

<https://doi.org/10.1038/s41531-025-01078-w>

Genome-wide association study of REM sleep behavior disorder in Parkinson's disease



Yuri L. Soser^{1,2}, Karl Heilbron³, Pierre Fontanillas³, Lucy Norcliffe-Kaufmann³, Eric Yu^{1,2}, Uladzislau Rudakou^{1,2}, Jennifer A. Ruskey², Kathryn Freeman², Farnaz Asayesh², Kajsa A. Brolin⁴, Maria Swanberg⁴, Huw R. Morris^{5,6}, Lesley Wu^{5,6}, Raquel Real^{5,6}, Lasse Pihlstrøm⁷, Manuela Tan⁷, Thomas Gasser⁸, Kathrin Brockmann⁸, Hui Liu⁸, Michele T. M. Hu^{9,10}, Donald G. Grosset¹¹, Simon J. G. Lewis¹², John B. Kwok¹², Pau Pastor^{13,14}, Ignacio Alvarez^{13,14}, Matej Skorvanek^{15,16}, Alexandra Lackova^{15,16}, Miriam Ostrozovicova^{15,16}, Mie Rizig¹⁷, 23andMe Research Team^{3*}, The International Parkinson's Disease Genomics Consortium^{*}, Lynne Krohn^{1,2} & Ziv Gan-Or^{1,2,18} ✉

REM sleep behavior disorder (RBD), is a prodromal synucleinopathy affecting a subset of Parkinson's disease (PD) patients and associated with neuropsychiatric symptoms. This study compared the genetic profiles of 13,020 PD patients with probable RBD (PD + RBD) and 5403 without (PD-RBD) using genome-wide association study (GWAS). RBD was assessed by questionnaires or self-reporting. Potential genetic correlations between neuropsychiatric traits and PD + RBD were assessed using linkage disequilibrium score regression. The top variant in the *SNCA* locus was associated with PD + RBD (rs10005233-T, OR = 1.21, 95% CI = 1.16–1.27, $p = 1.81 \times 10^{-15}$). PD risk variants in *SNCA* (rs5019538-G, OR = 0.85, 95% CI = 0.81–0.89, $p = 2.46 \times 10^{-10}$; rs356182-G, OR = 0.89, 95% CI = 0.84–0.95, $p = 0.0001$) and *LRRK2* loci (rs34637584, OR = 0.41, 95% CI = 0.28–0.61, $p = 1.04 \times 10^{-5}$) were associated with reduced PD + RBD risk. A suggestive genetic correlation between attention deficit hyperactivity disorder and PD + RBD was observed but was not statistically significant after correction. These findings highlight genetic distinctions between PD + RBD and PD-RBD, offering insights into PD stratification and potential subtype-specific treatments.

Rapid-eye-movement (REM) sleep behavior disorder (RBD) is a parasomnia characterized by the absence of muscle atonia during REM sleep and dreams enactment¹. When no neurological conditions or other concomitant factors are identified, it is referred to as isolated/idiopathic RBD (iRBD)². iRBD is typically considered a prodromal

stage of synucleinopathies, as about 80%–90% of the cases convert to either Parkinson's disease (PD), dementia with Lewy bodies (DLB) or, more rarely, multiple system atrophy (MSA)^{3,4}. These disorders are all characterized by the accumulation of alpha-synuclein, encoded by the *SNCA* gene⁵. RBD is therefore a key prodromal clinical marker of

¹Department of Human Genetics, McGill University, Montréal, QC, Canada. ²The Neuro (Montreal Neurological Institute-Hospital), McGill University, Montréal, QC, Canada. ³23andMe, Inc., Sunnyvale, CA, USA. ⁴Lund University, Translational Neurogenetics Unit, Department of Experimental Medical Science, Lund, Sweden. ⁵Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, University College London, London, UK. ⁶UCL Movement Disorders Centre, University College London, London, UK. ⁷Department of Neurology, Oslo University Hospital, Oslo, Norway. ⁸Department of Neurodegeneration at Hertie Institute for Clinical Brain Research, Tuebingen, Germany. ⁹Oxford Parkinson's Disease Centre (OPDC), University of Oxford, Oxford, UK. ¹⁰Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK. ¹¹Institute of Neurological Sciences, Queen Elizabeth University Hospital, Glasgow, UK. ¹²Parkinson's Disease Research Clinic, Brain and Mind Centre, School of Medical Sciences, University of Sydney, Sydney, NSW, Australia. ¹³Unit of Neurodegenerative Diseases, Department of Neurology, University Hospital Germans Trias i Pujol and The Germans Trias i Pujol Research Institute (IGTP) Badalona, Barcelona, Spain. ¹⁴Department of Neurology, Hospital Universitari Mutua de Terrassa, Barcelona, Spain. ¹⁵Department of Neurology, Pavol Jozef Šafárik University in Košice, Košice, Slovakia. ¹⁶Department of Neurology, University Hospital of L. Pasteur, Kosice, Slovakia. ¹⁷UCL Queen Square Institute of Neurology, London, UK. ¹⁸Department of Neurology and Neurosurgery, McGill University, Montreal, QC, Canada. *Lists of members and their affiliations appears in the Supplementary Information. ✉e-mail: ziv.gan-or@mcgill.ca

synucleinopathies, and its presence is also associated with a distinctive, more severe clinical presentation. In PD patients with RBD (approximately 25–58% of cases⁶), RBD is associated with a more malignant phenotype, characterized by faster progression⁷ and greater frequency and/or severity of neuropsychiatric manifestations, including cognitive decline, hallucinations, depression, anxiety and apathy^{8–11}. RBD can occur before (then it will be referred to as iRBD), or after the onset of PD and typically PD + RBD cohorts include similar frequencies of individuals who had RBD before and after the onset of RBD⁶.

In recent years, it was shown that the genetic background of iRBD only partially overlaps with that of PD or DLB. Genes such as *GBA1*¹², *TMEM175*¹³ and *SNCA*¹⁴ are important across all conditions^{15,16}, whereas other genes including *LRRK2*¹⁷, *APOE*¹⁸ and familial PD genes¹⁹, seem to not have a major role in iRBD. A recent RBD genome-wide association study (GWAS) identified 5 risk loci associated with RBD²⁰, namely *GBA1*, *TMEM175*, *INPP5F*, *SNCA* and *SCARB2*. Notably, the variants associated with RBD in the *SNCA* and *SCARB2* regions were different and independent to those associated with PD^{15,20}, supporting RBD as a distinctive subtype, with specific genetic and clinical correlates.

In the current study, we aimed to examine whether there are genetic differences between these two sub-groups of patients: PD patients with probable RBD (PD + RBD) and PD patients without RBD (PD-RBD). We used a case-only design, wherein 15 PD cohorts with available data on probable RBD (Table 1) were divided into these two sub-groups of patients. In total, the study included 18,423 patients, composed of 5403 PD + RBD patients and 13,020 PD-RBD patients. To further explore the relationships between RBD and neuropsychiatric manifestations in PD, we performed genetic correlation and Mendelian randomization (MR) analyses using the GWAS summary statistics of the current study and multiple neuropsychiatric conditions.

Results

Genome-wide association study identifies the *SNCA* and *LRRK2* loci as modifiers of risk for RBD in PD

To assess whether genetics can affect the risk of RBD in PD we performed GWAS between PD + RBD ($N = 5403$) and PD-RBD ($N = 13,020$). We

evaluated the genomic inflation using quantile-quantile plots (Q-Q plots) and the lambda factor, showing no inflation ($\lambda = 0.994$, $\lambda_{1000} = 0.999$) (Supplementary Fig. 1).

We found that rs10005233-T, in the 5' region of the *SNCA* locus, was associated with PD + RBD (OR = 1.21, 95% CI = 1.16–1.27, $p = 1.81E-15$, Fig. 1). No secondary signal was detected in the GCTA-COJO analysis at a GWAS significance level. We also examined the 92 variants associated with PD in the most recently published GWAS in Europeans¹⁵ and Asians²¹ (Table 2 and Supplementary Table 1). Using Bonferroni correction based on the number of these variants ($\alpha/\text{number of variants} = 0.00054$), we identified three associations. Two were variants in the *SNCA* locus, which were not on LD with the primary *SNCA* variant rs10005233-T variant nor with each other, whose minor alleles were associated with decreased risk for PD + RBD (rs5019538-G, OR = 0.85, 95% CI = 0.81–0.89, $p = 2.46E-10$ and rs356182-G, OR = 0.89, 95% CI = 0.84–0.95, $p = 0.0001$), and one was the *LRRK2* p.G2019S variant, also associated with a reduced risk for PD + RBD (rs34637584, OR = 0.41, 95% CI = 0.28–0.61, $p = 1.04E-5$, the carrier frequency for this variant across the different cohorts is detailed in Table 3). These three variants were associated with increased risk for PD in the most recent GWAS^{15,21}. *GBA1* variants did not show significant associations with increased risk of RBD in PD (Supplementary Table 2). Additional potential associations in the *SETD1A*, *SPPL2B*, *CRHR1* and *LINC00693* loci should be further studied (Table 2).

Genetic correlation and causative associations between PD with RBD and neuropsychiatric disorders

To examine potential genetic correlations between the risk of RBD in PD and multiple neuropsychiatric conditions, we performed LDSC (Fig. 2 and Supplementary Table 3). We found that genetic factors associated with the presence of RBD in PD are mildly genetically correlated with attention deficit hyperactivity disorder (ADHD, $rg = 0.30$, $SE = 0.14$, $p = 0.04$). The most recently published European PD GWAS was negatively genetically correlated with PD + RBD ($rg = -0.38$, $SE = 0.15$, $p = 0.01$). However, these correlations were not statistically significant after Bonferroni correction ($\alpha = 0.0025$).

To assess possible causative associations between neuropsychiatric conditions and PD + RBD we performed MR using neuropsychiatric disorders as exposures and PD + RBD as the outcome (Supplementary Figs. 2 and 3 and Supplementary Tables 4–7). No test showed a statistically

Table 1 | Demographic characteristics of PD patients in the individual cohorts

Center	PD + RBD, n	Age PD + RBD (SD)	%Fem PD + RBD	PD-RBD, n	Age PD-RBD (SD)	%Fem PD-RBD	Total
Oslo	130	65.6 (8.8)	25%	180	66.1 (10)	44%	310
Lund	365	71.4 (7)	39%	555	70.8 (9.2)	33%	920
McGill	285	67.4 (9.2)	31%	217	67 (8.8)	49%	502
AMP-PD	111	66.4 (9.4)	69%	269	65.5 (10.3)	62%	380
Sydney	105	59 (10.8)	32%	125	60 (10.7)	38%	230
Tuebingen	453	68.8 (9.1)	33%	659	67.7 (10.2)	38%	1112
Barcelona	133	69.5 (9.8)	44%	71	74.3 (10)	66%	204
PRoBaND	585	67.3 (8.9)	30%	1134	67.7 (9.3)	38%	1719
PFP	257	62.4 (12.5)	37%	339	61.4 (13)	47%	596
OPDC	274	67.3 (9.3)	27%	539	67.4 (9.6)	38%	813
Kosice	102	71.4 (8.3)	33%	225	69.2 (10.7)	42%	327
23andMe	2603	NA ^a	38%	8707	NA ^a	45%	11,310
Total	5403	/	/	13,020	/	/	18,423

PD + RBD, n number of Parkinson's disease patients with REM sleep behavior disorder, PD-RBD, n number of Parkinson's disease patients without REM sleep behavior disorder, Age mean age in the group, SD standard deviation, %Fem percentage of females, Tot total number of PD patients in the cohort, Oslo Oslo University Hospital, Lund Lund University, McGill McGill University, AMP-PD Accelerating Medicines Partnership Parkinson's disease, including the New Discovery of Biomarkers (BioFIND), the Harvard Biomarker Study (HBS) and the Parkinson's Disease Biomarkers Program (PDBP) cohorts, Sydney University of Sydney, Tuebingen University of Tuebingen, Barcelona Hospital Universitari Mutua de Terrassa, PRoBaND Parkinson's repository of biosamples and networked datasets, PFP Parkinson's Families Project, OPDC Oxford Parkinson's Disease Centre, Kosice Pavol Jozef Šafárik University in Kosice.

^a23andMe only provides age ranges (i.e., in cases: 5 individuals <30 years of age, 49 individuals 30–45, 346 individuals 45–60, 2203 individuals >60; in controls: 11 individuals <30, 146 individuals 30–45, 1302 individuals 45–60, 7248 individuals >60).

significant causative association between neuropsychiatric traits and PD + RBD. However, our power for this analysis was suboptimal (35.7%), therefore there could be associations that we could not detect. We were not able to conduct reverse MR using PD + RBD as the exposure since only one locus passed GWAS significance, preventing us from performing appropriate sensitivity analyses.

Discussion

In the current GWAS, we found that variants in the *SNCA* and *LRRK2* loci may be associated with the risk of RBD in PD. Additional loci (*SETD1A*, *SPPL2B*, *CRHR1* and *LINC00693*) require further studies to examine whether they have a role in PD + RBD. The top variant in the *SNCA* locus, rs10005233-T, was previously reported to be associated with iRBD in a candidate gene study (OR = 1.43, 95% CI = 1.27–1.62, $p = 1.1 \times 10^{-8}$)¹⁴. Another study, using the Oslo and Parkinson's Progression Marker Initiative cohorts, found another variant in the *SNCA* locus associated with

PD + RBD (rs3756063), which is in strong linkage disequilibrium (LD) with rs10005233-T ($D' = 0.97$, $r^2 = 0.91$)²². Furthermore, rs10005233-T is in LD with other 5' region *SNCA* variants associated with synucleinopathies, including rs7681440 ($D' = 0.99$, $r^2 = 0.94$), associated with DLB²³, rs763443 ($D' = 0.89$, $r^2 = 0.78$), a secondary PD GWAS signal^{14,15,24}, as well as rs2583988 ($D' = 0.99$, $r^2 = 0.40$), a variant located in the *SNCA*-*AS1* region (discussed below) and associated with Lewy body variant of Alzheimer's disease (ADLBV)²⁵. It is still unclear whether it is a specific variant in the *SNCA* locus or the presence of a specific *SNCA* haplotype that drives these associations with cognitive phenotypes across synucleinopathies¹⁴. The rs10005233 variant is also in LD ($D' = 0.97$, $r^2 = 0.91$) with the top signal of a recently published RBD GWAS, rs3756059²⁰, which was associated with reduced expression of *SNCA*-*AS1*, an antisense RNA molecule that could potentially reduce the translation of alpha-synuclein when it is over-expressed or increase the translation of alpha-synuclein when it is down-regulated. Notably, this reduced expression of *SNCA*-*AS1* is mainly in cortical areas²⁰, thus potentially increasing alpha-synuclein levels and exposing the cerebral cortex to a greater risk of neurodegeneration in carriers of this RBD-associated variant. The latter hypothesis should be tested in relevant animal models. Altogether, these data suggest that, depending on possible region-specific effects, different *SNCA* variants might play different roles in synucleinopathies.

We found that three of the 92 PD GWAS signals associated with increased PD risk in Europeans and Asians^{15,21,26}, the *LRRK2* variant p.G2019S and the *SNCA* variants rs5019538 and rs356182, were less frequent in PD + RBD compared to PD-RBD. The association between p.G2019S and PD-RBD is in line with a previously reported reduced frequency of RBD in PD patients carrying this variant (OR = 0.49, 95% CI 0.39–0.61, $p < 0.001$)^{27,28}, and with lack of p.G2019S carriers in about 1000 iRBD patients in another study¹⁹. In addition to a reduced occurrence of RBD, carriers of the p.G2019S *LRRK2* variant also present an overall more benign phenotype, including less frequent and milder cognitive decline^{29,30}. These findings, together with the negative correlation between the presence of RBD in PD and the most recent PD GWAS in Europeans¹⁵ may suggest that overall, the PD-RBD subgroup is more genetically similar to PD than the subgroup of PD + RBD. However, the latter group may still also be genetically similar to PD in general, and in order to perform proper genetic correlation between PD + RBD, PD-RBD and PD, each of these groups should be compared to separate control groups, followed by re-analyses of genetic correlations.

These findings further support a pathophysiological relationship between the manifestation of RBD in PD and cognitive decline, which is in line with the comorbidity of these two clinical entities. The rs10005233-T *SNCA* variant is associated with both RBD and increased cognitive decline, while the *LRRK2* p.G2019S variant is associated with reduced risk of RBD and reduced cognitive decline. These findings may suggest a potential pathophysiological relationship between the manifestation of RBD in PD and cognitive decline, which at least in part is affected by genetics. It was

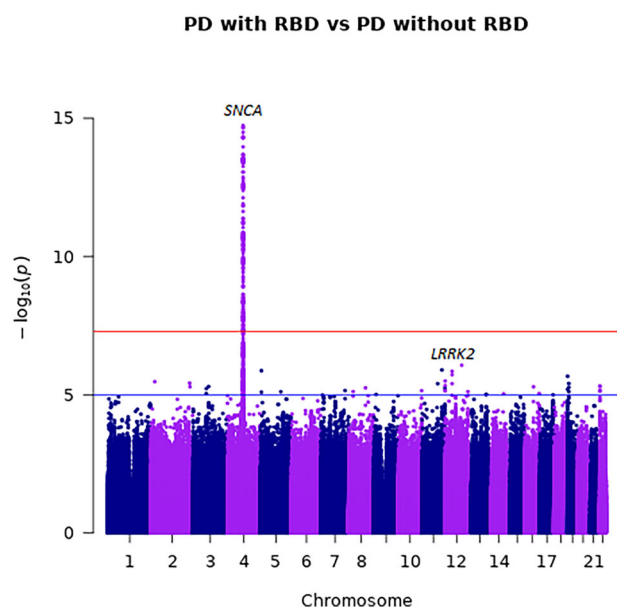


Fig. 1 | Manhattan plot of PD with RBD vs. PD without RBD. A Manhattan plot showing the results of the GWAS meta-analysis, comparing PD with RBD and PD without RBD, highlighting the *SNCA* and *LRRK2* loci. The Y axis represents the negative logarithm of p value, the X axis represents the chromosomal position of the variants and each dot on the figure represents a SNP. The red line represents the genome-wide Bonferroni-corrected statistical significance threshold (5×10^{-8}), whereas the blue line is the false-discovery rate-corrected significance threshold (5×10^{-5}). Chr chromosome, PD Parkinson's disease, RBD REM sleep behavior disorder.

Table 2 | Among PD-variants, associations with RBD in PD compared to prior PD-GWAS associations

Variant	Nearest gene	PD with and without RBD		PD ^{15,37}	
		OR (95% CI)	p value	OR (95% CI)	p value
4:90636630	<i>SNCA</i>	0.85 (0.81–0.89)	2.46×10^{-8}	1.17 (1.14–1.2)	1.13×10^{-36}
12:40734202	<i>LRRK2</i>	0.41 (0.28–0.61)	1.04×10^{-5}	11.35 (9.44–13.63)	3.61×10^{-148}
4:90626111	<i>SNCA</i>	0.89 (0.84–0.95)	0.0001365^a	1.32 (1.29–1.35)	3.89×10^{-154}
16:30977799	<i>SETD1A</i>	0.92 (0.88–0.97)	0.001952	1.09 (1.07–1.12)	5.12×10^{-20}
19:2341047	<i>SPPL2B</i>	1.09 (1.03–1.15)	0.002506	0.93 (0.91–0.95)	4.18×10^{-10}
17:43798308	<i>CRHR1</i>	1.19 (1.05–1.36)	0.008267	0.79 (0.75–0.84)	6.71×10^{-16}
3:28705690	<i>LINC00693</i>	0.95 (0.9–0.99)	0.02552	1.07 (1.05–1.09)	8.09×10^{-12}

PD Parkinson's disease, RBD REM sleep behavior disorder, OR (95% CI) odds ratio with relative 95% confidence interval.

^aVariant statistically significant after Bonferroni correction ($a/\text{number of variants} = 0.00054$).

Table 3 | Carriers of the *LRRK2* p.G2019S variant across different cohorts

Cohort	Non-carriers PD-RBD, N	Carriers PD-RBD, N (%)	Non-carriers PD + RBD, N	Carriers PD + RBD, N (%)
Oslo	156	1 (0.64%)	122	1 (0.81%)
Lund	552	3 (0.54%)	364	1 (0.27%)
McGill	212	5 (2.30%)	282	3 (1.05%)
AMP-PD	258	9 (3.37%)	110	1 (0.90%)
Sydney	115	1 (0.86%)	98	1 (1.01%)
Tuebingen	365	0 (0.00%)	641	0 (0.00%)
Barcelona	45	0 (0.00%)	100	0 (0.00%)
PRoBaND	1132	2 (0.18%)	582	3 (0.51%)
PFP	328	9 (2.67%)	246	6 (2.38%)
OPDC	537	2 (0.37%)	274	0 (0.00%)
Kosice	216	1 (0.46%)	101	0 (0.00%)
23andMe	8459	248 (2.85%)	2584	19 (0.73%)
Total ^a	12,010	281 (2.29%)	4863	35 (0.71%)

PD-RBD participants without REM sleep behavior disorder (RBD), PD + RBD participants with RBD, N number of participants, % percentage of participants, Oslo Oslo University Hospital, Lund Lund University, McGill McGill University, AMP-PD Accelerating Medicines Partnership Parkinson's disease, including the New Discovery of Biomarkers (BioFIND), the Harvard Biomarker Study (HBS) and the Parkinson's Disease Biomarkers Program (PDBP) cohorts, Sydney University of Sydney, Tuebingen University of Tuebingen, Barcelona Hospital Universitari Mutua de Terrassa, PRoBaND Parkinson's repository of biosamples and networked datasets, PFP Parkinson's Families Project, OPDC Oxford Parkinson's Disease Centre, Kosice Pavol Jozef Šafárik University in Kosice.

^aThe total excludes individuals with unknown carrier status for p.G2019S.

hypothesized that this clinical and pathophysiological correlation could reflect the two alternative directions of alpha-synuclein spreading, body-first or brain-first³¹. In body-first PD, alpha-synuclein pathology may start in the enteric nervous system, whereas in brain-first PD it may arise in the amygdala, entorhinal cortex and substantia nigra. These different neuropathological patterns correspond to two different subgroups of clinical progression. In body-first PD, RBD may manifest before the motor PD symptoms, and cognitive decline occurs faster, whereas in brain-first PD, RBD may occur after the onset of motor PD symptoms, if at all, and cognitive impairment develops more slowly^{31–34}. We can therefore speculate that the SNCA rs1005233 variant associated with PD + RBD might also be associated with the body-first subtype of PD, whereas the LRRK2 p.G2019S variant might be associated with the brain-first PD subtype, with less frequent RBD and milder cognitive decline. Since we cannot determine in our data which patients had RBD prior to PD diagnosis and which had it after PD diagnosis, this hypothesis should be studied in future genetic analyses of brain-first vs. body-first PD.

Similar to previous reports in iRBD and PD + RBD^{14,15,22}, in this study we did not observe any involvement of *APOE* variants in PD + RBD, suggesting that this gene does not affect RBD risk in PD patients. The rs117615688 variant (chromosomal position 17:43798308) in the *CRHR1* gene, located in the *MAPT* locus, was nominally associated with RBD (OR = 1.19, 95% CI = 1.05–1.36, $p = 0.008$) with an opposite direction of effect to that seen in PD (OR = 0.79, 95% CI = 0.75–0.84, $p = 6.71\text{E}^{-16}$) (Table 2).

There are several limitations in this study. All participants were Europeans, therefore our results might not fully apply to other ancestries. We cannot completely rule out survival bias as a confounder of our results, although PD patients with and without RBD had similar age. In addition, although we included a large number of patients with PD, insufficient power in our analysis might explain the lack of causative associations between PD + RBD and neuropsychiatric traits as well as of genome-wide significance of the LRRK2 p.G2019S and *GBA1* variants. It is possible that *GBA1* variants are strongly implicated also in the PD subtype without RBD, thus counterbalancing their previously reported

contribution to RBD risk^{12,20}. Another limitation is represented by the inclusion of patients who developed RBD both before and after PD, as they may represent body-first vs. brain-first subtypes of PD as discussed above. Future research with larger sample sizes could investigate possible genetic and biological differences between them and specifically differentiate brain-first and body-first PD in that sense. The aim of the current study was limited to the genetic differences between PD + RBD and PD-RBD. Future studies may compare each of these sub-groups to healthy controls, or to other subgroups, and other synucleinopathies such as DLB and MSA. Such analyses may also shed more light on the genetic basis of the different synucleinopathies. Additional studies that examine gene-gene and gene-environment interactions will also provide crucial information, but such studies will require much larger sample sizes than currently available.

Since the sex distribution between PD + RBD and PD-RBD is different, sex-stratified analysis is warranted in larger, better-powered studies. Some individuals with PD-RBD could still develop RBD, which is why adjustment for age has been performed. Yet this fact could still dilute the model and the effects of specific variants. There is more than one way to interpret the findings of our study. The most straightforward interpretation is that PD + RBD and PD-RBD largely share the same genetic background, which is common to PD in general, and individual factors such as SNCA and LRRK2 variants can tip the balance to one or the other subtype. However, other scenarios are also possible, in which the different phenotypes have more distinct genetic architecture, which in the present study design we could not yet identify. This potential interpretation may be consistent with the genetic correlation analyses, but much larger studies with additional complementary genetic analyses are needed to examine this possibility. It is also possible that while some PD-associated variants are less common in PD + RBD than in PD-RBD, they are still more common in PD + RBD than in the healthy population, and this should be explored in a separate study.

In conclusion, in this study we demonstrated that the risk of PD + RBD may be modified by variants in the SNCA and LRRK2 loci, and potentially other loci. These genetic associations may explain why cognitive decline is more frequently observed in PD + RBD compared to PD-RBD, with possible implications for therapeutic management of PD patients. Future research will need to further explore the relationship between genetics, biology and clinical comorbidities to define PD subtypes and implement a precision medicine guided by early markers.

Methods

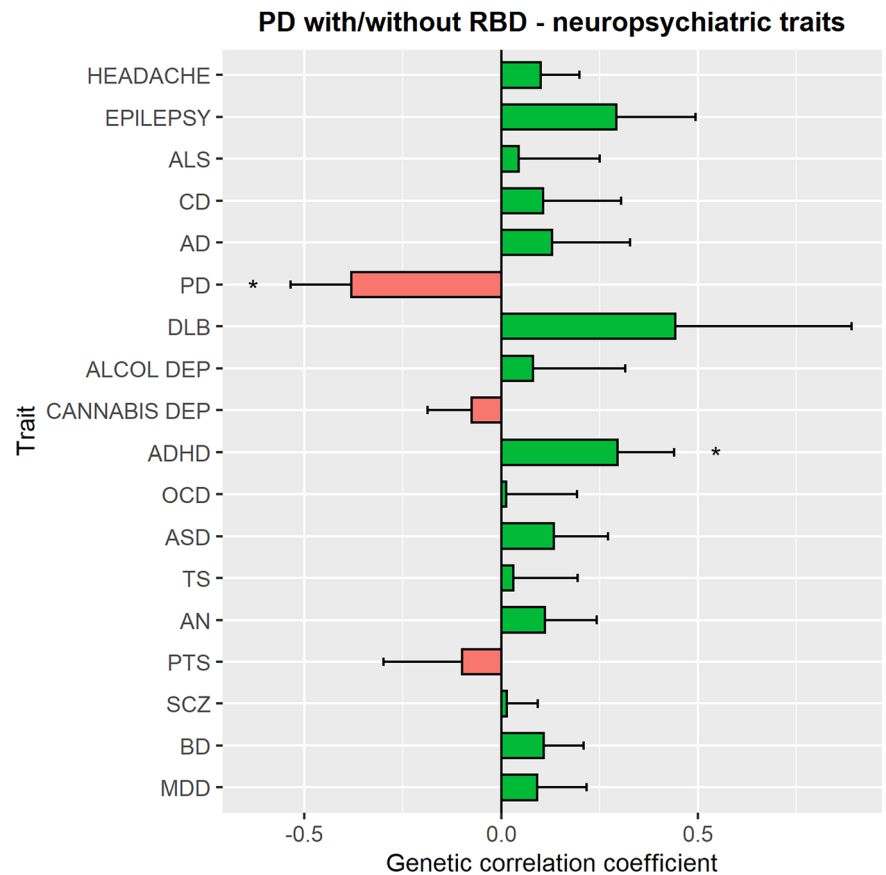
Study design

The aim of this study is to examine whether there are differences in frequencies of genetic variants when comparing PD + RBD to PD-RBD. Such differences might highlight specific genes, variants and pathways that are more involved in one sub-type compared to the other. For this purpose, we performed a case-only GWAS in 15 cohorts with available data on Probable RBD (Table 1). We performed a GWAS in each cohort separately, comparing PD + RBD and PD-RBD, followed by a meta-analysis. Specific details on the cohort, quality control and analysis are detailed below.

Population

The study population included 18,423 PD patients (detailed in Table 1), of whom 5403 had probable RBD (PD + RBD) and were treated as cases, whereas 13,020 did not (PD-RBD) and were treated as controls. Probable RBD was defined using either the RBD single-question screen (RBD1Q)³⁵ or the RBD screening questionnaire (RBDSQ)³⁶, both of which show high sensitivity and specificity in PD patients³⁷. We refer to iRBD when RBD occurs prior to the neurodegeneration and to RBD for subjects with RBD regardless of the time of onset of neurodegeneration. PD was diagnosed by movement disorder specialists according to the UK Brain Bank³⁸ or International Parkinson Disease and Movement Disorders Society criteria³⁹. The 23andMe cohort had self-reported a diagnosis of PD as well as RBD and/or dream enactment behaviors.

Fig. 2 | Genetic correlation between PD with RBD and neuropsychiatric traits. The bar plot shows the genetic correlations between PD with RBD and neuropsychiatric traits. The correlation coefficient is illustrated on the X axis. Green bars represent positive correlations whereas red bars negative ones (i.e., a positive correlation of the neuropsychiatric trait with PD without RBD). The asterisks highlight the nominally significant correlations. ALS amyotrophic lateral sclerosis, CD cognitive decline, AD Alzheimer's disease, PD Parkinson's disease, DLB dementia with Lewy bodies, Alcohol dep alcohol dependence, cannabis dep cannabis dependence, ADHD attention deficit hyperactivity disorder, OCD obsessive-compulsive disorder, ASD autism spectrum disorder, TS Tourette syndrome, AN anorexia nervosa, PTS post-traumatic syndrome, SCZ schizophrenia, BD bipolar disorder, MDD major depressive disorder.



The participants were of European ancestry and their clinical and genetic data were collected from 15 different cohorts (Table 1), 11 of which are from the International Parkinson's Disease Genomics Consortium (IPDGC), three cohorts are from the Accelerating Medicines Partnership Parkinson's disease (AMP-PD, <https://amp-pd.org/>) and one cohort was collected and analyzed by 23andMe Inc. (<https://www.23andme.com/research/>). The Central European Group on Genetics of Movement Disorders (CEGEMOD) contributed to the Kosice cohort.

Ethical compliance statement

IRB Study Number A11-M60-21A (21-11-023) was reviewed and approved by the Research Ethics Offices (REOs). Informed written patient consent was provided in each center before the inclusion of each in the study.

Genetic analysis

In the 23andMe cohort, participants provided informed consent and volunteered to participate in the research online, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent (E&I) Review Services. As of 2022, E&I Review Services is part of Salus IRB (<https://www.versiticlinicaltrials.org/salusirb>). DNA extraction and genotyping were performed on saliva samples by National Genetics Institute (NGI), a CLIA-licensed clinical laboratory and a subsidiary of Laboratory Corporation of America. Samples were genotyped on one of five genotyping platforms. The v1 and v2 platforms were variants of the Illumina HumanHap550 + BeadChip, including about 25,000 custom single nucleotide polymorphisms (SNPs) selected by 23andMe, with a total of about 560,000 SNPs. The v3 platform was based on the Illumina OmniExpress+ BeadChip, with custom content to improve the overlap with the 23andMe v2 array, with a total of about 950,000 SNPs. The v4 platform was a fully customized array, including a lower redundancy subset of v2 and v3 SNPs with additional coverage of lower-frequency coding variation, and about 570,000 SNPs. The v5 platform, in current use, is an

Illumina Infinium Global Screening Array (~640,000 SNPs) supplemented with ~50,000 SNPs of custom content. This array was specifically designed to better capture global genetic diversity and to help standardize the platform for genetic research. Samples that failed to reach 98.5% call rate were re-analyzed. Individuals whose analyses failed repeatedly were re-contacted by 23andMe customer service to provide additional samples.

Participants were restricted to European ancestry through an analysis of local ancestry⁴⁰. A support vector machine (SVM) to classify individual haplotypes into one of 31 reference populations was used (<https://www.23andme.com/ancestry-composition-guide/>). The SVM classifications are then fed into a hidden Markov model (HMM) that accounts for switch errors and incorrect assignments, and gives probabilities for each reference population in each window. Finally, we used simulated admixed individuals to recalibrate the HMM probabilities so that the reported assignments are consistent with the simulated admixture proportions. A maximal set of unrelated individuals was chosen for each analysis using a segmental identity-by-descent (IBD) estimation algorithm⁴¹.

We phased participant data using either an internally-developed tool, Finch (V1-V4 genotyping arrays) or Eagle2 (V5 genotyping array)⁴². Finch implements the Beagle haplotype graph-based phasing algorithm, modified to separate the haplotype graph construction and phasing steps⁴³. It extends the Beagle model to accommodate genotyping error and recombination, to handle cases where there are no consistent paths through the haplotype graph for the individual being phased. We constructed haplotype graphs for European and non-European samples on each 23andMe genotyping platform from a representative sample of genotyped individuals, and then performed out-of-sample phasing of all genotyped individuals against the appropriate graph. For the X-chromosome, we built separate haplotype graphs for the non-pseudoautosomal region and each pseudoautosomal region, and these regions were phased separately.

Imputation panels created by combining multiple smaller panels have been shown to give better imputation performance than the individual

constituent panels alone⁴⁴. To that end, we combined the May 2015 release of the 1000 Genomes Phase 3 haplotypes with the UK10K imputation reference panel to create a single unified imputation reference panel^{45,46}. Multiallelic sites with N alternate alleles were split into N separate biallelic sites. We then removed any site whose minor allele appeared in only one sample. For each chromosome, we used Minimac3 to impute the reference panels against each other, reporting the best-guess genotype at each site⁴⁷. This gave us calls for all samples over a single unified set of variants. We then joined these together to get, for each chromosome, a single VCF with phased calls at every site for 6,285 samples.

In preparation for imputation, we split each chromosome of the reference panel into chunks of no more than 300,000 variants, with overlaps of 10,000 variants on each side. We used a single batch of 10,000 individuals to estimate Minimac3 imputation model parameters for each chunk⁴⁷. We imputed phased participant data against the chunked merged reference panel using Minimac3, treating males as homozygous pseudo-diploids for the non-pseudoautosomal region. Throughout, we treated structural variants and small indels the same as SNPs.

We excluded SNPs that: (1) had a MAF < 0.01, (2) had a call rate < 90%, (3) had a Hardy-Weinberg $p < 10^{-20}$ in people with predominantly European ancestry, (4) were only genotyped on the V1 and/or V2 platforms, (5) were found on the mitochondrial chromosome or the Y-chromosome, (6) failed a test for parent-offspring transmission (specifically, we regressed the child's allele count against the mean parental allele count and excluded SNPs with fitted < 0.6 and $p < 10^{-20}$ for a test of < 1), (7) had an association with genotype date ($p < 10^{-50}$ by ANOVA of SNP genotypes against a factor dividing genotyping date into 20 roughly equal-sized buckets), (8) had a large sex effect (ANOVA of SNP genotypes, $r^2 > 0.1$), or (9) had probes matching multiple genomic positions in the reference genome.

We excluded SNPs with imputed $r^2 < 0.3$, as well as SNPs that had strong evidence of a platform batch effect. For each SNP we identified the largest sub-set of the data passing other quality control criteria based on their original genotyping platform—either v2 + v3 + v4 + v5, v4 + v5, v4, or v5 only—and computed association test results for the largest passing set. The batch effect test is an F test from an ANOVA of the SNP dosages against a factor representing the V4 or V5 platform; we excluded results with $p < 10^{-50}$.

Across both genotyped and imputed GWAS results, we excluded SNPs that had sample size of less than 20% of the total GWAS sample size. We also removed SNPs that did not converge during logistic regression, as identified by $\text{abs}(\text{effect}) > 10$ or $\text{stderr} > 10$ on the log-odds scale. If SNPs were both genotyped and imputed, and they passed QC for both, we used results from the imputed analysis. After quality control, we had analyzed 904,040 genotyped SNPs and 25,208,208 imputed SNPs.

GWAS was performed using logistic regression adjusted for age, sex, top five principal components as well as the genotype platform to account for genotype batch effects. The significance threshold was set at $p < 5 \times 10^{-8}$.

In the other centers, genotyping was performed using the OmniExpress, NeuroX or Global Screening (GSA) GWAS array according to the manufacturer's instructions (Illumina Inc.). Parkinson's Families Project (PFP) was genotyped with NeuroChip, Parkinson's repository of bio-samples and networked datasets (PROBaND) with HumanCoreExome array, and Oxford Parkinson's Disease Centre (OPDC) with either HumanCoreExome-12 v.1.1 or Infinium HumanCoreExome-24v.1.1 arrays. Quality control was performed following standard pipelines (detailed in <https://github.com/neurogenetics/GWAS-pipeline>) using plink 1.9⁴⁸. In brief, we filtered out heterozygosity outliers using an F -statistic cut-off of < -0.15 or > 0.15 . Individuals with a variant call rate < 95% and sex mismatch were excluded. Variants missing in > 5% of the participants, with disparate missingness between cases and controls ($p < 1 \times 10^{-4}$), or significantly deviating from the Hardy-Weinberg equilibrium in controls ($p < 1 \times 10^{-4}$) were also removed. We used GCTA to check for relatedness closer than first cousins between participants ($\text{pihat} > 0.125$). We performed imputation using the Michigan imputation server (<https://imputationserver.sph.umich.edu/index.html#>) with the Haplotype Reference Consortium reference panel r1.1 2016 under default settings. Ancestry outliers were detected using HapMap3 principal component analysis (PCA) data in R version 4.0.1. After imputation, we selected variants with $R^2 > 0.8$ and a minor allele frequency (MAF) > 0.01, while retaining variants that have strong pathogenic implications in PD (i.e., the LRRK2 p.G2019S variant and the GBA1 p.N370S, p.E326K and p.T369M variants). After QC filtering, a total of 9,979,381 SNPs were analyzed in the GWAS for these cohorts.

Statistical analysis

To test for genetic associations to RBD in PD, we performed GWAS using logistic regression comparing PD + RBD and PD-RBD adjusted for age at RBD questionnaire administration, sex and principal components. The significance threshold was set at $p < 5 \times 10^{-8}$. The analyses were performed separately in each cohort and the results were then meta-analyzed with a fixed-effect model using METAL (https://genome.sph.umich.edu/wiki/METAL_Documentation)⁴⁹. To identify any possible secondary associations hidden by the principal signals of the GWAS, we also performed Conditional and Joint-Genome-wide Complex Trait Analysis (COJO-GCTA), a method that harnesses a conditional stepwise regression approach to identify independent associations (<https://yanglab.westlake.edu.cn/software/gcta/#Overview>)⁵⁰.

Genetic correlation

To investigate the potential genetic correlation between the presence of RBD in PD and known neuropsychiatric conditions we used linkage-disequilibrium score regression (LDSC) on LDHub (<http://ldsc.broadinstitute.org/ldhub/>)⁵¹. The neuropsychiatric traits we analyzed include epilepsy, headache, amyotrophic lateral sclerosis, cognitive decline, Alzheimer's disease, Parkinson's disease, dementia with Lewy bodies, alcohol dependence, cannabis dependence, attention deficit hyperactivity disorder, Tourette syndrome, anorexia nervosa, post-traumatic syndrome, schizophrenia, bipolar disorder, obsessive-compulsive disorder, autism spectrum disorder and major depressive disorder. Summary statistics for the compared traits were accessed through the LDHub platform or downloaded from publicly available sources, then formatted and analyzed using LDHub python v2.7 scripts (<https://github.com/bulik/ldsc/wiki/>). Positive genetic correlations indicate positive association with genetic factors associated with RBD among individuals with PD.

Mendelian randomization

To assess any possible causal association between neuropsychiatric disorders and the presence of RBD in PD we performed Mendelian randomization (MR)⁵². The neuropsychiatric traits for this analysis were selected based on their known clinical relevance to RBD or PD, as they have been reported in either RBD patients, PD patients or both. In brief, this method harnesses summary statistics from an exposure (the neuropsychiatric traits, in this case) and an outcome (the presence of RBD in PD) and uses the statistically significant variants from the former as instrumental variables (IVs) to infer a potential causative association with the latter. This approach mimics randomized control trials, since genetics is randomly assigned at conception and unaffected by the environment^{53–56}. Differently from randomized control trials, however, MR relies on certain restrictive assumptions, varying based on the specific methods used, like the absence of horizontal pleiotropy and others, as well as on the quality of the GWASs they rely on. The neuropsychiatric traits for this analysis were selected based on relevance to RBD comorbidities, known neuropsychiatric manifestations in PD or with clinical relevance to PD. They include Alzheimer's disease, dementia with Lewy Bodies, schizophrenia, major depressive disorder and bipolar disorder. We used the TwoSampleMR R package (<https://mrcieu.github.io/TwoSampleMR/>)⁵⁷ to perform MR analyses, including sensitivity analyses, tests assessing pleiotropy and heterogeneity between IVs, in R version 4.0.1 according to protocols previously established⁵⁸. Sensitivity analyses included MR Egger, inverse variance weighted (IVW), weighted

median, simple mode and weighted mode. Steiger filtering was also performed to check for reverse causality. Summary statistics were downloaded by the MRBase GWAS catalog (<http://www.mrbase.org/>) and the Psychiatric Genomics Consortium (<https://pgc.unc.edu/>) publicly available database. To calculate the power to detect an odds ratio = 1.2 we used an online Mendelian Randomization power calculation tool (<https://sb452.shinyapps.io/power/>)⁵⁹.

Data availability

The full PD with and without RBD GWAS summary statistics is detailed on the GWAS catalog (<https://www.ebi.ac.uk/gwas/>). The full GWAS summary statistics for the 23andMe dataset will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit <https://research.23andme.com/collaborate/#dataset-access/> for more information and to apply to access the data. The GWAS summary statistics for the neuropsychiatric traits used in the study are available on the GWAS catalog and Psychiatric Genomics Consortium (<https://pgc.unc.edu/>).

Code availability

The codes used for the analyses are available on https://github.com/daskrohn/RBD_GWAS and <https://github.com/gan-orlab>.

Received: 30 June 2025; Accepted: 15 July 2025;

Published online: 25 September 2025

References

1. St Louis, E. K. & Boeve, B. F. REM sleep behavior disorder: diagnosis, clinical implications, and future directions. *Mayo Clin. Proc.* **92**, 1723–1736 (2017).
2. Högl, B., Stefani, A. & Videnovic, A. Idiopathic REM sleep behaviour disorder and neurodegeneration—an update. *Nat. Rev. Neurol.* **14**, 40–55 (2018).
3. Postuma, R. B. et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain* **142**, 744–759 (2019).
4. Postuma, R. B. et al. Quantifying the risk of neurodegenerative disease in idiopathic REM sleep behavior disorder. *Neurology* **72**, 1296–1300 (2009).
5. Ayers, J. I. et al. Different α -synuclein prion strains cause dementia with Lewy bodies and multiple system atrophy. *Proc. Natl. Acad. Sci. USA* **119**, e2113489119 (2022).
6. Hu, M. T. REM sleep behavior disorder (RBD). *Neurobiol. Dis.* **143**, 104996 (2020).
7. Barasa, A., Wang, J. & Dewey, R. B. Jr Probable REM sleep behavior disorder is a risk factor for symptom progression in Parkinson disease. *Front. Neurol.* **12**, 651157 (2021).
8. Diaconu, Ș., Falup-Pecurariu, O., Țiņ, D. & Falup-Pecurariu, C. REM sleep behaviour disorder in Parkinson's disease (Review). *Exp. Ther. Med.* **22**, 812 (2021).
9. Bargiotas, P. et al. Apathy in Parkinson's disease with REM sleep behavior disorder. *J. Neurol. Sci.* **399**, 194–198 (2019).
10. Liu, Y. et al. Longitudinal changes in Parkinson's disease symptoms with and without rapid eye movement sleep behavior disorder: the Oxford Discovery Cohort Study. *Mov. Disord.* **36**, 2821–2832 (2021).
11. Duarte Folle, A., Paul, K. C., Bronstein, J. M., Keener, A. M. & Ritz, B. Clinical progression in Parkinson's disease with features of REM sleep behavior disorder: a population-based longitudinal study. *Parkinsonism Relat. Disord.* **62**, 105–111 (2019).
12. Krohn, L. et al. GBA variants in REM sleep behavior disorder: a multicenter study. *Neurology* **95**, e1008–e1016 (2020).
13. Krohn, L. et al. Genetic, structural, and functional evidence link TMEM175 to synucleinopathies. *Ann. Neurol.* **87**, 139–153 (2020).
14. Krohn, L. et al. Fine-mapping of SNCA in rapid eye movement sleep behavior disorder and overt synucleinopathies. *Ann. Neurol.* **87**, 584–598 (2020).
15. 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).
16. Chia, R. et al. Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. *Nat. Genet.* **53**, 294–303 (2021).
17. Ouled Amar Bencheikh, B. et al. LRRK2 protective haplotype and full sequencing study in REM sleep behavior disorder. *Parkinsonism Relat. Disord.* **52**, 98–101 (2018).
18. Gan-Or, Z. et al. The dementia-associated APOE ϵ 4 allele is not associated with rapid eye movement sleep behavior disorder. *Neurobiol. Aging* **49**, 218.e13–e15 (2017).
19. Mufti, K. et al. Comprehensive analysis of familial Parkinsonism genes in rapid-eye-movement sleep behavior disorder. *Mov. Disord.* **36**, 235–240 (2021).
20. Krohn, L. et al. Genome-wide association study of REM sleep behavior disorder identifies polygenic risk and brain expression effects. *Nat. Commun.* **13**, 7496 (2022).
21. Foo, J. N. et al. Identification of risk loci for Parkinson disease in Asians and comparison of risk between Asians and Europeans: a genome-wide association study. *JAMA Neurol.* **77**, 746–754 (2020).
22. Björnarå, K. A., Pihlström, L., Dietrichs, E. & Toft, M. Risk variants of the α -synuclein locus and REM sleep behavior disorder in Parkinson's disease: a genetic association study. *BMC Neurol.* **18**, 20 (2018).
23. Guerreiro, R. et al. Investigating the genetic architecture of dementia with Lewy bodies: a two-stage genome-wide association study. *Lancet Neurol.* **17**, 64–74 (2018).
24. Chang, D. et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat. Genet.* **49**, 1511–1516 (2017).
25. Linnertz, C. et al. The genetic contributions of SNCA and LRRK2 genes to Lewy body pathology in Alzheimer's disease. *Hum. Mol. Genet.* **23**, 4814–4821 (2014).
26. Nalls, M. A. et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat. Genet.* **46**, 989–993 (2014).
27. Pont-Sunyer, C. et al. Sleep disorders in Parkinsonian and nonparkinsonian LRRK2 mutation carriers. *PLoS ONE* **10**, e0132368 (2015).
28. Huang, J., Cheng, Y., Li, C. & Shang, H. Genetic heterogeneity on sleep disorders in Parkinson's disease: a systematic review and meta-analysis. *Transl. Neurodegener.* **11**, 21 (2022).
29. Alcalay, R. N. et al. Neuropsychological performance in LRRK2 G2019S carriers with Parkinson's disease. *Parkinsonism Relat. Disord.* **21**, 106–110 (2015).
30. Healy, D. G. et al. Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. *Lancet Neurol.* **7**, 583–590 (2008).
31. Borghammer, P. et al. Neuropathological evidence of body-first vs. brain-first Lewy body disease. *Neurobiol. Dis.* **161**, 105557 (2021).
32. Horsager, J., Knudsen, K. & Sommerauer, M. Clinical and imaging evidence of brain-first and body-first Parkinson's disease. *Neurobiol. Dis.* **164**, 105626 (2022).
33. Ferri, R., Cosentino, F. I., Pizza, F., Aricò, D. & Plazzi, G. The timing between REM sleep behavior disorder and Parkinson's disease. *Sleep Breath.* **18**, 319–323 (2014).
34. Nomura, T., Kishi, M. & Nakashima, K. Differences in clinical characteristics when REM sleep behavior disorder precedes or comes after the onset of Parkinson's disease. *J. Neurol. Sci.* **382**, 58–60 (2017).

35. Postuma, R. B. et al. A single-question screen for rapid eye movement sleep behavior disorder: a multicenter validation study. *Mov. Disord.* **27**, 913–916 (2012).
36. Nomura, T., Inoue, Y., Kagimura, T., Uemura, Y. & Nakashima, K. Utility of the REM sleep behavior disorder screening questionnaire (RBDSQ) in Parkinson's disease patients. *Sleep Med.* **12**, 711–713 (2011).
37. Skorvanek, M., Feketeova, E., Kurtis, M. M., Ruzs, J. & Sonka, K. Accuracy of rating scales and clinical measures for screening of rapid eye movement sleep behavior disorder and for predicting conversion to Parkinson's disease and other synucleinopathies. *Front. Neurol.* **9**, 376 (2018).
38. Hughes, A. J., Daniel, S. E., Kilford, L. & Lees, A. J. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatry.* **55**, 181–184 (1992).
39. Postuma, R. B. et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord.* **30**, 1591–1601 (2015).
40. Durand, E. Y., Do, C. B., Mountain, J. L. & Macpherson, J. M. Ancestry composition: a novel, efficient pipeline for ancestry deconvolution. *bioRxiv* <https://doi.org/10.1101/010512> (2014).
41. Henn, B. M. et al. Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples. *PLoS ONE* **7**, e34267 (2012).
42. Loh, P. R. et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
43. Browning, S. R. & Browning, B. L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).
44. Huang, J. et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nat. Commun.* **6**, 8111 (2015).
45. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
46. Walter, K. et al. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82–90 (2015).
47. Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
48. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
49. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* **26**, 2190–2191 (2010).
50. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
51. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
52. Smith, G. D. & Ebrahim, S. Mendelian randomization: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22 (2003).
53. Burgess, S., Small, D. S. & Thompson, S. G. A review of instrumental variable estimators for Mendelian randomization. *Stat. Methods Med. Res.* **26**, 2333–2355 (2017).
54. Storm, C. S. et al. Finding genetically-supported drug targets for Parkinson's disease using Mendelian randomization of the druggable genome. *Nat. Commun.* **12**, 7342 (2021).
55. Bottigliengo, D. et al. A Mendelian randomization study investigating the causal role of inflammation on Parkinson's disease. *Brain* **145**, 3444–3453 (2022).
56. Kang, X. et al. Tumor necrosis factor inhibition and Parkinson disease: a Mendelian randomization study. *Neurology* **96**, e1672–e1679 (2021).
57. Hemani, G. et al. The MR-Base platform supports systematic causal inference across the human phenotype. *Elife* **7**, e34408 (2018).
58. Noyce, A. J. et al. The Parkinson's disease Mendelian randomization research portal. *Mov. Disord.* **34**, 1864–1872 (2019).
59. Burgess, S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int. J. Epidemiol.* **43**, 922–929 (2014).

Acknowledgements

We wholeheartedly thank the participants in this study. We would like to thank the research participants and all members of IPDGC for making this work possible. The AMP-PD cohort data used in this study included the Fox Investigation for New Discovery of Biomarkers (BioFIND), the Harvard Biomarker Study (HBS) and the Parkinson's Disease Biomarkers Program (PDBP) cohorts. For up-to-date information on the study, visit <https://www.amp-pd.org>. AMP PD—a public-private partnership—is managed by the FNIH and funded by Celgene, GSK, the Michael J. Fox Foundation for Parkinson's Research, the National Institute of Neurological Disorders and Stroke, Pfizer, Sanofi, and Verily. BioFIND is sponsored by The Michael J. Fox Foundation for Parkinson's Research (MJFF) with support from the National Institute for Neurological Disorders and Stroke (NINDS). The BioFIND Investigators have not participated in reviewing the data analysis or content of the manuscript. For up-to-date information on the study, visit <https://www.michaeljfox.org/news/biofind>. The Harvard Biomarker Study (HBS) is a collaboration of HBS investigators [full list of HBS investigators found at <https://www.bwhparkinsoncenter.org/biobank/> and funded through philanthropy and NIH and Non-NIH funding sources. The HBS Investigators have not participated in reviewing the data analysis or content of the manuscript. The Parkinson's Disease Biomarker Program (PDBP) consortium is supported by the National Institute of Neurological Disorders and Stroke (NINDS) at the National Institutes of Health. A full list of PDBP investigators can be found at <https://pdbp.ninds.nih.gov/policy>. The PDBP investigators have not participated in reviewing the data analysis or content of the manuscript. The PRoBaND cohort is primarily funded and supported by Parkinson's UK (<https://www.parkinsons.org.uk/>) and supported by the National Institute for Health Research (NIHR) Dementias and Neurodegenerative Diseases Research Network (DeNDroN) and by NHS Greater Glasgow and Clyde. PRoBaND has multicentre research ethics approval from the West of Scotland Research Ethics Committee: IRAS 70980, MREC 11/AL/0163. The Oxford Parkinson's Disease Centre (OPDC) Discovery cohort is also primarily funded by Parkinson's UK and additionally supported by the NIHR-DeNDroN and by the NIHR Oxford Biomedical Research Centre, based at the Oxford University Hospitals NHS Trust, and the University of Oxford. OPDC has multicentre research ethics approval from the South Central Oxford A Research Ethics Committee: 16/SC/0108. The Parkinson's Family Project (PFP) cohort has received funding from Parkinson's UK and the Janet Owens Bequest Fund. PFP has multicentre research ethics approval from the London—Camden and Kings Cross Research Ethics Committee: 5/LO/0097. R.R. is funded by Aligning Science Across Parkinson's (grant number ASAP-000478) through the Michael J. Fox Foundation for Parkinson's Research (MJFF). MS, MO and AL received funding from the Slovak Research and Development Agency under contract no. APVV-22-0279, and the EU Renewal and Resilience Plan “Large projects for excellent researchers” under grant No. 09I03-03-V03-00007. This work was financially supported by the Michael J. Fox Foundation, Parkinson's Society Canada, the Canadian Consortium on Neurodegeneration in Aging (CCNA), and the Canada First Research Excellence Fund (CFREF), awarded to McGill University for the Healthy Brains for Healthy Lives (HBHL) program. Z.G.O. is supported by the Fonds de recherche du Québec—Santé (FRQS) Chercheurs-boursiers award, and is a William Dawson Scholar. We would also like to thank the research participants and employees of 23andMe for making this work possible. K.H., P.F., L.K., and 23andMe Research Team are employed by and hold stock or stock options in 23andMe, Inc. A complete list of the members of IPDGC and 23andMe is detailed in the Supplementary Material.

Author contributions

Y.L.S., Z.G.O., and L.K.: conception and design of the study; Y.L.S., L.K., K.H., P.F., L.N.K., E.Y., U.R., J.A.R., K.F., F.A., K.B., M.S., H.M., L.W., R.R., L.P., M.T., T.G., K.B., H.L., M.T.M.H., D.G.G., S.J.G.L., J.B.K., P.P., I.A., M.S., A.L., M.O., M.R., Z.G.O.: data acquisition/analysis; Y.L.S. and Z.G.O.: manuscript drafting; All authors have read and approved this manuscript.

Competing interests

Z.G.O. has received consulting fees from Lysosomal Therapeutics Inc., Idorsia, Prevail Therapeutics, Denali, Ono Therapeutics, Neuron23, Handl Therapeutics, UBC, Bial Biotech Inc., Bial, Deerfield, Guidepoint, Lighthouse and Inception Sciences (now Ventus). None of these companies were involved in any parts of preparing, drafting and publishing this study. K.H., P.F., L.N.K., and the 23andMe Research Team are employed by and hold stock or stock options in 23andMe. H.M. is employed by UCL. In the last 12 months he reports paid consultancy from Roche, Aprinoia and Amylyx; lecture fees/honoraria—BMJ, Kyowa Kirin, Movement Disorders Society. Research Grants from Parkinson's UK, Cure Parkinson's Trust, PSP Association, Medical Research Council, Michael J Fox Foundation. H.R.M. is a co-applicant on a patent application related to C9ORF72—Method for diagnosing a neurodegenerative disease (PCT/GB2012/052140). The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41531-025-01078-w>.

Correspondence and requests for materials should be addressed to Ziv Gan-Or.

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) 2025