Assessment of private variants in *PRKN, PARK7* and *PINK1* in Parkinson's disease

Jing Hu[1,2], Cheryl H. Waters[3], Dan Spiegelman[4], Edward A. Fon[4,5], Eric Yu[4,6], Farnaz Asayesh[4,5], Lynne Krohn[4,6], Prabhjyot Saini[4,6], Roy N. Alcalay[3,7], Sharon Hassin-Baer[8,9], Ziv Gan-Or[4,5,6], Dimitri Krainc[2], BaoRong Zhang[1*], Bernabe I. Bustos[2*#], Steven J. Lubbe[2#], t

[1] *Department of Neurology, The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, 310009, China.*
[2] *Ken and Ruth Davee Department of Neurology and Simpson Querrey Center for Neurogenetics, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA.*
[3] *Department of Neurology, Columbia University Irving Medical Center*
[4] *The Neuro (Montreal Neurological Institute-Hospital), Montreal, Quebec, Canada.*
[5] *The Department of Neurology & Neurosurgery, McGill University, Montreal, Quebec, Canada.*
[6] *The Department of Human Genetics, McGill University, Montreal, Quebec, Canada.*
[7] *Center of Movement Disorders, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel.*
8 The Movement Disorders Institute, Department of Neurology, Chaim Sheba Medical Center, Tel- Hashomer, Israel.
9 The Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.

 # these authors contributed equally

Correspondence to: Bernabe I. Bustos (bernabe.bustos@northwestern.edu) and BaoRong Zhang (brzhang@zju.edu.cn).

Abstract

Recessive mutations in *PRKN*, *PARK7* and *PINK1* are established causes of early-onset Parkinson's disease (EOPD). Previous studies have interrogated the role of heterozygous variants in these genes but mainly focused on rare (minor allele frequency [MAF] <1%) damaging variants or established mutations. Here, we assessed heterozygous private *PRKN*, *PARK7* and *PINK1* variants in PD risk in four large-scale PD case-control datasets by performing gene-wise burden analyses using sequencing data totaling 5,831 PD cases and 7,221 controls, and summary allele counts from 9,506 PD cases and 48,207 controls. Results showed no significant burden in all three genes after meta-analyses. Burden in EOPD (age at onset <50 years) and late-onset PD (≥50 years) remained nonsignificant. In summary, we found no evidence to support the association of the excess burden

of heterozygous private variants in *PRKN*, *PARK7,* and *PINK1* with PD risk in the European population. Larger, more diverse cohorts are needed to accurately determine their role in PD.

Keywords: private variants, Parkinson's disease, *PRKN, PARK7*, *PINK1,* genetics

Introduction

Biallelic variants in *PRKN* (*Parkin*)*, PARK7* (*DJ-1*) and *PINK1* cause early-onset Parkinson's disease (EOPD, defined here as age at onset [AAO] <50 years) (Bonifati 2012; Kilarski et al. 2012; Kasten et al. 2018). The role of heterozygous mutations is more controversial. Several recent large-scale studies did not uncover any association for heterozygous mutations in these genes with PD (Yu et al. 2021; Krohn et al. 2020; Lubbe et al. 2021; Zhu et al. 2021). These studies focused on rare (minor allele frequency [MAF] <1%) damaging variants or established mutations, and largely ignored the specific role of private variants that are only seen in single individuals within a cohort and are absent or very rare in public population databases such as gnomAD (https://gnomad.broadinstitute.org; (Karczewski et al. 2020)). An increased burden of rare (MAF <1%) and private heterozygous *PRKN* variants in late-onset PD (LOPD, AAO ≥50 years) was recently reported (Hopfner et al. 2020). Here, we analyzed heterozygous private variants in *PRKN*, *PARK7* and *PINK1* using four large PD cohorts to assess their role in PD risk.

Methods

Sequencing data were obtained from three independent PD case-control cohorts totaling 5,831 PD cases and 7,221 controls, and summary count data from 9,506 PD cases and 48,207 controls was extracted from the *PD Variant Browser (Kim et al. 2021)* and two of the sequencing cohorts (International Parkinson's disease Genomics Consortium (IPDGC)-whole exome sequencing (WES) and McGill). Cohort information is described in **Supplementary Table 1**. Heterozygous private *PRKN*, *PARK7* and *PINK1* variants were prioritized using, as a guideline, a MAF<1% threshold in >141,000 population-based individuals (Karczewski et al. 2020). Damaging private

variants were identified as described in **Supplementary Methods**. To avoid overlap between the sequencing cohorts and those used to extract allele counts, different analytical methods were used. R package *MetaSKAT* (Lee et al. 2013) and *Metafor* (Viechtbauer 2010)were utilized for meta-analysis of genotyping data and allele counts data, respectively. Bonferroni correction based on the number of genes was applied for multiple testing. See **Supplementary Methods** for more details.

Results

Using available sequencing data, we first performed a gene-based Optimized Sequence-kernel-association test (SKAT-O) (Lee et al. 2012) to assess the burden of all heterozygous private and private damaging variants that yielded no significant association for all genes in individual cohort analysis, nor after meta-analysis (**Table 1**). Similarly, after stratifying PD cases into EOPD and LOPD, no significant enrichment was observed (**Supplementary Tables 2 and 3**).

To further assess the contribution of *PRKN*, *PARK7* and *PINK1* heterozygous private variants, we next extracted allele counts and performed an allelic gene-based Fisher's exact burden test. Again, no significant associations at the individual cohort level were observed for both private and private damaging variants (**Supplementary Table 4**). After meta-analysis, no significant results were found for private variants (**Supplementary Figure 1**) and private potentially damaging variants (**Figure 1**) in the three genes. After stratifying PD cases into EOPD and LOPD, private and private potentially damaging *PRKN* variants demonstrated some level of enrichment in the McGill EOPD New York (NY) cohort (**Supplementary Table 4**), however, these findings no longer remained after meta-analysis (**Supplementary Figures 2-5)**.

Discussion

Here we assessed the contribution of heterozygous private variants within *PRKN*, *PARK7* and *PINK1* to PD risk in several large European ancestry cohorts. Overall, no enrichment across all genes was observed for individual cohorts and upon meta-analysis for both the sequencing and allele count data, and also not after stratifying by early and late onset PD cases. Together this

suggests that private *PRKN*, *PARK7* and *PINK1* variants do not influence PD risk in the European population. While in agreement with previous studies assessing heterozygous mutations (Yu et al. 2021; Krohn et al. 2020; Lubbe et al. 2021; Zhu et al. 2021), our findings are not consistent with those of Hopfner et al. (2020). Differences in genotyping methodologies (pooled DNA sequencing and individual genotype simulation compared to whole-exome, genome, or targeted sequencing in our case) could explain this discordance (see **Supplementary Discussion**). Additionally, cohort characteristics and failure to assess all possible private variants, which can influence *PRKN* burden assessments (Yu et al. 2021; Krohn et al. 2020; Lubbe et al. 2021; Zhu et al. 2021), may also play a role. Despite our study being well powered to detect significant gene-based variant enrichment (power=85.7%; α=0.05; based on 3.4-fold enrichment and gene-based carriership of 0.03% in controls (Hopfner et al. 2020) and global disease prevalence of 0.2% (Keller et al. 2012)), we failed to uncover evidence that private *PRKN*, *PARK7* and *PINK1* variants influence PD risk in the European population. Additional studies of private variants in more diverse populations are needed to further understand their role in PD.

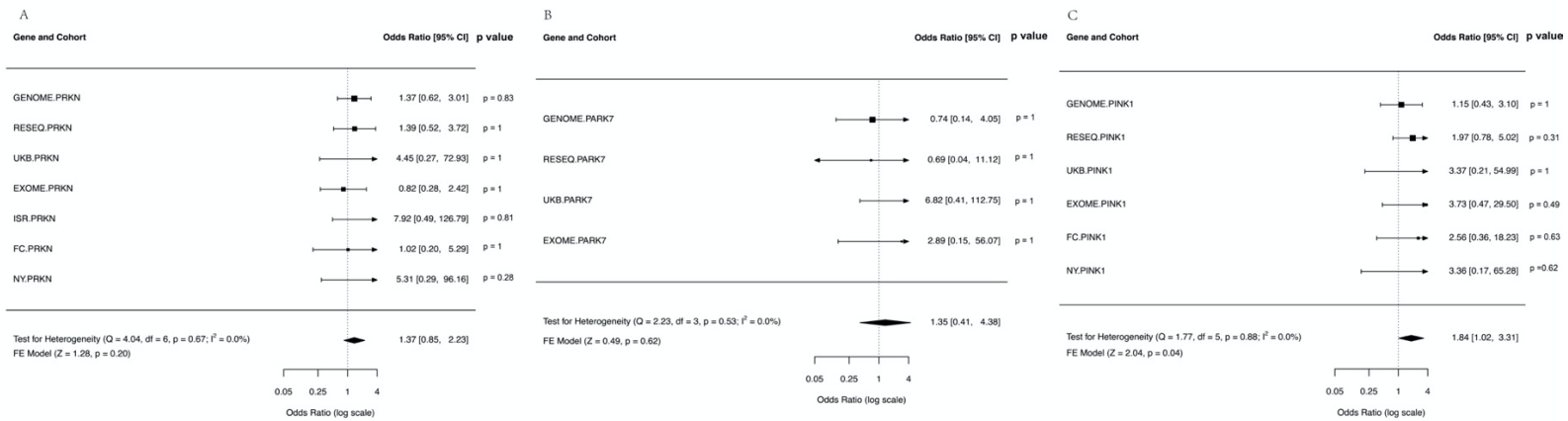

Figure 1. Meta-analysis of carriership burden of private damaging variants in *PRKN*, *PARK7* and *PINK1* genes in Parkinson's disease (PD) using allele count data.

Forest plots assessing private, potentially damaging variants (CADD >20) in (A) *PRKN* (*Parkin*), (B) *PARK7* (*DJ-1*) and (C) *PINK1* genes. Allele counts for each gene were extracted from the *PD Variant Browser,* International Parkinson's disease Genomics Consortium (IPDGC) Whole Exome Sequencing Project cohort and individual McGill cohort*s* separately. Meta-analyses were performed using *Metafor*. Panels show Odds ratio, confidence intervals (CI), corrected Fisher's exact test gene burden p-value per cohort (right side of each plot), and the overall meta-analysis p-value per gene (bottom left of each plot).

Key: CADD=*Combined Annotation Dependent Depletion* phred-scaled score; GENOME=PD Whole Genome Project cohort; RESEQ=IPDGC Resequencing Project cohort; UKB=UK Biobank cohort; EXOME=IPDGC Whole Exome Sequencing Project cohort; ISR=McGill Israel cohort; FC=McGill French/French-Canadian cohort; NY=McGill Columbia University Spot Study; FE=fixed effects; I^2=I-square test for heterogeneity.

Table 1. Assessment of private variants in *PRKN*, *PARK7* and *PINK1* with Parkinson's disease (PD) in cohorts with available sequencing data.

Gene	Variant Group	*SKAT-O* (p-values)					*MetaSKAT*
		AMP-PD	IPDGC-WES	ISR	FC	NY	Meta p-value (corrected)
PRKN	Private	0.41	0.51	0.65	0.81	0.19	0.76 (1.00)
	Private damaging	0.57	0.36	0.65	0.42	0.45	0.57 (1.00)
PINK1	Private	0.13	0.24	-	0.23	0.36	0.10 (0.30)
	Private damaging	0.21	0.24	-	0.23	0.36	0.08 (0.24)
PARK7	Private	0.53	0.26	-	-	0.73	0.44 (1.00)
	Private damaging	0.49	0.37	-	-	-	0.62 (1.00)

Sequence-kernel-association test (SKAT-O) was used to assess variant burden for individual cohort analysis and *MetaSKAT* was used to perform the meta-analysis. Uncorrected P-values for the SKAT-O tests are indicated. Corrected P-values for the meta-analyses are indicated in parentheses.

Key: AMP-PD=Accelerating Medicine Partnership in Parkinson's disease cohort; IPDGC-WES=International Parkinson's disease Genomics Consortium whole exome sequencing cohort; ISR=McGill Israel cohort; FC=McGill French-Canadian cohort; NY=McGill

Columbia University Spot Study.

Reference

Bonifati, Vincenzo. 2012. "Autosomal Recessive Parkinsonism." *Parkinsonism & Related Disorders* 18 Suppl 1 (January): S4–6.

Hopfner, Franziska, Stefanie H. Mueller, Silke Szymczak, Olaf Junge, Lukas Tittmann, Sandra May, Katja Lohmann, et al. 2020. "Private Variants in PRKN Are Associated with Late-Onset Parkinson's Disease." *Parkinsonism & Related Disorders* 75 (June): 24–26.

Karczewski, Konrad J., Laurent C. Francioli, Grace Tiao, Beryl B. Cummings, Jessica Alföldi, Qingbo Wang, Ryan L. Collins, et al. 2020. "The Mutational Constraint Spectrum Quantified from Variation in 141,456 Humans." *Nature* 581 (7809): 434–43.

Kasten, Meike, Corinna Hartmann, Jennie Hampf, Susen Schaake, Ana Westenberger, Eva-Juliane Vollstedt, Alexander Balck, et al. 2018. "Genotype-Phenotype Relations for the Parkinson's Disease Genes Parkin, PINK1, DJ1: MDSGene Systematic Review." *Movement Disorders: Official Journal of the Movement Disorder Society* 33 (5): 730–41.

Kilarski, Laura L., Justin P. Pearson, Victoria Newsway, Elisa Majounie, M. Duleeka W. Knipe, Anjum Misbahuddin, Patrick F. Chinnery, et al. 2012. "Systematic Review and UK-Based Study of PARK2 (parkin), PINK1, PARK7 (DJ-1) and LRRK2 in Early-Onset Parkinson's Disease." *Movement Disorders: Official Journal of the Movement Disorder Society* 27 (12): 1522–29.

Kim, Jonggeol J., Mary B. Makarious, Sara Bandres-Ciga, Jesse Raphael Gibbs, Jinhui Ding, Dena G. Hernandez, Janet Brooks, et al. 2021. "The Parkinson's Disease DNA Variant Browser." *Movement Disorders: Official Journal of the Movement Disorder Society* 36 (5): 1250–58.

Krohn, Lynne, Francis P. Grenn, Mary B. Makarious, Jonggeol Jeffrey Kim, Sara Bandres-Ciga, Dorien A. Roosen, Ziv Gan-Or, et al. 2020. "Comprehensive Assessment of PINK1 Variants in Parkinson's Disease." *Neurobiology of Aging* 91 (July): 168.e1–168.e5.

Lee, Seunggeun, Mary J. Emond, Michael J. Bamshad, Kathleen C. Barnes, Mark J. Rieder, Deborah A. Nickerson, NHLBI GO Exome Sequencing Project—ESP Lung Project Team, David C. Christiani, Mark M. Wurfel, and Xihong Lin. 2012. "Optimal Unified Approach for Rare-Variant Association Testing with Application to Small-Sample Case-Control Whole-Exome Sequencing Studies." *American Journal of Human Genetics* 91 (2): 224–37.

Lee, Seunggeun, Tanya M. Teslovich, Michael Boehnke, and Xihong Lin. 2013. "General Framework for Meta-Analysis of Rare Variants in Sequencing Association Studies." *American Journal of Human Genetics* 93 (1): 42–53.

Lubbe, Steven J., Bernabe I. Bustos, Jing Hu, Dimitri Krainc, Theresita Joseph, Jason Hehir, Manuela Tan, et al. 2021. "Assessing the Relationship between Monoallelic PRKN Mutations and Parkinson's Risk." *Human Molecular Genetics* 30 (1): 78–86.

Viechtbauer, Wolfgang. 2010. "Conducting Meta-Analyses inRwith themetaforPackage." *Journal of Statistical Software* 36 (3). https://doi.org/10.18637/jss.v036.i03.

Yu, Eric, Uladzislau Rudakou, Lynne Krohn, Kheireddin Mufti, Jennifer A. Ruskey, Farnaz Asayesh, Mehrdad A. Estiar, et al. 2021. "Analysis of Heterozygous PRKN Variants and Copy-Number Variations in Parkinson's Disease." *Movement Disorders: Official Journal of the Movement Disorder Society* 36 (1): 178–87.

Zhu, William, Xiaoping Huang, Esther Yoon, Sara Bandres-Ciga, Cornelis Blauwendraat, Joshua H. Cade, Beverly P. Wu, et al. 2021. "Heterozygous PRKN Mutations Are Common but Do Not Increase the Risk of Parkinson's Disease." *bioRxiv*. medRxiv. https://doi.org/10.1101/2021.08.11.21261928.