Meta-analysis of genetic association with diagnosed Alzheimer's disease identifies novel risk loci and implicates Abeta, Tau, immunity and lipid processing – Supplemental Note

INTERNATIONAL GENOMICS OF ALZHEIMERS PROJECT (IGAP) STAGE 1 DISCOVERY DATASET DESCRIPTIONS

IGAP is composed of datasets from the Alzheimer Disease Genetics Consortium (ADGC), Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE), Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES). Written informed consent was obtained from study participants or, for those with substantial cognitive impairment, from a caregiver, legal guardian, or other proxy, and the study protocols for all populations were reviewed and approved by the appropriate Institutional review boards (IRB's). Descriptions of cohorts within ADGC, CHARGE, EADI and GERAD/PERADES are provided below.

Alzheimer's Disease Genetics Consortium (ADGC)

The ADGC dataset comprises subjects from the Adult Changes in Thought (ACT)/ Electronic Medical Records and Genetics (eMERGE) Study, the National Institute on Aging (NIA) Alzheimer Disease Centers (ADCs), the Alzheimer Disease Neuroimaging Initiative (ADNI) Study, the Biomarkers of Cognitive Decline Among Normal Individuals: The BIOCARD Cohort (BIOCARD), the Chicago Health and Aging Project (CHAP), the Einstein Aging Study (EAS), the Multi-Site Collaborative Study for Genotype-Phenotype Associations in Alzheimers Disease (GenADA) Study, the Mayo Clinic (MAYO), the MIRAGE Study (MIRAGE), the NIA-LOAD Study (NIA-LOAD), the Netherlands Brain Bank (NBB), Oregon Health and Science University (OHSU), Pfizer, Rochester Mayo Clinic (RMAYO), the Rush University Religious Orders Study/Memory and Aging Project (ROSMAP), the Texas Alzheimer's Research and Care Consortium (TARCC), the Translational Genomics Research Institute series 2 (TGEN2) dataset, the University of Miami/Vanderbilt University University/Mt. Sinai School of Medicine (UM/VU/MSSM), Universitatsklinikum Saarlandes (UKS), the University of Pittsburgh (UPITT), Washington University (WASHU), and the Washington Heights-Inwood Community Aging Project (WHICAP). Descriptions of the ACT/eMERGE, ADC waves 1-6, ADNI, CHAP, GenADA, MAYO, MIRAGE, NIA-LOAD, NBB, OHSU, ROSMAP, TARCC, TGEN2, UM/VU/MSSM, UPITT, WASHU and WHICAP cohorts have been provided in previous ADGC and IGAP studies[1–5]. Here we update descriptions of these studies, where applicable, and provide descriptions for ADC wave 7, BIOCARD, EAS, MTV (comprised of subjects from UM, TARCC and VU), RMAYO, ROSMAP wave 2, WASHU wave 2 and UKS. All analyses were restricted to individuals of European ancestry. All subjects were recruited under protocols approved by the appropriate Institutional Review Boards (IRB's).

The ACT/eMERGE Studies (ACT): The ACT cohort is an urban and suburban elderly population from a stable HMO that includes 2,581 cognitively intact subjects age ≥ 65 who were enrolled between 1994 and 1998[6,7]. An additional 811 subjects were enrolled in 2000-2002 using the same methods except oversampling clinics with more minorities. More recently, a Continuous Enrollment strategy was initiated in which new subjects are contacted, screened and enrolled to keep 2000 active at-risk person-years accruing in each calendar year. This resulted in an enrollment of 4,146 participants as of May 2009. All clinical data are reviewed at a consensus conference. Dementia onset is assigned half way between the prior biennial and the exam that diagnosed dementia. Enrollment for eMERGE Study began in 2007. A waiver of consent was obtained from the IRB to enroll deceased ACT participants. In total, ACT/eMERGE contributed data on 532 individuals with probable or possible AD (70 with autopsy-confirmation) and on 1,571 CNEs (155 with autopsy-confirmation) who were included in the analyses.

*The NIA ADC Samples (ADC)***:** The NIA ADC cohort included subjects ascertained and evaluated by the clinical and neuropathology cores of the 32 NIA-funded ADCs. Data collection is coordinated by the National Alzheimer's Coordinating Center (NACC). NACC coordinates collection of phenotype data from the 32 ADCs, cleans all data, coordinates implementation of definitions of AD cases and controls, and coordinates collection of samples. 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. 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 NINCDS-ADRDA/DSMIV-V criteria[8,9] or Clinical Dementia Rating (CDR) ≥ 1[10]. Neuropathologic stratification of cases followed NIA/Reagan criteria explicitly or used a similar approach when NIA/Reagan criteria[4] were coded as not done, missing, or unknown. Cases were intermediate or high likelihood by NIA/Reagan criteria with moderate to frequent amyloid plaques[11] and neurofibrillary tangle (NFT) Braak stage of III-VI[12,13]. Persons with Down's syndrome, non-AD tauopathies and synucleinopathies were excluded. All autopsied controls had a clinical evaluation within two years of death. Controls did not meet NINCDS-ADRDA/DSMIV-V 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. ADCs sent frozen tissue from autopsied subjects and DNA samples from some autopsied subjects and from living subjects to the ADCs 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. ADC samples were genotyped and analyzed in separate batches (waves 1-7).

The ADC cohort used in the discovery stage (ADC waves 1-6) consists of 2,349 autopsy-confirmed and 913 clinically-confirmed AD cases, and 542 cognitively normal elders (CNEs) with complete neuropathology data who were older than 60 years at age of death, and 744 living CNEs

evaluated using the Uniform dataset (UDS) protocol[14,15] who were documented to not have mild cognitive impairment (MCI) and were between 60 and 100 years of age at assessment. ADC1 and ADC2 contributed 2,304 AD cases (1,761 autopsy-confirmed; 543 clinically-confirmed) and 675 CNEs (515 autopsy-confirmed; 160 clinically-confirmed). The ADC3 dataset contains 897 clinically-identified living cases (527 with autopsy-confirmation) and 588 CNEs (4 with autopsy-confirmation) who were genotyped between July and August 2010. Diagnoses for 99 AD cases and 32 CNEs, 115 AD cases and 45 CNEs, 59 AD cases and 12 CNEs in ADC waves 4-6 were confirmed pathologically, respectively.

The ADNI Study (ADNI): ADNI is a longitudinal, multi-site observational study including AD, mild cognitive impairment (MCI), and elderly individuals with normal cognition assessing clinical and cognitive measures, MRI and PET scans (FDG and 11C PIB) and blood and CNS biomarkers. For this study, ADNI contributed data on 268 AD cases with MRI confirmation of AD diagnosis and 173 healthy controls with AD-free status confirmed as of most recent follow-up. AD subjects were between the ages of 55–90, had an MMSE score of 20–26 inclusive, met NINCDS/ADRDA criteria for probable AD[8,9], and had an MRI consistent with the diagnosis of AD. Control subjects had MMSE scores between 28 and 30 and a Clinical Dementia Rating of 0 without symptoms of depression, MCI or other dementia and no current use of psychoactive medications. According to the ADNI protocol, subjects were ascertained at regular intervals over 3 years, but for the purpose of our analysis we only used the final ascertainment status to classify case-control status. Additional details of the study design are available elsewhere[5,16,17].

The BIOCARD Cohort (BIOCARD): BIOCARD is supported by a grant jointly funded by the National Institute on Aging (NIA) and the National Institute of Mental Health (NIMH). The overarching goal of the BIOCARD Study is to identify biomarkers associated with progression from normal cognitive status to cognitive impairment or dementia, with a particular focus on Alzheimer's Disease. Please see Albert et al. 2014[18] for a detailed description of the study. Briefly, A total of 354 individuals were initially enrolled in the study. Recruitment was conducted by the staff of the Geriatric Psychiatry Branch (GPB) of the intramural program of the NIMH, beginning in 1995 and ending in 2005. The domains of information collected as part of the study include: cognitive testing, magnetic resonance imaging (MRI), cerebrospinal fluid (CSF), amyloid imaging (using PET-PiB), and blood specimens. Investigators at the Johns Hopkins University School of Medicine began evaluating participants in 2009, and subjects are seen annually. At each visit there are assessments of medical and cognitive status, as well as acquisition of MRI, CSF, PET-PiB, and blood. Each subject in the analyses received a consensus diagnosis by a team of neurologists, neuropsychologists, research nurses and research assistants of the BIOCARD Clinical Core at Johns Hopkins with diagnoses based on evidence of clinical or cognitive dysfunction (i.e., individuals with a CDR score > 0 and/or evidence of decline on cognitive testing). To the extent possible, this diagnosis did not use the cognitive test scores. In brief, (1) clinical data relating to the medical, neurologic and psychiatric status of the subject were examined, (2) reports of changes in cognition by the subject and

other sources were examined, and (3) decline in cognitive performance was established. Cognitive test scores were used to: (1) determine whether the subject had become cognitively impaired, and (2) determine the likely etiology of such impairment. These diagnostic procedures are comparable to those implemented in the Alzheimer's Disease Centers (ADC) program, supported by the NIA.

The CHAP Project (CHAP): Chicago Health and Aging Project (CHAP): CHAP is an on-going community based study of individuals from a geographically defined community of 3 neighborhoods in Chicago, Illinois (Morgan Park, Washington Heights, and Beverly), with 6,158 participants in the first phase of the study (78.7% overall; 80.5% of the blacks, 74.6% of the whites)[19]. Data were collected in cycles of approximately 3 years; each consisting of an in-home interview of all participants and clinical evaluation of a random, stratified sample. The baseline cycle measured disease prevalence and provided risk factor data prior to incident disease onset. A cohort of 3,838 persons free of AD was identified; 729 persons were sampled for baseline clinical evaluation. Persons in the disease-free cohort had either good cognitive function at baseline, or if cognitive function was intermediate or poor, were free from AD at the baseline clinical evaluation. This disease-free cohort was evaluated for incident disease after an average of 4.1 years. Sampling for incident clinical evaluation was based on age, sex, race, and change in cognitive function (i.e., stable or improved, small decline, or large decline). The sample set available in the ADGC for genetic analyses included 32 AD cases and 197 persons free of AD at time of last assessment. All subjects were age 65 years or older at last assessment.

The Einstein Aging Study (EAS): EAS, based at the Albert Einstein College of Medicine, is an ongoing community based cohort study of cognitive aging and Alzheimer's disease in the elderly which began over four decades ago. Please see Barzilai et al. 2004[20] and Katz et al. 2012[21] for details. Briefly, the EAS cohort has employed systematic recruiting methods to reduce the selection biases that arise from clinic-based samples and to capture the racial diversity within the Bronx community. Since 1993, a total of 1944 participants have been enrolled. Between 1993 and 2004, Health Care Financing Administration/ Centers for Medicaid and Medicare Services (HCFA/CMS) rosters of Medicare eligible persons aged 70 and above were used to develop sampling frames of community residing participants in Bronx County. Since 2004, New York City Board of Elections registered voter lists for the Bronx have been used due to changes in policies for release of HCFA/CMS rosters. Individuals were mailed introductory letters regarding the study and were then telephoned to complete a brief screening interview. Eligible participants were at least 70 years of age, Bronx residents, non-institutionalized, and English speaking. Exclusion criteria included visual or auditory impairments that preclude neuropsychological testing, active psychiatric symptomatology that interfered with the ability to complete assessments, and non-ambulatory status. Written informed consent was obtained at the initial clinic visit. In-person evaluations were completed at baseline and at subsequent 12-month intervals. Functional status was assessed by the self-administered CERAD C1-ALT, a cognitive/functional impairment instrument, and the Instrumental

Activities of Daily Living scale(IADL), a subscale on the Lawton Brody Activities of Daily Living Scale. The score on the IADL was based on 5 domains of function that were common to both elderly men and women. Scores for each domain were dichotomized as impaired vs. not impaired and then the domain scores were summed. If the participant agreed, an informant completed the CERAD C2-ALT, a cognitive/functional impairment instrument, and the Informant Questionnaire on Cognitive Decline in the Elderly (IQ-CODE)14 forms. The standard neurological physical examination was adapted from the Unified Parkinson's Disease Rating Scale. The evaluation assessed the participant's memory for significant recent events in the news and personal events. The coherence and focus of responses, repetitiveness, and language were determined. When possible, informants were interviewed to ascertain whether they noted any cognitive changes in the participant, and to assess accuracy of the participant's responses. The neurologist also assessed each participant for abnormal behaviors, fluctuation in cognition, and history of sleep disturbance and visual/auditory hallucinations. The neurologist assigned an Hachinski Ischemic Score (HIS), the Clinical Dementia Rating (CDR), and provided a clinical impression of presence or absence of dementia. A diagnosis of dementia was based on standardized clinical criteria from the Diagnostic and Statistical Manual, Fourth Edition (DSM-IV) and required impairment in memory plus at least one additional cognitive domain, accompanied by evidence of functional decline. Diagnoses were assigned at consensus case conferences, which included comprehensive review of cognitive test results, relevant neurological signs and symptoms, and functional status. Memory impairment was defined as scores in the impaired range on any of the memory tests in the neuropsychological battery. (FCSRT ≤ 2430 or 1.5 standard deviations below the age-adjusted mean on Logical Memory) Functional decline was determined at case conference based on information from self or informant report, impairment score on the IADL Lawton Brody Scale, clinical evaluation, and informant questionnaires. AD was diagnosed in participants with dementia meeting clinical criteria for probable or possible disease established by the National Institute of Neurological and Communication Disorders and Stroke and the Alzheimer Disease and Related Disorders Association (NINCDS-ADRDA)[8,9]. Incident dementia and AD were diagnosed in persons free of dementia at baseline who met criteria at follow-up. A subset of individuals who participated in the clinical studies of the EAS came to autopsy, providing an important quality control for diagnostic accuracy. A clinical diagnosis of dementia had a positive predictive value (PPV) of 96% for significant pathology upon autopsy. A clinical diagnosis of possible or probable AD had a PPV of 79% for the presence of NIA-Reagan intermediate or high likelihood Alzheimer type pathology based on an autopsy sample of 175.

The GenADA Study (GenADA): Data from the GenADA cohort that were analyzed included 666 AD cases and 712 CNEs ascertained from nine memory referral clinics in Canada between 2002 and 2005. Patients and CNEs were of non-Hispanic White ancestry from Northern Europe. All patients with AD satisfied NINCDS-ADRDA and DSM-IV criteria for probable AD with Global Deterioration Scale scores of 3-7. CNEs had MMSE test scores higher than 25 (mean 29.2 ± 1.1), a Mattis Dementia Rating Scale score of

≥ 136, a Clock Test without error, and no impairments on seven instrumental activities of daily living questions from the Duke Older American Resources and Services Procedures test. Data were collected under an academic-industrial grant from Glaxo-Smith-Kline, Canada by Principal Investigator P. St George-Hyslop. Detailed characteristics of this cohort have been described previously[22].

Mayo Clinic (MAYO and RMAYO): All 671 cases and 1,279 controls consisted of non-Hispanic White 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[8,9]. Autopsy-confirmed samples (221 cases, 216 CNEs) 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. All AD brains analyzed in the study had a Braak score of 4.0 or greater. Brains employed as controls had a Braak score of 2.5 or lower but often had brain pathology unrelated to AD and pathological diagnoses that included vascular dementia, frontotemporal dementia, dementia with Lewy bodies, multi-system atrophy, amyotrophic lateral sclerosis, and progressive supranuclear palsy.

The MIRAGE Study (MIRAGE): The MIRAGE study is a family-based genetic epidemiological study of AD that enrolled AD cases and unaffected sibling controls at 17 clinical centers in the United States, Canada, Germany, and Greece (details elsewhere[22]), and contributed 1,262 subjects (491 AD cases and 738 CNEs), a subset of the cases and controls that were incorporated into our prior studies[1,5] which met more stringent QC criteria for this study. Briefly, families were ascertained through a proband meeting the NINCDS-ADRDA criteria for definite or probable AD[8,9]. Unaffected sibling controls were verified as cognitively healthy based on a Modified Telephone Interview of Cognitive Status score ≥ 86[23].

University of Miami/ Texas Alzheimer's Research Care Consortium Wave 2/Case Western Reserve University (MTC): The MTC sample included 256 cases and 189 controls from the University of Miami, the Texas Alzheimer's Research Care Consortium, and Case Western Reserve University. All AD cases had onset of disease symptoms after age 65 years and met NINCDS-ADRDA criteria for probable or possible AD[8,9]. Controls were adjudicated to have MMSE scores greater than 28 and no clinically identified signs of cognitive impairment. Additional details of subject recruitment at these sites are described in the UM/CWR/MSSM (formerly UM/VU/MSSM) and TARRC cohort descriptions in this supplement and elsewhere[1,3,4].

The NIA LOAD Family Study (NIA-LOAD): The NIA LOAD Family Study[24] recruited families with two or more affected siblings with LOAD and unrelated, CNEs similar in age and ethnic background. A total of 1,819 cases and 1,969 CNEs from 1,802 families were recruited through the NIA-LOAD study, NCRAD,

and the University of Kentucky, with 1,798 cases and 1,568 CNEs included for analysis. One case per family was selected after determining the individual with the strictest diagnosis (definite > probable > possible LOAD). If there were multiple individuals with the strictest diagnosis, then the individual with the earliest age of onset was selected. The controls included only those samples that were neurologically evaluated to be normal and were not related to a study participant.

The Netherlands Brain Bank (NBB): The NBB, which has been previously described[3], is a department of the Netherlands Institute for Neuroscience, an institute of the Royal Netherlands Academy of Arts and Sciences. The NBB is a non-profit organization that collects human brain tissue from donors with a variety of neurological and psychiatric disorders and brain tissue from non-diseased donors, as well as anonymized summaries of donors' medical records to be made available for neuroscience research[25]. The sample set available in the ADGC for genetic analyses included 80 pathologically-confirmed AD cases and 48 subjects free of Alzheimer's pathology at autopsy. All cases were age 65 years or older at time of diagnosis, and all controls were age 65 years or older at time of death.

Oregon Health and Science University (OHSU): The OHSU dataset includes 132 autopsy-confirmed AD cases and 153 deceased controls that were evaluated for dementia within 12 months prior to death (age at death > 65 years), which are a subset of the 193 cases and 451 controls examined in our previous study[1] meeting more stringent QC criteria in this study. Subjects were recruited from aging research cohorts at 10 NIA-funded ADCs, and did not overlap other samples assembled by the ADGC. A more extensive description of control samples can be found elsewhere[8].

Pfizer: The Pfizer sample collection comprises AD cases taken from the Lipitor's Effect in Alzheimer's Disease (LEADe) trial, including subjects who converted to AD after ascertainment as MCI, as well as 216 probable AD subjects enrolled by PrecisionMed for a case-control study and 149 subjects from a Phase II trial (#A3041005) of CP-457920 (a selective α5 GABAA receptor inverse agonist) in AD. Samples were collected from multiple clinical sites, and with appropriate IRB/ethics committee approvals at each individual site, with written and informed consent given by subjects for use in follow-up studies. All subjects were diagnosed with probable or possible AD if they met NINCDS[8,9] and/or DSM-IV criteria, and had mini-mental status exam (MMSE) scores < 25 at baseline. The control group included subjects from two studies: 1) the PrecisionMed case-control study (#A9010012), which recruited elderly subjects free of neurological or psychiatric conditions, and 2) 999-GEN-0583-001, which obtained a reference population of cognitively, neurologically and psychiatrically normal subjects. Controls have no neuropsychiatric conditions or diseases, and had MMSE>27 at the time of enrollment. For AD analysis, all cases with age-at-onset (AAO) less than 65 years were removed to exclude early-onset AD subjects. All controls were re-matched with remaining cases according to gender, age (all controls are older than cases), and ethnicity (only individuals with non-Hispanic White background were analyzed). The final Pfizer AD case-control

GWAS dataset included 696 cases and 762 controls. Cases from the PrecisionMed/A3041005 and LEADe studies and age-matched controls were genotyped using the Illumina HumanHap550 array. APOE genotypes were determined from genotypes for rs429358 and rs7412 obtained using Taqman assays.

The ROS/MAP Studies: ROS/MAP are two community-based cohort studies. The ROS has been on-going since 1993, with a rolling admission. Through July of 2010, 1,139 older nuns, priests, and brothers from across the United States initially free of dementia who agreed to annual clinical evaluation and brain donation at the time of death completed their baseline evaluation. The MAP has been on-going since 1997, also with a rolling admission. Through July of 2010, 1,356 older persons from across northeastern Illinois initially free of dementia who agreed to annual clinical evaluation and organ donation at the time of death completed their baseline evaluation. Details of the clinical and neuropathologic evaluations have been previously reported[26–28]. A total of 1,064 persons passed genotyping QC. Of these, 295 met clinical criteria for AD at the time of their last clinical evaluation or time of death and met neuropathologic criteria for AD for those on whom neuropathologic data were available, and 769 were without dementia or MCI at the time of their last clinical evaluation or time of death and did not meet neuropathologic criteria for AD for those on whom neuropathologic data were available. A second wave of ROSMAP (referred to as ROSMAP2 in this study included 59 persons who met clinical criteria for AD at the time of their last clinical evaluation or time of death and met neuropathologic criteria for AD for those on whom neuropathologic data were available, and 217 persons who were without dementia or MCI at the time of their last clinical evaluation or time of death and did not meet neuropathologic criteria for AD for those on whom neuropathologic data were available.

Texas Alzheimer's Research Care Consortium (TARCC): The TARCC is a collaborative Alzheimer's research effort directed and funded by the Texas Council on Alzheimer's Disease and Related Disorders (the Council), as part of the Darrell K Royal Texas Alzheimer's Initiative. Composed of Baylor College of Medicine (BCM), Texas Tech University Health Sciences Center (TTUHSC), University of North Texas Health Science Center (UNTHSC), the UT Southwestern Medical Center at Dallas (UTSW), University of Texas Health Science Center at San Antonio (UTHSCSA), Texas A&M Health Science Center (TAMHSC), and the University of Texas at Austin (UTA), this consortium was created to establish a comprehensive research cohort of well characterized subjects to address better diagnosis, treatment, and ultimately prevention of AD25. The resulting prospective cohort, the Texas Harris Alzheimer's Research Study, contains clinical, neuropsychiatric, genetic, and blood biomarker data on more than 3,000 participants diagnosed with Alzheimer's disease (AD), mild cognitive impairment (MCI), and cognitively normal individuals. Longitudinal data/sample collection and follow-up on participants occurs on an annual basis. Two waves of case-control data from TARCC were examined as part of genetic analyses in the ADGC. Data from the TARCC included 323 cases and 181 controls in the first wave, with 84 cases and

115 controls in the second wave. All TARCC subjects were greater than 65 years of age at disease onset (cases) or at last disease-free exam (non-cases).

The TGEN2 Study: Among the TGEN2 data analyzed were 864 clinically- and neuropathologically-characterized brain donors, and 493 CNEs without dementia or significant AD pathology. Of these cases and CNEs, 667 were genotyped as a part of the TGEN1 series[29]. Samples were obtained from twenty-one different National Institute on Aging-supported AD Center brain banks and from the Miami Brain Bank as previously described[29–31]. Additional individual samples from other brain banks in the United States, United Kingdom, and the Netherlands were also obtained in the same manner. The criteria for inclusion were as follows: self-defined ethnicity of European descent, neuropathologically confirmed AD or neuropathology present at levels consistent with status as a control, and age of death greater than 65. Autopsy diagnosis was performed by board certified neuropathologists and was based on the presence or absence of the characterization of probable or possible AD. Where it was possible, Braak and Braak staging and/or CERAD classification were employed. Samples derived from subjects with a clinical history of stroke, cerebrovascular disease, comorbidity with any other known neurological disease, or with the neuropathological finding of Lewy bodies were excluded.

UKS: The UKS cohort is a thoroughly diagnosed case-control cohort from Universitatsklinikum Saarlandes, consisting of individuals clinically diagnosed with sporadic AD (N = 596; mean age onset, 72.2 ± 6.6 years) and cognitively healthy, age, gender and ethnicity matched population-based controls (N = 170; 64.1 ± 3.0 years).

University of Miami/Case Western Reserve University/Mt. Sinai School of Medicine (UM/CWR/MSSM): The UM/CWR/MSSM dataset (formerly UM/VU/MSSM) contains 1,186 cases and 1,135 CNEs (new and previously published)[32–35] ascertained at the University of Miami, Case Western Reserve University and Mt. Sinai School of Medicine, including 409 autopsy-confirmed cases and 136 controls, primarily from the Mt. Sinai School of Medicine[36]. An additional 16 cases were included and 34 controls excluded from the data analyzed in the Jun et al. 2010 study[5]. Each affected individual met NINCDS-ADRDA criteria[8,9] 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.

University of Pittsburgh (UP): The University of Pittsburgh dataset contains1,251 non-Hispanic White AD cases (of which 277 were autopsy-confirmed) recruited by the University of Pittsburgh Alzheimer's Disease Research Center, and 829 non-Hispanic White, CNEs ages 60 and older (2 were autopsy-confirmed). All AD cases met NINCDS/ADRDA criteria for probable or definite AD[8,9]. Additional details of the cohort used for GWAS have been previously published[37].

Washington University (WU): A non-Hispanic White American LOAD case-control dataset consisting of 377 cases and 281 healthy elderly controls was used in analyses for this study. This dataset was split between two analysis datasets (WASHU and WASHU2). Participants were recruited as part of a longitudinal study of healthy aging and dementia. Diagnosis of dementia etiology was made in accordance with standard criteria and methods[15]. Severity of dementia was assessed using the Clinical Dementia Rating scale[10].

Washington Heights-Inwood Community Aging Project (WHICAP): WHICAP is a community-based longitudinal study of aging and dementia among elderly, urban-dwelling residents[38,39]. Beginning enrolment in 1989, WHICAP has followed more than 5,900 residents over 65 years of age, including white, African American, and Hispanic participants. Detailed clinical assessments were performed at approximately 24-month intervals over the 7 years of the initial study. All interviews were conducted in either English or Spanish. The choice of language was decided by the subject in order to ensure the best performance, and the majority of assessments were performed in the subject's home, which included medical, neurological, and neuropsychological evaluations. Results of the neurological, psychiatric and neuropsychological assessments were reviewed in a consensus conference comprised of neurologists, psychiatrists, and neuropsychologists. Based on this review all participants were assigned to one of three categories: dementia, cognitive impairment or normal cognitive function. The sample set available in the ADGC for genetic analyses included 73 AD cases and 560 subjects with normal cognitive function.

The Cohort for Heart and Ageing Research in Genomic Epidemiology (CHARGE) Consortium

The CHARGE consortium currently includes six large, prospective, community-based cohort studies that have genome-wide variation data coupled with extensive data on multiple phenotypes[40]. A neurology working-group arrived at a consensus on phenotype harmonization, covariate selection and analytic plans for within-study analyses and meta-analysis of results[41]. Consent procedures, examination and surveillance components, data security, genotyping protocols and study design at each study were approved by a local Institutional Review Board, details are provided below. Of the six studies, one, the Atherosclerosis Risk in Communities (ARIC) study, was excluded from these analyses as they had not systematically ascertained all their dementia cases at the time of this analysis. Three of the remaining five studies, the Cardiovascular Health Study (CHS), the Framingham Heart Study (FHS) and the Rotterdam

Study had data on both the baseline prevalence, and on incident Alzheimer's disease (AD), whereas the, the Age, Gene/Environment Susceptibility – Reykjavik Study (AGES-RS) and Austrian Stroke Prevention Study (ASPS) only had data on prevalent AD.

ASPS/PRODEM: The Austrian Stroke Prevention Study and The Prospective Dementia Register of the Austrian Alzheimer Society was supported by The Austrian Science Fond (FWF) grant number P20545-P05 (H Schmidt) and P13180; The Austrian Alzheimer Society; The Medical University of Graz

AGES-RS: The AGES RS is a single center prospective study based on the Reykjavik Study, which was initiated in 1967 by the Icelandic Heart Association to study cardiovascular disease and risk factors. The cohort included men and women born 1907-35 and living in Reykjavik in 1967. In 2002, the National Institute of Aging in collaboration with the Icelandic Heart Association started the AGES study that focuses on 4 biological systems: vascular, neurocognitive, musculoskeletal and metabolism[42] . Between 2002 and 2006, it enrolled 5764 participants (42% male) who were randomly selected among the survivors (n= 8030) of the Reykjavik Study. All cohort members were European non-Hispanic White ancestry. Participants were evaluated with a questionnaire and clinical exam, had a fasting sample of blood drawn, and underwent various bio-imaging measures. Of the 5764 participants, 3664 participants were randomly selected for the GWAS. Genotyping was undertaken using the HumanCNV370-Duo (Illumina) at the Laboratory of Neurogenetics, Intramural Research Program, at the National Institute of Aging, Bethesda, Maryland. A total of 2807 persons passed genotyping QC criteria, and, based on the methods outlined in Section 3, were categorized as either non-demented or having Alzheimer's disease (AD). The AGES-RS was approved by the Icelandic National Bioethics Committee (VSN 00-063), the Icelandic Data Protection Authority, and by the Institutional Review Board of the US National Institute on Aging, National Institutes of Health.

This study included all persons with prevalent AD detected between 2002 and 2006. The Folstein Mini Mental State Examination (MMSE) and the Digit Symbol Substitution Test (DSST) were administered to all participants and persons who scored below a pre-determined threshold on these tests (≤23 on the MMSE or ≤17 on the DSST) were administered a second, diagnostic test battery. Based on performance on the Trails B and the Rey Auditory Verbal Learning test (RAVLT), a subset of these individuals with a RAVLT score ≤18 or Trails B score≥8 (ratio of time taken for Trails B/Trails A corrected for the number correct) went on to a third step, which included a neurological examination and a structured informant interview about medical history and social, cognitive, and daily functioning. MRI was acquired as a part of the core study protocol. A panel that included a geriatrician, neurologist, neuropsychologist, and neuroradiologist reached a consensus diagnosis of dementia based on the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) guidelines. There were 319 cases of dementia diagnosed in the first 5764 AGES participants and of these 123 also had genotyping and brain MRI. International diagnostic guidelines, including the National Institute of Neurological and Communicative

Disorders and Stroke–Alzheimer Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable and possible Alzheimer Disease and the Alzheimer's Disease Diagnosis and Treatment Center's (ADDTC) State of California criteria for probable and possible vascular dementia (VaD) with or without AD, were followed. The AGES study identified 3 subtypes: possible/probable AD without VaD (n=55, included in analysis), mixed AD (n=23, cases that met criteria for both AD and VaD, also included in analysis), and, possible/probable VaD or other dementia without AD (n=45, excluded from this study).

Cardiovascular Health Study (CHS): The CHS is a prospective population-based cohort study of risk factors for vascular and metabolic disease that in 1989-90, enrolled adults aged ≥65 years, at four field centers located in North Carolina, California, Maryland and Pennsylvania[43]. The original predominantly non-Hispanic White ancestry cohort of 5201 persons was recruited from a random sample of people on Medicare eligibility lists and an additional 687 African-Americans were enrolled subsequently for a total sample of 5888.[34] Deoxyribonucleic acid (DNA) was extracted from blood samples drawn on all persons who consented to genetic testing at their baseline examination in 1989-90. In 2007-2008, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV Duo® BeadChip system on 3980 CHS participants who were free of cardiovascular disease (CVD) at baseline. The 1908 persons excluded for prevalent CVD had prevalent coronary heart disease (n=1195), congestive heart failure (n=86), peripheral vascular disease (n=93), valvular heart disease (n=20), stroke (n=166) or transient ischemic attack (n=56). Some persons had more than one reason to be excluded and for these individuals only the initial exclusionary event is listed. Because the other cohorts were predominantly white, the African American CHS participants were excluded from this analysis to limit errors secondary to population stratification. Among white participants genotyping was attempted in 3397 participants and was successful in 3295 persons. A total of 742 persons were excluded as they either died prior to the start of the CHS cognition study in 1992 (see section 3 for details), or could not be evaluated completely for baseline cognitive status, leaving a baseline sample of 2553 persons; an additional 31 persons were excluded for having dementia other than AD leaving a study sample of 2522 persons. The CHS study protocols were approved by the Institutional review boards at the individual participating centers.

The AD sample for this study included all prevalent cases identified in 1992 and incident events identified between 1992 and December 2006[44]. Briefly, persons were examined annually from enrollment to 1999, and the examination included a 30 minute screening cognitive battery. In 1992-94 and again, in 1997-99, participants were invited to undergo brain MRI and detailed cognitive and neurological assessment as part of the CHS Cognition Study. Persons with prevalent dementia were identified, and all others were followed until 1999 for the development of incident dementia and AD. Since then, CHS participants at the Maryland and Pennsylvania centers have remained under ongoing dementia surveillance[45].

Beginning in 1988/89, all participants completed the Modified Mini-Mental State Examination (3MSE) and the DSST at their annual visits, and the Benton Visual Retention Test (BVRT) from 1994 to 1998. The Telephone Interview for Cognitive Status (TICS) was used when participants did not come to the clinic. Further information on cognition was obtained from proxies using the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE), and the dementia questionnaire (DQ). Symptoms of depression were measured with the modified version of the Center for Epidemiology Studies Depression Scale (CES-D). In 1991-94, 3608 participants had an MRI of the brain and this was repeated in 1997-98. The CHS staff also obtained information from participants and next-of-kin regarding vision and hearing, the circumstances of the illness, history of dementia, functional status, pharmaceutical drug use, and alcohol consumption. Data on instrumental activities of daily living (IADL), and activities of daily living (ADL) were also collected.

Persons suspected to have cognitive impairment based on the screening tests listed above underwent a neuropsychological and a neurological evaluation. The neuropsychological battery included the following tests: the American version of the National Reading test (AMNART), Raven's Coloured Progressive Matrices, California Verbal Learning Test (CVLT), a modified Rey-Osterreith figure, the Boston Naming test, the Verbal fluency test, the Block design test, the Trails A and B tests, the Baddeley & Papagno Divided Attention Task, the Stroop, Digit Span and Grooved Pegboard Tests. The results of the neuropsychological battery were classified as normal or abnormal (>1.5 standard deviations below individuals of comparable age and education) based on normative data collected from a sample of 250 unimpaired subjects. The neurological exam included a brief mental status examination, as well as a complete examination of other systems. The examiner also completed the Unified Parkinson's Disease Rating Scale (UPDRS) and the Hachinski Ischemic Scale. After completing the neurological exam, the neurologist classified the participant as normal, having mild cognitive impairment (MCI), or dementia. International diagnostic guidelines, including the NINCDS-ADRDA criteria for probable and possible AD and the ADDTC's State of California criteria for probable and possible vascular dementia (VaD) with or without AD, were followed. CHS identified 3 subtypes: possible/probable AD without VaD (categorized as pure AD, included in all AD) and mixed AD (for cases that met criteria for both AD and VaD, included in all-AD), and, possible/probable VaD without AD (excluded from current study).

Framingham Heart Study (FHS): The FHS is a three-generation, single-site, community-based, ongoing cohort study that was initiated in 1948. It now comprises three generations of participants including the Original cohort followed since 1948 (n=5209),[46] their Offspring and spouses of the offspring (n=5216) followed since 1971[47]; and children from the largest Offspring families enrolled in 2000 (Gen 3)[48]. Participants in the Original and Offspring cohorts are used in these analyses, but Gen 3 participants were not included since they are young (mean age 40±9 years in 2000) and none had developed Alzheimer's Disease (AD). The Original cohort enrolled 5209 men and women who comprised two-thirds of the adult population then residing in Framingham, Massachusetts. Survivors continue to receive biennial

examinations. The Offspring cohort comprises 5124 persons (including 3514 biological offspring) who have been examined approximately once every 4 years. Almost all the FHS Original and Offspring participants are non-Hispanic White. FHS participants had DNA extracted and provided consent for genotyping in the 1990s. All available eligible participants were genotyped at Affymetrix (Santa Clara, CA) through an NHLBI funded SNP-Health Association Resource (SHARe) project using the Affymetrix GeneChip® Human Mapping 500K Array Set and 50K Human Gene Focused Panel.® In 272 persons, small amounts of DNA were extracted from stored whole blood and required whole genome amplification prior to genotyping. Cell lines were available for most of the remaining participants. Genotyping was attempted in 5293 Original and Offspring cohort participants, and 4425 persons met QC criteria. Failures (call rate<97%, extreme heterozygosity or high Mendelian error rate) were largely restricted to persons with whole-genome amplified DNA and DNA extracted from stored serum samples. In addition, since the persons with whole genome amplified DNA represent a group of survivors who may differ from the others we included whole genome amplified status as a covariate in FHS analyses. For the prevalent analyses, we also excluded 2268 participants who were less than 65 years old at the time of the DNA draw and 14 persons with dementia other than AD; the remaining 2143 subjects constitute the FHS sample for the prevalent study. A total of 806 well-genotyped persons from the Original cohort (which has been under ongoing surveillance for incident dementia since 1975) were included in the incident AD analyses. The FHS component of this study was approved by the Institutional Review Board of the Boston Medical Center.

The Original cohort of the FHS has been evaluated biennially since 1948, was screened for prevalent dementia and AD in 1974-76 and has been under surveillance for incident dementia and AD since then[49–51]. The Offspring have been examined once every 4 years and have been screened for prevalent dementia with a neuropsychological battery and brain MRI[52,53]. In order to be consistent with the sampling frame for the AGES and CHS samples, we excluded FHS subjects with a baseline age <65 yrs at the time of DNA draw which was in the 1990s. To minimize survival biases, Original cohort and Offspring participants who developed dementia prior to the date of DNA draw were treated as prevalent cases, and subsequent events in the Original cohort occurring prior to December 2006 were included in the incident analyses.

At each clinic exam, participants receive questionnaires, physical examinations and laboratory testing; between examinations they remain under surveillance (regardless of whether or not they live in the vicinity) via physician referrals, record linkage and annual telephone health history updates. Methods used for dementia screening and follow-up have been previously described[49,54]. Briefly, surviving cohort members who attended biennial examination cycles 14 and 15 (May 1975-November 1979) were administered a standardized neuropsychological test battery to establish a dementia-free cohort. Beginning at examination cycle 17 (1982), the MMSE was administered biennially to the cohort. A MMSE score below the education-specific cutoff score, a decline of 3 or more points on subsequent administrations, a decline of more than 5 points compared with any previous examination, or a physician

or family referral prompted further in-depth testing. The Offspring cohort that was enrolled in 1971 has undergone 8 re-examinations, one approximately every 4 years. Starting at the 2nd Offspring examination, participants were questioned regarding any subjective memory complaints and since the 5th Offspring examination participants have been administered the MMSE at each visit. In addition concurrent with the 7th and 8th Offspring examinations (between 1999 and 2004 and then again between 2005 and 2009) surviving Original cohort and all eligible and consenting Offspring participants have undergone volumetric brain MRI and neuropsychological testing[52,53]. The neuropsychological test battery included the Reading subtest of the Wide Range Achievement Test (WRAT-3), the Logical Memory and the Paired Associates Learning tests from the Wechsler Memory Scale, the Visual Reproduction and Hooper Visual Organization Tests, Trails A and B, the Similarities subtest from the Wechsler Adult Intelligence test, the 30-iterm version of the Boston Naming Test and at the second assessment only, the Digit Span, Controlled Word Association and Clock Drawing Tests. Offspring participants suspected to have cognitive impairment based on their MMSE scores, participant, family or physician referral, hospital records or performance in the neuropsychological test battery described above were referred for more detailed neuropsychological and neurological evaluation.

Each participant thus identified underwent baseline neurologic and neuropsychological examinations. Neurologists (trained in geriatric behavioral assessment) supplemented their clinical assessment with a few structured cognitive tests and administered the Clinical Dementia Rating (CDR). Persons were reassessed systematically for the onset of at least mild dementia. A panel consisting of at least 1 neurologist (S.A., P.A.W., or S.S.) and 1 neuropsychologist (R.A.) reviewed all available medical records to arrive at a final determination regarding the presence or absence of dementia, the date of onset of dementia, and the type of dementia. For this determination, we used data from the neurologist's examination, neuropsychological test performance, Framingham Study records, hospital records, information from primary care physicians, structured family interviews, computed tomography and magnetic resonance imaging records, and autopsy confirmation when available. All individuals identified as having dementia satisfied the DSM-IV criteria, had dementia severity equivalent to a CDR of 1 or greater, and had symptoms of dementia for at least 6 months. All individuals identified as having Alzheimer-related dementia met the NINCDS-ADRDA criteria for definite, probable, or possible AD. Vascular Dementia was diagnosed using the ADDTC criteria but the presence of vascular dementia did not disqualify a participant from obtaining a concomitant diagnosis of AD if indicated. The recruitment of Original cohort participants at FHS had occurred long before the DNA collection with the result that the majority of dementia events in the FHS (although ascertained prospectively) were prevalent at the time of DNA collection or these persons had died prior to DNA draw and were thus excluded from analyses of incident disease. Due to the limited number of incident dementia and AD events in the Framingham Offspring only the Original cohort were included in our analyses of incident events.

Rotterdam Study: The Rotterdam Study enrolled inhabitants from a district of Rotterdam (Ommoord) aged ≥55 years (N=7983, virtually all white) at the baseline examination in 1990-93 when blood was drawn for genotyping[55,56]. It aims to examine the determinants of disease and health in the elderly with a focus on neurogeriatric, cardiovascular, bone, and eye disease.8 All inhabitants of Ommoord aged ≥55 years (n = 10275) were invited and the participation rate was 78%. All participants gave written informed consent to retrieve information from treating physicians. Baseline measurements were obtained from 1990 to 1993 and consisted of an interview at home and two visits to the research center for physical examination. Survivors have been re-examined three times: in 1993-1995, 1997-1999, and 2002-2004. All persons attending the baseline examination in 1990-93 consented to genotyping and had DNA extracted. This DNA was genotyped using the Illumina Infinium II HumanHap550chip v3·0® array in 2007-2008 according to the manufacturer's protocols. Genotyping was attempted in persons with high-quality extracted DNA (n=6449). From these 6449, samples with low call rate (<97.5%, n=209), with excess autosomal heterozygosity (> 0.336, n=21), with sex-mismatch (n=36), or if there were outliers identified by the IBS clustering analysis (>3 standard deviations from population mean, n=102 or IBS probabilities > 97%, n=129) were excluded from the study population with some persons meeting more than one exclusion criterion; in total, 5974 samples were available with good quality genotyping data, 42 persons were excluded since they did not undergo cognitive screening at baseline, hence their cognitive status was uncertain. An additional 61 persons were excluded because they suffered from dementia other than AD at baseline. Thus there were 5871 persons included in the prevalent analysis and after exclusion of 171 persons with prevalent AD, 5700 persons were followed for incident AD and other dementia. The Rotterdam Study (including its brain magnetic resonance imaging (MRI) and neurological components) has been approved by the institutional review board (Medical Ethics Committee) of the Erasmus Medical Center and the Netherlands Ministry of Health, Welfare and Sports.

Participants were screened for prevalent dementia in 1990-93 using a three-stage process; those free of dementia remained under surveillance for incident dementia, a determination made using records linkage and assessment at three subsequent re-examinations[57]. We included all prevalent cases and all incident events up to 31st December 2004.

[58,59]Screening was done with the MMSE and GMS organic level for all persons. Screen-positives (MMSE <26 or Geriatric Mental Schedule (GMS) organic level >0) underwent the CAMDEX. Persons who were suspected of having dementia underwent more extensive neuropsychological testing. When available, imaging data were used. In addition, all participants have been continuously monitored for major events (including dementia) through automated linkage of the study database with digitized medical records from general practitioners, the Regional Institute for Outpatient Mental Health Care and the municipality. In addition physician files from nursing homes and general practitioner records of participants who moved out of the Ommoord district were reviewed twice a year. For suspected dementia events, additional information (including neuroimaging) was obtained from hospital records and research physicians discussed available information with a neurologist experienced in dementia diagnosis and

research to verify all diagnoses. Dementia was diagnosed in accordance with internationally accepted criteria for dementia (Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition (DSM-III-R), and AD using the NINCDS-ADRDA criteria for possible, probable and definite AD. The National Institute of Neurological Disorders and Stroke–Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) criteria were used to diagnose vascular dementia. The final diagnosis was determined by a panel of a neurologist, neurophysiologist, and research physician and the diagnoses of AD and VaD were not mutually exclusive.

EUROPEAN ALZHEIMER'S DISEASE INITIATIVE (EADI) CONSORTIUM

All the 2,243 AD cases were ascertained by neurologists from Bordeaux, Dijon, Lille, Montpellier, Paris, Rouen, and were identified as French non-Hispanic White ancestry. Clinical diagnosis of probable AD was established according to the DSM-III-R and NINCDS-ADRDA criteria. Controls were selected from the 3C Study[59]. This cohort is a population-based, prospective (7-years follow-up) study of the relationship between vascular factors and dementia. It has been carried out in three French cities: Bordeaux (southwest France), Montpellier (southeast France) and Dijon (central eastern France). A sample of non-institutionalised, over-65 subjects was randomly selected from the electoral rolls of each city. Between January 1999 and March 2001, 9,686 subjects meeting the inclusion criteria agreed to participate. Following recruitment, 392 subjects withdrew from the study. Thus, 9,294 subjects were finally included in the study (2,104 in Bordeaux, 4,931 in Dijon and 2,259 in Montpellier). Genomic DNA samples of 7,200 individuals were transferred to the French Centre National de Génotypage (CNG). First stage samples that passed DNA quality control were genotyped with Illumina Human 610-Quad BeadChips. At the end we removed 308 samples because they were found to be first- or second-degree relatives of other study participants, or were assessed non-European descent based on genetic analysis using methods described in [60]. In this final sample, at 7 years of follow-up, 459 individuals suffered from AD with 97 prevalent and 362 incident cases. These AD cases were included as cases in the EADI discovery dataset. We retained the other individuals as controls (n=6,017).

GENETIC AND ENVIRONMENTAL RISK IN AD (GERAD) CONSORTIUM/ DEFINING GENETIC, POLYGENIC AND ENVIRONMENTAL RISK FOR ALZHEIMER'S DISEASE CONSORTIUM (PERADES)

The GERAD/PERADES sample comprises 3177 AD cases and 7277 controls with available age and gender data[61]. Cases and elderly screened controls were recruited by the Medical Research Council (MRC) Genetic Resource for AD (Cardiff University; Institute of Psychiatry, London; Cambridge University; Trinity College Dublin), the Alzheimer's Research Trust (ART) Collaboration (University of Nottingham; University of Manchester; University of Southampton; University of Bristol; Queen's

University Belfast; the Oxford Project to Investigate Memory and Ageing (OPTIMA), Oxford University); Washington University, St Louis, United States; MRC PRION Unit, University College London; London and the South East Region AD project (LASER-AD), University College London; Competence Network of Dementia (CND) and Department of Psychiatry, University of Bonn, Germany; the National Institute of Mental Health (NIMH)AD Genetics Initiative. 6129 population controls were drawn from large existing cohorts with available GWAS data, including the 1958 British Birth Cohort (1958BC) (http://www.b58cgene.sgul.ac.uk), the KORA F4 Study and the Heinz Nixdorf Recall Study. All AD cases met criteria for either probable (NINCDS-ADRDA, DSM-IV) or definite (CERAD) AD. All elderly controls were screened for dementia using the MMSE or ADAS-cog, were determined to be free from dementia at neuropathological examination or had a Braak score of 2.5 or lower. Genotypes from all cases and 4617 controls were previously included in the AD GWAS by Harold and colleagues[52]. Genotypes for the remaining 2660 population controls were obtained from WTCCC2.

STAGE 2 DATASET DESCRIPTIONS

Stage 2 consisted of 8,362 cases and 10,483 controls genotyped using the Illumina iSelect microarray. For a detailed description of the genotyping and quality control please see Lambert et al. 2013[2]. Briefly, variants associated with Alzheimer's disease risk with $P < 1 \times 10^{-3}$ in the Lambert et al. 2013 stage 1 analysis were selected for replication genotyping. A total of 11,632 variants passed the array design process and quality control and were available for analysis in Stage 2 of the current study. Cohorts genotyped were from centers in Belgium (1 center), Finland (1 center), Germany (4 centers), Greece (1 center), Hungary (1 center), Italy (8 centers), Spain (7 centers), Sweden (2 centers), the UK (5 centers) and the USA (1 center) (see **Supplementary Table 20** for descriptions by center). Clinical diagnoses of probable AD were all established according to the DSM-III-R and NINCDS-ADRDA criteria. Controls were defined as subjects without DMS-III-R dementia criteria and with integrity of their cognitive functions (MMS > 25). German individuals from Bonn center were genotyped using the Illumina Omni1-Quad chip and imputed with IMPUTE2 using 1,000 Genomes reference data (Feb 2012 release).

STAGE 3 DATASET DESCRIPTIONS

Stage 3A consisted of 3,348 cases and 3,650 controls from GERAD and CHARGE genotyped using Taqman genotyping. Cohorts genotyped were from the UK and Spain. An additional 514 cases and 790 controls from the ADC7 cohort from the ADGC were genotyped using the Illumina OmniExpress and imputed following the same procedures described in the imputation methods for Stage 1. Stage 3B included the persons genotyped in Stage 2A and 3A genotyped using Sequenom MassArray iPLEX or Taqman genotyping. Clinical diagnoses of probable AD were all established according to the DSM-III-R and NINCDS-ADRDA criteria. Controls for the cohorts from the UK and Spain were defined as subjects without DMS-III-R dementia criteria and with integrity of their cognitive functions (MMS > 25). Controls from ADC7 did not meet NINCDS-ADRDA/DSMIV-V criteria for dementia, did not have a diagnosis of mild

cognitive impairment (MCI), and had a CDR of 0, if performed. ADC7 controls also 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. 29 AD cases and 9 controls from ADC7 were confirmed pathologically.

REFERENCES

1. Naj, A. C. *et al.* Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* **43,** 436–441 (2011).
2. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45,** 1452–8 (2013).
3. Sims, R. C. *et al.* Novel rare coding variants in PLCG2, ABI3 and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat. Genet.* 1373–1387 (2017). doi:10.1038/ng.3916
4. Jun, G. *et al.* A novel Alzheimer disease locus located near the gene encoding tau protein. *Mol. Psychiatry* **21,** 108–117 (2016).
5. Jun, G. *et al.* Meta-analysis confirms CR1, CLU, and PICALM as alzheimer disease risk loci and reveals interactions with APOE genotypes. *Arch Neurol* **67,** 1473–1484 (2010).
6. Kukull, W. A. *et al.* Dementia and Alzheimer disease incidence: a prospective cohort study. *Arch Neurol* **59,** 1737–1746 (2002).
7. Larson, E. B. *et al.* Exercise is associated with reduced risk for incident dementia among persons 65 years of age and older. *Ann Intern Med* **144,** 73–81 (2006).
8. McKhann, G. *et al.* Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34,** 939–944 (1984).
9. McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* **7,** 263–269 (2011).
10. Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A. & Martin, R. L. A new clinical scale for the staging of dementia. *Br. J. Psychiatry* **140,** 566–72 (1982).
11. Mirra, S. S., Hart, M. N. & Terry, R. D. Making the diagnosis of Alzheimer's disease. A primer for practicing pathologists. *Arch Pathol Lab Med* **117,** 132–144 (1993).
12. Nagy, Z. *et al.* Assessment of the pathological stages of Alzheimer's disease in thin paraffin sections: a comparative study. *Dement Geriatr Cogn Disord* **9,** 140–144 (1998).
13. Braak, H. & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol* **82,** 239–259 (1991).
14. Beekly, D. L. *et al.* The National Alzheimer's Coordinating Center (NACC) database: the Uniform Data Set. *Alzheimer Dis Assoc Disord* **21,** 249–258 (2007).

15. Morris, J. C. *et al.* The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Dis Assoc Disord* **20,** 210–216 (2006).
16. Petersen, R. C. *et al.* Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* **74,** 201–209 (2010).
17. Saykin, A. J. *et al.* Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. *Alzheimers Dement* **6,** 265–273 (2010).
18. Albert, M. *et al.* Cognitive changes preceding clinical symptom onset of mild cognitive impairment and relationship to ApoE genotype. *Curr. Alzheimer Res.* **11,** 773–84 (2014).
19. Bienias, J. L., Beckett, L. A., Bennett, D. A., Wilson, R. S. & Evans, D. A. Design of the Chicago Health and Aging Project (CHAP). *J. Alzheimers. Dis.* **5,** 349–55 (2003).
20. Barzilai, N., Rossetti, L. & Lipton, R. B. Einstein's institute for aging research: Collaborative and programmatic approaches in the search for successful aging. *Exp. Gerontol.* **39,** 151–157 (2004).
21. Katz, M. J. *et al.* Age-specific and sex-specific prevalence and incidence of mild cognitive impairment, dementia, and Alzheimer dementia in blacks and whites: a report from the Einstein Aging Study. *Alzheimer Dis. Assoc. Disord.* **26,** 335–43 (2012).
22. Green, R. C. *et al.* Risk of dementia among white and African American relatives of patients with Alzheimer disease. *JAMA* **287,** 329–336 (2002).
23. Roccaforte, W. H., Burke, W. J., Bayer, B. L. & Wengel, S. P. Validation of a telephone version of the mini-mental state examination. *J Am Geriatr Soc* **40,** 697–702 (1992).
24. Lee, J. H. *et al.* Analyses of the National Institute on Aging Late-Onset Alzheimer's Disease Family Study: implication of additional loci. *Arch Neurol* **65,** 1518–1526 (2008).
25. Ravid, R. & Swaab, D. F. The Netherlands brain bank--a clinico-pathological link in aging and dementia research. *J. Neural Transm. Suppl.* **39,** 143–53 (1993).
26. Bennett, D. A. *et al.* The Rush Memory and Aging Project: study design and baseline characteristics of the study cohort. *Neuroepidemiology* **25,** 163–175 (2005).
27. Bennett, D. A. *et al.* Natural history of mild cognitive impairment in older persons. *Neurology* **59,** 198–205 (2002).
28. Bennett, D. A., Schneider, J. A., Bienias, J. L., Evans, D. A. & Wilson, R. S. Mild cognitive impairment is related to Alzheimer disease pathology and cerebral infarctions. *Neurology* **64,** 834–841 (2005).
29. Reiman, E. M. *et al.* GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron* **54,** 713–720 (2007).
30. Caselli, R. J. *et al.* Cognitive domain decline in healthy apolipoprotein E epsilon4 homozygotes before the diagnosis of mild cognitive impairment. *Arch Neurol* **64,** 1306–1311 (2007).
31. Webster, J. A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet* **84,** 445–458 (2009).
32. Scott, W. K. *et al.* Complete genomic screen in Parkinson disease: evidence for multiple genes.

JAMA **286,** 2239–2244 (2001).

33. Beecham, G. W. *et al.* Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. *Am J Hum Genet* **84,** 35–43 (2009).
34. Edwards, T. L. *et al.* Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann Hum Genet* **74,** 97–109 (2010).
35. Naj, A. C. *et al.* Dementia Revealed: Novel Chromosome 6 Locus for Late-Onset Alzheimer Disease Provides Genetic Evidence for Folate-Pathway Abnormalities. *PLoS Genet* **6,** (2010).
36. Haroutunian, V. *et al.* Regional distribution of neuritic plaques in the nondemented elderly and subjects with very mild Alzheimer disease. *Arch Neurol* **55,** 1185–1191 (1998).
37. Kamboh, M. I. *et al.* Association of CLU and PICALM variants with Alzheimer's disease. *Neurobiol Aging* **33,** 518–521 (2010).
38. Tang, M.-X. *et al.* Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. *Neurology* **56,** 49–56 (2001).
39. Mayeux, R., Small, S. A., Tang, M.-X., Tycko, B. & Stern, Y. Memory performance in healthy elderly without Alzheimer's disease: effects of time and apolipoprotein-E. *Neurobiol. Aging* **22,** 683–689 (2001).
40. Psaty, B. M. *et al.* Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium design of prospective meta-analyses of genome-wide association studies from 5 Cohorts. *Circ. Cardiovasc. Genet.* **2,** 73–80 (2009).
41. Ikram, M. A. *et al.* Genomewide Association Studies of Stroke. *N. Engl. J. Med.* **360,** 1718–28 (2009).
42. Harris, T. B. *et al.* Age, gene/environment susceptibility-reykjavik study: Multidisciplinary applied phenomics. *Am. J. Epidemiol.* **165,** 1076–1087 (2007).
43. Fried, Linda P., Nemat O. Borhani, Paul Enright, Curt D. Furberg, Julius M. Gardin, Richard A. Kronmal, L. H. K. et al. The cardiovascular health study: design and rationale. *Ann. Epidemiol.* **1,** 263–276 (1991).
44. Fitzpatrick, A. L. *et al.* Incidence and Prevalence of Dementia in the Cardiovascular Health Study. *J. Am. Geriatr. Soc.* **52,** 195–204 (2004).
45. Lopez, O. L. *et al.* Evaluation of Dementia in the Cardiovascular Health Cognition Study. *Neuroepidemiology* **22,** 1–12 (2003).
46. Dawber, T. R. & Kannel, W. B. The Framingham study. An epidemiological approach to coronary heart disease. *Circulation* **34,** 553–555 (1966).
47. Feinleib, M., Kannel, W. B., Garrison, R. J., McNamara, P. M. & Castelli, W. P. The Framingham Offspring Study. Design and preliminary data. *Prev Med* **4,** 518–525 (1975).
48. Splansky, G. L. *et al.* The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* **165,** 1328–1335 (2007).

49. Beiser, A., D'Agostino, R. B., Seshadri, S., Sullivan, L. M. & Wolf, P. A. Computing Estimates of Incidence, Including Lifetime Risk: Alzheimer's Disease in the Framingham Study. The Practical Incidence Estimators (PIE) Macro. *Stat. Med.* **19,** 1945–1522 (2000).
50. Bachman, D. L. *et al.* Incidence of dementia and probable Alzheimer's disease in a general population: the Framingham Study. *Neurology* **43,** 515–9 (1993).
51. Farmer, M. E. *et al.* Neuropsychological test performance in Framingham: A descriptive study. *Psychol. Rep.* **60,** 1023–1040 (1987).
52. DeCarli, C. *et al.* Measures of brain morphology and infarction in the framingham heart study: Establishing what is normal. *Neurobiol. Aging* **26,** 491–510 (2005).
53. Au, R. *et al.* New Norms for a New Generation: Cognitive Performance in the Framingham Offspring Cohort. *Exp. Aging Res.* **30,** 333–358 (2004).
54. Seshadri, S. *et al.* Lifetime risk of dementia and Alzheimer's disease. The impact of mortality on risk estimates in the Framingham Study. *Neurology* **49,** 1498–504 (1997).
55. Hofman, A. *et al.* The rotterdam study: 2010 objectives and design update. *Eur. J. Epidemiol.* **24,** 553–572 (2009).
56. Hofman, A. *et al.* The Rotterdam Study: Objectives and design update. *Eur. J. Epidemiol.* **22,** 819–829 (2007).
57. Ott, A., Breteler, M. M., van Harskamp, F., Stijnen, T. & Hofman, A. Incidence and risk of dementia. The Rotterdam Study. *Am. J. Epidemiol.* **147,** 574–80 (1998).
58. Dreses-Werringloer, U. *et al.* A Polymorphism in CALHM1 Influences Ca2+ Homeostasis, Aβ Levels, and Alzheimer's Disease Risk. *Cell* **133,** 1149–1161 (2008).
59. 3C Study Group, M. *et al.* Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* **22,** 316–25 (2003).
60. Heath, S. C. *et al.* Investigation of the fine structure of European populations with applications to disease association studies. *Eur. J. Hum. Genet.* **16,** 1413–1429 (2008).
61. Harold, D. *et al.* Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* **41,** 1088–1093 (2009).