Novel associations of *BST1* and *LAMP3* with rapid eye movement sleep behavior disorder

Running title: Association of *BST1* and *LAMP3* with iRBD

Kheireddin Mufti, BPharm[1,2], Eric Yu[1,2], BSc, Uladzislau Rudakou, BSc[1,2], Lynne Krohn, MSc[1,2], Jennifer A. Ruskey, MSc[2,3], Farnaz Asayesh, MSc[2,3], Sandra B. Laurent, BTS[2,3], Dan Spiegelman, MSc[2,3], Isabelle Arnulf, MD, PhD[4], Michele T.M. Hu, MBBS, PhD[5,6], Jacques Y. Montplaisir, MD, PhD[7,8], Jean-François Gagnon, PhD[7,9], Alex Desautels, MD, PhD[7,10], Yves Dauvilliers, MD, PhD[11], Gian Luigi Gigli, MD[12,13], Mariarosaria Valente, MD[12,14], Francesco Janes, MD, PhD[12], Andrea Bernardini, MD[12], Birgit Högl, MD[15], Ambra Stefani, MD[15], Evi Holzknecht, MD[15], Karel Sonka, MD, PhD[16], David Kemlink, MD, PhD[16], Wolfgang Oertel, MD[17], Annette Janzen, MD[17], Giuseppe Plazzi, MD[18,19], Elena Antelmi, MD, PhD[20], Michela Figorilli, MD, PhD[21], Monica Puligheddu, MD, PhD[21], Brit Mollenhauer, MD[22,23], Claudia Trenkwalder, MD[22.23], Friederike Sixel-Döring, MD[17,22], Valérie Cochen De Cock, MD, PhD[24,25] Christelle Charley Monaca, MD,PhD[26], Anna Heidbreder, MD[27], Luigi Ferini-Strambi, MD[28], Femke Dijkstra, MD[29,30,31], Mineke Viaene, MD, PhD[29,30], Beatriz Abril, MD[32], Bradley F. Boeve, MD[33], Jean-François Trempe, PhD[34], Guy A. Rouleau, MD, PhD, FRCPC, FRSC[1,2,3], Ronald B. Postuma, MD, MSc[2,3,7], Ziv Gan-Or, MD, PhD[1,2,3].

Affiliations:
1. Department of Human Genetics, McGill University, Montréal, QC, Canada
2. Montreal Neurological Institute, McGill University, Montréal, QC, Canada
3. Department of Neurology and Neurosurgery, McGill University, Montréal, QC, Canada
4. Sleep Disorders Unit, Pitié Salpêtrière Hospital, Paris Brain Institute and Sorbonne University, Paris, France
5. Oxford Parkinson's Disease Centre (OPDC), University of Oxford, Oxford, United Kingdom
6. Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom
7. Center for Advanced Research in Sleep Medicine, Centre Intégré Universitaire de Santé et de ServicesSociaux du Nord-de-l'Île-de-Montréal – Hôpital du Sacré-Coeur de Montréal, Montreal, Quebec, Canada
8. Department of Psychiatry, Université de Montréal, Montréal, QC, Canada
9. Department of Psychology, Université du Québec à Montréal, Montréal, QC, Canada
10. Department of Neurosciences, Université de Montréal, Montréal, QC, H3T 1J4, Canada
11. National Reference Center for Narcolepsy, Sleep Unit, Department of Neurology, Gui-de-Chauliac Hospital, CHU Montpellier, University of Montpellier, Inserm U1061, Montpellier, France
12. Clinical Neurology Unit, Department of Neurosciences, University Hospital of Udine, Udine, Italy.
13. DMIF, University of Udine, Udine, Italy
14. Department of Medicine (DAME), University of Udine, Udine, Italy

15. Sleep Disorders Clinic, Department of Neurology, Medical University of Innsbruck, Innsbruck, Austria
16. Department of Neurology and Centre of Clinical Neuroscience, Charles University, First Faculty of Medicine and General University Hospital, Prague, Czech Republic
17. Department of Neurology, Philipps University, Marburg, Germany
18. Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio-Emilia, Modena, Italy
19. IRCCS, Institute of Neurological Sciences of Bologna, Bologna, Italy
20. Neurology Unit, Movement Disorders Division, Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy
21. Department of Medical Sciences and Public Health, Sleep Disorder Research Center, University of Cagliari, Cagliari, Italy. 22. Paracelsus-Elena-Klinik, Kassel, Germany
23. Department of Neurosurgery, University Medical Centre Göttingen, Göttingen, Germany
24. Sleep and Neurology Unit, Beau Soleil Clinic, Montpellier, France
25. EuroMov Digital Health in Motion, University of Montpellier IMT Mines Ales, Montpellier, France
26. University Lille North of France, Department of Clinical Neurophysiology and Sleep Center, CHU Lille, Lille, France
27. Department of Sleep Medicine and Neuromuscular Disorders, University of Müenster, Müenster, Germany
28. Department of Neurological Sciences, Università Vita-Salute San Raffaele, Milan, Italy
29. Laboratory for Sleep Disorders, St. Dimpna Regional Hospital, Geel, Belgium
30. Department of Neurology, St. Dimpna Regional Hospital, Geel, Belgium
31. Department of Neurology, University Hospital Antwerp, Edegem, Antwerp, Belgium
32. Sleep disorder Unit, Carémeau Hospital, University Hospital of Nîmes, France
33. Department of Neurology, Mayo Clinic, Rochester, MN, USA
34. Centre de Recherche en Biologie Structurale and Department of Pharmacology and Therapeutics, McGill University, Montréal, QC, Canada

Corresponding author:

Ziv Gan-Or
Montreal Neurological Institute,
McGill University
1033 Pine Avenue, West,
Ludmer Pavilion, room 312
Montréal, QC, H3A 1A1,
Phone: +1-514-398-5845,
e-mail: ziv.gan-or@mcgill.ca

Abstract

Isolated rapid-eye-movement (REM) sleep behavior disorder (iRBD) is a parasomnia, characterized by loss of muscle atonia and dream enactment occurring during REM sleep phase. Since a large subgroup of iRBD patients will convert to Parkinson's disease, and since previous genetic studies have suggested common genes, it is likely that there is at least a partial overlap between iRBD and Parkinson's disease genetics. To further examine this potential overlap and to identify genes specifically involved in iRBD, we fully sequenced 25 genes previously identified in genome-wide association studies of Parkinson's disease. The genes were captured and sequenced using targeted next-generation sequencing in a total of 1,039 iRBD patients and 1,852 controls of European ancestry. The role of rare heterozygous variants in these genes was examined using burden tests and optimized sequence Kernel association tests (SKAT-O), adjusted for age and sex. The contribution of biallelic (homozygous and compound heterozygous) variants was further tested in all genes. To examine the association of common variants in the target genes, we used logistic regression adjusted for age and sex. We found a significant association between rare heterozygous nonsynonymous variants in *BST1* and iRBD (p=0.0003 at coverage >50X and 0.0004 at >30X), mainly driven by three nonsynonymous variants (p.V85M, p.I101V and p.V272M) found in a total of 22 (1.2%) controls vs. two (0.2%) patients. Rare non-coding heterozygous variants in *LAMP3* were also found to be associated with reduced iRBD risk (p=0.0006 at >30X). We found no statistically significant association between rare heterozygous variants in the rest of genes and risk of iRBD. Several carriers of biallelic variants were identified, yet there was no overrepresentation in iRBD patients vs. controls. To examine the potential impact of the rare nonsynonymous *BST1* variants on the protein structure, we performed *in silico* structural analysis. All three variants seem to be loss-of-

function variants significantly affecting the protein structure and stability. Our results suggest that rare coding variants in *BST1* and rare non-coding variants in *LAMP3* are associated with iRBD, and additional studies are required to replicate these results and examine whether loss-of-function of *BST1* could be a therapeutic target.

Keywords (5): Parkinson's disease, Rapid-eye-movement sleep behavior disorder, association study, BST1, LAMP3

Abbreviations:

AAO: age at onset

AAS: age at sampling

CADD: Combined Annotation Dependent Depletion

CI: confidence interval

DOC: depth of coverage

F_A: frequency in affected

F_C: frequency in control

GATK: Genome Analysis Toolkit

GnomAD: Genome Aggregation Database

GTEx: Genotype-tissue expression

GWAS: genome-wide association study

hg19: human genome version 19

iRBD: isolated rapid-eye-movement sleep behavior disorder

LD: linkage disequilibrium

LOF: loss-of-function

MAF: minor allele frequency

MIPs: molecular inversion probes

NGS: next-generation sequencing

NS: non-synonymous

OR: odds ratio

PCA: principle component analysis

QC: quality control

Qs: quality score

REM: rapid eye movement

SKAT: sequence kernel association test

SKAT-O: optimized SKAT

UTR: untranslated region

vPSG: video polysomnography

Introduction

Isolated rapid-eye-movement sleep behavior disorder (iRBD) is a prodromal synucleinopathy, as more than 80% of iRBD patients will eventually convert to an overt neurodegenerative syndrome associated with α-synuclein pathology. Typically, iRBD patients will convert to Parkinson's disease (about 40-50% of patients), dementia with Lewy bodies or unspecified dementia (40-50%), or, in much fewer cases, to multiple system atrophy (5-10%) (Iranzo *et al.*, 2016; Postuma *et al.*, 2019). While our understanding of the genetic background of dementia with Lewy bodies or multiple system atrophy is limited, the rapid development of various genetic methods during the recent decades has led to wealth of data on the role of common and rare genetic variants in Parkinson's disease. To date, there are 80 genetic loci found to be associated with Parkinson's disease risk discovered through genome-wide association studies (GWASs) (Nalls *et al.*, 2019; Foo *et al.*, 2020), and several genes have been implicated in familial Parkinson's disease (Gan-Or *et al.*, 2018; Gan-Or and Rouleau, 2019; Blauwendraat *et al.*, 2020).

In order to study the genetic background of iRBD and its conversion to α-synucleinopathies, recent studies have examined whether Parkinson's disease- or dementia with Lewy bodies-related genes are also associated with iRBD. These studies have suggested that while there is some overlap between the genetic backgrounds of iRBD and Parkinson's disease or dementia with Lewy bodies, this overlap is only partial. For example, it was demonstrated that *GBA* variants are associated with iRBD risk, Parkinson's disease and dementia with Lewy bodies (Gan□Or *et al.*, 2015; Gan-Or and Rouleau, 2019), but pathogenic *LRRK2* variants are found to only be associated with Parkinson's disease, and not with iRBD and dementia with Lewy bodies (Heckman *et al.*, 2016; Bencheikh *et al.*, 2018; Blauwendraat *et al.*, 2020). We have recently reported that the familial Parkinson's disease and atypical parkinsonism genes *PRKN*, *PARK7*,

GCH1, *VPS35*, *ATP13A2*, *VPS13C*, *FBXO7* and *PLA2G6* are not likely to be involved in iRBD (Mufti *et al.*, 2020). Heterozygous variants in *SMPD1* have been reported to be associated with Parkinson's disease risk (Gan-Or *et al.*, 2013; Alcalay *et al.*, 2019), yet no association was found with iRBD (Rudakou *et al.*, 2020). Whereas variants in *MAPT* are associated with Parkinson's disease and *APOE* haplotypes are important risk factors of dementia with Lewy bodies, (Dickson *et al.*, 2018; Li *et al.*, 2018), neither are linked to iRBD (Gan-Or *et al.*, 2017; Li *et al.*, 2018). Furthermore, there are independent risk variants of Parkinson's disease, dementia with Lewy bodies and iRBD within *SNCA* locus; specific variants in the 3' untranslated region (UTR) are associated with Parkinson's disease but not with iRBD, and other, independent variants at 5' UTR are associated with Parkinson's disease, iRBD and dementia with Lewy bodies (Krohn *et al.*, 2020b). In the *TMEM175* locus, there are two independent Parkinson's disease risk variants, but only one of them, p.M393T, has also been associated with iRBD risk (Krohn *et al.*, 2020a).

Thus far, the role of most Parkinson's disease GWAS genes has not been thoroughly studied in iRBD. In the current study, we aimed to examine whether rare and common variants in 25 Parkinson's disease-related GWAS genes are associated with iRBD. The entire coding regions with the exon-intron boundaries as well as the regulatory 3' and 5' UTRs were captured and sequenced. We then performed different genetic analyses to investigate the association of rare and common variants in these genes with iRBD.

Materials and methods

Study population

This study included a total of 2,891 subjects, composed of 1,039 unrelated individuals diagnosed with iRBD (according to the International Classification of Sleep Disorders criteria, version 2 or 3) and 1,852 controls. Details on age and sex of patients and controls have been previously described (Mufti *et al.*, 2020) and can be found in Supplementary Table 1. Differences in age and sex were taken into account as needed in the statistical analysis. All patients and controls were of European ancestry (confirmed by principal component analysis [PCA] of GWAS data compared to data from HapMap v.3 and hg19/GRCh37).

Standard protocol approvals, registrations, and patient consents

All study participants signed an informed consent form before entering the study, and the study protocol was approved by the institutional review boards.

Selection of genes and genetic analysis

The current study was designed and performed before the publication of the recent Parkinson's disease GWAS (Nalls *et al.*, 2019), therefore, the genes for analysis were selected from previous GWASs (Nalls *et al.*, 2014; Chang *et al.*, 2017). A total of 25 genes were selected for analysis, including: *ACMSD*, *BST1*, *CCDC62*, *DDRGK1*, *DGKQ*, *FGF20*, *GAK*, *GPNMB*, *HIP1R*, *ITGA8*, *LAMP3*, *MAPT*, *MCCC1*, *PM20D1*, *RAB25*, *RAB29*, *RIT2*, *SETD1A*, *SLC41A1*, *STK39*, *SIPA1L2*, *STX1B*, *SYT11*, *TMEM163* and *USP25*. These genes were selected based on the

presence of one or more of the following: quantitative trait loci, expression in brain, potential interaction with known Parkinson's disease-associated genes and involvement in pathways implicated in Parkinson's disease, such the autophagy-lysosomal pathway, mitochondria quality control and endolysosomal recycling. The 25 genes were fully captured (coding sequence and 3'- and 5'- untranslated regions) using molecular inversion probes (MIPs) designed as previously described (O'Roak *et al.*, 2012). The full protocol is available upon request. Supplementary Table 2 details the probes used in the current study for the MIPs capture. Targeted next-generation sequencing (NGS) was performed post-capture using illumina HiSeq 2500\4000 platform at the McGill University and Génome Québec Innovation Centre. Sequencing reads were aligned to the human reference genome (hg19) using the Burrows-Wheeler Aligner (Li and Durbin, 2009). Genome Analysis Toolkit (GATK, v3.8) was used for post-alignment quality control and variant calling (McKenna *et al.*, 2010), and ANNOVAR for annotation (Wang *et al.*, 2010). The Frequency of each variant was extracted from the Genome Aggregation Database (GnomAD) (Lek *et al.*, 2016).

Quality control

Quality control (QC) was performed using PLINK v1.9 (Purcell *et al.*, 2007). We excluded variants that deviated from Hardy-Weinberg equilibrium in controls with a threshold set at p=0.001, and those identified in <25% of the reads for a specific variant. We also filtered out variants with genotyping rate lower than 90%. The same genotyping rate cut-off was used for exclusion of individual samples. Threshold for rate of missingness difference between patients and controls was set at p=0.05, and variants below this threshold were excluded from the analysis. To be included in the analysis, the minimum genotype quality score was set to 30. We

used two coverage thresholds for rare variants (minor allele frequency [MAF] <0.01), >30X and >50X, and all analyses were repeated using these thresholds. For the analysis of common variants, coverage of >15X was used.

Statistical analysis

To test whether rare heterozygous variants (defined by MAF<0.01) in each of our target genes are associated with iRBD, we performed sequence kernel association test (SKAT, R package) (Lee *et al.*, 2012) and optimized sequence kernel association test (SKAT-O) on different groups of variants: all rare variants, potentially functional rare variants (including nonsynonymous, frame-shift, stop-gain and splicing), rare loss-of-function variants (frame-shift, stop-gain and splicing), and rare nonsynonymous variants only. In addition, we further tested whether rare variants that are predicted to be pathogenic based on Combined Annotation Dependent Depletion (CADD) score of ≥12.37 (representing the top 2% of potentially deleterious variants) are enriched in iRBD patients. To test the association between biallelic variants and iRBD risk, we compared the frequencies of carriers of two vary rare (MAF<0.001) nonsynonymous, splice-site, frame-shift and stop-gain variants between patients and controls using Fisher's exact test. Bonferroni correction for multiple comparisons was applied as necessary. We tested the association between common variants (MAF>0.01) in the target genes and iRBD risk using logistic regression adjusted for age and sex using PLINK v1.9. Linkage disequilibrium between the discovered variants and the respective GWAS top hits was examined using the non-Finnish European reference cohort on LDlink (https://ldlink.nci.nih.gov/) (Machiela and Chanock, 2015). Effects of common variants on expression was viewed using the genotype-tissue expression database (GTEx - https://www.gtexportal.org). We further performed in silico structural analysis

of *BST1* to test whether the rare coding variants that were found to be associated with iRBD in our analysis could potentially affect the enzyme structure and activity. The atomic coordinates of human *BST1* bound to ATP-γ-S were downloaded from the Protein Data Bank (ID 1isg). The steric clashes induced by each variant were evaluated using the "mutagenesis" toolbox in PyMol v. 2.2.0.

Data availability

Data used for the analysis is available in the supplementary tables. Anonymized raw data can be shared upon request.

Results

Coverage and identified variants

The average coverage of the 25 genes analyzed in this study was >647X (range 73-1162, median 790). An average of 95% of the target regions were covered with >15X, 93% with >30X and 90% with >50X. The average coverage of each gene and the percentage of the nucleotides covered at 15X, 30X and 50X are detailed in Supplementary Table 3. Finally, there were no differences in the coverage between patients and controls. A total of 1,189 rare variants were found with coverage of > 30X, and 570 rare variants with > 50X (Supplementary Table 4). We identified 125 common variants across all genes (Supplementary Table 5) with a coverage of >15X.

Rare heterozygous variants in *BST1* and *LAMP3* are associated with iRBD

To examine whether rare heterozygous variants in our genes of interest may be associated with iRBD risk, we performed SKAT and SKAT-O tests, repeated twice for variants detected at depths of coverage of >30X and >50X (see methods). Supplementary Table 4 details all rare heterozygous variants identified in each gene and included in the analysis. We applied both SKAT and SKAT-O on five different groups of variants: all rare variants, all potentially functional variants (nonsynonymous, splice site, frame-shift and stop-gain), loss-of-function variants (frame-shift, stop-gain and splicing), nonsynonymous variants only, and variants with CADD score ≥12.37 (Table 1). The Bonferroni-corrected *p*-value threshold for statistical significance was set at $p<0.001$ after correcting for the number of genes and depths of coverage.

We found a statistically significant association of rare heterozygous functional variants in *BST1* (SKAT p=0.0004 at >30X and p=0.0003 at >50X for rare functional variants), found more in controls than in iRBD patients. This association is mainly driven by the nonsynonymous variants p.V85M (rs377310254, found in five controls and none in patients), p.I101V (rs6840615, found in seven controls and none in patients), and p.V272M (rs144197373, found in 10 controls and two patients). Overall, these variants were found in 22 (1.2%) controls vs. 2 (0.2%) patients. Another statistically significant association was found between rare variants in *LAMP3* gene and reduced iRBD risk in SKAT-O analysis. This association is driven by two non-coding variants (one intronic [location - chr3:182858302] and one at the 3' UTR of *LAMP3* [rs56682988, c.*415T>C]) found only in controls (15 and nine controls, respectively). In order to further examine whether these variants indeed drive the association in both *BST1* and *LAMP3*, we excluded them and repeated the analysis (SKAT and SKAT-O), which resulted in loss of statistical significance for both genes (Supplementary table 6). There were no additional statistically significant associations of the remaining genes with iRBD after correcting for multiple comparisons (p<0.001).

Structural analysis of *BST1* variants suggests that loss-of-function may be protective in iRBD

To investigate the potential impact of the three *BST1* nonsynonymous variants (p.V85M, p.I101V and p.V272M) on the structure and activity of the enzyme, we performed *in silico* mutagenesis and evaluated potential clashes with surrounding residues. Figure 1 depicts the structure of *BST1* with the respective locations of the three nonsynonymous variants that drive

the *BST1* association detected in our analysis. The structure of human *BST1* was solved by X-ray crystallography in complex with five substrate analogues (Yamamoto-Katayama *et al.*, 2002). All structures revealed a homodimeric assembly, with the catalytic clefts facing the cavity at the interface of the two chains (Figure 1A).

The sidechain of p.V85M points towards the hydrophobic core of the protein, behind a helix facing the nucleotide binding site. The amino-acid change from valine to the bulkier sidechain of methionine results in clashes with other residues in the core, for all rotamers (Figure 1B). This variant would therefore likely destabilize the enzyme active site and potentially unfold the protein. The sidechain of the variant p.I101V is located underneath the active site towards the hydrophobic core. Although the amino-acid change from isoleucine to the smaller sidechain of valine does not create clash (Figure 1C), it reduces the packing in the core, which could also destabilize the enzyme. Finally, the p.V272M variant is located in a helix at the C-terminus of the protein that forms symmetrical contacts with the same helix in the other chain of the dimer. The p.V272M variant would create clashes with sidechain and main-chain atoms located in the other chain of the dimer (Figure 1D). As p.V272M resides at the dimer interface of the enzyme and probably helps maintaining the two subunits together, this variant would most likely lead to the disruption of the dimer. Overall, all the disease-associated nonsynonymous variants in *BST1* (p.V85M, p.I101V, and p.V272M) appear to be “loss-of-function”, suggesting that reduced BST1 activity may be protective in iRBD. This is supported by the top Parkinson’s disease GWAS hit in the *BST1* locus, the rs4698412 G allele, which is associated with reduced risk of Parkinson’s disease (Nalls *et al.*, 2019). This allele is also associated with reduced expression of BST1 in blood in GTEx (normalized effect size =-0.07, p=1.5e-6), suggesting that reduced expression might be protective.

Very rare bi-allelic variants are not enriched in iRBD patients

In order to examine whether bi-allelic variants in our genes of interest are enriched in iRBD, we compared the carrier frequencies of very rare (MAF<0.001) homozygous and compound heterozygous variants between iRBD patients and controls. To analyze compound heterozygous variants, since phasing could not be performed, we considered carriers of two very rare variants as compound heterozygous carriers, with the following exceptions: 1) when variants were physically close (less than 112 base pairs [bp]; probes' target length) and we could determine their phase based on the sequence reads, and 2) if the same combination of very rare variants appeared more than once across samples, we assumed that the variants are most likely to be on the same allele. We found five (0.5%) iRBD patients and seven (0.4%) controls presumably carriers of bi-allelic variants in the studied genes (Table 2, p=0.731, Fisher test).

Association of common variants in the target genes with iRBD

To test whether common variants in our target genes are associated with iRBD, we performed logistic regression (using PLINK v1.9 software) adjusted for age and sex for common variants (MAF>0.01) detected at coverage depth of >15X. A nominal association was observed in 12 variants across all genes (Supplementary table 5), but no association remained statistically significant after Bonferroni correction for multiple comparisons (set at p<0.0005).

Of the variants with nominal associations, one variant in the *ITGA8* 3' UTR (rs896435, OR=1.15, 95% CI = 1.01-1.32, p=0.04) is the top hit from the most recent Parkinson's disease

GWAS (Nalls *et al.*, 2019), and two other *ITGA8* 3' UTR variants are almost in perfect LD (D'=1.0, R^2>0.99, p<0.0001) with rs896435. Four variants in the 3' UTR of *RAB29* were almost in perfect LD (Supplementary table 5) and are associated with expression of RAB29 in multiple tissues in GTEx, including the brain. Three *MAPT* variants were in partial LD with Parkinson's disease GWAS hits in the *MAPT* locus and were associated with expression of multiple genes in multiple tissues in GTEx, demonstrating the complexity of this genomic region.

Discussion

In the current study, we studied a large cohort of iRBD patients by fully-sequencing and analyzing 25 Parkinson's disease-related GWAS genes and their association with iRBD. Our results identify *BST1* and *LAMP3* as novel genes potentially associated with iRBD. Based on *in silico* models, the three nonsynonymous *BST1* variants that drive the association with iRBD may be loss-of-function variants, suggesting that reduced *BST1* activity may reduce the risk of developing iRBD. The variants driving the association of *LAMP3* are in noncoding regions and could be regulatory. These hypotheses will require confirmation in functional studies in relevant models. While some common variants were nominally associated with iRBD, none of them remained statistically significant after correction for multiple comparisons.

BST1, also called CD157, is a glycosyl phosphatidylinositol (GPI) anchored membrane protein initially found in bone marrow stromal cells and is essential for B-lymphocyte growth and development. It has an extracellular enzymatic domain that produces cyclic ADP-ribose (cADPR). This metabolite acts as a second messenger that can trigger Ca^{2+} release from intracellular stores (Ishihara and Hirano, 2000), a process that plays a role in cellular function and death. Specific features of calcium homeostasis have been suggested to be responsible for the specific vulnerability of dopaminergic neurons in Parkinson's disease (Chan *et al.*, 2009), yet whether BST1 is involved in calcium homeostasis in human neurons is still unclear, as most work was done in non-human models. Another mechanism by which BST1 may be involved in Parkinson's disease is immune response and neuroinflammation, which are likely important in the pathogenesis of the disease (Wang *et al.* 2015). BST1 serves as a receptor which regulates leukocyte adhesion and migration, and plays a role in inflammation (Malavasi *et al.*, 2006). However, its potential role in microglia activation and neuroinflammation is yet to be

determined. Our *in-silico* analysis suggested that the *BST1* variants found mostly in controls are loss-of-function variants. We can therefore hypothesize that these variants may reduce immune response and lead to a reduced risk of iRBD, and that inhibition of *BST1* could be a therapeutic target for iRBD and Parkinson's disease treatment or prevention.

LAMP3 encodes the lysosomal-associated membrane protein 3, which plays a role in the unfolded protein response (UPR) that contributes to protein degradation and cell survival during proteasomal dysfunction (Dominguez-Bautista *et al.*, 2015). Furthermore, LAMP3 knockdown impairs the ability of the cells to complete the autophagic process, and high LAMP3 expression is associated with increased basal autophagy levels (Nagelkerke *et al.*, 2014). Numerous Parkinson's disease-related genes have been implicated in the autophagy-lysosomal pathway (Senkevich and Gan-Or, 2019), and genes associated with iRBD such as *GBA* (Krohn *et al.* 2020), *TMEM175* (Krohn *et al.* 2020) and *SNCA* (Krohn *et al.* 2020) are all involved in this pathway (Senkevich and Gan-Or, 2019). Our current findings further strengthen the potential association between the autophagy-lysosomal pathway and iRBD.

Our study has several limitations. First, despite being the largest genetic study of iRBD to date, it may be still underpowered to detect rare variants in GWAS Parkinson's disease-related genes, as well as common variants with a small effect size. Therefore, we cannot completely rule out the possibility that rare and common variants in these genes may contribute to iRBD risk. A second limitation is the younger age and the differences in sex distribution between iRBD patients and controls, for which we adjusted in the statistical analysis as needed. Another potential limitation is the possibility that there were undiagnosed iRBD patients among the control population. However, since iRBD is found in only ~1% of the population (Postuma *et al.*,

2019), the effect of having undiagnosed iRBD patients in the controls would be minimal, given the large sample size.

To conclude, our results suggest two novel genetic associations with iRBD; an association with rare functional variants in *BST1*, and with rare non-coding variants in *LAMP3*. All the association-driving coding variants found in *BST1*, mainly in controls, appear to potentially cause loss-of-function, suggesting that reduced BST1 activity may reduce the risk of iRBD. Further studies would be required to confirm our results and to examine the biological mechanism underlying the effect of disease-associated variants in both *LAMP3* and *BST1*. The absence of evidence of association between rare and common variants in the remaining genes and iRBD risk suggests that these genes either have no effect in iRBD or have a minor effect that we could not detect with this sample size. Environmental factors and environment-gene interactions are likely to play a major role on iRBD, and larger studies that include carefully collected epidemiological data and more extensive genetic data such as whole-exome or whole-genome sequencing will be required to clarify these issues.

Acknowledgments

We thank the participants for contributing to this study. JYM holds a Canada Research Chair in Sleep Medicine. JFG holds a Canada Research Chair in Cognitive Decline in Pathological Aging. WO is Hertie Senior Research Professor, supported by the Hertie Foundation. EAF holds a Canada Research Chair (Tier 1) in Parkinson disease. GAR holds a Canada Research Chair in Genetics of the Nervous System and the Wilder Pen field Chair in Neurosciences. JFT holds a Canada Research Chair (Tier 2) in Structural Pharmacology. ZGO is supported by the Fonds de recherche du Québec–Santé Chercheur-Boursier award and is a Parkinson Canada New Investigator awardee. We thank D. Rochefort, H. Catoire, and V. Zaharieva for their assistance.

Funding

This work was financially supported by the Michael J. Fox Foundation, the Canadian Consortium on Neurodegeneration in Aging (CCNA), Parkinson Canada, the Canada First Research Excellence Fund (CFREF), awarded to McGill University for the Healthy Brains for Healthy Lives (HBHL) program. The Montreal cohort was funded by the Canadian Institutes of Health Research (CIHR) and the W. Garfield Weston Foundation. The Oxford Discovery study is funded by the Monument Trust Discovery Award from Parkinson's UK and supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre based at Oxford University Hospitals NHS Trust and University of Oxford, the NIHR Clinical Research Network and the Dementias and Neurodegenerative Diseases Research Network (DeNDRoN).

Competing interests

JYM received fees from Takeda, Eisai and Paladin Pharma for consultancies in unrelated fields. ZGO received consultancy fees from Lysosomal Therapeutics Inc. (LTI), Idorsia, Prevail Therapeutics, Inceptions Sciences (now Ventus), Ono Therapeutics, Denali, Deerfield, Neuron23 and Handl Therapeutics. None of these companies were involved in any parts of preparing, drafting and publishing this review.

Supplementary material

Supplementary Table 1. Details on age and sex of samples

Supplementary Table 2. MIPs used for each gene

Supplementary Table 3. Average coverage of genes

Supplementary Table 4. All variants included in SKAT-O and SKAT analyses

Supplementary Table 5. All common variants detected at 15x

Supplementary Table 6. Burden and SKAT-O results without *BST1* and *LAMP3* association-driving variants

References

Alcalay RN, Mallett V, Vanderperre B, Tavassoly O, Dauvilliers Y, Wu RY*, et al.* SMPD1 mutations, activity, and α-synuclein accumulation in Parkinson's disease. Movement Disorders 2019; 34(4): 526-35.

Bencheikh BOA, Ruskey JA, Arnulf I, Dauvilliers Y, Monaca CC, De Cock VC*, et al.* LRRK2 protective haplotype and full sequencing study in REM sleep behavior disorder. Parkinsonism & related disorders 2018; 52: 98-101.

Blauwendraat C, Nalls MA, Singleton AB. The genetic architecture of Parkinson's disease. The Lancet Neurology 2020; 19(2): 170-8.

Chan CS, Gertler TS, Surmeier DJ. Calcium homeostasis, selective vulnerability and Parkinson's disease. Trends in neurosciences 2009; 32(5): 249-56.

Chang D, Nalls MA, Hallgrímsdóttir IB, Hunkapiller J, van der Brug M, Cai F*, et al.* A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. Nature genetics 2017; 49(10): 1511.

Dickson DW, Heckman MG, Murray ME, Soto AI, Walton RL, Diehl NN*, et al.* APOE ε4 is associated with severity of Lewy body pathology independent of Alzheimer pathology. Neurology 2018; 91(12): e1182-e95.

Dominguez-Bautista JA, Klinkenberg M, Brehm N, Subramaniam M, Kern B, Roeper J*, et al.* Loss of lysosome-associated membrane protein 3 (LAMP3) enhances cellular vulnerability against proteasomal inhibition. European journal of cell biology 2015; 94(3-4): 148-61.

Foo JN, Chew EGY, Chung SJ, Peng R, Blauwendraat C, Nalls MA*, 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 neurology 2020.

Gan-Or Z, Alcalay RN, Rouleau GA, Postuma RB. Sleep disorders and Parkinson disease; lessons from genetics. Sleep medicine reviews 2018; 41: 101-12.

Gan-Or Z, Montplaisir JY, Ross JP, Poirier J, Warby SC, Arnulf I*, et al.* The dementia-associated APOE ε4 allele is not associated with rapid eye movement sleep behavior disorder. Neurobiology of aging 2017; 49: 218. e13-. e15.

Gan-Or Z, Ozelius LJ, Bar-Shira A, Saunders-Pullman R, Mirelman A, Kornreich R*, et al.* The p. L302P mutation in the lysosomal enzyme gene SMPD1 is a risk factor for Parkinson disease. Neurology 2013; 80(17): 1606-10.

Gan-Or Z, Rouleau GA. Genetics of REM Sleep Behavior Disorder. Rapid-Eye-Movement Sleep Behavior Disorder: Springer; 2019. p. 589-609.

Gan-Or Z, Mirelman A, Postuma RB, Arnulf I, Bar-Shira A, Dauvilliers Y*, et al.* GBA mutations are associated with rapid eye movement sleep behavior disorder. Ann Clin Transl Neurol 2015; 2(9): 941-5.

Heckman MG, Soto-Ortolaza AI, Contreras MYS, Murray ME, Pedraza O, Diehl NN*, et al.* LRRK2 variation and dementia with Lewy bodies. Parkinsonism & related disorders 2016; 31: 98-103.

Iranzo A, Santamaria J, Tolosa E. Idiopathic rapid eye movement sleep behaviour disorder: diagnosis, management, and the need for neuroprotective interventions. The Lancet Neurology 2016; 15(4): 405-19.

Ishihara K, Hirano T. BST-1/CD157 regulates the humoral immune responses in vivo. Chemical immunology 2000; 75: 235-55.

Krohn L, Öztürk TN, Vanderperre B, Ouled Amar Bencheikh B, Ruskey JA, Laurent SB*, et al.* Genetic, Structural, and Functional Evidence Link TMEM175 to Synucleinopathies. Annals of neurology 2020a; 87(1): 139-53.

Krohn L, Wu RY, Heilbron K, Ruskey JA, Laurent SB, Blauwendraat C*, et al.* Fine-Mapping of SNCA in Rapid Eye Movement Sleep Behavior Disorder and Overt Synucleinopathies. Annals of Neurology 2020b; 87(4): 584-98.

Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA*, et al.* Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. The American Journal of Human Genetics 2012; 91(2): 224-37.

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T*, et al.* Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016; 536(7616): 285-91.

Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. bioinformatics 2009; 25(14): 1754-60.

Li J, Ruskey JA, Arnulf I, Dauvilliers Y, Hu MT, Högl B*, et al.* Full sequencing and haplotype analysis of MAPT in Parkinson's disease and rapid eye movement sleep behavior disorder. Movement Disorders 2018; 33(6): 1016-20.

Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 2015; 31(21): 3555-7.

Malavasi F, Deaglio S, Ferrero E, Funaro A, Sancho J, Ausiello CM*, et al.* CD38 and CD157 as receptors of the immune system: a bridge between innate and adaptive immunity. Molecular medicine; 2006: BioMed Central; 2006. p. 334-41.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A*, et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 2010; 20(9): 1297-303.

Mufti K, Rudakou U, Yu E, Ruskey J, Asayesh F, Laurent S*, et al.* A comprehensive analysis of dominant and recessive parkinsonism genes in REM sleep behavior disorder. medRxiv 2020.

Nagelkerke A, Sieuwerts AM, Bussink J, Sweep F, Look MP, Foekens JA*, et al.* LAMP3 is involved in tamoxifen resistance in breast cancer cells through the modulation of autophagy. Endocr Relat Cancer 2014; 21(1): 101-12.

Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D*, et al.* Expanding Parkinson's disease genetics: novel risk loci, genomic context, causal insights and heritable risk. BioRxiv 2019: 388165.

Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M*, et al.* Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nature genetics 2014; 46(9): 989.

O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG*, et al.* Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science 2012; 338(6114): 1619-22.

Postuma RB, Iranzo A, Hu M, Högl B, Boeve BF, Manni R*, et al.* Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. Brain 2019; 142(3): 744-59.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D*, et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics 2007; 81(3): 559-75.

Rudakou U, Futhey NC, Krohn L, Ruskey JA, Heilbron K, Cannon P*, et al.* SMPD1 variants do not have a major role in REM sleep behavior disorder. Neurobiology of Aging 2020.

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 2010; 38(16): e164-e.

Yamamoto-Katayama S, Ariyoshi M, Ishihara K, Hirano T, Jingami H, Morikawa K. Crystallographic studies on human BST-1/CD157 with ADP-ribosyl cyclase and NAD glycohydrolase activities. Journal of molecular biology 2002; 316(3): 711-23.

Figure legends

Figure 1. Structural analysis of human *BST1* variants

This figure was produced using the software PyMOL v.2.2.0, and represents: (A) Structure of the BST1 dimer bound to ATP-γS (pdb 1ISG). The position of each variant sites is indicated. The ATP-γS molecule in the active site is shown as sticks. (B) Close-up view of the p.V85M variant site. The mutated residue is shown in white. The variant would create clashes (red disks) with nearby Ala77 in the core. (C) Close-up view of the p.I101V variant site. The residue is located in the core of the protein, but the variant to a smaller residue results in no clash. (D) Close-up view of the p.V272M variant site. Primed (‘) residues correspond to chain B. This residue is located at the dimer interface and the variant would create clashes with the other chain resulting in a destabilization of the dimer.

Table 1. Summary of results from burden analyses of rare heterozygous variants

DOC	Gene	All rare (*p*-value)		Rare functional (*p*-value)		Rare LOF (*p*-value)		Rare NS (*p*-value)		Rare CADD (*p*-value)	
		SKAT-O	SKAT Burden	SKAT-O	SKAT Burden	SKAT-O	SKAT Burden	SKAT-O	SKAT Burden	SKAT-O	SKAT Burden
30x	***ACMSD***	0.124	0.788	0.101	0.475	NV	NV	0.101	0.475	0.098	0.299
	BSTI	0.040	0.020	**0.0009**	**0.0004**	0.348	0.653	0.001	**0.0007**	0.011	0.005
	CCDC62	0.006	0.357	0.048	0.975	0.080	0.596	0.282	0.203	0.134	0.074
	DDRGK1	0.152	0.112	0.745	0.527	0.765	0.587	0.746	0.742	0.919	0.677
	DGKQ	0.067	0.062	0.067	0.062	NV	NV	0.183	0.188	NV	NV
	FGF20	0.059	0.020	0.142	0.552	0.162	0.149	0.162	0.149	0.052	0.041
	GAK	0.195	0.836	0.386	0.565	0.885	0.642	0.781	0.613	0.701	0.684
	GPNMB	0.189	0.798	0.311	0.205	0.547	0.665	0.447	0.304	0.530	0.348
	HIP1R	0.166	0.940	0.219	0.609	NV	NV	0.072	0.993	0.056	0.671
	ITGA8	0.282	0.726	0.379	0.945	0.346	0.648	0.382	0.997	0.288	0.873
	LAMP3	**0.0006**	0.478	0.189	0.322	0.787	0.300	0.318	0.724	0.524	0.601
	MAPT	0.063	0.132	0.003	0.001	1	1	0.063	0.037	0.165	0.101
	MCCC1	0.426	0.413	0.303	0.743	0.347	0.649	1	0.866	0.886	0.690
	PM20D1	0.333	0.463	0.844	0.643	0.674	0.464	0.645	0.479	0.569	0.416
	RAB25	0.561	0.395	0.807	0.802	0.777	0.968	0.565	0.637	NV	NV
	RAB29	0.252	0.222	0.660	0.425	NV	NV	0.777	0.967	0.777	0.967
	RIT2	0.395	0.242	0.023	0.266	NV	NV	0.576	0.495	0.576	0.495
	SETD1A	0.667	0.500	0.241	0.452	0.171	0.166	0.300	0.940	0.310	0.399
	SLC41A1	0.073	0.167	0.055	0.031	0.060	0.065	0.754	0.433	0.257	0.213
	STK39	0.098	0.789	0.055	0.094	0.174	0.171	0.160	0.247	0.160	0.247
	SIPA1L2	0.024	0.095	0.028	0.176	0.886	0.806	0.032	0.379	0.044	0.084
	STX1B	0.504	0.682	NV	NV	NV	NV	NV	NV	NV	NV
	SYT11	0.365	0.887	0.737	0.582	1	1	0.673	0.465	0.673	0.465
	TMEM163	0.598	0.411	0.170	0.165	NV	NV	0.170	0.165	0.170	0.165
	USP25	0.012	0.325	0.102	0.150	NV	NV	0.048	0.067	0.052	0.096
50x	***ACMSD***	0.127	0.538	0.102	0.188	NV	NV	0.102	0.188	0.087	0.093

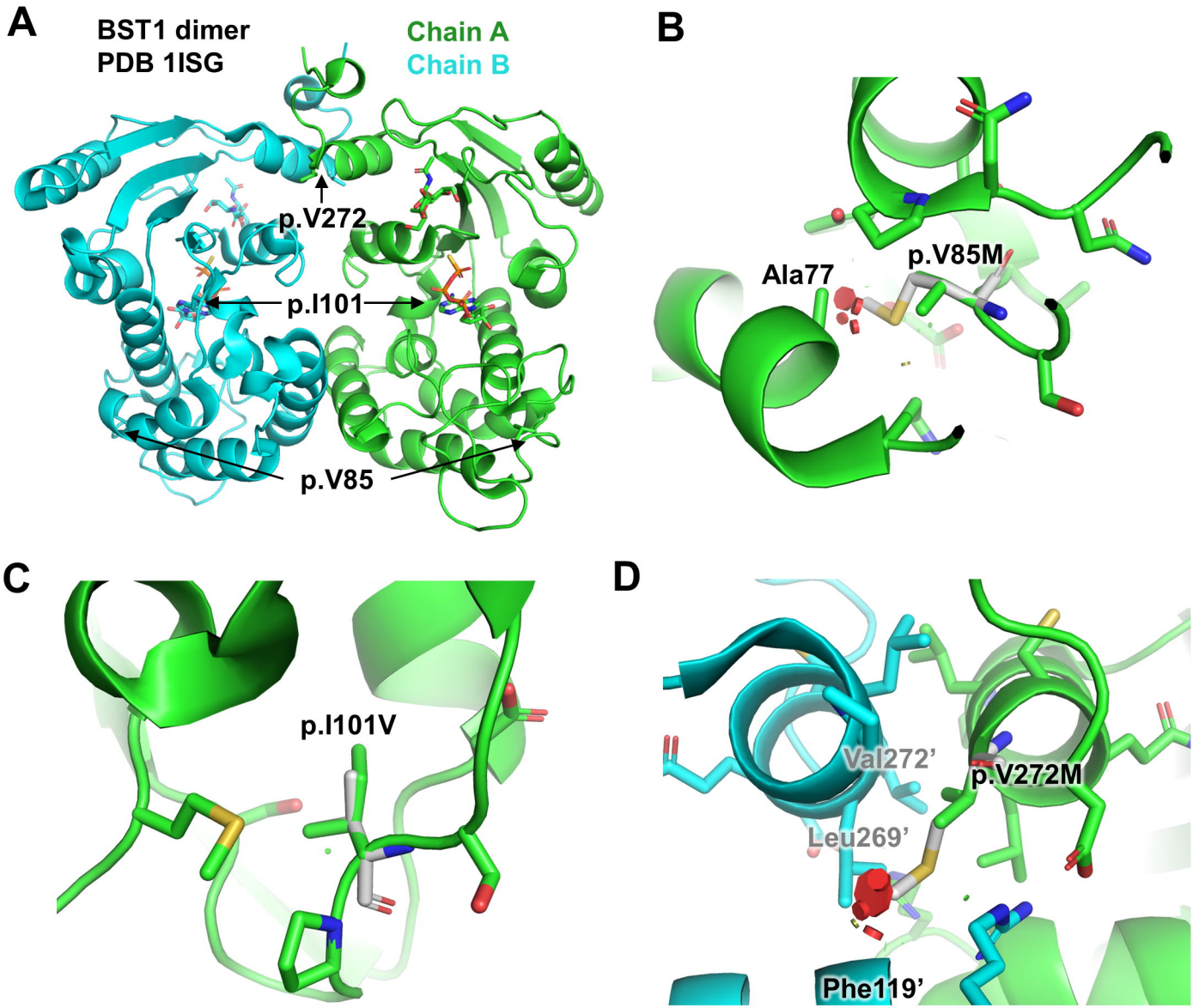
	BSTI	0.018	0.009	**0.0007**	**0.0003**	0.671	0.470	0.001	**0.0009**	0.010	0.006
	CCDC62	0.011	0.012	0.031	0.018	0.056	0.047	0.027	0.014	0.005	0.004
	DDRGK1	0.368	0.314	0.899	0.811	NV	NV	0.899	0.811	0.899	0.811
	DGKQ	NV	NV	NV	NV	NV	NV	NV	NV	NV	NV
	FGF20	0.015	0.030	0.164	0.153	0.164	0.153	NV	NV	0.164	0.153
	GAK	0.195	0.836	0.386	0.565	0.885	0.642	0.781	0.613	0.701	0.684
	GPNMB	0.046	0.995	0.177	0.127	0.108	0.413	0.343	0.215	0.283	0.189
	HIP1R	NV	NV	NV	NV	NV	NV	NV	NV	NV	NV
	ITGA8	0.184	0.315	0.402	0.540	0.340	0.634	0.428	0.584	0.345	0.942
	LAMP3	0.004	0.930	0.047	0.458	0.785	0.304	0.020	0.016	0.458	0.513
	MAPT	0.605	0.503	0.619	0.659	NV	NV	0.619	0.659	0.619	0.659
	MCCC1	0.26	0.151	0.972	0.752	0.334	0.618	0.951	0.745	0.958	0.800
	PM20D1	0.903	0.934	0.873	0.758	0.662	0.487	1	0.758	0.816	0.666
	RAB25	0.173	0.169	NV	NV	NV	NV	NV	NV	NV	NV
	RAB29	NV	NV	NV	NV	NV	NV	NV	NV	NV	NV
	RIT2	0.953	0.874	0.667	0.478	NV	NV	0.667	0.478	0.667	0.478
	SETD1A	0.212	0.147	0.671	0.469	NV	NV	0.671	0.469	0.671	0.469
	SLC41A1	0.132	0.189	NV	NV	NV	NV	NV	NV	NV	NV
	STK39	0.073	0.808	0.051	0.087	0.160	0.146	0.160	0.146	0.160	0.146
	SIPA1L2	0.163	0.098	0.598	0.861	NV	NV	0.777	0.963	0.777	0.963
	STX1B	NV	NV	NV	NV	NV	NV	NV	NV	NV	NV
	SYT11	0.0896	0.671	0.672	0.467	NV	NV	NV	NV	NV	NV
	TMEM163	0.308	0.237	0.169	0.163	NV	NV	0.169	0.163	0.169	0.163
	USP25	0.127	0.418	0.027	0.019	NV	NV	0.027	0.019	0.069	0.037

The table shows *p*-values of the SKAT and SKAT-O analysis in the specified subgroups; in bold - result significant after Bonferroni correction for the number of genes and depths; DOC: Depth of coverage; LOF: Loss of function; NS: Nonsynonymous; CADD: Combined annotation dependent depletion; SKAT-O: Optimized sequence kernel association test; SKAT: Kernel association test; NV: No variants were found for this filter.

Table 2. Summary of all samples carrying two nonsynonymous variants detected in the present study

Gene	Status	Sex	AAS/ AAO	dbSNP	Allele*	Substitution	F_A	F_C	gnomAD ALL	gnomAD NFE
GPNMB	A	M	75/50	7:23300122	T/C	p.S250P	0.000528	0	-	-
				7:23313795	G/T	p.Q545H	0.0005285	0	-	-
MAPT	A	M	85/66	17:44051771	G/T	p.D81Y	0.0005656	0	-	-
				rs63750612	G/A	p.A120T	0.0005423	0	0.0035	7.17E-05
SIPA1L2	A	F	64/48	rs184013125	G/A	p.S1482L	0.002665	0.008451	0.0057	0.0092
				rs200917620	G/A	p.R1089W	0.000547	0.000293	9.44E-05	0.0002
SIPA1L2	A	F	66/-	rs184013125	G/A	p.S1482L	0.002665	0.008451	0.0057	0.0092
				rs200293380	C/T	p.D1088N	0.006031	0.004695	0.0028	0.0041
USP25	A	M	68/-	rs377694221	C/T	p.A56V	0.000533	0	5.35E-05	7.23E-05
				rs200059109	G/T	p.D815Y	0.000533	0	0.0002	1.90E-05
MCCC1	C	M	35	rs149017703	C/T	p.G648S	0.001063	0.001683	0.0005	0.0009
				3:182810207	T/G	p.H88P	0	0.0002804	8.12E-06	8.95E-06
PM20D1	C	M	50	rs145195839	G/A	p.A332V	0.001059	0.000844	0.0003	0.0004
				rs14160575	G/T	p.P281Q	0.005325	0.005936	0.0057	0.0058
SIPA1L2	C	M	48	1:232615440		p.Y673F	0	0.0002817	-	-
				rs761063595		p.S266N	0	0.0002817	4.07E-06	8.99E-06
SIPA1L2	C	M	-	rs61729754	T/G	p.M1427L	0.009062	0.005634	0.0058	0.0088
				rs200293380	C/T	p.D1088N	0.006031	0.004695	0.0028	0.0041
SIPA1L2	C	M	28	rs184013125	G/A	p.S1482L	0.002665	0.008451	0.0057	0.0092
				1:232581435	T/C	p.T1065A	0	0.000282	-	-
SIPA1L2	C	M	50	rs184013125	G/A	p.S1482L	0.002665	0.008451	0.0057	0.0092
				rs966761148	C/T	p.V822M	0	0.000282	2.03E-05	2.69E-05

The table represents all the carriers of biallelic nonsynonymous variants in the target genes; A: Affected; C: Control; M: Male; F: Female; AAS: Age at sampling; AAO: Age at onset of iRBD; dbSNP: Single nucleotide polymorphism database; *Allele: Reference allele/mutant allele; F_A: Frequency in affected; F_C: Frequency in controls; gnomAD ALL: Exome allele frequency in all populations; gnomAD NFE: Exome allele frequency in non-Finnish European.


A
BST1 dimer
PDB 1ISG
Chain A
Chain B
p.V272
p.I101
p.V85
B
Ala77
p.V85M
C
p.I101V
D
Val272'
p.V272M
Leu269'
Phe119'