Parkinson's disease genetics across diverse ancestries: an observational genetic study of causal and risk variants with translational implications

Lara M Lange, Zih-Hua Fang, Mary B Makarious, Nicole Kuznetsov, Kajsa Atterling Brolin, Shannon Ballard, Soraya Bardien, Maria Leila Doquenia, Peter Heutink, Henry Houlden, Hirotaka Iwaki, Simona Jasaityte, Lietsel Jones, Johanna Junker, Rauan Kaiyrzhanov, Mathew J Koretsky, Kishore R Kumar, the Latin American Research Consortium on the Genetics of Parkinson's Disease (LARGE-PD), Hampton L Leonard, Kristin S Levine, Shen-Yang Lim, Niccoló E Mencacci, Wael M Y Mohamed, Mike A Nalls, Alastair J Noyce, Rajeev Ojha, Njideka U Okubadejo, Shoaib ur Rehman, Laurel Screven, Chingiz Shashkin, Sophia Sopromadze, Eleanor J Stafford, Ai Huey Tan, Manuela Tan, Zaruhi Tavadyan, Joanne Trinh, Bayasgalan Tserensodnom, Enza Maria Valente, Dan Vitale, Nazira Zharkinbekova, Katja Lohmann, Sara Bandres-Ciga, Cornelis Blauwendraat, Andrew B Singleton, Huw R Morris*, Christine Klein*, on behalf of the Global Parkinson's Genetics Program (GP2)†*

Summary

Background The genetic architecture of Parkinson's disease varies considerably across ancestries, yet most previous genetic studies have focused on individuals of European ancestry. We aimed to characterise the distribution of established Parkinson's disease causal variants, as well as risk-associated variants with clinical implications (ie, variants in genes involved in pathways targeted by ongoing clinical trials), across ancestrally diverse populations.

Methods We conducted a multi-ancestry, observational, cross-sectional genetic study using retrospective data from the Global Parkinson's Genetics Program (GP2) release 11 (released in December, 2025). The study investigated causal and risk variants, including copy number variants, in established Parkinson's disease and parkinsonism-associated genes, following the recommendations of the Movement Disorder Society (MDS) Task Force on the Nomenclature of Genetic Movement Disorders, including *GBA1, LRRK2, SNCA, VPS35, RAB32, PINK1, PRKN, PARK7, ATP13A2, DCTN1, DNAJC6, FBXO7, JAM2, RAB39B, SLC20A2, SYNJ1, VPS13C*, and *WDR45*. Individuals with Parkinson's disease were diagnosed based on established clinical criteria, including the Parkinson's UK Brain Bank or MDS diagnostic criteria (or both), and healthy control participants were defined as individuals without evidence of neurodegenerative disease and unrelated to participants with Parkinson's disease. We analysed genome and exome sequencing and array genotyping data of 99 783 individuals, including 58 559 individuals with Parkinson's disease and 41 224 controls, from 11 genetically inferred ancestries (African, African admixed, Ashkenazi Jewish, Latino and Indigenous people of the Americas, central Asian, complex admixture, east Asian, European, Finnish, Middle Eastern, and south Asian), defined using reference population-based ancestry inference methods. We calculated allele frequencies for all investigated variants in individuals with Parkinson's disease and controls, both overall and stratified by ancestry.

Findings Approximately 29% of individuals (29 001 of 99 783; 15 443 [26·4%] of 58 559 individuals with Parkinson's disease and 13 558 [32·9%] of 41 224 controls) were from under-represented populations (ie, non-European and non-Ashkenazi Jewish). Our findings indicated both shared genetic contributors across ancestries as well as ancestry-specific differences in variant frequencies and the spectrum of variants within Parkinson's disease-associated genes. Overall, 1217 (2·1%) of 58 559 individuals with Parkinson's disease carried a causal variant, with substantial variations across ancestries ranging from ten (0·4%) of 2844 African individuals to 251 (10·7%) of 2343 individuals of Ashkenazi Jewish ancestry. Risk variants in *GBA1* and *LRRK2* were identified in 6893 (11·8%) of 58 559 individuals with Parkinson's disease and 3578 (8·7%) of 41 224 controls. *GBA1* risk variants were most frequent overall and identified across all ancestries, but variant frequency and spectra differed substantially between ancestries, from 195 (4·1%) of 4773 in the east Asian ancestry group to 1505 (52·9%) of 2844 in the African ancestry group. Similarly, *LRRK2* causal and risk variants showed ancestry-specific enrichment, with the highest frequencies of causal variants in the Ashkenazi Jewish (250 [10·7%] of 2343) and Middle Eastern (59 [4·4%] of 1347) ancestry groups, whereas risk variants were predominantly identified in the east Asian ancestry group (601 [12·6%] of 4773). Carriers of biallelic causal variants in *PRKN*, commonly including deletions and duplications, were also identified across all ancestries except Ashkenazi Jewish; the highest frequency was in the Middle Eastern ancestry group (17 [1·3%] of 1347), and frequencies in all other ancestries were less than 1%.

Interpretation This large-scale, multi-ancestry genetic study offers crucial insights into the population-specific genetic architecture of Parkinson's disease. Whereas clinical trials targeting *GBA1* and *LRRK2* variant carriers are primarily performed in Europe and the USA, increased ancestral diversity in Parkinson's disease research will be crucial to improve diagnostic accuracy, enhance our understanding of disease mechanisms across populations, and ensure equitable application of and access to emerging genetically informed therapies.

Lancet Neurol 2026; 25: 741–54

See **Comment** page 705

*Joint last authors

†Collaborators listed in appendix 1 (pp 37–41)

Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA (L M Lange MD); **Institute of Neurogenetics, University of Luebeck, Lübeck, Germany** (L M Lange, J Trinh PhD, Prof K Lohmann PhD, Prof C Klein MD); **German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany** (Z-H Fang PhD); **DataTecnica, Washington, DC, USA** (M B Makarious PhD, N Kuznetsov MPS, S Ballard PhD, H Iwaki PhD, L Jones MSc, M J Koretsky BSc, H L Leonard MSc, K S Levine MSc, M A Nalls PhD, D Vitale MSc); **Center for Alzheimer's and Related Dementias (CARD), National Institute on Aging and National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA** (M B Makarious, N Kuznetsov, S Ballard, H Iwaki, L Jones, M J Koretsky, H L Leonard, K S Levine, M A Nalls, D Vitale, S Bandres-Ciga PhD); **Translational Neurogenetics Unit, Department of Experimental Medical Science, Lund University, Lund, Sweden** (K Atterling Brolin PhD); **Centre for Preventive Neurology, Wolfson Institute of Population Health, Queen Mary University of London, London, UK** (K Atterling Brolin, Prof A J Noyce PhD, M Tan PhD); **Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch**

University (Prof S Bardien PhD); South African Medical Research Council, Stellenbosch University Genomics of Brain Disorders Research Unit, Cape Town, South Africa (Prof S Bardien); Department of Neuroscience and Brain Health, Metropolitan Medical Center, Manila, Philippines (M L Doquenia MD); Department of Clinical Neurosciences, Mary Mediatrix Medical Center, Lipa, Batangas, Philippines (M L Doquenia); Department of Anatomy, University of the East Ramon Magsaysay Memorial Medical Center, Quezon City, Philippines (M L Doquenia); The Global Parkinson's Genetics Program (GP2), Chevy Chase, MD, USA (Prof P Heutink PhD, L Screven PhD, C Blauwendraat PhD, A B Singleton PhD); Department of Neuromuscular Diseases (Prof H Houlden MD, R Kaiyrzhanov MD) and Department of Clinical and Movement Neurosciences (S Jasaityte MSc, E J Stafford MSc, Prof H R Morris PhD), UCL Queen Square Institute of Neurology, University College London, London, UK; University Hospital Schleswig Holstein, Campus Lübeck, Lübeck, Germany (J Junker MD); Center for Movement Disorders and Neuromodulation, University Hospital Düsseldorf, Düsseldorf, Germany (J Junker); Department of Neurology, South Kazakhstan Medical Academy, Shymkent, Kazakhstan (R Kaiyrzhanov, N Zharkinbekova MD); Neurology Department and Molecular Medicine Laboratory, Concord Repatriation General Hospital and University of Sydney, Concord, NSW, Australia (K R Kumar PhD); Translational Neurogenomics Group, Genomics and Inherited Disease Program, The Garvan Institute of Medical Research, Darlinghurst, NSW, Australia (K R Kumar); Faculty of Medicine and Health, University of New South Wales, Sydney, NSW, Australia (K R Kumar); Division of Neurology, Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia (Prof S-Y Lim MD, Prof A H Tan MD); Department

Funding Aligning Science Across Parkinson's (ASAP) through the Global Parkinson's Genetics Program (GP2).

Copyright © 2026 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Introduction

Genetic discoveries have transformed our understanding of Parkinson's disease. Genetic factors substantially contribute to Parkinson's disease risk and progression.[1] The genetic factors range from rare variants with large effect sizes to common variants with moderate effect sizes. More than 130 independent risk loci have been identified by genome-wide association studies (GWAS),[2] and over 15 genes have been linked to monogenic Parkinson's disease and parkinsonism.[1,3]

A crucial gap in Parkinson's disease genetics research is the lack of ancestral diversity.[4] Although genetic forms of Parkinson's disease are found in all regions across the world, approximately 75% of all genetic Parkinson's disease studies to date focused on European-ancestry individuals.[5,6] Research in diverse populations is critical to uncovering unique genetic contributions to disease and improve diagnosis, risk assessment, and genetic counselling. This need for greater ancestral diversity in Parkinson's disease genetics research is particularly relevant for clinical trials driven by genetic discoveries, such as targeting glucocerebrosidase or LRRK2 kinase (encoded by the Parkinson's disease-linked genes *GBA1* and *LRRK2*, respectively).[7] The absence of diversity in genetic studies could limit the broader applicability of emerging opportunities, because it is unclear whether current genetic targets are relevant across ancestries or whether population-specific variants exist.

The Global Parkinson's Genetics Program (GP2)[8] is a large-scale initiative actively addressing this gap by

Research in context

Evidence before this study

We searched PubMed from Jan 1, 2020, to Dec 30, 2025, for Articles reporting genetic screening studies in Parkinson's disease, available in English, using search terms "Gene*", "Parkinson's disease", and "screen". Most genetic data have been generated in Parkinson's disease cohorts of European ancestry. Two comprehensive, large-scale genetic screening studies, the PD GENEration North America and the Rostock International Parkinson's disease study, each comprised more than 85% White participants. In a global survey of monogenic Parkinson's disease, published in 2023, 91% of individuals with identified genetic causal or risk variants in Parkinson's disease-associated genes were White, underscoring the major gap in ancestral representation in Parkinson's disease genetics research. Although ancestry-specific variants and founder effects have been reported for some Parkinson's disease-associated genes such as *GBA1*, *LRRK2*, and *PINK1*, comprehensive and ancestry-diverse studies are missing.

Added value of this study

This study, conducted within the Global Parkinson's Genetics Program (GP2), represents the largest multi-ancestry genetic investigation of Parkinson's disease to date, including around 100 000 individuals from around the world, with approximately 29 000 participants of non-European and non-Ashkenazi Jewish ancestry. This study represents the largest systematically curated genetic resource for multiple populations that have been historically under-represented in genetic studies, for whom sample collection and data curation are logistically challenging and genetic reference data remain sparse. Even where absolute sample sizes are smaller than in European cohorts, these data provide essential information for ancestry-specific variant interpretation and disease biology. By exploring causal variants in established Parkinson's disease-linked genes and Parkinson's disease risk-associated variants across 11 diverse genetically defined ancestry groups, this study provides a comprehensive view of the genetic architecture of Parkinson's disease. This work highlights both shared and ancestry-specific contributions. Pathogenic and risk variants in *GBA1* are globally relevant contributors to Parkinson's disease risk, with variant carriers identified across all investigated populations, reinforcing the relevance of GBA1-targeted therapeutic approaches across ancestries. Similarly, *LRRK2* variant carriers, currently being recruited into LRRK2-targeted clinical trials, and individuals with *PRKN* causal variants, for whom PRKN-targeted approaches are being explored in experimental studies, were identified across multiple ancestral groups, supporting the broad applicability of genetically targeted approaches. Substantial ancestry-specific differences regarding variant frequencies and spectra underscore the limited transferability of current Parkinson's disease gene and variant findings, which are largely derived from European-ancestry datasets.

Implications of all the available evidence

This study establishes a foundation for future ancestry-specific analyses, functional studies, and clinically actionable genetic research. As genetically informed therapeutic approaches are integrated into Parkinson's disease research and clinical trials, it is crucial to ensure that emerging therapies are applicable and accessible to individuals of all ancestral backgrounds. Current genetically stratified trials predominantly enrol participants of European or Ashkenazi Jewish ancestry and trial sites are predominantly located across Europe and North America, raising concerns about equity, generalisability, and global clinical impact. This study underscores the importance of inclusive genetic efforts such as GP2 to close these existing gaps, improve diagnostic equity, inform trial design, and enable equitable and effective implementation of genetically informed therapeutics in Parkinson's disease worldwide.

investigating the genetic underpinnings of Parkinson's disease and parkinsonism across ancestry-diverse populations.[5,9] As part of these efforts, several GWAS were performed in populations historically underrepresented in genetic studies, including African,[10] Latin American,[11] and south Asian populations,[12] leading to the identification of both shared and ancestry-specific risk loci. In parallel, studies in diverse populations have identified marked ancestry enrichment of specific disease-related variants, such as *LRRK2* p.R1067Q in east Asia[13] and *GBA1* p.K198E in Colombia.[14]

These observations highlight the need for systematic, large-scale, multi-ancestry investigations. Accordingly, the aim of our study was to comprehensively and descriptively characterise the carrier frequencies and ancestry-specific presence and spectrum of causal and risk variants across diverse populations, and to evaluate their translational relevance for genetic screening and genetically informed therapies.

Methods

Study design and participants

The study workflow is shown in figure 1. We analysed short-read genome sequencing, clinical exome sequencing, and Illumina NeuroBooster Array genotyping data from GP2 release 11, comprising all data available and processed within GP2 up to Dec 1, 2025. A detailed data overview is provided in appendix 1 (p 21). The total sample comprised 99 783 individuals, including 58 559 individuals with Parkinson's disease and 41 224 healthy controls. This analysis includes participant data from 144 different cohorts at 129 sites. Our analyses included a subset of previously published individuals from PD GENEration (n=8902)[15] and ROPAD (n=3113),[16] now both part of the harmonised GP2 dataset.

All participants in GP2 underwent neurological assessment according to local study protocols. A diagnosis of Parkinson's disease was based on established clinical criteria, including one or both of the Parkinson's UK Brain Bank and the Movement Disorder Society (MDS) diagnostic criteria. Control participants were defined as individuals without evidence of neurodegenerative disease and were unrelated to participants with Parkinson's disease included in the study.

This study was conducted in accordance with the ethical standards of the institutional and national research committees and approved by the ethics committees or institutional review boards at all GP2 sites that provided samples and data for this study. Informed consent for study participation was obtained from all participants by each local site. All cohorts recruited to the GP2 initiative underwent a thorough review of the consent forms in the Operations and Compliance working group, ensuring that each contributing study abides by the ethics guidelines set out by their institutional review boards.

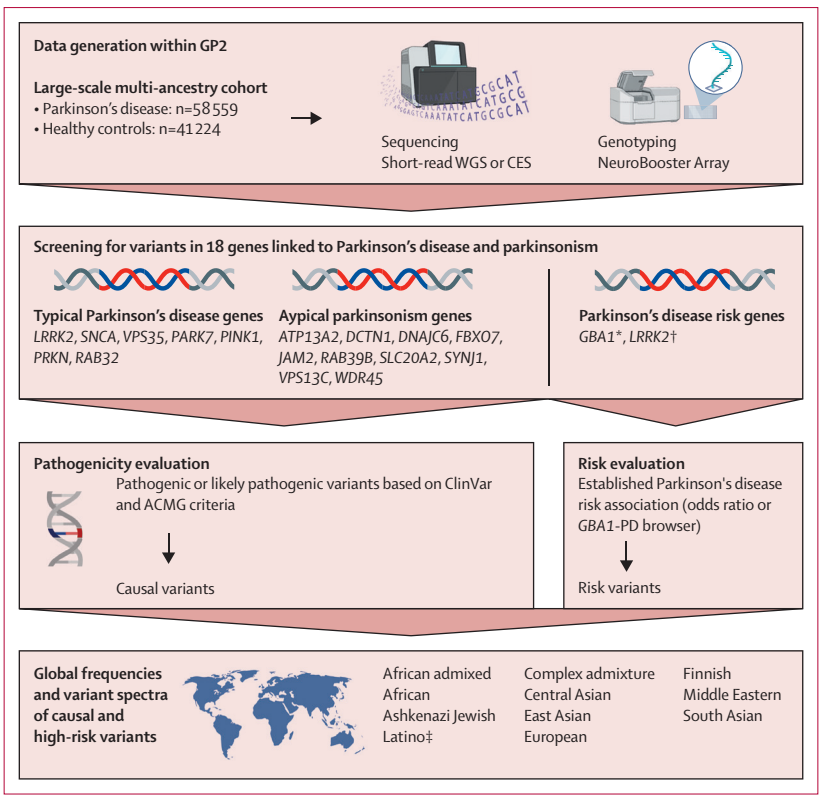

Figure 1: **Study workflow**
Individuals were screened for causative (pathogenic or likely pathogenic according to ClinVar criteria, ACMG criteria, or both) and selected risk-associated variants in 18 distinct Parkinson's disease-linked genes using short-read sequencing and genome-wide genotyping data. Figure created with BioRender.com. ACMG=American College of Medical Genetics and Genomics. CES=clinical exome sequencing. GP2=Global Parkinson's Genetics Program. WGS=whole-genome sequencing. *For the purpose of this study and in the context of Parkinson's disease, all variants in *GBA1* (including variants causal for Gaucher disease) are considered Parkinson's disease risk variants (regardless of their Gaucher disease severity). †*LRRK2* variants with established Parkinson's disease-risk associations included rs33949390 (chr12:40320043:G>C, p.R1628P) and rs34778348 (chr12:40363526:G>A, p.G2385R). ‡Latino people and Indigenous people of the Americas.

Procedures

Data processing and quality control were performed using GenoTools, as described elsewhere;[17] details are provided in appendix 1 (p 4). Genetic ancestry was inferred using GenoTools, which leverages high-quality variants shared with curated reference panels from the 1000 Genomes Project, Human Genome Diversity Project, and Ashkenazi reference panels. Reference variants have undergone quality control for minor allele frequency, missingness, Hardy–Weinberg equilibrium, and linkage disequilibrium. Based on this pipeline, samples were divided into the following ancestries: African admixed (n=550 individuals with Parkinson's disease and n=881 healthy controls), African (n=2844 and n=4489, respectively), Ashkenazi Jewish (n=2343 and n=913, respectively), Latino and Indigenous people of the Americas (n=2809 and n=1654, respectively), complex admixture (n=902 and n=444, respectively), central Asian (n=1284 and n=1480, respectively), east Asian (n=4773 and n=2809, respectively), European

of Neurology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA (N E Mencacci MD); **Department of Basic Medical Sciences, Kulliyyah of Medicine, International Islamic University Malaysia (IIUM), Kuantan, Malaysia** (Prof W M Y Mohamed PhD); **Clinical Pharmacology Department, Menoufia Medical School, Menoufia University, Shebin El-Kom, Menoufia, Egypt** (Prof W M Y Mohamed); **Department of Neurology, Tribhuvan University Teaching Hospital, Kathmandu, Nepal** (R Ojha MD); **Neurology Unit, Department of Medicine, College of Medicine, University of Lagos, Lagos State, Nigeria** (Prof N U Okubadejo); **Department of Biotechnology, University of Science and Technology Bannu, Bannu,**

Pakistan (S ur Rehman PhD); **International Research Institute of Postgraduate Education, Department of Neurosurgery and Neurology, Almaty, Kazakhstan** (C Shashkin MD); **Shashkin Clinic, Almaty, Kazakhstan** (C Shashkin); **Department of Neurology, Ivane Javakhishvili Tbilisi State University, Tbilisi, Georgia** (S Sopromadze MD); **East European University, Tbilisi, Georgia** (S Sopromadze); **Department of Neurology, Oslo University Hospital, Oslo, Norway** (M Tan); **National Institute of Health, Yerevan, Armenia** (Z Tavadyan MD); **Department of Neurology, School of Medicine, Mongolian National University of Medical Sciences, Ulaanbaatar, Mongolia** (B Tserensodnom MD); **Department of Molecular Medicine, University of Pavia, Pavia, Italy** (Prof E M Valente PhD); **Neurogenetics Research Center, IRCCS Mondino Foundation, Pavia, Italy** (Prof E M Valente); **Coalition for Aligning Science, Chevy Chase, MD, USA** (C Blauwendraat); **UCL Movement Disorders Centre, University College London, London, UK** (Prof H R Morris)

Correspondence to:
Dr Lara M Lange, Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, USA
lara@gp2.org

or

Prof Christine Klein, Institute of Neurogenetics, University of Lübeck, 23538 Lübeck, Germany
christine.klein@uni-luebeck.de

See Online for appendix 1

For more on **GP2** see https://gp2.org/

For **ClinVar** see https://www.ncbi.nlm.nih.gov/clinvar/

See Online for appendix 2

(n=40 773 and n=26 753, respectively), Finnish (n=127 and n=22, respectively), Middle Eastern (n=1347 and n=1339), and south Asian (n=807 and n=440, respectively). Because few publicly available reference samples are available for certain populations and highly admixed groups, GenoTools incorporates an approach to identify such individuals and assign them to a distinct category not represented in the reference panel, termed complex admixture. Because ancestry-specific interpretation is limited for these individuals, this group is not discussed in this report, but all corresponding cohort information and results for this ancestry group are provided in the tables and appendices.

We focused on variants in genes linked to Parkinson's disease and parkinsonism following the recommendations of the MDS Task Force on the Nomenclature of Genetic Movement Disorders,[3] including variants in: *LRRK2, SNCA,* and *VPS35* linked to autosomal-dominant Parkinson's disease; *PARK7, PINK1,* and *PRKN* linked to autosomal-recessive Parkinson's disease; and *ATP13A2, DCTN1, DNAJC6, FBXO7, JAM2, RAB39B, SLC20A2, SYNJ1, VPS13C,* and *WDR45* linked to atypical or complex parkinsonism. We included only PARK-designated genes—ie, all genes confidently linked to Parkinson's disease and parkinsonism based on the available evidence from the literature and following criteria defined by the MDS Task Force. Genes linked to combined phenotypes including Parkinson's disease or parkinsonism (eg, dystonia parkinsonism or frontotemporal dementia parkinsonism) were beyond the scope of this study. We added *RAB32* p.S71R, which was identified as causal for Parkinson's disease after the Task Force recommendations were published. We included pathogenic or likely pathogenic (according to ClinVar) variants in these genes and also investigated Parkinson's disease risk-associated variants in *GBA1* and *LRRK2*. Details on pathogenicity evaluation are provided in appendix 1 (pp 4–5, 22–27). A full list of all identified and investigated variants, including their pathogenicity evaluation and detectability by genotyping or sequencing, is also provided in appendix 1 (pp 22–27). Copy number variation analyses for *SNCA* and *PRKN* were performed for samples with available genotyping data, as described previously.[18] Carriers of two causal variants in recessive genes and an age at onset of 60 years or less were considered likely compound heterozygous without further validation, based on findings from previous studies.[19]

Statistical analysis

We calculated allele frequencies for all investigated variants in individuals with Parkinson's disease and healthy controls, both overall and stratified by ancestry. For selected variants meeting predefined criteria, we performed single-variant association analyses using logistic regression. Variants were included if they met the following criteria: available from imputed genotyping data; observed in at least two ancestry groups; and at least one carrier per group (cases and controls) across these ancestry groups. We used sex and the first five principal components as covariates. Age was not included as a covariate because it was missing for a considerable proportion of individuals; sensitivity analyses including age in the subset of samples with available data yielded similar results to analyses performed without age as a covariate. Analyses were conducted using PLINK version 2.

Role of the funding source

The funders of the study had no role in data collection, data analysis, data interpretation, or writing of the report.

Results

Sample characteristics are summarised in table 1. Approximately 29% of individuals (29 001 of 99 783; 15 443 [26·4%] of 58 559 individuals with Parkinson's disease and 13 558 [32·9%] of 41 224 controls) were from under-represented populations (ie, non-European and non-Ashkenazi Jewish). Across all ancestries, 1217 (2·1%) of 58 559 individuals with Parkinson's disease carried a causal variant and 6893 (11·8%) of 58 559 carried a risk-associated variant (table 2). By comparison, causal variants were observed in 108 (0·3%) of 41 224 and risk-associated variants in 3578 (8·7%) of 41 224 controls (table 2).

We identified 113 distinct causal single nucleotide variants (SNVs) in the 16 genes studied (including single heterozygous variants in recessive genes), alongside two *LRRK2* risk-associated variants, 95 *GBA1* risk-associated variants, and structural variations such as *SNCA* multiplications and *PRKN* exon deletions and duplications (appendix 1 pp 22–27). The allele frequencies of all identified variants as well as number of variant carriers, stratified by phenotype and ancestry, are provided in appendix 2. Only *PRKN* deletions and duplications and three *GBA1* SNVs (ie, p.E365K, p.T408M, and p.L483P) were detected across all ancestries, whereas the remaining variants were only identified in some ancestries. All variants identified in healthy controls, along with their age ranges, are provided in appendix 1 (pp 29–31).

Association analyses confirmed several well established ancestry-specific associations (appendix 1 pp 32–34). *GBA1* p.E365K and p.N409S were significantly associated with Parkinson's disease in Ashkenazi Jewish and European populations, whereas *GBA1* p.T408M reached significance only in the European population. *GBA1* rs3115534-G was significantly associated with Parkinson's disease in African and African admixed populations. For *LRRK2*, significant associations were observed for p.R1628P and p.G2385R in the east Asian population and for p.G2019S in European, Ashkenazi Jewish, and Middle Eastern populations.

The relative contribution of variants in different Parkinson's disease-associated genes in individuals with Parkinson's disease across ancestries is summarised in table 2, figure 2, and appendix 1 (pp 7–8), illustrating both

	Individuals with Parkinson's disease					Healthy controls			
	Samples, n	Male (%)	Age in years*, median (IQR); range	Age at onset in years†, median (IQR); range	Positive family history of Parkinson's disease‡ (%)	Samples, n	Male (%)	Age in years*, median (IQR); range	Positive family history of Parkinson's disease‡ (%)
African admixed	550	334 (60·7%)	67 (60–73); 29–95	60 (51–68); 13–90	75 (13·6%)	881	314 (35·6%)	65 (58–71); 18–91	3 (0·3%)
Missing	..	..	86	147	102	..	..	39	450
African	2844	1983 (69·7%)	65 (57–72); 22–96	58 (54–66); 12–92	326 (11·5%)	4489	2303 (51·3%)	63 (56–69); 18–99	11 (0·2%)
Missing	..	..	177	182	1006	..	..	71	1878
Ashkenazi Jewish	2343	1570 (67·0%)	71 (64–77); 32–94	64 (56–71); 12–93	578 (24·7%)	913	530 (58·1%)	67 (60–74); 23–95	31 (3·4%)
Missing	..	..	182	466	385	..	..	284	586
Latino and Indigenous people of the Americas	2809	1618 (57·6%)	64 (54–71); 18–97	54 (44–63); 10–92	485 (17·3%)	1654	582 (35·2%)	59 (53–65); 18–92	84 (5·1%)
Missing	..	..	192	1689	793	..	..	71	210
Complex admixture	902	527 (58·4%)	63 (54–70); 18–92	55 (45–64); 13–88	164 (18·2%)	444	184 (41·4%)	55 (32–66); 18–94	5 (1·1%)
Missing	..	..	151	313	245	..	..	53	321
Central Asian	1284	599 (46·7%)	62 (54–69); 18–94	53 (43–61); 14–89	108 (8·4%)	1480	612 (41·4%)	59 (52–65); 29–96	2 (0·1%)
Missing	..	..	464	667	635	..	..	689	1454
East Asian	4773	2530 (53·0%)	67 (58–73); 19–96	56 (47–65); 10–92	943 (19·8%)	2809	1890 (67·3%)	63 (55–70); 20–98	11 (0·4%)
Missing	..	..	671	1078	625	..	..	1166	1737
European	40 773	25 633 (62·9%)	68 (60–75); 19–100	60 (51–68); 10–98	8300 (20·4%)	26 753	13 148 (49·1%)	64 (57–71); 18–99	1013 (3·8%)
Missing	..	..	6052	8635	15 436	..	..	4507	19 914
Finnish	127	59 (46·5%)	67 (58–73); 38–86	57 (50–66); 35–76	20 (15·7%)	22	11 (50·0%)	72 (64–78); 49–85	1 (4·5%)
Missing	..	..	59	60	77	..	..	6	16
Middle Eastern	1347	797 (59·2%)	64 (56–70); 21–94	53 (44–62); 11–85	337 (25·0%)	1339	603 (45·0%)	63 (53–70); 22–96	0
Missing	..	..	339	485	485	..	..	601	1263
South Asian	807	528 (65·4%)	63 (53–71); 21–91	56 (45–64); 16–85	207 (25·7%)	440	304 (69·1%)	57 (50–63); 19–84	1 (0·2%)
Missing	..	..	68	93	131	..	..	92	228
Total	58 559	36 178 (61·8%)	67 (59–74); 18–100	60 (50–67); 10–98	11 543 (19·7%)	41 224	20 481 (49·7%)	63 (56–70); 18–99	1162 (2·8%)
Missing	..	..	8437	13 791	19 893	..	..	7579	28 057

*Age refers to the age at recruitment or age at sample collection; for controls, age <18 years or >100 years were considered outliers and excluded from these statistics; for individuals with Parkinson's disease, age <10 years or >100 years were considered outliers and excluded from these statistics. †Individuals with an age of onset <10 years or >100 years were considered outliers and excluded from these statistics. ‡Includes related individuals.

Table 1: **Cohort characteristics**

shared genetic contributors and ancestry-specific distributions of carriers across Parkinson's disease-associated genes. *GBA1* variant carriers were most frequent overall and detected across all ancestries, albeit with substantially different variant spectra (figure 3). Similarly, *LRRK2* variants were identified across multiple ancestries with ancestry-specific variant spectra (figure 4). Rarer *PRKN* SNVs and structural variations were also observed across most ancestries, whereas causal variants in other genes were only identified in certain populations. Variants in genes linked to atypical parkinsonism were only identified in a small subset of individuals (appendix 1 p 35).

Individuals with *GBA1* risk variant-associated Parkinson's disease showed a significantly earlier median age at onset across multiple ancestries compared with those with idiopathic Parkinson's disease (difference of 3–8 years; appendix 1 pp 9, 36). The difference in age at onset between *LRRK2* Parkinson's disease and idiopathic Parkinson's disease showed similar trends in some ancestries, with median age at onset 1–3 years earlier with *LRRK2* Parkinson's disease (appendix 1 pp 9, 36). Furthermore, the distribution of age at onset among *GBA1* and *LRRK2* carriers showed heterogeneity across ancestries (appendix 1 p 9).

Among 550 individuals with Parkinson's disease of African admixed ancestry, causal or risk variants were identified in 199 (36·2%) individuals. Almost all carried *GBA1* variants (198 [99·5%] of 199), predominantly driven by the common intronic risk variant, rs3115534-G (figure 3). One individual (0·2%) with early-onset Parkinson's disease (age of onset 27 years) carried a homozygous *PRKN* deletion, and another one (0·2%) harboured both a *LRRK2* causal variant and a *GBA1* risk variant. In controls, *GBA1* risk variants were observed in 210 (23·8%) of 881 individuals, again largely attributable to the rs3115534-G risk variant; *LRRK2* causal or risk variants were identified in four (0·5%) of 881 individuals.

Causal or risk variants were detected in 1510 (53·1%) of 2844 individuals with Parkinson's disease of African

ancestry. *GBA1* variants accounted for the majority of genetic findings, again predominantly driven by rs3115534-G (figure 3). Less frequent findings included variants in *LRRK2* (five [0·2%] of 2844) and *SNCA* (two [<0·1%] of 2844), as well as variants in *PRKN* (four [0·1%] of 2844) and *PINK1* (one [<0·1%] of 2844) observed in individuals with early age of onset (ranging from 27–54 years); most of these individuals (seven of 12) carried a co-occurring *GBA1* variant. In controls, genetic findings were identified in 1691 [37·7%] of 4489 individuals, predominantly in *GBA1*, largely attributable to rs3115534-G, with causal or risk *LRRK2* variants observed in a few individuals (four [<0·1%] of 4489).

In the Ashkenazi Jewish ancestry group, causal or risk variants were identified in 620 [26·5%] of 2343 individuals with Parkinson's disease. Almost all genetic findings involved *GBA1* and *LRRK2*, each characterised by a predominant recurrent variant (*GBA1* p.N409S and *LRRK2* p.G2019S), alongside a small number of additional variants (figures 3, 4). Variants in *GBA1* were observed in 394 (16·8%) individuals and in *LRRK2* in 250 (10·7%) individuals; 25 of these individuals were dual

	Investigated samples, n	Parkinson's disease causal variants									Parkinson's disease risk variants*			Dual carriers†, n	Total risk and causal variant carriers, yield (%)
		Typical autosomal-dominant Parkinson's disease				Early-onset recessive Parkinson's disease			Atypical Parkinson's disease genes	Total causal variant carriers, yield (%)	Risk-associated		Total risk variant carriers, yield (%)		
		LRRK2‡	*SNCA*	*VPS35*	*RAB32*	*PINK1*	*PRKN*	*PARK7*			*GBA1*§	*LRRK2*§			
African admixed															
Parkinson's disease	550	1	0	0	0	0	1	0	1	3/550 (0·5%)	171/3/24	0	198/550 (36·0%)	2	199/550 (36·2%)
Controls	881	3	0	0	0	0	0	0	0	3/881 (0·3%)	201/3/6	1	211/881 (24·0%)	1	213/881 (24·2%)
African															
Parkinson's disease	2844	3	2	0	0	1	4	0	0	10/2844 (0·4%)	1133/33/339	2	1507/2844 (53·0%)	7	1510/2844 (53·1%)
Controls	4495	3	0	0	0	0	0	0	1	4/4489 (0·1%)	1503/3/183	1	1690/4489 (37·6%)	3	1691/4489 (37·7%)
Ashkenazi Jewish															
Parkinson's disease	2343	245/5	1	0	0	0	0	0	0	251/2343 (10·7%)	383/6/5	0	394/2343 (16·8%)	25	620/2343 (26·5%)
Controls	913	20	0	0	0	0	0	0	0	20/913 (2·2%)	73/1/2	0	76/913 (8·3%)	1	95/913 (10·4%)
Latino and Indigenous people of the Americas															
Parkinson's disease	2809	59	2	0	0	0	24	0	0	85/2809 (3·0%)	151/8/3	0	162/2809 (5·8%)	2	245/2809 (8·7%)
Controls	1654	0	0	0	0	0	0	0	0	0/1654	30/0/0	0	30/1654 (1·8%)	0	30/1654 (1·8%)
Complex admixture															
Parkinson's disease	902	21	1	1	0	0	12	0	0	35/902 (3·9%)	130/2/7	9	148/902 (16·4%)	4	179/902 (19·8%)
Controls	444	2	0	0	0	0	0	0	0	2/444 (0·5%)	40/1/2	3	46/444 (10·4%)	0	48/444 (10·8%)
Central Asian															
Parkinson's disease	1284	1	0	0	1	0	7	0	0	9/1284 (0·7%)	73/1/1	23	98/1284 (7·6%)	2	105/1284 (8·2%)
Controls	1480	0	0	0	0	0	1	0	0	1/1480 (0·1%)	29/0/0	14	43/1480 (2·9%)	0	44/1480 (3·0%)
East Asian															
Parkinson's disease	4773	15	4	5	0	17	17	0	1	69/4773 (1·4%)	190/5/0	574/16/11	796/4773 (16·7%)	25	840/4773 (17·6%)
Controls	2809	2	0	0	0	0	0	0	0	2/2809 (0·1%)	14/1/0	165/4/1	185/2809 (6·6%)	0	187/2809 (6·7%)

(Table 2 continues on next page)

	Investigated samples, n	Parkinson's disease causal variants									Parkinson's disease risk variants*			Dual carriers†, n	Total risk and causal variant carriers, yield (%)
		Typical autosomal-dominant Parkinson's disease				Early-onset recessive Parkinson's disease			Atypical Parkinson's disease genes	Total causal variant carriers, yield (%)	Risk-associated		Total risk variant carriers, yield (%)		
		LRRK2‡	*SNCA*	*VPS35*	*RAB32*	*PINK1*	*PRKN*	*PARK7*			*GBA1*§	*LRRK2*§			
(Continued from previous page)															
European															
Parkinson's disease	40773	451/2	31	8	17	5	126	3	8	651/40773 (1·6%)	3312/99/20	21	3452/40773 (8·5%)	49	4054/40773 (9·9%)
Controls	26761	65	2	0	0	0	2	0	2	71/26753 (0·3%)	1249/11/2	6	1268/26753 (4·7%)	6	1333/26753 (5·0%)
Finnish															
Parkinson's disease	127	0	0	0	0	0	1	0	0	1/127 (0·8%)	25/1/0	0	26/127 (20·5%)	1	26/127 (20·5%)
Controls	22	0	0	0	0	0	0	0	0	0/22	4/0/0	0	4/22 (18·2%)	0	4/22 (18·2%)
Middle Eastern															
Parkinson's disease	1347	56/3	1	0	4	11	17	0	0	92/1347 (6·8%)	70/1/2	0	73/1347 (5·4%)	7	158/1347 (11·7%)
Controls	1340	4	0	0	0	0	1	0	0	5/1339 (0·4%)	20/0/0	0	20/1339 (1·5%)	0	25/1339 (1·9%)
South Asian															
Parkinson's disease	807	1	0	1	0	4	5	0	0	11/807 (1·4%)	38/0/1	0	39/807 (4·8%)	1	49/807 (6·1%)
Controls	440	0	0	0	0	0	0	0	0	0/440	4/0/0	0	4/440 (0·9%)	0	4/440 (0·9%)
Total															
Parkinson's disease	58559	853/10	42	15	22	48	214	3	10	1217/58559 (2·1%)	5676/159/402	629/16/11	6893/58559 (11·8%)	125	7985/58559 (13·6%)
Controls	41224	99	2	0	0	0	4	0	3	108/41224 (0·3%)	3167/20/195	191/4/1	3578/41224 (8·7%)	11	3675/41224 (8·9%)

This table reports gene-specific numbers of variant carriers across ancestries (genetic findings), with the total reflecting the number of unique individuals. For autosomal-dominant genes, counts reflect heterozygous carriers, whereas for autosomal-recessive genes, counts reflect (likely) biallelic carriers. Furthermore, this table summarises the total counts and proportions (genetic yield) of individuals carrying risk variants only, causal variants only, or both. The combined category represents unique individuals with any qualifying variant. Because some individuals carry both risk and causal variants, combined counts and percentages are lower than the sum of category-specific values. *All pathogenic or likely pathogenic GBA1 variants are considered risk variants in the context of Parkinson's disease, including Gaucher disease-causing variants of all severities (eg, severe and mild), as well as variants only associated with an increased risk of Parkinson's disease. †Dual carriers refer to individuals harbouring pathogenic or likely pathogenic or Parkinson's disease risk variants in two different genes; the total number reflects the number of individuals rather than variants; therefore, dual carriers are counted only once in the total, and column-wise sums might exceed the total. ‡For columns with two different values, the first denotes heterozygous and the second denotes homozygous carriers. §For columns with three different values, the first denotes heterozygous carriers, the second denotes individuals carrying two different heterozygous variants, and the third denotes homozygous variant carriers.

Table 2: **Summary of genetic findings across ancestries**

GBA1–LRRK2 variant carriers. Only one individual (<0·1%) with Parkinson's disease carried a *SNCA* multiplication. Among controls, 95 (10·4%) of 913 carried causal and risk variants in *GBA1* and *LRRK2*.

Among Latino people and Indigenous people of the Americas, 245 (8·7%) of 2809 individuals with Parkinson's disease carried causal and risk variants across several Parkinson's disease-linked genes, including *GBA1* (162 [5·8%] of 2809), *LRRK2* (59 [2·1%]), *PRKN* (24 [0·9%]), and *SNCA* (two [<0·1%]); two individuals were dual *GBA1–LRRK2* carriers. Variant spectra for *GBA1* and *LRRK2* are shown in figures 3 and 4, respectively. Among controls, causal or risk variants were identified in 30 (1·8%) of 1654 individuals, exclusively involving *GBA1*.

Among individuals with Parkinson's disease of central Asian ancestry, genetic variants were identified in 105 (8·2%) of 1284 individuals, primarily attributable to Parkinson's disease risk-associated variants in *GBA1* (75 [5·8%]) and *LRRK2* (23 [1·8%]; figures 3, 4). Causal variants were rare and included variants in *LRRK2* and *RAB32* (<0·1% each, respectively), as well as *PRKN* variants identified in seven individuals (0·5%). Two individuals carried variants in more than one Parkinson's disease-associated gene (*PRKN–GBA1* and *LRRK2*

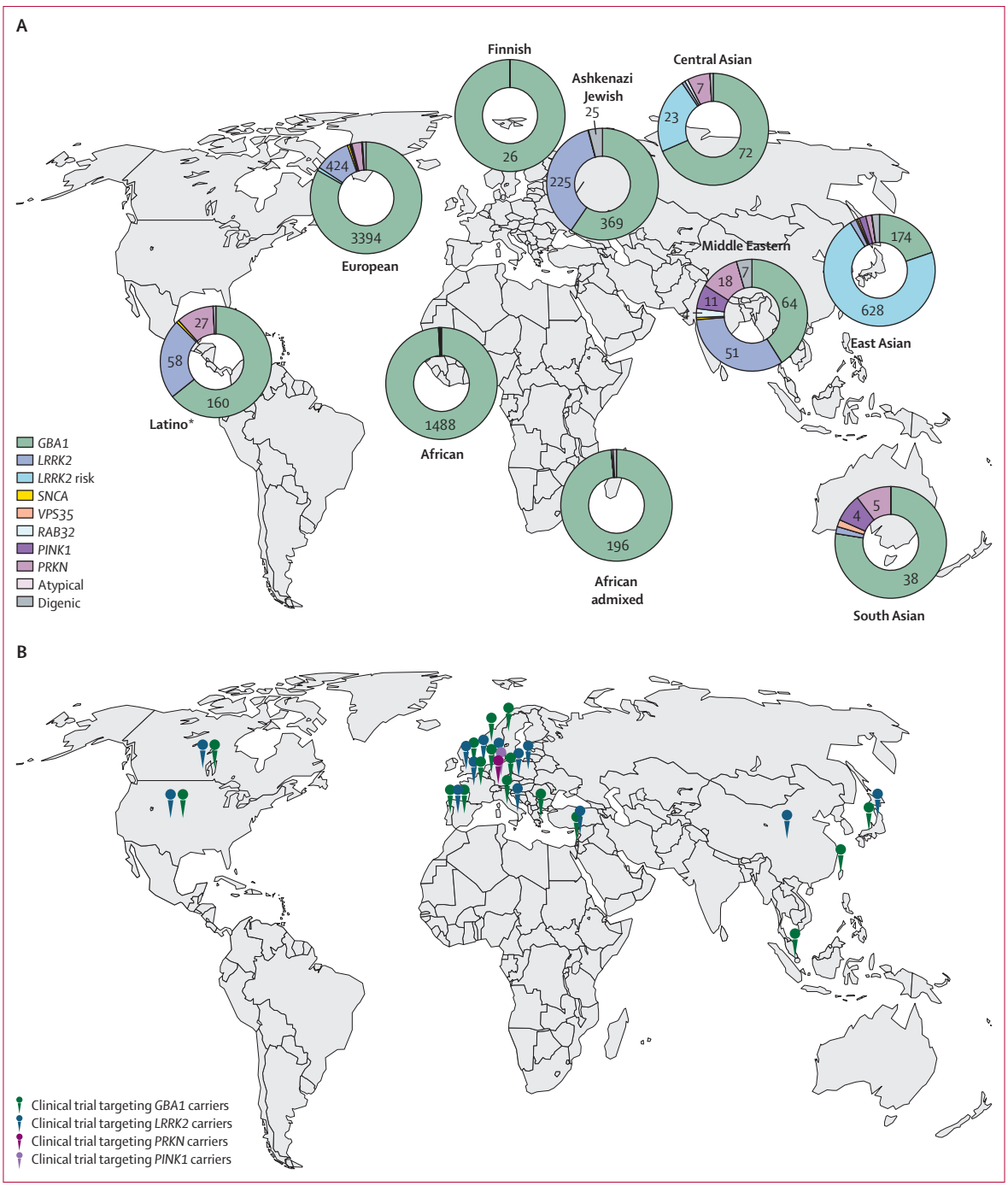

Figure 2: **Global genetic spectrum of Parkinson's disease and locations of clinical trials targeting gene variant carriers**
(A) Charts displaying the percentage of identified carriers per gene among all carriers per ancestry. Dual refers to carriers of variants in two different genes (eg, *GBA1* and a causal gene). (B) Global map of clinical trial locations recruiting gene variant carriers (*GBA1*, *LRRK2*, *PINK1*, and *PRKN*) and targeting proteins encoded by Parkinson's disease-linked genes or pathways in which these genes are involved, obtained from ClinicalTrials.gov. Only one pin per gene and country is shown, regardless of the number of study sites within that country. *Latino people and Indigenous people of the Americas.

risk–*GBA1*). In controls, causal or risk variants were detected in 44 (3·0%) of 1480, including *GBA1* variants (n=29), Parkinson's disease risk-associated *LRRK2* variants (n=14), and a single biallelic *PRKN* carrier.

In the east Asian ancestry group, 840 (17·6%) of 4773 individuals with Parkinson's disease carried causal or risk variants, with the majority harbouring one or both of the two *LRRK2* risk variants p.R1628P and p.G2385R,

Figure 3: **Mutational spectrum and pathogenicity of *GBA1* variants across ancestries**

Charts illustrate the distribution of *GBA1* variants in individuals with Parkinson's disease stratified by variant severity (inner rings) and specific variants (outer rings) across ancestries. Variant severity was assigned using the GBA1-PD browser. The size of each segment reflects the number of variant carriers within each ancestry. Inner-ring colours denote severity categories: severe (purple), mild (blue), risk (green), and unknown (grey). Outer-ring segments represent individual variants, with colour shades corresponding to their severity classification. To improve readability, variants identified in only a single individual were grouped as "Other" for central Asian, Middle Eastern, and south Asian populations; for African admixed, African, Latino and Indigenous people of the Americas, east Asian, and European populations, variants identified in five or fewer individuals were grouped. Numbers within the circles indicate the total number of *GBA1* variant carriers identified per ancestry; individuals carrying two different *GBA1* variants contributed to both variant counts. *Latino people and Indigenous people of the Americas.

identified in a total of 601 (12·6%) individuals. *GBA1* variants were the next most frequent in 190 (4·0%) individuals, with a broad and heterogeneous variant spectrum (figure 3), including over 40 distinct variants. Causal *LRRK2* variants were rare (15 [0·3%]; figure 4); other causal variants were found in *SNCA* (four [<0·1%]), *VPS35* (five [0·1%]), *PINK1* (17 [0·4%]), and *PRKN* (17 [0·4%]). All biallelic *PINK1* variant carriers harboured

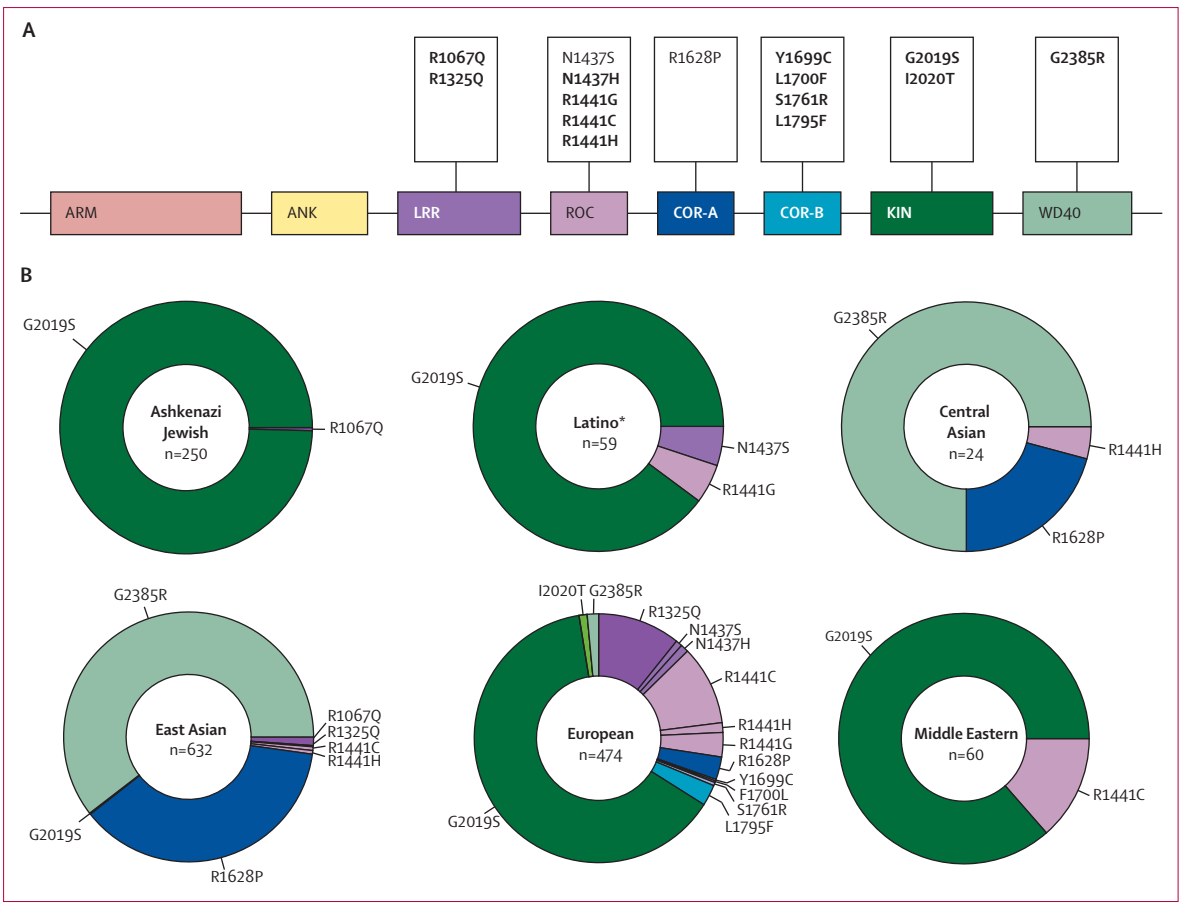

Figure 4: **Distribution of *LRRK2* variants across protein domains and ancestries**
(A) Schematic representation of the LRRK2 protein with functional domains. Boxes indicate locations of the identified *LRRK2* variants within the respective functional domains of the protein. Variants highlighted in bold have been previously shown to increase LRRK2 kinase activity. (B) Charts illustrating the ancestry-specific spectrum of *LRRK2* variants in individuals with Parkinson's disease. Each segment represents a specific variant; the segment size is proportional to the number of individuals carrying that variant. If two variants were present in one individual, both variants were counted. Segment colours correspond to the LRRK2 protein domain in which the variant is located, as indicated in part A. Only ancestries with more than five variant carriers are shown. *Latino people and Indigenous people of the Americas.

the same p.L347P variant in the homozygous state, whereas this variant was also observed in the heterozygous state in Parkinson's disease (13 [0·3%] of 4773) and controls (six [0·2%] of 2809). 25 individuals with Parkinson's disease were dual carriers, most frequently harbouring an *LRRK2* risk variant in combination with a *GBA1* risk variant. Similarly, among the 187 (6·7%) of 2809 controls identified as carriers of a Parkinson's disease-associated causal or risk variant, *LRRK2* risk variants were most frequent (n=170).

In the European ancestry group, genetic findings were identified in 4054 (9·9%) of 40773 individuals with Parkinson's disease, across a range of different Parkinson's disease-associated genes. *GBA1* variants accounted for the largest proportion, with 3431 carriers (8·4%). Although the established Parkinson's disease risk variants p.E365K and p.T408M were the most frequent, a broad spectrum of additional *GBA1* variants was identified across the European population (figure 3). Causal *LRRK2* variants were the second most frequent finding, identified in 453 (1·1%) individuals, and encompassed a broad range of variants, including several predominantly or exclusively observed in this ancestry (figure 4). *LRRK2* risk variants were detected in a small subset (21 [<0·1%]). 49 (0·1%) of 40773 were dual carriers, most frequently involving a *GBA1* risk variant in combination with an *LRRK2* causal variant. In addition to *GBA1* and *LRRK2*, variants were detected in several other Parkinson's disease-associated genes, including *SNCA* (31 [<0·1%]), *VPS35* (eight [<0·1%]), and *RAB32* (17 [<0·1%]). Among recessive genes, biallelic variants were most frequently identified in *PRKN* (126 [0·3%] of 40773), whereas *PINK1* (n=5) and *PARK7* (n=3) variants were rare (both <0·1%, respectively). By comparison, 1333 (5·0%) of 26753 controls carried causal or risk variants, most often involving *GBA1*, with fewer causal *LRRK2* or other variants.

In the Finnish ancestry group, we observed a causal or risk variant in 26 (20·5%) of 127 individuals with Parkinson's disease. One individual carried a biallelic *PRKN* with a co-occurring *GBA1* variant, whereas the

remaining individuals carried *GBA1* variants only. Similarly, genetic findings were confined to *GBA1* among controls (four [18·2%] of 22).

Among individuals with Parkinson's disease in the Middle Eastern ancestry group, Parkinson's disease-associated causal and risk variants were identified in 158 (11·7%) of 1347 individuals and were distributed across multiple Parkinson's disease-associated genes. Most frequent were risk variants in *GBA1* (73 [5·4%]) and causal variants in *LRRK2* (59 [4·4%]), with the *LRRK2* variant spectrum largely driven by p.G2019S (figure 4), whereas *GBA1* showed a more heterogeneous spectrum (figure 3). Additional findings included variants in *PRKN* (17 [1·3%]), *PINK1* (11 [0·8%]), *RAB32* (four [0·3%]), and *SNCA* (n=1, <0·1%). Seven individuals were dual carriers of *GBA1* and *LRRK2* variants. Among controls, causal or risk variants were detected in 25 (1·9%) of 1339, most frequently involving *GBA1*.

In the south Asian ancestry group, causal or risk variants were identified in 49 (6·1%) of 807 individuals with Parkinson's disease. *GBA1* variants were the most frequent findings (39 [4·8%]), with p.L483R being the most frequent variant (figure 3). Variants in recessive Parkinson's disease genes were also detected, including *PRKN* (five [0·6%]) and *PINK1* (four [0·5%]), whereas causal variants in *LRRK2* and *VPS35* were each identified in a single individual only (0·1% each, respectively). Among controls (four [0·9%] of 440), only *GBA1* variants were identified.

Across all ancestries, 1182 (2·0%) of 58 559 individuals with Parkinson's disease carried single heterozygous causal variants in recessive genes, most frequently single heterozygous variants in *PRKN* (n=1064), followed by *PINK1* (n=40), *VPS13C* (n=23), *ATP13A2* (n=22), and *FBXO7* (n=21). Variants in *JAM2, PARK7*, and *SYNJ1* were rarer (n<20). A substantial proportion of these (377 [41·7%] of 904) had an age of onset of 50 years or younger (age of onset missing for n=278). The overall proportion of individuals with an age of onset of 50 years or younger among the entire Parkinson's disease group included in this study was 11 814 (26·4%) of 44 768 (age of onset missing for n=13791). By comparison, the frequency of single heterozygous causal variants in recessive genes in controls was 583 (1·4%) of 41 224. 11 individuals carried two heterozygous *PRKN* variants but had an age of onset greater than 60 years or the age of onset was unavailable, so we did not interpret them as likely compound heterozygous without further confirmation.

Discussion

In this multi-ancestry investigation, we delineated shared and ancestry-specific patterns of genetic risk and causation across established Parkinson's disease genes, extending previous work that has largely focused on European ancestry and smaller, ancestry-restricted cohorts. Our work represents the largest ancestry-informed assessments to date, including about 100 000 individuals with around 30% of non-European and non-Ashkenazi Jewish ancestry, offering insights into the ancestry-specific genetic architecture of Parkinson's disease. This characterisation of genetic Parkinson's disease across diverse ancestries is important because multiple ongoing or planned clinical trials target proteins encoded by Parkinson's disease-linked genes.[7,20] Ensuring these advances are inclusive requires a detailed understanding of genetic variation beyond European ancestry.[21]

GBA1 risk variants were most frequent overall and identified across all populations, albeit with distinct ancestry-specific variant spectra: p.N409S was the most common variant in the Ashkenazi Jewish group, whereas p.E365K and p.T408M were more frequent in the European, Finnish, and central Asian groups, but rare in African, African admixed, and east Asian populations. The variant spectrum in the east Asian group was broad, but p.L483P appeared more frequent than in other populations. Similarly, we newly highlight p.L483R as a recurrent variant in south Asian individuals. The intronic variant rs3115534-G was the most frequently observed in the African admixed and African populations, as previously reported.[10] This variant represents a clear example of an ancestry-specific *GBA1* risk allele that would have remained undetected in studies restricted to populations of European ancestry, underscoring the necessity of ancestry-diverse genetic investigations. Collectively, these findings emphasise the importance of *GBA1* as a globally relevant therapeutic target, an important observation given that most trials targeting glucocerebrosidase are conducted in countries with a high representation of European or Ashkenazi Jewish populations.[4]

LRRK2 variants also showed ancestral variability. p.G2019S was detected across multiple ancestries, with the highest frequencies in the Middle Eastern (partially reflecting North African Berbers), Latino and Indigenous people of the Americas, and Ashkenazi Jewish populations.[22,23] Although p.G2019S was also the most frequent *LRRK2* variant in the European population, the mutational spectrum was notably broader and included p.R1441C, p.R1441G, and p.L1795F, with suggested founder effects in European sub-populations.[24–27] By contrast, p.G2019S was largely absent from Asian populations, where the risk variants p.R1628P and p.G2385R were most frequent,[28] especially in the east Asian group. The most frequent causal *LRRK2* variant identified in the east Asian population was p.R1067Q, only recently demonstrated to be disease-relevant and shown to be enriched in east Asians.[13] Our data newly document the first causal *LRRK2* variant carriers of African ancestry, supporting the relevance of LRRK2-mediated Parkinson's disease in populations that have been largely absent from genetic and trial efforts to date. Together, these observations illustrate that, although *LRRK2* is a globally relevant therapeutic target, ancestry-specific studies are essential to capture distinct variant spectra. *LRRK2* variants differ in their impact on kinase activity,[29] which might inform clinical trial design and therapeutic targeting. Investigating *LRRK2*

variants that affect kinase activity and could affect eligibility for, or response to, emerging LRRK2 kinase-directed therapies is therefore important globally, particularly given the ongoing LRRK2-targeted trials.

SNCA variants were rare overall but were identified across multiple ancestries, in line with previous reports.[30] Although *SNCA* multiplications were identified in several populations, missense variants (p.A53T and p.G51D) were predominantly observed in the European population. Other known autosomal dominant forms of Parkinson's disease caused by *VPS35* p.D620N or *RAB32* p.S71R were rare. *RAB32* variants were only detected in European and Middle Eastern individuals (more frequent Middle Eastern individuals), whereas *VPS35* variant carriers were observed across three different populations, with an apparent enrichment in Asian populations (east and south Asian) compared with Europeans. This ancestry pattern has not previously been described at this scale. Given the small sample size of some populations, especially with sequencing data, these findings need cautious interpretation. Nonetheless, detecting *VPS35* and *RAB32* variants across multiple ancestries underscores the importance of continued investigation in larger, diverse cohorts.

Finally, investigating recessive genes confirms *PRKN* as the most frequently implicated cause of recessive Parkinson's disease across nearly all ancestries and indicates a wider-than-appreciated distribution of *PINK1* variants, underscoring the global relevance of approaches based on the PINK1–Parkin pathway. In Asian populations (east, central, and south Asian), *PRKN* variants were more frequent than causal *LRRK2* variants. Of note, identifying *PRKN* carriers across multiple ancestries reinforces its global relevance, as preclinical efforts increasingly explore therapeutic strategies targeting Parkin or its associated pathways.[7] Although *PINK1* variants were rare, we identified a notable proportion of Middle Eastern, east Asian, and south Asian carriers. All east Asian *PINK1* carriers harboured p.L347P, a variant known to be enriched in east Asian populations.[31] We further identified numerous individuals carrying single heterozygous causal variants in recessive Parkinson's disease genes. Although these are not causal, many carriers had an early age of onset of 50 years or younger, suggesting a second causal variant, potentially a complex structural variant not captured by genotyping or short-read sequencing.

This global analysis underscores the clinical and translational importance of extending Parkinson's disease genetics research beyond European ancestry. Ancestry-informed genetic studies improve diagnostic accuracy, variant interpretation, and genotype–phenotype correlations, with direct implications for prognosis and genetic counselling. Our findings also show that variants predominantly identified in European populations often contribute only modestly in other ancestries, suggesting additional ancestry-specific variants and highlighting that our current understanding of the genetic architecture of Parkinson's disease remains incomplete. From a translational perspective, such diversity is essential for equitable clinical trial design, accurate identification of trial-eligible individuals, and the global applicability of emerging genetically targeted therapies. Together, these data provide a robust foundation for future in-depth ancestry-specific analyses and genetically informed therapeutic development efforts in Parkinson's disease.

We consider our efforts complementary to ongoing studies investigating common genetic risk through GWAS in diverse populations,[10–12,32] which continue to expand our understanding of Parkinson's disease susceptibility across ancestries. In this context, we focused specifically on established causal variants and risk-associated variants in genes involved in pathways targeted by ongoing clinical trials, while acknowledging that other, well established Parkinson's disease risk loci also contribute substantially to disease susceptibility (eg, common risk variants at the *SNCA* locus and the *MAPT* H1 haplotype). Although rare causal variants in *SNCA* were included in our analyses because of their established direct role in disease pathogenesis, common GWAS-associated risk variants at the *SNCA* and *MAPT* loci were not included, because they are not currently used for genetically stratified clinical trial recruitment or therapeutic targeting in the same way as *GBA1* and *LRRK2* risk variants; however, they are increasingly explored in large-scale GWAS and dedicated haplotype analyses across diverse populations.[33]

Our study has limitations. It focused on variants in known Parkinson's disease genes, primarily identified in European populations. This specific focus and the predominance of European-ancestry data in resources like ClinVar could bias variant interpretation in non-European groups. Although our study allowed for a large-scale, clinically relevant assessment of those known variants, it was not designed to lead to the discovery of novel Parkinson's disease genes and variants that could be more relevant in other populations. Systematically studying large, well powered non-European ancestry cohorts using sequencing data will be important to capture the full global spectrum of novel variation. Another important area will be investigating variants of uncertain significance in known Parkinson's disease genes. Some variants might be causal in specific populations but are currently classified as of uncertain significance due to scarce ancestry-specific reference data and low statistical power, whereas other variants might be overinterpreted in populations for which adequate reference data are lacking. Ancestry-aware classification (or reclassification) using variant frequency, functional, and segregation data will be key for establishing pathogenicity, improving genetic counselling, and enabling trial access.[13,27]

An additional limitation of this study is the small sample size of some populations. This limitation might inflate the relative contribution of known genes in well powered populations and is particularly relevant for groups with limited whole-genome sequencing data, because some

variants are not captured by genotyping (eg, *RAB32* p.S71R and select *GBA1* variants).

The detection of compound-heterozygous variants was limited by the absence of phased data, preventing us from determining whether multiple variants were located on the same or opposite parental alleles. We conservatively considered carriers of two heterozygous causal variants in recessive genes with an age of onset of 60 years or less as likely compound heterozygous, but validation is needed. Although genetically inferred ancestry provides a more robust framework than geographical labels, the broad ancestry categories still encompass substantial heterogeneity and might reflect cohort-specific recruitment patterns. Nonetheless, we believe this approach represents the most transparent and scalable strategy currently feasible for large, harmonised cross-ancestry analyses.

Lastly, findings were generated in a research setting; although we employed quality control measures and whole-genome sequencing validation where possible, further confirmation of identified variants in certified diagnostic laboratories is required before clinical implementation.

In conclusion, this study represents the largest global assessment of the genetic spectrum of Parkinson's disease, offering crucial insights into the population-specific genetic architecture of Parkinson's disease and underscoring the crucial need to expand Parkinson's disease genetics research beyond European ancestry. Ancestry-informed analyses enhance diagnostic precision and risk interpretation and are essential to ensure equitable access to genetically stratified clinical trials and emerging precision therapies. Addressing current gaps in data diversity and global genetic representation is not optional but fundamental to translating genetic discovery into effective and inclusive genetically informed therapies in Parkinson's disease on a global scale.

Contributors

LML, Z-HF, and NK accessed and verified the data reported in this manuscript. LML, SoB, HH, RK, KRK, S-YL, NEM, WMYM, AJN, RO, NUO, SuR, CS, SS, AHT, ZT, BT, EMV, NZ, HRM, and CK carried out study assessments for participants, including clinical data and biospecimen collection. Z-HF, MBM, MJK, HLL, KSL, MAN, and DV were responsible for processing, performing quality control, and harmonising genotyping and sequencing data from the Global Parkinson's Genetics Program (GP2). HI and LJ were responsible for processing and harmonising clinical data from GP2. KAB, ShB, MLD, SJ, JJ, EJS, and MT contributed to the acquisition of data. MBM, HLL, KL, SB-C, CB, and ABS contributed to the analysis of genetic data. LML and CK conceptualised the study. PH, LS, and JT supported the execution of the study. CB, ABS, HRM, and CK supervised the study. LML drafted the manuscript. All authors had full access to all the data in the study and had final responsibility for the decision to submit the manuscript for publication.

Declaration of interests

LML received faculty honoraria from the Movement Disorder Society in the past 12 months, unrelated to this manuscript. Z-HF is supported by Aligning Science Across Parkinson's (ASAP) GP2 and receives GP2 salary support from The Michael J Fox Foundation for Parkinson's Research (MJFF). The participation of ZH-F, MBM, NK, ShB, HI, LJ, MJK, HLL, KSL, DV, and MAN in this project was part of a competitive contract awarded to Datatecnica by the National Institutes of Health (NIH) to support open science research. MAN also serves on the scientific advisory board for Character Bio, and is a scientific founder at Neuron23 Inc and owns stock. KAB and MT are supported by ASAP GP2. PH serves as an advisor to Alector, the Global Parkinson's Genetics Consortium (MJFF), LSP Advisory, and Neuro.VC. HH and RK received essential funding from the Wellcome Trust, the Medical Research Council (MRC), the Multiple System Atrophy Trust, the National Institute for Health Research (NIHR) University College London Hospitals Biomedical Research Centre, MJFF, The Fidelity Trust, Rosetrees Trust, the Dolby Family Fund, Alzheimer's Research UK, MSA Coalition, The Guarantors of Brain, Cerebral Palsy Alliance, Friedreich's Ataxia Research Alliance (FARA), European Academy of Neurology (EAN) and the NIH NeuroBioBank, Queen Square BrainBank, and the MRC Brainbank Network. JJ was supported by MJFF (Data Community Innovators Program). KRK is supported by the Ainsworth 4 Foundation and the Medical Research Future Fund. S-YL received consultancies and grants from MJFF and ASAP GP2. Unrelated to this manuscript, S-YL received honoraria for participating as a member of the Neurotorium Editorial Board and honoraria for lecturing or teaching from the International Parkinson and Movement Disorder Society (MDS) and Medtronic. S-YL also received stipends from the MDS as Chair of the Asian-Oceanian Section, and from *npj Parkinson's Disease* as Associate Editor. NEM receives NIH funding (1K08NS131581), is supported by ASAP GP2, and is a member of the steering committee of the PD GENEration study for which he receives an honorarium from the Parkinson's Foundation. WMYM is the founding member and Chair of the AfrAbia Parkinson's Disease Genomic Consortium (AfrAbia PD-GC), which is supported by funding from GP2. AJN reports grants from Parkinson's UK, Barts Charity, Cure Parkinson's, NIHR, Innovate UK, the Medical College of Saint Bartholomew's Hospital Trust, Alchemab, ASAP GP2, and MJFF, and consultancy and personal fees from AstraZeneca, AbbVie, Profile, Bial, Charco Neurotech, Alchemab, Sosei Heptares, Umedeor, and Britannia. AJN is an Associate Editor for the *Journal of Parkinson's Disease*. RO has received travel grants from the Movement Disorder Society in the past 12 months, unrelated to this manuscript, and received research grants and travel support to attend annual GP2 meetings from ASAP GP2. NUO reports funding from MJFF and UK NIHR institutional grant funding, both for Parkinson's disease research. AHT receives support from MJFF and GP2. Unrelated to this manuscript, AHT received speaker honoraria from International Parkinson and Movement Disorder Society, Eisai, and Orion Pharma, and reports consultancies from Elsevier as Section Editor for *Parkinsonism and Related Disorders*. JT is supported by the German Research Foundation (DFG), ASAP GP2, and is a consultant for Acurex. BT acknowledges support from GP2. EMV is supported by ASAP GP2. Unrelated to this manuscript, BT received research grants from Telethon Foundation Italy (GGP20070), the Italian Ministry of Health (Ricerca Finalizzata RF-2019-12369368 and ERA-NET Neuron NDCil project EUR002), the Italian Ministry of University and Research (GENERA, MNESYS), and the Silverstein Foundation. KL received grants from the Dystonia Medical Research Foundation and DFG, unrelated to this manuscript. CB is an employee of the Coalition for Aligning Science. ABS is the lead investigator for a grant from MJFF and has a contract for work on GP2. Unrelated to this manuscript, ABS is named as an inventor on patents for a diagnostic for stroke and for molecular testing for C9orf72 repeats; is on the scientific advisory board of the Lewy Body Disease Association and for Cajal Neuroscience (both unpaid positions); and received an honorarium for speaking at the World Laureates Association. HRM reports grant support from Parkinson's UK, Cure Parkinson's Trust, PSP Association, MRC, and MJFF. Unrelated to this manuscript, HRM is a co-applicant on a patent application related to C9orf72 (Method for diagnosing a neurodegenerative disease; PCT/GB2012/052140) and received honoraria from MDS. CK received grants from MJFF and ASAP GP2. Unrelated to this manuscript, CK received grants from DGF and speakers' honoraria from Bial. CK has royalties at Oxford University Press and Springer Nature and serves as a medical advisor to Centogene and Biogen. All other authors declare no competing interests.

Data sharing

Data used in the preparation of this article were obtained from GP2 (https://gp2.org). Specifically, we used Tier 2 data from GP2 release 11 (https://doi.org/10.5281/zenodo.17753486). Tier 1 data can be accessed by completing a form on the Accelerating Medicines Partnership in

Parkinson's Disease (AMP-PD) website (https://amp-pd.org/register-for-amp-pd). Tier 2 data access requires approval and a Data Use Agreement signed by your institution. Qualified researchers are encouraged to apply for direct access to the data through AMP-PD. All code generated for this Article, and the identifiers for all software programs and packages used, are available on GitHub (https://github.com/GP2code/GP2-global-genetic-variant-landscape) and were given a persistent identifier via Zenodo (https://doi.org/10.5281/zenodo.15699539). A detailed list of all identified variant carriers, including corresponding anonymised GP2-IDs, variant details, and basic demographic and clinical characteristics are available to qualified researchers upon reasonable request made to the corresponding authors.

Acknowledgments

This project was supported by GP2. GP2 is funded by the ASAP initiative and implemented by MJFF. For a complete list of GP2 members see https://doi.org/10.5281/zenodo.7904831. This research was supported in part by the Intramural Research Program of the NIH. The contributions of the NIH authors are considered Works of the US Government. The findings and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of the NIH or the US Department of Health and Human Services. We thank Yuliia Kanana for her valuable support in processing and handling the samples used in this study. We are grateful for the important support from all study participants and their families, as well as for contributions of all collaborators, brain banks, and biobanks.

Editorial note: The Lancet Group takes a neutral position with respect to territorial claims in published maps.

References

1 Lim S-Y, Klein C. Parkinson's disease is predominantly a genetic disease. *J Parkinsons Dis* 2024; **14:** 467–82.
2 The Global Parkinson's Genetics Program (GP2), Leonard HLL. Novel Parkinson's disease genetic risk factors within and across European populations. *medRxiv* 2025; published online March 17. https://doi.org/10.1101/2025.03.14.24319455 (preprint).
3 Lange LM, Gonzalez-Latapi P, Rajalingam R, et al. Nomenclature of genetic movement disorders: recommendations of the International Parkinson and Movement Disorder Society Task Force—an update. *Mov Disord* 2022; **37:** 905–35.
4 Lim S-Y, Tan AH, Ahmad-Annuar A, et al. Uncovering the genetic basis of Parkinson's disease globally: from discoveries to the clinic. *Lancet Neurol* 2024; **23:** 1267–80.
5 Junker J, Lange LM, Vollstedt EJ, et al. Team science approaches to unravel monogenic Parkinson's disease on a global scale. *Mov Disord* 2024; **39:** 1868–73.
6 Vollstedt E-J, Schaake S, Lohmann K, et al. Embracing monogenic Parkinson's disease: the MJFF Global Genetic PD cohort. *Mov Disord* 2023; **38:** 286–303.
7 Senkevich K, Rudakou U, Gan-Or Z. New therapeutic approaches to Parkinson's disease targeting GBA, LRRK2 and Parkin. *Neuropharmacology* 2022; **202:** 108822.
8 Global Parkinson's Genetics Program. GP2: the Global Parkinson's Genetics Program. *Mov Disord* 2021; **36:** 842–51.
9 Lange LM, Avenali M, Ellis M, et al. Elucidating causative gene variants in hereditary Parkinson's disease in the Global Parkinson's Genetics Program (GP2). *NPJ Parkinsons Dis* 2023; **9:** 100.
10 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:** 1015–25.
11 Leal TP, Waldo E, Duarte-Zambrano F, et al. Genotype-phenotype association study conducted on LARGE-PD reveals novel loci associated with Parkinson's Disease. *medRxiv* 2025; published online July 18. https://doi.org/10.1101/2025.07.18.25331793 (preprint).
12 Kishore A, Kumar Sreelatha AA, Tenghe AMM, et al. Deciphering the genetic architecture of Parkinson's disease in India. *medRxiv* 2025; published online Feb 21 https://doi.org/10.1101/2025.02.17.25322132 (preprint).
13 Lim S-Y, Toh TS, Hor JW, et al. Clinical and functional evidence for the pathogenicity of the LRRK2 p.Arg1067Gln variant. *NPJ Parkinsons Dis* 2025; **11:** 34.
14 Velez-Pardo C, Lorenzo-Betancor O, Jimenez-Del-Rio M, et al. The distribution and risk effect of GBA variants in a large cohort of PD patients from Colombia and Peru. *Parkinsonism Relat Disord* 2019; **63:** 204–08.
15 Cook L, Verbrugge J, Schwantes-An TH, et al. Parkinson's disease variant detection and disclosure: PD GENEration, a North American study. *Brain* 2024; **147:** 2668–79.
16 Westenberger A, Skrahina V, Usnich T, et al. Relevance of genetic testing in the gene-targeted trial era: the Rostock Parkinson's disease study. *Brain* 2024; **147:** 2652–67.
17 Vitale D, Koretsky MJ, Kuznetsov N, et al. GenoTools: an open-source Python package for efficient genotype data quality control and analysis. *G3 (Bethesda)* 2024; **15:** jkae268.
18 Kuznetsov N, Daida K, Makarious MB, et al. CNV-Finder: streamlining copy number variation discovery. *bioRxiv* 2025; published online July 26. https://doi.org/10.1101/2024.11.22.624040 (preprint).
19 Lesage S, Lunati A, Houot M, et al. Characterization of recessive Parkinson's disease in a large multicenter study. *Ann Neurol* 2020; **88:** 843–50.
20 McFarthing K, Buff S, Rafaloff G, et al. Parkinson's disease drug therapies in the clinical trial pipeline: 2023 update. *J Parkinsons Dis* 2023; **13:** 427–39.
21 Schumacher-Schuh AF, Bieger A, Okunoye O, et al. Underrepresented populations in Parkinson's genetics research: Current landscape and future directions. *Mov Disord* 2022; **37:** 1593–604.
22 El Otmani H, Daghi M, Tahiri Jouti N, Lesage S. An overview of the worldwide distribution of *LRRK2* mutations in Parkinson's disease. *Neurodegener Dis Manag* 2023; **13:** 335–50.
23 Lesage S, Leutenegger AL, Ibanez P, et al. LRRK2 haplotype analyses in European and North African families with Parkinson disease: a common founder for the G2019S mutation dating from the 13th century. *Am J Hum Genet* 2005; **77:** 330–32.
24 Gorostidi A, Ruiz-Martínez J, Lopez de Munain A, Alzualde A, Martí Massó JF. LRRK2 G2019S and R1441G mutations associated with Parkinson's disease are common in the Basque Country, but relative prevalence is determined by ethnicity. *Neurogenetics* 2009; **10:** 157–59.
25 Mata IF, Taylor JP, Kachergus J, et al. LRRK2 R1441G in Spanish patients with Parkinson's disease. *Neurosci Lett* 2005; **382:** 309–11.
26 Nuytemans K, Rademakers R, Theuns J, et al. Founder mutation p.R1441C in the leucine-rich repeat kinase 2 gene in Belgian Parkinson's disease patients. *Eur J Hum Genet* 2008; **16:** 471–79.
27 Lange LM, Levine K, Fox SH, et al. The LRRK2 p.L1795F variant causes Parkinson's disease in the European population. *NPJ Parkinsons Dis* 2025; **11:** 58.
28 Lim S-Y, Tan AH, Ahmad-Annuar A, et al. Parkinson's disease in the Western Pacific region. *Lancet Neurol* 2019; **18:** 865–79.
29 Kalogeropulou AF, Purlyte E, Tonelli F, et al. Impact of 100 LRRK2 variants linked to Parkinson's disease on kinase activity and microtubule binding. *Biochem J* 2022; **479:** 1759–83.
30 Trinh J, Zeldenrust FMJ, Huang J, et al. Genotype-phenotype relations for the Parkinson's disease genes SNCA, LRRK2, VPS35: MDSGene systematic review. *Mov Disord* 2018; **33:** 1857–70.
31 Patel SG, Buchanan CM, Mulroy E, et al. Potential PINK1 founder effect in Polynesia causing early-onset Parkinson's disease. *Mov Disord* 2021; **36:** 2199–200.
32 Kim JJ, Vitale D, Otani DV, et al. Multi-ancestry genome-wide association meta-analysis of Parkinson's disease. *Nat Genet* 2024; **56:** 27–36.
33 Reyes-Pérez P, Hor JW, Toh TS, et al. Exploring MAPT-containing H1 and H2 haplotypes in Parkinson's disease across diverse populations. *medRxiv* 2025; published online Oct 19. https://doi.org/10.1101/2025.10.17.25337033 (preprint).