Published in final edited form as:

Mov Disord. 2018 November ; 33(11): 1821–1823. doi:10.1002/mds.103.

Coding Variation in *GBA* Explains the Majority of the SYT11-*GBA* Parkinson's Disease GWAS Locus

Cornelis Blauwendraat, PhD¹, Jose M. Bras, PhD², Mike A. Nalls, PhD^{2,3}, Patrick A. Lewis, PhD^{4,5}, Dena G. Hernandez, PhD¹, Andrew B. Singleton, PhD^{1,*}, and International Parkinson's Disease Genomics Consortium

¹Laboratory of Neurogenetics, National Institute on Aging, Bethesda, Maryland, USA

²UK Dementia Research Institute at UCL and Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK

³Founder/consultant with Data Tecnica International, Glen Echo, Maryland, USA

⁴School of Pharmacy, University of Reading, Reading, UK

⁵Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK

Keywords

genome-wide association; Parkinson's disease; *GBA*; risk allele; glucocerebrosidase

A landmark discovery in the genetics of Parkinson's disease (PD) was the determination that mutations in the gene encoding glucocerebrosidase (*GBA*) convey risk for PD and dementia with Lewy bodies.^{1,2} Mutations in *GBA* are also causing the autosomal-recessive disorder Gaucher's disease (GD). GD is a lysosomal storage disorder that presents with a variety of features, including parkinsonism. Although it has been clear that some GD-causing *GBA* mutations such as p.N409S are risk factors for PD, there has been some controversy regarding *GBA* mutations that do not cause GD and the risk for PD. Resolving this issue and the effect of such mutations on, for example, glucocerebrosidase activity will have important implications for our understanding of the pathobiology of PD.

We present here a large genetic assessment of *GBA* using NeuroX genotyping of 6249 PD cases and 6032 controls.³ In total, we detected 11 *GBA* coding variants. No large deletions or duplications at the *GBA* locus were identified. As expected p.N409S and p.E365K were enriched in cases compared with controls (Table 1). Unfortunately, another relatively common disease associated, L483P, could not be genotyped.¹ Interestingly, variant p.T408M, reported as uncertain pathological significance, was also significantly overrepresented in cases. In total, the majority of identified *GBA* variants were more

*Correspondence to: Andrew B. Singleton, PhD, Laboratory of Neurogenetics, NIA IRP NIH, 35 Convent Drive, Bethesda, MD 20892; singleton@mail.nih.gov.

Relevant conflicts of interest/financial disclosures: Drs. Nalls, Blauwendraat, Lewis, Hernandez, Bras, and Singleton have nothing to disclose.

Supporting Data

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

common in cases compared with controls, the exception being p.K13R, which is also reported as benign (Table 1). Although some individuals were carrying multiple *GBA* mutations, no high linkage disequilibrium patterns were identified between coding variants (Supplementary Results).

In previous PD genome-wide association studies (GWAS), a locus approximately 150 kb from *GBA* was implicated as a risk factor for PD (nominated as *SYT11-GBA*).⁴ Analysis of the non-coding variant rs71628662, previously implicated, revealed a strong association for disease (Table 1). Previous analyses suggested that this was independent of known GD-causing mutations, and the mechanism of action at this locus was therefore unclear. Conditioned analysis of rs71628662 on each of the identified coding variants revealed that the association at p.E365K completely accounted for the previously identified GWAS signal (conditional p.E365K rs71628662, $P = 0.6234$). Within the current data set, the linkage disequilibrium between p.E365K and rs71628662 is very high ($D' = 0.987$, $r^2 = 0.84$).

Previously, another noncoding variant (rs114138760) was also identified as an independent signal.⁵ In the NeuroX data set, this variant also showed association with disease ($P = 0.0001853$) and remained significant following adjustment for the primary GWAS variant ($P = 0.0001511$). In addition, this signal also remained significant following adjustment for p.E365K, p.N409S, and p.T408M ($P > 0.007146$), and of the 15 rs114138760 risk variant carriers, none carry p.L483P.

In summary, this work affirms that functional alleles within *GBA* (p.E365K and p.T408M) that are insufficient to cause GD variants are robustly associated with PD. Further, we have shown that a previously reported GWAS signal at this locus can be largely explained by protein coding variability within *GBA*. These data are consistent with p.E365K being the functional effector allele underlying the original GWA locus (*GBA-SYT11* locus), as suggested previously.⁶ The results from this current study suggest that efforts directed at replacing or augmenting glucocerebrosidase activity already under way for PD may have wider-ranging applicability than just for individuals with GD-associated mutations in glucocerebrosidase.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding agencies: This work was supported in part by the Intramural Research Programs of the National Institute of Neurological Disorders and Stroke (NINDS), the National Institute on Aging (NIA) both part of the National Institutes of Health, Department of Health and Human Services; project numbers 1ZIA-NS003154 and Z01-AG000949-02.

References

1. Sidransky E, Nalls MA, Aasly JO, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med* 2009; 361(17):1651–1661. [PubMed: 19846850]

2. Aharon-Peretz J, Rosenbaum H, Gershoni-Baruch R. Mutations in the glucocerebrosidase gene and Parkinson's disease in Ashkenazi Jews. *N Engl J Med* 2004;351(19):1972–1977. [PubMed: 15525722]
3. Nalls MA, Bras J, Hernandez DG, et al. NeuroX, a fast and efficient genotyping platform for investigation of neurodegenerative diseases. *Neurobiol Aging* 2015;36(3):1605.e1607–1612.
4. Nalls MA, Pankratz N, Lill CM, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* 2014;46(9):989–993. [PubMed: 25064009]
5. Mazzulli JR, Zunke F, Tsunemi T, et al. Activation of β -glucocerebrosidase reduces pathological α -synuclein and restores lysosomal function in Parkinson's patient midbrain neurons. *J Neurosci* 2016;36(29):7693–7706. [PubMed: 27445146]
6. Berge-Seidl V, Pihlstrøm L, Maple-Grødem J, et al. The GBA variant E326K is associated with Parkinson's disease and explains a genome-wide association signal. *Neurosci Lett* 2017; 658:48–52. [PubMed: 28830825]

Table 1.

GBA variant frequencies in cases and controls

Amino acid	GRCh38 (bp)	rs Number	Clinvar	HGMD pathogenicity	Cases				Controls				ExAC (EUR)			P
					M/M	M/W	W/W	MAF	M/M	M/W	W/W	MAF	M/M	MAF	OR	
K13R (K>27R)	155240707	rs150466109	Benign	DM?	0	0	6249	0	0	2	6030	0.00017	0	0.00021	NA	0.999
C62Y (C23Y)	155240008	rs145888253	na	na	0	1	6248	0.00008	0	0	6032	0	0	0.00001	NA	1
E150K (E111K)	155239622	none	na	DM?	0	1	6248	0.00008	0	0	6031	0	0	0	NA	0.9993
F255Y (F216Y)	155237576	rs74500255	Likely	Pathogenic	0	2	6247	0.00016	0	1	6031	0.00008	0	0	NA	0.5827
F298L (F259L)	155237446	none	na	pathogenic	0	1	6248	0.00008	0	0	5990	0	0	0	NA	0.9993
A348V (A309V)	155236426	rs78396650	Pathogenic	Pathogenic	0	5	6244	0.00040	0	0	6032	0	0	0	NA	0.9984
E365K (E326K)	155236376	rs2230288	Conflicting interpretations	DM?	4	256	5989	0.02158	0	121	5911	0.01013	7	0.01196	2.037	1.64E-09
T408M (T369M)	155236246	rs75548401	Conflicting interpretations	DM?	2	169	6078	0.01404	0	107	5925	0.00895	5	0.00976	1.53	0.001188
N409S (N370S)	155235843	rs76763715	Likely pathogenic	Pathogenic	0	87	6161	0.00701	0	41	5991	0.00341	1	0.00363	1.882	0.002014
D419N/Y	155235814	none	na	Pathogenic	0	2	6247	0.00016	0	0	6032	0	0	0	NA	0.9991
(D380N/Y) E427K (E388K)	155235790	rs149171124	Uncertain significance	DM?	0	10	6239	0.00080	0	5	6027	0.00041	0	0.00031	NA	0.2157
na	155390201	rs71628662	na	na	1	187	6059	0.01513	0	88	5935	0.007305	na	na	2.07	1.324E-07
na	154925709	rs114138760	na	na	1	166	6082	0.01344	0	107	5925	0.008869	na	na	1.52	0.0001853

Indicated are uncorrected *P*-values (Bonferroni correction for 0.05 lies at > 0.0045) and odds ratios. Tests were carried out in plink using logistic regression with covariates age, sex, first 20 principal components based on genome-wide genotype data, and sample provenance. Na, not applicable; bp, base pair; W, wild type; M, mutant; MAF, minor allele frequency; EUR, European; ExAC, Exome Aggregation Consortium (<http://exac.broadinstitute.org/>); OR, odds ratio; HGMD, Human Gene Mutation Database.