Parkinsonism and Related Disorders xxx (2017) 1–5

Contents lists available at ScienceDirect

Parkinsonism and Related Disorders

journal homepage: www.elsevier.com/locate/parkreldis

Genetic risk factors in Finnish patients with Parkinson's disease

Susanna Ylönen [a, b], Ari Siitonen [a, b], Michael A. Nalls [c], Pauli Ylikotila [d, e], Jaana Autere [f], Johanna Eerola-Rautio [g], Raphael Gibbs [c], Mikko Hiltunen [h], Pentti J. Tienari [g], Hilkka Soininen [h], Andrew B. Singleton [c], Kari Majamaa [a, b, *]

[a] Institute of Clinical Medicine, Department of Neurology, University of Oulu, Oulu, Finland
[b] Department of Neurology and Medical Research Center, Oulu University Hospital, Oulu, Finland
[c] Laboratory for Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA
[d] Institute of Clinical Medicine, Department of Neurology, University of Turku, Turku, Finland
[e] Division of Clinical Neurosciences, Turku University Hospital, Turku, Finland
[f] Neuro Center, Kuopio University Hospital, Kuopio, Finland
[g] Department of Neurology, Helsinki University Hospital, Research Programs Unit, Molecular Neurology, Biomedicum, University of Helsinki, Helsinki FIN-02900, Finland
[h] Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland

ARTICLE INFO

Article history:
Received 18 July 2017
Received in revised form
30 August 2017
Accepted 28 September 2017

Keywords:
Molecular epidemiology
Neurodegenerative diseases
Mutation
Mitochondrial
Gene

ABSTRACT

Introduction: Variation contributing to the risk of Parkinson's disease (PD) has been identified in several genes and at several loci including *GBA*, *SMPD1*, *LRRK2*, *POLG1*, *CHCHD10* and *MAPT*, but the frequencies of risk variants seem to vary according to ethnic background. Our aim was to analyze how variation in these genes contributes to PD in the Finnish population.
Methods: The subjects consisted of 527 Finnish patients with early-onset PD, 325 patients with late-onset PD and 403 population controls. We screened for known genetic risk variants in *GBA*, *SMPD1*, *LRRK2*, *POLG1*, *CHCHD10* and *MAPT*. In addition, DNA from 225 patients with early-onset Parkinson's disease was subjected to whole exome sequencing (WES).
Results: We detected a significant difference in the length variation of the CAG repeat in *POLG1* between patients with early-onset PD compared to controls. The p.N370S and p.L444P variants in *GBA* contributed to a relative risk of 3.8 in early-onset PD and 2.5 in late-onset PD. WES revealed five variants in *LRRK2* and *SMPD1* that were found in the patients but not in the Finnish ExAC sequences. These are possible risk variants that require further confirmation. The p.G2019S variant in *LRRK2*, common in North African Arabs and Ashkenazi Jews, was not detected in any of the 849 PD patients.
Conclusions: The *POLG1* CAG repeat length variation and the *GBA* p.L444P variant are associated with PD in the Finnish population.

1. Introduction

Genetic variation contributing to the risk of Parkinson's disease (PD) has been identified. One of the genes involved is glucosylceramidase beta (*GBA*) that codes for a lysosomal enzyme. It is traditionally associated with Gaucher's disease, but heterozygous mutations are an established risk factor for Parkinson's disease. The two most common mutations in *GBA* include p.L444P and p.N370S. *SMPD1* encodes another gene involved in lysosomal function. It codes for the protein sphingomyelin phosphodiesterase 1 that generates ceramide by cleaving the phosphocholine group of sphingomyelin. Mutations in *SMPD1* are found in Niemann-Pick disease and the variant p.L302P has been suggested to be associated with PD [1].

The *LRRK2* (leucine-rich repeat kinase 2) gene encodes a 2527-amino acid protein with several structural and functional domains such as leucine-rich repeat and kinase. The common p.G2019S mutation is located in the kinase domain and it has been shown to increase the kinase activity [2]. The frequency of the p.G2019S mutation varies greatly in different populations being most common in North African Arabs, where it is found in 39% of sporadic

* Corresponding author. University of Oulu, Department of Neurology, P.O. Box 5000, FIN-90014 Oulu, Finland.
E-mail address: kari.majamaa@oulu.fi (K. Majamaa).

https://doi.org/10.1016/j.parkreldis.2017.09.021
1353-8020/

and 36% of familial cases, and Ashkenazi Jews, where it is found in 10% of sporadic and 28% of familial cases. Among Europeans, the frequency is 4% in sporadic cases in Portugal and 14% in familial cases, whereas in Spain, Italy and France it is 2–4%, and in Sweden and Norway the frequency is 1%. In Asian countries the mutation is very rare being found in <0.1% of patients [3]. The penetrance of p.G2019S is high, but not complete, so that the risk of PD is 74% at the age of 79 years [3].

Other genes, where variation may contribute to the risk of PD, include *POLG1*, *CHCHD10* and *MAPT*. Polymerase γ (*POLG1*) is responsible for the replication of the mitochondrial genome. The second exon of the gene harbors a microsatellite consisting of ten CAG trinucleotide repeats that is translated into a tract of ten glutamine moieties (10Q). Genotypes different from the common 10Q variant have been associated with a higher risk of PD [4]. *CHCHD10* encodes a coiled-coil helix coiled-coil helix protein that is involved in the maintenance of mitochondrial cristae integrity [5]. The variant p.S59L in *CHCHD10* has been associated with a complex phenotype consisting of amyotrophic lateral sclerosis, frontotemporal lobar degeneration, cerebellar ataxia, parkinsonism, and myopathy [5] and, subsequently, another variant p.G66V has been identified [6]. *MAPT* (microtubule-associated protein tau) is involved in the assembly and stabilization of microtubules. Of the two haplotypes H1 and H2, haplotype H1 has been reported to be slightly more common in PD patients than in population controls [7].

Previous data have shown that the frequencies of these mutations in patients with PD vary according to ethnic background [8]. Therefore, we screened for the most common known mutations in the genes *GBA*, *SMPD1*, *LRRK2* and *CHCHD10*, determined the length variation in *POLG1* and determined the *MAPT* haplotype in Finnish patients with Parkinson's disease and population controls.

2. Subjects and methods

2.1. Patients and controls

The study group consisted of a national cohort of 441 patients with EOPD defined by age of onset <55 years [9]. In addition, a case series collected at the Kuopio University Hospital during one year [10] comprised of 209 patients with late-onset Parkinson's disease (LOPD) and 55 patients with EOPD, and a case series collected at the Helsinki University Hospital [4] comprised 116 patients with LOPD and 31 patients with EOPD. Previous analyses on these cohorts have not revealed causative mutations in the common PD genes. Controls consisted of 403 healthy blood donors from three different regions of Finland. Part of the samples in the national cohort of EOPD (N = 225) were subjected to whole exome sequencing (WES). The study has been approved by the Ethics Committee of Turku University Hospital, the Ethics Committee of the Medical Faculty of the University of Kuopio, the Ethics Committee of Helsinki University Hospital and the Ethics committee of the Finnish Red Cross. Written informed consent was obtained from all patients prior to participating the study.

2.2. Molecular methods

Genomic DNA was extracted from peripheral blood using standard protocols. The variations in *GBA* (p.N370S, p.L444P), *SMPD1* (p.L302P), *LRRK2* (p.R1441C/G/H, p.G2019S), *CHCHD10* (p.S59L, p.G66V) and *MAPT* (haplotype H1) were detected using a PCR-restriction digestion protocol. The primers and restriction enzymes used are presented in supplementary table. The detected variations were confirmed by sequencing.

The CAG repeat in exon 2 of the *POLG1* gene was analyzed by fragment length analysis. A fragment containing the CAG repeat was amplified using a FAM-labeled forward primer CTCCGAGGATAGCACTTGC and a reverse primer CTGGGTCTCCAGCTCCGT (CHLC.GCT14A01.P16693.1 and CHLC.GCT14A01.P16693.2, respectively; GenBank accession number G16014). The fragment length was 124 bp in the presence of the most common allele containing ten CAG repeats. The fluorescent-labeled PCR products in the EOPD group were separated on an *ABI PRISM*® 3100 Genetic Analyzer (Perkin Elmer, Foster City, CA, U.S.A.). GeneScan™ −500 LIZ® Size Standard (Applied Biosystems, Foster City, CA, U.S.A.) was used as an internal standard and was run along with each sample. The alleles were identified using the Peak Scanner™ Software Version 1.0 (Applied Biosystems). In the LOPD group and the controls an ABI PRISM™ 377 DNA Sequencer (Perkin Elmer) was used with Genescan version 2.1 fragment analysis software (Perkin Elmer). GENESCAN-500™ TAMRA (Applied Biosystems) was used as an internal standard and the Genotyper 2.0 program (Perkin Elmer) was used for identifying the alleles.

WES was carried out as described previously [11]. Annovar [12], SNPEff [13] and SNPSift [14] were used for functional annotations. The data were unfiltered so that even the unlikely SNPs could be detected. As controls we used 563 STAMPEED population controls from the North Finland Birth Cohort 1966. WES has recently revealed several mutations, but no splice mutations, in established EOPD genes among the 225 EOPD patients [11]. Clinical significance of the mutations is, however, unknown. Nonsynonymous mutations in *LRRK2* and *SMPD1* that were more frequent among the patients than among the controls were verified by direct sequencing and were analysed for their pathogenic potential by using PredictSNP [15].

2.3. Statistical analysis

Fisher exact test was used to compare allele frequencies between patients and controls. The frequency distribution of the CAG repeat alleles in the *POLG1* gene was analyzed with the exact test of population differentiation using the Arlequin software [16].

3. Results

3.1. Genetic variants contributing to the risk of EOPD

None of the EOPD patients harbored *SMPD1* p.L302P, *LRRK2* p.R1441C/G/H, *LRRK2* p.G2019S, *CHCHD10* p.S59L or *CHCHD10* p.G66V. The frequency of haplotype H1 in *MAPT* was 92.6% in the patients and 95.0% in the 292 controls (p = 0.068 for difference, Table 1). The length of the CAG repeat in the *POLG1* gene varied between 8Q and 13Q in the EOPD patients and between 8Q and 12Q in the controls. The frequency distribution of the CAG repeat alleles

Table 1
MAPT H1/H2 genotype and allele distribution.

(A) Genotypes	H1/H1 N (%)	H1/H2 N (%)	H2/H2 N (%)	Total N
Controls	263 (90.1)	29 (9.9)	0	292
EOPD	452 (86.1)	68 (13.0)	5 (1.0)	525
LOPD	290 (89.8)	30 (9.3)	3 (0.9)	323
(B) Alleles	**H1 N (%)**		**H2 N (%)**	
Controls	555 (95.0)		29 (5.0)	584
EOPD	972 (92.6)		78 (7.4)	1050
LOPD	610 (94.4)		36 (5.6)	646

EOPD, early-onset Parkinson's disease; LOPD, late-onset Parkinson's disease.

Please cite this article in press as: S. Ylönen, et al., Genetic risk factors in Finnish patients with Parkinson's disease, Parkinsonism and Related Disorders (2017), https://doi.org/10.1016/j.parkreldis.2017.09.021