

UCSF

UC San Francisco Previously Published Works

Title

Assessing the lack of diversity in genetics research across neurodegenerative diseases:
A systematic review of the GWAS Catalog and literature

Permalink

<https://escholarship.org/uc/item/6c73k7kh>

Journal

Alzheimer's & Dementia, 20(8)

ISSN

1552-5260

Authors

Jonson, Caroline

Levine, Kristin S

Lake, Julie

et al.

Publication Date

2024-08-01

DOI

10.1002/alz.13873

Peer reviewed

REVIEW ARTICLE

Assessing the lack of diversity in genetics research across neurodegenerative diseases: A systematic review of the GWAS Catalog and literature

Caroline Jonson^{1,2,3,4}  | Kristin S. Levine^{1,2} | Julie Lake^{1,5} | Linnea Hertslet¹ |
Lietsel Jones^{1,2} | Dhairya Patel⁶ | Jeff Kim^{1,5} | Sara Bandres-Ciga¹ | Nancy Terry⁷ |
Ignacio F. Mata⁸ | Cornelis Blauwendraat^{1,6} | Andrew B. Singleton^{1,5} |
Mike A. Nalls^{1,2,5} | Jennifer S. Yokoyama^{3,4,9} | Hampton L. Leonard^{1,2,5,10}

¹Center for Alzheimer's and Related Dementias, National Institutes of Health, Bethesda, Maryland, USA

²DataTecnica LLC, Washington, District of Columbia, USA

³Pharmaceutical Sciences and Pharmacogenomics Graduate Program, University of California, San Francisco, California, USA

⁴Memory and Aging Center, Department of Neurology, Weill Institute for Neurosciences, University of California, San Francisco, California, USA

⁵Laboratory of Neurogenetics, National Institutes on Aging, National Institutes of Health, Bethesda, Maryland, USA

⁶Integrative Neurogenomics Unit, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA

⁷Division of Library Services, Office of Research Services, National Institutes of Health, Bethesda, Maryland, USA

⁸Genomic Medicine Institute, Lerner Research Institute, Genomic Medicine, Cleveland Clinic Foundation, Cleveland, Ohio, USA

⁹Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, California, USA

¹⁰German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany

Correspondence

Mike A. Nalls and Hampton L. Leonard, at the Center for Alzheimer's and Related Dementias, Building T44, Room 102, NIH Campus, Bethesda, MD 20892, USA.
Email: nallsm@nih.gov and hampton.leonard@nih.gov

Funding information

Intramural Research Program; National Institute on Aging; Department of Health and Human Services, Grant/Award Number: Z01AG000534; National Institute of Neurological Disorders and Stroke, Grant/Award Numbers: R01AG062588, R01AG057234, P30AG062422, P01AG019724, U19AG079774, U54NS123985, 75N95022C00031; Rainwater Charitable Foundation; Alzheimer's Association; Global Brain Health Institute; and the Mary Oakley Foundation; National Institute on Drug Abuse

Abstract

The under-representation of non-European cohorts in neurodegenerative disease genome-wide association studies (GWAS) hampers precision medicine efforts. Despite the inherent genetic and phenotypic diversity in these diseases, GWAS research consistently exhibits a disproportionate emphasis on participants of European ancestry. This study reviews GWAS up to 2022, focusing on non-European or multi-ancestry neurodegeneration studies. We conducted a systematic review of GWAS results and publications up to 2022, focusing on non-European or multi-ancestry neurodegeneration studies. Rigorous article inclusion and quality assessment methods were employed. Of 123 neurodegenerative disease (NDD) GWAS reviewed, 82% predominantly featured European ancestry participants. A single European study identified over 90 risk loci, compared to a total of 50 novel loci identified in all non-European or multi-ancestry studies. Notably, only six of the loci have been replicated. The significant under-representation of non-European ancestries in NDD GWAS

Caroline Jonson, Kristin S. Levine, and Julie Lake contributed equally to this work and considered as co-first authors.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

hinders comprehensive genetic understanding. Prioritizing genomic diversity in future research is crucial for advancing NDD therapies and understanding.

KEYWORDS

Alzheimer's disease, amyotrophic lateral sclerosis, ancestral diversity, genetic research disparities, genome-wide association study, multi-ancestry cohorts, neurodegenerative diseases, Parkinson's disease, population genetics, precision medicine

Highlights

- Eighty-two percent of neurodegenerative genome-wide association studies (GWAS) focus on Europeans.
- Only 6 of 50 novel neurodegenerative disease (NDD) genetic loci have been replicated.
- Lack of diversity significantly hampers understanding of NDDs.
- Increasing diversity in NDD genetic research is urgently required.
- New initiatives are aiming to enhance diversity in NDD research.

1 | INTRODUCTION

Genome-wide association studies (GWAS) have shown a significant bias toward individuals of European ancestry, despite comprising only 16% of the global population.¹ This under-representation issue is particularly salient in the realm of neurodegenerative disease (NDD) studies. For instance, although a recent Alzheimer's disease (AD) GWAS including $\approx 800,000$ individuals of European descent identified 75 disease-associated loci,² no GWAS studies on AD currently exist for Admixed American or Native American populations. Similarly, Parkinson's disease (PD) research exhibits a glaring imbalance, with Black individuals included in just $\approx 4\%$ of published PD studies.³

The prevalence of NDDs varies significantly among global populations and racial/ethnic groups. This warrants a critical examination of the disparity in genetic research efforts over time. In this article, we present a systematic review spanning 2012 through 2022, focusing on NDD GWAS research.

Our analysis encompasses common NDDs such as AD, PD, and amyotrophic lateral sclerosis (ALS), as well as less common atypical dementias. Our objective is to quantify the disparity in participant recruitment for genetic studies, shed light on genetic findings in under-represented populations, and discuss ongoing initiatives aimed at addressing this pervasive issue.

2 | METHODS

2.1 | Search strategy

The systematic review was conducted in two phases on seven of the most common NDDs: Alzheimer's disease (or AD), Parkinson's disease (or PD), amyotrophic lateral sclerosis (or ALS), multiple sclerosis (MS), frontotemporal dementia (FTD), myasthenia gravis (MG), Lewy body

dementia (LBD), and vascular dementia (VaD). First, we reviewed the GWAS Catalog; then, because the GWAS Catalog does not include all GWAS studies, we performed a formal literature review in collaboration with the National Library of Medicine (NLM). The keywords used in both searches are included in Table S1. The GWAS Catalog uses precise indexing terms, and our targeted keyword strategy was intended to maximize the retrieval of studies, thereby ensuring a comprehensive review. By including specific onset-related terms, we aimed to cover all types of NDD GWAS research relevant to our review, acknowledging that the correct keywords can significantly enhance study inclusion.

Results from both the GWAS Catalog and the NLM search were uploaded to Covidence,⁴ a web-based software platform, for further review. We removed duplicate studies and any studies published before 2012 or after 2022.

All studies were filtered to include only genome-wide associations examining neurological disease risk factors, family history of disease, disease progression, age at onset, or survival genome-wide, excluding exome-wide studies and those that focused on a targeted set of single nucleotide polymorphisms (SNPs) or genetic loci. Studies investigating disease subtypes, biomarkers, or non-English-language studies, and those investigating only rare or structural variation were also excluded.

Studies were assessed for eligibility by two independent reviewers and all conflicts were resolved by a third independent reviewer. Publication date, phenotype, and cohort information were extracted from each publication. If multiple phenotypes of interest were analyzed in the same study, information was included in both phenotype categories. The number of samples per ancestry was extracted manually from each study. We looked at seven ancestry groupings: European (EUR), East Asian (EAS), Middle-Eastern (MDE), African (AFR), African American and Caribbean (AAC), Latino and Indigenous Americas populations (AMR), and South Asian (SAS). If more than one ancestry group was present, the study was labeled multi-ancestry (MULTI). In this article, multi-ancestry refers to a combination of ancestral groups rather

than individuals who are mixtures of ancestral backgrounds. A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram of our filtering process can be found in Figure S1. All 123 studies passing our filters can be found in Table S2.

Finally, results were examined manually for all studies passing implemented filtering methods, flagging novel loci discovered in non-European or multi-ancestry populations with a *p*-value below 5E-8. We checked for replication of novel loci using the [Open Targets](#) platform.

3 | RESULTS

3.1 | Search results

We identified **123** eligible GWAS studies. Unsurprisingly, we found that European populations were over-represented in GWAS pertaining to NDDs. When non-European populations were included, the sample sizes were on average 15× smaller than the European ancestry samples included in the same disease category. The under-representation of non-European populations was particularly evident among the less common NDDs, including LBD, FTD, and VaD, where we did not identify any non-European or multi-ancestry GWAS studies using the outlined search methods. We have summarized the lack of diversity in genetic studies of NDDs in Table 1 and Figure 1.

We found 50 novel NDD loci that were identified in non-European or multi-ancestry populations (Table 2). Of these 50 loci, 24 were found in multi-ancestry studies, 23 were found in East Asian studies, and only 3 were found in other populations (AAC and AMR). No loci were discovered in AFR, MDE, or SAS ancestries. Recent studies that combined individuals of multiple ancestries by using standard random-effects and some custom meta-analytic techniques⁵ have succeeded in identifying novel disease loci that reach genome-wide significance, including two novel AD loci⁶ and 12 novel PD loci.⁷ However, these studies leverage existing European sample sizes as a backbone for much of the statistical power needed for discovery.

Of the 50 novel loci, we were could find replication for only 6 (Table 2). Of those six loci, five were found in multi-ancestry cohorts, which are predominantly powered by European samples. One locus was found in an East Asian population and later replicated in another East Asian population. Genome wide replication is an issue for non-European studies, because they are often very small and include all samples available. This further highlights the importance of conducting more and larger non-European studies. As data sets become larger and more inclusive, the genetic architecture of these diseases may grow and change.

In the following sections, we briefly summarize the results of our systematic review in a disease-specific manner.

Alzheimer's Disease

Largest European GWAS: Wightman 2021

Total samples: 1,126,563

Largest multi-ancestry GWAS: Lake 2022

Total samples: 644,188

Total non-European samples: 18,246

RESEARCH IN CONTEXT

- 1. Systematic review:** We conducted a systematic review utilizing existing genome-wide association studies (GWAS) results and publications, curated from the GWAS Catalog and with the National Library of Medicine, to assess the inclusion of diverse ancestry groups in neurodegeneration and neurogenetics studies. We employed rigorous methods for the inclusion of identified articles and quality assessment.
- 2. Interpretation:** Our findings underscore the pressing need for increased diversity in neurodegenerative research. The significant under-representation of non-European ancestry participants in NDD GWAS limits our understanding of the genetic underpinnings of these diseases.
- 3. Future directions:** Our work underscores the need for more inclusive research approaches in neurodegenerative diseases that emphasize multi-ancestry and non-European populations to advance precision medicine and develop treatments effective for diverse populations. Our review showed that diverse GWAS in non-European populations have enhanced power, led to the discovery of new loci, and provided a more comprehensive understanding of NDD genetic architecture beyond just European contexts.

% non-European: 2.83%

Largest non-European GWAS: Sherva 2022

Ancestry: AAC

Total samples: 75,058

Loci discovered in non-EUR or multi-ancestry studies: 14

Loci replicated: 1 (SORL1)

The largest AD GWAS of European populations included ≈1.1 million individuals and identified a total of 38 associated loci.⁸ Another recent GWAS included ≈800,000 individuals of European ancestry and identified a total of 75 loci.² The discrepancy between identified loci in these studies could be due to many factors, including differences in neuropathological/diagnostic criteria.⁹

A 2013 GWAS conducted in African Americans replicated an association at *ABCA7* previously identified in European populations. That GWAS found that rs115550680, rare in European populations, was associated with an increased risk for AD in African Americans comparable to the highly pathogenic apolipoprotein E (*APOE*) ε4 variant observed in Europeans.¹⁰ A 2017 GWAS in African Americans identified two novel loci at *COBL* and *SLC10A2*.¹¹ The most extensive African American GWAS to date, drawing from a military cohort of ≈22,000 individuals and a proxy GWAS involving ≈50,000 individuals, identified significant associations with established AD risk genes such as *TREM2*, *CD2AP*, and *ABCA7*. Notably, distinct lead variants

TABLE 1 Largest GWAS sample size by NDD and ancestry for single and multi-ancestry studies.

NDD	Ancestry	Author	Year	Total samples (% non-European*)	N _{proxycases} N _{cases} N _{controls}	Minor Allele Frequency threshold (MAF)	Study (DOI)
AD	AAC	Sherva	2022	75,058	6641 4012 64,405	1%	African ancestry GWAS of dementia in a large military cohort identifies significant risk loci (10.1038/s41380-022-01890-3)
	EAS	Hirano	2015	17,031	0 1827 15,204	1%	A genome-wide association study of late-onset Alzheimer's disease in a Japanese population (10.1097/YPG.0000000000000090)
	EUR	Wightman	2021	1,126,563	46,613 43,725 1,036,225	~0.06%	A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease (10.1038/s41588-021-00921-z)
	MULTI	Lake	2022	644,188 (2.83%)	46,828 54,233 543,127	1%	Multi-ancestry meta-analysis and fine-mapping in Alzheimer's disease (10.1038/s41380-023-02089-w)
	AMR, MDE		NO STUDIES				
PD	AMR	Loesch	2021	1497	0 807 690	1%	Characterizing the genetic architecture of Parkinson's disease in Latinos (10.1002/ana.26153)
	EAS	Foo	2017	14,006	0 779 13,227	1%	Genome-wide association study of Parkinson's disease in East Asians (10.1093/hmg/ddw379)
	EUR	Nalls	2019	1,474,097	18,618 37,688 1,417,791	1% & 5%	Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies (10.1016/S1474-4422(19)30320-5)
	MULTI	Kim	2022	2,525,730 (38.12%)	18,618 49,049 2,458,063	0.1%	Multi-ancestry genome-wide meta-analysis in Parkinson's disease (10.1038/s41588-023-01584-8)
	AAC, MDE	NO STUDIES					
ALS	EAS	Wei	2019	4727	0 700 4027	5%	Identification of TYW3/CRYZ and FGD4 as susceptibility genes for amyotrophic lateral sclerosis (10.1212/NXG.0000000000000375)
	EUR	van Rheenen	2016	41,398	0 15,156 26,242	1%	Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis (10.1038/ng.3622)
	MULTI	van Rheenen	2021	152,268 (12.00%)	0 29,612 122,656	0.1%	Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology (10.1038/s41588-021-00973-1)
	AAC, AMR, MDE	NO STUDIES					
MS	AAC	Isobe	2015	2319	0 803 1516	1%	An ImmunoChip study of multiple sclerosis risk in African Americans (10.1093/brain/awv078)
	AMR	Ordoñez	2015	161	0 29 132	1%	Genome wide admixture study in Mexican Mestizos with multiple sclerosis (10.1016/j.clineuro.2014.11.026)
	EUR	Patsopoulos	2019	115,803	0 47,429 68,374	1%	Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility (10.1126/science.aav7188)
	EAS, MDE, MULTI	NO STUDIES					

(Continues)

TABLE 1 (Continued)

NDD	Ancestry	Author	Year	Total samples (% non-European*)	$N_{\text{proxycases}}$ N_{cases} N_{controls}	Minor Allele Frequency threshold (MAF)	Study (DOI)
FTD	EUR	Ferrari	2014	12,928	0 3526 9402		Frontotemporal dementia and its subtypes: a genome-wide association study (10.1016/S1474-4422(14)70065-1)
	AAC, AMR, EAS, MDE, MULTI	NO STUDIES					
MG	EAS	Na	2014	259	0 109 150	1%	Whole-genome analysis in Korean patients with autoimmune myasthenia gravis (10.3349/ymj.2014.55.3.660)
	EUR	Chia	2022	45,675	0 2227 43,448	0.1%	Identification of genetic risk loci and prioritization of genes and pathways for myasthenia gravis: a genome-wide association study (10.1073/pnas.2108672119)
	MULTI	Sakaue	2021	533,853 (33.48%)	0 278 533,575	0.01%	A cross-population atlas of genetic associations for 220 human phenotypes (10.1038/s41588-021-00931-x)
	AAC, AMR, MDE	NO STUDIES					
LBD	EUR	Chia	2021	7372	0 2981 4391	1%	Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture (10.1038/s41588-021-00785-3)
	AAC, AMR, EAS, MDE, MULTI	NO STUDIES					
VaD	EUR	Moreno-Grau	2019	4830	0 1541 3289	1%	Genome-wide association analysis of dementia and its clinical endophenotypes reveal novel loci associated with Alzheimer's disease and three causality networks: The GR@ACE project (10.1016/j.jalz.2019.06.4950)
	MULTI	Fongang	2022	482,088 (2.40%)	0 4138 477,950	1%	A meta-analysis of genome-wide association studies identifies new genetic loci associated with all-cause and vascular dementia (10.1002/alz.056081)
	AAC, AMR, EAS, MDE	NO STUDIES					

Note: Minor allele frequency (MAF) thresholds are included as inappropriate selection of MAF thresholds can lead to spurious findings and misleading reports of heterogeneity.

Abbreviations: AAC, African American and Caribbean; AD, Alzheimer's disease; AFR, African; ALS, amyotrophic lateral sclerosis; AMR, Latino and Indigenous Americas populations; EAS, East Asian; EUR, European; FTD, frontotemporal dementia; GWAS, genome-wide association studies; LBD, Lewy body dementia; MDE, Middle Eastern; MG, myasthenia gravis; MS, multiple sclerosis; NDD, neurodegenerative disease; PD, Parkinson's disease; SAS, South Asian; VaD, vascular dementia.

*If applicable.

were observed in these loci compared to those found in European cohorts.¹²

The only study conducted in Caribbean Hispanic individuals was a 2017 study with 2451 cases and 2063 controls. Investigators found a novel and population specific locus near *FBXL7*.¹³ The lead SNP, rs75002042, is much more common in individuals with African ancestry compared to individuals of European ancestry, with minor allele

frequencies around 20% and 0.009%, respectively. This study also replicated six loci previously reported in European populations, including *FRMD4A*, *CELF1*, *FERMT2*, *SLC24A4/RIN3*, *ABCA7*, and *CD33*.¹³

The largest AD study in East Asian populations was conducted in Japanese participants with 1827 cases and 15,204 controls (discovery + replication),¹⁴ but authors did not nominate any genome-wide significant loci. More recent but smaller studies have since been

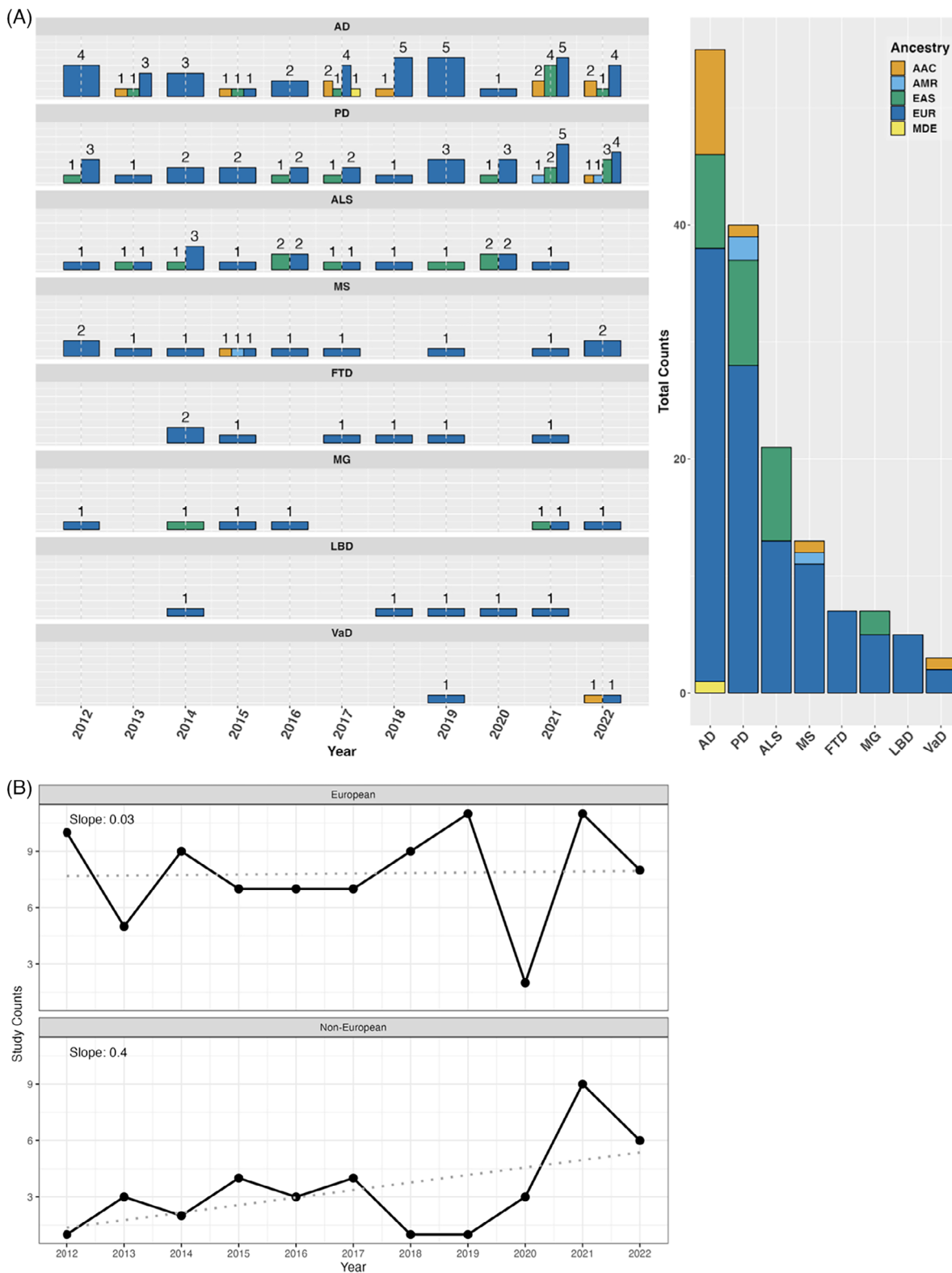


FIGURE 1 Number of studies over time from 2012 to 2022. (A) Bar plot of study counts by NDD (left) with cumulative counts for each ancestry (right). Data in this figure include both single and multi-ancestry studies. (B) Time series of the annual study counts in European and non-European populations from 2012 to 2022. The slope from a linear regression is also displayed to highlight the rate of change in the number of study counts over time. AAC, African American and Caribbean; AD, Alzheimer’s disease; AFR, African; ALS, amyotrophic lateral sclerosis; AMR, Latino and Indigenous Americas populations; EAS, East Asian; EUR, European; FTD, frontotemporal dementia; LBD, Lewy body dementia; MDE, Middle-Eastern; MG, myasthenia gravis; MS, multiple sclerosis; NDD, neurodegenerative disease; PD, Parkinson’s disease; SAS, South Asian; VaD, vascular dementia.

TABLE 2 Genome-wide significant (p -value < 5E-8) novel loci nominated in non-European populations or multi-ancestry studies.

NDD	Author	Year	Study	Ancestry	Nominated loci	p -value	Replicated (open targets)
AD	Tosto	2015	F-box/LRR-repeat protein 7 is genetically associated with Alzheimer's disease (10.1002/acn3.223)	AMR	FBXL7	6.19E-09	no
AD	Mez	2017	Two novel loci, COBL and SLC10A2, for Alzheimer's disease in African Americans (10.1016/j.jalz.2016.09.002)	AAC	COBL SLC10A2	3.80E-08 4.60E-08	no no
AD	Jia	2021	Prediction of Alzheimer's disease using multi-variants from a Chinese genome-wide association study (10.1093/brain/awaa364)	EAS	RHOBTB3/GLRX CTC/278L1.1 (rs6859823) CTD/2506/14.1 (rs234434) CHODL	3.07E-19 2.49E-23 1.35E-67 4.81E-09	no no no no
AD	Kang	2021	Potential Novel Genes for Late-Onset Alzheimer's Disease in East-Asian Descent Identified by APOE-Stratified Genome-Wide Association Study (10.3233/JAD-210145)	EAS	CACNA1A LRIG1	2.49E-08 1.51E-08	no no
AD	Shigemizu	2021	Ethnic and trans-ethnic genome-wide association studies identify new loci influencing Japanese Alzheimer's disease risk (10.1038/s41398-021-01272-3)	EAS MULTI	FAM47E OR2B2	5.34E-09 2.14E-08	no no
AD	Miyashita	2013	SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians (10.1371/journal.pone.0058618)	MULTI	SORL1	1.04E-08	Schwartzentruber J (2021), Wightman DP (2021), Marioni RE (2018), Jansen IE (2019), Moreno-Grau S (2019)
AD	Lake	2022*	Multi-ancestry meta-analysis and fine-mapping in Alzheimer's Disease (10.1038/s41380-023-02089-w)	MULTI	TRANK1 VWA5B2	3.49E-08 3.75E-08	no no
ALS	Deng	2013	Genome-wide association analyses in Han Chinese identify two new susceptibility loci for amyotrophic lateral sclerosis (10.1038/ng.2627)	EAS	CAMK1G CABIN1/SUSD2	2.92E-08 2.35E-09	Wei L (2019) Wei L (2019)
ALS	Xie	2014	Genome-wide association study combining pathway analysis for typical sporadic amyotrophic lateral sclerosis in Chinese Han populations (10.1016/j.neurobiolaging.2014.01.014)	EAS	INPP5B IQCF5/IQCF1 ITGA9 PFKP MYO18B ALCAM OPCML GPR133	2.24E-08 2.06E-09 2.55E-08 2.46E-09 2.28E-10 4.00E-08 8.43E-09 8.45E-10	no no no no no no no no

(Continues)

TABLE 2 (Continued)

NDD	Author	Year	Study	Ancestry	Nominated loci	p-value	Replicated (open targets)
ALS	Wei	2019	Identification of TYW3/CRYZ and FGD4 as susceptibility genes for amyotrophic lateral sclerosis (10.1212/NXG.0000000000000375)	EAS	TYW3/CRYZ	2.10E-14	no
ALS	Benyamin	2017	Cross-ethnic meta-analysis identifies association of the GPX3-TNIP1 locus with amyotrophic lateral sclerosis (10.1038/s41467-017-00471-1)	MULTI	FGD4	5.19E-09	no
ALS	Benyamin	2017	Cross-ethnic meta-analysis identifies association of the GPX3-TNIP1 locus with amyotrophic lateral sclerosis (10.1038/s41467-017-00471-1)	MULTI	GPX3/TNIP1	1.30E-08	van Rheenew W (2021), Nicolas A (2018)
ALS	Nakamura	2020	A multi-ethnic meta-analysis identifies novel genes, including ACSL5, associated with amyotrophic lateral sclerosis (10.1038/s42003-020-01251-2)	MULTI	ACSL5	2.97E-08	Iacoangeli (2020)
ALS	van Rheenew	2021	Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology (10.1038/s41588-021-00973-1)	MULTI	SLC9A8/SPATA2	3.20E-10	no
PD	Li	2022	Genetic Determinants of Survival in Parkinson's Disease in the Asian Population (10.1002/mds.29069)	EAS	ERGIC1 NEK1	5.60E-09 6.90E-09	no no
PD	Li	2021	Genetic Modifiers of Age at Onset for Parkinson's Disease in Asians: A Genome-Wide Association Study (10.1002/mds.28621)	EAS	COG3 PTPRN2	1.20E-08 1.80E-08	no no
PD	Foo	2020	Identification of Risk Loci for Parkinson Disease in Asians and Comparison of Risk Between Asians and Europeans: A Genome-Wide Association Study (10.1001/jama.2020.0428)	EAS	SV2C WBSCR17	3.48E-08 2.53E-08	no no
PD	Li	2021	Genetic Modifiers of Age at Onset for Parkinson's Disease in Asians: A Genome-Wide Association Study (10.1002/mds.28621)	EAS	NDN/PWRN4	3.14E-09	no

(Continues)

TABLE 2 (Continued)

NDD	Author	Year	Study	Ancestry	Nominated loci	p-value	Replicated (open targets)
PD	Li	2022	Genetic Determinants of Survival in Parkinson's Disease in the Asian Population (10.1002/mds.29069)	EAS	RPL3	2.72E-08	no
PD	Lill	2012	Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database (10.1371/journal.pgen.1002548)	MULTI	ITGA8	1.30E-08	Nalls MA (2019), Chang D (2017)
PD	Kim	2022*	Multi-ancestry genome-wide meta-analysis in Parkinson's disease (10.1038/s41588-023-01584-8)	MULTI	MTF2 PIK3CA ADD1 SYBU IRS2 USP8 PIGL FASN MYLK2 USP25 EP300 PPP6R2	1.15E-10 1.65E-10 4.11E-09 3.62E-09 2.30E-09 6.45E-10 2.93E-09 2.61E-09 3.86E-09 1.12E-09 3.81E-09 4.09E-10	no no no no no no no no no no no no

Note: We found no diverse or multi-ancestry loci for LBD, FTD, MG, or VaD. p-values are given in parentheses. Nominated loci were determined as the nearest gene or genomic context within 1 B of the significant SNP. We looked in Open Targets to determine replicated loci. The * indicates the study was originally available as a preprint.
Abbreviations: AAC, African American and Caribbean; AD, Alzheimer's disease; ALS, amyotrophic lateral sclerosis; AMR, Latino and Indigenous Americas populations; EAS, East Asian; LBD, Lewy body dementia; NDD, neurodegenerative disease; PD, Parkinson's disease.

conducted, including a 2021 GWAS in a Chinese cohort that reported four novel loci near *RHOBTB3/GLRX*, *CTC/278L1.1*, *CTD/2506J14.1*, and *CHODL*¹⁵; a study in Japanese participants that nominated a locus in *FAM47E*¹⁶; and a study including both Korean and Japanese participants that nominated two novel loci at *CACNA1A* and *LRIG1*.¹⁷

Multi-ancestry studies have nominated additional AD loci; however, these studies still rely on Europeans as the majority population. *SORL1* was first identified as a risk locus for AD in a GWAS that included East Asian and European ancestry populations.¹⁸ Other multi-ancestry GWAS replicated loci nominated in previous studies as well as identified *OR2B2*,¹⁶ *TRANK1*, and *VWA5B2*¹⁹ as novel loci for AD.

Although the inclusion of diverse populations in genetic research for AD is arguably better than what is seen for some of the atypical dementias, the largest study size for a non-European population¹² was still only 7% of the total sample size for the largest European AD GWAS. Only one locus discovered in a non-European or multi-ancestry study has been replicated (*SORL1*).

Parkinson's Disease

Largest European GWAS: Nalls 2019

Total samples: 1,456,306

Largest multi-ancestry GWAS: Kim 2022

Total samples: 2,525,730

Total non-European samples: 962,735

% non-European: 38.12%

Largest non-European GWAS: Foo 2017

Ancestry: EAS

Total samples: 14,006

Loci discovered in non-EUR or multi-ancestry studies: 17

Loci replicated: 1 (*ITGA8*)

The largest meta-GWAS of PD risk in individuals of European ancestry found 90 significant risk signals across 78 genomic regions. The 90 nominated risk variants collectively explain roughly 16%–36% of the heritable risk of non-monozygotic or complex PD.²⁰

The largest study in East Asian populations (with exception of a study done in Japan before our review period²¹) was conducted with Han Chinese participants, replicating loci identified previously in European populations including *SNCA*, *LRKK2*, and *MCCC1* in their discovery GWAS of 14,006 participants.²² More recent studies in Chinese populations have nominated a locus on *NDN/PWRN4* associated with age at onset and a locus on *RPL3* associated with reduced survival.^{23,24} The first and most recent PD GWAS of a South American population was conducted in 2021, replicating an association at *SNCA* with 1497 participants.²⁵

Recently, more multi-ancestry studies have been conducted in PD, nominating novel loci for disease risk and age at onset including *ITGA8*, *SV2C*, and *WBSCR17*^{20,26,27} The largest meta-GWAS for PD, which included four ancestral populations, replicated 66 loci previously nominated in European studies as well as identified 12 novel loci: *MTF2*, *PIK3CA*, *ADD1*, *SYBU*, *IRS2*, *USP8*, *PIGL*, *FASN*, *MYLK2*, *USP25*, *EP300*, and *PPP6R2*.⁷

The largest non-European PD GWAS was in East Asian populations; however, only a few novel loci have been nominated in that ancestry.

Multi-ancestry studies have nominated more novel variants in recent studies, but much more work is needed to better understand the risk for PD in non-European populations. Only one locus discovered in a non-European or multi-ancestry study has been replicated (*ITGA8*).

Amyotrophic Lateral Sclerosis

Largest European GWAS: van Rheenen 2016

Total samples: 41,398

Largest multi-ancestry GWAS: van Rheenen 2021

Total samples: 152,268

Total non-European samples: 18,266

% non-European: 12.00%

Largest non-European GWAS: Wei 2019

Ancestry: EAS

Total samples: 4727

Loci discovered in non-EUR or multi-ancestry studies: 19

Loci replicated: 4 (*CAMK1G*, *CABIN1/SUSD2*, *GPX3/TNIP1*, *ACSL5*)

ALS GWAS in European populations have nominated a number of risk loci including *C9ORF72*, *UNC13A*, *C21orf2*, *SARM1*, *MOBP*, *SCFD1*, *TBKK1*, and *KIF5A*.^{28,29}

ALS is the only disease in our review for which more genome-wide significant novel loci have been identified in a non-European population than in the largest European-only study. The first GWAS of individuals with Chinese Han ancestry identified *CAMK1G* and *CABIN1/SUSD2* as susceptibility loci for ALS.³⁰ Later studies in the Han Chinese population nominated additional novel loci including *INPP5B*, *IQCF5/IQCF1*, *ITGA9*, *PFKP*, *MYO18B*, *ALCAM*, *OPCML*, *GPR133*, *TYW3/CRYZ*, and *FGD4*.^{31,32} With a total of 12 genome-wide significant loci, East Asian ancestry GWAS for ALS have nominated the most of any single non-European population covered in our review.

Multi-ancestry GWAS for ALS, which typically consist of European and East Asian ancestry populations, have been successful at nominating additional risk loci including *GPX3/TNIP1* and *ACSL5*.^{29,33,34} The largest ALS GWAS to date was a multi-ancestry study including more than 150,000 individuals of European and East Asian ancestry. This study identified a total of 15 risk loci for ALS, replicating previously identified and nominating 5 novel loci: *SLC9A8/SPATA2*, *ERGIC1*, *NEK1*, *COG3*, and *PTPRN2*.³⁵

Similar to PD, ALS GWAS including or focused on East Asian populations have made more progress than other non-European populations for these diseases. However, much more work is still needed in all populations to progress potential precision medicine initiatives for ALS. Of the 19 loci discovered in non-European or multi-ancestry studies, only 4 have been replicated (*CAMK1G*, *CABIN1/SUSD2*, *GPX3/TNIP1*, *ACSL5*).

Multiple Sclerosis

Largest European GWAS: Patsopoulos 2019

Total samples: 115,803

There are no multi-ancestry studies in MS.

Largest non-European GWAS: Isobe 2015

Ancestry: AAC

Total samples: 2319

Loci discovered in non-EUR or multi-ancestry studies: 0

Loci replicated: 0

The largest GWAS meta-analysis for MS included 115,803 individuals of European ancestry and found 82 significant genome-wide associations with MS. This study was also the first to identify a risk locus for MS on chromosome X, and the identified genetic markers accounted for nearly 50% of the hereditary risk for MS.³⁶

Studies in non-European populations were more limited in MS than in the previous diseases discussed. The largest GWAS in African Americans was successful at replicating 21 of the loci previously identified in European populations but did not nominate any new risk loci at a genome-wide significant level.³⁷ The only other study nominated by our review process for MS in non-European populations was conducted in a Mexican population. The study found four significant variants; however, these variants had limited regional support and the study was severely underpowered, with only 29 cases and 132 controls.³⁸ Due to these limitations, we concluded that the variants identified in this study could not be classified as novel.

Sample sizes for MS GWAS are still relatively small, even for European populations. In addition, we did not find any multi-ancestry studies through our search methods, highlighting a potential opportunity for further discovery for this disease.

Frontotemporal Dementia

Largest European GWAS: Ferrari 2014

Total samples: 12,928

There are no multi-ancestry or non-European studies in FTD.

Common risk loci nominated by previous European FTD studies include *C9ORF72*, *GRN*, and *MAPT*.³⁹ The largest FTD GWAS in our review date range included ≈13,000 participants of European ancestry and nominated an additional locus in the *HLA-DRA/HLA-DRB5* region.⁴⁰ This study was conducted in 2014, and although more recent GWAS of FTD have been performed, none have surpassed the sample size from the Ferrari study, and many have focused on smaller FTD subtypes.^{41,42}

No non-European or multi-ancestry GWAS were identified in our systematic review for FTD. Investigation of known genetic risk factors in non-Europeans suggest that *C9ORF72* expansions may be quite rare in Chinese populations,⁴³ highlighting the need for further research in this area.

Myasthenia Gravis

Largest European GWAS: Sakaue 2021

Total samples: 355,142

Largest Multi-ancestry GWAS: Sakaue 2021

Total samples: 533,853

Total non-European samples: 178,711

% non-European: 33.47%

Largest non-European GWAS: Na 2014

Ancestry: EAS

Total samples: 259

Loci discovered in non-EUR or multi-ancestry studies: 0

Loci replicated: 0

Known loci for MG include *PTPN22*, *CTLA4*, *HLA-DQA1*, *ZBTB10*, and *TNFRSF11A*,^{44,45} all nominated in European-based GWAS. The most recent GWAS for MG nominated additional loci at *CHRNA1*, *SFMBT2*, and *FAM76B*, although the latter two did not replicate.⁴⁶ The largest

European and multi-ancestry GWAS for MG to date were both performed in the same study, leveraging 533,853 total samples from Japanese, UK, and Finnish-based biobanks. However, with only 278 cases, the effective sample size ($4/(1/n_{\text{case}}+1/n_{\text{control}})$) for the meta-analysis was insufficiently powered and no new loci were nominated for MG.⁴⁷

In non-European populations, the literature review identified one Korean GWAS for MG. However, this study was small and did not identify any loci meeting genome-wide significance.⁴⁸ Other studies have found that there is earlier onset of MG in Asian populations, and higher prevalence of the ocular form in Asian children, highlighting the importance of continued discovery efforts for MG in non-European populations.⁴⁹

Lewy Body Dementia

Largest European GWAS: Chia 2021

Total samples: 7372

There are no multi-ancestry or non-European studies in LBD.

Previously nominated risk loci for LBD include *GBA*, *APOE*, and *SNCA*.^{50,51} LBD can be hard to diagnose as there are a number of clinical and genetic overlaps with AD and PD, which may be one of the reasons why there is still limited genetic research for LBD in both European and non-European populations.^{50–52}

We found no LBD GWAS in any single non-European ancestry populations or any multi-ancestry studies through our search methods. Concrete data on the prevalence of LBD in diverse ancestries is difficult to acquire, showing a potential opportunity for valuable future research.

Vascular dementia

Largest European GWAS: Moreno-Grau 2019

Total samples: 4830

Largest Multi-ancestry GWAS: Fongang 2022

Total samples: 482,088

Total non-European samples: 11,590

% non-European: 2.40%

There are no non-European studies in VaD.

Despite an approximated prevalence of about 15%–20% in all dementia cases,⁵³ vascular dementia (or VaD) remains difficult to study because of the uncertainty of diagnosis. In fact, only two studies on VaD passed our criteria and only one of these found genome-wide significant novel loci. The first study was a European GWAS that looked at vascular, mixed, and pure AD phenotypes and nominated loci at *ANKRD31* and *NDUFAF6*.⁵⁴ The second study that passed our criteria was a multi-ancestry GWAS for all-cause and VaD including participants from European, African, Asian, and Hispanic ancestries, but did not find any significant novel loci.⁵⁵

VaD prevalence and risk appears to be higher in South Asian ancestries compared to European or Chinese populations.^{56,57} Additional studies have suggested that African Americans are most likely to be admitted to inpatient care with a VaD primary diagnosis.⁵⁸ Despite these findings, there are still limited genetic studies for VaD, and we found no single non-European GWAS, highlighting the need for future research.

4 | DISCUSSION

This review highlights the lack of ancestral diversity in genetic research across NDD GWAS over the past decade. Current research suggests that including non-European populations can improve our understanding of the genetic architecture of disease through novel ancestry-specific discoveries, increased statistical power awarded by studying diverse haplotype structures, and the identification of loci with heterogeneous effects across populations.⁵⁹ Although we identified 50 novel loci discovered in diverse populations in our review, only six of the loci have been replicated, highlighting the need for larger numbers of non-Europeans to be included in genetic studies.

Furthermore, discrepancies in how diseases are diagnosed across populations introduce additional heterogeneity, complicating the replication process. Variations in diagnostic criteria, health care access, and cultural factors can lead to differences in disease classification and severity, influencing the observed genetic associations. For instance, in NDDs, diagnostic criteria may vary between populations due to differences in symptom presentation, age at onset, or cultural perceptions of cognitive decline. These differences not only affect the phenotypic characterization of the disease but also may impact the genetic underpinnings identified through GWAS. Consequently, the failure to replicate GWAS findings across populations may stem not only from genetic diversity or insufficient power in smaller studies but also from variations in disease definitions and diagnostic practices. Addressing these challenges necessitates not only larger and more diverse sample sizes but also efforts to harmonize diagnostic criteria and standardize phenotypic assessments across populations.

In addition, although we looked at seven genetic ancestry groups, these “buckets” do not capture the true diversity of global populations. The African continent is known to have high genetic diversity, yet individuals of African ancestry are routinely grouped into a single category.⁶⁰ In fact, we found no studies investigating South Asian (SAS) or continental African (AFR) populations. After investigating the cohorts in our review, we noted that although there were multiple “African” labeled studies, none of them directly investigated individuals in continental Africa, instead looking at African American or other African admixed populations. It is critical to mention the reference population used to define the specific population, to prevent the misattribution of genetic features across ancestries. Grouping all participants with any African ancestry into a generalized African category obscures the significant issue of inadequate representation of continental Africans.

In addition, there has been very little research done on admixed populations, and how the combinations of different ancestries affect SNP frequencies and/or gene expression. A GWAS in a Caribbean Hispanic admixed population found that the frequency of a novel locus spanning *FBXL7* varied greatly, from 1% in those with European ancestry to 20% in African Americans.¹³

Furthermore, previous research has shown that the transferability of polygenic risk scores from African Americans to various African

populations is highly unreliable.⁶¹ The substantial genetic and environmental disparities among individuals of African descent underscore the urgent need to improve diversity in genetic studies.

Our review is not without limitations. One limitation is that due to studies using different versions of summary statistics, different release versions of cohorts, and different filtering and quality control pipelines, we could not accurately calculate the percentage of sample overlap across studies. It is important to note that disease cohorts or summary statistics are often included in multiple GWAS. This is especially common in multi-ancestry meta-analyses, which often rely on publicly available summary statistics to increase power.⁷ Potentially high sample overlap between studies such as Wightman et al. and Schwartzentruber et al., which both include cohorts like the UK Biobank, may spark debate about whether they can be counted as completely “unique” GWAS. However, we chose to include any GWAS in our review despite potential overlap if they included any additional or varying cohorts to paint an overall picture of the state of diversity in GWAS. The number of unique participants included in each subsequent GWAS across ancestries would be an interesting comparison for a future study, although it will likely highlight very little overlap in diverse ancestry studies as there are still too few studies being conducted in those populations.

Another limitation to our review is that we chose to focus on a 10-year time period. This time-span was chosen to cover the trends in the most recent decade in GWAS research in-depth. However, many older GWAS, including some in diverse ancestry populations, were conducted prior to 2012. In addition, important diverse ancestry studies have since been published beyond our search date of the National Institutes of Health (NIH) Library of Medicine and the GWAS Catalog. We chose to highlight a few of these foundational studies in our review despite being outside of the designated 10-year period. Future studies may choose a wider time span for reviewing GWAS studies, which may highlight an even larger disparity between European and diverse ancestry studies prior to 2012.

4.1 | Diversity in SNP discovery

Although six of the eight NDDs we investigated had non-European representation, only PD had >1000 cases and >30% non-European samples (Table 1). Furthermore, no new significant loci have been identified in diverse population studies for MS, LBD, FTD, VaD, or MG. Of the 50 novel significant loci that have been discovered, only six have been replicated. Although the largest non-European cohort in MG included almost 180,000 samples, only 81 MG cases were included. A GWAS with fewer than 1000 cases is unlikely to achieve sufficient statistical power for SNP discovery in polygenic diseases where multiple loci with small effect sizes are generally expected.^{62–64} Alzheimer's, the most well-funded of all the NDDs, has less than 3% diversity among cases in genetic studies. The incorporation of studies that lack statistical power and replicability minimizes the true imbalance between European and non-European studies, maintaining a Eurocentric bias.

Furthermore, the variation in minor allele frequency (MAF) thresholds across studies, as detailed in Table 1, illustrates a methodological challenge in such genetic research. The ability to detect rarer genetic variants in smaller studies is often limited by higher MAF thresholds, potentially omitting crucial findings in these populations. This limitation is particularly significant, as it suggests that some variants currently identified as present only in European populations may, in fact, be found in other ancestries as more comprehensive and diverse samples become available. The current perception of ancestry-specific genetic markers may thus evolve with the inclusion of broader, more representative data sets. Such developments could lead to the identification of previously unrecognized genetic diversity and the re-evaluation of the geographical and ancestral specificity of certain genetic variants.

In addition, many genetic association studies in East Asian populations did not meet our review criteria because they were not conducted on a genome-wide scale. Instead, these studies often investigated only one or a small group of SNPs that had been previously associated with disease in European populations, potentially missing associations that are specific to non-European populations.

In addition, our study may not have captured some non-European research, as we excluded studies not published in English. This decision was based on the understanding that over 95% of scientific literature is published in English; however, we acknowledge that this represents a slight limitation of our work.⁶⁵

Many loci identified in European populations have heterogeneous effects or ancestry-specific SNP associations. For example, although the *APOE* alleles account for around a quarter of overall heritability for AD in Europeans,^{8,16} several studies suggest that the *APOE4* allele has a weaker effect in African ancestry^{66,67} and Caribbean Hispanic⁶⁸ populations. The effect has been found to be greater in Japanese populations.^{66,67} Heterogeneity of risk at *APOE* ϵ 4 has been quantified in a recent multi-ancestry meta-analysis, with an estimated variability in effect size (I^2) across ancestries of up to 85%. Fifty percent or more of that risk heterogeneity was found to be attributable to genetic ancestry differences, suggesting that although *APOE* is still an important locus for AD risk across populations, the strength of the effect of *APOE* on AD risk varies with genetic ancestral background.⁶ We believe that examination of local ancestry at loci with such global differences may help discern whether locus-specific inheritance patterns modulate disease risk.

Similarly, *C9ORF72* is one of the most common risk factors for ALS. However, the frequency of the *C9ORF72* expansion is lower in Chinese populations (0.3%) as compared to European populations (7%).²⁹ Recent research suggests that commonly used genetic tests to diagnose ALS may be less accurate in non-European ancestry patients because they are less likely to carry the *C9ORF72* structural variant.⁶⁹

Some SNPs with large effect sizes do not exist or are extremely rare in certain ancestry groups. Variants in *ABCA7*, for example, increase AD risk more in individuals of African ancestry than in those of European ancestry.⁷⁰ In fact, *ABCA7* has a comparable effect size to *APOE* in individuals of African ancestry.¹⁰ Genetic variants in *LRRK2*, *GBA*,

and *SNCA*, which have been associated with increased risk of PD in European ancestry populations, appear to have a negligible effect in individuals from India.⁷¹⁻⁷⁴ Without studying diverse populations, researchers would miss the population-specific effects of these loci and potential therapeutic targets that modify their effects.

4.2 | Looking forward

Despite the inequalities highlighted above, progress is being made. Researchers in AD are taking a strong multi-modal approach to increasing diversity. The Multi-Partner Consortium to Expand Dementia Research in Latin America (ReDLat) is leveraging “on the ground” connections with research communities in Latin America and the Caribbean to grow a diverse database of dementia resources.^{75,76} The Alzheimer's Disease Neuroimaging (ADNI) study is growing more inclusive cell lines and generating partner data for multiple ancestrally diverse samples.⁷⁷ The NIH's Center for Alzheimer's and Related Dementias (CARD) is filling diversity gaps by creating training materials, generating data to complement existing efforts, and providing open science support for researchers in diverse communities.

Multiple efforts are also underway in the PD space. The Genetic Architecture of Parkinson disease in India (GAP-India) plans to develop a large clinical/genomic biobank in India.⁷¹ The Latin American Research Consortium on the Genetics of PD (LARGE-PD) aims to address inclusivity and genomic differences within and across Latino populations. Finally, the Global Parkinson's Genetics Program (GP2) aims to genotype >150,000 individuals from around the world. GP2-funded projects include the Black and African Americans Connections to Parkinson's Disease Study (BLAACPD), which seeks to assess the genetic architecture of Black and African American individuals with PD, as well as healthy subjects, from across the United States.⁷⁸ GP2 is motivated to increase diversity not only just among samples recruited into studies, but also in the investigators making use of the data, providing training and resources to ensure that all researchers are on an open and equal field of play.⁷⁹ A list of ongoing efforts for increasing diversity in NDD genetic research, including atypical dementias, can be found in Table S3.

These efforts are paying off. More than 60% of the NDD-associated loci discovered in non-European or multi-ancestry populations were identified in the period between 2020 and 2022 (Figure 2). Over the past 10 years, there has been a steady increase in the proportion of non-European samples included in genetic studies (Figure 1B). With the increase in diverse samples in recent years, there has also been a growing interest in the use of multi-ancestry analyses to discover, fine-map, and assess heterogeneity at disease risk loci, particularly in AD, PD, and ALS (Figure 1A).

In fact, we are already seeing the benefits of increased diversity on genetic discovery in NDD research. A 2023 GWAS using African and African American samples collected by GP2 and 23andMe, and co-lead by researchers in Nigeria and NIH, found a novel *GBA1* locus that is rare in other populations.⁸⁰ We anticipate that in the future, leveraging multiple ancestries will continue to improve fine-mapping resolution

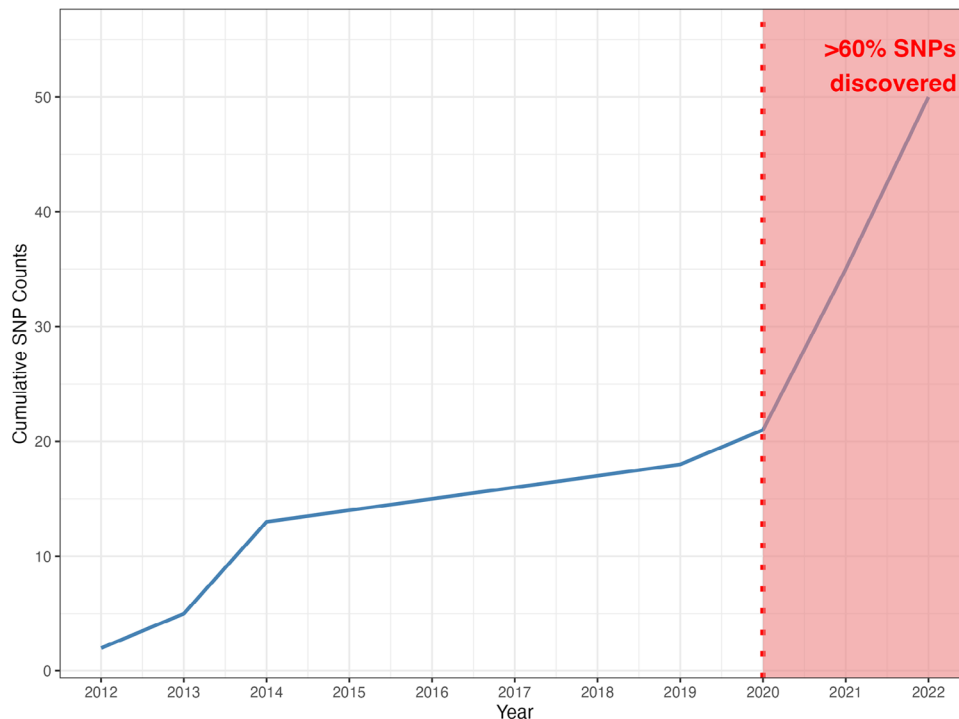


FIGURE 2 Cumulative count of discovered SNPs from 2012 through 2022. Notably, more than 60% of the SNPs were identified in the period between 2020 through 2022. SNPs, single nucleotide polymorphisms.

to prioritize causal variants,⁵ increase access to and reduce bias in precision medicine practices such as polygenic risk prediction,¹ and drive many new discoveries in the genetics of NDDs.

5 | CONCLUSION

Our systematic review highlights a striking disparity in the representation of diverse genetic ancestry populations in NDD research. We also show the variability among the novel loci, noting that no locus found in a single non-European population has been replicated in another population, emphasizing the urgent need for greater inclusivity to advance our understanding of these complex conditions and develop more equitable precision medicine approaches. Efforts to bridge this gap and promote diversity in genetic studies are vital for achieving meaningful progress in the diagnosis, treatment, and prevention of NDDs across global populations.

ACKNOWLEDGMENTS

We would like to thank the National Library of Medicine, including Nancy Terry and Alicia Livinski, for their assistance with the review and Covidence. This research was supported in part by the Intramural Research Program of the National Institutes of Health (NIH), National Institute on Aging (NIA), NIH, Department of Health and Human Services; project number ZO1 AG000534, as well as the National Institute of Neurological Disorders and Stroke (NINDS) and the National Institute on Drug Abuse (NIDA). J.S.Y. is sup-

ported by NIH-NIA R01AG062588, R01AG057234, P30AG062422, P01AG019724, U19AG079774; NIH-NINDS U54NS123985; NIH-NIDA 75N95022C00031; the Rainwater Charitable Foundation; the Alzheimer's Association; the Global Brain Health Institute; and the Mary Oakley Foundation.

CONFLICT OF INTEREST STATEMENT

C.J., K.S.L., L.J., H.L.L., and M.A.N.'s participation in this project was part of a competitive contract awarded to DataTecnica LLC by the National Institutes of Health to support open science research. M.A.N. also currently serves on the scientific advisory board for Character Bio Inc. and is a scientific founder at Neuron23 Inc. J.S.Y. serves on the scientific advisory board for the Epstein Family Alzheimer's Research Collaboration. J.L., L.H., D.P., J.K., S.B., N.T., I.M., C.B., and A.B.S. have no conflicts of interest. Author disclosures are available in the [Supporting Information](#).

ORCID

Caroline Jonson  <https://orcid.org/0000-0001-7049-6281>

REFERENCES

- Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet.* 2019;51(4):584-591.
- Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet.* 2022;54(4):1-25.
- Schumacher-Schuh AF, Bieger A, Okunoye O, et al. Diversity in Parkinson's disease genetics research: current landscape and future

- directions. *bioRxiv*. Preprint posted online December 8, 2021. doi:[10.1101/2021.12.07.21266995](https://doi.org/10.1101/2021.12.07.21266995)
4. Innovation VH. Covidence systematic review software. Published online 2019.
 5. Mägi R, Horikoshi M, Sofer T, et al. Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Hum Mol Genet*. 2017;26(18):3639-3650.
 6. Lake J, Warly Solsberg C, Kim JJ, et al. Multi-ancestry meta-analysis and fine-mapping in Alzheimer's disease. *Mol Psychiatry*. 2023;28(7):1-12. doi:[10.1038/s41380-023-02089-w](https://doi.org/10.1038/s41380-023-02089-w)
 7. Kim JJ, Vitale D, Véliz Otani D, et al. Multi-ancestry genome-wide meta-analysis in Parkinson's disease. *bioRxiv*. Preprint posted online August 6, 2022. doi:[10.1101/2022.08.04.22278432](https://doi.org/10.1101/2022.08.04.22278432)
 8. 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(9):1276-1282.
 9. Beach TG, Monsell SE, Phillips LE, Kukull W. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. *J Neuropathol Exp Neurol*. 2012;71(4):266-273.
 10. Reitz C, Jun G, Naj A, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ϵ 4, and the risk of late-onset Alzheimer disease in African Americans. *JAMA*. 2013;309(14):1483-1492.
 11. Mez J, Chung J, Jun G, et al. Two novel loci, COBL and SLC10A2, for Alzheimer's disease in African Americans. *Alzheimers Dement*. 2017;13(2):119-129.
 12. Sherva R, Zhang R, Sahelijo N, et al. African ancestry GWAS of dementia in a large military cohort identifies significant risk loci. *bioRxiv*. Preprint posted online May 26, 2022. doi:[10.1101/2022.05.25.22275553](https://doi.org/10.1101/2022.05.25.22275553)
 13. Tosto G, Fu H, Vardarajan BN, et al. F-box/LRR-repeat protein 7 is genetically associated with Alzheimer's disease. *Ann Clin Transl Neurol*. 2015;2(8):810-820.
 14. Hirano A, Ohara T, Takahashi A, et al. A genome-wide association study of late-onset Alzheimer's disease in a Japanese population. *Psychiatr Genet*. 2015;25(4):139-146.
 15. 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(3):924-937.
 16. 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(1):151.
 17. Kang S, Gim J, Lee J, et al. Potential novel genes for late-onset Alzheimer's disease in east-asian descent identified by apoe-stratified genome-wide association study. *J Alzheimers Dis*. 2021;82(4):1451-1460.
 18. 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(4):e58618.
 19. Lake J, Warly Solsberg C, Kim JJ, et al. Multi-ancestry meta-analysis and fine-mapping in Alzheimer's Disease. *bioRxiv*. Preprint posted online August 6, 2022. doi:[10.1101/2022.08.04.22278442](https://doi.org/10.1101/2022.08.04.22278442)
 20. Nalls MA, Blauwendraat C, Vallerga CL, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol*. 2019;18(12):1091-1102.
 21. Satake W, Nakabayashi Y, Mizuta I, et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat Genet*. 2009;41(12):1303-1307.
 22. Foo JN, Tan LC, Irwan ID, et al. Genome-wide association study of Parkinson's disease in East Asians. *Hum Mol Genet*. 2017;26(1):226-232.
 23. Li C, Ou R, Chen Y, et al. Genetic Modifiers of Age at Onset for Parkinson's Disease in Asians: A Genome-Wide Association Study. *Mov Disord*. 2021;36(9):2077-2084.
 24. Li C, Hou Y, Ou R, et al. Genetic determinants of survival in Parkinson's disease in the Asian population. *Mov Disord*. 2022;37(8) doi:[10.1002/mds.29069](https://doi.org/10.1002/mds.29069)
 25. Loesch DP, Horimoto ARVR, Heilbron K, et al. Characterizing the genetic architecture of Parkinson's disease in Latinos. *Ann Neurol*. 2021;90(3):353-365.
 26. Lill CM, Roehr JT, McQueen MB, et al. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLoS Genet*. 2012;8(3):e1002548.
 27. Foo JN, Chew EGY, Chung SJ, et al. Identification of Risk Loci for Parkinson Disease in Asians and Comparison of Risk Between Asians and Europeans: A Genome-Wide Association Study. *JAMA Neurol*. 2020;77(6):746-754.
 28. Nicolas A, Kenna KP, Renton AE, et al. Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron*. 2018;97(6):1268-1283.e6.
 29. Benyamin B, He J, Zhao Q, et al. Cross-ethnic meta-analysis identifies association of the GPX3-TNIP1 locus with amyotrophic lateral sclerosis. *Nat Commun*. 2017;8(1):611.
 30. Deng M, Wei L, Zuo X, et al. Genome-wide association analyses in Han Chinese identify two new susceptibility loci for amyotrophic lateral sclerosis. *Nat Genet*. 2013;45(6):697-700.
 31. Wei L, Tian Y, Chen Y, et al. Identification of TYW3/CRYZ and FGD4 as susceptibility genes for amyotrophic lateral sclerosis. *Neuro Genet*. 2019;5(6):e375.
 32. Xie T, Deng L, Mei P, et al. Genome-wide association study combining pathway analysis for typical sporadic amyotrophic lateral sclerosis in Chinese Han populations. *Neurobiol Aging*. 2014;35(7):1778.e9-e1778.e23.
 33. Iacoangeli A, Lin T, Al Khleifat A, et al. Genome-wide meta-analysis finds the ACSL5-ZDHHC6 locus is associated with ALS and links weight loss to the disease genetics. *Cell Rep*. 2020;33(4):108323.
 34. Nakamura R, Misawa K, Tohna G, et al. A multi-ethnic meta-analysis identifies novel genes, including ACSL5, associated with amyotrophic lateral sclerosis. *Communications Biology*. 2020;3(1):1-9.
 35. van Rheenen W, van der Spek RAA, Bakker MK, et al. Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. *Nat Genet*. 2021;53(12):1636-1648.
 36. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science*. 2019;365(6460):eaav7188. doi:[10.1126/science.aav7188](https://doi.org/10.1126/science.aav7188)
 37. Isobe N, Madireddy L, Khankhanian P, et al. An ImmunoChip study of multiple sclerosis risk in African Americans. *Brain*. 2015;138(Pt 6):1518-1530.
 38. Ordoñez G, Romero S, Orozco L, et al. Genome wide admixture study in Mexican Mestizos with multiple sclerosis. *Clin Neurol Neurosurg*. 2015;130:55-60.
 39. Sirkis DW, Geier EG, Bonham LW, Karch CM, Yokoyama JS. Recent advances in the genetics of frontotemporal dementia. *Curr Genet Med Rep*. 2019;7(1):41-52.
 40. Ferrari R, Hernandez DG, Nalls MA, et al. Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet Neurol*. 2014;13(7):686-699.
 41. Reus LM, Jansen IE, Mol MO, et al. Genome-wide association study of frontotemporal dementia identifies a C9ORF72 haplotype with a median of 12-G4C2 repeats that predisposes to pathological repeat expansions. *Transl Psychiatry*. 2021;11(1):451.
 42. Pottier C, Zhou X, Perkerson RB 3rd, et al. Potential genetic modifiers of disease risk and age at onset in patients with frontotemporal lobar degeneration and GRN mutations: a genome-wide association study. *Lancet Neurol*. 2018;17(6):548-558.

43. Che XQ, Zhao QH, Huang Y, et al. Genetic Features of MAPT, GRN, C9orf72 and CHCHD10 Gene Mutations in Chinese Patients with Frontotemporal Dementia. *Curr Alzheimer Res.* 2017;14(10):1102-1108.
44. Renton AE, Pliner HA, Provenzano C, et al. A genome-wide association study of myasthenia gravis. *JAMA Neurol.* 2015;72(4):396-404.
45. Seldin MF, Alkhairy OK, Lee AT, et al. Genome-wide association study of late-onset myasthenia gravis: confirmation of TNFRSF11A and identification of ZBTB10 and three distinct HLA associations. *Mol Med.* 2016;21(1):769-781.
46. Chia R, Saez-Atienzar S, Murphy N, et al. Identification of genetic risk loci and prioritization of genes and pathways for myasthenia gravis: a genome-wide association study. *Proc Natl Acad Sci U S A.* 2022;119(5). doi:10.1073/pnas.2108672119
47. Sakaue S, Kanai M, Tanigawa Y, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet.* 2021;53(10):1415-1424.
48. Na SJ, Lee JH, Kim SW, et al. Whole-genome analysis in Korean patients with autoimmune myasthenia gravis. *Yonsei Med J.* 2014;55(3):660-668.
49. Murai H, Yamashita N, Watanabe M, et al. Characteristics of myasthenia gravis according to onset-age: Japanese nationwide survey. *J Neurol Sci.* 2011;305(1-2):97-102.
50. Rongve A, Witoelar A, Ruiz A, et al. GBA and APOE ε4 associate with sporadic dementia with Lewy bodies in European genome wide association study. *Sci Rep.* 2019;9(1):7013.
51. Chia R, Sabir MS, Bandres-Ciga S, et al. Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. *Nat Genet.* 2021;53(3):294-303.
52. Guerreiro R, Ross OA, Kun-Rodriguez C, et al. Investigating the genetic architecture of dementia with Lewy bodies: a two-stage genome-wide association study. *Lancet Neurol.* 2018;17(1):64-74.
53. Schrijvers EMC, Schürmann B, Koudstaal PJ, et al. Genome-wide association study of vascular dementia. *Stroke.* 2012;43(2):315-319.
54. Moreno-Grau S, de Rojas I, Hernández I, et al. Genome-wide association analysis of dementia and its clinical endophenotypes reveal novel loci associated with Alzheimer's disease and three causality networks: The GR@ACE project. *Alzheimers Dement.* 2019;15(10):1333-1347.
55. Fongang B, Weinstein G, Guðjónsson A, et al. A meta-analysis of genome-wide association studies identifies new genetic loci associated with all-cause and vascular dementia. *Alzheimers Dement.* 2021;17(Suppl 3):e056081.
56. Singh V, Dharmoon MS, Alladi S. Stroke risk and vascular dementia in South Asians. *Curr Atheroscler Rep.* 2018;20(9):43.
57. Sahadevan S, Saw SM, Gao W, et al. Ethnic differences in Singapore's dementia prevalence: the stroke, Parkinson's disease, epilepsy, and dementia in Singapore study. *J Am Geriatr Soc.* 2008;56(11):2061-2068.
58. Simpkins AN. Impact of race-ethnic and economic disparities on rates of vascular dementia in the national inpatient sample database from 2006-2014. *J Stroke Cerebrovasc Dis.* 2020;29(5):104731.
59. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747-753.
60. Campbell MC, Tishkoff SA. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annu Rev Genomics Hum Genet.* 2008;9:403-433.
61. Kamiza AB, Toure SM, Vujkovic M, et al. Transferability of genetic risk scores in African populations. *Nat Med.* 2022;28(6):1163-1166.
62. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A.* 2009;106(23):9362-9367.
63. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res.* 2007;17(10):1520-1528.
64. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. *Nat Rev Genet.* 2014;15(5):335-346.
65. Ramírez-Castañeda V. Disadvantages in preparing and publishing scientific papers caused by the dominance of the English language in science: the case of Colombian researchers in biological sciences. *PLoS ONE.* 2020;15(9):e0238372.
66. Naslavsky MS, Suemoto CK, Brito LA, et al. Global and local ancestry modulate APOE association with Alzheimer's neuropathology and cognitive outcomes in an admixed sample. *bioRxiv.* Preprint posted online February 4, 2022. doi:10.1101/2022.02.02.22270331
67. Jun GR, Chung J, Mez J, et al. Transethnic genome-wide scan identifies novel Alzheimer's disease loci. *Alzheimers Dement.* 2017;13(7):727-738.
68. Blue EE, Horimoto ARVR, Mukherjee S, Wijsman EM, Thornton TA. Local ancestry at APOE modifies Alzheimer's disease risk in Caribbean Hispanics. *Alzheimers Dement.* 2019;15(12):1524-1532.
69. Mesaros M. *Investigating the Genetic Profile of Amyotrophic Lateral Sclerosis in Patients of Diverse Race, Ethnicity, and Ancestry (REA).* The Ohio State University; 2021. https://rave.ohiolink.edu/etdc/view?acc_num=osu1618934315464032
70. Stepler KE, Gillyard TR, Reed CB, Avery TM, Davis JS, Robinson RAS. ABCA7, a genetic risk factor associated with Alzheimer's disease risk in African Americans. *J Alzheimers Dis.* 2022;86(1):5-19. doi:10.3233/JAD-215306
71. Rajan R, Divya KP, Kandadai RM, et al. Genetic architecture of Parkinson's disease in the Indian population: harnessing genetic diversity to address critical gaps in Parkinson's disease research. *Front Neurol.* 2020;11:524.
72. Punia S, Behari M, Govindappa ST, et al. Absence/rarity of commonly reported LRRK2 mutations in Indian Parkinson's disease patients. *Neurosci Lett.* 2006;409(2):83-88.
73. Sun X, Dou K, Xue L, Xie Y, Yang Y, Xie A. Comprehensive analysis of clinical and biological features in Parkinson's disease associated with the LRRK2 G2019S mutation: Data from the PPMI study. *Clin Transl Sci.* 2024;17(1):e13720.
74. Vishwanathan Padmaja M, Jayaraman M, Srinivasan AV, Srikumari Srisailapathy CR, Ramesh A. The SNCA (A53T, A30P, E46K) and LRRK2 (G2019S) mutations are rare cause of Parkinson's disease in South Indian patients. *Parkinsonism Relat Disord.* 2012;18(6):801-802.
75. Ibanez A, Yokoyama JS, Possin KL, et al. The multi-partner consortium to expand dementia research in Latin America (ReDLat): driving multicentric research and implementation science. *Front Neurol.* 2021;12:631722.
76. Ibanez A, Parra MA, Butler C, Latin America and the Caribbean consortium on dementia (LAC-CD): from networking to research to implementation science. *J Alzheimers Dis.* 2021;82(s1):S379-S394.
77. Crane PK, Choi S, Gibbons LE, et al. Cognitive assessments in ADNI: lessons learned from the ADNI psychometrics project. *Alzheimers Dement.* 2021;17(S5). doi:10.1002/alz.056474
78. Bandres-Ciga S. Black and African American connections to Parkinson's disease study: addressing missing diversity in Parkinson's disease genetics. *Mov Disord.* 2022;37(7):1559-1561.
79. Program TGPG, The Global Parkinson's Genetics Program. GP2: the global Parkinson's genetics program. *Movement Disord.* 2021;36(4):842-851. doi:10.1002/mds.28494
80. Rizig M, Bandres-Ciga S, Makarious MB, et al. Identification of genetic risk loci and causal insights associated with Parkinson's disease in

African and African admixed populations: a genome-wide association study. *Lancet Neurol.* 2023;22(11):1015-1025.

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: Jonson C, Levine KS, Lake J, et al. Assessing the lack of diversity in genetics research across neurodegenerative diseases: A systematic review of the GWAS Catalog and literature. *Alzheimer's Dement.* 2024;20:5740–5756. <https://doi.org/10.1002/alz.13873>