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
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

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 × 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 × 10⁻⁹). 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 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.

JAMA. 2013;309(14):1483-1492

www.jama.com
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. As much as 20% of the disease-attributable risk is related to the ε4 variant in APOE. 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/MS4A4E, and ABCA7. In addition, SORL1 was identified as a susceptibility gene in candidate gene and functional studies. However, LOAD heritability estimates are high (h² ≈ 60%-80%), and a large part of the genetic contribution to LOAD remains unexplained. The incidence of LOAD among African Americans is higher than among whites living in the same community, and the reported risk for the disease associated with APOE ε4 heterozygosity is inconsistent in African Americans compared with whites. 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. 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. Classification of participants as African American was based on self-report using the format of the 1990 US census. 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 GenerAAttions cohorts, APOE genotypes were determined using the Roche Diagnostics LightCycler 480 instrument (Roche Diagnostics) 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 or analysis of restriction fragment length polymorphisms; 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.

**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.

1484 JAMA, April 10, 2013—Vol 309, No. 14

©2013 American Medical Association. All rights reserved.
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-chromosome 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 ($\pi$) 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 ($\pi$) using PLINK. Where $P$ indicates probability. One participant from each duplicate male 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). 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 were 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. 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. 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 ($0 \sim 2$). 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 $x^2$ distributed $Q$ statistic. 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 ($\lambda$) 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 (Figure 1). All findings with $P \leq 10^{-5}$ in the fully adjusted model were compared with results obtained in whites.
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 R², 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 17332474 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 × 10⁻¹⁰) 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 × 10⁻⁷), rs885330 in SOX13 (NCBI Entrez Gene 9580) (OR, 1.25 [95% CI, 1.17-1.33]; P = 3.9 × 10⁻⁷), an intergenic SNP (rs145848414) at 174014114 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 × 10⁻⁷), and rs115500680 in ABCA7 (NCBI Entrez Gene 10347) (OR, 1.78 [95% CI, 1.28-1.82]; P = 1.4 × 10⁻⁶). After adjustment for APOE, the associations with ELMO1 and SOX13 SNPs diminished, whereas the association for rs115500680 in ABCA7 (OR, 1.79 [95% CI, 1.47-2.12]; P = 2.21 × 10⁻⁸) 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 × 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 (rs115500680) is in linkage disequilibrium 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] and rs3752246 [Naj et al], 0.8 < D’ = 0.9) (FIGURE 1) and showed the same direction of effect. The effect size for rs115500680 in ABCA7 (OR, 1.79 [95% CI, 1.47-2.12]; P = 2.21 × 10⁻⁸) was

Table 1. Characteristics of Data Sets

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ACT</th>
<th>ADC1/ ADC2</th>
<th>ADC3</th>
<th>CHAP</th>
<th>Indianapolis</th>
<th>NIA-LOAD/ NCRAD</th>
<th>ADGC</th>
<th>Mirage 300K</th>
<th>Mirage 660K</th>
<th>GenerAAttions</th>
<th>Total No. of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected</td>
<td>32 (33.0)</td>
<td>59 (44.7)</td>
<td>166 (59.7)</td>
<td>115 (20.9)</td>
<td>173 (14.7)</td>
<td>35 (36.5)</td>
<td>907 (35.1)</td>
<td>61 (56.0)</td>
<td>236 (55.7)</td>
<td>204 (45.7)</td>
<td>1968</td>
</tr>
<tr>
<td>Unaffected</td>
<td>65 (67.0)</td>
<td>73 (55.3)</td>
<td>112 (40.3)</td>
<td>435 (79.1)</td>
<td>1002 (85.3)</td>
<td>61 (63.5)</td>
<td>1675 (64.9)</td>
<td>242 (54.3)</td>
<td>175 (39.2)</td>
<td>2062 (60.3)</td>
<td>4095</td>
</tr>
<tr>
<td>Women</td>
<td>62 (63.9)</td>
<td>94 (71.2)</td>
<td>211 (75.9)</td>
<td>362 (65.8)</td>
<td>771 (65.6)</td>
<td>70 (72.9)</td>
<td>1879 (72.8)</td>
<td>81 (69.1)</td>
<td>305 (71.9)</td>
<td>260 (58.3)</td>
<td>4095</td>
</tr>
<tr>
<td>Age at last evaluation, mean (SD)</td>
<td>80.5 (6.1)</td>
<td>74.2 (7.6)</td>
<td>77.6 (7.8)</td>
<td>78.8 (6.7)</td>
<td>83.0 (5.5)</td>
<td>73.9 (6.8)</td>
<td>75.6 (6.5)</td>
<td>69.5 (13.9)</td>
<td>71.4 (9.4)</td>
<td>79.4 (6.7)</td>
<td></td>
</tr>
<tr>
<td>APOE genotype</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/4</td>
<td>57 (58.8)</td>
<td>59 (44.7)</td>
<td>101 (36.3)</td>
<td>328 (59.6)</td>
<td>748 (63.7)</td>
<td>46 (47.9)</td>
<td>1362 (52.7)</td>
<td>42 (36.2)</td>
<td>190 (44.8)</td>
<td>206 (46.2)</td>
<td>3139</td>
</tr>
<tr>
<td>4/4</td>
<td>32 (33.0)</td>
<td>58 (43.9)</td>
<td>117 (42.1)</td>
<td>194 (35.3)</td>
<td>373 (31.7)</td>
<td>39 (40.6)</td>
<td>810 (31.4)</td>
<td>61 (52.6)</td>
<td>183 (43.2)</td>
<td>175 (39.2)</td>
<td>2042</td>
</tr>
<tr>
<td>4/4</td>
<td>4 (4.1)</td>
<td>10 (7.6)</td>
<td>21 (7.8)</td>
<td>17 (3.1)</td>
<td>54 (4.8)</td>
<td>11 (11.5)</td>
<td>131 (5.1)</td>
<td>13 (11.2)</td>
<td>49 (11.5)</td>
<td>32 (7.2)</td>
<td>342</td>
</tr>
<tr>
<td>Missing</td>
<td>4 (4.1)</td>
<td>5 (3.8)</td>
<td>39 (14.0)</td>
<td>11 (2.0)</td>
<td>0</td>
<td>0</td>
<td>225 (8.7)</td>
<td>0</td>
<td>2 (0.5)</td>
<td>33 (7.4)</td>
<td>319</td>
</tr>
</tbody>
</table>

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.

aSamples 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.

bAll no-APOE 4–containing genotypes (APOE 3/3, APOE 3/2, APOE 2/2).
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

### Table 2. Genome-Wide Meta-analysis Results of Fully Adjusted Model for Single-Nucleotide Polymorphisms with $P \leq 10^{-8}$ Excluding the APOE Region

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Chromosome</th>
<th>Base-Pair Position</th>
<th>Function</th>
<th>Allele 1</th>
<th>Allele 2</th>
<th>MAF</th>
<th>OR (95% CI)</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA7</td>
<td>rs115550680</td>
<td>19</td>
<td>1,050,420</td>
<td>Intron</td>
<td>G</td>
<td>A</td>
<td>0.07</td>
<td>1.79 (1.47-2.12)</td>
<td>2.21 x 10^{-3}</td>
</tr>
<tr>
<td>HMHA1</td>
<td>rs115553053</td>
<td>19</td>
<td>1,082,844</td>
<td>Coding-synonymous</td>
<td>T</td>
<td>C</td>
<td>0.06</td>
<td>1.86 (1.49-2.32)</td>
<td>3.14 x 10^{-3}</td>
</tr>
<tr>
<td>GRIN3B</td>
<td>rs115882880</td>
<td>19</td>
<td>1,001,777</td>
<td>Intron</td>
<td>A</td>
<td>G</td>
<td>0.11</td>
<td>1.55 (1.32-1.81)</td>
<td>6.34 x 10^{-6}</td>
</tr>
<tr>
<td>-</td>
<td>rs145848414</td>
<td>5</td>
<td>1,740,114</td>
<td>Intergenic</td>
<td>A</td>
<td>G</td>
<td>0.04</td>
<td>2.29 (1.69-3.09)</td>
<td>6.90 x 10^{-6}</td>
</tr>
</tbody>
</table>

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.

---

**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$^{4,5}$ (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 rs11559680 associated with Alzheimer disease in the present study. The SNPs shown in blue are not represented in HapMap. Kb indicates kilobase.
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 ≤ 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 disequilibrium (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 and rs3752246) did not reach genome-wide significance in this African American data set. However, the MAFs for these 2 variants 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 (CRI, BIN1, PICALM, CLU, EPHA1, MS4A4 cluster, CD2AP, CD33), were further explored in gene-based analyses adding 50 kb to both sides of each gene.3,5 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 CRI (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 e4. 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]),5 as do all other genes reported in whites (CRI, BIN1, PICALM, CLU, EPHA1, MS4A4 cluster, CD2AP, CD33).5,6

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...
Table 3. Known Genetic Loci Associated With Alzheimer Disease in the African American Data Set in a Versatile Gene-Based Association Study (VEGAS)\(^a\)

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Chromosome</th>
<th>Base-Pair Location(^b)</th>
<th>Risk Allele</th>
<th>MAF</th>
<th>No. of Tagging SNPs</th>
<th>(P) Value Threshold for Significance</th>
<th>Smallest (P) Value Detected in Current Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR1</td>
<td>rs146366639</td>
<td>1</td>
<td>207 649 473-207 835 110</td>
<td>G</td>
<td>0.26</td>
<td>44</td>
<td>0.001</td>
<td>0.005</td>
</tr>
<tr>
<td>BIN1</td>
<td>rs5563820</td>
<td>2</td>
<td>127 785 603-127 884 931</td>
<td>G</td>
<td>0.02</td>
<td>38</td>
<td>0.001</td>
<td>0.0007</td>
</tr>
<tr>
<td>EPHA1</td>
<td>rs6973770</td>
<td>7</td>
<td>143 067 382-143 125 985</td>
<td>G</td>
<td>0.06</td>
<td>30</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>CD33</td>
<td>rs11428264</td>
<td>19</td>
<td>51 708 320-51 763 274</td>
<td>G</td>
<td>0.03</td>
<td>24</td>
<td>0.002</td>
<td>0.0007</td>
</tr>
</tbody>
</table>

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.

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 disequilibrium 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 transporter that belongs to the ABC family proteins and that mediates the biogenesis of high-density lipoprotein with cellular lipid and helical apolipoproteins.\(^47\) It binds apolipoprotein A1 and functions in apolipoprotein-mediated phospholipid and cholesterol efflux from cells.\(^48\)

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.\(^30\) Moreover, dyslipidemia and cardiovascular and cerebrovascular diseases are well-recognized risk factors for LOAD.\(^31,32\) 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\(^3\) 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,\(^3\) 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

©2013 American Medical Association. All rights reserved.
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, rs11550680 is in linkage disequilibrium with the 2 ABCA7 SNPs reported by Naj et al. and Hollingworth et al. in non-Hispanic whites of European ancestry (rs3764650 and rs3752246, $0.8 < R^2 < 0.9$) 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 rs11550680 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. The recent study by Hancock et al, 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 sequencing and validation by functional studies would allow refinement of risk estimates and diagnostic and predictive testing protocols specific for African Americans.

Author Affiliations: Taub Institute for Research on Alzheimer’s Disease and the Aging Brain (Dr Reitz, Manly, and Mayeux); Gertrude H. Sergievsky Center (Dr Reitz, and Mayeux), and Departments of Neurology (Dr 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) (Dr Jun, Vardarajan, Logue, Baldwin, and Farrer), Biostatistics (Dr Jun, Vardarajan, Lunetta, and Farrer), Pathology (Dr Jun and Farrer), and Neurology (Dr Farrer), Boston University, Boston, Massachusetts; The John P. Hussman Institute for Human Genetics, University of Miami, Miami, Florida (Dr 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 Mt. Valladares and Mr. Cantwell); Department of Medicine (Dr 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 (Dr Graff-Radford, and Ertekin-Taner) and Neurology (Dr Graff-Radford and Ertekin-Taner), Mayo Clinic, Jacksonville, Florida; Reid Institute for Healthy Aging, Department of Internal Medicine (Dr Evans), Departments of Neurological Sciences (Drs Barnes and Bennett) and Sciences (Dr Barnes), and Rush Alzheimer’s Disease Center (Dr Bennett), Rush University Medical Center, Chicago, Illinois; Program in Translational Neuropsychiatric Genomics, Department of Psychiatry & Neurology, Massachusetts General 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 Medicine, Baltimore, Maryland (Dr Fallin); School of Public Health, University of Alabama at Birmingham (Dr Go); Department of Neurology, Meharry Medical College, Nashville, Tennessee (Dr Griffin); 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; Center for Psychiatric 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).

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, Reitz, Griffith, Obisesan, Manly, Bennett, Haines, Farrer, Pericak-Vance, Schellenberg, Mayeux.

Acquisition of data: Reitz, Vardarajan, Lin, Larson, Graff-Radford, Evans, Crane, Buxbaum, Murrell, Ertekin-Taner, Baldwin, Green, Barnes, Cantwell, Fallin, Go, Griffith, Obisesan, Mann, Lunetta, Kamboh, Lopez, Buxbaum, Murrell, Reitz, Foroud, Haines, Farrer, Pericak-Vance, Schellenberg, Mayeux.


Drafing 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, Vardarajan, Buxbaum, Hall, Goate, Byrd, Murrell, Rajbhandary, Foroud, Haines, Farrer, Pericak-Vance, Schellenberg, Mayeux.

Statistical analysis: Reitz, Jun, Rajbhandary, Buxbaum, Manly, Lunetta, 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, Murrell, Foroud, Haines, Farrer, Pericak-Vance, 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. Dr Lopez reported grants or contracts 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 Merz 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 royalties from Taconic for a tau mutation. Dr Pericak-Vance reported receiving revenues from Athena Diagnostics. No other authors reported disclosures.

Funding/Support: The NIH supported this work through grants U01-AG303984, R2C-AG36528, U01-AG106761 (Dr Kukull); U24 AG026395, U24 AG026390, R01 AG037212, R01 AG015473 (Dr Mayeux); K23 AG09049 (Dr Reitz); U24 AG021869 (Dr Foroud); R01 AG009956, RC2 AG036650 (Dr Hall); U01 AG006781, U01 HG004610 (Dr Larson); R01 AG009029 (Dr Farrer); R01 AG020688 (Dr Manly); P50 AG005133, AG030653 (Dr Kukull); R01 AG019085 (Dr Haines); R01 AG1101, R01 AG030146, R01 AG36560 (Dr Evans); P30 AG010161, R01 AG15819, R01 AG03146, R01 AG17917, R01 AG15819 (Dr Bennett); R01 AG028786 (Dr Manly); R01 AG20218, R01 AG10161 (Dr Barnes); P30 AG16574 (Dr Ertekin-Taner, Dr Graff-Radford), R01 AG32990 (Dr Ertekin-Taner), KL2 RR024515 (Dr Ertekin-Taner), R01 AG027944, R01 AG028786 (Dr Pericak-Vance); P20 MD005546, R01 AG28786-01 A1 (Dr Byrd); R0501538 (Dr Buxbaum); P50 AG05681, P01 AG03991, P01 AG02676 (Dr Goate); P01 AG031860, P30 AG031860, R01 AG031860 (Dr Bal; R01 CA129769, R01 HM080295, R01 AG17173, R01 AG025259, R01 AG031933, P05 AG080702, 1490 JAMA, April 10, 2013—Vol 309, No. 14

Corrected on April 9, 2013

©2013 American Medical Association. All rights reserved.
LATE-ONSET ALZHEIMER DISEASE IN AFRICAN AMERICANS

References


©2013 American Medical Association. All rights reserved.

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.


