Transethnic genome-wide scan identifies novel Alzheimer's disease loci

Carlos Cruchaga
Washington University School of Medicine in St. Louis
et al

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Featured Article

Transethnic genome-wide scan identifies novel Alzheimer’s disease loci

Gyungah R. Jun, a, b, Jaeyoon Chung, b, Jesse Mez, c, Robert Barber, d, Gary W. Beecham, e, David A. Bennett, f, Joseph D. Buxbaum, g, h, i, Goldie S. Byrd, j, Minerva M. Carrazquillo, k, Paul K. Crane, l, Carlos Cruchaga, m, n, Philip De Jager, o, p, Nilufer Ertekin-Taner, k, Denis Evans, q, M. Danielle Fallin, r, Tatiana M. Foroud, s, Robert P. Friedland, t, Alison M. Goate, u, Neill R. Graff-Radford, v, Hugh Hendrie, w, Kathleen S. Hall, w, Kara L. Hamilton-Nelson, e, Rivka Inzelberg, x, M. Ilyas Kamboh, y, John S. K. Kauwe, z, Walter A. Kukull, aa, bh, Brian W. Kunkle, e, Ryozo Kuwano, cc, Eric B. Larson, p, dd, Mark W. Logue, b, le, Jennifer J. Manly, ff, gg, Eden R. Martin, e, Thomas J. Montine, hh, Shubhabrata Mukherjee, i, Adam Naj, ii, Eric M. Reiman, jj, kk, ll, mm, Christiane Reitz, mn, oo, pp, Richard Sherva, b, Peter H. St. George-Hyslop, qq, rr, Timothy Thornton, l, Steven G. Younkin, k, Badri N. Vardarajan, ff, gg, nn, Li-San Wang, ii, Jens R. Wendlund, ss, Ashley R. Winslow, ss, Alzheimer’s Disease Genetics Consortium, Jonathan Haines, tt, Richard Mayeux, ff, gg, nn, oo, pp, Margaret A. Pericak-Vance, e, Gerard Schellenberg, jj, Kathryn L. Lunetta, uu, Lindsay A. Farrer, v, w, vv, ww, xx, yy

aNeurogenetics and Integrated Genomics, Andover Innovative Medicines Institute, Eisai Inc, Andover, MA, USA
bDepartment of Medicine (Biomedical Genetics), Boston University Schools of Medicine, Boston, MA, USA
cDepartment of Neurology, Boston University Schools of Medicine, Boston, MA, USA
dDepartment of Pharmacology and Neuroscience, University of North Texas Health Science Center, Fort Worth, TX, USA
eThe John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, USA
fDepartment of Neurological Sciences and Rush Alzheimer’s Disease Center, Chicago, IL, USA
gDepartment of Neuroscience, Mount Sinai School of Medicine, New York, NY, USA
hDepartment of Psychiatry, Mount Sinai School of Medicine, New York, NY, USA
iDepartment of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY, USA
jDepartment of Biology, North Carolina A&T State University, Greensboro, NC, USA
kDepartment of Neuroscience, Mayo Clinic, Jacksonville, FL, USA
lDepartment of Medicine, University of Washington, Seattle, WA, USA
mHope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St Louis, MO, USA
nDepartment of Psychiatry, Washington University School of Medicine, St Louis, MO, USA
oProgram in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology & Psychiatry, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
pProgram in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA
qRush Institute for Healthy Aging, Department of Internal Medicine, Rush University Medical Center, Chicago, IL, USA
rDepartment of Mental Health, Johns Hopkins University School of Medicine, Baltimore, MD, USA
sDepartment of Medical & Molecular Genetics, Indiana University, Indianapolis, IN, USA
tDepartment of Neurology, University of Louisville, Louisville, KY, USA
uDepartment of Psychology, Indiana University, Indianapolis, IN, USA
vRegenstrief Institute, Inc, Indianapolis, IN, USA
wDepartment of Medicine, Indiana University, Indianapolis, IN, USA
xDepartment of Neurology and Neurosurgery, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel
yUniversity of Pittsburgh Alzheimer’s Disease Research Center and Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA, USA
zDepartment of Biology, Brigham Young University, Provo, UT, USA
aaDepartment of Epidemiology, University of Washington, Seattle, WA, USA

*Corresponding author. Tel.: (617) 638-5393; Fax: (617) 638-4275.
E-mail address: farrer@bu.edu

http://dx.doi.org/10.1016/j.jalz.2016.12.012
1552-5260/ © 2017 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer’s Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Abstract

Introduction: Genetic loci for Alzheimer’s disease (AD) have been identified in whites of European ancestry, but the genetic architecture of AD among other populations is less understood.

Methods: We conducted a transethnic genome-wide association study (GWAS) for late-onset AD in Stage 1 sample including whites of European Ancestry, African-Americans, Japanese, and Israeli-Arabs assembled by the Alzheimer’s Disease Genetics Consortium. Suggestive results from Stage 1 from novel loci were followed up using summarized results in the International Genomics Alzheimer’s Project GWAS dataset.

Results: Genome-wide significant (GWS) associations in single-nucleotide polymorphism (SNP)–based tests ($P < 5 \times 10^{-8}$) were identified for SNPs in $PFDN1/HBEGF$, $USP6NL/ECHIDC3$, and $B2RAP1$ and for the interaction of the (apolipoprotein E) $APOE$ e4 allele with $NFIC$ SNP. We also obtained GWS evidence ($P < 2.7 \times 10^{-8}$) for gene-based association in the total sample with a novel locus, $TPBG$ ($P = 1.8 \times 10^{-6}$).

Discussion: Our findings highlight the value of transethnic studies for identifying novel AD susceptibility loci.

© 2017 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer’s Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Transethnic; Alzheimer’s disease; Genome-wide association; $APOE$ interaction

1. Introduction

Alzheimer’s disease (AD) is the most prevalent neurodegenerative disease in persons aged 65 years and older and the sixth leading cause of death in the United States [1]. Total healthcare payments in 2014 for people aged 65 years and older with dementia are estimated at $214 billion [1]. By the middle of the century, the number of Americans with AD is projected to be 13.8 million with one new case developing every 33 seconds or almost one million new cases per year. The global burden of AD or dementia in 2015 is more daunting with new cases of dementia in every 3 seconds, and the estimated worldwide costs of dementia are about $818 billion, rising to $2 trillion by 2030 [2]. The number of people living with dementia in 2015 is estimated to be 9.4 million in the Americas, 10.5 million in Europe, 4.0 million in Africa, and 22.9 million in Asia [2]. This is a tremendous global epidemic in elderly persons regardless of ethnic background.

AD with onset age after 65 years is highly heritable with an estimated 74% of the liability explained by genetic factors [3]. A major genetic risk factor for AD is $APOE$ genotype [4] that accounts for approximately 35% of the genetic variance [5]. The three common apolipoprotein E ($APOE$) alleles ($\varepsilon2$, $\varepsilon3$, and $\varepsilon4$) are determined by combinations of polymorphic amino acid residues at Arg112 (rs429358) and Cys158 (rs7412) [6]. Among non-Hispanic whites of European ancestry (EA), $\varepsilon4$ heterozygotes have a 2.5- to 3.0-fold increased risk and $\varepsilon4$ homozygotes have a 10- to 12-fold increased risk, compared with persons with the $\varepsilon3/\varepsilon3$ genotype [4]. The $\varepsilon2$ allele is protective [7] such that carriers of this allele have a
40% reduction in AD risk compared with ε3/ε3 individuals [4]. The effect of APOE genotype to AD risk is highly variable in other populations. The ε4 frequency is lower in Asians [8] and associated with higher AD risk among Japanese (JPN) compared with EAs [9]. In contrast, the effect of ε4 on AD risk is much less in African-Americans (AAs) among whom the ε4 frequency is about 50% higher than in EAs [10]. It is noteworthy that the ε4 allele is virtually absent among Arabs living in northern Israeli community where the prevalence of dementia is roughly double than in EA populations [11].

More than 20 loci have been robustly associated with AD [12] and are enriched in immune response, regulation of endocytosis, cholesterol transport, and protein ubiquitination pathways [13]. A recent genome-wide association study (GWAS) identified significant association of AD with multiple single-nucleotide polymorphisms (SNPs) in the MAPT-KANSL1 region among EAs lacking an APOE ε4 allele [14]. Genetic studies in other populations have increased our understanding of the genetic architecture of AD. For example, the effect of the APOE ε4 allele is much greater in JPN and substantially weaker in AA and some Hispanic groups, due in part to varying frequencies of this allele across populations [4]. Three loci (SORL1, ABCA7, and ACE) whose association with AD attained genome-wide significance in EAs [12] were found to have larger effects on AD risk in AAs (ABCA7) [15], JPN (SORL1) [9], and Israeli-Arabs (IAs) (ACE) [16]. Some loci including PLXNA4 [17] and SORL1 [18] demonstrate allelic heterogeneity among genetically diverse populations. In the present study, we leveraged genetic diversity across ethnic groups to increase discovery of additional AD risk loci by combining GWAS results obtained from samples of EAs, AAs, JPN, and IAs.

2. Methods

2.1. Subjects, genotyping, and data processing

Details of subject recruitment and genotyping for individual case-control and family-based datasets, genotype imputation, quality control, population substructure, and statistical methods for association analyses were reported previously for Alzheimer’s Disease Genetics Consortium (ADGC) datasets containing EAs [5], AAs [15], JPN [9], and IAs [11]. Characteristics of the 33,269 ADGC subjects (26,320 EAs, 4983 AAs, 1845 JPN, and 115 IAs) used for discovery in Stage 1 were shown in Supplementary Table 1. Summarized results achieved in the NIA Genetics of Alzheimer’s Disease Data Storage Site (https://www.niagads.org) that are from a previous GWAS of EAs conducted by the International Genomics Alzheimer’s Project (IGAP) including 5813 AD cases and 20,474 controls after excluding the ADGC datasets [12] were used in Stage 2 follow-up analyses (Supplementary Table 1).

2.2. Genome-wide association analysis in Stage 1

2.2.1. Design and power considerations

The primary analysis was a single GWAS including all discovery datasets. Analyses were performed separately for each dataset, and the results were pooled sequentially, first within ethnic groups and then across ethnic groups. The minimum detectable genotype relative risk for EAs range from 1.16 for minor allele frequency (MAF) = 0.5 to 1.73 for MAF = 0.01. The corresponding ranges for AAs and JPN are 1.40–2.69 and 1.74–3.78, respectively. Genotype relative risks (GRRs) of <5 are not detectable with 80% power in the small IA sample. However, the goal of this study was not for novel discovery within ethnic groups but rather in the total transethnic sample. Prompted by findings of the previous studies [14], we also conducted separate GWAS in subgroups of subjects who have or lack an APOE ε4 allele. We also applied a complementary approach for assessing a differential effect of association by APOE genotype by evaluating association of AD with an interaction of SNP and ε4 status.

2.2.2. SNP-based association

Within each dataset, genome-wide association analyses were conducted using more than 7 million imputed SNPs in the total sample and in subgroups of subjects with and without the APOE ε4 allele, using regression models including age, sex, and the first three PCs. An additive effect of a SNP was included in the model as a quantitative estimate between 0 and 2 representing the probability score of the effect allele to incorporate the uncertainty of the imputation estimates. Models were evaluated using a logistic generalized linear model in case-control datasets and a logistic generalized estimating equation in family-based datasets. We also evaluated models including a term for the interaction of the SNP dosage with the APOE ε4 status and models among subgroups stratified by APOE ε4 status. Results for each model across datasets were combined by meta-analysis separately within each ethnic group using a fixed-effects, inverse-variance weighted meta-analysis in the METAL program [19]. SNPs with a minor allele frequency ≥1% and imputation quality ≥0.4 that were available in at least 50% of the datasets were included in the meta-analysis. The meta-analysis P-value for association was estimated by the summarized test statistic, after applying genomic control within each individual study. Meta-analysis was also conducted using Han and Eskin–modified random-effects (RE-HE) model that is optimized to detect associations under effect heterogeneity, as implemented in METASOFT [20]. This model has similar power to the fixed effects model when heterogeneity is modest, for example, when the standard deviation of the different ethnicities log odds ratios (ORs) is ≤0.5 times the mean log OR, but has better power than the fixed effects model for substantial heterogeneity. Thus, we do not expect the RE-HE model to produce substantially different results from the fixed effects
model unless substantial heterogeneity among ethnicities exists.

2.2.3. Gene-based association

We conducted genome-wide gene-based tests using ethnic-specific association results from SNP-based tests. Intragenic SNPs and SNPs within 30 kilobases (kb) of transcription start and stop sites were included in each gene-based test. We used the GATES [21] method, which computes a gene-based \( P \)-value using SNP-based \( P \) values and SNP-SNP correlations by penalizing lack of association in correlated SNPs. Ethnic-specific gene-based results for EA, AA, JPN, and IA groups were combined using the sample-size weighted Z-score method in METAL assuming the same direction of effect.

2.3. Follow-up association analysis

In Stage 2, we attempted to replicate Stage 1 top-ranked SNP-based \( P < 10^{-5} \) results and validate gene-based \( P < 10^{-4} \) results from each ethnic subgroup. Previously known AD genes were evaluated in Stage 2 only when both SNP-based and gene-based \( P \) values met threshold criteria for follow up. These analyses incorporated summarized results for the Stage 2 ADGC datasets and previously reported results for IGAP datasets excluding those from the ADGC that are described in Supplementary Table 1. The genome-wide significance threshold was set at \( P < 5 \times 10^{-8} \) for individual SNPs and \( P < 2 \times 10^{-6} \) for gene-based tests in the Stage 1 + 2 analyses.

3. Results

3.1. Findings with individual SNPs

There was little evidence for genomic inflation in SNP-based GWA results in the total sample with main effect (\( \lambda = 1.02 \)) and interaction effect of an SNP with APOE e4 status on AD risk (\( \lambda = 1.02 \)) and in APOE e4+ subjects (\( \lambda = 0.99 \)) and APOE e4−subjects (\( \lambda = 0.99 \)) (Supplementary Fig. 1). In the total sample, we confirmed genome-wide significant (GWS) association \(( P < 5 \times 10^{-8} \) with SNPs in several previously implicated AD loci including \( CR1, BIN1, PTK2B, MS4A2/MS4A6A, PICALM \) (Supplementary Table 2 and Supplementary Fig. 2). GWS association was also observed with SNPs in \( NFIC \) and \( PRKCE \) through interaction with APOE (Supplementary Fig. 2B) and with SNPs between \( USP6N \) and \( ECHDC3 \) among subjects lacking APOE e4 (Supplementary Fig. 2D). Top-ranked SNPs in EA for \( PICALM, SORLI \), and \( ABCA7 \) had strong support for association in JPN, whereas the top-ranked SNPs in \( CR1, BIN1, \) and \( EPHAI \) were consistently associated in EAs and AAs (Supplementary Table 2). In contrast, the effect direction was significantly opposite in EAs versus AAs for \( NME8, ABCA7, \) and \( CASS4 \) SNPs (Supplementary Table 2). A total of 35 SNPs from 9 novel loci met criteria for follow up in Stage 2 (Supplementary Table 3). Extensive evaluation of SNPs from the APOE region across the different ethnic groups demonstrated that only the APOE e2 SNP (rs7412) remained genome-wide significant among APOE e4−subjects (Supplementary Table 4), confirming our prior observation that APOE accounts for all association signals in this region [22]. SNPs in other loci showed suggestive evidence for association \(( P < 10^{-6} \) in EAs or AAs (Supplementary Table 5), but these results were much less significant in the transethnic meta-analyses. Analysis of models including an interaction term for each SNP with APOE e4 status identified a GWS significant interaction (interaction: \( P = 1.5 \times 10^{-8} \) for \( NFIC \) SNP rs9749589 (Table 1). This SNP appeared protective in e4+ subjects (OR = 0.83, \( P = 6.4 \times 10^{-6} \)) but slightly increased risk of AD in e4−subjects (OR = 1.11, \( P = 6.0 \times 10^{-3} \)) (Supplementary Table 6).

In the combined Stage 1 + 2 sample, GWS association was observed with SNPs in several previously established AD loci (\( CR1, BIN1, PTK2B, MS4A4A, \) and \( PICALM \))

### Table 1

<table>
<thead>
<tr>
<th>SNP and model</th>
<th>CH Locus</th>
<th>EAF</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stages 1 + 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11168036</td>
<td>PFDN1/</td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>HBEGF</td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>rs7920721</td>
<td>USP6NL/</td>
<td>0.4</td>
<td>0.2</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>ECHDC3</td>
<td>0.4</td>
<td>0.2</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>rs2632516</td>
<td>BZRAP1-ASI</td>
<td>0.4</td>
<td>0.6</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.6</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>rs9749589</td>
<td>NFIC</td>
<td>0.1</td>
<td>0.2</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>0.2</td>
<td>0.8</td>
<td>0.8</td>
</tr>
</tbody>
</table>

**Abbreviations:** SNP, single-nucleotide polymorphism; CH, chromosome; EAF, effect allele; EA, European ancestry; AA, African-American; JPN, Japanese; IA, Israeli-Arab; OR, odds ratio; CI, confidence interval; \( P \), \( P \)-value; NA, not applicable.

*Results for interaction term (\( NFIC \) rs9749589 \( \times \) APOE e4) and main effect of rs9749589.
In the combined Stage 1 + 2 sample, GWS association was attained with two intergenic SNPs between PFDN1 and HBEGF (best SNP: rs11168036, $P = 7 \times 10^{-5}$), six intergenic SNPs between USP6NL and ECHDC3 (best SNP: rs7920721, $P = 3 \times 10^{-8}$), and BZRAP1-ASI SNP rs2632516 ($P = 4 \times 10^{-5}$) (Table 1, Fig. 1, and Supplementary Table 6). Analyses of models that conditioned on the top SNP at the PFDN1/HBEGF, USP6NL/ECHDC3, and BZRAP1-ASI loci confirmed a single association signal in each region (Supplementary Fig. 5). The significant interaction between NFIC SNP rs9749589 and ε4 status in Stage 1 was not significant in Stage 2 ($P = .2$); however, the magnitude and direction of effect were the same, and the interaction $P$ value in the total sample was not diminished (Table 1 and Fig. 1). These GWS associations, except for rs7920721, were supported by evidence in multiple ethnic groups (Fig. 2). Further evaluation of the Stage 1 + 2 findings revealed that the association with the USP6NL/ECHDC3 SNPs was exclusive to subjects lacking APOE ε4 (e.g., rs7920721: ε4+, $P = .07$, OR = 1.05; ε4−, $P = 2.7 \times 10^{-9}$, OR = 1.14; interaction $P = .01$) and comparable in terms of effect size and direction in the non-EA groups (Supplementary Table 6 and Supplementary Fig. 6). All GWS findings were similar using the METASOFT RE-HE model (Supplementary Table 7).

3.2. Gene-based test findings

In Stage 1 analyses, there was strong evidence of association (gene-based $P < 10^{-4}$) with previously established loci and novel loci in the total sample (Supplementary Fig. 7 and Supplementary Table 8) but only seven known genes (CRI, BIN1, PTK2B, CLU, MS4A4A, PICALM, and ABCA7) and one novel one (TPBG) were GWS ($P < 2.7 \times 10^{-6}$) in the combined Stage 1 + 2 sample (Table 2 and Supplementary Table 8). Both EAs and AAs contributed to the association with TPBG. No additional genes were identified as GWS in interaction models or APOE genotype subgroups.

4. Discussion

In this large transethnic genetic study of AD, we identified robust associations with several novel loci at the individual SNP level (PFDN1/HBEGF, USP6NL/ECHDC3, BZRAP1-ASI, and NFIC) and gene level (TPBG) in a sample of AD subjects and cognitively normal elders in cohorts containing whites of EAs, AAs, JPN, and IAs. Most of these findings are supported by evidence in more than one ethnic group (Fig. 2 and Table 2). Previous GWAS using the EA discovery cohorts in this study did not detect genome-wide significant association with any of these loci, although there was suggestive evidence of association ($P > 10^{-7}$) for the top SNPs in the PFDN1/HBEGF and USP6NL/ECHDC3 regions in EAs [12,14]. The other novel genes identified in this study were not previously reported to be associated with AD in any ethnic groups. The association with SNPs in the USP6NL/ECHDC3 region was specific to persons lacking the APOE ε4 allele. Our study also showed that associations for several genes that have previously been robustly implicated in AD in Caucasians of European descent (CRI, BIN1, PTK2B, MS4A4A, and PICALM) were evident in other populations even at the SNP level.

HBEGF, heparin epidermal growth factor (EGF)-like growth factor, has roles in wound healing, cardiac hypertrophy, and heart development [23]. Although the biological role for this gene in AD is not obvious, an HBEGF knockout mouse that does not express HBEGF in cortex and hippocampus has psychiatric and cognitive dysfunctions that accompany downregulated N-methyl-D-aspartate receptors [24]. Another study showed that rats exposed to the pesticide cypermethrin had a reduction of HBEGF expression leading to upregulation of GSK3b-dependent Aβ and phosphorylated tau [25].

A recent GWAS demonstrated pleiotrophic effects of SNPs in the USP6NL/ECHDC3 (including rs7920721) and BZRAP1-ASI loci for AD and plasma C-reactive protein and lipid levels [26]. The pleiotropy at USP6NL/ECHDC3 may be related to the association finding at this locus among persons lacking the APOE ε4 allele. USP6NL, ubiquitin-specific peptidase 6 N-terminal like, has a role in the EGF receptor (EGFR) signaling pathway by acting as a GTPase-activating protein and inhibiting internalization of EGFR [27]. Insight for a role of USP6NL may be gained from information about USP6 that regulates ubiquitylation and trafficking of cargo protein by clathrin-mediated endocytosis [28]. There is a growing body of evidence from studies in humans and mice supporting a role for clathrin-independent endocytosis in AD [29–31]. In addition, the association of the phosphatidylinositol-binding clathrin assembly protein (PICALM) gene to AD is well established [12].

ECHDC3, enoyl CoA hydratase domain containing 3, is involved in fatty acid biosynthesis in mitochondria, and its expression is increased in patients with acute myocardial infarction [32]. It has been observed that ECHD3 expression is altered in brains from persons with AD compared with controls [26]. Although rs7920721 is closer to ECHDC3 than USP6NL, it is located on USP6NL side of a recombination hotspot between these two genes (Fig. 1B). Therefore, we cannot rule out either of these genes, or even one not adjacent to rs7920721, as explaining the association signal in this region.

BZRAP1, benzodiazepine-associated protein 1 (renamed as TSPO-associated protein 1, TSPOAP1), is a subunit of the benzodiazepine receptor complex in
Fig. 1. Regional association plots in the combined Stage 1 and Stage 2 samples including main effects at (A) PFDN1/HBEGF, (B) USP6NL/ECHDC3, (C) BZRAP1-AS1, and (D) single-nucleotide polymorphism (SNP) × APOE ε4 interaction near NFIC.
Fig. 1. (continued).
mitochondria and a marker of neuroinflammation [33]. A recent prospective cohort study of 8240 individuals aged 65 years and older showed an increased risk of dementia with use of long half-life benzodiazepines [34], a drug often prescribed for treatment of anxiety. A TSPO ligand (Ro5-4864) has been shown to reverse β-amyloid accumulation and behavioral impairment in 3xTgAD mice [35]. A recent positron-emission tomography imaging study demonstrated that the change over time of TSPO binding to radioligand 11C-PBR28 is correlated with progression of AD [36].

The relationship of AD to the other novel loci identified in this study is less clear. PFDN1, a prefoldin subunit, is upregulated in colorectal cancer [37]. NFIC is a CCAAT-binding transcription factor. A study comparing brain gene expression profiles between HIV seropositive individuals with cognitive impairment and AD cases identified NFIC as having significant high co-expression connectivity in white matter [38]. Trophoblast glycoprotein (TPBG), also known as ST4, regulates development of the olfactory bulb GABAergic interneurons and its overexpression in newborns is associated with abnormal dendrites [39].

Our study highlights the benefit of combining results obtained from genetically diverse populations. The trans-ethnic approach applied here identified three novel loci (BZRAP1-AS1, NFIC, and TPBG) and GWS association for the first time with two other loci (PFDN1/HBEGF and USP6NL/ECHDC3) noting that the size of the discovery sample in this study was less than 45% of the one included in a previous GWAS that contained more than 74,000 EA subjects. The improved power in our smaller sample can be ascribed to allele frequency differences and allelic heterogeneity among the ethnic groups.

Table 2
Genome-wide significant results ($P < 2.7 \times 10^{-6}$) from gene-based tests in Stage 1 + 2

<table>
<thead>
<tr>
<th>Gene</th>
<th>CH</th>
<th>EA</th>
<th>AA</th>
<th>JPN</th>
<th>IA</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 1 + 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRI</td>
<td>1</td>
<td>$3.4 \times 10^{-9}$</td>
<td>0.84</td>
<td>0.35</td>
<td>0.10</td>
<td>$4.8 \times 10^{-9}$</td>
<td>$6.4 \times 10^{-8}$</td>
<td>$1.4 \times 10^{-12}$</td>
</tr>
<tr>
<td>BIN1</td>
<td>2</td>
<td>$1.4 \times 10^{-14}$</td>
<td>0.08</td>
<td>0.35</td>
<td>0.10</td>
<td>$3.7 \times 10^{-9}$</td>
<td>$2.5 \times 10^{-9}$</td>
<td>$8.8 \times 10^{-23}$</td>
</tr>
<tr>
<td>TPBG</td>
<td>6</td>
<td>$2.2 \times 10^{-3}$</td>
<td>$3.5 \times 10^{-3}$</td>
<td>0.29</td>
<td>0.70</td>
<td>$6.8 \times 10^{-6}$</td>
<td>$8.2 \times 10^{-3}$</td>
<td>$1.8 \times 10^{-6}$</td>
</tr>
<tr>
<td>PTK2B</td>
<td>8</td>
<td>$4.7 \times 10^{-6}$</td>
<td>0.26</td>
<td>0.49</td>
<td>0.81</td>
<td>$2.2 \times 10^{-6}$</td>
<td>$1.9 \times 10^{-3}$</td>
<td>$7.6 \times 10^{-9}$</td>
</tr>
<tr>
<td>CLU</td>
<td>7</td>
<td>$1.0 \times 10^{-6}$</td>
<td>0.11</td>
<td>0.59</td>
<td>0.10</td>
<td>$1.1 \times 10^{-6}$</td>
<td>$1.1 \times 10^{-9}$</td>
<td>$1.4 \times 10^{-12}$</td>
</tr>
<tr>
<td>MS4A4A</td>
<td>11</td>
<td>$5.4 \times 10^{-13}$</td>
<td>0.18</td>
<td>0.03</td>
<td>0.95</td>
<td>$3.6 \times 10^{-14}$</td>
<td>0.04</td>
<td>$6.1 \times 10^{-14}$</td>
</tr>
<tr>
<td>PICALM</td>
<td>11</td>
<td>$1.9 \times 10^{-8}$</td>
<td>0.71</td>
<td>$1.8 \times 10^{-3}$</td>
<td>0.93</td>
<td>$1.6 \times 10^{-9}$</td>
<td>$3.6 \times 10^{-3}$</td>
<td>$2.7 \times 10^{-11}$</td>
</tr>
<tr>
<td>ABCA7</td>
<td>19</td>
<td>$2.3 \times 10^{-4}$</td>
<td>$1.6 \times 10^{-3}$</td>
<td>0.07</td>
<td>0.02</td>
<td>$6.6 \times 10^{-7}$</td>
<td>$4.9 \times 10^{-3}$</td>
<td>$1.1 \times 10^{-8}$</td>
</tr>
</tbody>
</table>

Abbreviations: CH, chromosome; EA, European ancestry; AA, African-American; JPN, Japanese; IA, Israeli-Arab; P, gene-based P-value.
As an example highlighting the importance of these differences, the top SNPs from BZRAP1-AS1 and NFIC had different minor allele frequencies across ethnic groups, but the effect sizes were similar and association signals were greater in fixed-effect meta-analysis. In addition, gene-based tests, which consider association patterns with all SNPs in the locus, identified TPBG. Importantly, the most significant SNPs in these two regions differed among the ethnic groups. Gene-based tests also indicated potential allelic heterogeneity among ethnic groups for previously established AD genes including TREM2 and ABCA7. The novel GWS SNP associations were robust in analyses allowing for heterogeneity across different ethnic groups, and the P-values for the RE-HE approach were slightly larger than for the fixed-effect model, suggesting that the effect size heterogeneity across the groups is modest.

Our study also revealed that the effect direction for several SNPs vary across ethnic groups. For example, the top-ranked SNPs in NME8, ABCA7, and CASS4 (Supplementary Table 2) were nominally significant in EAs and AAs, but the referent allele was associated with increased risk in one group and decreased risk in the other. One explanation for these differences is that the SNPs are tagging different functional variants across groups. This idea is consistent with our findings from gene-based tests showing that the constellation of variants contributing to the association with some genes was different across ethnic groups. Alternatively, when examining a large number of variants, it is expected that a few will show nominal significance in opposite directions among groups.

There are several limitations associated with our study. The sample size imbalance between the EAs and the other populations weakened the opportunity to identify association patterns that may be unique to the non-EA groups. The small size of the non-EA groups also reduced power to detect novel gene associations if the functional variants (and the SNPs that tag them) differ among ethnic groups. An additional weakness is the lack of replication samples for the non-EA populations. Despite these limitations, our study highlights the importance of investigating the genetic architecture for AD in ethnically diverse populations.

Our findings warrant further replication in independent samples, deep sequencing and bioinformatics studies to identify the potentially functional variants, and experimental validation. We expect that additional novel gene discoveries will emerge in future transethnic studies including larger samples from non-EA populations.

Acknowledgments

We thank Drs. D. Stephen Snyder and Marilyn Miller from National Institute on Aging (NIA) who are ex-officio Alzheimer’s Disease Genetics Consortium (ADGC) members. Funding support: The National Institutes of Health, National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01 AG032984, RC2 AG036528; NACC, U01 AG016976; NCRRAD, U24 AG021886; NIA LOAD, U24 AG026395, U24 AG026390; MIRAGE: R01 AG025259; Banner Sun Health Research Institute P30 AG019610; Boston University, P30 AG013846, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG017173, R01 AG33193, R01 AG048927; Columbia University, P50 AG008702, R37 AG015473; Duke University, P30 AG028377, AG05128; Emory University, AG025688; Group Health Research Institute, U01 AG06781, U01 HG004610; U01 HG006375; Indiana University, P30 AG10133; Johns Hopkins University, P50 AG005146, R01 AG020688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG016574; Mount Sinai School of Medicine, P50 AG005138, P01 AG002219; New York University, P30 AG08051, MO1RR00096, and U1L RR029893; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG30146; TGen, R01 NS059873; University of Alabama at Birmingham, P50 AG016582, UL1RR02777; University of Arizona, R01 AG031581; University of California, Davis, P30 AG010129; University of California, Irvine, P50 AG016573, P50, P50 AG016575, P50 AG016576, P50 AG016577; University of California, Los Angeles, P50 AG016570; University of California, San Diego, P50 AG05131; University of California, San Francisco, P50 AG023501, P01 AG019724; University of Kentucky, P30 AG028383; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124; University of Pittsburgh, P50 AG005133, AG030653; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG010491, AG027944, AG021547, AG019757; University of Washington, P50 AG005136; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant #NS39764, NIMH MH60451, and by GlaxoSmithKline. Genotyping of the TGEN2 cohort was supported by Kronos Science. The TGen series was also funded by NIA grant AG034504 to AJM, The Banner Alzheimer’s Foundation, The Johnnie B. Byrd Sr. Alzheimer’s Institute, the Medical Research Council, and the State of Arizona and also includes samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts, and Newcastle University), MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council), Southwest Dementia Brain Bank
(funding via numerous sources including the Higher Education Funding Council for England, Alzheimer’s Research Trust [ART], BRACE, and North Bristol NHS Trust Research and Innovation Department and DEnDRoN), The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Herenstichting Nederland Breinbrekkend Werk, International Parkinson Fonds, Internationale Stichting Alzheimer Onderzoek), Instituto de Neuropatología, Servei Anatomia Patològica, Universitat de Barcelona. Marcelle Morrison-Bogorad, PhD, Tony Phelps, PhD, and Walter Kukull PhD are thanked for helping to coordinate this collection. Alzheimer’s Disease Neuroimaging Initiative funding is through the Northern California Institute for Research and Education by grants from Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Merck & Co., Inc., Novartis AG, Pfizer, F. Homan-La Roche, Schering-Plough, Synarc, Alzheimer’s Association, Alzheimer’s Drug Discovery Foundation, the Dana Foundation, and by the National Institute of Biomedical Imaging and Bioengineering and NIA grants U01 AG024904, RC2 AG036535, K01 AG030514. Support was also from the Alzheimer’s Association’s (LAF, IIRG-08-89720; MP-V, IIRG-05-14147) and the US Department of Veterans Affairs Administration, Office of Research and Development, Biomedical Laboratory Research Program. P.S.G.-H. is supported by the Wellcome Trust, Howard Hughes Medical Institute, and Canadian Institute of Health. ADGC members are as follows: Perrie M. Adams, PhD; Marilyn S. Albert, PhD; Roger L. Albin, MD; R. Barbieri, MD; J. Barlow, MD; J. Baroldi, PhD; M. Barmada, PhD; Marilyn S. Albert, PhD; Roger L. Albin, MD; Liana G. Adam, MD; Craig S. Atwood, PhD; Michel J. Barmada, PhD; Lisa L. Barnes, PhD; Thomas G. Beach, MD, PhD; James T. Becker, PhD; Eileen H. Bigio, MD; Thomas D. Bird, MD; Deborah Blacker, MD; Bradley F. Boeve, MD; James D. Bowen, MD; Adam Boxer, MD, PhD; James R. Burke, MD, PhD; Nigel J. Cairns, PhD, FRCPath; Chuanhai Cao, PhD; Chris S. Carlson, PhD; Cynthia M. Carlsson, MD, Regina M. Carney, MD; Minerva M. Carraquillo, PhD; Steven L. Carroll, MD, PhD; Helena C. Chui, MD; David G. Clark, MD; Jason Corneveaux, BS; David H. Cribs, PhD; Elizabeth A. Crocco, MD; Carlos Cruchaga, PhD; Philip L. De Jager, MD, PhD; Charles DeCarli, MD; Steven T. DeKosky, MD; F. Yesim Demirci, MD; Malcolm Dick, PhD; Dennis W. Dickson, MD; Rachelle S. Doody, MD, PhD; Ranjan Duara, MD; Nilufer Ertekin-Taner, MD, PhD; Kelley M. Faber, MS; Thomas J. Fairchild, PhD; Kenneth B. Fallon, MD; Martin R. Farlow, MD; Steven Ferris, PhD; Matthew P. Frosch, MD, PhD; Douglas R. Galasko, MD; Marla Gearing, PhD; Daniel H. Geschwind, MD, PhD; Bernardino Ghatti, MD; John R. Gilbert PhD; Jonathan D. Glass, MD; Neill R. Graff-Radford, MD; Robert C. Green, MD, MPH; John H. Growdon, MD; Hakon Hakonarson, MD, PhD; Ronald L. Hamilton, MD; John Hardy, PhD; Lindy E. Harrell, MD, PhD; Elizabeth Head, PhD; Lawrence S. Honig, MD, PhD; Ryan M. Huebinger, PhD, Matthew J. Huentelman, PhD; Christine M. Hulette, MD; Bradley T. Hyman, MD, PhD; Gail P. Jarvik, MD, PhD; Gregory A. Jicha, MD, PhD; Lee-Way Jin, MD, PhD; Anna Karydas, BA; John S.K. Kauwe, PhD; Jeffrey A. Kaye, MD; Ronald Kim, MD; Edward H. Koo, MD; Neil W. Kowall, MD; Joel H. Kramer, PsyD; Frank M. LaFerla, PhD; James J. Lah, MD, PhD; James B. Leverenz, MD; Allan I. Levey, MD, PhD; Ge Li, MD, PhD; Andrew P. Lieberman, MD, PhD; Chiao-Feng Lin, PhD; Oscar L. Lopez, MD; Constantine G. Lyketsos, MD, MHS; Wendy J. Mack, PhD; Daniel C. Marson, JD, PhD; Frank Martiniuk, PhD; Deborah C. Marsh, PhD; Eliezer Masliah, PhD; Wayne C. McCormick, MD, MPH; Susan M. McCurry, PhD; Andrew N. McDaid, BA; Ann C. McKee, MD; Marcel Meola, MD; Bruce L. Miller, MD; Carol A. Miller, MD; Joshua W. Miller, PhD; John C. Morris, MD; Shubbabrata Mukherjee, PhD; Jill R. Murrell, PhD, Amanda J. Myers, PhD; Sid O’Bryant, PhD; John M. Olichney, MD; Vernon S. Pankratz, PhD; Joseph E. Parisi, MD; Amanda Parutch, MS; Henry L. Paulson, MD, PhD; William Perry, MPH; Elaine Peskind, MD; Ronald C. Petersen, MD, PhD; Aimee Pierce, MD; Wayne W. Poon, PhD; Huntington Potter, PhD; Joseph F. Quinn, MD; Ashok Raj, MD; Murray Raskind, MD; Barry Reisberg, MD; Joan S. Reisch, PhD; Christiane Reitz, MD, PhD; John M. Ringman, MD; Erik D. Roberson, MD, PhD; Ekaterina Rogaeva, PhD; Howard J. Rosen, MD; Roger N. Rosenberg, MD; Donald R. Royall, MD; Mark A. Sager, MD; Mary Sano, PhD; Andrew J. Saykin, PsyD; Julie A. Schneider, MD; Lon S. Schneider, MD; William W. Seeley, MD; Amanda G. Smith, MD; Joshua A. Sonnen, MD; Salvatore Spina, MD; Robert A. Stern, PhD; Rudolph E. Tanzi, PhD; Tricia A. Thornton-Wells, PhD; John Q. Trojanowski, MD, PhD; Juan C. Troncoso, MD; Debby W. Tsuang, MD; Vivianna M. Van Deerlin, MD, PhD; Linda J. Van Eldik, PhD; Badri N. Vardarajan, PhD; Harry V. Vinters, MD; Jean Paul Vonsattel, MD; Sandra Weintraub, PhD; Kathleen A. Welsh-Bohmer, PhD; Jennifer Williamson, MS; Sarah Wishnek, MPH; Randall L. Wolter, MD, PhD; Clinton B. Wright, MD, MS; Chuang-Kuo Wu, MD, PhD; Chang-En Yu, PhD; Lei Yu, PhD; and Xiaoling Zhang, PhD. Conflicts of interest: The authors of this article have no conflicts of interest to report. Web resources: NIA Genetics of Alzheimer’s Disease Data Storage Site: https://www.niagads.org/.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jalz.2016.12.012.
RESEARCH IN CONTEXT

1. Systematic review: We reviewed previously published genome-wide association studies (GWAS) for late-onset Alzheimer’s disease (AD) including reports for non-white populations. Few GWAS have been conducted in populations of non-white European ancestry.

2. Interpretation: Tranethnic meta-analysis of GWAS results for whites of European Ancestry, African-Americans, Japanese, and Israeli-Arabs identified novel genome-wide significant associations with single-nucleotide polymorphisms in PFDN1/ HBEGF, USP6NL/ECHDC3, and BZRAP1-AS1 and with TPBG using a gene-based test. These findings further our understanding of the genetic basis of AD and provide insight about mechanisms leading to AD.

3. Future directions: These results should be confirmed in independent samples including subjects from the same ethnic populations and tested in populations of other genetic backgrounds. DNA sequencing studies are needed to identify the functional variants in these genes and their biological roles in AD should be determined experimentally.

References


