Published in final edited form as:
JAMA. 2013 April 10; 309(14): 1483–1492. doi:10.1001/jama.2013.2973.

Corresponding Author: Richard Mayeux, MD, MSc, Gertrude H. Sergievsky Center, Columbia University, 630 W 168th St, New York, NY 10032 (rpm2@columbia.edu)..

Author Contributions: Dr Mayeux had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Reitz, Buxbaum, Fallin, Griffith, Obisesan, Manly, Bennett, Haines, Farrer, Pericak-Vance, Schellenberg, Mayeux.
Acquisition of data: Reitz, Valladares, Lin, Larson, Graff-Radford, Evans, Crane, Buxbaum, Murrell, Ertekin-Taner, Baldwin, Green, Barnes, Cantwell, Fallin, Go, Obisesan, Manly, Kamboh, Lopez, Bennett, Hendrie, Hall, Goate, Byrd, Kukull, Foroud, Haines, Farrer, Pericak-Vance, Schellenberg, Mayeux.
Analysis and interpretation of data: Reitz, Jun, Naj, Rajbhandary, Vardarajan, Wang, Valladares, Graff-Radford, De Jager, Crane, Buxbaum, Raj, Ertekin-Taner, Logue, Manly, Lunetta, Bennett, Hendrie, Farrer, Pericak-Vance, Schellenberg, Mayeux.
Drafting of the manuscript: Reitz, Naj, De Jager, Raj, Cantwell, Pericak-Vance, Mayeux.
Critical revision of the manuscript for important intellectual content: Reitz, Jun, Naj, Rajbhandary, Vardarajan, Wang, Valladares, Lin, Larson, Graff-Radford, Evans, Crane, Buxbaum, Murrell, Ertekin-Taner, Logue, Baldwin, Green, Barnes, Fallin, Go, Griffith, Obisesan, Manly, Lunetta, Kamboh, Lopez, Bennett, Hendrie, Hall, Goate, Byrd, Kukull, Foroud, Haines, Farrer, Pericak-Vance, Schellenberg, Mayeux.
Statistical analysis: Reitz, Jun, Naj, Rajbhandary, Vardarajan, Raj, Logue, Hall, Lunetta, Hendrie, Haines, Farrer, Pericak-Vance.
Obtained funding: Larson, Evans, Murrell, Green, Barnes, Fallin, Go, Manly, Kamboh, Bennett, Goate, Byrd, Haines, Pericak-Vance, Schellenberg, Mayeux.
Administrative, technical, or material support: Naj, Wang, Valladares, Lin, Larson, Crane, Buxbaum, Murrell, Ertekin-Taner, Baldwin, Green, Cantwell, Fallin, Obisesan, Kamboh, Lopez, Bennett, Hall, Goate, Byrd, Kukull, Foroud, Farrer, Schellenberg, Mayeux.
Study supervision: Go, Obisesan, Bennett, Hendrie, Byrd, Foroud, Farrer, Schellenberg, Mayeux.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Baldwin reported serving as a consultant for the Center for Human Genetics Inc. Dr Go reported receiving travel support from the National Institutes of Health (NIH). Dr Griffith reported receiving payment for lectures from Eisai and Pfizer. Dr Manly reported serving as a board member for the International Neuropsychological Society; receiving grants or grants pending from the Alzheimer's Association and the National Institute on Aging (NIA); and receiving travel expenses from the Alzheimer's Association. Dr Kamboh reported receiving travel support from the NIH. Dr Lopez reported receiving consulting fees or honoraria from Mertz and Lundbeck. Dr Bennett reported receiving travel support from the NIH. Dr Goate reported serving as a consultant for Finnegan; providing expert testimony in cases involving the genetics of Alzheimer disease; receiving grants or grants pending from Genentech and Pfizer; receiving payment for lectures from Pfizer, Genentech, and Amgen; and receiving patent royalties from Taconic for a tau mutation. Dr Pericak-Vance reported receiving revenues from Athena Diagnostics. No other authors reported disclosures.

Members of the Alzheimer's Disease Genetics Consortium: Marilyn S. Albert, Roger L. Albin, Liana G. Apostolova, Steven E. Arnold, Robert Barber, Michael M. Barmada, Thomas G. Beach, Gary W. Beecham, Duane Beekly, Eileen H. Bigio, Thomas D. Bird, Deborah Blacker, Bradley F. Boeve, James D. Bowen, Adam Boxer, James R. Burke, Guiqing Cai, Nigel J. Cairns, Chuanhai Cao, Chris S. Carlson, Regina M. Carney, Steven L. Carroll, Helena C. Chui, David G. Clark, David H. Cribbs, Elizabeth A. Crocco, Carlos Cruchaga, Charles DeCarli, Steven T. DeKosky, F. Yesim Demirci, Malcolm Dick, Kelley M. Faber, Kenneth B. Fallon, Martin R. Farlow, Steven Ferris, Matthew P. Frosch, Douglas R. Galasko, Mary Ganguli, Marla Gearing, Daniel H. Geschwind, Bernardino Ghetti, John R. Gilbert, Sid Gilman, Jonathan D. Glass, John H. Growdon, Hakon Hakonarson, Ronald L. Hamilton, Kara L. Hamilton-Nelson, Vahram Haroutunian, Lindy E. Harrell, Elizabeth Head, Lawrence S. Honig, Christine M. Hulette, Bradley T. Hyman, Gail P. Jarvik, Gregory A. Jicha, Lee-Way Jin, Anna Karydas, John S. K. Kauwe, Jeffrey A. Kaye, Ronald Kim, Edward H. Koo, Neil W. Kowall, Joel H. Kramer, Patricia Kramer, Frank M. LaFerla, James J. Lah, Rosalyn Lang-Walker, James B. Leverenz, Allan I. Levey, Ge Li, Andrew P. Lieberman, Constantine G. Lyketsos, Wendy J. Mack, Daniel C. Marson, Eden R. Martin, Frank Martiniuk, Deborah C. Mash, Eliezer Masliah, Wayne C. McCormick, Susan M. McCurry, Andrew N. McDavid, Ann C. McKee, Marsel Mesulam, Bruce L. Miller, Carol A. Miller, Joshua W. Miller, Thomas J. Montine, John C. Morris, John M. Olichney, Joseph E. Parisi, Elaine Peskind, Ronald C. Petersen, Aimee Pierce, Wayne W. Poon, Huntington Potter, Joseph F. Quinn, Ashok Raj, Murray Raskind, Eric M. Reiman, Barry Reisberg, John M. Ringman, Erik D. Roberson, Howard J. Rosen, Roger N. Rosenberg, Mary Sano, Andrew J. Saykin, Julie A. Schneider, Lon S. Schneider, William W. Seeley, Amanda G. Smith, Joshua A. Sonnen, Salvatore Spina, Robert A. Stern, Rudolph E. Tanzi, John Q. Trojanowski, Juan C. Troncoso, Debby W. Tsuang, Vivi-anna M. Van Deerlin, Linda J. Van Eldik, Harry V. Vinters, Jean Paul Vonsattel, Sandra Weintraub, Kathleen A. Welsh-Bohmer, Jennifer Williamson, Randall L. Woltjer, Clinton B. Wright, Steven G. Younkin, Chang-En Yu, and Lei Yu.

Online-Only Material: eTables 1-3, eFigures 1 and 2, eMethods, and Author Video Interview are available at http://www.jama.com.

Variants in the ATP-Binding Cassette Transporter (*ABCA7*), Apolipoprotein E ε4, and the Risk of Late-Onset Alzheimer Disease in African Americans

Christiane Reitz, MD, PhD, **Gyungah Jun, PhD**, **Adam Naj, PhD**, **Ruchita Rajbhandary, MPH**, **Badri Narayan Vardarajan, PhD**, **Li-San Wang, PhD**, **Otto Valladares, MS**, **Chiao-Feng Lin, PhD**, **Eric B. Larson, MD, MPH**, **Neill R. Graff-Radford, MD**, **Denis Evans, MD**, **Philip L. De Jager, MD, PhD**, **Paul K. Crane, MD, MPH**, **Joseph D. Buxbaum, PhD**, **Jill R. Murrell, PhD**, **Towfique Raj, PhD**, **Nilufer Ertekin-Taner, MD, PhD**, **Mark Logue, PhD**, **Clinton T. Baldwin, PhD**, **Robert C. Green, MD, MPH**, **Lisa L. Barnes, PhD**, **Laura B. Cantwell, MPH**, **M. Daniele Fallin, PhD**, **Rodney C. P. Go, PhD**, **Patrick Griffith, MD**, **Thomas O. Obisesan, MD**, **Jennifer J. Manly, PhD**, **Kathryn L. Lunetta, PhD**, **M. Ilyas Kamboh, PhD**, **Oscar L. Lopez, MD**, **David A. Bennett, MD**, **Hugh Hendrie, MB, ChB, DSc**, **Kathleen S. Hall, PhD**, **Alison M. Goate, PhD**, **Goldie S. Byrd, PhD**, **Walter A. Kukull, PhD**, **Tatiana M. Foroud, PhD**, **Jonathan L. Haines, PhD**, **Lindsay A. Farrer, PhD**, **Margaret A. Pericak-Vance, PhD**, **Gerard D. Schellenberg, PhD**, **Richard Mayeux, MD, MSc**, and **the Alzheimer Disease Genetics Consortium**
Taub Institute for Research on Alzheimer's Disease and the Aging Brain (Drs Reitz, Manly, and Mayeux), Gertrude H. Sergievsky Center (Drs Reitz, and Mayeux), and Departments of Neurology (Drs Reitz, Manly, and Mayeux), Psychiatry (Dr Mayeux), and Epidemiology (Dr Mayeux), College of Physicians and Surgeons, Columbia University, New York, New York; Departments of Medicine (Genetics Program) (Drs Jun, Vardarajan, Logue, Baldwin, and Farrer), Biostatistics (Drs Jun, Vardarajan, Lunetta, and Farrer), Ophthalmology (Drs Jun and Farrer), Epidemiology (Dr Farrer), and Neurology (Dr Farrer), Boston University, Boston, Massachusetts; The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida (Drs Naj and Pericak-Vance and Ms Rajbhandary); Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia (Drs Wang, Lin, and Schellenberg and Mr Valladares and Ms Cantwell); Department of Medicine (Drs Larson and Crane) and National Alzheimer's Coordinating Center and Department of Epidemiology (Dr Kukull), University of Washington, Seattle; Group Health Research Institute, Group Health, Seattle (Dr Larson); Departments of Neuroscience (Drs Graff-Radford, and Ertekin- Taner) and Neurology (Drs Graff-Radford and Ertekin- Taner), Mayo Clinic, Jacksonville, Florida; Rush Institute for Healthy Aging, Department of Internal Medicine (Dr Evans), Departments of Neurological Sciences (Drs Barnes and Bennett) and Behavioral Sciences (Dr Barnes), and Rush Alzheimer's Disease Center (Dr Bennett), Rush University Medical Center, Chicago, Illinois; Program in Translational Neuropsychiatric Genomics, Department of Neurology, Brigham & Women's Hospital, Boston, Massachusetts (Dr De Jager); Harvard Medical School, Boston (Drs De Jager and Raj); Program in Medical and Population Genetics, The Broad Institute, Cambridge, Massachusetts (Dr De Jager); Department of Medical and Molecular Genetics, Indiana University (Drs Murrell and Foroud), Indiana University Center for Aging Research (Dr Hendrie), and Department of Psychiatry, Indiana University School of Medicine (Drs Hendrie and Hall), Indianapolis; Division of Genetics, Department of Medicine, and Partners Center for Personalized Genetic Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston (Dr Green); Departments of Psychiatry (Dr Buxbaum), Genetics and Genomics Sciences (Dr Buxbaum), and Neuroscience (Dr Buxbaum) and the Friedman Brain Institute (Dr Buxbaum), Mount Sinai School of Medicine, New York, New York; Department of Epidemiology, Johns Hopkins University School of Public Health, Baltimore, Maryland (Dr Fallin); School of Public Health, University of Alabama at Birmingham (Dr Go); Department of Neurology, Meharry Medical College, Nashville, Tennessee (Dr Griffith); Division of Geriatrics, Howard University Hospital, Washington, DC (Dr Obisesan); Department of Human Genetics (Dr Kamboh) and Alzheimer's

Disease Research Center (Drs Kamboh and Lopez), University of Pittsburgh, Pittsburgh, Pennsylvania; Department of Psychiatry and Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St Louis, Missouri (Dr Goate); Department of Biology, North Carolina A & T University, Winston-Salem (Dr Byrd); Department of Molecular Physiology and Biophysics and Vanderbilt Center for Human Genetics Research, Vanderbilt University, Nashville (Dr Haines); and Regenstrief Institute Inc, Indianapolis (Dr Hendrie).

Abstract

Importance—Genetic variants associated with susceptibility to late-onset Alzheimer disease are known for individuals of European ancestry, but whether the same or different variants account for the genetic risk of Alzheimer disease in African American individuals is unknown. Identification of disease-associated variants helps identify targets for genetic testing, prevention, and treatment.

Objective—To identify genetic loci associated with late-onset Alzheimer disease in African Americans.

Design, Setting, and Participants—The Alzheimer Disease Genetics Consortium (ADGC) assembled multiple data sets representing a total of 5896 African Americans (1968 case participants, 3928 control participants) 60 years or older that were collected between 1989 and 2011 at multiple sites. The association of Alzheimer disease with genotyped and imputed single-nucleotide polymorphisms (SNPs) was assessed in case-control and in family-based data sets. Results from individual data sets were combined to perform an inverse variance–weighted meta-analysis, first with genome-wide analyses and subsequently with gene-based tests for previously reported loci.

Main Outcomes and Measures—Presence of Alzheimer disease according to standardized criteria.

Results—Genome-wide significance in fully adjusted models (sex, age, *APOE* genotype, population stratification) was observed for a SNP in *ABCA7* (rs115550680, allele = G; frequency, 0.09 cases and 0.06 controls; odds ratio [OR], 1.79 [95% CI, 1.47-2.12]; $P = 2.2 \times 10^{-9}$), which is in linkage disequilibrium with SNPs previously associated with Alzheimer disease in Europeans ($0.8<D'<0.9$). The effect size for the SNP in *ABCA7* was comparable with that of the *APOE* ε4–determining SNP rs429358 (allele = C; frequency, 0.30 cases and 0.18 controls; OR, 2.31 [95% CI, 2.19-2.42]; $P = 5.5 \times 10^{-47}$). Several loci previously associated with Alzheimer disease but not reaching significance in genome-wide analyses were replicated in gene-based analyses accounting for linkage disequilibrium between markers and correcting for number of tests performed per gene (*CR1, BIN1, EPHA1, CD33;* $0.0005<\text{empirical } P < .001$).

Conclusions and Relevance—In this meta-analysis of data from African American participants, Alzheimer disease was significantly associated with variants in *ABCA7* and with other genes that have been associated with Alzheimer disease in individuals of European ancestry. Replication and functional validation of this finding is needed before this information is used in clinical settings.

Late-onset Alzheimer disease (LOAD) is the most common cause of dementia, increasing in frequency from 1% at age 65 years to more than 30% for people older than 80 years.[1] As much as 20% of the disease-attributable risk is related to the ε4 variant in *APOE*.[2] A series of large genome-wide association studies (GWASs) identified several additional variants that affect disease susceptibility in non-Hispanic whites of European ancestry, including *CR1*, *CLU*, *PICALM*, *BIN1, CD2AP, CD33, EPHA1, MS4A6A/MS4E4*, and *ABCA7*.[3-7] In addition, *SORL1* was identified as a susceptibility gene in candidate gene and functional

studies.[8,9] However, LOAD heritability estimates are high (h^2 ≈60%-80%), and a large part of the genetic contribution to LOAD remains unexplained.[10]

The incidence of LOAD among African Americans is higher than among whites living in the same community,[11] and the reported risk for the disease associated with *APOE* ε4 heterozygosity is inconsistent in African Americans compared with whites.[12] African Americans and other minorities are understudied, and it is unclear whether any of the recently identified loci modify risk of LOAD in racial or ethnic groups other than whites.

To identify genetic variants associated with LOAD in African Americans, the Alzheimer Disease Genetics Consortium (ADGC) performed a GWAS among the largest sample, to our knowledge, of African Americans ever assembled for genetic studies of Alzheimer disease.

METHODS

Study Samples

Participants were recruited from several independent community-based case-control and family studies of African Americans collected over a period of approximately 30 years between 1989 and 2011.[12-35] All participants underwent rigorous phenotyping for LOAD, and diagnoses were made by National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria.[36] Classification of participants as African American was based on self-report using the format of the 1990 US census.[37] A detailed description of the original cohorts contributing samples is provided in the eMethods, available at http://www.jama.com. A glossary of terms used in this article is provided in the box.

All participants provided written informed consent, and the data sets for the study were approved for analysis by the relevant institutional review boards.

Censoring Age

Information on age at onset for case participants and age at examination or death for control participants was available for most cohorts. However, surrogate age information was available for other data sets including age at ascertainment (Indiana University), age at diagnosis (Chicago Health and Aging Project [CHAP], Minority Aging Research Study/ Clinical Minority Core [MARS/CORE]), or age at death (subset of autopsy-confirmed samples in the University of Miami/Vanderbilt University [UM/VU] cohort). Age at death was used for autopsied participants. To restrict the analyses to case participants with LOAD, persons younger than 60 years at last evaluation, symptom onset, or death were excluded.

Genotyping

GWAS genotypes were from a variety of Illumina arrays (eTable 1). For all data sets, case and control samples were randomly plated to minimize potential batch effects. For the Alzheimer Disease Centers, Adult Changes in Thought, National Institute in Aging–LOAD/ National Cell Repository for Alzheimer Disease (NIA-LOAD/NCRAD), UM/VU, CHAP, Columbia University, and Mayo Clinic cohorts, *APOE* genotypes were based on haplotypes derived from single-nucleotide polymorphisms (SNPs) rs7412 and rs429358. For the MIRAGE and GenerAAtions cohorts, *APOE* genotypes were determined using the Roche Diagnostics LightCycler 480 instrument (Roche Diagnostics)[38] and LightMix Kit ApoE C112R R158 (TIB MOLBIOL); for the University of Pittsburgh, Washington Heights Columbia Aging Project, and Indianapolis cohorts, they were determined by pyrosequencing[39] or analysis of restriction fragment length polymorphisms[40,41]; for the Religious Orders Study/Rush Memory and Aging Project (ROS/MAP) and MARS/CORE

they were determined by high-throughput sequencing of codons 112 and 158 in *APOE* by Agencourt Bioscience Corporation; for the Washington University samples they were determined using a taqman-based assay from Applied Biosystems. Single-nucleotide polymorphisms were annotated based on the National Center for Biotechnology Information (NCBI) Reference Sequence database and the GRCh37/hg19 genome build; genes were annotated using NCBI Entrez Gene accession number.

Quality Control Procedures

Single-nucleotide polymorphisms with minor allele frequencies (MAFs) less than 0.01, call rates less than 98%, or not in Hardy-Weinberg equilibrium ($P < 10^{-6}$ in controls) were excluded. Participants whose reported sex differed from the sex assignment determined by analysis of the X-chromo-some SNPs using PLINK version 1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/) were excluded. For cohorts genotyped on multiple chips (MIRAGE, UM/VU), quality control was performed separately for the subsets of individuals genotyped using different chips. Latent relatedness among participants within and across the case-control cohorts was identified by the estimated proportion of alleles (π) shared identical by descent (IBD) using PLINK. The proportion IBD is calculated by estimating the probability of sharing 0, 1, or 2 alleles IBD for any 2 individuals (π = P [IBD = 2] + 0.5 × P [IBD = 1], where P indicates probability). One participant from each duplicate pair ($\pi > 0.95$) or relative pair ($0.4 \leq \pi < 0.95$) was included in the sample used for association analyses, prioritizing based on nonmissing disease status and then higher SNP call rate. Relationships among individuals in the family-based cohorts (MIRAGE, NIA-LOAD/NCRAD) were confirmed by pairwise genome-wide estimates of IBD allele sharing. All discrepancies were reviewed with clinical and pedigree data to determine the most likely relationship consistent with IBD estimates.

Population Substructure

Population substructure was evaluated in each cohort separately using EIGENSTRAT (EIGENSOFT version 3.0) (http://genepath.med.harvard.edu/~reich/EIGENSTRAT.htm).[42] First, genetic profiles for all participants in the case-control data sets and a group of unrelated participants in the MIRAGE family-based data set were compared with those in the HapMap reference panel of African Americans (African ancestry in the Southwest USA), and outliers with respect to African American ancestry were removed from the sample. Then, the data were reevaluated using EIGENSTRAT to derive loadings for the first 10 principal components. Principal component analysis was used to model for each assessed marker ancestry differences in frequency between case and control participants. The resulting information can be used to adjust for population substructure, which minimizes spurious associations and maximizes power to detect true associations.

Genotype Imputation

Genome-wide imputation of allele dosages was performed using the June 2011 panel from 1000 Genomes build 37 for imputation of genotypes (http://www.1000genomes.org/announcements/june-2011-data-release-2011-06-23) and the IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) software applying strict prephasing, preimputation filtering, and variant position and strand alignment control.[43] The reference panel used is a multi-reference panel specifically developed for imputation of nonwhite populations and shown to impute genotypes in African Americans with high accuracy.[43,44] Only imputed SNP dosages with an imputation quality estimate of $R^2 \geq 0.50$ were included in the final SNP set for analysis.

Association Analyses

Association of LOAD with genotyped and imputed SNPs (allele dosages) that had passed quality control was assessed using logistic regression methods for case-control data sets and logistic generalized estimating equations for family data sets as implemented in PLINK. All analyses were performed using an additive genetic model (ie, genotyped SNPs were coded 0, 1, or 2 based on the number of minor alleles [with 0 being homozygous for the reference allele, 1 being heterozygous, and 2 being homozygous for the minor allele], and imputed SNPs were coded based on the posterior probability of the minor allele [0–2]). The primary association analyses were adjusted for age, sex, and population substructure (using for each cohort the appropriate number of principal components) (**Table 1**).

Results from the individual data sets were combined using an inverse variance–weighted meta-analysis approach implemented in METAL (http://genome.sph.umich.edu/wiki/METAL). The meta-analysis *P* value was estimated by the summarized test statistic after applying a genomic control within each individual study. Heterogeneity of effect estimates across data sets (I^2) was tested with the χ^2 distributed Q statistic.[45,46] All analyses were repeated adjusting for the number of *APOE*-ε4 alleles (0, 1, or 2). The threshold for genome-wide significance was calculated as $P \leq 5 \times 10^{-8}$, taking linkage disequilibrium between markers into account. The genomic inflation factors (λ) for each model are estimated based on the concept that apart from a small number of SNPs showing a true association with the disease, the test statistics for other SNPs should follow the distribution under the null hypothesis of no association and thus reflect cryptic population stratification, relatedness, or genotyping errors. These factors were between 0.87 and 1.03, indicating that there was no substantial inflation of the test statistics in either meta-analysis (eFigure 1). All findings with $P \leq 10^{-6}$ in the fully adjusted model were compared with results obtained in whites.[5]

Because of the a priori hypothesis of an involvement with LOAD, associations of SNPs in previously reported LOAD genes (*CR1* [NCBI Entrez Gene 1378], *BIN1* [NCBI Entrez Gene 274], *PICALM* [NCBI Entrez Gene 8301], *CLU* [NCBI Entrez Gene 1191], *EPHA1* [NCBI Entrez Gene 2041], *MS4A6A* cluster [NCBI Entrez Gene 64231], *CD2AP* [NCBI Entrez Gene 23607], and *CD33* [NCBI Entrez Gene 945]), were analyzed with a versatile gene-based association study (VEGAS[47]), adding 50 kilobases (kb) to each side. Gene-based tests for association are a useful complement to GWASs, because gene-based tests consider association between a trait and all SNPs within a gene rather than each marker individually. Depending on the underlying genetic architecture, gene-based approaches can be more powerful than traditional single-SNP–based GWASs, in particular if a gene contains several SNPs with marginal levels of significance that are often indistinguishable from random noise in the initial GWAS. For the specific gene assessed, VEGAS incorporates linkage disequilibrium information from a set of reference individuals from HapMap, determines the number of tagging SNPs, and calculates the empirical *P* value for the gene by using simulations from the multivariate normal distribution.[47] Accordingly, the *P* value threshold for significance differs between genes depending on the linkage disequilibrium structure and number of tagging SNPs assessed.

Strength of linkage disequilibrium—which is a measure of the association of 2 alleles at different loci—between different SNPs observed in the same gene in this African American sample and the white samples was determined by estimating D'. D' ranges from 0 to 1, with 0 indicating no linkage (ie, fully independent transmission from parent to offspring) and 1 indicating perfect linkage (ie, completely linked transmission from parent to offspring) between 2 markers. In contrast to R^2, D' is not influenced by differences in allele frequencies between ethnic groups.

RESULTS

We performed the GWAS using data from 1968 African American case participants with LOAD and 3928 cognitively normal elderly control participants. Fifty percent of the cohort had preexisting genome-wide genotyping, and another 1074 cases and 1908 controls were genotyped specifically for this project. Several characteristics of the individual data sets are shown in Table 1.

The final SNP set included a total of 17 332 474 genotyped and imputed variants. The association with the lowest *P* value was with *APOE* (NCBI Entrez Gene 348). In models adjusting for age, sex, and population stratification, numerous SNPs in the *APOE* region were significant (eg, rs429358, $P = 5.5 \times 10^{-47}$) for association with LOAD. Excluding SNPs in the *APOE* region, the strongest associations were observed for rs10247412 in *ELMO1* (NCBI Entrez Gene 9844) (odds ratio [OR], 0.66 [95% CI, 0.56-0.77]; $P = 2.9 \times 10^{-7}$), rs885330 in *SOX13* (NCBI Entrez Gene 9580) (OR, 1.25 [95% CI, 1.17-1.33]; $P = 3.9 \times 10^{-7}$), an intergenic SNP (rs145848414) at 174 014 114 base pairs on chromosome 5q35.2 that is not near any genes with a known function (OR, 2.03 [95% CI, 1.54-2.67]; $P = 5.1 \times 10^{-7}$), and rs115550680 in *ABCA7* (NCBI Entrez Gene 10347) (OR, 1.78 [95% CI, 1.28-1.82]; $P = 1.4 \times 10^{-6}$). After adjustment for *APOE*, the associations with *ELMO1* and *SOX13* SNPs diminished, whereas the association for rs115550680 in *ABCA7* (OR, 1.79 [95% CI, 1.47-2.12]; $P = 2.21 \times 10^{-9}$) became stronger (**Table 2**). The association of rs145848414 on chromosome 5q35.2 with LOAD also became stronger but did not fully reach genome-wide significance (OR, 2.29 [95% CI, 1.69-3.09]; $P = 6.9 \times 10^{-8}$). The increases in effect size were accompanied by decreases in *P* value, which were most pronounced in the larger data sets (ADGC, CHAP, MIRAGE660, Indianapolis).

In African Americans, the SNP in *ABCA7* (rs115550680) is in linkage disquilibrium with 2 other *ABCA7* SNPs previously associated with LOAD at the genome-wide significance level in non-Hispanic whites of European ancestry (rs3764650 [Hollingworth et al[3]] and rs3752246 [Naj et al[5]], 0.8<D' = 0.9) (**Figure 1**) and showed the same direction of effect. The effect size for rs115550680 in *ABCA7* (OR, 1.79 [95% CI, 1.47-2.12]; $P = 2.21 \times 10^{-9}$) was comparable with that observed for *APOE* (OR, 2.31 [95% CI, 2.19-2.42]; $P = 5.5 \times 10^{-47}$). Comparison of regional association plots for *ABCA7* in this African American sample and the non-Hispanic white sample described in Naj et al[5] showed more widespread associations among African Americans (**Figure 2**). Consistent with this finding, in this African American sample, SNPs at 2 adjacent loci on chromosome 19p (*GRIN3B* [NCBI Entrez Gene 116444] and *HMHA1* [NCBI Entrez Gene 23526]) were associated with LOAD at $P \leq 10^{-8}$ in the fully adjusted model (Table 2). *ABCA7*, *GRIN3B*, and *HMHA1* span a 81-kb region on chromosome 19p and are in linkage dis-equilibrium (0.8<D'<0.95) (Table 2). Further analyses conditioned on rs115550680 in *ABCA7* revealed that the associations in *GRIN3B* and *HMHA1* were not independent (eTable 2)

The imputation quality (R^2) for rs115550680 in *ABCA7*, the significant SNPs in *GRIN3B*, *HMHA1*, and the novel locus on chromosome 5q35.2 was high (0.87-0.99) across all data sets included in the analyses (eTable 3). Forest plots (eFigure 2) indicated the consistency of results across data sets. The *ABCA7* SNPs previously reported in whites (rs3764650[3] and rs3752246[5]) did not reach genome-wide significance in this African American data set. However, the MAFs for these 2 vari ants largely differ between populations of European and African descent (MAF for rs3764650, 0.25 in African Americans and 0.11 in Europeans; MAF for rs3752246, 0.04 in African Americans and 0.19 in Europeans). In turn, rs115550680, significant in this African American data set, is monomorphic in Europeans. However, as described above, the direction of effects of rs115550680, rs3764650, and rs3752246 were similar.

The susceptibility loci previously associated with LOAD in whites, which did not reach the *P* value cutoff for genome-wide significance in this African American data set (*CR1, BIN1, PICALM, CLU, EPHA1, MS4A* cluster, *CD2AP, CD33)*, were further explored in gene-based analyses adding 50 kb to both sides of each gene.[3-7] **Table 3** shows the genes significant in these gene-based tests and reports information on the number of tagging SNPs assessed in each gene and the corresponding *P* value threshold needed to reach statistical significance. After correcting for the number of independent tests per gene, SNPs in *CR1* (rs146366639: OR, 0.82 [95% CI, 0.73-0.92]; empirical *P* = .0005), *BIN1* (rs55636820: OR, 1.89 [95% CI, 1.31-2.75]; empirical *P* = .0007), *EPHA1* (rs6973770: OR, 0.70 [95% CI, 0.56-0.87]; empirical *P* = .001), and *CD33* (rs114282264: OR, 0.61 [95% CI, 0.47-0.81]; empirical *P* = .0007) were significantly associated with LOAD, although the most significant SNPs differed from the top-ranked SNPs in Europeans.

DISCUSSION

To our knowledge, the present study is the largest GWAS for the study of LOAD in African Americans ever assembled. Aside from SNPs associated with APOE, the top-ranked SNP observed in this study was located in *ABCA7* (rs115550680) and had an effect size comparable with that of *APOE* ε4. This observation differs from the previous GWAS in whites. The reported *ABCA7* SNPs in non-Hispanic whites have lower effect sizes (rs3752246: OR, 1.13 [95% CI, 1.03-1.25]; rs3764650: OR, 1.23 [95% CI, 1.17-1.28]),[3,5] as do all other genes reported in whites (*CR1, BIN1, PICALM, CLU, EPHA1, MS4A* cluster, *CD2AP, CD33).*[3-6]

It remains possible that this could be attributable to population differences in the frequencies of the causative variant(s) tagged by the associated SNPs (rs115550680 in *ABCA7* is monomorphic in non-Hispanic whites; the MAF for rs3752246 is 0.04 in African Americans and 0.19 in non-Hispanic whites; the MAF for rs3764650 is 0.25 in African Americans and 0.11 in non-Hispanic whites) or the result of a bias in the estimated effect of a newly identified allele on disease (also termed “winner's curse”).

However, it is also possible that the large difference between whites and African Americans in the effect size of the *ABCA7* locus on the relative odds of being diagnosed with LOAD is explained by population-specific causative variants with variable influence on protein structure or function. The linkage disequilibrium block in which rs115550680 is located spans across several introns and exons (Figure 1), implying that rs115550680 is in dis-equilibrium with exonic variants that could be potentially causative. Thus, although the findings of this study require replication in an independent African American sample with enough power to detect small ORs as well as functional confirmation, support for our findings comes from the previous studies in whites observing *ABCA7* as a risk locus in Alzheimer disease, albeit with marginal effects.[3,5]

If validated by future replication and functional studies, identification of *ABCA7* as a risk gene in LOAD among African Americans not only may help elucidate the disease etiology but also may have major implications for developing targets for genetic testing, prevention, and treatment. *ABCA7* is an integral transmembrane adenosine triphosphate–binding cassette transporter that belongs to the ABC family proteins and that mediates the biogenesis of high-density lipoprotein with cellular lipid and helical apolipoproteins.[48] It binds apolipoprotein A1 and functions in apolipoprotein-mediated phospholipid and cholesterol efflux from cells.[49]

The findings of the current study suggest that lipid metabolism is a prominent pathway of LOAD in African Americans. This is consistent with the fact that cardiovascular and

cerebrovascular diseases are more prominent in African Americans than in non-Hispanic whites.[50] Moreover, dyslipidemia and cardiovascular and cerebrovascular diseases are well-recognized risk factors for LOAD,[51,52] and the LOAD-related genes *SORL1*, *CLU*, and *APOE* are also involved in lipid metabolism. If confirmed, focusing on the role of lipid metabolism in LOAD may have significant effects on disease management.

ABCA7 also affects the transport of other important proteins, including amyloid precursor protein,[49] through the cell membrane and is involved in host defense through effects on phagocytosis by macrophages of apoptotic cells.[48] Thus, there are multiple ways in which *ABCA7* might affect risk of LOAD.

Compared with the findings described in Naj et al[5] among non-Hispanic whites, the area including significant SNP associations in the *ABCA7* region was broader in the African American sample. It is possible that this broad region of association in African Americans is attributable to a large, ancestral risk haplotype recently introduced by admixture with white ("European") Americans and has remained substantially intact within African Americans because of the relatively short time since its introduction. In contrast, the risk allele may exist on several different haplotypes in non-Hispanic whites (ie, may be older), only a subset of which was introduced into the African American population.

In a previous study,[5] the ADGC reported genome-wide associations for variants in *MSA4*, *CD2AP*, *CD33*, and *EPHA1* among individuals of white European ancestry. A cohort-based consortium comprising whites from the United Kingdom, Europe, and the United States had similar findings and first reported the association between SNPs in *ABCA7* and Alzheimer disease.[3] Logue et al[29] reported nominal significance for the *ABCA7* SNP rs3764650 reported by Hollingworth et al[3] in a well-characterized cohort of 513 African American persons with Alzheimer disease and 496 cognitively normal controls. As described above, the effect sizes for the association between *ABCA7* and LOAD in these studies is small compared with the effect size observed in the current study. In the current study *CR1*, *BIN1*, *EPHA1*, and *CD33* were replicated with significance in gene-based analyses. Differences in disease-associated SNPs in these loci between the white and African American consortium data sets also reflect differences in degree of variation and size of haplotype blocks, which in turn is helpful in identifying the true causative variants.

This study has limitations. Because of the paucity of available African American data sets for LOAD, we could not divide the assembled data sets into discovery and replication data sets but rather used the ADGC white race data set for replication. Thus, this study requires replication in an independent African American sample. In addition, we had limited power to detect associations with small effect sizes and associations with rare variants. Although all data sets included in the analytic sample used accepted clinical or pathological criteria to define LOAD, phenotypic heterogeneity between samples may have limited our ability to detect some associations.

In addition, the top-ranked SNP observed in *ABCA7* was not directly genotyped but imputed in all data sets. However, several facts make it unlikely that the observed association was caused by imputation error. First, as stated above and shown in Figure 1, rs115550680 is in linkage disequilibrium with the 2 *ABCA7* SNPs reported by Naj et al[5] and Hollingworth et al[3] in non-Hispanic whites of European ancestry (rs3764650 and rs3752246, 0.8<D' = 0.9) that make this finding plausible. Second, the imputation quality (R^2) of this SNP is high across all data sets (0.89<R^2<0.99) (eTable 3). Third, the MAF of rs115550680 in our African American sample is 7%. Although in general the imputation error rate increases with decreasing MAF, several recent studies suggest that SNPs with MAFs less than 5% are especially prone to imputation errors.[36] The recent study by

Hancock et al,[53] which specifically assessed genotype imputation performance using 1000 Genomes reference panels in African Americans, determined that the threshold for high imputation lies at MAF 2% or greater, applying the software and reference panel used in the present study.

The variant associations reported herein reflect a portion of the genetic influences of common alleles on LOAD in African Americans. Among these, *ABCA7* and *APOE* genotype were the strongest risk factors that both substantially increased the risk of LOAD (OR, 1.79 and 2.31, respectively). Identification of the genetic risk variants by resequencing and validation by functional studies would allow refinement of risk estimates and diagnostic and predictive testing protocols specific for African Americans.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support: 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.

Role of the Sponsor: The funding organizations and sponsors had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

REFERENCES

1. Fratiglioni L, De Ronchi D, Agüero-Torres H. Worldwide prevalence and incidence of dementia. Drugs Aging. 1999; 15(5):365–375. [PubMed: 10600044]
2. Slooter AJ, Cruts M, Kalmijn S, et al. Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. Arch Neurol. 1998; 55(7):964–968. [PubMed: 9678314]
3. Hollingworth P, Harold D, Sims R, et al. Alzheimer's Disease Neuroimaging Initiative; CHARGE Consortium; EADI1 Consortium. Common variants at *ABCA7*, *MS4A6A/MS4A4E*, *EPHA1*, *CD33*

and *CD2AP* are associated with Alzheimer's disease. Nat Genet. 2011; 43(5):429–435. [PubMed: 21460840]

4. Lambert JC, Heath S, Even G, et al. European Alzheimer's Disease Initiative Investigators. Genome-wide association study identifies variants at *CLU* and *CR1* associated with Alzheimer's disease. Nat Genet. 2009; 41(10):1094–1099. [PubMed: 19734903]
5. Naj AC, Jun G, Beecham GW, et al. Common variants at *MS4A4/MS4A6E*, *CD2AP*, *CD33* and *EPHA1* are associated with late-onset Alzheimer's disease. Nat Genet. 2011; 43(5):436–441. [PubMed: 21460841]
6. Seshadri S, Fitzpatrick AL, Ikram MA, et al. CHARGE Consortium; GERAD1 Consortium; EADI1 Consortium. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA. 2010; 303(18):1832–1840. [PubMed: 20460622]
7. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at *CLU* and *PICALM* associated with Alzheimer's disease. Nat Genet. 2009; 41(10):1088–1093. [PubMed: 19734902]
8. Reitz C, Cheng R, Rogaeva E, et al. Genetic and Environmental Risk in Alzheimer Disease 1 Consortium. Meta-analysis of the association between variants in *SORL1* and Alzheimer disease. Arch Neurol. 2011; 68(1):99–106. [PubMed: 21220680]
9. Rogaeva E, Meng Y, Lee JH, et al. The neuronal sortilin-related receptor *SORL1* is genetically associated with Alzheimer disease. Nat Genet. 2007; 39(2):168–177. [PubMed: 17220890]
10. Gatz M, Pedersen NL, Berg S, et al. Heritability for Alzheimer's disease: the study of dementia in Swedish twins. J Gerontol A Biol Sci Med Sci. 1997; 52(2):M117–M125. [PubMed: 9060980]
11. Tang MX, Cross P, Andrews H, et al. Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. Neurology. 2001; 56(1):49–56. [PubMed: 11148235]
12. Tang MX, Stern Y, Marder K, et al. The APOE-epsilon4 allele and the risk of Alzheimer disease among African Americans, whites, and Hispanics. JAMA. 1998; 279(10):751–755. [PubMed: 9508150]
13. Barnes LL, Shah RC, Aggarwal NT, Bennett DA, Schneider JA. The Minority Aging Research Study: ongoing efforts to obtain brain donation in African Americans without dementia. Curr Alzheimer Res. 2012; 9(6):734–745. [PubMed: 22471868]
14. Beecham GW, Martin ER, Li YJ, et al. Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. Am J Hum Genet. 2009; 84(1):35–43. [PubMed: 19118814]
15. Bennett DA, Schneider JA, Bienias JL, Evans DA, Wilson RS. Mild cognitive impairment is related to Alzheimer disease pathology and cerebral infarctions. Neurology. 2005; 64(5):834–841. [PubMed: 15753419]
16. Bennett DA, Schneider JA, Buchman AS, Mendes de Leon C, Bienias JL, Wilson RS. The Rush Memory and Aging Project: study design and baseline characteristics of the study cohort. Neuroepidemiology. 2005; 25(4):163–175. [PubMed: 16103727]
17. Bennett DA, Wilson RS, Schneider JA, et al. Natural history of mild cognitive impairment in older persons. Neurology. 2002; 59(2):198–205. [PubMed: 12136057]
18. Berg L, McKeel DW Jr, Miller JP, et al. Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype. Arch Neurol. 1998; 55(3):326–335. [PubMed: 9520006]
19. Bienias JL, Beckett LA, Bennett DA, Wilson RS, Evans DA. Design of the Chicago Health and Aging Project (CHAP). J Alzheimers Dis. 2003; 5(5):349–355. [PubMed: 14646025]
20. Carrasquillo MM, Zou F, Pankratz VS, et al. Genetic variation in *PCDH11X* is associated with susceptibility to late-onset Alzheimer's disease. Nat Genet. 2009; 41(2):192–198. [PubMed: 19136949]
21. Edwards TL, Scott WK, Almonte C, et al. Genome-wide association study confirms SNPs in *SNCA* and the *MAPT* region as common risk factors for Parkinson disease. Ann Hum Genet. 2010; 74(2):97–109. [PubMed: 20070850]

22. Evans DA, Bennett DA, Wilson RS, et al. Incidence of Alzheimer disease in a biracial urban community: relation to apolipoprotein E allele status. Arch Neurol. 2003; 60(2):185–189. [PubMed: 12580702]

23. Green RC, Cupples LA, Go R, et al. MIRAGE Study Group. Risk of dementia among white and African American relatives of patients with Alzheimer disease. JAMA. 2002; 287(3):329–336. [PubMed: 11790212]

24. Haroutunian V, Perl DP, Purohit DP, et al. Regional distribution of neuritic plaques in the nondemented elderly and subjects with very mild Alzheimer disease. Arch Neurol. 1998; 55(9): 1185–1191. [PubMed: 9740112]

25. Jun G, Naj AC, Beecham GW, et al. Alzheimer's Disease Genetics Consortium. Meta-analysis confirms *CR1*, *CLU*, and *PICALM* as Alzheimer disease risk loci and reveals interactions with *APOE* genotypes. Arch Neurol. 2010; 67(12):1473–1484. [PubMed: 20697030]

26. Kamboh MI, Minster RL, Demirci FY, et al. Association of *CLU* and *PICALM* variants with Alzheimer's disease. Neurobiol Aging. 2012; 33(3):518–521. [PubMed: 20570404]

27. Kukull WA, Higdon R, Bowen JD, et al. Dementia and Alzheimer disease incidence: a prospective cohort study. Arch Neurol. 2002; 59(11):1737–1746. [PubMed: 12433261]

28. Lee JH, Cheng R, Graff-Radford N, Foroud T, Mayeux R, National Institute on Aging Late-Onset Alzheimer's Disease Family Study Group. Analyses of the National Institute on Aging Late-Onset Alzheimer's Disease Family Study: implication of additional loci. Arch Neurol. 2008; 65(11): 1518–1526. [PubMed: 19001172]

29. Logue MW, Schu M, Vardarajan BN, et al. Multi-Institutional Research on Alzheimer Genetic Epidemiology (MIRAGE) Study Group. A comprehensive genetic association study of Alzheimer disease in African Americans. Arch Neurol. 2011; 68(12):1569–1579. [PubMed: 22159054]

30. Morris JC, Roe CM, Xiong C, et al. *APOE* predicts amyloid-beta but not tau Alzheimer pathology in cognitively normal aging. Ann Neurol. 2010; 67(1):122–131. [PubMed: 20186853]

31. Morris JC, Weintraub S, Chui HC, et al. The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer disease centers. Alzheimer Dis Assoc Disord. 2006; 20(4):210–216. [PubMed: 17132964]

32. Murrell JR, Price B, Lane KA, et al. Association of apolipoprotein E genotype and Alzheimer disease in African Americans. Arch Neurol. 2006; 63(3):431–434. [PubMed: 16533971]

33. Scott WK, Nance MA, Watts RL, et al. Complete genomic screen in Parkinson disease: evidence for multiple genes. JAMA. 2001; 286(18):2239–2244. [PubMed: 11710888]

34. Storandt M, Grant EA, Miller JP, Morris JC. Longitudinal course and neuropathologic outcomes in original vs revised MCI and in pre-MCI. Neurology. 2006; 67(3):467–473. [PubMed: 16894109]

35. Meier IB, Manly JJ, Provenzano FA, et al. White matter predictors of cognitive functioning in older adults. J Int Neuropsychol Soc. 2012; 18(3):414–427. [PubMed: 22390883]

36. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984; 34(7):939–944. [PubMed: 6610841]

37. US Census Bureau. 1990 Census of Population and Housing Summary File 1. US Census Bureau; Washington, DC: 1991.

38. Wittwer CT, Ririe KM, Andrew RV, David DA, Gundry RA, Balis UJ. The LightCycler: a microvolume multisample fluorimeter with rapid temperature control. Biotechniques. 1997; 22(1): 176–181. [PubMed: 8994665]

39. Ahmadian A, Gharizadeh B, Gustafsson AC, et al. Single-nucleotide polymorphism analysis by pyrosequencing. Anal Biochem. 2000; 280(1):103–110. [PubMed: 10805527]

40. Hixson JE, Vernier DT. Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with HhaI. J Lipid Res. 1990; 31(3):545–548. [PubMed: 2341813]

41. Lai E, Riley J, Purvis I, Roses A. A 4-mb high-density single nucleotide polymorphism−based map around human *APOE*. Genomics. 1998; 54(1):31–38. [PubMed: 9806827]

42. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38(8): 904–909. [PubMed: 16862161]

43. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012; 44(8):955–959. [PubMed: 22820512]
44. Chanda P, Yuhki N, Li M, et al. Comprehensive evaluation of imputation performance in African Americans. J Hum Genet. 2012; 57(7):411–421. [PubMed: 22648186]
45. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002; 21(11): 1539–1558. [PubMed: 12111919]
46. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003; 327(7414):557–560. [PubMed: 12958120]
47. Liu JZ, McRae AF, Nyholt DR, et al. AMFS Investigators. A versatile gene-based test for genome-wide association studies. Am J Hum Genet. 2010; 87(1):139–145. [PubMed: 20598278]
48. Tanaka N, Abe-Dohmae S, Iwamoto N, Yokoyama S. Roles of ATP-binding cassette transporter A7 in cholesterol homeostasis and host defense system. J Atheroscler Thromb. 2011; 18(4):274–281. [PubMed: 21173549]
49. Chan SL, Kim WS, Kwok JB, et al. ATP-binding cassette transporter A7 regulates processing of amyloid precursor protein in vitro. J Neurochem. 2008; 106(2):793–804. [PubMed: 18429932]
50. Roger VL, Go AS, Lloyd-Jones DM, et al. American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2012 update: a report from the American Heart Association. Circulation. 2012; 125(1):e2–e220. [PubMed: 22179539]
51. Breteler MM. Vascular risk factors for Alzheimer's disease: an epidemiologic perspective. Neurobiol Aging. 2000; 21(2):153–160. [PubMed: 10867200]
52. Shepardson NE, Shankar GM, Selkoe DJ. Cholesterol level and statin use in Alzheimer disease, I: review of epidemiological and preclinical studies. Arch Neurol. 2011; 68(10):1239–1244. [PubMed: 21987540]
53. Hancock DB, Levy JL, Gaddis NC, et al. Assessment of genotype imputation performance using 1000 Genomes in African American studies. PLoS One. 2012; 7(11):e50610. [PubMed: 23226329]

Box. Glossary of Terms

Genome-wide analysis: A genetic study evaluating the potential linkage of genetic markers located throughout the genome to a specific trait. This approach has been used for mendelian (single-gene) disorders as well as complex traits (genome-wide association study).

Haplotype: The combination of linked marker alleles (may be polymorphisms or mutations) for a given region of DNA on a single chromosome.

Imputation: A statistical method for inferring genotypes that are not directly measured.

Linkage disequilibrium: Refers to alleles at loci close enough together that they remain inherited together through many generations because their extreme close proximity makes recombination (crossing over) between them highly unlikely.

For a complete list of genomic terms, see the Appendix in this issue.

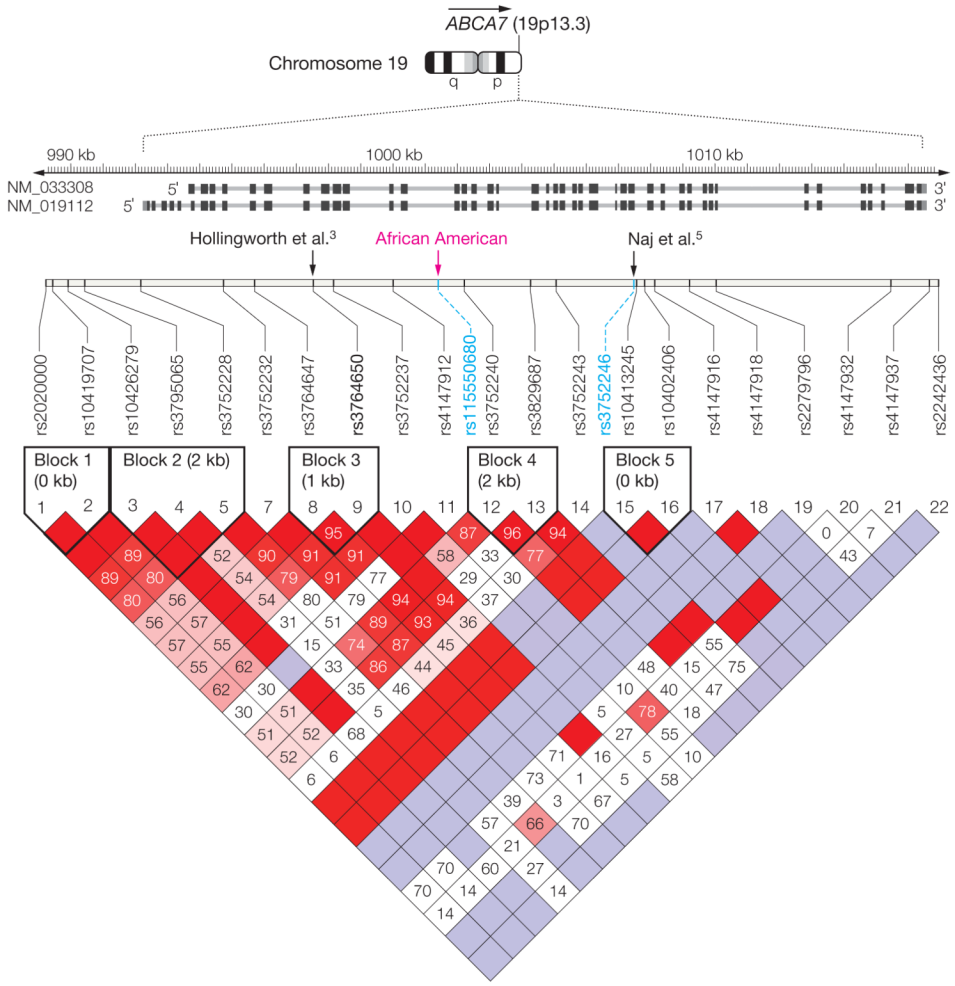

Figure 1.

Linkage Disequilibrium Pattern of Single-Nucleotide Polymorphisms in *ABCA7* Based on the HapMap Reference Sample (African Americans in the Southwest USA) and NCBI36/hg18 Genome Build

Black arrows indicate single-nucleotide polymorphisms (SNPs) previously reported to be associated with Alzheimer disease in whites[3,5] (the top hit reported by Hollingworth et al [rs3752228] may have changed if the entire cohort had been genotyped in stages 2 and 3). Pink arrow indicates the location of rs115550680 associated with Alzheimer disease in the present study.The SNPs shown in blue are not represented in HapMap. Kb indicates kilobase.

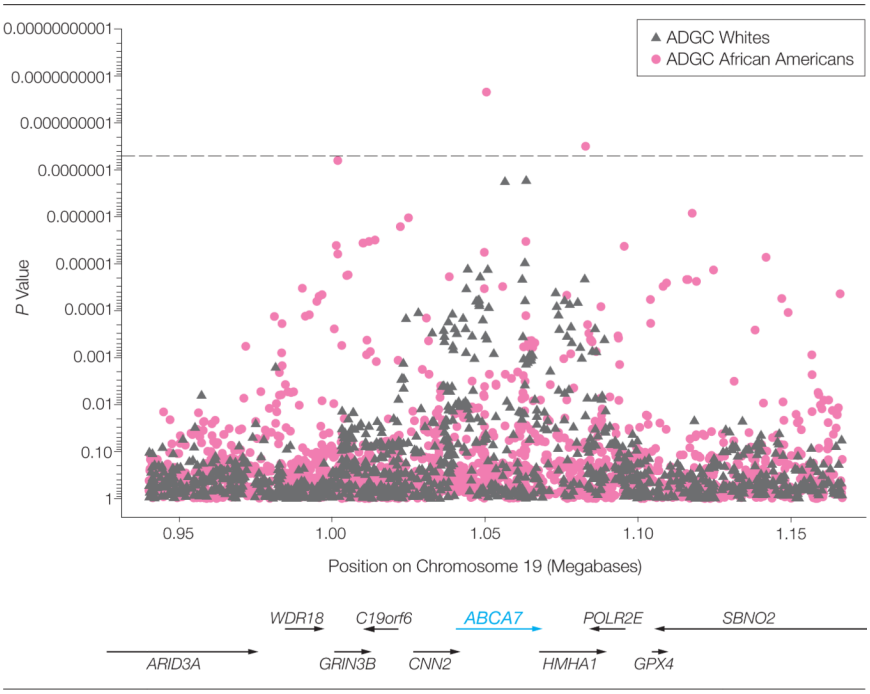

Figure 2.

Regional Association Plot of the *ABCA7* Region (±100 Kilobases) in the African American Sample and the White Sample Described in Naj et al[5] Based on the GRCh37/hg19 Genome Build

Dashed black line indicates the threshold typically applied in genome-wide association studies for genome-wide significance ($P \leq 5 \times 10^{-8}$).

Table 1

Characteristics of Data Sets[12-35]

Characteristic	No. (%)										
	ACT	ADC1/ADC2	ADC3	CHAP	Indianapolis	NIA-LOAD/NCRAD	ADGC[a]	Mirage 300k	Mirage 660k	GenerAAtions	Total No. of Participants
Individuals											
Affected	32 (33.0)	59 (44.7)	166 (59.7)	115 (20.9)	173 (14.7)	35 (36.5)	907 (35.1)	51 (44.0)	188 (44.3)	242 (54.3)	1968
Unaffected	65 (67.0)	73 (55.3)	112 (40.3)	435 (79.1)	1002 (85.3)	61 (63.5)	1675 (64.9)	65 (56.0)	236 (55.7)	204 (45.7)	3928
Women	62 (63.9)	94 (71.2)	211 (75.9)	362 (65.8)	771 (65.6)	70 (72.9)	1879 (72.8)	81 (69.1)	305 (71.9)	260 (58.3)	4095
Age at last evaluation, mean (SD)	80.5 (6.1)	74.2 (7.6)	77.6 (7.8)	78.8 (6.7)	83.0 (5.5)	73.9 (6.8)	75.6 (8.5)	69.5 (13.9)	71.4 (9.4)	79.4 (6.7)	
APOE genotype											
–/–[b]	57 (58.8)	59 (44.7)	101 (36.3)	328 (59.6)	748 (63.7)	46 (47.9)	1362 (52.7)	42 (36.2)	190 (44.8)	206 (46.2)	3139
–/4	32 (33.0)	58 (43.9)	117 (42.1)	194 (35.3)	373 (31.7)	39 (40.6)	810 (31.4)	61 (52.6)	183 (43.2)	175 (39.2)	2042
4/4	4 (4.1)	10 (7.6)	21 (7.6)	17 (3.1)	54 (4.6)	11 (11.5)	131 (5.1)	13 (11.2)	49 (11.5)	32 (7.2)	342
Missing	4 (4.1)	5 (3.8)	39 (14.0)	11 (2.0)	0	0	225 (8.7)	0	2 (0.5)	33 (7.4)	319

Abbreviations: ACT, Adult Changes in Thought; ADC, Alzheimer Disease Center; ADGC, Alzheimer Disease Genetics Consortium; APOE, apolipoprotein E; CHAP, Chicago Health and Aging Project; NIA-LOAD/NCRAD, National Institute on Aging–Late-Onset Alzheimer Disease/National Cell Repository for Alzheimer's Disease.

[a]Samples genotyped by the ADGC for this project were received from the African American Genetics Study, the ADCs, CHAP, Mayo Clinic, Mount Sinai School of Medicine, NIA-LOAD/NCRAD, Religious Orders Study/Rush Memory and Aging Project/Minority Aging Research Study/Clinical Minority Core, University of Miami/Vanderbilt University, University of Pittsburgh, Washington Heights Columbia Aging Project, and Washington University.

[b]All no-APOE*4–containing genotypes (*APOE* 3/3, *APOE* 2/3, *APOE* 2/2).

Table 2

Genome-Wide Meta-analysis Results of Fully Adjusted Model for Single-Nucleotide Polymorphisms with $P \leq 10^{-8}$ Excluding the *APOE* Region[a,b]

Gene	SNP	Chromosome	Base-Pair Position	Function	Allele 1	Allele 2	MAF	OR (95% CI)	*P* Value
ABCA7	rs115550680	19	1 050 420	Intron	G	A	0.07	1.79 (1.47-2.12)	2.21×10^{-9}
HMHA1	rs115553053	19	1082 844	Coding-synonymous	T	C	0.06	1.86 (1.49-2.32)	3.14×10^{-8}
GRIN3B	rs115882880	19	1 001 777	Intron	A	C	0.11	1.55 (1.32-1.81)	6.34×10^{-8}
–	rs145848414	5	174 014 114	Intergenic	A	G	0.04	2.29 (1.69-3.09)	6.90×10^{-8}

Abbreviations: MAF, minor allele frequency; OR, odds ratio; SNP, single-nucleotide polymorphism.

[a]Adjusted for age, sex, *APOE* genotype, and population stratification.

[b]Odds ratio greater than 1 for all data sets except Mirage300k and Mirage660k, which were not included in the meta-analyses because rs145848414 on chromosome 5 did not pass the minor allele frequency cutoff during the postimputation quality control. The direction of effects in the individual data sets is in the following order: ACT, ADC1 + 2, ADC3, CHAP, ADGC, GenerAAtions, Indianapolis, NIA-LOAD/NCRAD, Mirage300k, Mirage660k.

Table 3

Known Genetic Loci Associated With Alzheimer Disease in the African American Data Set in a Versatile Gene-Based Association Study (VEGAS[47])[a]

Gene	SNP	Chromosome	Base-Pair Location[b]	Risk Allele	MAF	No. of Tagging SNPs	*P* Value Threshold for Significance	Smallest *P* Value Detected in Current Data Set
CR1	rs146366639	1	207 649 473-207 835 110	G	0.26	44	.001	.0005
BIN1	rs55636820	2	127 785 603-127 884 931	G	0.02	38	.001	.0007
EPHA1	rs6973770	7	143 067 382-143 125 985	G	0.06	30	.002	.001
CD33	rs114282264	19	51 708 320-51 763 274	G	0.03	24	.002	.0007

Abbreviations: MAF, minor allele frequency; SNP, single-nucleotide polymorphism.

[a]Adjusted for age, sex, *APOE* genotype, and population stratification. VEGAS incorporates linkage disequilibrium information for the assessed genes from a set of reference individuals from HapMap, determines the number of tagging SNPs, and calculates the empirical *P* value for the gene by using simulations from the multivariate normal distribution.

[b]Including 20 kilobases to each side; based on genome build 37.3.