33	2.50E-10	9.60
CDR-SB at baseline	1.14	0.01	86.30	0	1619.35
Age at baseline	-0.22	0.01	-26.07	8.22E-150	149.08
Age	0.17	0.01	22.58	6.82E-113	112.17
R145C	-0.70	1.78	-0.40	0.69	0.16
APOE2 dosage	-0.17	0.10	-1.73	0.08	1.08
APOE4 dosage	-5.34	0.48	-11.17	5.53E-29	28.26
Sex (female)	-0.18	0.09	-2.03	0.04	1.37
Years of education	0.00	0.01	-0.41	0.68	0.17
R145C*APOE4	-9.90	3.07	-3.22	1.28E-03	2.89
Age*APOE4	0.08	0.01	11.73	8.75E-32	31.06
Age*R145C	0.01	0.02	0.42	0.67	0.17
Age*R145C*APOE4	0.17	0.04	3.86	1.15E-04	3.94

eReferences

1. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015;4(1):7. doi:10.1186/s13742-015-0047-8

2. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7

3. Le Guen Y, Belloy ME, Napolioni V, et al. A novel age-informed approach for genetic association analysis in Alzheimer's disease. *Alzheimer's Research & Therapy*. 2021;13(1):72. doi:10.1186/s13195-021-00808-5

4. Le Guen Y, Napolioni V, Belloy ME, et al. Common X-Chromosome Variants Are Associated with Parkinson Disease Risk. *Annals of Neurology*. 2021;90(1):22-34. doi:10.1002/ana.26051

5. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26(22):2867-2873. doi:10.1093/bioinformatics/btq559

6. Sherva R, Zhang R, Sahelijo N, et al. African Ancestry GWAS of Dementia in a Large Military Cohort Identifies Significant Risk Loci. Published online May 26, 2022:2022.05.25.22275553. doi:10.1101/2022.05.25.22275553

7. Seripa D, Matera MG, Daniele A, et al. The Missing ApoE Allele. *Annals of Human Genetics*. 2007;71(4):496-500. doi:10.1111/j.1469-1809.2006.00344.x

8. Bis JC, Jian X, Kunkle BW, et al. Whole exome sequencing study identifies novel rare and common Alzheimer's-Associated variants involved in immune response and transcriptional regulation. *Molecular Psychiatry*. 2020;25(8):1859-1875. doi:10.1038/s41380-018-0112-7

9. Zhu C, Farrell J, Kuzma A, Schellenberg G, Farrer L. Precision genotyping of APOE from whole genome sequencing. In: *American Society of Human Genetics Congress*. ; 2019.

10. Kunkle BW, 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. *Nature Genetics*. 2019;51(3):414-430. doi:10.1038/s41588-019-0358-2

11. Kunkle BW, Schmidt M, Klein HU, et al. Novel Alzheimer Disease Risk Loci and Pathways in African American Individuals Using the African Genome Resources Panel: A Meta-analysis. *JAMA Neurol*. 2021;78(1):102. doi:10.1001/jamaneurol.2020.3536

12. Loh PR, Danecek P, Palamara PF, et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*. 2016;48(11):1443-1448. doi:10.1038/ng.3679

13. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: A Discriminative Modeling Approach for Rapid and Robust Local-Ancestry Inference. *The American Journal of Human Genetics*. 2013;93(2):278-288. doi:10.1016/j.ajhg.2013.06.020

14. Byrska-Bishop M, Evani US, Zhao X, et al. *High Coverage Whole Genome Sequencing of the Expanded 1000 Genomes Project Cohort Including 602 Trios*.; 2021:2021.02.06.430068. doi:10.1101/2021.02.06.430068

15. Bergström A, McCarthy SA, Hui R, et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science*. 2020;367(6484):eaay5012. doi:10.1126/science.aay5012

16. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*. 2006;38(8):904-909. doi:10.1038/ng1847