

<https://doi.org/10.1038/s41531-024-00752-9>

Loss-of-function variants in *ITSN1* confer high risk of Parkinson's disease

Check for updates

Astros Th. Skuladottir^{1,2} ✉, Vinicius Tragante¹, Gardar Sveinbjornsson¹, Hannes Helgason¹, Arni Sturluson¹, Anna Bjornsdottir³, Palmi Jonsson⁴, Vala Palmadottir⁵, Olafur A. Sveinsson², Brynjar O. Jensson¹, Sigurjon A. Gudjonsson¹, Erna V. Ivarsdottir¹, Rosa S. Gisladottir^{1,6}, Arni F. Gunnarsson¹, G. Bragi Walters¹, Gudrun A. Jonsdottir¹, Thorgeir E. Thorgeirsson¹, Gyda Bjornsdottir¹, Hilma Holm¹, Daniel F. Gudbjartsson^{1,7}, Patrick Sulem¹, Hreinn Stefansson¹ & Kari Stefansson^{1,2} ✉

Parkinson's disease (PD) is a debilitating neurodegenerative disorder and its rising global incidence highlights the need for the identification of modifiable risk factors. In a gene-based burden test of rare variants (8647 PD cases and 777,693 controls) we discovered a novel association between loss-of-function variants in *ITSN1* and PD. This association was further supported with burden data from the Neurodegenerative Disease Knowledge Portal and the Accelerating Medicines Partnership Parkinson's Disease Knowledge Platform. Our findings show that Rho GTPases and disruptions in synaptic vesicle transport may be involved in the pathogenesis of PD, pointing to the possibility of novel therapeutic approaches.

Parkinson's disease (PD) is a complex neurodegenerative disorder with a large effect on individuals and society¹. The disorder is mainly characterized by death of dopaminergic neurons in the pars compacta of the substantia nigra^{1,2} as a result of aberrant α -synuclein accumulation in Lewy bodies^{3,4}, dysfunction of mitochondria, lysosomes or vesicle transport, and synaptic transport issues¹.

PD develops from a complicated interplay between genetics and environment¹. Genome-wide association studies (GWAS) have associated over 90 common variants with PD⁵. In GWAS, associations with rare variants often go undetected due to the low number of carriers. However, high-impact variants in coding regions often undergo negative selection and thus, tend to be rare. Unlike common variants, rare variants are often located within coding regions and may have substantial effects. The use of large-scale whole-genome sequence (WGS) data enable a more comprehensive search for rare variants. Combining variants predicted to cause a loss of gene function increases the statistical power^{6,7} and may allow us to pinpoint genes involved in the pathogenesis of PD. Previously, a meta-analysis found that the burden of rare variants in nine genes associated with PD using 7184 PD cases, 6701 proxy cases, and 51,650 controls of European ancestry⁸.

In a gene-based burden test, we collapsed rare variants (MAF < 0.1%) predicted to cause loss-of-function (LOF) as estimated with transcript effect

(LOFTEE)⁹ using WGS data. We tested for association between PD and burden of LOF in 11,976 genes ($P < 0.05/11,976 = 4.2 \times 10^{-6}$) in Iceland and the UK (Supplementary Data 1 and Supplementary Fig. 1). We combined the results in a meta-analysis using 8647 PD cases and 777,693 controls (Fig. 1, Supplementary Data 2 for demographics) of European descent, of which 62% have been WGS and the rest chip-genotyped and imputed. For the top finding, we included the Accelerating Medicines Partnership Parkinson's Disease dataset (AMP-PD) in the meta-analysis, adding 3538 PD cases and 2365 controls, in total 12,185 PD cases and 780,058 controls.

Rare LOF variants in *ITSN1* (Cumulative Variant Frequency [CVF] = 0.012% in Iceland, CVF = 0.013% in the UK, and CVF = 0.10% in AMP-PD) were associated with a large effect on the risk of PD (OR = 7.3, 95% Confidence Interval [CI] = 3.5–15.2, $P = 1.5 \times 10^{-7}$, Table 1). The effects in the datasets were similar (P -het = 0.86). In Iceland, we found five LOF variants in *ITSN1* and the association was mainly driven by the two most common ones, a splice acceptor (chr21:33781560-A, c.1597-1 G > A, MAF = 0.006%, $P = 0.028$, OR = 6.1) and a stop-gain variant (rs1569244177-A, p.Trp944Ter, MAF = 0.005%, $P = 0.034$, OR = 10.0, Supplementary Data 3). In the UK, we found 64 LOF variants in *ITSN1* (Supplementary Data 3). Among the 154 carriers, 13 (8.4%) were diagnosed with PD compared to 1.1% of individuals diagnosed with PD in the two population-based datasets. In the AMP-PD dataset ($N = 5903$ WGS), we

¹deCODE genetics/Amgen Inc., Reykjavik, Iceland. ²Faculty of Medicine, University of Iceland, Reykjavik, Iceland. ³Heilsuklasinn Clinic, Reykjavik, Iceland.

⁴Department of Geriatric Medicine, Landspítali University Hospital, Reykjavik, Iceland. ⁵Department of Internal Medicine, Landspítali University Hospital, Reykjavik, Iceland. ⁶Faculty of Icelandic and Comparative Cultural Studies, University of Iceland, Reykjavik, Iceland. ⁷Faculty of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland. ✉ e-mail: astros.skuladottir@decode.is; kstefans@decode.is



found six carriers of six rare LOF variants (MAF < 0.1%, Supplementary Data 3), thereof 5 with PD.

We find support for the association ($P = 8.9 \times 10^{-6}$) in a comparable LOFTEE rare variant (MAF < 0.1%) burden test including over 750,000 individuals from the UK dataset ($N \sim 450$ K), All of Us ($N \sim 250$ K), and Massachusetts General Brigham Biobank ($N \sim 53$ K) - results shown in the Neurodegenerative Diseases Knowledge Portal¹⁰.

The youngest *ITSN1* carrier with PD, diagnosed at 27 years of age, did not have variants in other genes described as pathogenic by OMIM¹¹. Neither did any of the other PD diagnosed carriers. The average age at onset was 8 years earlier among *ITSN1* carriers ($P = 0.022$) than non-carriers (62 years, SD = 12.2 vs 70 years, SD = 9.8, respectively, Supplementary Data 2 for each dataset).

We note that *ITSN1* is intolerant of LOF variants (probability of LOF intolerance [pLI] = 1), according to the gnomAD database¹², and both heterozygous LOF and missense variants in the gene have been associated with neurodevelopmental disorders¹³.

ITSN1 encodes Intersectin-1, a cytoplasmic membrane-associated protein involved in two main pathways; (1) actin cytoskeleton rearrangements through the Rho GTPase cycle^{14,15} and (2) clathrin-mediated endo- and exocytosis, including the synaptic vesicle cycle^{14,16}.

One of the pathological hallmarks of PD is the early death of dopaminergic neurons in the pars compacta of substantia nigra¹. The growth and degeneration of dopaminergic neurons depend on the actin cytoskeleton which is regulated by small GTPases of the Rho family^{15,17}, encompassing Rho, Rac, and CDC42 subfamilies. *ITSN1* acts as a guanine nucleotide exchange factor that specifically activates CDC42¹⁸, Cell Division Control Protein 42, and can, thus, modulate regulation of the actin cytoskeleton. Several findings suggest the involvement of CDC42 in PD. Downregulation of *CDC42* has been observed in various brain regions, including substantia nigra, in postmortem brain tissues from PD patients¹⁹. In PD mouse models, *Cdc42* signaling is decreased in the caudate and putamen, accompanied by impairments in motor coordination and cognitive function²⁰ and *Cdc42* knockout mice display similar phenotype as parkinsonian mice²⁰.

Aggregation of α -synuclein in Lewy bodies is another pathological hallmark of PD¹. Numerous proteins have been shown to interact with α -synuclein, including CDC42²¹. Activating Rho GTPases is followed by

substantial reduction in α -synuclein expression in dopaminergic neurons¹⁷ and neurite extension²².

In conclusion, we report a novel gene associated with PD and suggest that the loss of *ITSN1* function may be involved in the pathogenesis of PD in at least one of three ways; (1) Inactive CDC42 and its downstream pathway results in degeneration of dopaminergic neurons, (2) inactive CDC42 cannot regulate the vesicle exocytosis of α -synuclein, and (3) disrupted synaptic vesicle transport, a known feature in the pathogenesis of PD¹, via clathrin-mediated endo- and exocytosis. Thus, direct modulation of CDC42 or its upstream regulator, *ITSN1*, could be exploited as therapeutic avenue for PD.

The results for all genes tested are available (Supplementary Data 1) and will hopefully become a useful resource in future PD studies.

Methods

Study sample and ethics statement

Icelandic PD cases were identified by ICD-10 code G20 and ICD-9 code 332.0 from medical records, filed from 1985 to 2023, through collaboration with physicians at Landspítali—National University Hospital in Reykjavik, the Registry of Primary Health Care Contacts, and the Registry of Contacts with Medical Specialists in Private Practice. The data in this study was approved by the National Bioethics Committee (NBC, VSN-17-142; VSNb2017060004/03.01) following review by the Icelandic Data Protection Authority. All genotyped participants signed a broad informed consent allowing the use of their samples and data in projects at deCODE genetics approved by the NBC. Personal identifiers of the participants' data were encrypted in accordance with the regulations of the Icelandic Data Protection Authority.

The UKB resource consists of extensive phenotype and genotype data from ~500,000 participants, who enrolled in the study between 2006 and 2010 throughout the UK and were aged 40 to 69 years at recruitment²³. PD cases were identified by ICD-10 code G20 in General Practice clinical event records (Field ID 42040), ICD-10 code G20 and ICD-9 code 332.0 in UK hospital inpatient data (Field ID 41270 and 41271), and self-report in non-cancer self-reported illness records (Field ID 20002). This study was conducted under application number 42256. All participants provided an informed consent for the use of their genotype data and the link to electronic health records. The North West Research Ethics Committee reviewed and approved the UKB protocol (ref. 06/MRE08/65).

Genotyping

The genomes of 58,346 Icelanders were WGS^{24,25} using GAIIX, HiSeq, HiSeqX, and NovaSeq Illumina technology to a mean depth of at least 38x. Joint variant calling was performed with GraphTyper (v2.7.1)^{26,27}. Roughly 155,000 Icelanders were genotyped (of which all were WGS) using various Illumina SNP arrays^{24,25}. The genotypes were long-range phased²⁸, allowing for improved genotype calls using haplotype sharing information. Subsequently, familial imputation of genotypes in first and second degree relatives was used to increase the sample size²⁹.

In the UKB, 431,079 white British/Irish individuals (identified by PCA analyses)²³ were WGS using Illumina NovaSeq sequencing machines at deCODE for 214,548 individuals and Wellcome Trust Sanger Institute for 216,531 individuals. The average genome-wide sequencing coverage was 32.4x and joint variant calling was performed using GraphTyper (v.2.7.5)^{26,27}. The protocol for a preliminary release of the WGS of the UK Biobank dataset has been described in detail³⁰.

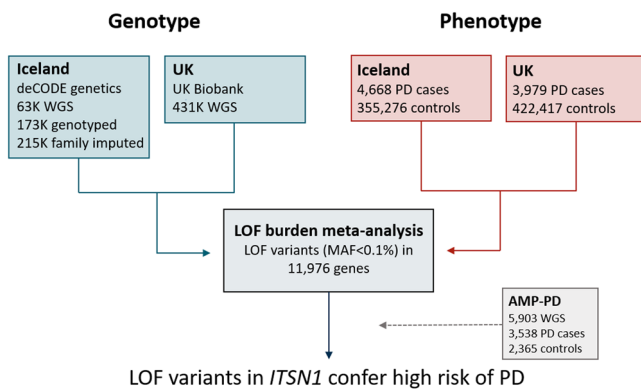


Fig. 1 | Overview of study design and summary of results.

Table 1 | Association results for burden of LOF in *ITSN1*

Gene	Dataset	Variants (N)	CVF cases/controls (%)	N carriers (cases)	P value	OR (95% CI)	P-het
<i>ITSN1</i>	Meta-analysis	75	0.15/0.018	160 (18)	1.5×10^{-7}	7.3 (3.5–15.2)	0.86
	Iceland	5	0.11/0.010	41 (5)	2.4×10^{-3}	7.5 (2.0–27.4)	
	UK	64	0.20/0.025	113 (8)	2.2×10^{-5}	7.8 (3.0–20.1)	
	AMP-PD	6	0.14/0.042	6 (5)	0.41	3.35 (0.37–158.2)	

Variant quality control and annotation

Only high-quality sequence variants were considered for selection. To estimate the quality of the sequence variants we regressed the alternative allele counts on the depth conditioned on the genotypes reported by GraphTyper^{26,27}. For a well-behaving sequence variant, the mean alternative allele count should be 0 for a homozygous genotype, depth/2 for a heterozygous genotype, and depth for homozygous alternative genotype. Assuming no sequencing or genotyping errors, the expected value of alternative allele count should be depth conditioned on the genotype, represented by an identity line (slope 1, intercept 0). Deviations from this line suggest a spurious or somatic sequence variant. We filtered variants with slope less than 0.5. Additionally, GraphTyper assigns each variant a score (AAScore), predicting the probability that it is a true positive. We only included variants with AAScore > 0.8.

Variant Effect Predictor (VEP)³¹ was used to attribute predicted consequences to the variants sequenced in the two datasets. LOF variants were classified as those predicted as start-lost, stop-gain, stop-lost, splice donor, splice acceptor, or frameshift. LOF variants with MAF < 0.1% were further evaluated by LOFTEE⁹, which determined which high-confidence LOF variants were used in this study.

Association and meta-analysis

We applied logistic regression under an additive model to test for association between gene-based burden of LOF variants and PD where the disease status was the dependent variable and genotype counts were the independent variable. To compute two-sided *P* values, we used likelihood ratio test. Individuals were coded 1 if they carry any LOF variants in the autosomal gene being tested and 0 otherwise. In the Icelandic dataset, we adjusted for sex, current age or age at death, county of origin (equivalent to principal components), blood sample availability, and an indicator function for the overlap of the lifetime of the individual with the time span of phenotype collection. In the UK dataset, we adjusted for sex, age, the first 20 principal components and three variables indicating sequencing batches, to remove batch effects. We used LD score regression intercepts³², 1.14 for Iceland and 1.00 for the UK, to adjust the χ^2 statistics and avoid inflation due to cryptic relatedness and stratification, using a set of 1.1 million variants. *P*-values were calculated from the adjusted χ^2 results.

Meta-analysis was performed on the summary results from the two datasets when available, using a fixed-effects inverse variance weighted method³³, in which the datasets were allowed to have different population frequencies for alleles and genotypes but were assumed to have a common odds ratio (OR) and weighted with the inverse of the variance of the effect estimate derived from the logistic regression. The significance threshold was based on the number of genes tested (Bonferroni significance). In a random-effects method, a likelihood ratio test was performed to test the heterogeneity of the effect estimate in the datasets; the null hypothesis is that the effects are the same in the two datasets and the alternative hypothesis is that the effects differ between datasets.

Additional datasets

The Neurodegenerative Disease Knowledge Portal (NDKP) framework is an open platform developed by a team of scientists and software engineers at the Broad Institute of MIT and Harvard. The NDKP includes results generated via consortia based science focusing on neurodegenerative diseases such as ALS, PD, and Alzheimer's disease. More information on <https://ndkp.hugeamp.org/>.

The Accelerating Medicines Partnership® (AMP®) program is a public-private partnership between the National Institutes of Health (NIH), multiple biopharmaceutical and life sciences companies, and non-profit organizations. Release 3.0 cohorts include the Michael J. Fox Foundation for Parkinson's Research (MJFF) Parkinson's Progression Marker Initiative (PPMI), The National Institute of Neurological Disorders and Stroke (NINDS) BioFIND study, Harvard Biomarkers Study (HBS), the NINDS Parkinson's Disease Biomarkers Program (PDBP), the MJFF LRRK2 Cohort Consortium (LCC), the NINDS Study of Isradipine as a Disease Modifying Agent in Subjects With Early Parkinson Disease, Phase 3

(STEADY-PD3), the MJFF and NINDS Study of Urate Elevation in Parkinson's Disease, Phase 3 (SURE-PD3), and the Global Parkinson's Genetics Program (GP2). WGS was performed by Macrogen and the Uniformed Services University of Health Sciences using the Illumina HiSeq XTen sequencer with samples coming from whole blood. The data were processed using GATK Best Practices guidelines set by the Broad Institute's joint discovery pipeline and elaborated on elsewhere³⁴. Variant annotations were generated on the joint genotyped variants with VEP. All individuals were of European ancestry as confirmed by principal component analysis using HapMap3 European ancestry populations. The *P* value in Table 1 was generated with a two-sided Fisher's Exact Test in R.

Data availability

Summary data of the gene-burden meta-analysis are in Supplementary Data 1.

Received: 29 January 2024; Accepted: 12 July 2024;

Published online: 15 August 2024

References

- Kalia, L. V. & Lang, A. E. Parkinson's disease. *Lancet* **386**, 896–912 (2015).
- Dickson, D. W. et al. Neuropathological assessment of Parkinson's disease: refining the diagnostic criteria. *Lancet Neurol.* **8**, 1150–1157 (2009).
- Spillantini, M. G. et al. α -Synuclein in Lewy bodies. *Nature* **388**, 839–840 (1997).
- Polymeropoulos, M. H. et al. Mutation in the α -synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047 (1997).
- Nalls, M. A. et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* **18**, 1091–1102 (2019).
- Li, B. & Leal, S. M. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am. J. Hum. Genet.* **83**, 311–321 (2008).
- Lee, S., Abecasis, G. R., Boehnke, M. & Lin, X. Rare-variant association analysis: study design and statistical tests. *Am. J. Hum. Genet.* **95**, 5 (2014).
- Makarious, M. B. et al. Large-scale rare variant burden testing in Parkinson's disease. *Brain* **146**, 4622–4632 (2023).
- Karczewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
- Neurodegenerative Diseases Portal. Available at: <https://ndkp.hugeamp.org/> accessed 4 January 2024.
- Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD. MIM Number: PS168600. World Wide Web URL: <https://omim.org/> Available at: <https://www.omim.org/> accessed 3 November 2023.
- ITSN1 | gnomAD v4.0.0 | gnomAD. Available at: https://gnomad.broadinstitute.org/gene/ENSG00000205726?dataset=gnomad_r4 accessed 20 November 2023.
- Bruel, A. L. et al. ITSN1: a novel candidate gene involved in autosomal dominant neurodevelopmental disorder spectrum. *Eur. J. Hum. Genet.* **30**, 111–116 (2021).
- Pechstein, A., Shupliakov, O. & Haucke, V. Intersectin 1: a versatile actor in the synaptic vesicle cycle. *Biochem. Soc. Trans.* **38**, 181–186 (2010).
- Hall, A. & Lalli, G. Rho and Ras GTPases in axon growth, guidance, and branching. *Cold Spring Harb. Perspect. Biol.* **2**, a001818 (2010).
- Katsu, M. et al. MicroRNA expression profiles of neuron-derived extracellular vesicles in plasma from patients with amyotrophic lateral sclerosis. *Neurosci. Lett.* **708**, 134176 (2019).
- Zhou, Z. et al. Rho GTPase regulation of α -synuclein and VMAT2: implications for pathogenesis of Parkinson's disease. *Mol. Cell. Neurosci.* **48**, 29–37 (2011).

18. Hussain, N. K. et al. Endocytic protein intersectin-1 regulates actin assembly via Cdc42 and N-WASP. *Nat. Cell Biol.* **3**, 927–932 (2001).
19. Zhang, Y., James, M., Middleton, F. A. & Davis, R. L. Transcriptional analysis of multiple brain regions in Parkinson's disease supports the involvement of specific protein processing, energy metabolism, and signaling pathways, and suggests novel disease mechanisms. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **137B**, 5–16 (2005).
20. Ying, L. et al. Regulation of Cdc42 signaling by the dopamine D2 receptor in a mouse model of Parkinson's disease. *Aging Cell* **21**, e13588 (2022).
21. Schnack, C., Danzer, K. M., Hengerer, B. & Gillardon, F. Protein array analysis of oligomerization-induced changes in alpha-synuclein protein–protein interactions points to an interference with Cdc42 effector proteins. *Neuroscience* **154**, 1450–1457 (2008).
22. Luo, L. RHO GTPASES in neuronal morphogenesis. *Nat. Rev. Neurosci.* **1**, 173–180 (2000).
23. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
24. Gudbjartsson, D. F. et al. Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435–444 (2015).
25. Jónsson, H. et al. Data descriptor: whole genome characterization of sequence diversity of 15,220 Icelanders. *Sci. Data* **4**, 1–9 (2017).
26. Eggertsson, H. P. et al. GraphTyper enables population-scale genotyping using pangenome graphs. *Nat. Genet.* **49**, 1654–1660 (2017).
27. Eggertsson, H. P. et al. GraphTyper2 enables population-scale genotyping of structural variation using pangenome graphs. *Nat. Commun.* **10**, 5402 (2019).
28. Kong, A. et al. Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat. Genet.* **40**, 1068–1075 (2008).
29. Gudbjartsson, D. F. et al. Sequence variants from whole genome sequencing a large group of Icelanders. *Sci. Data* **2**, 1–11 (2015).
30. Halldorsson, B. V. et al. The sequences of 150,119 genomes in the UK Biobank. *Nature* **607**, 732–740 (2022).
31. McLaren, W. et al. The Ensembl variant effect predictor. *Genome Biol.* **17**, 1–14 (2016).
32. Bulik-Sullivan, B. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
33. Mantel, N. & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl Cancer Inst.* **22**, 719–748 (1959).
34. Iwaki, H. et al. Accelerating medicines partnership: Parkinson's disease. genetic resource. *Mov. Disord.* **36**, 1795 (2021).

Acknowledgements

We thank the participants in this study for their valuable contribution to research. We thank all investigators and colleagues who contributed to data collection, phenotypic characterization of clinical samples, genotyping, and analysis of the whole-genome association data. This research was conducted using the UK Biobank Resource under Application Number 42256 and used data provided by patients collected by the NHS as part of their care and support. Data was obtained from the AMP-PD Knowledge Platform, a public-private partnership, managed by the Foundation for the National Institutes of Health (FNIH) and funded by the National Institute of Neurological Disorders and Stroke (NINDS), Aligning Science Across Parkinson's initiative (ASAP), Celgene, GlaxoSmithKline, Michael J. Fox Foundation for Parkinson's Research (MJFF), AbbVie, Pfizer, Sanofi US Services, and Verily Life Sciences. For up-to-date information on the study, visit <https://www.amp-pd.org>. AMP are registered service marks of the U.S. Department of Health and Human Services. We acknowledge AMP-PD for the whole-genome sequencing data made available from MJFF, NINDS BioFIND study, Harvard Biomarker Study (HBS), MJFF LRRK2 Cohort Consortium (LCC), NINDS Parkinson's Disease Biomarkers Program (PDBP), MJFF Parkinson's Progression Markers Initiative (PPMI), NINDS Study of Isradipine as a Disease-modifying Agent in Subjects With Early Parkinson Disease, Phase 3 (STEADY-PD3), and NINDS Study of Urate Elevation in Parkinson's Disease, Phase 3 (SURE-PD3). BioFIND is sponsored by MJFF with the support from NINDS. For up-to-date information

on the study, visit <https://michaeljfox.org/biofind>. HBS is a collaboration of HBS investigators (full list found at <https://www.bwhparkinsoncenter.org/biobank/>) and funded through philanthropy and NIH and non-NIH funding sources. Data was obtained from MJFF LCC. For up-to-date information on the study, visit <https://www.michaeljfox.org/biospecimens>. PPMI is sponsored by MJFF and supported by a consortium of scientific partners (list of PPMI funding partners found at <https://www.ppmi-info.org/about-ppmi/who-we-are/study-sponsors>). For up-to-date information on the study, visit www.ppmi-info.org. PDBP consortium is supported by NINDS at NIH. A full list of PDBP investigators can be found at <https://pdbp.ninds.nih.gov/policy>. STEADY-PD3 is funded NINDS at NIH with support from MJFF and the Parkinson Study Group. For additional study information, visit <https://clinicaltrials.gov/ct2/show/study/NCT02168842>. SURE-PD3 is funded by NINDS at NIH with support from MJFF and the Parkinson Study Group. For additional study information, visit <https://clinicaltrials.gov/ct2/show/NCT02642393>. Data was obtained from Global Parkinson's Genetics Program (GP2). GP2 is funded by ASAP Initiative and implemented by MJFF. For a complete list of GP2 members see <http://www.gp2.org>. The BioFIND, HBS, LCC, PPMI, PDBP, STEADY-PD3, SURE-PD3, and GP2 investigators have not participated in reviewing the data analysis or content of the manuscript.

Author contributions

A.T.S., V.T., P.S., and K.S. designed the study. Data collection and subject ascertainment and recruitment was carried out by A.T.S., A.B., P.J., V.P., and H.S. A.T.S., V.T., G.S., Ha.H., A.S., B.O.J., S.A.G. analyzed the data and A.T.S., V.T., and P.S. interpreted the results. A.T.S. drafted the manuscript with input from V.T., G.S., A.B., O.A.S., B.O.J., E.V.I., R.S.G., A.F.G., G.B.W., G.A.J., T.E.T., G.B., Hi.H., D.F.G., P.S., H.S., and K.S. All authors read and approved the final version of the manuscript.

Competing interests

A.T.S., V.T., G.S., Ha.H., A.S., B.O.J., S.A.G., E.V.I., R.S.G., A.F.G., G.B.W., G.A.J., T.E.T., G.B., Hi.H., D.F.G., P.S., H.S., and K.S. are employees of deCODE genetics/Amgen Inc. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at

<https://doi.org/10.1038/s41531-024-00752-9>.

Correspondence and requests for materials should be addressed to Astros Th. Skuladottir or Kari Stefansson.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024