Supplementary Material

An Exploratory Study Using Electronic Medical Records to Assess the Feasibility of Establishing Cohorts of Patients with Genetic Causes of Parkinson's Disease

Supplementary Table 1. Parkinson's disease family history questionnaire

1. Are your biological grandparents (maternal and paternal) of Ashkenazi Jewish ancestry?

2. Do you have any blood relatives (parents, grandparents, siblings, or children) with Parkinson's Disease?

3. Which of your blood relatives (grandparents, biological mother, biological father, full siblings, half siblings, biological children) have Parkinson's Disease?

Response options for all questions:	Yes	No	unknown	prefer not to respond

Supplementary Table 2. Participants characteristics

	N (%)		
Participants in population	837 (100%)		
Gender	N (%)		
Female	353 (42.2%)		
Male	484 (57.8%)		
Age (y)	Female (%)	Male (%)	Total (%)
<18	0 (0%)	0 (0%)	0 (0%)
18 to 29	0 (0%)	0 (0%)	0 (0%)
30 to 39	0 (0%)	0 (0%)	0 (0%)
40 to 49	6 (0.7%)	6 (0.7%)	12 (1.4%)
50 to 59	12 (1.4%)	38 (4.5%)	50 (5.9%)
60 to 64	38 (4.5%)	44 (5.3%)	82 (9.8%)
≥ 65	297 (35.5%)	396 (47.3%)	693 (82.8%)

Supplementary Table 3. Family history of Ashkenazi ancestry among genotyped subjects with Parkinson's disease

Self-reported Ashkenazi Ancestry by Number of Grandparents	Fraction of Total Genotyped (n/N)	Percent Genotyped (%)
4	540/788	68.5
1-3	31/788	3.9
0	213/788	27.0
Unknown	4/788	0.5

N, Total number of subjects with definitive or clinically established Parkinson's disease (788); n, number of subjects.

Genotyping Assays

LRRK2 pP.(Gly2019Ser) variant (rs34637584)

An allelic discrimination assay from the pre-designed TaqMan SNP Genotyping Assay Collection (http://www.lifetechnologies.com/taqmansnp) was used for rs34637584. A brief overview of the reaction set up is as follows: Into each well of a 384-well reaction plate, 3 µl assay master mix cocktail (genotyping master mix and TaqMan primers/probes) was combined with 2 µl DNA sample (10 ng per reaction well). PCR was performed in an Eppendorf Master cycler thermal cycler using the following protocol: 95°C, 10 min hot start and 40 cycles of 92°C for 15 s and 60°C for 1 min. After PCR amplification, an endpoint post-PCR plate reading on ViiA7 real-time PCR system was performed. All p.(Gly2019Ser)G2019S mutations were confirmed using Sanger sequencing.

GBA genotyping of 7 variants (rs387906315, rs2230288, rs421016, rs76763715, rs104886460, rs80356769, and rs80356773)

To eliminate the problem of pseudogene contamination, a selective PCR amplification of the whole *GBA* DNA sequence by long range PCR (LR-PCR) was used, as described by Finckh et al. [1]. Three Sanger sequencing reactions of the long-range PCR products were performed to identify seven SNPs using 3 pairs of forward and reverse primers:

Fragment 1: Length 293 bp;
Primers – Fw: 5'-CCA GGA GAG TAG TTG AGG GGT GG-3'
Rv: 5'-CCC CAA AGT TGG TCT CAG TCA CTC-3'
Contains SNPs: rs387906315 and rs104886460.

Fragment 2: Length 830 bp;
Primers – Fw: 5'-CCT GTG TGC AAG GTC CAG GAT C-3'
Rv: 5'-CCT GAA GTG GCC AAG GTG GTA G-3'
Contains SNPs: rs2230288, rs76763715 and rs80356769.

Fragment 3: Length 505 bp;
Primers – Fw: 5'-CTTAGATGAGGGTTTCATGGGAGGTAC-3'

Rv: 5’-GGGCTTACGTCGCTGTAAGCTC-3’
Contains SNPs: rs421016 and rs80356773.

DNA sequencing data generated as a result of Sanger sequencing was analyzed using Sequencher 5.1 software. SNP alleles were called from both forward and reverse sequencing reactions.

REFERENCE

[1] Finckh U, Seeman P, von Widdern OC, Rolfs A (1998) Simple PCR amplification of the entire glucocerebrosidase gene (GBA) coding region for diagnostic sequence analysis. *DNA Seq* **8**, 349-356.