Received: 17 November 2023 | Revised: 14 February 2024 | Accepted: 23 February 2024

DOI: 10.1002/alz.13796

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

On the effect heterogeneity of established disease susceptibility loci for Alzheimer's disease across different genetic ancestries

Sanghun Lee[1,2,3] | **Julian Hecker**[2] | **Georg Hahn**[3] | **Kristina Mullin**[4] | **Alzheimer's Disease Neuroimaging Initiative (ADNI)**[#] | **Sharon M. Lutz**[3,5] | **Rudolph E. Tanzi**[4] | **Christoph Lange**[2,3] | **Dmitry Prokopenko**[4]

[1]Department of Medical Consilience, Division of Medicine, Graduate school, Dankook University, Yongin-si, Gyeonggi-do, South Korea

[2]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA

[3]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

[4]Genetics and Aging Unit and McCance Center for Brain Health, Department of Neurology, Massachusetts General Hospital, Charlestown, Massachusetts, USA

[5]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute, Boston, Massachusetts, USA

Correspondence
Rudolph E. Tanzi and Dmitry Prokopenko, Genetics and Aging Unit and McCance Center for Brain Health, Department of Neurology, Massachusetts General Hospital, 114 16th St, Charlestown, MA 02129, USA.
Email: rtanzi@mgh.harvard.edu and dprokopenko@mgh.harvard.edu

Christoph Lange, Department of Biostatistics, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA 02115, USA.
Email: clange@hsph.harvard.edu

[#]Data used in preparation of this article were in part obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Funding information
Cure Alzheimer's Fund, Grant/Award Number: R01MH129337; NIH, Grant/Award Number: R01MH129337

Abstract

INTRODUCTION: Genome-wide association studies have identified numerous disease susceptibility loci (DSLs) for Alzheimer's disease (AD). However, only a limited number of studies have investigated the dependence of the genetic effect size of established DSLs on genetic ancestry.

METHODS: We utilized the whole genome sequencing data from the Alzheimer's Disease Sequencing Project (ADSP) including 35,569 participants. A total of 25,459 subjects in four distinct populations (African ancestry, non-Hispanic White, admixed Hispanic, and Asian) were analyzed.

RESULTS: We found that nine DSLs showed significant heterogeneity across populations. Single nucleotide polymorphism (SNP) rs2075650 in translocase of outer mitochondrial membrane 40 (*TOMM40*) showed the largest heterogeneity (Cochran's $Q = 0.00$, $I^2 = 90.08$), followed by other SNPs in apolipoprotein C1 (*APOC1*) and apolipoprotein E (*APOE*). Two additional loci, signal-induced proliferation-associated 1 like 2 (*SIPA1L2*) and solute carrier 24 member 4 (*SLC24A4*), showed significant heterogeneity across populations.

DISCUSSION: We observed substantial heterogeneity for the *APOE*-harboring 19q13.32 region with *TOMM40/APOE/APOC1* genes. The largest risk effect was seen among African Americans, while Asians showed a surprisingly small risk effect.

KEYWORDS
Alzheimer's disease, disease susceptibility loci, effect size, genetic ancestry, heterogeneity

1 | INTRODUCTION

Genetic studies, including genome-wide association studies (GWASs), have made significant progress in identifying disease-associated genes and risk variants in Alzheimer's disease (AD).[1] To date, GWAS have identified numerous disease susceptibility loci (DSLs) associated with AD risk. Previous GWASs suggested over 70 DSLs for AD across world-wide populations.[2–13] These DSLs encompass a range of genes involved in various biological processes, including amyloid processing, tau pathology, inflammation, lipid metabolism, synaptic function, and immune response.

Diverse genetic architectures among different ethnic groups can influence how genetic factors contribute to the pathogenesis of AD.[12,14,15] Genetic factors play a complex role in AD, and variations in genetic architectures across populations can result in differences in disease susceptibility, progression, and presentation.[16] Genetic variations, including single nucleotide polymorphisms (SNPs), can vary in frequency across different ethnic groups. Certain alleles that are associated with increased or decreased risk of AD in one population may have a different effect size in another population. Therefore, an interaction between ethnicity and the effect of the genotypes on AD risk has gained much attention. Genetic factors associated with AD can show effects of different magnitude (or even opposite effects) in the presence or absence of another genetic factor or in different environments, resulting in complex gene–gene or gene–environment interactions.[17] These interactions may differ among ethnic groups and thus impact the overall genetic risk profile of AD across populations.

For example, the apolipoprotein E (*APOE*) locus, especially the *APOE* ε4 allele, is the most well-established and robust genetic risk factor for late-onset AD in multiple GWASs. The frequency of the *APOE* ε4 allele varies widely across ethnic groups (African ancestry [AA] 19%, non-Hispanic White [NHW] 14%, Hispanic [HISP] 12%, Asian [ASN] 9%, other/mixed 23%) but drastically rises in AD patients (AA 35%, NHW 38%, HISP 24%, ASN 28%, other/mixed 45%) based on the AlzGene database (www.alzgene.org; last accessed August 2023).[18] Among NHW, the heterozygotes (*APOE* ε4/ε3) have a 2.7 to 2.8 times greater risk in comparison to those with the *APOE* ε3/ε3 genotype, and *APOE* ε4 homozygotes have a 12-fold increased risk.[12,19] The odds ratio (OR) is weaker in those of AA and Hispanics despite a higher frequency of the *APOE* ε4 allele compared to NHW populations.[19] However, the risk among Japanese populations is much stronger compared to NHW (ε3/ε4: OR = 3.9 to 5.6, ε4/ε4: OR = 21.8 to 33.1), while the ε4 allele is less common.[12,18] Interestingly, the report showed that another ASN population, Chinese, have ORs similar to those of NHW (ε3/ε4: OR = 3.08, ε4/ε4: OR = 11.76).[20] Furthermore, the dose effect of the ε4 allele on AD risk was not observed in another ASN study of Indians (ε3/ε4: OR = 4.18, ε4/ε4: OR = 4.81).[21]

Investigating these interactions between DSL for AD and populations is vital for advancing our understanding of AD and tailoring prevention, diagnosis, and treatment strategies for diverse populations. To address the knowledge gap in understanding genetic factors in AD across diverse populations, we tried to compare the effects of the established loci with AD in a multi-ethnic AD cohort.

2 | METHODS

2.1 | NIA genetics of Alzheimer's disease data storage site (NIAGADS)

The genetic, genomic, and phenotypic data, including clinical and neuropathology data from National Institutes of Health (NIH)-funded genetic studies are organized and shared by the National Institute on Aging Genetics of Alzheimer's Disease Data Storage (NIAGADS) site. The NIAGADS site currently hosts 76 high-quality human genetics datasets in addition to Alzheimer's Disease Sequencing Project (ADSP) data, corresponding to 55,241 participants, and has a genomics database for cross-referencing.

2.2 | Whole genome sequencing (WGS) data

The Variant Call Format (VCF) files for the NIA ADSP cohort were obtained from the NIAGADS site under accession no. NG00067.v9. WGS data were released in 2022 (ADSP R4) from the participants in AD and AD-related dementia (ADRD) studies who were from ethnically diverse populations (35,569 unique subjects – 5218 AA, 2791 A, 10,398 HISP, 16,191 NHW, and 971 Other/Unknown) as part of the ADSP Follow Up Study (FUS) to expand the diversity of ancestries that are included in the ADSP's sequencing and analysis.

After the sample quality control, a total of 26,243 subjects with a valid AD phenotype remained. The genetic principal components (PCs) (PC1-PC10) were based on the genetic relationship matrix with common variants for population substructure. After PC analysis, we excluded the PC outliers to their population, and a total of 25,459 subjects in four different populations such as AA (n = 4265), NHW (n = 10,184), HISP (n = 8467), and ASN (n = 2543) were included in the final analysis.

The case affection status for most subjects was based on the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA) workgroup criteria. Unaffected subjects were at least 60 years old and were free of dementia by cognitive

assessment.[22] It is important to note that the reported diagnosis in the ASN subpopulation, which was predominantly from the Longitudinal Aging Study in India (LASI)-Diagnostic Assessment of Dementia cohort, is based on either a consensus Clinical Dementia Rating (CDR) or a predicted CDR using a machine learning method.[23] Full descriptions of related study cohorts can be found at https://dss.niagads.org/datasets/ng00067/#studies.

2.3 | Statistical analysis

Among the reported DSLs (n = 308) associated with AD in GWASs (Table S1), 259 SNPs were available in our WGS data. For the loci, we ran a logistic regression using a generalized linear model (GLM) in each population (AA, NHW, HISP, ASN, and all of them) on AD affection status with covariates such as age, sex, sequencing center, and each population-specific PC, PC1 to PC10. Both the p value for Cochrane's Q statistic and I^2 heterogeneity index (0 to 100) were calculated from fixed effects and random effects in meta-analysis using the previous results in each subpopulation.

We performed simulations to illustrate the effect of AD affection rate in different populations on the effect estimate of heterogeneous DSLs in meta-analysis. Due to the lowest AD affection rate (7.12%) in the ASN population, we extracted cases (N = 181) and controls (N = 2362) with the same ratio by random sampling with replacement from AA, NHW, or HISP populations (50 simulations, respectively). The genotype × PC terms in the combined analysis were added to the regression model to assess the association of ethnicity–gene interaction with AD affection status using the same set of covariates as above. All analyses were conducted using PLINK software (https://www.cog-genomics.org/plink/2.0/, accessed in October 2022) using an additive genetic model.

Clustering by country of the ASN subpopulation was performed with the bigsnpr package using ancestry grouping to reference populations.[24] Clinical features between populations were compared using t tests or chi-squared tests as appropriate. The R statistical software version 4.10 (http://www.R-project.org) was used.

3 | RESULTS

3.1 | Study subjects

The characteristics of the 25,459 subjects included in our analysis are summarized in Table 1. The median age of study participants was 72.90 years with a sex ratio of 1:1.62 (male: female). The AD affection rate was different between populations as follows: 63.79% in NHW, 32.99% in AA, 30.22% in HISP, and 7.12% in ASN population. The population structure in the subjects analyzed was shown in the principal component analysis (PCA) plots (Figure 1A &1B), where the ASN population was clearly isolated against the other populations on the third PC. Most of our ASN population were from the LASI from South Asia.

RESEARCH IN CONTEXT

1. **Systematic review**: The authors reviewed the previous genome-wide association studies (GWASs) for numerous disease susceptibility loci (DSLs) associated with Alzheimer's disease (AD) risk to date. Only a limited number of studies have investigated the dependence of the genetic effect size of established DSLs across genetic ancestry.
2. **Interpretation**: Substantial heterogeneity was observed specifically for the APOE-harboring 19q13.32 region with *TOMM40/APOE/APOC1* genes between populations, where a much smaller effect of *APOE*-ε4 allele was revealed in the South Asian population. Two additional loci, *SIPA1L2* and *SLC24A4*, showed significant heterogeneity across populations.
3. **Future directions**: Ancestry-specific genetic involvement for AD leads to better understanding of these genes implicated in AD development. Future studies are aimed to assess the biological and environmental implications for our findings.

3.2 | Heterogeneity of reported DSLs for AD across four populations

The GLM results for the reported DSL, which were available in the dataset (n = 259), are summarized in Table S2. The number of replicated DSLs having a nominally significant p value (< 5%) in our cohort was 79 (30.50%). The replication rate differed among populations as follows: 20.46% in NHW, 11.97% in AA, 13.51% in HISP, and 2.32% in ASN population. We observed that some DSLs were population-specific. Specifically, a total of 9 DSLs had significant heterogeneity based on Cochrane's Q statistic (< 0.01), which were summarized in Table 2. The odds ratios (ORs) are also compared in Figure 2A.

In particular, rs2075650 and rs157582 in linkage disequilibrium (LD) located on translocase of outer mitochondrial membrane 40 (*TOMM40*) in chromosome 19 were not only strongly significant (OR = 1.72 and 1.60, and p value = 8.18×10^{-88} and 1.05×10^{-99}, respectively) but also had significant heterogeneity between populations (Cochran's Q = 0.00 and 0.00, I^2 = 90.08 and 87.99, respectively; Table 2). Moreover, the SNPs on *APOE*, including rs429358 (*APOE* ε4) and rs7412 (*APOE* ε2), showed significant heterogeneity. In the ASN population, the OR was not significant, while AA and NHW showed a highly significant OR and HISP was followed (Figure 2A). Significant heterogeneity was also observed for rs4420638 located on apolipoprotein C1 *APOC1* where NHW showed the largest OR compared to AA and HISP (Figure 2A). Taken together, *TOMM40-APOE-APOC1* DSLs in the *APOE*-harboring 19q13.32 region had significant heterogeneity in AD between populations; in particular, ASN subjects

TABLE 1 Demographic characteristics of final subjects in each population. Mean (SD) or number (%) is shown.

Population	Subjects (*n*)	Age (mean ± SD)	Female (*n*)	AD cases (*n*)	Age in cases (mean ± SD)	AD controls (*n*)	Age in controls (mean ± SD)
All	25,459	72.90 ± 9.98	15,728 (61.78%)	10,643 (41.80%)	72.26 ± 10.62	14,816 (58.20%)	73.35 ± 9.47
NHW	10,184	74.30 ± 11.09	5746 (56.42%)	6496 (63.79%)	70.89 ± 11.28	3688 (36.21%)	80.02 ± 7.97
AA	4265	74.28 ± 8.40	3083 (72.29%)	1407 (32.99%)	74.60 ± 9.09	2858 (67.01%)	74.12 ± 8.03
HISP	8467	71.60 ± 9.63	5556 (65.62%)	2559 (30.22%)	74.17 ± 9.21	5908 (69.78%)	70.45 ± 9.59
ASN	2543	69.42 ± 7.12	1343 (52.81%)	181 (7.12%)	74.15 ± 9.08	2362 (92.88%)	69.06 ± 6.82

Abbreviations: AA, African ancestry; ASN, Asian, HISP, Hispanic; NHW, non-Hispanic White.

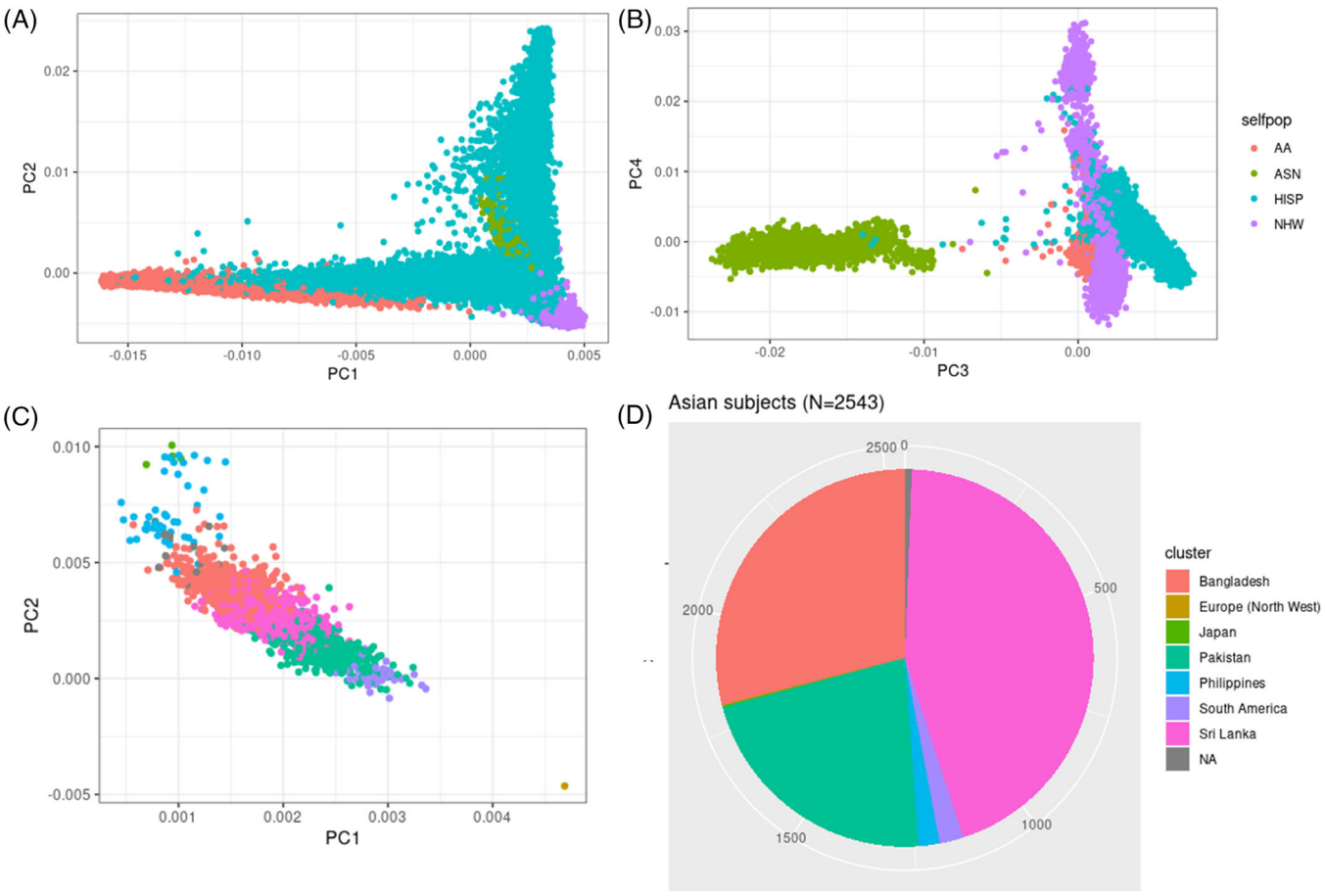

FIGURE 1 The PCA plots show the population structure according to NHW, AA, HISP, and ASN subjects (A: PC1 vs PC2 and B: PC3 vs PC4). The PCA plot (C) in the only ASN population shows the population substructure according to clustering by country and most of them are from South Asia in the pie chart (D).

were very different. To show that the effect sizes are not affected by the low case rate in the ASN population, we performed a simulation study. In the simulations where we used the affection rate as in ASN for other populations, we observed similar effect estimates with similar large confidence intervals (Figure 2B)

Two other loci showed significant heterogeneity across populations. The SNP rs115684722 on signal-induced proliferation-associated 1 like 2 (*SIPA1L2*) showed a significant effect (OR = 3.18, *p* value = 8.6×10^{-4}) in the AA population and rs7401792 on solute carrier family 24 member 4 (*SLC24A4*) showed a significant effect in the HISP population (OR = 1.12, *p* value = 2.76×10^{-3}).

Finally, we performed a SNP-by-PC interaction association test for the nine significant loci in the combined dataset. When including all 10 SNP-by-PC interaction terms into the regression model for each SNP, we observed multiple significant *p* values of SNP × PC interaction. For example, the SNP rs115684722 on *SIPA1L2* had the most significant interaction *p* value for PC1 separating AA from the others (Figure S1 and Table S3).

TABLE 2 Results from logistic regression in each population (ALL, NHW, AA, HISP, and ASN) on AD affection status with covariates such as age, sex, sequencing center, and PC1-PC10, which are sorted according to I^2 heterogeneity index.

rsID	Nearest gene	SNP ID	Population	A1	A1_FREQ	No. allele observations	Odds ratio	LOG(OR)_SE	ZSTAT	*P*
rs2075650	*TOMM40*	19:44892362:A:G	ALL	G	0.171	24,789	1.72	0.03	19.87	8.18E-88
			NHW	G	0.242	9828	1.93	0.04	15.22	2.45E-52
			AA	G	0.135	4205	1.34	0.07	4.25	2.11E-05
			HISP	G	0.116	8213	1.51	0.05	7.71	1.22E-14
			ASN	G	0.135	2543	1.15	0.17	0.86	3.90E-01
							Q statistics[a]	0	I^2 index[a]	90.08
rs157582	*TOMM40*	19:44892962:C:T	ALL	T	0.335	24,788	1.60	0.02	21.20	1.05E-99
			NHW	T	0.328	9828	1.76	0.04	14.59	3.10E-48
			AA	T	0.483	4203	1.58	0.05	9.21	3.32E-20
			HISP	T	0.306	8214	1.37	0.04	8.15	3.54E-16
			ASN	T	0.212	2543	1.17	0.14	1.17	2.40E-01
							Q statistics	0	I^2 index	87.99
rs4420638	*APOC1*	19:44919689:A:G	ALL	G	0.217	24,779	1.89	0.03	25.07	9.83E-139
			NHW	G	0.318	9824	2.05	0.04	18.02	1.32E-72
			AA	G	0.215	4200	1.58	0.06	7.97	1.58E-15
			HISP	G	0.124	8213	1.65	0.05	9.74	2.13E-22
			ASN	G	0.132	2542	1.27	0.16	1.47	1.40E-01
							Q statistics	0	I^2 index	87.4
rs429358	*APOE*	19:44908684:T:C	ALL	C	0.217	24,620	2.38	0.03	33.15	4.90E-241
			NHW	C	0.287	9668	2.52	0.04	21.07	1.66E-98
			AA	C	0.259	4196	2.64	0.06	16.82	1.61E-63
			HISP	C	0.145	8213	2.10	0.05	15.28	9.69E-53
			ASN	C	0.111	2543	1.35	0.17	1.80	7.26E-02
							Q statistics	0	I^2 index	86.71
rs769450	*APOE*	19:44907187:G:A	ALL	A	0.336	24,780	0.75	0.02	−13.11	2.72E-39
			NHW	A	0.346	9824	0.74	0.03	−8.78	1.67E-18
			AA	A	0.366	4200	0.67	0.05	−7.84	4.53E-15
			HISP	A	0.310	8214	0.81	0.04	−5.40	6.57E-08
			ASN	A	0.333	2543	1.11	0.12	0.83	7.91E-01
							Q statistics	0.0003	I^2 index	84.16

(Continues)

TABLE 2 (Continued)

rsID	Nearest gene	SNP ID	Population	A1	A1_FREQ	No. allele observations	Odds ratio	LOG(OR)_SE	ZSTAT	*P*
rs7412	*APOE*	19:44908822:C:T	ALL	T	0.052	24,531	0.57	0.05	−11.83	2.80E-32
			NHW	T	0.043	9567	0.47	0.08	−9.20	3.48E-20
			AA	T	0.096	4207	0.51	0.09	−7.26	3.96E-13
			HISP	T	0.042	8214	0.71	0.09	−3.74	1.86E-04
			ASN	T	0.047	2543	1.07	0.27	0.26	7.91E-01
							Q statistics	0.0004	I^2 index	83.63
rs71352238	*TOMM40*	19:44891079:T:C	ALL	C	0.146	24,766	1.89	0.03	20.75	1.11E-95
			NHW	C	0.242	9804	1.91	0.04	15.01	6.48E-51
			AA	C	0.031	4207	2.31	0.13	6.29	3.22E-10
			HISP	C	0.093	8213	1.66	0.06	8.49	2.08E-17
			ASN	C	0.134	2542	1.13	0.17	0.75	4.56E-01
							Q statistics	0.0018	I^2 index	80.05
rs115684722	*SIPA1L2*	1:232240417:A:T	ALL	T	0.010	24,784	0.91	0.09	−1.04	2.98E-01
			NHW	T	0.016	9826	0.94	0.13	−0.47	6.40E-01
			AA	T	0.005	4205	3.18	0.35	3.33	8.60E-04
			HISP	T	0.010	8212	0.73	0.19	−1.64	1.00E-01
			ASN	T	0.000	2541	3.47	2.34	0.53	5.95E-01
							Q statistics	0.0026	I^2 index	78.96
rs7401792	*SLC24A4*	14:92464917:G:A	ALL	G	0.477	24,790	1.04	0.02	1.84	6.57E-02
			NHW	G	0.367	9828	1.05	0.03	1.36	1.74E-01
			AA	A	0.240	4206	1.11	0.06	1.86	6.33E-02
			HISP	G	0.445	8214	1.12	0.04	2.99	2.76E-03
			ASN	A	0.462	2542	1.14	0.12	1.10	2.72E-01
							Q statistics	0.0064	I^2 index	75.64

[a] Cochrane's *Q* statistic and I^2 heterogeneity index (0 to 100) were calculated in meta-analysis using previous outcomes in each subpopulation.

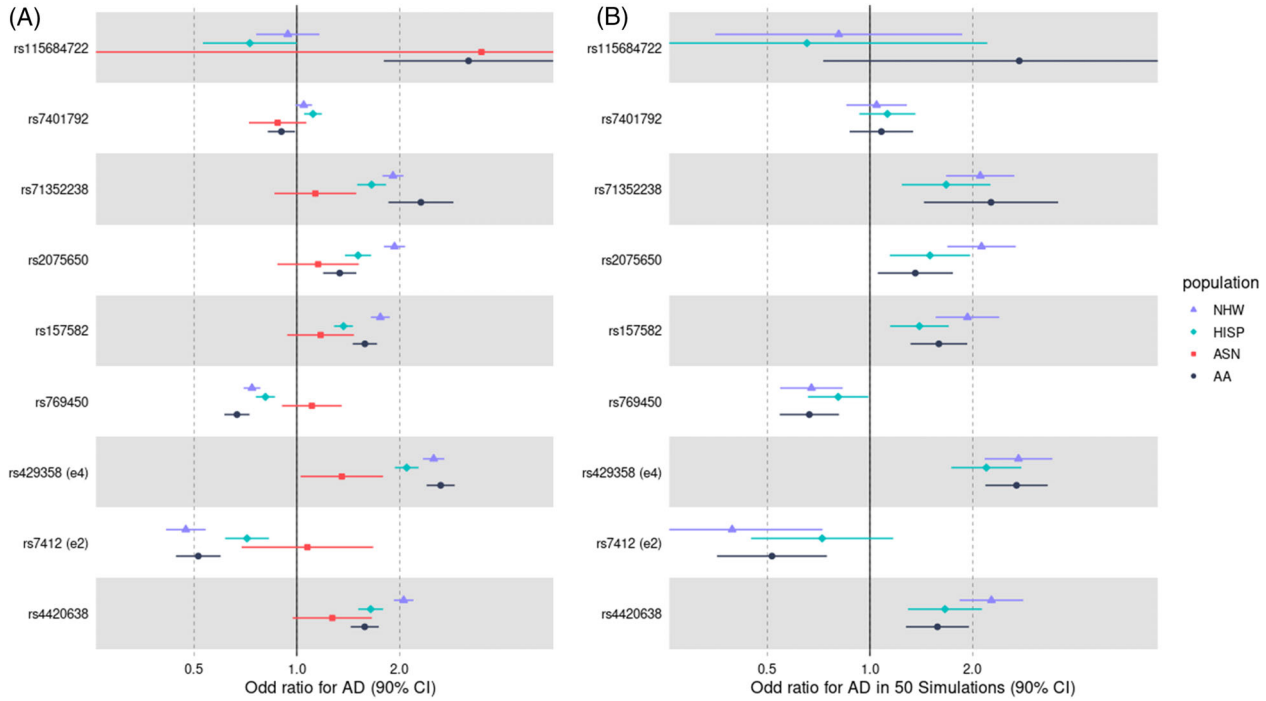

FIGURE 2 The nine DSLs for AD had significant heterogeneity based on Cochrane's *Q* statistic across four populations suggesting little effect of the DSLs in the ASN population (A). The simulation based on the cases ($N = 181$) and the controls ($N = 2362$) with the same AD affection rate (7.12%) as in the ASN population by randomly sampling with replacement from AA, NHW, or HISP populations (B, 50 simulations, respectively). They are sorted by the chromosome and base position.

4 | DISCUSSION

Among 259 DSLs associated with AD in previous reports, we observed significant heterogeneity between populations in nine variants. Seven out of nine variants were located in the *TOMM40-APOE-APOC1* region, where the LD structure is different among subpopulations in 1000 Genomes Project (Table S4). Both *APOE* variants rs7412 (ε2 allele) and rs429358 (ε4 allele) showed a strong protective and risk association, respectively, as expected. However, we observed significant heterogeneity across four continental populations. In particular, the strongest effect of both alleles was observed in the AA and NHW populations, followed by a lower effect in the HISP population, and lowest, although not significant, effect in the ASN population. Previously, their effects on AD in the ASN population were reported to be dependent on subpopulation. Even though the largest ORs of ε4 were reported in Koreans and Japanese compared to those of NHW,[12] the Chinese subpopulation showed similar effects to NHW. It is important to note that in our analysis, most of our ASN population was derived from the LASI (South Asia). Previous reports failed to find a significant association between *APOE* variants and memory scores or AD in a subset of this cohort.[25] The LD structure in 1000 Genomes Project differs between East ASN and South ASN (Table S4). In South ASN Indians, small sample size studies suggest that both ε2 and ε3 alleles are protective against AD.[21] Therefore, the genetic risk for AD in the Indian/South ASN population is not well explored.

Except for rs7412 and rs429358, the other DSLs in the *TOMM40-APOE-APOC1* region are non-coding changes affecting gene expression. The genes in the region are all transcribed in the same direction, which raises the possibility that cis regulatory elements are co-regulating these genes.[26] *TOMM40* encodes the TOM40 protein, a subunit of the translocase of the outer membrane (TOM) complex.[27] TOM is crucial for mitochondrial functions including lipid synthesis, energy metabolism, cell apoptosis, and cellular homeostasis. Mitochondrial dysfunction has been recently emphasized in both pathological and non-pathological aging, and *TOMM40* plays a role in this process by affecting mitochondrial neurotoxicity.[28] *APOC1* encodes a member of the apolipoprotein C1 family, which is involved in regulating the metabolism of triglyceride-rich lipoproteins such as high-density lipoprotein and very-low-density lipoprotein metabolism.[28] The changes in these genes' expression due to the non-coding variants are implicated with AD through oxidative stress processes based on shared lipid metabolism and mitochondrial function.[29]

In *TOMM40* and *APOC1*, DSLs such as rs2075650, rs71352238, rs157582, and rs4420638 are in moderate LD with rs429358 (ε4 allele), except AA (Table S3). The associations with AD are not independent of the ε4 allele. In NHW population, the association failed to reach significance after *APOE* adjustment.[30] Interestingly, the SNPs, including rs2075650, rs71352238, and rs157582, in LD located on *TOMM40* showed significant heterogeneity among populations (Table 2). Interestingly, AA had the largest OR for rs71352238, the second largest OR for rs157582, and the third largest OR for rs71352238 compared to the other populations. The lowest correlation between them was confirmed in AA based on subpopulations in 1000 Genomes Project (Table S3). Even though rs2075650 was reported to have a significantly

strong association with AD across NHW, HISP, and ASN populations, the high heterogeneity (Cochran's $Q = .000$, $I^2 = 93.3\%$) was previously suggested in the meta-analysis with the populations.[31] Also, the strong association in the Korean population was not replicated in the Chinese population.[32]

In signal-induced proliferation-associated 1 like 2 (*SIPA1L2*), rs115684722 had been discovered in AA-specific ancestry,[33] even though the function of the gene has yet to be determined. Our result confirmed that the loci had a significant OR only in AA. We would like to note that the signal was not significant in the combined analysis of all populations, thereby showing the importance of heterogeneity and/or stratified analysis in multi-ethnic studies.

Solute carrier family 24 (sodium/potassium/calcium exchanger), member 4 (*SLC24A4*) plays a role in calcium transport and lipid and glucose metabolism. Thus, it is involved in amyloid beta (Aβ) loading and tau pathology and contributes to AD risk.[34,35] Although rs7401792 was identified in a large meta-analysis, dominated by a NHW population, in our analysis this SNP had a significant OR in the admixed HISP population only, while the heterogeneity across populations had not been reported previously.

Our study had several limitations. First, most of the DSLs, which were selected from large GWASs or meta-analysis studies of AD, were not significant in NIAGADS due to small sample size. Additionally, different subsets of the analyzed NIAGADS cohort had often been included in the discovery GWAS of those DSL. Second, our ASN cohort consisted of predominantly subjects from LASI, which originated in India. The AD affection rate (7.12%) in the ASN population is much smaller compared to the other populations. However, this would only affect the size of the confidence interval, but not the effect estimate itself (Figure 2B). Interestingly, recent studies suggest several AD subtypes based on omics data, such as genomics, transcriptomics, or proteomics, where the *APOE* genotype is insufficient to fully account for the differences observed among AD subtypes.[36–39] Indeed, variations in race and ethnicity among AD subtypes may support the heterogeneity of genetic risk factors within specific populations.[40] However, the studies of non-European ancestries, especially in the Indian/South ASN population, remain very limited. The distribution of biological AD subtypes could vary across different genetic ancestries and should be more intensively investigated in large multi-ethnic populations.

In conclusion, our finding for the heterogeneity of *TOMM40/APOE/APOC1* genes between populations supports the idea that the genetic architecture of AD is partially ancestry-specific. We observed a much smaller effect of *APOE* ε4 in the South ASN population. All three genes are involved in multiple processes related to lipid metabolism, mitochondrial function, and AD risk; research into these genes and how they interact is ongoing. In addition, we highlight the importance of ancestry-specific GWASs by showing that *SIPA1L2* was only significant in the AA and *SLC24A4* only in the HISP, but not in the combined analysis. Therefore, ancestry-specific genetic involvement for AD may lead to better understanding of these genes implicated in AD development. More investigation with larger and independent cohorts is required for the validation.

AUTHOR CONTRIBUTIONS

Christoph Lange conceptualized and designed the project. Sanghun Lee and Dmitry Prokopenko performed statistical analyses and interpretation and drafted the manuscript. Julian Hecker and Georg Hahn assisted the analyses and manuscript preparation. Kristina Mullin assisted with project administration, data access, and data curation. Christoph Lange and Rudolph E. Tanzi obtained funding. Christoph Lange, Dmitry Prokopenko, and Rudolph E. Tanzi supervised this work. All authors contributed to the critical revision of the manuscript. All authors contributed to the relevant sections and approved the final manuscript.

ACKNOWLEDGEMENTS

The computations in this paper were run in part on the Faculty of Arts & Sciences Research Computing Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University. The funding body had no role in the design of the study, in the collection, analysis, and interpretation of data, or in writing the manuscript. Please refer to the Supplementary Note for full acknowledgements. This work was supported by Cure Alzheimer's Fund. SML was supported by NIH R01MH129337.

CONFLICT OF INTEREST STATEMENT

All authors declare that they have no potential conflicts of interest related to this work. Author disclosures are available in the supporting information.

CONSENT STATEMENT

This study was conducted in accordance with the revised Declaration of Helsinki and Good Clinical Practice guidelines and was approved by the relevant Institutional Review Boards (IRBs, protocol no. IRB20-0028 from Harvard School of Public Health and protocol no. 2019P001879 from Massachusetts General Hospital). Informed consent was obtained from all subjects.

ORCID

Sanghun Lee https://orcid.org/0000-0002-0573-9555
Dmitry Prokopenko https://orcid.org/0000-0002-1844-5652

REFERENCES

1. Bellenguez C, Grenier-Boley B, Lambert JC. Genetics of Alzheimer's disease: where we are, and where we are going. *Curr Opin Neurobiol*. 2020;61:40-48.
2. Miyashita A, Koike A, Jun G, et al. SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians. *PLoS One*. 2013;8:e58618.
3. Reitz C, Jun G, Naj A, Rajbhandary R, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ε4,and the risk of late-onset Alzheimer disease in African Americans. *JAMA*. 2013;309:1483-1492.
4. Jun GR, Chung J, Mez J, et al. Transethnic genome-wide scan identifies novel Alzheimer's disease loci. *Alzheimers Dement*. 2017;13:727-738.
5. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet*. 2019;51:414-430.

6. Jian X, Sofer T, Tarraf W, et al. Genome-wide association study of cognitive function in diverse Hispanics/Latinos: results from the Hispanic Community Health Study/Study of Latinos. *Transl Psychiatry*. 2020;10:245.
7. Horimoto A, Xue D, Thornton TA, Blue EE. Admixture mapping reveals the association between Native American ancestry at 3q13.11 and reduced risk of Alzheimer's disease in Caribbean Hispanics. *Alzheimers Res Ther*. 2021;13:122.
8. Jia L, Li F, Wei C, et al. Prediction of Alzheimer's disease using multi-variants from a Chinese genome-wide association study. *Brain*. 2021;144:924-937.
9. Shigemizu D, Mitsumori R, Akiyama S, et al. Ethnic and trans-ethnic genome-wide association studies identify new loci influencing Japanese Alzheimer's disease risk. *Transl Psychiatry*. 2021;11:151.
10. Wightman DP, Jansen IE, Savage JE, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat Genet*. 2021;53:1276-1282.
11. Bellenguez C, Kucukali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet*. 2022;54:412-436.
12. Miyashita A, Kikuchi M, Hara N, Ikeuchi T. Genetics of Alzheimer's disease: an East Asian perspective. *J Hum Genet*. 2023;68:115-124.
13. Sherva R, Zhang R, Sahelijo N, et al. African ancestry GWAS of dementia in a large military cohort identifies significant risk loci. *Mol Psychiatry*. 2023;28:1293-1302.
14. Zhang DF, Xu M, Bi R, Yao YG. Genetic analyses of Alzheimer's disease in China: achievements and perspectives. *ACS Chem Neurosci*. 2019;10:890-901.
15. Andrews SJ, Renton AE, Fulton-Howard B, Podlesny-Drabiniok A, Marcora E, Goate AM. The complex genetic architecture of Alzheimer's disease: novel insights and future directions. *EBioMedicine*. 2023;90:104511.
16. Reitz C, Pericak-Vance MA, Foroud T, Mayeux R. A global view of the genetic basis of Alzheimer disease. *Nat Rev Neurol*. 2023;19:261-277.
17. Choi KY, Lee JJ, Gunasekaran TI, et al. APOE promoter polymorphism-219T/G is an effect modifier of the influence of APOE epsilon4 on Alzheimer's disease risk in a multiracial sample. *J Clin Med*. 2019;8:1236.
18. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet*. 2007;39:17-23.
19. Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer disease meta analysis consortium. *JAMA*. 1997;278:1349-1356.
20. Liu M, Bian C, Zhang J, Wen F. Apolipoprotein E gene polymorphism and Alzheimer's disease in Chinese population: a meta-analysis. *Sci Rep*. 2014;4:4383.
21. Agarwal R, Tripathi CB. Association of apolipoprotein E genetic variation in Alzheimer's disease in Indian population: a meta-analysis. *Am J Alzheimers Dis Other Demen*. 2014;29:575-582.
22. Beecham GW, Bis JC, Martin ER, et al. The Alzheimer's disease sequencing project: study design and sample selection. *Neurol Genet*. 2017;3:e194.
23. Jin H, Chien S, Meijer E, Khobragade P, Lee J. Learning from clinical consensus diagnosis in india to facilitate automatic classification of dementia: machine learning study. *JMIR Ment Health*. 2021;8:e27113.
24. Prive F. Using the UK Biobank as a global reference of worldwide populations: application to measuring ancestry diversity from GWAS summary statistics. *Bioinformatics*. 2022;38:3477-3480.
25. Smith JA, Zhao W, Yu M, et al. Association between episodic memory and genetic risk factors for Alzheimer's disease in south Asians from the longitudinal aging study in india-diagnostic assessment of dementia (LASI-DAD). *J Am Geriatr Soc*. 2020;68:S45-S53. Suppl 3.
26. Yashin AI, Fang F, Kovtun M, et al. Hidden heterogeneity in Alzheimer's disease: insights from genetic association studies and other analyses. *Exp Gerontol*. 2018;107:148-160.
27. Humphries AD, Streimann IC, Stojanovski D, et al. Dissection of the mitochondrial import and assembly pathway for human Tom40. *J Biol Chem*. 2005;280:11535-11543.
28. Cervantes S, Samaranch L, Vidal-Taboada JM, et al. Genetic variation in APOE cluster region and Alzheimer's disease risk. *Neurobiol Aging*. 2011;32:2107. e7-17.
29. Prendecki M, Florczak-Wyspianska J, Kowalska M, et al. Biothiols and oxidative stress markers and polymorphisms of TOMM40 and APOC1 genes in Alzheimer's disease patients. *Oncotarget*. 2018;9:35207-35225.
30. Blue EE, Cheng A, Chen S, Yu CE. Alzheimer's disease genetics c. association of uncommon, noncoding variants in the apoe region with risk of Alzheimer disease in adults of European ancestry. *JAMA Netw Open*. 2020;3:e2017666.
31. Huang H, Zhao J, Xu B, et al. The TOMM40 gene rs2075650 polymorphism contributes to Alzheimer's disease in Caucasian, and Asian populations. *Neurosci Lett*. 2016;628:142-146.
32. He Y, Li C, Yang Y, et al. Meta-analysis of the rs2075650 polymorphism and risk of Alzheimer disease. *Aging Clin Exp Res*. 2016;28:805-811.
33. Kunkle BW, Schmidt M, Klein HU, et al. Novel Alzheimer disease risk loci and pathways in African American individuals using the African genome resources panel: a meta-analysis. *JAMA Neurol*. 2021;78:102-113.
34. Rosenthal SL, Kamboh MI. Late-Onset Alzheimer's disease genes and the potentially implicated pathways. *Curr Genet Med Rep*. 2014;2:85-101.
35. Stage E, Duran T, Risacher SL, et al. The effect of the top 20 Alzheimer disease risk genes on gray-matter density and FDG PET brain metabolism. *Alzheimers Dement (Amst)*. 2016;5:53-66.
36. Elman JA, Schork NJ, Rangan AV. Alzheimer's Disease Neuroimaging I. Exploring the genetic heterogeneity of Alzheimer's disease: evidence for genetic subtypes. *medRxiv*. 2023.
37. Ferreira D, Nordberg A, Westman E. Biological subtypes of Alzheimer disease: a systematic review and meta-analysis. *Neurology*. 2020;94:436-448.
38. Neff RA, Wang M, Vatansever S, et al. Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets. *Sci Adv*. 2021;7:eabb5398.
39. Tijms BM, Gobom J, Reus L, et al. Pathophysiological subtypes of Alzheimer's disease based on cerebrospinal fluid proteomics. *Brain*. 2020;143:3776-3792.
40. Wang D, Ma X, Schulz PE, Jiang X, Kim Y. Knowledge-guided deep temporal clustering for Alzheimer's disease subtypes in completed clinical trials. *medRxiv*. 2023.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lee S, Hecker J, Hahn G, et al. On the effect heterogeneity of established disease susceptibility loci for Alzheimer's disease across different genetic ancestries. *Alzheimer's Dement*. 2024;20:3397–3405.
https://doi.org/10.1002/alz.13796