Investigation of the role of genetic architecture in Parkinsonism

PhD thesis

Anett Illés

Doctoral School of Molecular Medicine
Semmelweis University

Supervisor: Mária Judit Molnár, MD, D.Sc.

Official reviewers: Ildikó Sipos, MD, Ph.D.
Eszter Hidasi, MD, Ph.D.

Head of the Final Examination Committee: György Nagy, MD, D.Sc.

Members of the Final Examination Committee:
Barbara Molnár-Érsek, Ph.D.
Róbert Szegedi, MD, Ph.D.

Budapest
2020

Introduction

Parkinson's disease (PD) is the second most common progressive neurodegenerative disorder. According to the latest literature, around 6 million people are affected worldwide. Approximately 2% of the population are affected over the age of 60, compared to 4% of the population over the age of 80. However, 10% of the disease occurs at young age (20-50 years), which is more likely to be caused by genetic factors.

The main cause of motor dysfunction is the destruction of dopaminergic neurons in the substantia nigra pars compacta. In PD patients, the prevalence of non-motor symptoms increases with both the severity of disease and the degree of progression and may also predict the onset of motor symptoms. This is criucial in therapeutic practice, as 70-80% of the dopaminergic neurons in the substantia nigra is already dead at the onset of motor symptoms. Furthermore, predicting the development of Lewy bodies has significant clinical benefit, in which case α-synuclein-targeted therapies may be effective. In addition to abnormal protein aggregation, mitochondrial

dysfunction, neuroinflammation and excitotoxicity also play an important role in the progression of PD.

Several studies have found significant differences between clinical and pathological diagnoses in many cases. In addition, genetic heterogeneity makes difficult to understand the correlation, as numerous different genetic causes could exist behind the similar phenotype. Moreover, individual differences of the phenotype could be manifest in patients with the same genetic variant.

In general, monogenic familial forms are characterized by rare, highly penetrant pathogenic mutations, whereas sporadic forms result from the co-existence of environmental and genetic risk factors. However, in most cases the two forms cannot be clearly distinguished based on clinical symptoms. Further researches are warranted to clarify the clinical significance of the identified variants.

Despite increasing genetic information, the genetic etiology of PD is still undetected in 40% of cases. The better availability of next-generation sequencing (NGS) has made it possible to identify several potential genes that may play an important role in the development of PD. As

various neurological disorders are manifested with overlapping clinical symptomps, a growing number of genes have been linked to PD, which have been previously associated with other neurodegenerative diseases. The deeper understanding of the genetic forms of PD has highlighted the importance of synaptic transmission, vesicular recycling, mitochondrial and protein quality control in the pathogenesis of the disease.

For environmental risk factors of PD, there are evidence that dysregulation of the dopamine transporter (DAT) also play a role in the disease patomechanism. Cocaine enhances dopaminergic signaling because it binds to DAT and prevents dopamine reuptake from synaptic clefts. Despite the evidence of a higher risk of PD in cocaine users is controversial, it has already been proved that the structure of the brain changes and the conformation of α-synuclein becomes more compact.

In summary, if the genetically heterogeneous nature of PD is taken into consideration, clarifying the genetic architecture of sporadic and familial PD could improve diagnostic accuracy. Consequently, it enables

presymptomatic diagnosis and testing of at-risk individuals in affected families. Furthermore, it may expand our understanding about the genetic and neuropathological mechanisms of the disease, which may be crucial in the development of new therapies. Finally, it improves the ability to categorize various PD patients into genetic subtypes. This can facilitate the effective treatment of patients through adequate intervention during the disease.

Objectives

Our aims can be summarized as follows:

1. Investigate the genetic background of Parkinsonism in the Hungarian population. To clarify the inheritance underlying the heritability and to develop an algorithm to be used in the genetic diagnosis of PD.
2. To investigate the presence and frequency of genetic disorders which is associated with Parkinsonism in Hungarian patients.
3. Study the correlation between the genetic architecture and the phenotype. To develop guidelines for genetic counseling for the communication of genetic results, and the interpretation of gene-environment interactions.
4. Developing a registry and biological sample collection of patients with Parkinson's disease, creating a trial-ready cohort to ensure our genetically stratified patients the greatest probability to participate in clinical trials.

Methods

Patients involved in genetic studies

Patients with clearly identifiable Parkinsonism according to the Movement Disorder Society (MDS) criteria, i.e. bradykinesis was present in either with resting tremor or with rigidity, or both, were included in our study. To the study, 152 EOPD (AOO 25-50 years), 30 familial LOPD (AOO> 50 years), and four JOPD (AOO <25 years) patients were selected. The genetic architecture of 66 patients were analyzed with NGS.

Methods of the genetic studies

For genetic testing, DNA was isolated from peripheral blood sample using the Qiagen DNA Blood Mini Kit. The most common PD-associated genes (*PRKN*, *PINK1*, *LRRK2*, *SNCA* and *PARK7*) were analyzed by Sanger bidirectional sequencing on an ABI PRISM 3500 Genetic Analyzer. The exonic copy number variations were examined by MLPA (Multiplex Ligation-Dependent Assay Amplification). The number of *C9orf72* GGGGCC hexanucleotide repeat was determined by repeat-primed PCR and subsequent fragment analysis. For NGS panel

sequencing a custom Agilent SureSelectQXT Target Enrichment kit, and for whole exome sequencing Agilent SureSelectQXT Human All Exon v5 reagent were used. Panel sequencing was performed on the Illumina MiSeq, and whole exome sequencing was carried out on the Hiseq 2500 platform. For the analysis process we used GATK HaplotypeCaller, SnpEff, ClinVar, dbNSFP, and VariantAnalyzer software. Variants were classified according to the ACMG (American College of Medical Genetics and Genomics) guidelines.

Odds ratios (ORs) were calculated using Medcalc software. The quantitative variables were described using mean ± standard deviation. The effect of the rare variant burden on the age of onset was analyzed by ANOVA. To investigate the possibility of an oligogenic effect, the variant burden was calculated in patients who were analyzed with NGS.

Results

Sequencing and MLPA

Rare, deleterious variants in PD-associated genes, which are potentially compatible with the monogenic PD

Duplication of exon 7 in the *PRKN* gene was confirmed in 4 sporadic cases and duplication of the *SNCA* gene in one sporadic case. In the *LRRK2* gene, a previously described pathogenic L1795F substitution was detected in two cases. Three new, potentially damaging variants were identified (*LRRK2*-Y1649S, *EIF4G1*-M1357T, *VPS35*-K552I). The *DNAJC13*-L2170W amino acid change, which is only a minor genetic risk factor according to the literature, has been identified in one patient. Variants, which cause amino acid changes, were described in familial PD patients (mean age of onset 41.5 ± 11.6 years).

Genetic risk variants, which were previously associated with PD

Genetic risk variants of PD, which were previously described in the literature, were identified in 51 cases. In the *GBA* gene, five heterozygous genetic risk variants (H294Q, E365K, T408M, N409S and L483P) have been

identified in our cohort. Several distinctive clinical features have been identified among patients with *GBA* alterations. In the *LRRK2* gene, heterozygous M1647T mutation was detected in six patients, while homozygous S1647T mutation was found in 27 patients. We identified *PINK1* A340T alteration in seven patients and G411S substitution in two patients in our cohort.

All detected differences, which were previously reported as a genetic risk factor, had an OR>1, excluding *PINK1* A340T, which had an OR=0.4. However, only the *GBA* T408M variant showed a significant correlation ($p < 0.05$).

„Multiple hit" mechanism that affects PD risk

By analyzing the data, we found 12 cases where a heterozygous deleterious mutation occurs in an AR-PD-associated gene, either with a previously described risk variant or with a rare substitution in another PD-associated gene. Risk variants were identified in the *LRRK2*, *GBA*, and *DNAJC13* genes. Probably damaging alterations were found in *PRKN*, *PINK1*, *SYNJ1*, *TMEM230*, *DNAJC6*, *C19orf12*, and *PLA2G6* genes. In some cases, a known risk factor (*LRRK2*-S1647T) was also found along with a

variant, which was identified in an AD-PD gene (*DNAJC13*-S790Y or *LRRK2*-I803T).

During the test of the oligogenic inheritance pattern, both the mean and the median of the age of onset were lower with increasing number of co-occurring variants, but this did not reach the level of statistical significance.

The Kruskal-Wallis statistical analysis showed that in patients diagnosed with Parkinsonism, tightening the screening criteria increase the number of variants in PD-associated genes compared to non-PD-diagnosed neurodegenerative patients and control subjects. However, this did not reach the level of statistical significance either.

Monoallelic heterozygous variant in AR-PD-associated genes

In this study, five genes with AR inheritance pattern haboured monoallelic heterozygous, probably damaging substitutions (*PRKN*-R234Q, D243N, R275W; *VPS13C*-S2904L, *DNAJC6*-F414Y, *CP*-1889M; *PLA2G6*-R396W) which might increase the susceptibility of PD as potential risk factors. Interestingly, the age of onset is like those

cases where a previously described risk factor was identified.

Prevalence of *POLG* mutation in the Hungarian PD cohort

The complete coding region of the *POLG* gene was analyzed in 67 patients with Parkinson's disease. During the analysis, 6 missense alteration were identified. Two of these have been previously described as polymorphisms (E1143G and Q1236H). Based on the literature, potentially pathogenic variants were found in two cases (compound heterozygous T251I+P587L alteration, and heterozygous G737R substitution, respectively). Moreover, the previously unpublished H613D variant was identified in a third case.

Prevalence of *C9orf72* repeat expansion in the Hungarian PD cohort

In our study, the *C9orf72* hexanucleotide repeat expansion was investigated using repeat-primed PCR in 147 Hungarian non-related patients with Parkinsonism. In two cases, the repeat number was above 30 (31 and 32 repeat number, respectively), which is 1.4% of the cases. The 7.5% (n=11) of our cases had intermedier expansion

between 23 and 30 repeats. The mean age of onset of patients with patogenic repeat expansion was 55.67±12.68 years. One of the positive cases had a positive family history. Cognitive impairment was confirmed in one positive case and in two intermedier cases. Levodopa treatment resulted in a positive response in 10 cases (one positive, nine intermedier cases) and no response was observed in one intermedier case (juvenile male patient). No significant correlation was found between *C9orf72* hexanucleotide repeat expansion and the age of onset.

Analyzing the interaction of environmental and genetic risk factors through a case study

A 44-year-old male patient with hand tremor, who used nasal cocaine regularly (~ 1 g / day - 15 mg / kg / day) for 18 months before onset of the symptom, was examined. He had stopped using cocaine 10 months before the examination.

Neurological examination suggested early Parkinsonism, but the tremor was atypical. It is important to highlight from the familial anamnesis that both the patient's father

and son had movement disorders (tremor, restless-leg syndrome). The brain MRI (lack of swallow tail sign) and the DaTscan suggested PD pathology. During the genetic analysis, in the *LRRK2* gene the homozygous S1647T risk factor was identified.

Clinical and DaTscan results suggested a Parkinson's syndrome, which is the result of the combined presence of a toxic environmental and a genetic risk factor. After one year of cocaine abstinence, the tremor decreased significantly. One and a half years after the first DaTscan, a follow-up study was performed which showed normal radiopharmacone uptake in the striatum with only slight asymmetry.

Conclusions

We were the first to perform a comprehesive genetic analysis of Parkinson's disease in Hungary.

In 8.1% of the whole PD cohort the cause of the disease was confirmed as monogenic, and in 32.3% of the patients a genetic risk factor was described. Monoallelic alterations, which were identified in AR-PD genes, were likely to increase the risk of developing PD as a susceptibility factor. Our results have contributed to increase the knowledge about the role of these monoallelic variants. Potentially pathogenic alterations, which were identified in the *POLG* gene, emphasized the importance of mitochondrial gene screening in patients with Parkinsonism.

The significance of "multiple hit" mechanism was also raised, as a “multiple hit” effect was assumed in 12 cases in the background of the symptoms. Statistical analysis of the oligogenic effect suggested a tendency that an increase of the potentially harmful, rare variants could be observed in the cohort. There is also a tendency that the age of onset

is changed inversely proportional to the number of potentially damaging variants.

In the Hungarian PD cohort, we were the first to demonstrate that *C9orf72* hexanucleotide repeat expansion may be associated with Parkinsonism. Based on these results, we have recommended that this gene should be included in the genetic diagnosis of PD.

The combined effect of cocaine, as an environmental risk factor, and a genetic risk factor in a patient with early Parkinsonism was demonstrated. Based on this case we assume that the molecular changes, resulting from cocaine use, are partially reversible. The relationship between cocaine use and Parkinsonism is a very complex issue, and the interpretation of each association contributes significantly to the better understanding of the pathomechanism of PD.

Based on our results, we develop a new genetic diagnostic recommendation for cost-effective diagnosis of Hungarian PD patients.

Bibliography of the candidate's publications

Publications related to the PhD thesis

1.) Illés A, Balicza P, Gál A, Pentelényi K, Csabán D, Gézsi A, Molnár V, Molnár MJ. (2020) Az örökletes Parkinson-kór mint a POLG-gén károsodásának új klinikai megjelenési formája [Hereditary Parkinson's disease as a new clinical manifestation of the damaged POLG gene]. Orvosi Hetilap, 161.20: 821–828. **IF: 0.564**

2.) Illés A, Csabán D, Grosz Z, Balicza P, Gézsi A, Molnár V, Bencsik R, Gál A, Klivényi P, Molnár MJ. (2019) The role of genetic testing in the clinical practice and research of early-onset Parkinsonian disorders in a Hungarian cohort: Increasing challenge in genetic counselling, improving chances in stratification for clinical trials. Frontiers in Genetics, 10:1061. **IF: 3.517**

3.) Illés A, Balicza P, Molnár V, Bencsik R, Szilvási I, Molnar MJ. (2019) Dynamic interaction of genetic risk factors and cocaine abuse in the background of

Parkinsonism–a case report. BMC neurology, 19.1:260. **IF: 2.233**

Publications not related to the PhD thesis

1.) Fekete B, Pentelényi K, Rudas G, Gál A, Grosz Z, Illés A, Jimoh I, Csukly G, Domonkos A, Molnar MJ. (2019). Broadening the phenotype of the TWNK gene associated Perrault syndrome. BMC Medical Genetics, 20(1), 1-8. **IF: 1.740**

2.) Melicher D, Illés A, Littvay L, Tárnoki ÁD, Tárnoki DL, Bikov A, Kunos L, Csabán D, Buzás EI, Molnár MJ, Falus A. (2019) Positive association and future perspectives of mitochondrial DNA copy number and telomere length – a pilot twin study. Archives of Medical Science, 15.1 **IF: 2.380**
Dóra Melicher and Anett Illés contributed equally to this work.

3.) Balicza P, Grosz Z, Molnár V, Illés A, Csabán D, Gézsi A, Dézsi L, Zádori D, Vécsei L, Molnár MJ. (2018)

NKX2-1 New Mutation Associated with Myoclonus, Dystonia, Pituitary Dysfunction and Empty Sella. Frontiers in genetics, 9:335. **IF: 3.517**

4.) Kecskeméti N, Szönyi M, Gáborján A, Küstel M, Milley GM, Süveges A, Illés A, Kékesi A, Tamás L, Molnár MJ, Szirmai Á, Gál A. (2018) Analysis of GJB2 mutations and the clinical manifestation in a large Hungarian cohort. Eur Arch Otorhinolaryngol. 275.10:2441-2448. **IF: 1.750**

5.) Balicza P, Grosz Z, Bencsik R, Illés A, Gál A, Gézsi A, Molnár MJ. (2018) Significance of whole exome sequencing in the diagnostics of rare neurological diseases - own experiences through a case presenting with ataxia]. Orvosi Hetilap, 159.28:1163-1169. **IF: 0.564**

6.) Varga NÁ, Pentelényi K, Balicza P, Gézsi A, Reményi V, Hársfalvi V, Bencsik R, Illés A, Prekop C, Molnár MJ. (2018) Mitochondrial dysfunction and autism: Comprehensive genetic analyses of children with autism

and mtDNA deletion. Behavioral and Brain Functions, 14.1:4. **IF: 2.457**

7.) Melicher D, Illés A, Pállinger É, Kovács ÁF, Littvay L, Tárnoki ÁD, Tárnoki DL, Bikov A, Molnár MJ, Buzás EI, Falus A. (2018) Tight co-twin similarity of monozygotic twins for hTERT protein level of T cell subsets, for telomere length and mitochondrial DNA copy number, but not for telomerase activity. Cellular and molecular life sciences, 75.13:2447-2456. **IF: 7.014**