

The LRRK2 p.L1795F variant causes Parkinson's disease in the European population

Lara Lange

l.a.lange@uni-luebeck.de

University of Lübeck <https://orcid.org/0000-0002-7162-9821>

Kristin Levine

Data Tecnica International <https://orcid.org/0000-0002-5702-0980>

Susan Fox

Edmond J. Safra Program in Parkinson's Disease and the Morton and Gloria Shulman Movement Disorders Clinic, Toronto Western Hospital, University Health Network, University of Toronto, Toronto, Ontar

Connie Marras

University of Toronto

Nazish Ahmed

University Health Network

Nicole Kuznetsov

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

Dan Vitale

Data Tecnica International LLC

Hiroataka Iwaki

Data Tecnica International <https://orcid.org/0000-0002-8982-7885>

Katja Lohmann

Institute of Neurogenetics, University of Luebeck, Luebeck, Germany

Luca Marsili

University of Cincinnati, Cincinnati, Ohio, USA

Alberto Espay

Gardner Family Center for Parkinson's Disease and Movement Disorders, Department of Neurology, University of Cincinnati College of Medicine, Cincinnati, Ohio <https://orcid.org/0000-0002-3389-136X>

Peter Bauer

Centogene GmbH, Rostock, Germany

Christian Beetz

CENTOGENE GmbH, Am Strande 7, 18055 Rostock, Germany

Jessica Martin

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
<https://orcid.org/0009-0000-0830-9289>

Stewart Factor

Emory University <https://orcid.org/0000-0002-0449-973X>

Lenora Higginbotham

Emory

Honglei Chen

Michigan State University

Hampton Leonard

National Institute on Aging

Mike Nalls

National Institute on Aging, National Institutes of Health <https://orcid.org/0000-0003-0319-4325>

Niccolo E. Mencacci

Huw Morris

University College London <https://orcid.org/0000-0002-5473-3774>

Christine Klein

Institute of Neurogenetics and Department of Neurology

Cornelis Blauwendraat

NIH <https://orcid.org/0000-0001-9358-8111>

Zih-Hua Fang

German Center for Neurodegenerative Diseases (DZNE) <https://orcid.org/0000-0002-0225-8772>

Article

Keywords:

Posted Date: September 20th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4772543/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: (Not answered)

The *LRRK2* p.L1795F variant causes Parkinson's disease in the European population

Lara M. Lange, MD,^{1,2} Kristin Levine,^{3,4} Susan H. Fox, MB, PhD,⁵ Connie Marras, MD, PhD,⁵ Nazish Ahmed, BSc, LMCC,⁵ Nicole Kuznetsov,^{3,4} Dan Vitale,^{3,4} Hirotaka Iwaki, MD, PhD,^{3,4,6} Katja Lohmann, PhD,¹ Luca Marsili, MD, PhD,⁷ Alberto J. Espay, MD,⁷ Peter Bauer, MD,⁸ Christian Beetz, PhD,⁸ Jessica Martin,⁴ Stewart A. Factor, DO,⁹ Lenora A. Higginbotham, MD,⁹ Honglei Chen, MD, PhD,¹⁰ Hampton Leonard,^{3,4} Mike Nalls,^{3,4} PhD, Niccolo E. Mencacci, MD, PhD,¹¹ Huw R. Morris, MB, PhD,^{12,13} Christine Klein, MD,^{1,2} Cornelis Blauwendraat, PhD,^{4,6} Zih-Hua Fang, PhD,¹⁴ and *the Global Parkinson's Genetics Program (GP2)*

¹ Institute of Neurogenetics, University of Luebeck, Luebeck, Germany

² Department of Neurology, University Hospital Schleswig-Holstein, Luebeck, Germany

³ DataTecnica, Washington, DC, USA

⁴ 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

⁵ Edmond J. Safra Program in Parkinson's Disease and the Morton and Gloria Shulman Movement Disorders Clinic, Toronto Western Hospital, University Health Network, University of Toronto, Toronto, Ontario, Canada.

⁶ Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA

⁷ University of Cincinnati, Cincinnati, Ohio, USA

⁸ CENTOGENE GmbH, Am Strande 7, 18055 Rostock, Germany

⁹ Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA

¹⁰ Department of Epidemiology and Biostatistics, Michigan State University, MI, USA

¹¹ Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

¹² Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, London, UK

¹³ UCL Movement Disorders Centre, University College London, London, UK

¹⁴ German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany

Corresponding authors:

Lara M. Lange

Institute of Neurogenetics, University of Luebeck

Ratzeburger Allee 160, 23562 Luebeck, Germany

email: la.lange@uni-luebeck.de

ORCID: [0000-0002-7162-9821](https://orcid.org/0000-0002-7162-9821)

Zih-Hua Fang

German Center for Neurodegenerative Diseases

Otfried-Müller-Straße 23, 72076 Tübingen, Germany

email: Zih-Hua.Fang@dzne.de

ORCID: [0000-0002-0225-8772](https://orcid.org/0000-0002-0225-8772)

Abstract

Pathogenic variants in the *LRRK2* gene represent the most common cause of autosomal dominant Parkinson's disease (PD) worldwide. We identified the *LRRK2* p.L1795F variant in 14 White/European ancestry PD patients, including two families with multiple affected carriers and seven additional affected individuals with familial PD using genotyping and sequencing data from more than 50,000 individuals through GP2, AMP-PD, PDGENERation, and CENTOGENE. All variant carriers were of White/European ancestry, and those with available genotyping data shared a common haplotype. The clinical presentation of p.L1795F carriers resembles that of other *LRRK2* pathogenic variant carriers. Combined with published functional evidence showing strongly enhanced LRRK2 kinase activity, our findings provide conclusive evidence that the *LRRK2* p.L1795F variant is pathogenic. It represents a rare cause of PD in the European population but needs to be included in genetic testing efforts and considered for ongoing gene-specific clinical trials.

Introduction

Parkinson's disease (PD) has a complex and multifactorial etiology that includes genetic, environmental, and lifestyle factors, and age¹. The prevalence of monogenic forms of PD, including pathogenic *GBA1* variants, among relatively unselected PD patients is estimated to be ~15 %^{2,3} with variants in the *LRRK2* gene representing one of the most common causes of autosomal dominant PD, although with reduced penetrance. Since its discovery roughly twenty years ago^{4,5}, more than 1,000 different missense variants in the *LRRK2* gene have been identified⁶⁻⁹, only a small fraction of which are considered disease-causing. Evaluating newly identified variants in established PD genes can be challenging, but determining pathogenicity is crucial for diagnosis and even more for treatment, particularly now that *LRRK2*-specific clinical trials are underway^{10,11}. Key aspects of evaluating newly identified genetic variants are evidence of segregation, absence or very low frequencies in controls, support from *in-silico* prediction tools, and evidence from functional studies showing altered protein function^{12,13}.

The *LRRK2* gene encodes the eponymous enzyme Leucine-rich repeat kinase 2 that contains different functional domains, including the N-terminal armadillo, ankyrin, and leucine-rich repeats domains, followed by a C-terminal Roco type GTPase, protein kinase, and WD40 domain. The Roco GTPase domain consists of three subdomains, the ROC GTPase and two scaffolding domains termed COR-A and COR-B^{14,15}. Notably, pathogenic variants in *LRRK2* are thought to cause PD through a gain-of-function mechanism leading to increased kinase activity, which impairs endosomal-lysosomal trafficking, promotes neuroinflammation, and affects ciliogenesis in the striatum¹⁴. Previously, several variants within the interacting ROC, COR-B, and kinase domains have been shown to robustly enhance *LRRK2* kinase activity (defined as >1.5-fold above the wild type), measured as the phosphorylation potential of target proteins such as Rab10, supporting their pathogenic role¹⁵. Amongst those variants was *LRRK2* p.L1795F (chr12:40322386:G:T, hg38), located in the COR-B domain. This variant was reported in a pair of siblings in 2007¹⁶ and two additional singleton cases with PD were identified in 2016 and 2019^{17,18}, nominating it as a possibly causative variant in PD. Moreover, it was recently nominated as a genetic risk factor for PD, with an estimated odds ratio (OR) of 2.5¹⁹. However, the lack of additional reports of variant carriers and evidence of strong segregation precluded this variant from being considered "pathogenic".

In our study, we provide conclusive evidence that the *LRRK2* p.L1795F variant is pathogenic by leveraging genome-wide genotyping and whole-genome sequencing data from the Global Parkinson's Genetics Program (GP2, <http://gp2.org/>), along with additional data from the

Accelerating Medicines Partnership in Parkinson's Disease (AMP-PD), PDGENERation (PDGENE) and the CENTOGENE database.

Methods

Study design and participants

Our study workflow is highlighted in Figure 1. Three sources of data were included in this study (Table 1). First, we used the multi-ancestry whole-genome sequencing and genotyping data from the study participants recruited as part of GP2²⁰ as previously described^{21,22}. Individual-level demographic and clinical data were obtained from participating principal investigators and publicly available databases (e.g., for Coriell samples included in GP2). Second, we incorporated whole-genome sequencing data from AMP-PD. Participants in this initiative were recruited through multiple studies, including BioFIND, the Harvard Biomarkers Study (HBS), the Lewy Body Dementia Case-Control Cohort (LBD), the Parkinson's Disease Biomarkers Program (PDBP), the Parkinson's Progression Markers Initiative (PPMI), the LRRK2 Cohort Consortium (LCC), the Study of Isradipine as a Disease-Modifying Agent in Subjects with Early Parkinson Disease, Phase 3 (STEADY-PD3), and the Study of Urate Elevation in Parkinson's Disease, Phase 3 (SURE-PD3). Clinical information and genetic samples from participants were obtained with appropriate written consent and local institutional and ethical approvals. Detailed information about these studies is available on the AMP-PD website (<https://amp-pd.org>) and the respective study websites. Third, we obtained the clinical exome sequencing data from PDGENE³, a large multi-center study in North America providing genetic testing and counseling to more than 15,000 participants.

Whole-genome sequencing (WGS) data

AMP-PD

We included 9,974 samples with the sequence alignment data available from BioFIND, HBS, LBD, PDBP, PPMI, STEADY-PD3, and SURE-PD3 cohorts through the AMP-PD release for joint genotyping with the GP2 cohort. Due to the unavailability of sequence alignment data from the LCC cohort, we used AMP-PD release 4 data to screen for potential pathogenic variants in this cohort.

GP2

The DNA samples from 5,926 participants were genome sequenced to an average of 30x coverage with 150bp paired-end reads following Illumina's TruSeq PCR-free library

preparation protocol. We followed the same functional equivalence pipeline²³ as AMP-PD to produce the sequence alignment against the GRCh38DH reference genome.

We used DeepVariant v.1.6.1²⁴ to generate the single-sample variant calls for a total of 15,900 samples in GP2 and AMP-PD and performed joint-genotyping using GLnexus v1.4.3 with the preset DeepVariant WGS configuration²⁵. We set genotypes to be missing after variant quality control defined as genotype quality ≥ 20 , read depth ≥ 10 , and heterozygous allele balance between 0.2 and 0.8, and retained high-quality variants with a call rate > 0.95 after quality control. After the sample quality control following the quality metrics defined by AMP-PD²⁶, we retained 15,752 samples (AMP-PD and GP2 combined) for the downstream analyses (Supplementary Table 1). Variant annotation was performed with Ensembl Variant Effect Predictor v111²⁷. We used KING v.2.3.0²⁸ to infer relatedness up to the second-degree relatives to confirm the known relationships and identify cryptic familial relationships. Genetic ancestry was determined using GenoTools v1.2.3 with the default settings²⁹.

Genome-wide genotyping with the Neurobooster Array (GP2)

We screened the genotyping data published as part of GP2's Data Release 7³⁰ (Supplementary Table 2). Genotyping was performed by GP2 using the NeuroBooster Array³¹ (NBA; v.1.0, Illumina, San Diego, CA). Raw genotyping data underwent quality control and genetic ancestry prediction using GenoTools v1.2.3 with the default settings²⁹. The *LRRK2* p.L1795F variant was directly genotyped using NBA, and the quality of genotype calls was assessed by examining the signal intensity plots.

Clinical exome sequencing (PDGENErATION)

We included 9,759 samples with clinical exome data available from PDGENE³. The sequence data processing followed the same pipeline of WGS data as mentioned above. We performed joint-genotyping using GLnexus v1.4.3 with the preset DeepVariant WES configuration and followed the same criteria for sample and variant quality control as for the WGS data.

Querying additional databases (CENTOGENE)

We queried the CENTOGENE proprietary Databank CentoMD®³² to identify potential additional variant carriers. CENTOGENE is a globally operating genetic diagnostic lab. Genetic data included in this manuscript was generated by exon-wise PCR amplification followed by Sanger sequencing.

Statistical analyses

To estimate the allele frequency of *LRRK2* p.L1795F variant in multi-ancestral populations, we analyzed the GP2 genotyping data, the largest available dataset in this study. We excluded related individuals and samples from targeted recruitment, such as *LRRK2* and *GBA1* variant carriers within specific efforts of PPMI and LCC. Subsequently, we performed an association analysis of this variant with PD using the European population. We fitted the logistic regression model with PD status as binary outcome variable and the covariates as the genotype of *LRRK2* p.L1795F variant, sex, age, family history, and the first six principal components to account for the population stratification. For cases, age at onset (AAO) or age at diagnosis was used, while for controls, age at sampling was used. Additionally, we merged GP2 genotyping data with the combined AMP-PD and GP2 WGS data, resulting in a cohort of 23,276 PD cases of European ancestry after excluding duplicated, related, and targeted recruitment samples as mentioned above. This allowed us to compare the carrier distribution between PD cases and non-Finnish European population from the Genome Aggregation Database (gnomAD v4.1) as external population controls using Fisher's exact test. We excluded the PDGENE clinical exome data from this analysis as we could not estimate the genetic ancestry in the same manner as with the other datasets. The *P* value ≤ 0.05 was considered statistically significant for all the analyses.

To determine if carriers of the *LRRK2* p.L1795F variant shared recent common ancestry, we phased the genotyping data from chromosome 12 in the European population using Beagle 5.4 with default settings³³ and searched for identical-by-descent (IBD) segments with the length ≥ 2 cM shared across the carriers using hap-ibd with default setting³⁴.

Results

Identification of the *LRRK2* p.L1795F variant segregating with disease in a family

Our discovery cohort consisted of 16,351 individuals from GP2 and AMP-PD with WGS data, including 15,752 samples from the joint-genotyping sample set and 599 samples from the LCC cohort from AMP-PD release 4 (Table 1). Searching for recurrent rare variants, we identified nine carriers of the *LRRK2* p.L1795F variant (ENST00000298910:c.5385G>T; chr12:40322386:G:T; Supplementary Figures 1-6). Of these carriers, we identified two families based on kinship inference (Figure 2). The larger family (GP2-FAM-1) consisted of four affected individuals showing segregation of this variant with PD. The second family (AMP-

FAM-1) consisted of three variant carriers, including one clinically affected with PD, while the other two were reported as asymptomatic at age 55 and 76 years. The remaining two carriers were singleton cases with familial PD. Next, we screened the genotyping data of 54,153 affected and unaffected individuals generated within GP2 and identified three additional variant carriers, all clinically affected with PD (Supplementary Figure 7). Further, screening the clinical exome data from 9,759 individuals available from PDGENE resulted in one additional variant carrier (Supplementary Figure 8). Finally, querying the CENTOGENE proprietary Databank CentoMD®³², we identified another family with four individuals carrying the *LRRK2* p.L1795F variant, three of whom were clinically affected with PD and one being an asymptomatic carrier. In total, we identified 17 individuals carrying this variant across all the datasets, of which 14 had PD and three were asymptomatic.

Evaluation of the *LRRK2* p.L1795F variant

The p.L1795F (c.5385G>T) missense variant is rare and confined to European populations in several investigated databases, including gnomAD v4.1 ($MAF_{\text{Europeans(non-Finnish)}}=0.000001695$) and the Regeneron Genetics Center Million Exome Variant Browser³⁵ (RGC-ME, $MAF_{\text{Europeans}}=0.000009515$). In comparison, it was not present in the UK Biobank³⁶ 500K genomes. Evaluation using various *in-silico* prediction tools and databases presents conflicting results. ClinVar, Varsome, and Franklin (the latter two based on the ACMG criteria¹²) categorize this variant as a variant of uncertain significance. Notably, Varsome and Franklin do not take the existing functional evidence¹⁵ into account. Furthermore, this variant is currently not included in the list of genetic variants reported to the clinician or the participant in PDGENE³. MutationTaster predicts the variant to be disease-causing, and the leucine at position 1795 is conserved across different species (Supplementary Figure 9). In contrast, other *in-silico* tools for predicting missense pathogenicity did not support pathogenicity following ACMG recommended thresholds³⁷, including CADD (19.94) and REVEL (0.638), with the exception of VEST4 (0.928). Finally, this variant is located in the COR-B subdomain of the C-terminal Roco GTPase domain and has previously been shown to strongly increase *LRRK2* kinase activity¹⁵.

Allele frequency of *LRRK2* p.L1795F in multi-ancestral populations and the founder effect

All identified *LRRK2* p.L1795F carriers in this study were of European ancestry, whereas the variant was absent in other ancestral populations (n=15,316) within the GP2 genotyping cohort. It had an allele frequency of 0.00012 among PD cases (5 heterozygous carriers and 20,812 noncarriers) while being absent in controls (n=9,032) in the European population of

the GP2 genotyping cohort (Table 2). The logistic regression analysis using the European population of the GP2 genotyping cohort did not reveal a significant association between this variant and PD, possibly due to insufficient controls available in the dataset ($P>0.8$, Supplementary Table 3). However, when comparing the carrier distribution between PD cases from the combined genotyping and WGS dataset (6 heterozygous carriers and 23,270 noncarriers) with the non-Finnish European population from Genome Aggregation Database (gnomAD v4.1) as external population controls (2 heterozygous carriers and 589,826 noncarriers), this variant showed a significant association with PD ($P<7.84e-08$, two-tailed, Fisher's exact test).

Given this variant was observed only in the European population from the GP2 genotyping cohort, we searched for the overlapping IBD segments among the variant carriers using the genotyping data (Figure 3). The median length of an IBD segment over *LRRK2* in these individuals was 7.05 cM (range: 2.1-96.3 cM, Supplementary Table 4). All genotyped carriers shared a core haplotype of 2.825 Mbp at this locus (Supplementary Table 5), suggesting that the p.L1795F variant descended from a common founder.

Clinical features of identified *LRRK2* p.L1795F variant carriers

The demographic and clinical details of all 17 identified variant carriers of White/European ancestry, including 14 affected and three unaffected individuals, are displayed in Table 3. More than two-thirds were females (70.6 %; $n=12/17$). All affected and unaffected carriers had a positive family history of PD. Ages of motor symptom onset in affected individuals ranged from 36 to 66 years. The median AAO was 54.5 years (interquartile range 47-60 years). The asymptomatic carriers were 55, 76 and 76 years old, respectively, at the time of sample collection. Based on the available clinical data, the majority of affected individuals had classical PD with an asymmetric onset of symptoms and a good response to dopaminergic medication, and without obvious atypical signs suggestive of other diagnoses (missing data for up to 30%); detailed data on non-motor symptoms and neuropsychiatric comorbidities were scarce. Cognition was reported to be unaffected in the majority of affected carriers with good scores in cognition tests (including Montreal Cognitive Assessment [MoCA] and Mini Mental State Examination [MMSE]); however, one clinically affected individual had significant cognitive impairment (MoCA score of 17/30 points) and one unaffected carrier also showed some cognitive deficits (MoCA score of 23/30 points). The characteristics of the individuals from the three identified families will be reported in more detail.

Family GP2-FAM-1

Family GP2-FAM-1 is of European ancestry with Ukrainian and Polish origin. Seven individuals are known to be clinically affected by PD, including the index patient (GP2-ID-3), his sister, mother, three maternal aunts, and a maternal cousin, consistent with autosomal dominant inheritance (Figure 2). Further, additional maternal aunts and uncles were reported to have PD but a detailed history was not available. We identified the *LRRK2* p.L1795F variant segregating within all four tested family members from both NBA and WGS data. No unaffected family samples were available. Screening variants segregating within this family from the WGS data did not reveal any other potential causal variants, including known pathogenic variants in the established dominant PD genes *SNCA* and *VPS35* as well as other variants in *LRRK2* and pathogenic *GBA1* variants.

All family members with available data were reported to have bradykinesia, rigidity, resting and action tremor, and motor symptoms that were responsive to dopaminergic treatment. Disease progression was mild to moderate in three of four individuals with low to moderate UPDRS (part III) motor scores and a Hoehn & Yahr stage 2 after 9+ years of disease duration. Only one individual, GP2-ID-1 (deceased), seemed to have had a more progressive disease course with a high UPDRS (part III) motor score and Hoehn & Yahr stage 5, though over a disease duration of more than 20 years. Neuropsychiatric comorbidities or severe autonomic features were not reported in those with available data. Cognition was unaffected in all family members. All but one affected family member, including those without genetic testing, had an AAO in their fifties (ranging from 50 to 59 years), and only one individual had a lower AAO of 40 years. Videos of individuals GP3-ID-3 and GP2-ID-4 are available in the supplementary materials.

Family AMP-FAM-1

Family AMP-FAM-1 included three individuals available for genetic testing, all of whom carried the p.L1795F variant. The index case was clinically affected by PD, while her sister and mother were both asymptomatic. The family history of PD was strongly positive, with multiple additional affected family members, including two maternal aunts and the maternal grandfather of the index, suggesting autosomal dominant inheritance with reduced penetrance (Figure 3). The index patient had a reported AAO of 46 years and a very low UPDRS (part III) motor score, indicating a rather mild disease course. Clinical details for the additional affected family members were unavailable. The two asymptomatic carriers were 55 and 76 years old at sample collection and showed no signs of PD. We did not identify other potential disease-causing variants in this family by WGS.

Family TORONTO-FAM-1

Family TORONTO-FAM-1 included four individuals available for genetic testing, and all four carried the *LRRK2* p.L1795F variant. The index case as well as her sister and a maternal uncle were clinically affected with PD whereas the mother of both siblings was an unaffected carrier. There were additional family members clinically affected with PD, including another maternal uncle and the maternal grandfather, both of which were unavailable for genetic testing within this study (Figure 2).

The reported AAO of the index case was 44 years and thereby younger than for the other two tested family members, which were 65 and 66, respectively. The index patient and her sister have been followed up for almost 12 years, whereas the other two individuals (TORONTO-ID-1 and TORONTO-ID-2) were only clinically assessed once in 2012. Notably, the index patient had a more progressive PD disease course than her sister, as indicated by a higher UPDRS motor score of 43 points and Hoehn & Yahr stage 3, compared to only 7 points in the UPDRS (part III) in her sister. All three affected individuals had a diagnosis of classical PD without any atypical features; however, one individual (TORONTO-ID-1) had significant cognitive impairment with a low MoCA score of only 17 out of 30 points. Interestingly, also the unaffected carrier showed some cognitive impairment (MoCA score of 23/30 points) but no motor symptoms of PD.

Discussion

Our study was carried out under the umbrella of GP2, a large international collaborative effort aimed at better understanding the genetic architecture of PD at a global scale by generating large-scale genetic data from diverse ancestries. Additionally, we leveraged data from AMP-PD and PDGENE and queried the CENTOGENE database. To investigate monogenic causes of the disease, we screened WGS data from our discovery cohort for recurrent rare variants. We identified the *LRRK2* p.L1795F variant segregating with the disease in four members of a large European ancestry family, with multiple additional affected family members not available for genetic analyses. Moreover, the variant was identified in a second family with one affected and two asymptomatic carriers, alongside multiple affected family members not available for genetic analyses. Further, we identified four affected carriers by analyzing additional datasets, including NBA genotyping data from GP2 and clinical exome data from PDGENE. Finally, we identified another family with 3 affected individuals and one unaffected carrier by querying the CENTOGENE database. All identified variant carriers in this study were of non-Ashkenazi Jewish, non-Finnish European ancestry and had a strongly positive family history with at least one but more often multiple additional affected family members. A previous rare-variant association analysis further supports the role of p.L1795F in PD pathogenesis, identifying it

as a genetic risk factor with an estimated OR of 2.5¹⁹. However, it should be noted that the number of identified carriers was quite small, likely resulting in an underestimation of the actual OR. Most importantly, our findings provide family segregation evidence missing from the previous reports^{16–18}. Furthermore, the previously reported increase in kinase activity of this variant aligns with the disease mechanism established for several pathogenic *LRRK2* variants¹⁵. When applying the ACMG criteria¹² in light of our findings, the variant can now be classified as pathogenic based on: i) the very low frequency in population databases (PM2), ii) established functional studies supporting a damaging effect consistent with the established disease mechanism (PS3), iii) observation of the variant in multiple unrelated individuals with the same phenotype (no specific criterion, may be considered as moderate evidence), and iv) strong evidence of segregation (at least PP1, based on our findings upgraded to strong evidence by segregation in three families with two generations of family members each). Taken together, we thereby propose the *LRRK2* p.L1795F variant to be considered pathogenic and causative of PD.

Interestingly, the *LRRK2* p.L1795F variant had an estimated allele frequency of 8.37×10^{-5} (5 observations in 59,698 alleles) in the European population and was absent from all other ancestral populations in the GP2 genotyping cohort. This finding was consistent with several public frequency databases, such as gnomAD v4.1 and RGC-ME, contrasting with the *LRRK2* p.G2019S variant. Globally, the *LRRK2* p.G2019S variant is the most common and well-studied genetic cause of PD. Due to independent founder effects^{38–40}, the highest frequencies of this variant were observed in the Ashkenazi Jewish population⁴¹, ranging from 10% in sporadic to 26% in familial PD, and Arab-Berber populations, ranging from 30% in sporadic to 41% in familial PD⁴². The variant was also commonly reported in individuals of Portuguese, Brazilian, Spanish, and Italian ancestry but is much rarer in individuals of other European, Asian, or Indian descent⁴³. In comparison, only four individuals carrying the p.L1795F were reported^{16–18}, and two additional carriers were identified through AMP-PD^{6,19}. To our knowledge, we provide the largest number of p.L1795F variant carriers thus far, including 14 carriers clinically affected with PD and three asymptomatic carriers. The available data on AAO and family pedigrees from these previously reported carriers^{16–18} do not align with our data, making an overlap of individuals between the different studies unlikely. Including those reported in the literature, this brings the total to 18 clinically affected carriers of European ancestry. In our GP2 genotyping cohort, the observed allele frequency of the p.L1795F variant among affected European individuals (n=20,817) was 0.00012, while the allele frequency for the p.G2019S variant was 0.003266 (Table 2). This indicates that p.G2019S is a more common cause of PD in the European population compared to p.L1795F. However, we

acknowledge that the overall number of p.L1795F carriers is still limited, and higher frequencies might be observed in specific European subpopulations. Our haplotype analysis further supports this hypothesis, where all genotyped carriers shared a core haplotype of 2.83 Mbp. We were able to determine the geographical origin of only one family of carriers in this study, which was of Ukrainian and Polish descent. Additionally, this variant is more prevalent in Northern Europe according to ancestry estimates of the carriers from RGC-ME but was not found in the 500K genomes of the UK Biobank. Consequently, investigating the *LRRK2* p.L1795F variant within the Central-Eastern European population could offer additional insights into a possible founder event.

Comparing the clinical phenotypes of the p.L1795F carriers with those of other pathogenic *LRRK2* variants, particularly p.G2019S¹³, revealed similarities among them and with idiopathic PD (iPD). *LRRK2*-PD is clinically indistinguishable from iPD on an individual level. Most individuals with *LRRK2*-PD, including p.L1795F carriers, exhibit a classic PD phenotype with asymmetric disease onset and display all the cardinal motor signs of PD with a good response to dopaminergic treatment. Atypical presentations have been described in single cases but are overall rare⁴³. Furthermore, the AAO was comparable between p.L1795F carriers and other *LRRK2*-PD genetic subtypes. Most individuals exhibited first motor symptoms in their 50s and 60s (53% in *LRRK2*-PD overall¹³ and 70 % of all known p.L1795F carriers including our study). However, a broader range of age at onset has been described, spanning from 20 to 95 years for *LRRK2*-PD overall¹³ and from 25 to 66 years for p.L1795F carriers. Non-motor features and neuropsychiatric comorbidities haven't been specifically reported for the majority of p.L1795F carriers, but the overall data is limited, making it difficult to draw meaningful conclusions. While group differences in clinical phenotypes among *LRRK2* variants may exist⁴³, they do not enable meaningful genotype-phenotype correlations at an individual level. Overall, the p.L1795F phenotype aligns well with the general characteristics of *LRRK2*-PD and appears comparable to other *LRRK2* variants with cautious interpretation given the limited number of identified carriers. The most significant differences between the genetic subtypes are their ancestral and geographical variability.

Notably, we identified three asymptomatic p.L1795F carriers from two different families who might still develop PD symptoms later in life. However, this seems unlikely for at least two individuals, who were 76 years old at the most recent follow-up, given that the oldest reported age at onset (AAO) for the affected p.L1795F carriers is currently 66 years¹⁶. Alternatively, reduced penetrance, a common phenomenon in monogenic forms of PD, including other pathogenic *LRRK2* variants, might explain the finding. All three asymptomatic p.L1795F carriers were first-degree relatives of an affected carrier. Additionally, several other family

members with PD were reported in these families, suggesting that these affected members might also carry the same variant. However, they were not available for genetic testing in this study. *LRRK2* penetrance depends on age⁴⁴, environmental and lifestyle factors⁴⁵, ancestral background^{46,47}, and the specific variant as well as additional genetic modifiers⁴⁸⁻⁵⁰. For example, penetrance of the most common p.G2019S variant is estimated at around 25-30% in the Ashkenazi Jewish population and up to 42% in non-Jewish individuals by the age of 80 years, and 45% in the North African Berber population over their life course⁴³. However, the current data on *LRRK2* p.L1795F is still limited, and the number of tested affected and unaffected family members is too low to estimate the penetrance for this variant accurately.

In conclusion, this is the first study providing evidence of the *LRRK2* p.L1795F variant segregating with disease in large multiplex families. Taken together with published functional data, showing strongly enhanced LRRK2 kinase activity, our findings support the *LRRK2* p.L1795F variant to be considered pathogenic. Our study demonstrates that large-scale studies can be helpful to identify novel rare causes of PD but also to re-evaluate previously identified variants by providing additional evidence of pathogenicity through an increased number of variant carriers and segregation. We, therefore, propose *LRRK2* p.L1795F as a cause of PD, especially in the European population. Including this variant in the genetic screening of PD patients may be beneficial for the variant carriers to be included in ongoing gene-specific clinical trials.

Data availability

GP2 partnered with the online cloud computing platform Accelerating Medicines Partnership - Parkinson's Disease (AMP PD; <https://amp-pd.org>) to share data generated by GP2. Anonymized data can be shared upon request and qualified researchers are encouraged to apply for direct access to the data through AMP PD.

References

1. Wüllner, U. *et al.* The heterogeneity of Parkinson's disease. *J. Neural Transm.* **130**, 827–838 (2023).
2. Westenberger, A. *et al.* Relevance of genetic testing in the gene-targeted trial era: The Rostock Parkinson's Disease Study. *Brain* (2024, in press).
3. Cook, L. *et al.* Parkinson's Disease Variant Detection and Disclosure: PD GENERation, a North American Study. *Brain* (2024, in press).
4. Zimprich, A. *et al.* Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **44**, 601–607 (2004).
5. Paisán-Ruiz, C. *et al.* Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron* **44**, 595–600 (2004).
6. Bryant, N. *et al.* Identification of LRRK2 missense variants in the accelerating medicines partnership Parkinson's disease cohort. *Hum. Mol. Genet.* **30**, 454–466 (2021).
7. Ross, O. A. *et al.* Association of LRRK2 exonic variants with susceptibility to Parkinson's disease: a case-control study. *Lancet Neurol.* **10**, 898–908 (2011).
8. Tratzuni, D. *et al.* Fine-mapping, gene expression and splicing analysis of the disease associated LRRK2 locus. *PLoS One* **8**, e70724 (2013).
9. Chen, S. *et al.* A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **625**, 92–100 (2024).
10. McFarthing, K. *et al.* Parkinson's Disease Drug Therapies in the Clinical Trial Pipeline: 2022 Update. *J. Parkinsons. Dis.* **12**, 1073–1082 (2022).
11. Kluss, J. H., Lewis, P. A. & Greggio, E. Leucine-rich repeat kinase 2 (LRRK2): an update on the potential therapeutic target for Parkinson's disease. *Expert Opin. Ther. Targets* **26**, 537–546 (2022).
12. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).
13. MDSGene. *Movement Disorder Society Genetic mutation database (MDSGene)* <https://www.mdsgene.org/>.
14. Taylor, M. & Alessi, D. R. Advances in elucidating the function of leucine-rich repeat protein kinase-2 in normal cells and Parkinson's disease. *Curr. Opin. Cell Biol.* **63**, 102–113 (2020).
15. Kalogeropoulou, A. F. *et al.* Impact of 100 LRRK2 variants linked to Parkinson's disease on kinase activity and microtubule binding. *Biochem. J* **479**, 1759–1783 (2022).
16. Nichols, W. C. *et al.* LRRK2 mutation analysis in Parkinson disease families with evidence of linkage to PARK8. *Neurology* **69**, 1737–1744 (2007).
17. Benitez, B. A. *et al.* Resequencing analysis of five Mendelian genes and the top genes from genome-wide association studies in Parkinson's Disease. *Mol. Neurodegener.* **11**, 29 (2016).
18. Illés, A. *et al.* The Role of Genetic Testing in the Clinical Practice and Research of Early-Onset Parkinsonian Disorders in a Hungarian Cohort: Increasing Challenge in Genetic Counselling, Improving Chances in Stratification for Clinical Trials. *Front. Genet.* **10**, 1061 (2019).
19. Pitz, V. *et al.* Analysis of rare Parkinson's disease variants in millions of people. *NPJ Parkinsons Dis* **10**, 11 (2024).
20. Global Parkinson's Genetics Program. GP2: The Global Parkinson's Genetics Program. *Mov. Disord.* **36**, 842–851 (2021).
21. Lange, L. M. *et al.* Elucidating causative gene variants in hereditary Parkinson's disease in the Global Parkinson's Genetics Program (GP2). *NPJ Parkinsons Dis* **9**, 100 (2023).
22. Towns, C. *et al.* Defining the causes of sporadic Parkinson's disease in the global Parkinson's genetics program (GP2). *NPJ Parkinsons Dis* **9**, 131 (2023).
23. Regier, A. A. *et al.* Functional equivalence of genome sequencing analysis pipelines enables harmonized variant calling across human genetics projects. *Nat. Commun.* **9**,

- 1–8 (2018).
24. Poplin, R. *et al.* A universal SNP and small-indel variant caller using deep neural networks. *Nat. Biotechnol.* **36**, 983–987 (2018).
 25. Yun, T. *et al.* Accurate, scalable cohort variant calls using DeepVariant and GLnexus. *Bioinformatics* **36**, 5582–5589 (2021).
 26. Iwaki, H. *et al.* Accelerating Medicines Partnership: Parkinson’s Disease. Genetic Resource. *Mov. Disord.* **36**, 1795–1804 (2021).
 27. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17:122**, (2016).
 28. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
 29. Vitale, D. *et al.* GenoTools: An Open-Source Python Package for Efficient Genotype Data Quality Control and Analysis. *bioRxiv* (2024) doi:10.1101/2024.03.26.586362.
 30. Leonard, H. *et al.* Global Parkinson’s Genetics Program data release 7. Zenodo <https://doi.org/10.5281/ZENODO.10962119> (2024).
 31. Bandres-Ciga, S. *et al.* NeuroBooster Array: A Genome-Wide Genotyping Platform to Study Neurological Disorders Across Diverse Populations. *medRxiv* (2023) doi:10.1101/2023.11.06.23298176.
 32. Trujillano, D. *et al.* A comprehensive global genotype-phenotype database for rare diseases. *Mol Genet Genomic Med* **5**, 66–75 (2017).
 33. Browning, B. L., Tian, X., Zhou, Y. & Browning, S. R. Fast two-stage phasing of large-scale sequence data. *Am. J. Hum. Genet.* **108**, 1880–1890 (2021).
 34. Zhou, Y., Browning, S. R. & Browning, B. L. A Fast and Simple Method for Detecting Identity-by-Descent Segments in Large-Scale Data. *Am. J. Hum. Genet.* **106**, 426–437 (2020).
 35. Sun, K. Y. *et al.* A deep catalog of protein-coding variation in 985,830 individuals. *bioRxiv* (2023) doi:10.1101/2023.05.09.539329.
 36. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
 37. Pejaver, V. *et al.* Calibration of computational tools for missense variant pathogenicity classification and ClinGen recommendations for PP3/BP4 criteria. *Am. J. Hum. Genet.* **109**, 2163–2177 (2022).
 38. Lesage, S. *et al.* Parkinson’s disease-related LRRK2 G2019S mutation results from independent mutational events in humans. *Hum. Mol. Genet.* **19**, 1998–2004 (2010).
 39. Bar-Shira, A., Hutter, C. M., Giladi, N., Zabetian, C. P. & Orr-Urtreger, A. Ashkenazi Parkinson’s disease patients with the LRRK2 G2019S mutation share a common founder dating from the second to fifth centuries. *Neurogenetics* **10**, 355–358 (2009).
 40. Zabetian, C. P. *et al.* LRRK2 G2019S in families with Parkinson disease who originated from Europe and the Middle East: evidence of two distinct founding events beginning two millennia ago. *Am. J. Hum. Genet.* **79**, 752–758 (2006).
 41. Ozelius, L. J. *et al.* LRRK2 G2019S as a cause of Parkinson’s disease in Ashkenazi Jews. *N. Engl. J. Med.* **354**, 424–425 (2006).
 42. Hulihan, M. M. *et al.* LRRK2 Gly2019Ser penetrance in Arab-Berber patients from Tunisia: a case-control genetic study. *Lancet Neurol.* **7**, 591–594 (2008).
 43. Saunders-Pullman, R., Raymond, D. & Elango, S. LRRK2 Parkinson Disease. in *GeneReviews®* (eds. Adam, M. P. *et al.*) (University of Washington, Seattle, Seattle (WA), 2006).
 44. Marder, K. *et al.* Age-specific penetrance of LRRK2 G2019S in the Michael J. Fox Ashkenazi Jewish LRRK2 Consortium. *Neurology* **85**, 89–95 (2015).
 45. Lüth, T. *et al.* Age at Onset of LRRK2 p.Gly2019Ser Is Related to Environmental and Lifestyle Factors. *Mov. Disord.* **35**, 1854–1858 (2020).
 46. Lee, A. J. *et al.* Penetrance estimate of LRRK2 p.G2019S mutation in individuals of non-Ashkenazi Jewish ancestry. *Mov. Disord.* **32**, 1432–1438 (2017).
 47. Hentati, F. *et al.* LRRK2 parkinsonism in Tunisia and Norway: a comparative analysis of disease penetrance. *Neurology* **83**, 568–569 (2014).

48. Trinh, J. *et al.* DNMT3 and genetic modifiers of age of onset in LRRK2 Gly2019Ser parkinsonism: a genome-wide linkage and association study. *Lancet Neurol.* **15**, 1248–1256 (2016).
49. Lai, D. *et al.* Genomewide Association Studies of LRRK2 Modifiers of Parkinson's Disease. *Ann. Neurol.* **90**, 76–88 (2021).
50. Kmieciak, M. J. *et al.* Genetic analysis and natural history of Parkinson's disease due to the LRRK2 G2019S variant. *Brain* **147**, 1996–2008 (2024).

Acknowledgment

Genotyping data (DOI 10.5281/zenodo.10962119, release 7) and whole-genome sequencing data used in the preparation of this article were obtained from Global Parkinson's Genetics Program (GP2). GP2 is funded by the Aligning Science Against Parkinson's (ASAP) Initiative and implemented by The Michael J. Fox Foundation for Parkinson's Research (<https://www.gp2.org>). For a complete list of GP2 members see <http://www.gp2.org>. Whole-genome sequencing data used in the preparation of this article were obtained from the Accelerating Medicine Partnership® (AMP®) Parkinson's Disease (AMP PD) Knowledge Platform. For up-to-date information on the study, visit <https://www.amp-pd.org>. The AMP® PD program is a public-private partnership managed by the Foundation for the National Institutes of Health and funded by the National Institute of Neurological Disorders and Stroke (NINDS) in partnership with the Aligning Science Across Parkinson's (ASAP) initiative; Celgene Corporation, a subsidiary of Bristol-Myers Squibb Company; GlaxoSmithKline plc (GSK); The Michael J. Fox Foundation for Parkinson's Research ; Pfizer Inc.; AbbVie Inc.; Sanofi US Services Inc.; and Verily Life Sciences. ACCELERATING MEDICINES PARTNERSHIP and AMP are registered service marks of the U.S. Department of Health and Human Services. Clinical data used in preparation of this article were obtained from the MJFF-sponsored LRRK2 Cohort Consortium (LCC). For up-to-date information on the study, visit www.michaeljfox.org/lcc. The LRRK2 Cohort Consortium is coordinated and funded by The Michael J. Fox Foundation for Parkinson's Research. The investigators within the LCC provided data, but did not participate in the analysis or writing of this report. The full list of LCC investigators can be found at www.michaeljfox.org/lccinvestigators. PD GENERation is a study funded by the Parkinson's Foundation and supported by GP2, a program of Aligning Science Across Parkinson's (ASAP). Variant queries in the UKBB were conducted under approved project 82590 (to Z-H.F). Additional data used for this article was obtained from PD GENERation. PD GENERation is a study funded by the Parkinson's Foundation and supported by GP2, a program of Aligning Science Across Parkinson's (ASAP). This work was supported in part by the Intramural Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services, project ZIA AG000949.

Author contributions

L.M.L. and Z-H.F. were responsible for the study conceptualization and execution. They analyzed and interpreted the generated genotyping, clinical-exome, and whole-genome sequencing data and wrote the first draft of the manuscript. L.M.L. analyzed and interpreted the clinical data. Z-H.F. performed sequencing data processing. H.I., J.M., N.K., K.Le., D.V., H.L., M.N., and C.B. were involved in sample and genotyping data acquisition and access to

raw data. H.I., H.R.M., and C.K. contributed to clinical data collection and analysis. L.M., A.E., and H.C. contributed samples from affected individuals to GP2 that were identified to carry the *LRRK2* variant and their respective demographic and clinical data included in this manuscript. S.A.F. and L.A.H. contributed clinical and genetic data generated as part of the PDGENERation study. S.F., N.A., and C.M. contributed clinical data for individuals included in this study. P.B. and C.B. contributed genetic data generated by CENTOGENE GmbH. All co-authors were involved in data interpretation, read and approved the final version of the manuscript.

Competing interests

L.M.L., N.A., K. Lo., and L.A.H. declare no competing interests.

C.M. receives research funding from the Michael J Fox Foundation, the Parkinson's Foundation (US) and holds the Catherine Manson Chair in Movement Disorders, funded by the Mayvon Foundation. D.V., H.L.L., H.I., K.Le., 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 plus is a scientific founder at Neuron23 Inc and owns stock. L.M. has received honoraria from the International Association of Parkinsonism and Related Disorders (IAPRD) Society for social media and web support, and personal compensation as a consultant/scientific advisory board member for Acadia. He has received a grant (collaborative research agreement) from the International Parkinson and Movement Disorders Society for the MDS-UTRS Validation Program (Role: PI), Non-Profit. A.J.E. has received grant support from the NIH and the Michael J Fox Foundation; personal compensation as a consultant/scientific advisory board member for Neuroderm, Amneal, Acadia, Avion Pharmaceuticals, Acorda, Kyowa Kirin, Supernus (formerly, USWorldMeds), NeuroDiagnostics, Inc (SYNAPS Dx), Intrance Medical Systems, Inc., Praxis Precision Medicines, and Herantis Pharma; and publishing royalties from Lippincott Williams & Wilkins, Cambridge University Press, and Springer. He co-founded REGAIN Therapeutics and is co-inventor of the patent "Compositions and methods for treatment and/or prophylaxis of proteinopathies. P.B. and C.B. are employees of CENTOGENE GmbH. S.A.F. received honoraria from Lundbeck, Biogen, Takeda, and Neurocrine and grants from Medtronic, Boston Scientific, Sun Pharmaceuticals Advanced Research Company, Aspen, Biohaven, Neurocrine, Voyager, Prilenia Therapeutics, CHDI Foundation, Michael J. Fox Foundation, NIH 1 P50 NS123103-01, NIH 1R01NS125294-01, and the Parkinson Foundation. Finally, he reports royalties from Demos, Blackwell Futura, Springer for textbooks, and Uptodate. N.E.M. receives salary and research support from the NIH (1K08NS131581), the Parkinson's Foundation and the Aligning Science Across Parkinson's (ASAP) Global Parkinson's Genetics Program (GP2). He serves as a member of the PDGENERation steering committee.

H.R.M. is employed by UCL. In the last months, he reports paid consultancies from Roche, Aprinolia, AI Therapeutics and Amylyx; lecture fees/honoraria from BMJ, Kyowa Kirin, and the Movement Disorders Society; and research Grants from Parkinson's UK, Cure Parkinson's Trust, PSP Association, Medical Research Council, and the Michael J Fox Foundation. H.R.M. is also a co-applicant on a patent application related to C9ORF72 - Method for diagnosing a neurodegenerative disease (PCT/GB2012/052140). C.K. has received grant support from The Michael J. Fox Foundation for Parkinson's Research and the Aligning Science Across Parkinson's Initiative. She serves as a medical advisor to Centogene, Retromer Therapeutics, and Takeda and received speakers' honoraria from Bial and Desitin. Z-H.F is supported by the Aligning Science Across Parkinson's (ASAP) Global Parkinson's Genetics Program (GP2) and receives GP2 salary support from The Michael J. Fox Foundation for Parkinson's Research.

Figures

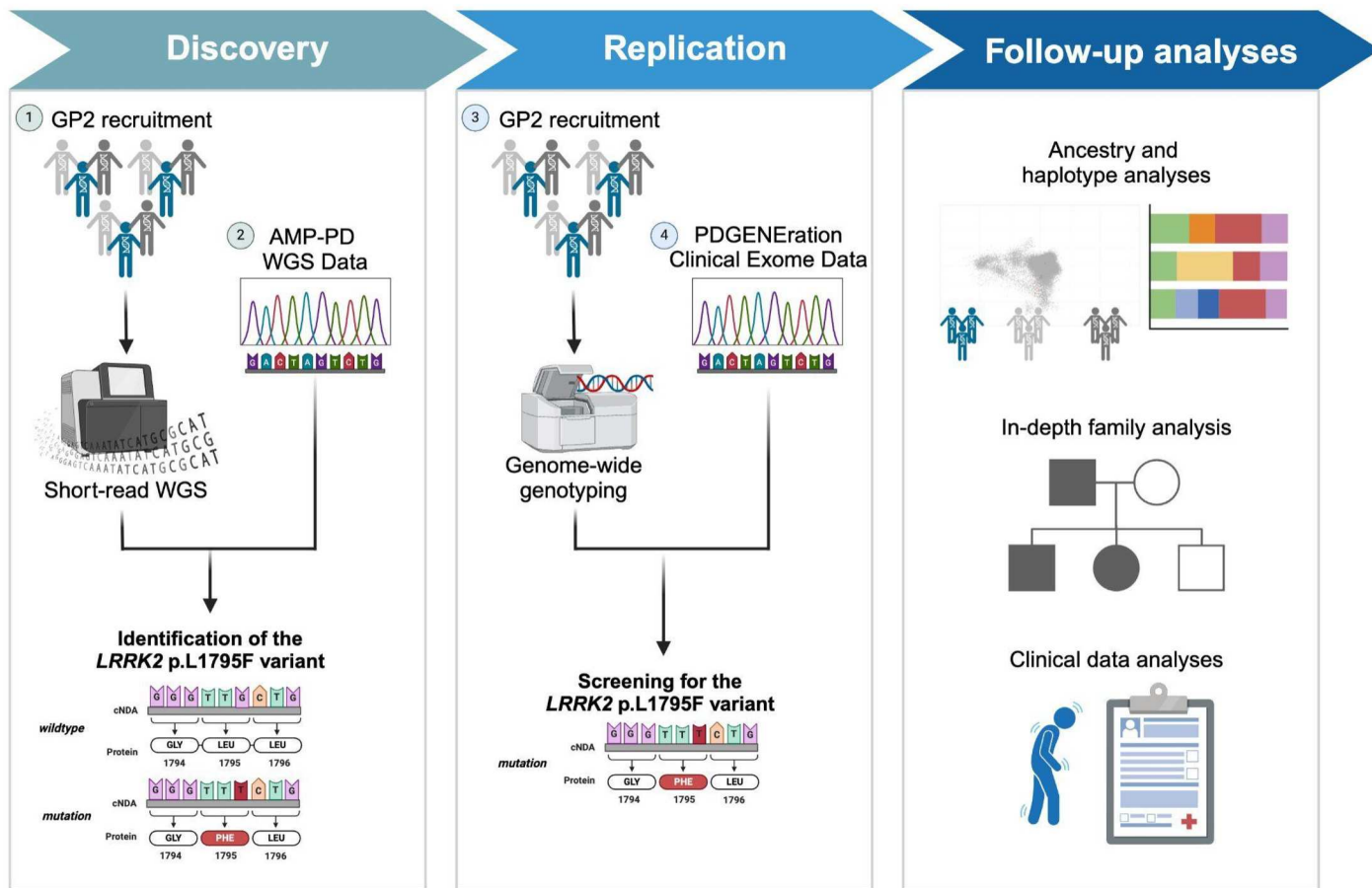


Figure 1. Study design and workflow. Figure created with BioRender.com.

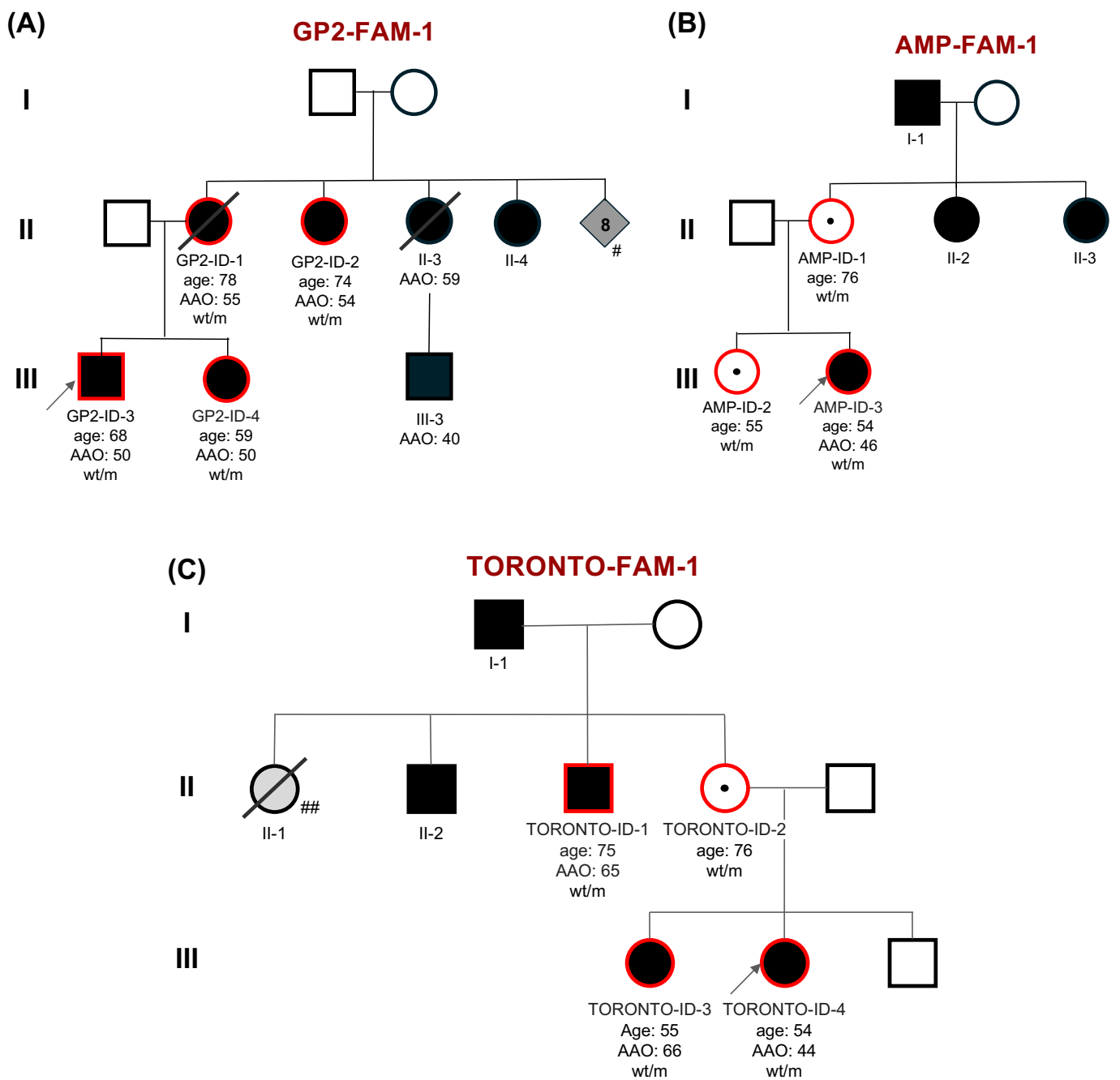


Figure 2. Pedigree of Family GP2-FAM-1 (A), AMP-FAM-1 (B), and TORONTO-FAM-1 (C) with the *LRRK2* p.L1795F variant. The pedigrees were drawn based on reported family history and may be incomplete. The index cases are indicated with arrows. Affected individuals are indicated by black symbols: circles (female) and squares (male). Diamond is where sex is undefined. Unaffected individuals are indicated by open symbols. Unaffected variant carriers are indicated by open symbols with a dot in the middle. A diagonal line indicates deceased individuals. Red circle indicates individuals with genetic data available (WGS data for GP2-FAM-1 and AMP-FAM-1, single gene testing for TORONTO-FAM-1). Heterozygous mutant (m) and wild-type (wt) genotypes are indicated with corresponding age at the sample collection (age) and age at motor symptom onset (if known; AAO). (A) The mother of GP2-FAM-1 index was reported to have eight additional siblings (#), several of whom are clinically affected with PD; however, no detailed family history is available for these relatives. (C) One maternal aunt (II-1) of the TORONTO-FAM-1 index was reported to have had Alzheimer's disease (##).

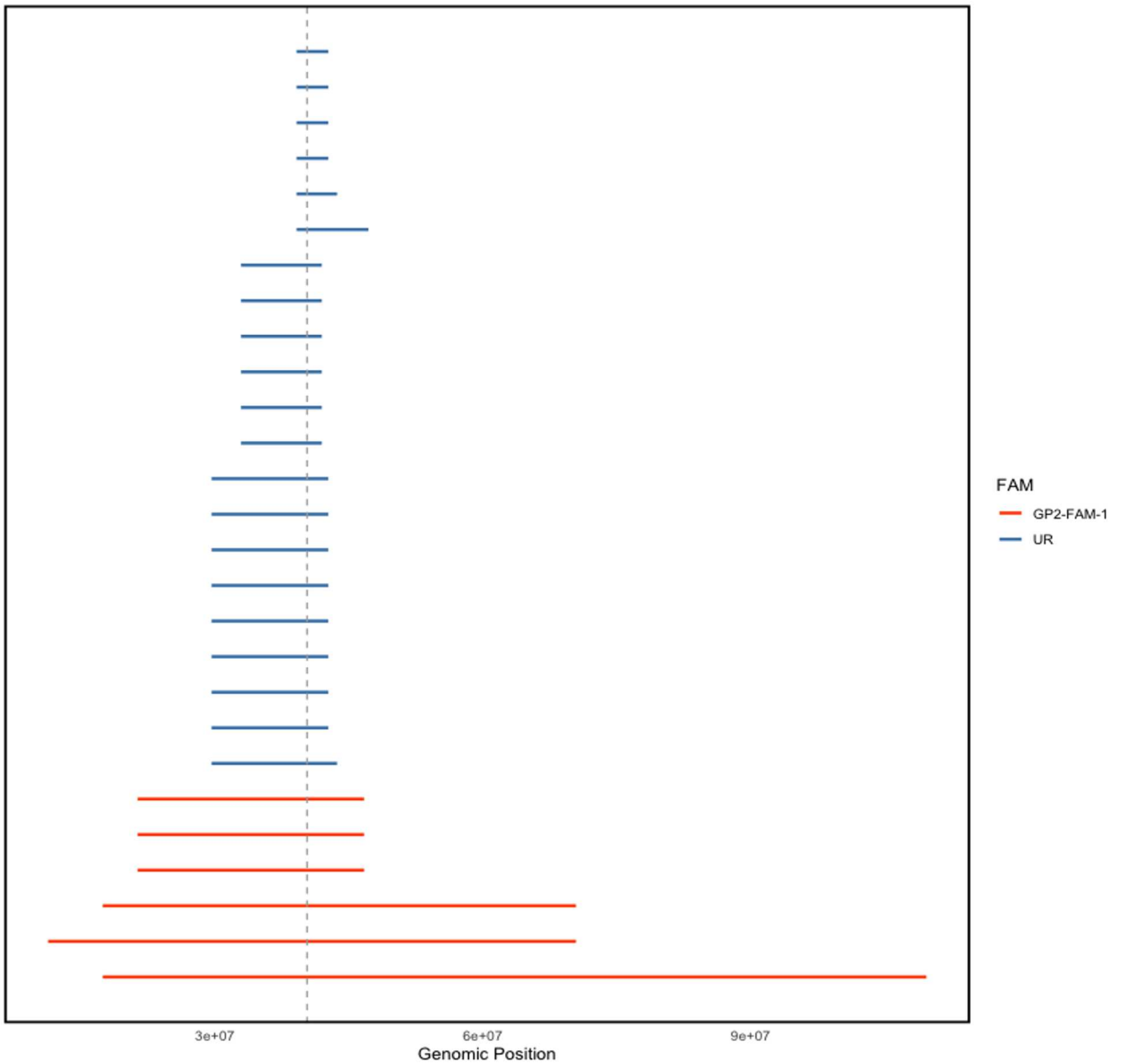


Figure 3. Overlapping identity-by-descent segments spanning *LRRK2* p.L1795F variant among the variant carriers with genotyping data. Each line represents an IBD segment inferred between a unique pair. IBD segments have been coloured according to whether both individuals within a pair belong to the same family (GP2-FAM-1) or are otherwise considered unrelated (UR). Vertical grey line represents the genomic position of *LRRK2* p.L1795F.

Table 1: Overview of the investigated cohorts.				
Cohort	Discovery		Replication	
	GP2	AMP-PD [3]	GP2	PDGENE
Data type	WGS	WGS	NBA	CES
Total number of samples	5,796	9,956 (599)	54,180	9,759
PD cases	5,283	3,442	28,729	9,759
Other phenotypes [1]	161	2,903	15,834	NA
Controls [2]	342	4,210	9,617	NA
AMP-PD = Accelerating Medicines Partnership Parkinson's disease, GP2 = Global Parkinson's Genetics Program, NA = not available, NBA = NeuroBooster Array, PD = Parkinson's disease, PD GENE = PD GENERation study, CES = clinical-exome sequencing, WGS = whole-genome sequencing				
[1] Other phenotypes include atypical parkinsonism, e.g., progressive supranuclear palsy (PSP), multi system atrophy (MSA), corticobasal degeneration/syndrome (CBD/CBS), and dementia with Lewy bodies (DLB), as well as prodromal PD.				
[2] Controls include asymptomatic carriers of known pathogenic variants.				
[3] Joint-genotyping using the 9,956 samples from BioFIND, Harvard Biomarkers Study (HBS), Lewy body dementia case-control cohort (LBD), Parkinson's disease Biomarkers Program (PDBP), Parkinson's Progression Markers Initiative (PPMI), Study of Isradipine as a Disease-modifying Agent in Subjects With Early Parkinson Disease, Phase 3 (STEADY-PD3), Study of Urate Elevation in Parkinson's Disease, Phase 3 (SURE-PD3), and Postmortem Cohort. AMP-PD release 4 was used to screen for potential pathogenic variants for the 599 samples from the LRRK2 Cohort Consortium (LCC).				

Table 2: Frequency of the *LRRK2* p.L1795F and p.G2019S variants across ancestries in the GP2 genotyping cohort.

Variant	Ancestry	AF in cases	AF in controls	Number of alleles in cases (AN_case)	Number of alleles in controls (AN_control)
chr12:40322386:G:T (<i>LRRK2</i> p.L1795F)	EUR	0.0001201	0	41634	18064
chr12:40340400:G:A (<i>LRRK2</i> p.G2019S)	AAC	0	0.0006281	568	1592
	AFR	0	0	1876	3252
	AJ	0.07081	0.01098	2556	820
	AMR	0.01339	0.003247	896	308
	CAH	0.006783	0.003436	1032	582
	CAS	0	0	1104	688
	EAS	0	0	5122	4752
	EUR	0.003266	0.000166	41636	18074
	FIN	0	0	192	14
	MDE	0.02805	0	606	446
	SAS	0	0	732	412

AF = Allele frequency, AAC= African admixed, AFR= African, AJ = Ashkenazi Jewish, AMR = Latino and Indigenous people of the Americas, CAH = Complex Admixture History, CAS= Central Asian, EAS = East Asian, EUR = European, FIN = Finnish, MDE = Middle Eastern, SAS = South Asian

LRRK2: ENST00000298910.12; ENSP00000298910.7

Table 3: Demographic and clinical characteristics of identified LRRK2 p.L1795F variant carriers.

Cohort	GP2								AMP-PD				PD GENE	TORONTO			
	GP2-FAM-1				NA				AMP-FAM-1			NA	TORONTO-FAM-1				
Family ID	GP2-ID-1	GP2-ID-2	GP2-ID-3	GP2-ID-4	GP2-ID-5	GP2-ID-6	GP2-ID-7	GP2-ID-8	AMP-ID-1	AMP-ID-2	AMP-ID-3	AMP-ID-4	PDGENE-ID-1	TORONTO-ID-1	TORONTO-ID-2	TORONTO-ID-3	TORONTO-ID-4
Sample ID	GP2-ID-1	GP2-ID-2	GP2-ID-3	GP2-ID-4	GP2-ID-5	GP2-ID-6	GP2-ID-7	GP2-ID-8	AMP-ID-1	AMP-ID-2	AMP-ID-3	AMP-ID-4	PDGENE-ID-1	TORONTO-ID-1	TORONTO-ID-2	TORONTO-ID-3	TORONTO-ID-4
Genetic method	NBA, WGS	NBA, WGS	NBA, WGS	NBA, WGS	NBA, WGS	NBA	NBA	NBA	WGS	WGS	WGS	WGS	CES	Single gene testing (LRRK2)			
Demographics																	
Gender	Female	Female	Male	Female	Male	Male	Female	Male	Female	Female	Female	Female	Female	Male	Female	Female	Female
Genetic ancestry	EUR	EUR	EUR	EUR	EUR	EUR	EUR	EUR	EUR	EUR	EUR	EUR	EUR	White	White	White	White
Age at sample collection	78	74	68	59	42	72	62	76	76	55	54	69	57	75	76	55	54
Family history of PD	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Family history details	two children, three sisters, one nephew, several aunts and uncles	three sisters, one niece and two nephews, several aunts and uncles	sister, mother, three maternal aunts	brother, mother, three maternal aunts	aunt, two great uncles	brother, mother	mother, sister	mother	father, two siblings, child	sibling, maternal grandparent, maternal aunt	maternal grandparent, two maternal aunts	mother	maternal grandmother	father, two siblings, two nieces	father, two siblings, two children	sibling, mother, two maternal uncles, maternal grandfather	sibling, mother, two maternal uncles, maternal grandfather
Clinical data																	
Diagnosis	PD	PD	PD	PD	PD	PD	PD	PD	Control*	Control*	PD*	PD	PD	PD	Control**	PD	PD
Age at motor symptom onset	55	54	50	50	36	60	57	55	NA	NA	46	65	47	65	NA	66	44
Bradykinesia	+	+	+	+	+	+	+	+	NA	NA	+	+	+	+	-	+	+
Rigidity	+	+	+	+	+	-	+	+	NA	NA	+	+	+	-	-	+	+
Resting Tremor	+	+	+	+	+	-	+	+	NA	NA	+	+	+	-	-	-	-
Action/Kinetic Tremor	+	+	+	+	-	+	+	NA	NA	NA	-	+	-	-	-	+	+
Postural Instability	+	+	-	+	+	-	+	+	NA	NA	-	-	-	-	-	-	+
Gait Disturbance	+	+	-	+	+	-	-	NA	NA	NA	-	+	-	-	-	-	+
Asymmetric onset of symptoms	+	+	+	+	+	+	+	NA	NA	NA	+	NA	+	+	NA	NA	+
Responsive to dopaminergic medication	+	+	+	+	+	+	+	NA	NA	NA	+	NA	+	NA	NA	NA	+
Fluctuations	NA	NA	+	+	-	NA	NA	NA	NA	NA	+	NA	+	-	-	-	+
UPDRS Part III (motor score)	70	NA	10	22	24	6	11	NA	NA	NA	3	32	6	6	0	7	43
Hoehn & Yahr	5	2	2	2	2	1	1.5	NA	NA	NA	2	2	2	1	0	0	3
Cognition	MMSE 29	MMSE 29	MMSE 30	MMSE 30	MMSE 30	MMSE 30	MMSE 30	NA	NA	NA	MoCA 28	NA	-	MoCA 17	MoCA 23	MoCA 29	MoCA 28
Neuropsychiatric features	NA	NA	-	-	NA	NA	NA	NA	NA	NA	NA	NA	-	NA	NA	-	+
Dysautonomia	-	-	-	constipation	-	-	-	NA	NA	NA	NA	NA	-	-	-	-	-
Atypical Features or signs suggestive of other diagnosis (#)	history of head trauma with loss of consciousness	-	-	-	history of head trauma with loss of consciousness	-	-	NA	NA	NA	NA	NA	-	-	-	-	-

+ present; - absent

EUR = European, MMSE = Mini Mental State Examination, MoCA = Montreal Cognitive Assessment, NA = Not available or applicable, NBA = NeuroBooster Array, PD = Parkinson's disease, CES = clinical-exome sequencing, WGS = Whole-genome sequencing

* Individuals were recruited through the LCC as "Genetically enriched" study arm.

** Recruited as unaffected family member, not population control.

(#) These include: history of strokes or stepwise deterioration, history of head injury with loss of consciousness, history of encephalitis, Oculogyric crisis, neuroleptic treatment at time of symptom onset, sustained remission, gaze palsy, Cerebellar signs (other than activation tremor), Fluctuations, hallucinations, dysautonomia, Memory loss, axial rigidity, Other

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.pdf](#)