BRIEF REPORT

Polygenic Parkinson's Disease Genetic Risk Score as Risk Modifier of Parkinsonism in Gaucher Disease

Cornelis Blauwendraat, PhD,[1,2*] Nahid Tayebi, PhD,[3] Elizabeth Geena Woo, BA,[3] Grisel Lopez, MD,[3] Luca Fierro, MS, CGC,[4] Marco Toffoli, MD,[5] Naomi Limbachiya, MSc,[5] Derralynn Hughes, MD,[6] Vanessa Pitz, PhD,[1] Dhairya Patel, BS,[1] Dan Vitale, MS,[2,7] Mathew J. Koretsky, BS,[2] Dena Hernandez, PhD,[8] Raquel Real, PhD,[5] Roy N. Alcalay, MD,[9,10] Mike A. Nalls, PhD,[2,7,8] Huw R. Morris, FRCP, PhD,[5] Anthony H.V. Schapira, MD,[5] Manisha Balwani, MD, MS,[4] and Ellen Sidransky, MD[3*]

[1]*Integrative Neurogenomics Unit, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA* [2]*Center for Alzheimer's and Related Dementias, National Institute on Aging and National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland, USA* [3]*Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA* [4]*Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA* [5]*Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, University College London, London, United Kingdom* [6]*Lysosomal Storage Diseases Unit, Royal Free London Hospital NHS Foundation Trust, and Department of Hematology, University College London, London, United Kingdom* [7]*Data Tecnica International, Washington, District of Columbia, USA* [8]*Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA* [9]*Department of Neurology, Columbia University Irving Medical Center, New York, New York, USA* [10]*Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel*

ABSTRACT: Background: Biallelic pathogenic variants in *GBA1* are the cause of Gaucher disease (GD) type 1 (GD1), a lysosomal storage disorder resulting from deficient glucocerebrosidase. Heterozygous *GBA1* variants are also a common genetic risk factor for Parkinson's disease (PD). GD manifests with considerable clinical heterogeneity and is also associated with an increased risk for PD.

Objective: The objective of this study was to investigate the contribution of PD risk variants to risk for PD in patients with GD1.

Methods: We studied 225 patients with GD1, including 199 without PD and 26 with PD. All cases were genotyped, and the genetic data were imputed using common pipelines.

Results: On average, patients with GD1 with PD have a significantly higher PD genetic risk score than those without PD ($P = 0.021$).

Conclusions: Our results indicate that variants included in the PD genetic risk score were more frequent in patients with GD1 who developed PD, suggesting that common risk variants may affect underlying biological pathways. © 2023 The Authors. *Movement Disorders* published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

Key Words: genetics; Parkinson's; Gaucher; *GBA1*

Introduction

Pathogenic variants in *GBA1*, the gene encoding the enzyme glucocerebrosidase (GCase), are associated with the recessively inherited lysosomal storage disorder Gaucher disease (GD). *GBA1* mutations are also the most common known genetic risk factor for Parkinson's disease (PD) and dementia with Lewy bodies.[1,2] Traditionally, GD is divided into three clinical types based on the severity of the disease and the degree of neurological involvement: type 1 (GD1) is non-neuronopathic, type 2 is the acute neuronopathic form presenting in infancy, and type 3 is chronic neuronopathic. There is a wide spectrum of disease manifestations and disease severity, even within the different types.[3]

*__Correspondence to:__ Dr. Cornelis Blauwendraat, Laboratory of Neurogenetics, Integrative Genomics Unit, National Institute on Aging, National Institutes of Health, Bld 35, 35 Convent Drive, Bethesda, MD 20892, USA; E-mail: cornelis.blauwendraat@nih.govDr. Ellen Sidransky, Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bld 35A, 35A Convent Drive, Bethesda, MD 20892, USA; E-mail: sidranse@mail.nih.gov

Relevant conflicts of interest/financial disclosures: M.A.N. and D.V. are consultants employed by Data Tecnica International; their participation in this work is part of a consulting agreement between the National Institutes of Health and said company.

Full financial disclosures and author roles may be found in the online version of this article.

Received: 21 September 2022; **Revised:** 5 December 2022; **Accepted:** 3 January 2023

Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.29342

Approximately 600 GD-causing *GBA1* variants have been described, of which the vast majority in the heterozygous state also confer an increased risk for PD. Certain *GBA1* variants (including p.E365K and p.T408M[4,5]) are considered PD risk factors and affect PD progression, but they do not cause GD, even in homozygotes, although they may produce a modest reduction in GCase activity.[6] Genotype/phenotype correlations for GD are somewhat limited, although some pathogenic variants are more likely to be associated with a specific type of GD. For example, the common N370S (p.N409S) mutation is exclusively seen in GD1. By definition, patients with GD1 do not have CNS involvement, although they are at increased risk for development of PD and DLB. However, the great majority of patients with GD do not exhibit parkinsonism, even as they age. Although there are no reliable large-scale studies, estimates are that 8% to 12% of patients with GD1 at age 80 years report PD symptoms,[7,8] indicating that a majority of patients with GD1 escape PD. Thus, patients with the lysosomal storage disorder GD1 provide a unique cohort for exploring other secondary (potentially genetic) factors predisposing to or protecting from the development of PD.[5] Although these patients have a major deficiency of GCase and often multisystem involvement, it is essential to understand why only a subset develop PD. In this study, we performed genome-wide genotyping on a relatively large cohort of subjects with GD1 and assessed the contribution of common variants included in the most recent PD risk score[9] to PD risk in patients with GD1.

Subjects and Methods

Cohort Information

GD1 cases were collected from three different sites: National Human Genome Research Institute, Mt. Sinai, and University College London (UCL)/Royal Free London Hospital NHS Foundation Trust. The National Human Genome Research Institute GD dataset was collected over the past two decades and includes patients with a verified diagnosis of GD evaluated at the Clinical Center of the National Institutes of Health under an institutional review board (IRB)-approved clinical protocol. *GBA1* genotyping, performed on all individuals, was done by Sanger sequencing of all exons of the gene. Patients were evaluated by a movement disorder specialist for signs of PD. A subset of the cohort was followed longitudinally for up to 15 years. Mt. Sinai cases were collected via an IRB-approved study beginning January 2019. Patients were evaluated at the Lysosomal Storage Disease program, had a confirmed diagnosis of GD, and were screened for PD, with the diagnosis confirmed by a movement disorder specialist. UCL GD cases were collected through UK Lysosomal Storage Diseases Centres, clinically assessed, and all *GBA1* exons genotyped through a project approved by the local ethics (IRB) committee. PD symptoms and diagnosis were all confirmed by a movement disorder specialist. Information on age of onset of motor symptoms for GD1 with PD and age at last examination for GD1 was also collected. Additional PD and control data were obtained from the Global Parkinson's Genetics Program (GP2) (release 1), which is genotyped on the same genotyping array, facilitating comparison with GD1 and GD1-PD data.[10] More cohort details are provided in Table S1.

Genotyping of DNA and Data Analysis

Genotyping was performed at National Institutes of Health/Laboratory of Neurogenetics and UCL Genomics centers using the new Global Diversity Array with NeuroBooster content (https://github.com/GP2code/Neuro_Booster_Array). Genetic data were cleaned and imputed using standard GP2 pipelines and protocols (https://github.com/GP2code/). In brief, sample-level quality-control steps was performed using PLINK (1.9 and 2) and includes genotype missingness (<0.02), genetic sex confirmation, duplicate check, and relatedness confirmation. We excluded a random sample if relatedness higher than 0.2 was identified, although if one of the related samples was a GD1-PD case, that sample was retained. Variants were filtered for missingness and other standard parameters and imputed using TOPMed Imputation Server.[11,12] Ancestry was defined using 1000 Genomes and an Ashkenazi Jewish (AJ) reference population.[13,14] Principal component (PC) analysis was performed using PLINK, and when plotting, genetic ancestries were color coded (Fig. S1). Additional PCs were generated per ancestry and included in subsequent analyses as covariates, together with age of onset for PD cases and age of last examination for non-PD subjects. Age was missing for 27 patients, and these were imputed to the mean. The PD genetic risk score was calculated using imputed data. We used weights from known genome-wide association study (GWAS) variants from Nalls et al,[9] excluding the full *GBA1* region, and two additional variants were filtered out because of high missingness >5% (chr10:119776815:G:A and chr19:2341049:C:T), leaving 85 variants. The genetic risk score was normalized based on GP2 controls from each population (AJ and European [EUR]), with one unit of change representing a single standard deviation increase in the control risk score. Case–control analysis was performed using a logistic regression using biological sex and five PCs as covariates. Meta-analysis was performed across ancestries using a conservative random effects model. All calculations and figures were made with R (3.6.1 or

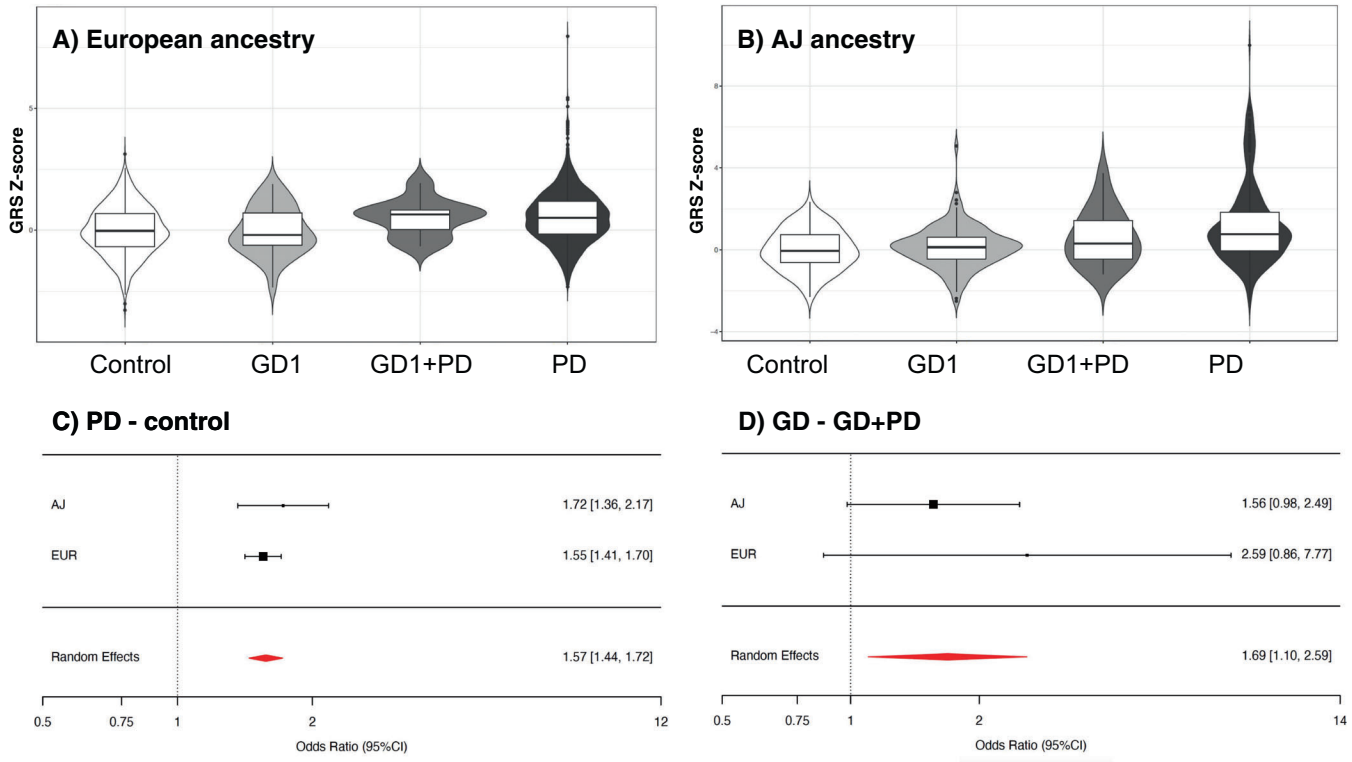

FIG. 1. Parkinson's disease (PD) genetic risk score comparisons across cohorts and ancestries. (**A**) Violin plots of genetic risk score of European (EUR) ancestry control, Gaucher disease type 1 (GD1) without PD, GD1 with PD, and PD cases show a higher risk score (normalized *Z* score) in both GD1 with PD and PD cases versus control and GD1, respectively. (**B**) Identical plots for subjects with Ashkenazi Jewish (AJ) ancestry. (**C**) Forest plot of PD and control risk score analysis of EUR and AJ ancestries shows a significant effect. (**D**) Forest plot of GD1 without PD and GD1 with PD risk score analysis of EUR and AJ ancestries shows a significant effect. Bracket end lines denote 95% confidence interval (CIs) of the study-specific estimates, with squares for effect estimates (odds ratios), and the size of these squares is proportionate to sample size. The red diamonds denote effect and 95% CI of the meta-analyzed results on the odds ratio scale. [Color figure can be viewed at wileyonlinelibrary.com]

4.2.0) using packages ggplot2, metafor, and rmeta. All code used can be found online at: https://github.com/neurogenetics/Gaucher_PD_GRS_modifiers.

Results

To investigate genetic modifiers of risk for PD symptoms in patients with GD1, we genotyped three GD1 cohorts totaling 266 GD1 cases, of which 27 were also diagnosed with PD. Unfortunately, the analysis was possible only in the EUR and AJ subgroups due to insufficient GD1 with PD cases for analysis in the other ancestral groups. After further sample-level quality control, this resulted in 26 GD1 with PD cases (18 AJ and 8 EUR) and 199 GD1 cases without PD (134 AJ and 65 EUR). As additional data for comparison, we included PD cases (335 AJ and 2050 EUR) and controls (109 AJ and 933 EUR) from GP2 (Table S2).

Using the imputed data, we calculated the PD genetic risk score for each participant and normalized the values based on the GP2 controls of each ancestry. The genetic risk score is the cumulative dosage of risk alleles with each SNP's contribution weighted by its identified GWAS effect estimate as a means of predicting risk. First, we compared the genetic risk score for PD versus controls per ancestry and meta-analyzed results. In this study, we identified a significant difference ($P = 1.18E{-}24$) with a highly similar effect size (odds ratio [OR] = 1.575, confidence interval [CI] = 1.444–1.717, $I^2 = 0$) as previously described[9] (Fig. 1A–C, Table S3). Next, we performed a similar analysis using GD1 with PD cases and GD1 without PD. We identified a significant effect ($P = 0.0168$, OR = 1.687, CI = 1.099–2.589) between GD1with PD cases and GD1 without PD (Fig. 1A,C,D, Table S4). No significant heterogeneity of effect was detected ($I^2 = 0.0$), and the effect size was remarkably similar to the PD versus control groups. Notably, two AJ individuals were identified as carriers of *LRRK2* p.G2019S, one with GD1 with PD and one GD1 without reported PD diagnosis.

Discussion

Current estimates of patients with GD1 who report PD symptoms by age 80 range from 8% to 12%, which is significantly higher than the general risk for PD, estimated at 2%.[15] It is clear that the majority of patients

with GD1 do not develop PD symptoms, suggesting that there are other modifiers of risk. In this study, we assessed potential genetic modifiers by investigating the PD genetic risk score and determined that the PD genetic score contributes to the risk for PD symptoms in patients with GD1. Interestingly, we previously identified a similar contribution of the PD genetic risk score in *LRRK2* p.G2019S and *GBA1* heterozygous carriers.[16,17]

Other lysosomal storage disorder genes have also been implicated in PD by rare variant analysis,[18] and several lysosomal pathways are associated with PD,[9] emphasizing the importance of the lysosome in PD pathogenesis. Currently, the most prevalent hypothesis linking *GBA1* mutations to PD posits a connection between α-synuclein aggregation and reduced GCase activity.[19,20] Our results suggest that the common risk variant component of PD contributes to this complex interplay between *GBA1* and α-synuclein aggregation and can increase the risk for development of PD even in GD1 cases, which are already at higher risk for developing PD.

Naturally this study has some limitations. The number of patients with both GD1 and PD included is relatively small (n = 28). Because GD1 is a rare disease with a prevalence of approximately 1 in 100,000 in the general population[21] and with an estimate of ∼10% developing PD symptoms, obtaining more participants is challenging. Due to our relatively small numbers, we also could not investigate the effect of specific *GBA1* variants on PD risk in patients with GD1, as well as on the onset and severity of symptoms, although this is an area of interest that requires further study. All three sites conducted longitudinal assessments over the past years/decades, with evaluations not specifically designed for this study. Therefore, harmonized clinical data regarding specific clinical measurements such as motor signs or prodromal nonmotor symptoms of PD are not available. We also lack ancestral diversity because of both the low number of participants and the demographics of the referral centers participating. We are planning to expand this cohort and welcome future collaborations. Furthermore, with limited power to perform single-variant testing, we could not define GD1 PD-specific risk estimates and used the previously established PD risk estimates from previous GWASs. In addition, we could not identify which specific PD GWAS loci contribute to the PD risk in the participating patients with GD1.

In summary, for the first time, we show that a PD genetic risk score is a risk modifier of PD symptoms in patients with GD1, which has potentially interesting biological implications. Together with our previous work,[16] our findings provide clear evidence of substantial overlap in the genetic risk affecting *GBA1*-associated PD and typical PD. Thus, identifying the specific factors impacting risk in *GBA1*-PD may yield shared mechanisms underlying other forms of PD. Future studies should carefully phenotype larger cohorts of patients with GD1 to accurately estimate age-specific penetrance of PD and identify other modifiers.

Acknowledgments: This work was supported in part by the Intramural Research Program of the National Human Genome Research Institute, the Center for Alzheimer's and Related Dementias, within the Intramural Research Program of the National Institute on Aging and the National Institute of Neurological Disorders and Stroke, National Institutes of Health, Department of Health and Human Services, and the National Institutes of Health and Cure Parkinson's (UK). Data used in the preparation of this article were obtained from 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). For a complete list of GP2 members, see https://gp2.org.

Data availability statement

The algortihms and tools that were used in this study are openly available a https://github.com/GP2code/Neuro_Booster_Array and https://github.com/GP2code/. The code used can be found online at: https://github.com/neurogenetics/Gaucher_PD_GRS_modifiers. The PD case and control data is available via GP2 (https://gp2.org).

References

1. Sidransky E et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. N Engl J Med 2009;361:1651–1661.
2. Nalls MA et al. A multicenter study of glucocerebrosidase mutations in dementia with Lewy bodies. JAMA Neurol 2013;70:727–735.
3. Sidransky E. Gaucher disease: complexity in a 'simple' disorder. Mol Genet Metab 2004;83:6–15.
4. Davis MY et al. Association of GBA Mutations and the E326K Polymorphism With Motor and Cognitive Progression in Parkinson Disease. JAMA Neurol 2016;73:1217–1224.
5. Davidson BA, Hassan S, Garcia EJ, Tayebi N, Sidransky E. Exploring genetic modifiers of Gaucher disease: the next horizon. Hum Mutat 2018;39:1739–1751.
6. Alcalay RN et al. Glucocerebrosidase activity in Parkinson's disease with and without GBA mutations. Brain 2015;138:2648–2658.
7. Bultron G et al. The risk of Parkinson's disease in type 1 Gaucher disease. J Inherit Metab Dis 2010;33:167–173.
8. Thaler A et al. A 'dose' effect of mutations in the GBA gene on Parkinson's disease phenotype. Parkinsonism Relat Disord 2017;36:47–51.
9. Nalls MA et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet Neurol 2019;18:1091–1102.
10. Global Parkinson's Genetics Program. GP2: The Global Parkinson's Genetics Program. Mov Disord 2021;36:842–851.
11. Taliun D et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature 2021;590:290–299.
12. Das S et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–1287.
13. Bray SM et al. Signatures of founder effects, admixture, and selection in the Ashkenazi Jewish population. Proc Natl Acad Sci U S A 2010;107:16222–16227.
14. Fairley S, Lowy-Gallego E, Perry E, Flicek P. The international genome sample resource (IGSR) collection of open human genomic variation resources. Nucleic Acids Res 2020;48:D941–D947.

15. Gasser T. Genetics of Parkinson's disease. Curr Opin Neurol 2005; 18:363–369.

16. Blauwendraat C et al. Genetic modifiers of risk and age at onset in GBA associated Parkinson's disease and Lewy body dementia. Brain 2020;143:234–248.

17. Iwaki H et al. Penetrance of Parkinson's Disease in LRRK2 p.G2019S Carriers Is Modified by a Polygenic Risk Score. Mov Disord 2020;35:774–780.

18. Robak LA et al. Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. Brain 2017;140:3191–3203.

19. Stojkovska I, Krainc D, Mazzulli JR. Molecular mechanisms of α-synuclein and GBA1 in Parkinson's disease. Cell Tissue Res 2018; 373:51–60.

20. Aflaki E, Westbroek W, Sidransky E. The complicated relationship between Gaucher disease and parkinsonism: insights from a rare disease. Neuron 2017;93:737–746.

21. Nalysnyk L, Rotella P, Simeone JC, Hamed A, Weinreb N. Gaucher disease epidemiology and natural history: a comprehensive review of the literature. Hematology 2017;22:65–73.

Supporting Data

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

SGML and CITI Use Only
DO NOT PRINT

Author Roles

1. Research project: A. Conception, B. Organization, C. Execution;
2. Cohort generation: A. Cohort recruitment, B. Sample processing, C. Data generation:
3. Statistical Analysis: A. Design, B. Execution, C. Review and Critique;
4. Manuscript Preparation: A. Writing of the first draft, B. Review and Critique;

C.B.: 1A, 1B, 1C, 3A, 3B, 3C, 4A.
N.T.: 1B, 2A, 2B, 2C, 4B.
E.G.W.: 2A, 2B, 2C, 4B.
G.L.: 2A, 2B, 2C, 4B.
L.F.: 2A, 2B, 2C, 4B.
M.T.: 2A, 2B, 2C, 4B.
N.L.: 2A, 2B, 2C, 4B.
D. Hughes: 2A, 2B, 2C, 4B.
V.P.: 3B, 4B.
D.P.: 3B, 4B.
D.V.: 3B, 3C, 4B.
M.J.K.: 3B, 3C, 4B.
D. Hernandez: 2C, 4B.
R.R.: 2C, 4B.
R.N.A.: 2A, 2B, 2C, 4B.
M.A.N.: 3A, 3B, 3C, 4B.
H.R.M.: 2C, 4B.
A.H.V.S.: 2A, 2B, 2C, 4B.
M.B.: 2A, 2B, 2C, 4B.
E.S.: 1A, 1B, 1C, 2A, 2B, 2C, 4A.

Financial Disclosures

M.A.N. and D.V. are consultants employed by Data Tecnica International; their participation in this work is part of a consulting agreement between the National Institutes of Health and said company.