Investigating the Protective Role of the Mitochondrial 2158 T > C Variant in Parkinson's Disease

A considerable portion of the risk for Parkinson's disease (PD) is attributed to genetic factors.[1,2] Several monogenic forms of PD have been associated with mutations in genes encoding proteins involved in mitochondrial function including *PRKN* and *PINK1*.[1,3,4] Furthermore, human cell culture studies and animal models have offered evidence supporting the presence of mitochondrial disturbances in PD.[5]

Hudson et al. proposed a protective role of two mitochondrial DNA variants in PD etiology.[6] In an array-based genotyping study, the authors showed that the m.2158 T > C (p.Lys4Arg, rs41349444) variant in *SHLP2* is associated with reduced risk for PD (P-value $= 2 \times 10^{-2}$, OR = 0.32). A follow-up functional study by Kim et al. demonstrated that the mutated protein was protective against mitochondrial dysfunction in both *in vitro* and *in vivo* models of PD.[7] Nevertheless, the association of this variant with reduced risk of PD has not been confirmed in large-scale sequencing datasets.

To further investigate the association between m.2158 T > C and PD, we conducted an extensive genetic characterization utilizing large-scale genome sequencing (GS) datasets, totaling 4358 PD cases and 16,609 controls. Additionally, we included 779 maternal PD proxies from All of Us, considering the maternal transmission of mitochondrial DNA. The homoplasmic allele frequency (AF) of m.2158 T > C variant is reported as 0.0066 in gnomAD v.4.0.0.[8] Considering the limited capture of rare variants by genotyping arrays, the challenge becomes more substantial for a variant in mitochondrial DNA. GS offers a comprehensive, accurate, and high-resolution approach to explore mitochondrial DNA, making it the preferred method for researchers studying the complexity of mitochondrial genetics and associated diseases. Worldwide and extensive efforts, exemplified by initiatives such as the Global Parkinson's Genetics Program (GP2; https://gp2.org/), enable us to conduct large-scale and unbiased screenings, facilitating genetic associations with significant statistical power.

First, we genotyped the m.2158 T > C variant from alignment files using the mitochondrial mode.[9] Details regarding sequencing, which includes sample and variant-level quality control procedures, are presented in the supplementary materials. The homoplasmic AF of the m.2158 T > C variant was 0.012, 0.010, 0.010, and 0.013 in All of Us, AMP-PD, GP2, and 100KGP, respectively. Subsequently, we performed per-cohort logistic regression analyses adjusted by age at onset for cases and age for controls, sex, and the first 10 principal components using PLINK v.2.0 (https://www.cog-genomics.org/plink/2.0/).[10] Our inverse-variance weighted meta-analysis[11] did not identify an association between *SHLP2* m.2158 T > C and reduced risk of developing PD in the cohorts under study (Table 1).

Our study, which utilized large-scale GS data from various datasets while considering covariates such as sex and age, did not support the findings reported by Kim et al. in 2024, suggesting that previous associations may represent a type 1 error. Our investigation focused on evaluating the association of the m.2158 T > C variant with PD, accounting for potential confounders. Utilizing genomes of more than 20,000 individuals provided a statistical power of over 95% to detect an association with a minimum relative risk of 1.5 (https://csg.sph.umich.edu/abecasis/cats/gas_power_calculator/). Our results underscore the significance of leveraging multiple datasets encompassing diverse populations to validate genetic associations before embarking on extensive functional follow-up studies. ■

Key Words: Parkinson's disease, *SHLP2*, mtDNA, genome sequencing, All of Us, AMP PD, GP2

***Correspondence to:** Dr. Sara Bandres-Ciga, 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 20892, USA; E-mail: sara.bandresciga@nih.gov

Relevant conflicts of interest/financial disclosures: Genomics England Limited is a wholly owned Department of Health and Social Care company created in 2013 to introduce genome sequencing into healthcare in conjunction with NHS England. All Genomics England affiliated authors are, or were, salaried by or seconded to Genomics England. The authors have no financial interests to disclose.

Funding agency: This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging (NIA), National Institutes of Health, Department of Health and Human Services; project number ZO1 AG000535/ZIA AG000534 and ZIA AG000949, as well as the National Institute of Neurological Disorders and Stroke (NINDS) and the National Human Genome Research Institute (NHGRI). Data used in the preparation of this article were obtained from the Global Parkinson's Genetics Program (GP2; https://gp2.org). GP2 is funded by the Aligning Science Across Parkinson's (ASAP) initiative and implemented by The Michael J. Fox Foundation (MJJF) for Parkinson's Research. Additional funding was provided by The Michael J. Fox Foundation for Parkinson's Research through grant MJFF-009421/17483. 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.; Sanofi US Services Inc.; and Verily Life Sciences.

Received: 8 May 2024; **Accepted:** 28 May 2024

Published online 28 June 2024 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.29892

TABLE 1 *Association between the m.2158 T > C variant in Parkinson's disease cases (n = 4358), proxies (n = 779), and controls (n = 16,609)*

Dataset	PD cases (F:M)	Controls (F:M)	Mean (SD) AO	Total	*P*-value	OR (95% CI)
All of Us	1021 (0.64)	12,787 (1.69)	60.0 (4.0)	14,587	0.45	0.50 (0.08–3.05)
All of Us maternal PD proxy	779 (1.86)	12,787 (1.69)	–		0.64	0.69 (0.14–3.33)
AMP-PD	1914 (0.55)	902 (1.05)	60.35 (10.33)	2816	0.31	3.89 (0.29–52.93)
100KGP early-onset familial PD	501 (0.69)	2691 (1.14)	44.37 (12.61)	3192	0.81	1.15 (0.37–3.56)
GP2 PD cohort	922 (0.62)	229 (0.39)	61.0 (12.86)	1151	0.51	0.32 (0.01–9.29)
Meta-analysis	4358 (0.60)	16,609 (1.51)	58.57 (–)	20,967	1	1.56 (0.42–2.40)

Maternal proxies are individuals whose mothers have Parkinson's disease but do not have the condition themselves. Logistic regression was performed adjusted by age at onset for cases, age for controls, sex, and the first 10 principal components. $I^2 = 0$, heterogeneity *P*-value = 0.55. AO, age at onset; CI, confidence interval; OR, odds ratio; PD, Parkinson's disease; SD, standard deviation.

Acknowledgments: This work was supported in part by the Intramural Research Program of the National Institute on Aging (NIA), and the Center for Alzheimer's and Related Dementias, within the Intramural Research Program of the NIA and the National Institute of Neurological Disorders and Stroke. Data used in the preparation of this article were obtained from the Global Parkinson's Genetics Program (GP2). GP2 is funded by the Aligning Science Across Parkinson's initiative and implemented by The Michael J. Fox Foundation for Parkinson's Research (https://gp2.org). This work used the computational resources of the National Institutes of Health high-performance computing Biowulf cluster (https://hpc.nih.gov). Data used in the preparation of this article were obtained from the Accelerating Medicines Partnership® (AMP®) Parkinson's Disease (AMP® PD) Knowledge Platform. 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. For up-to-date information on the study visit https://www.amp-pd.org. The All of Us Research Program is supported by the National Institutes of Health, Office of the Director: Regional Medical Centers: 1 OT2 OD026549; 1 OT2 OD026554; 1 OT2 OD026557; 1 OT2 OD026556; 1 OT2 OD026550; 1 OT2 OD 026552; 1 OT2 OD026553; 1 OT2 OD026548; 1 OT2 OD026551; 1 OT2 OD026555; IAA #: AOD 16037; Federally Qualified Health Centers: HHSN 263201600085 U; Data and Research Center: 5 U2C OD023196; Biobank: 1 U24 OD023121; The Participant Center: U24 OD023176; Participant Technology Systems Center: 1 U24 OD023163; Communications and Engagement: 3 OT2 OD023205; 3 OT2 OD023206; and Community Partners: 1 OT2 OD025277; 3 OT2 OD025315; 1 OT2 OD025337; 1 OT2 OD025276. In addition, the All of Us Research Program would not be possible without the partnership of its participants. This research was made possible through access to the data and findings generated by the 100,000 Genomes Project. The 100,000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The 100,000 Genomes Project is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK, and the Medical Research Council have also funded research infrastructure. The 100,000 Genomes Project uses data provided by patients and collected by the National Health Service as part of their care and support. Genomics England is a wholly owned Department of Health and Social Care company created in 2013 to introduce genome sequencing into healthcare in conjunction with NHS England. All Genomics England affiliated authors are, or were, salaried by or seconded to Genomics England. The authors have no financial interests to disclose.

Data Availability Statement

All GP2 data are hosted in collaboration with the Accelerating Medicines Partnership in Parkinson's Disease and are available via application on the website. The GP2 PD case and control data are available via the GP2 website (https://gp2.org; release 6 https://doi.org/10.5281/zenodo.10472143). Genotyping imputation, quality control, ancestry prediction, and processing were performed using GenoTools (version 10), publicly available on GitHub. The All of Us genomic data are available under restricted access for human subject data. Access can be obtained by following the instructions under the All of Us workbench at https://workbench.researchallofus.org/. Primary data from the 100KGP, which are held in a secure Research Environment, are available to registered users. Please see https://www.genomicsengland.co.uk/ for further information. The algorithms and tools that were used in this study are openly available at https://github.com/GP2code/. The code used can be found online at https://zenodo.org/records/11037328, https://github.com/GP2code/chrM.2158-analysis/.

Fulya Akçimen, PhD,[1] Vesna van Midden, MD,[2] S. Can Akerman, PhD,[3,4] Mary B. Makarious, BSc,[1,5] Global Parkinson's Genetics Program, Jeffrey D. Rothstein, MD, PhD,[3,4] Zih-Hua Fang, PhD,[6] and Sara Bandres-Ciga, PhD,[7*]

[1]*Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA,* [2]*Department of Neurology, University Medical Centre Ljubljana, Ljubljana, Slovenia,* [3]*Brain Science Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA,* [4]*Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA,* [5]*UCL Movement Disorders Centre, University College London, London, United Kingdom,* [6]*German Center for Neurodegenerative Diseases, DZNE, Tübingen, Germany, and* [7]*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, Maryland, USA*

References

1. Bandres-Ciga S, Diez-Fairen M, Kim JJ, Singleton AB. Genetics of Parkinson's disease: an introspection of its journey towards precision medicine. Neurobiol Dis 2020;137:104782. https://doi.org/10.1016/j.nbd.2020.104782
2. Blauwendraat C, Nalls MA, Singleton AB. The genetic architecture of Parkinson's disease. Lancet Neurol 2020;19(2):170–178. https://doi.org/10.1016/S1474-4422(19)30287-X
3. Valente EM, Abou-Sleiman PM, Caputo V, et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science 2004;304(5674):1158–1160. https://doi.org/10.1126/science.1096284
4. Kitada T, Asakawa S, Hattori N, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature 1998;392(6676):605–608. https://doi.org/10.1038/33416
5. Winklhofer KF, Haass C. Mitochondrial dysfunction in Parkinson's disease. Biochim Biophys Acta 2010;1802(1):29–44. https://doi.org/10.1016/j.bbadis.2009.08.013

6. Hudson G, Nalls M, Evans JR, et al. Two-stage association study and meta-analysis of mitochondrial DNA variants in Parkinson disease. Neurology 2013;80(22):2042–2048. https://doi.org/10.1212/WNL.0b013e318294b434
7. Kim SJ, Miller B, Hartel NG, et al. A naturally occurring variant of SHLP2 is a protective factor in Parkinson's disease. Mol Psychiatry 2024;29(2):505–517. https://doi.org/10.1038/s41380-023-02344-0
8. Chen S, Francioli LC, Goodrich JK, et al. A genomic mutational constraint map using variation in 76,156 human genomes. Nature 2024; 625(7993):92–100. https://doi.org/10.1038/s41586-023-06045-0
9. Laricchia KM, Lake NJ, Watts NA, et al. Mitochondrial DNA variation across 56,434 individuals in gnomAD. Genome Res 2022; 32(3):569–582. https://doi.org/10.1101/gr.276013.121
10. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7. https://doi.org/10.1186/s13742-015-0047-8
11. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010; 26(17):2190–2191. https://doi.org/10.1093/bioinformatics/btq340

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Long-Read Sequencing Unravels the Complexity of Structural Variants in *PRKN* in Two Individuals with Early-Onset Parkinson's Disease

About 5% to 10% of Parkinson's disease (PD) cases are monogenic; otherwise PD is generally known to be idiopathic. Although more than a dozen genes that contain disease-causing mutations have been identified to date, *PRKN* is the most frequently mutated gene in autosomal recessive early-onset PD (EOPD).[1] However, the genetic cause of patients with a typical *PRKN* phenotype is sometimes elusive because of the limitations of traditional genetic methods to detect complex structural mutations that are frequent in *PRKN*.[2]

The phenotype is usually specific, consisting of a slowly progressive EOPD with a good and long-standing response to levodopa. Dystonia, dyskinesia, and motor fluctuations are typical, whereas autonomic dysfunction, psychotic symptoms, and cognitive decline are usually absent.[3] We report 2 siblings of European ancestries exhibiting *PRKN* phenotype left undiagnosed for years after multiple genetic investigations (Fig. 1).

Siblings II-2 and II-4 presented at age 31 and 33 years, respectively, with asymmetrical limb akinesia associated with resting tremor with no medical history and no parental consanguinity. Cerebral magnetic resonance imaging was normal, and Wilson's disease biomarkers were negative. Focal and paroxysmal dystonia was present in II-4. The disease slowly evolved with a low *off*-medication state UPDRS (Unified Parkinson's Disease Rating Scale) 13 and 16 years after disease onset (scores of 33 and 35 for II-2 and II-4, respectively). Initial response to levodopa was remarkable for both (90% and 80%). At last examination, II-4 had dyskinesia and motor fluctuations. Of note, at the most recent examination (age 45 and 47 years), cognitive impairment, postural instability, neurogenic bladder, and bowel dysfunction were absent.

Because this presentation was consistent with *PRKN*-PD, we first performed *PRKN* multiple ligation probe amplification (MLPA) and Sanger sequencing, which revealed one copy of exon 4 for both individuals and the absence of pathogenic single-nucleotide variant, interpreted as a heterozygous exon 4 deletion (Fig. S1). Multiple genetic investigations, including another MLPA, digital droplet polymerase chain reaction, and targeted and exome sequencing, confirmed the presence of one copy of exon 4, without any additional pathogenic variant. Thus, this result was not sufficient to explain the phenotype.

Next we performed Oxford Nanopore long-read sequencing (LRS) for one individual using a protocol reported previously (https://www.protocols.io/view/processing-frozen-cells-for-population-scale-sqk-l-6qpvr347bvmk/v1). LRS detected a large compound heterozygous 178-kb deletion and 106-kb duplication, encompassing exons 3 and 4 and exon 3, respectively (Fig. 1). Both DNA loss and gain of the same exons 3 and 4 are described in typical *PRKN*-PD individuals, as reported in the movement disorders society gene database (https://www.mdsgene.org/d/1/g/4). Breakpoint junction PCR confirmed the presence of the two structural variants and revealed both variants in the second individual (Fig. S2). LRS did not identify any additional variants in PD known genes. Because both deletion and duplication breakpoints were located in deep intronic regions and genetic dosage of exon 3 was normal, short-read sequencing and other methods could not detect the complex and balanced rearrangement. Overall, these results demonstrated that biallelic *PRKN* variants were the cause of PD in this family.

As shown by a previous study, we here confirm the potential of LRS to determine complex *PRKN* structural variants in unsolved *PRKN*-PD cases.[4]

Key Words: Parkinson's disease, *PRKN*, long-read sequencing

***Correspondence to:** Dr. Alexis Brice, Institut National de la Recherche Médicale-U1127, Centre National de la Recherche Scientifique-UMR7225, APHP, Paris, France; E-mail: alexis.brice@icm-institute.org

Relevant conflicts of interest/financial disclosures: The authors declare that they have no conflicts of interest.

Funding agencies: This work was supported in part by the Intramural Research Programs of the National Institute on Aging (NIA). This work was also supported by the Fondation pour la Recherche Médicale (FRM, MND202004011718), the Fondation de France, la Fédération pour la Recherche sur le Cerveau (FRC), France-Parkinson Association, and the French program "Investissements d'avenir" (ANR-10-IAIHU-06).

Received: 29 April 2024; **Accepted:** 17 June 2024

Published online 28 June 2024 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.29914