Diagnostic exome sequencing in early-onset Parkinson's disease confirms *VPS13C* as a rare cause of autosomal-recessive Parkinson's disease

Barbara Schormair[1,2,*], David Kemlink[3,*], Brit Mollenhauer[4,5], Ondrej Fiala[3,6], Gerrit Machetanz[4], Jan Roth[3], Riccardo Berutti[7], Tim M. Strom[2,7], Bernhard Haslinger[8], Claudia Trenkwalder[4], Daniela Zahorakova[9], Pavel Martasek[9], Evzen Ruzicka[3], Juliane Winkelmann[1,2,8,10,†]

* Both authors contributed equally
† corresponding author
1) Institute of Neurogenomics, Helmholtz Zentrum München, Munich, Germany
2) Institute of Human Genetics, Technische Universität München, Munich, Germany
3) Department of Neurology and Center of Clinical Neuroscience, First Faculty of Medicine and General University Hospital in Prague, Charles University in Prague, Prague, Czech Republic
4) Paracelsus-Elena-Klinik, Kassel, Germany
5) University Medical Center, Institute of Neuropathology and Department of Neurosurgery, Göttingen, Germany
6) Institute of Neuropsychiatric Care (INEP), Křižíkova 264/22, 186 00, Prague, Czech Republic
7) Institute of Human Genetics, Helmholtz Zentrum München, Munich, Germany
8) Neurologische Klinik und Poliklinik, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany
9) Department of Pediatrics and Adolescent Medicine, First Faculty of Medicine and General University Hospital in Prague, Charles University in Prague, Prague, Czech Republic
10) Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

Corresponding author:

Prof. Juliane Winkelmann, Institute of Neurogenomics, Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)
Ingolstädter Landstraße 1
85764 Neuherberg, Germany

E-mail: juliane.winkelmann@tum.de
Phone: 0049-89-3187-1884
Fax: 0049-89-3187-3297

Conflict of interest: The authors report no relevant conflicts of interests related to the manuscript.

Acknowledgments: We thank all patients who participated in this study. We thank all colleagues, technical assistants, and study nurses at the participating centers for their work. Part of the study was supported by grants UNCE-204011/12 (Charles University Prague) and NT-11331-6/2010 (Czech Ministry of Health).

Abstract

Parkinson's disease (PD) is a genetically heterogeneous disorder and new putative disease genes are discovered constantly. Therefore, whole exome sequencing could be an efficient approach to genetic testing in PD. To evaluate its performance in early-onset sporadic PD, we performed diagnostic exome sequencing in 80 individuals with manifestation of PD symptoms at age 40 or earlier and a negative family history of PD. Variants in validated and candidate disease genes and risk factors for PD and atypical Parkinson syndromes were annotated, followed by further analysis for selected variants. We detected pathogenic variants in Mendelian genes in 6.25% of cases and high-impact risk factor variants in *GBA* in 5% of cases, resulting in overall maximum diagnostic yield of 11.25%. One individual was compound heterozygous for variants affecting canonical splice sites in *VPS13C*, confirming the causal role of protein-truncating variants in this gene linked to autosomal-recessive early-onset PD. Despite the low diagnostic yield of exome sequencing in sporadic early-onset Parkinson's disease, the confirmation of the recently discovered *VPS13C* gene highlights its advantage over using predefined gene panels.

Key Words: Exome, Parkinson disease, *VPS13C*, genetic testing

Introduction

Among the neurodegenerative disorders, Parkinson's disease (PD) (MIM 168600) is the second most prevalent, affecting overall 1.6% of the population of Caucasian descent with 20% of all individuals reporting a positive family history.(1) Only 3–5% of all patients with PD are diagnosed before age 40.(2) This rare form of the disease is referred to as early-onset PD (EOPD) and a higher genetic contribution to disease risk is expected in this group. There is considerable genetic heterogeneity in PD with to date 11 confirmed genes with Mendelian inheritance and several putative Mendelian genes and risk factors for which either conflicting evidence exists or no replication in independent datasets has been reported yet.(3, 4) The strongest genetic risk factors for PD to date are variants in *GBA*, the gene implicated in autosomal-recessive Gaucher disease (MIM 606463). Heterozygous carriers of mutations in *GBA* were shown to have a four- to five-fold increase in their risk of developing PD.(5) Finally, new disease-related genes are still discovered for PD, promoted especially by the introduction of new sequencing technologies.

Whole exome sequencing (WES) has become a cornerstone of molecular diagnosis in genetic disorders.(6-8) Pre-defined gene panels and whole exome sequencing are routinely used for a large range of different disease types. Diagnostic yield is reported for different disorders to be in the range of 20 to 40%.(9) For PD patients in general and also EOPD, reports on exome sequencing are sparse and so far these studies have focused on using WES to discover new disease genes.(10-14) Only two studies specifically sequenced individuals with EOPD.(14, 15) Our aim was to assess the performance of WES in the challenging subset of PD patients with sporadic EOPD. The early onset of the disease suggests a genetic basis, but the lack of a family history of PD complicates the analysis in this group. Therefore, we screened a series of 80 sporadic EOPD cases of European ancestry with WES to investigate the diagnostic yield using stringent variant classification criteria.

Materials and Methods

Samples

DNA from 80 participants from movement disorder specialty clinics in Czech Republic (Prague; n = 49, mean age at onset 34.9 years, 63.3% male) and Germany (Munich and Kassel; n = 31, mean age at onset 35.8 years, 67.8% male) was analyzed. Inclusion criterion was a diagnosis of PD with onset of symptoms at the age of ≤ 40 years and a negative family history of PD. Diagnosis of PD was based on the UK Brain Bank diagnostic criteria for PD. Study approval was obtained from the respective ethics review boards and all participants gave written informed consent.

Definition of gene sets

To assess the diagnostic yield, we included validated genes causing PD or atypical Parkinson syndromes with Mendelian inheritance. These include three loci with autosomal dominant (*SNCA, LRRK2, VPS35*) and nine with autosomal recessive or X-linked manner of transmission (*PARK2, PINK1, PARK7, DNAJC6, ATP13A2, FBXO7, PLA2G6, ATP6AP2, SYNJ1, RAB39B*). (3, 4, 16) Furthermore, we included *GBA* as a validated risk factor for PD.(5)

As Parkinson syndromes can also be observed as part of the phenotype in other disorders, we screened a list of candidate genes for such phenotypes based on OMIM (supplementary table 1).

For exploratory analysis, we included not yet validated genes for Mendelian inheritance identified by sequencing in families (*TMEM230, VPS13C, RIC3, PODXL, PLXNA4, LRRK1, ADORA1*).(12, 13, 17-21) In addition, we screened genes with conflicting evidence for Mendelian inheritance (*DNAJC13, EIF4G1, CHCHD2, UCHL1, GIGYF2*, and *HTRA2*) and candidate genes from recent exome sequencing studies(*PTEN*, *VAPB*, *ASNA1, TNR, TNK2, SCAPER, HYDIN, UBE2H, EZR, MMRN2,* and *OGFOD1*). (3, 4, 10, 11, 15, 22-26)

Whole-exome sequencing and variant calling

Exomes were enriched with SureSelect XT Human All Exon 50Mb kit version 5 (Agilent Technologies, USA) and sequenced with 100-bp paired-end reads on a HiSeq2500 (Illumina, USA) to an average depth of coverage of 139.1 (s.d ± 28.0, supplementary table 2). Reads were mapped to the human reference genome (hg19) with BWA (v.0.6.2.). SAMtools (v.0.1.18), PINDEL (v.0.2.4t), ExomeDepth (v.1.0.0), and custom Perl scripts were employed for variant detection and annotation. Only variant calls with a minimum read depth of 20x and a minimum quality score of 30 were included in downstream analyses. To identify candidate pathogenic variants in the gene sets described above, we required a minor allele frequency (MAF) ≤ 0.001 for heterozygous and ≤ 0.01 for homozygous or compound heterozygous variants in the ExAC non-Finnish European ancestry dataset or an in-house exome database, focused on missense and loss-of-function variants (nonsense, splice, stoploss, frameshift, and copy number variants (CNVs) called by ExomeDepth) and based our classification on ACMG recommendations.(27) Triplett repeat expansions were not analyzed. Tools used in variant interpretation were the Munich Exome Variant and Annotation Database (https://ihg4.helmholtz-muenchen.de/cgi-bin/mysql/snv-vcf/login.pl), the M-CAP pathogenicity classifier (http://bejerano.stanford.edu/mcap/), and the Variant Effect Predictor (VEP) tool (http://www.ensembl.org/info/docs/tools/vep/index.html). Moreover, we used public genetic variation databases (ExAC, Exome Aggregation Consortium; gnomAD, genome Aggregation Database; HGMD, Human Gene Mutation Database Professional release 2016.2; ClinVar, www.ncbi.nlm.nih.gov/clinvar/; MDSGene, http://www.mdsgene.org/; PDmutDB, http://www.molgen.vib-ua.be/PDMutDB/), and frequency data of the Wellderly Study as a cohort of individuals who are >80 years old with no chronic diseases.(28, 29)

Sanger sequencing

Variants detected in GBA were validated by Sanger sequencing of exons 8 to 11 using BigDyeTerminator chemistry 3.1 (ABI) on an ABI 3730 sequencer with primers and conditions as described previously.(30) Splice variants in *VPS13C* were validated and assessed for *cis*- or *trans*-configuration by Sanger sequencing of the affected individual and one unaffected child.

Molecular analysis of splice variants

The consequences of splice variants in *VPS13C* were analyzed on the cDNA level by Sanger sequencing and agarose gel electrophoresis. Total RNA was extracted from whole blood of the affected individual and a healthy unrelated control sample (PAXgene Blood miRNA kit, Qiagen). RNA quality and concentration was assessed with Agilent Bioanalyzer 2100. Reverse transcription was performed using SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen). Subsequently, cDNA fragments covering exon 19– 23 or exon 30–34 of *VPS13C* (reference sequence NM_020821.2), respectively, were amplified with specific primers. PCR products were visualized on an agarose gel (4%) and sequenced directly or following extraction (QIAquick Gel Extraction Kit,Qiagen).

Results

Screening of validated Mendelian PD genes and risk factors

Assuming an autosomal-recessive mode of inheritance, we detected bi-allelic pathogenic variants in known Mendelian PD genes (*PINK1* and *PARK2*) in two individuals (table 1a). In both cases, the variants are expected to lead to a complete loss of function. They have been observed in PD patients and annotated as pathogenic (MDSGene and PDmutDB databases). No bi-allelic pathogenic or likely pathogenic variants were detected in *PARK7, DNAJC6, ATP13A2, FBXO7, PLA2G6, ATP6AP2, SYNJ1*, and *RAB39B*. With regard to autosomal-dominant forms of PD, we identified two individuals carrying the most frequent known pathogenic variant p.Gly2019Ser in *LRRK2* in heterozygous state (table 1a).(31) No pathogenic or likely pathogenic variants were found in *VPS35* and *SNCA*. We detected four

heterozygous variants in known autosomal-dominant Mendelian PD genes, which we classified as variants of unknown significance (VUS) or benign after reviewing currently available genetic, functional and *in-silico* data (table 1a and detailed discussion in supplementary table 3). The overall diagnostic yield for validated Mendelian PD genes in our cohort was 5% (4/80 individuals, basic demographic and clinical information is given in table 2).

In *GBA,* we initially detected four different heterozygous rare variants in five individuals passing our filtering criteria. We identified two carriers of p.Asn409Ser (formerly annotated as N370S, not based on HGVS recommendations), one carrier of p.Leu483Pro (formerly annotated as L444P), and one carrier of p.Asp448His (formerly annotated as D409H). In addition, we found one carrier of the singleton protein-truncating variant p.Trp432Ter. We also detected carriers of two variants with frequencies of about 1% in the ExAC and gnomAD non-Finnish European ancestry dataset or in > 8,000 in-house control exomes: 10 individuals were heterozygous for p.Glu365Lys (formerly annotated as E326K) and three individuals were heterozygous for p.Thr408Met (formerly annotated as T369M). Due to the presence of a highly homologous pseudogene (*GBAP1* with 98% sequence homology), sequencing of GBA is challenging. We therefore used Sanger sequencing of exons 8 to 11 with primers specific for GBA amplification in all carriers of GBA variants in our dataset to validate the variants detected by WES. We were able to confirm all variant calls with the exception of p.Leu483Pro, where Sanger sequencing determined the carrier as homozygous for the wildtype allele. Overall, we were able to confirm 17 of 18 variant calls in GBA (94.4%). Concerning the sensitivity of our WES approach, Sanger sequencing revealed no additional variants not present in the exome sequencing data.

Overall, we identified and confirmed three different heterozygous rare variants in *GBA* which we classify as strong risk factors for PD (table 1b). According to ClinVar, both missense variants have been described as pathogenic in neuropathic Gaucher disease and as risk factors for PD. Nonsense variants in *GBA* are rare, but have been identified in neuropathic

Gaucher disease and in PD patients.(32-34) Further support for a role of such loss of function variants in PD risk is provided by evidence of reduced GBA enzymatic activity in PD patients.(35, 36) The common missense variants p.Glu365Lys and p.Thr408Met do not cause Gaucher disease, but functional studies have shown a reduction of GBA enzymatic activity in carriers of these variants.(35, 37) Additional studies have shown an influence on the progression of PD symptoms and larger case-control studies reported an association with PD.(38-41) However, effect sizes were smaller than those observed for Gaucher-related *GBA* variants (odds ratios (ORs) of 1.7 compared to ≥ 3). Therefore, we classify these as risk factors with mild effect (table 1b). The overall diagnostic yield for GBA as a validated risk factor in our cohort was 5% (4/80 individuals) for rare variants with larger effects. When including common low-impact variants, we detect variants modulating disease risk or progression in 21.25% of individuals (17/80 individuals). Age of onset was not significantly different between carriers and non-carriers of variants in *GBA* (35.35 years (s.d. ± 4.20) vs 35.42 (s.d. ± 5.02)). We also found a statistically significant enrichment of rare variants in *GBA* in our EOPD sample compared to > 8,000 in-house control exomes (Fisher exact test, $p = 3.77 \times 10^{-7}$, OR = 5.21 [95% confidence interval (CI) 3.02–8.98]).

We also looked for heterozygous variants in *PARK2* in our sample and detected a total of 8 individuals carrying such variants (10%). No second variant was found in *PARK2* in these individuals even with relaxed filtering criteria, although on average 95.7% (s.d. ± 3.5) of the exonic sequence of *PARK2* (NM_004562.2) had a coverage ≥ 20x. In total, three CNVs and five different missense variants were identified (table 3 and supplementary table 4). Our EOPD samples showed an enrichment for such variants within an OR of 2.52 compared to > 8,000 in-house exomes, but it was not statistically significant (Fisher exact test, p = 0.0578, 95% CI of OR = 1.01–6.29). Two CNVs (EX2DEL and EX2-5DEL) correspond to exon deletions reported as pathogenic in homozygous or compound heterozygous state in PD (HGMD, ClinVar, MDSGene, and PDMutDB databases). The third CNV (EX1-2DEL) was not found in these databases, but deletions of exon 1, exon 1–4, and of promotor and exon 1 of

PARK2 have been reported as pathogenic. The missense variants in *PARK2* were rare variants with MAFs < 0.008 in the ExAC and gnomAD non-Finnish European datasets. Variant p.Arg275Trp is classified as pathogenic according to ClinVar and has been reported in 76 PD patients in homozygous or compound heterozygous state according to the MDSGene database. We classified the remaining missense variants as variants of unknown significance due to the lack of a second variant in *PARK2* in the carriers. For these variants, database annotations also did not provide consistent support for a pathogenic role.

Our screening of genes linked to Parkinson syndromes in OMIM did not identify variants which could be classified as clearly or likely pathogenic.

Confirmation of VPS13C as causal in autosomal-recessive PD

Recently, homozygous or compound heterozygous protein-truncating variants in vacuolar protein sorting 13C (*VPS13C)* were described as a rare cause of autosomal-recessive EOPD.(12) We identified one individual who was a compound heterozygote for variants affecting canonical splice sites in *VPS13C* (table 1a). Segregation analysis confirmed the compound heterozygous condition of the variants (figure 1). Variant c.2029+2T>G is predicted to alter the splice donor site in intron 21, variant c.3215-1G>T the splice acceptor site in intron 31 of the longest transcript of *VPS13C* (NM_020821.2). The variants are predicted to affect splicing in all known protein-coding isoforms of *VPS13C*. Sanger sequencing of cDNA confirmed the suspected splicing defects with multiple mutant splice isoforms only present in the affected individual (figure 1). At the protein level, one splice isoform has an in-frame deletion of 38 amino acids, whereas all others contain premature stop codons leading to truncated proteins. Both variants were not present in >8,000 in-house control exomes sequenced on the same platform and only a single heterozygous carrier of c.2029+2T>G was found in the non-Finnish European ExAC and gnomAD populations. The affected individual presented with overall milder motor symptoms and disease progression than the initially reported cases for *VPS13C*, but still showed a rather rapid deterioration with

response fluctuations and severe dyskinesias, psychiatric symptoms, and impaired cognition (table 2).

Including *VPS13C* in our diagnostic gene set raises the diagnostic yield for Mendelian PD genes to 6.25% and the overall diagnostic yield to 11.25%.

Exploratory screening of putative PD genes

Apart from *VPS13C*, we did not identify pathogenic or likely pathogenic variants in any of the other not yet validated genes for Mendelian inheritance identified by sequencing in families. We also did not identify pathogenic or likely pathogenic variants in the candidate genes put forward by recent exome sequencing studies in PD. None of the variants previously reported in PD patients in these genes were detected in our study sample.

Concerning genes with conflicting evidence for Mendelian inheritance, we detected variants passing our initial filtering criteria in *DNAJC13*, *EIF4G1*, and *GIGYF2* (supplementary table 5). Interestingly, two of the three variants in *DNAJC13* were carried by an individual tested positive for the p.Gly2019Ser variant in *LRRK2*. This supports the observation by Deng et al. who recently suggested variants in *TMEM230* as the actual cause of PD in the original *DNAJC13* family.(13) We therefore classify these variants in *DNAJC13* as benign. In *EIF4G1*, four different missense variants were identified with frequencies < 0.001 in ExAC, gnomAD and in-house control exomes. Variant p.Ile806Val was also found in the Wellderly cohort, supporting a classification of this variant as benign. In *GIGYF2*, two singleton variants and two rare variants with a frequency < 0.001 in the non-Finnish European ExAC, gnomAD, and in-house control exomes were detected. None of the variants we detected in *DNAJC13*, *EIF4G1*, and *GIGYF2* have been reported in PD patients or as disease-related variants in public databases (HGMD, ClinVar), despite screening efforts in larger PD case and case-control studies. In addition, the majority of follow-up studies on these genes do not clearly support a causal role for them in PD. Therefore we classify the detected variants as either benign or likely benign (supplementary table 4).

Discussion

We applied diagnostic WES to a series of 80 sporadic EOPD cases and detected clearly pathogenic variants in known or candidate Mendelian PD genes in 6.25% of patients. Bi-allelic pathogenic variants with autosomal-recessive inheritance were detected in three individuals (3.75%). Heterozygous variants with autosomal-dominant inheritance were found in two individuals (2.5%). This diagnostic yield is lower compared to the application of WES in molecular diagnostics of rare disorders such as childhood-onset Mendelian disorders or ataxia with reported yields of 20–30%.(6-8) However, it is comparable to a recent study which used targeted next-generation sequencing to screen eight PD genes in a population from southern Spain.(42) In their sample of 28 sporadic EOPD patients they detected two individuals with pathogenic mutations in Mendelian PD genes, resulting in a diagnostic yield of 7.1%.

Different factors contribute to the lower diagnostic yield for PD observed in our study. As we only included sporadic cases, our sample likely contains less monogenic cases than a sample of familial PD would. DNA of the parents of the sporadic cases was not available, so we could not screen trios *for de-novo* variation. Importantly, we used very strict criteria for frequency filtering and attributing pathogenicity to a variant, which likely reduced our yield compared to other studies. Currently, the classification of low penetrance variants, which may reach higher frequencies than our frequency cutoff, is challenging, because in-silico prediction tools are still prone to misclassification and high-throughput functional testing of putative causal variants is not readily available. Finally, WES technology is not perfect. A small set of genes and exons is not covered sufficiently, triplet repeat expansions cannot be analyzed, and the detection scope for structural variants and noncoding variants is limited.(43) However, all heterozygous deletions in *PARK2* in our study were in accordance with results of a previous MLPA analysis, indicating that high coverage WES is able to detect CNVs reliably.(44) The location of CNV calls by ExomeDepth overlapped with the MLPA results, and manual inspection of the sequencing reads confirmed the exact size of the

CNVs. This is in line with a previous study showing high sensitivity of up to 88% for depth-of-coverage-based calling of CNVs containing three or more exons.(45)

GBA is an important risk factor and disease modifier for PD; therefore we included it in the diagnostic evaluation. Based on our experience, we recommend focusing on extremely high-quality variants for WES analysis of *GBA* to avoid false-positive variant calls due to the homologous pseudogene. In our study, the only false-positive variant had a considerably lower quality score in WES compared to all other variants detected in *GBA* (SAMtools variant quality score of 78 vs. 225). This filtering did not lower the discovery rate as Sanger sequencing revealed no variants not seen in the WES high-quality variant calls. Overall, 5% of the patients carried rare variants in *GBA* with a strong effect and 16.25% carried low-frequency variants with small effects. Interestingly, we saw a population-specific difference: 28.6% of cases from the Czech Republic carried variants in *GBA*, whereas only 9.7% of cases recruited in German centers did. The difference was mainly driven by the low-risk variant p.Glu365Lys, which was found in 9 Czech patients (18.4%), and only 1 German patient (3.2%), but also all carriers of rare variants in *GBA* were Czech patients For none of the patients in our study, a Jewish background was reported. Our observations are in line with previous studies, which scanned the entire geneand showed population-specific frequencies between 4.2% and 9.8% for *GBA* carriers in PD cases of non-Ashkenazi Jewish ancestry.(5, 46) The targeted next-generation sequencing study from Spain had detected GBA variants in 17.8% of their EOPD cases and other studies focused on EOPD showed frequencies of 11.4% in a Greek sample and 25% in a sample from the United Kingdom.(47, 48) To our knowledge, we are the first study to report a high frequency of *GBA* variants in Czech EOPD patients. To fully evaluate the role of these variants as risk factors for PD in the Czech population, sequencing of *GBA* in matched controls is required to determine allele frequencies in unaffected individuals. For *GBA*, we found a statistically significant enrichment of variants passing our filtering criteria in our EOPD samples compared to in-house control exomes with an OR of 5.2, in line with previous observations.(5)

The role of heterozygous variants in *PARK2* in PD is still a matter of debate. Targeted sequencing studies have reported an increased frequency of carriers of heterozygous variants in *PARK2* in PD patients compared to healthy controls.(31) This is in contrast to results of larger case-control sequencing efforts which did not detect an enrichment of such variants in PD patients.(49) We did not observe a statistically significant enrichment in our sample. However, due to its small sample size it has only moderate power for such association analyses and also may lead to biased frequency estimates. Therefore, definite conclusions on the importance of heterozygous variants in *PARK2* for PD should be drawn based on larger datasets.

Finally, our study highlights the advantage of WES compared to using pre-defined gene panels, which are lagging behind in integrating newly discovered genes. We were able to confirm for the first time the recently identified autosomal-recessive Mendelian PD gene *VPS13C*. Bi-allelic truncating variants in this gene were described as a very rare cause of rapidly progressing EOPD with early cognitive decline.(12) We detected one individual who carried compound heterozygous canonical splice site mutations in *VPS13C*.

In conclusion, we achieved a diagnostic yield of 6.25% when focusing solely on Mendelian PD genes. We could increase the yield to 11.25% when including rare high-impact risk variants in *GBA*. Adding also low-frequency low-impact variants in *GBA* would raise the yield to 27.5%. However, the high frequency of *GBA* variants was specific to the Czech EOPD sample and is therefore not generalizable to WES in other populations. In terms of genetic counselling, *GBA* should be considered a risk factor for PD rather than a causal Mendelian gene. Penetrance estimates of 15%–29.7% at the age of 80 years for heterozygous variants in *GBA* indicate that the majority of carriers of *GBA* variants never develop PD.(50-52) At present, it is not understood which additional genetic and environmental factors are necessary to develop PD in carriers of *GBA* variants. Counselling of affected individuals and families should be provided by specialized clinical geneticists who are qualified to evaluate and interpret the available clinical and genetic data.

Despite the comparably low diagnostic yield for Mendelian PD genes, exome sequencing should be the preferred approach for genetic testing of a patient presenting with early onset PD. WES enables fast integration of new genetic findings into molecular diagnostics and the data can be reanalyzed at any time, which has been shown to improve diagnostic rate dramatically.(53) Finally, the inclusion of WES data in large-scale research projects is straightforward and has the potential to contribute to not only a better understanding of PD genetics and etiology, but also to accelerate the development of new and personalized treatment options.

References

1. Pringsheim T, Jette N, Frolkis A et al. The prevalence of Parkinson's disease: A systematic review and meta-analysis. *Mov Disord* 2014: **29**: 1583–90.
2. Schrag A, Schott JM. Epidemiological, clinical, and genetic characteristics of early-onset parkinsonism. *Lancet Neurol* 2006: **5**: 355–63.
3. Bras J, Guerreiro R, Hardy J. SnapShot: Genetics of Parkinson's disease. *Cell* 2015: **160**: 570–570.e1.
4. Verstraeten A, Theuns Ja. Progress in unraveling the genetic etiology of Parkinson disease in a genomic era. *Trends Genet* 2015: **31**: 140–149.
5. Sidransky E, Nalls MA, Aasly JO et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med* 2009: **361**: 1651–1661.
6. Retterer K, Juusola J, Cho MT et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med* 2016: **18**: 696–704.
7. Yang Y, Muzny DM, Reid JG et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 2013: **369**: 1502–1511.
8. Pyle A, Smertenko T, Bargiela D et al. Exome sequencing in undiagnosed inherited and sporadic ataxias. *Brain* 2015: **138**: 276–283.
9. Need AC, Goldstein DB. Neuropsychiatric genomics in precision medicine: diagnostics, gene discovery, and translation. *Dialogues Clin Neurosci* 2016: 18: 237–252.
10. Farlow JL, Robak LA, Hetrick K et al. Whole-Exome Sequencing in Familial Parkinson Disease. *JAMA Neurol* 2016: **73**: 68–75.
11. Quadri M, Yang X, Cossu G et al. An exome study of Parkinson's disease in Sardinia, a Mediterranean genetic isolate. *Neurogenetics* 2015: **16**: 55–64.
12. Lesage S, Drouet VrV, Majounie E et al. Loss of VPS13C Function in Autosomal-Recessive Parkinsonism Causes Mitochondrial Dysfunction and Increases PINK1/Parkin-Dependent Mitophagy. *Am J Hum Genet* 2016: **98**: 500–513.
13. Deng H-XH, Shi Y, Yang Y et al. Identification of TMEM230 mutations in familial Parkinson's disease. *Nat Genet* 2016: **48**: 733–739.
14. Siitonen A, Nalls MA, Hernandez D et al. Genetics of early-onset Parkinson's disease in Finland: exome sequencing and genome-wide association study. *Neurobiol Aging* 2017: **53**: 195.e7-195.e10.
15. Kun-Rodrigues C, Ganos C, Guerreiro R et al. A systematic screening to identify de novo mutations causing sporadic early-onset Parkinson's disease. *Hum Mol Genet* 2015: **24**: 6711–6720.
16. Wilson GR, Sim JCH, McLean C et al. Mutations in RAB39B cause X-linked intellectual disability and early-onset parkinson disease with alpha-synuclein pathology. *Am J Hum Genet* 2014: **95**: 729–735.
17. Sudhaman S, Muthane UB, Behari M et al. Evidence of mutations in RIC3 acetylcholine receptor chaperone as a novel cause of autosomal-dominant Parkinson's disease with non-motor phenotypes. *J Med Genet* 2016: **53**: 559–566.
18. Sudhaman S, Prasad K, Behari M et al. Discovery of a frameshift mutation in podocalyxin-like (PODXL) gene, coding for a neural adhesion molecule, as causal for autosomal-recessive juvenile Parkinsonism. *J Med Genet* 2016: **53**: 450–456.
19. Schulte EC, Ellwanger DC, Dihanich S et al. Rare variants in LRRK1 and Parkinson's disease. *Neurogenetics* 2013: **15**: 49–57.
20. Schulte EC, Stahl I, Czamara D et al. Rare variants in PLXNA4 and Parkinson's disease. *PLoS ONE* 2013: **8**: e79145.
21. Jaberi E, Rohani M, Shahidi GA et al. Mutation in ADORA1 identified as likely cause of early-onset parkinsonism and cognitive dysfunction. *Mov Disord* 2016: **31**: 1004-1011.
22. Vilario-Guell C, Rajput AHA, Milnerwood AJ et al. DNAJC13 mutations in Parkinson disease. *Hum Mol Genet* 2014: **23**: 1794–1801.
23. Chartier-Harlin MC, Dachsel JC, Vilario-Guell C et al. Translation initiator EIF4G1 mutations in familial parkinson disease. *Am J Hum Genet* 2011: **89**: 398–406.

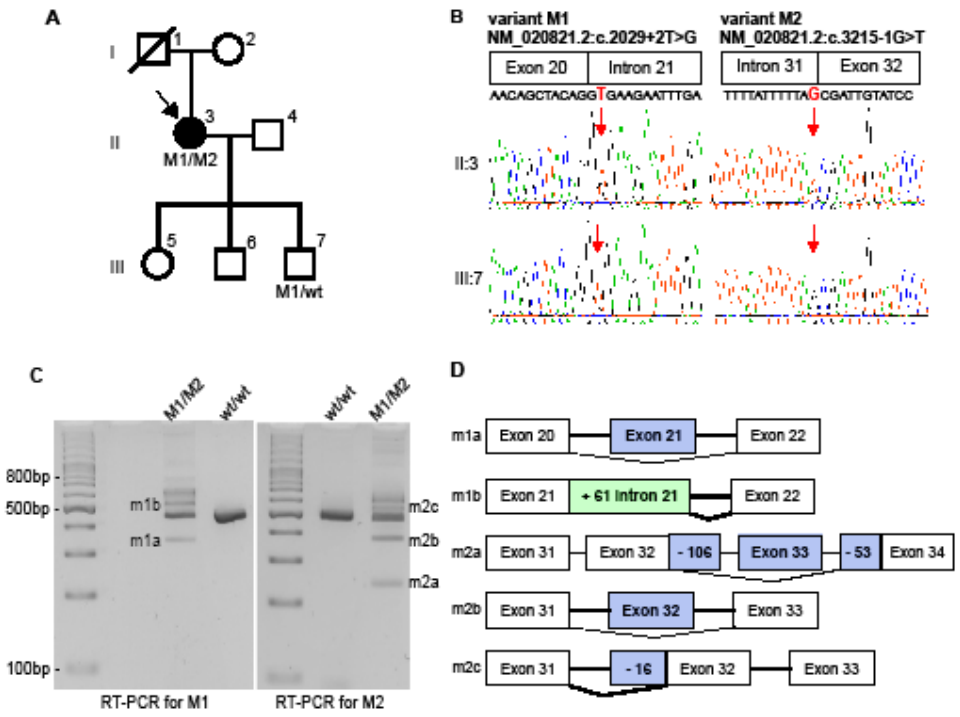
24. Funayama M, Ohe K, Amo T et al. CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: A genome-wide linkage and sequencing study. *Lancet Neurol* 2015: **14**: 274–282.
25. Lautier C, Goldwurm S, Dürr A et al. Mutations in the GIGYF2 (TNRC15) Gene at the PARK11 Locus in Familial Parkinson Disease. *Am J Hum Genet* 2008: **82**: 822–833.
26. Strauss KM, Martins LM, Plun-Favreau H et al. Loss of function mutations in the gene encoding Omi/HtrA2 in Parkinson's disease. *Hum Mol Genet* 2005: **14**: 2099–2111.
27. Richards S, Aziz N, Bale S et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015: **17**: 405–424.
28. Lek M, Karczewski KJ, Minikel EV et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016: **536**: 285–291.
29. Erikson GA, Bodian DL, Rueda M et al. Whole-Genome Sequencing of a Healthy Aging Cohort. *Cell* 2016: **165**: 1002–1011.
30. Tammachote R, Tongkobpetch S, Srichomthong C et al. A common and two novel GBA mutations in Thai patients with Gaucher disease. *J Hum Genet* 2013: **58**: 594–599.
31. Hernandez DG, Reed X, Singleton AB. Genetics in Parkinson disease: Mendelian versus non-Mendelian inheritance. *J Neurochem* 2016: **139**: 59-74.
32. Beutler E, Gelbart T, Kuhl W et al. Mutations in Jewish patients with Gaucher disease. *Blood* 1992: **79**: 1662–1666.
33. Liu G, Boot B, Locascio JJ et al. Specifically neuropathic Gaucher's mutations accelerate cognitive decline in Parkinson's. *Ann Neurol* 2016: **80**: 674-685.
34. Lesage S, Anheim M, Condroyer C et al. Large-scale screening of the Gaucher's disease-related glucocerebrosidase gene in Europeans with Parkinson's disease. *Hum Mol Genet* 2011: **20**: 202–210.
35. Alcalay RN, Levy OA, Waters CC et al. Glucocerebrosidase activity in Parkinson's disease with and without GBA mutations. *Brain* 2015: **138**: 2648–2658.
36. Ortega RA, Torres PA, Swan M et al. Glucocerebrosidase enzyme activity in GBA mutation Parkinson's disease. *J Clin Neurosci* 2016: **28**: 185–186.
37. Horowitz M, Pasmanik-Chor M, Ron I et al. The enigma of the E326K mutation in acid beta-glucocerebrosidase. *Mol Genet Metab* 2011: **104**: 35–38.
38. Davis MY, Johnson CO, Leverenz JB et al. Association of GBA Mutations and the E326K Polymorphism With Motor and Cognitive Progression in Parkinson Disease. *JAMA Neurol* 2016: **73**: 1217–1224.
39. Mata IF, Leverenz JB, Weintraub D et al. GBA Variants are associated with a distinct pattern of cognitive deficits in Parkinson's disease. *Mov Disord* 2016: **31**: 95–102.
40. Pankratz N, Beecham GW, DeStefano AL et al. Meta-analysis of Parkinson's disease: identification of a novel locus, RIT2. *Ann Neurol* 2012: **71**: 370–384.
41. Mallett V, Ross JP, Alcalay RN et al. GBA p.T369M substitution in Parkinson disease: Polymorphism or association? A meta-analysis. *Neurol Genet* 2016: **2**: e104.
42. Bandres-Ciga S, Mencacci NE, Duran R et al. Analysis of the genetic variability in Parkinson's disease from Southern Spain. *Neurobiol Aging* 2016: **37**: 210.e1–5.
43. Bras JM, Singleton AB. Exome sequencing in Parkinson's disease. *Clin Genet* 2011: **80**: 104–109.
44. Fiala O, Zahorakova D, Pospisilova L et al. Parkin (PARK 2) mutations are rare in Czech patients with early-onset Parkinson's disease. *PLoS One* 2014: **9**: e107585.
45. de Ligt J, Boone PM, Pfundt R et al. Detection of clinically relevant copy number variants with whole-exome sequencing. *Hum Mutat* 2013: **34**: 1439–1448.
46. Zhao F, Bi L, Wang W et al. Mutations of glucocerebrosidase gene and susceptibility to Parkinson's disease: An updated meta-analysis in a European population. *Neuroscience* 2016: **320**: 239–246.
47. Kalinderi K, Bostantjopoulou S, Paisan-Ruiz C et al. Complete screening for glucocerebrosidase mutations in Parkinson disease patients from Greece. *Neurosci Lett* 2009: **452**: 87–89.

48. Duran R, Mencacci NE, Angeli AV et al. The glucocerobrosidase E326K variant predisposes to Parkinson's disease, but does not cause Gaucher's disease. *Mov Disord* 2013: **28**: 232–236.
49. Kay DM, Moran D, Moses L et al. Heterozygous parkin point mutations are as common in control subjects as in Parkinson's patients. *Ann Neurol* 2007: **61**: 47–54.
50. Alcalay RN, Dinur T, Quinn T, et al. Comparison of Parkinson risk in Ashkenazi Jewish Gaucher patients and GBA heterozygotes. JAMA Neurol 2014: **71**: 752–757.
51. Anheim M, Elbaz A, Lesage S, et al. Penetrance of Parkinson disease in glucocerebrosidase gene mutation carriers. *Neurology* 2012: **78**: 417–420.
52. McNeill A, Duran R, Hughes DA, Mehta A, Schapira AH. A clinical and family history study of Parkinson's disease in heterozygous glucocerebrosidase mutation carriers. *J Neurol Neurosurg Psychiatry* 2012: **83**: 853–854.
53. Need AC, Shashi V, Schoch K et al. The importance of dynamic re-analysis in diagnostic whole exome sequencing. *J Med Genet* 2017: **54**: 155–156.

Figure Legends:

Figure 1: Segregation and splicing analysis of variants in *VPS13C*

(A) Family pedigree. Black symbol represents affected individual, open symbols unaffected individuals. The arrow indicates the individual who underwent exome sequencing. M1 indicates presence of splice-site variant c.2029+2T>G, M2 presence of splice-variant c.3215-1G>T, and wt presence of reference allele.



(B) Sanger sequencing chromatograms of both splice-site variants in the affected individual (II:3) and an unaffected child (III:7).. Sequences represent the reverse strand.

(C) For both variants, consequences on splicing were analyzed using reverse-transcription PCR (RT-PCR). For each variant, several mutant transcripts are detected, denoted as m1a, m1b, m2a, m2b and m2c. For variant M1, the PCR had an expected wildtype product size of 466bp. For variant M2, the PCR had an expected wildtype product size of 509bp. M1/M2 refers to the affected individual, wt/wt to an unrelated healthy control. The higher molecular

size bands were also sequenced and were found to contain both wt and m1b or m2c sequence, respectively.

(D) Structure of mutant transcripts and consequence on protein level. Retained intronic sequence is indicated with a "+", the number of affected nucleotides, and green shading. Deleted exonic sequence is indicated with a "-", the number of affected nucleotides, and blue shading. Exons and introns are not drawn to scale. Protein level consequences refer to reference NP_065872.1 as follows: m1a = p.Leu678GlufsTer26, m1b = p.Lys639AspfsTer14, m2a = p.Ala1072GlufsTer14, m2b = p.Ala1072_Gln1110del, m2c = p.Ser1076ArgfsTer4.

Table Legends:

Table 1: Diagnostic classification of variants detected in Mendelian PD genes and in GBA.

ExAC = Exome Aggregation Consortium, gnomAD = genome Aggregation Database. Frequencies in ExAC, gnomAD and in-house exomes are reported as allele count / total allele number. [a]numbering according to NCBI accessions NM_032409.2 and NP_115785.1 for PINK1, NM_004562.2 and NP_004553.2 for PARK2, NM_198578.3 and NP_940980.3 for LRRK2, NM_020821.2 and NP_065872.1for VPS13C, NM_018206.5 and NP_060676.2 for VPS35, NM_000157.3 and NP_000148.2 for GBA. [b]consisting of 8,493 exomes with no neurological disease phenotype. [c]classification of variant in ClinVar referring to PD or parkinsonism. hom = homozygous, het = heterozygous. VUS = variant of unknown significance. not found = variant not present in database. na = exact validated CNV call data is not available in ExAC, gnomAD, and in in-house exomes.

Table 2: Demographic and clinical data of individuals with pathogenic variants in Mendelian PD genes

y = years; GIT = gastrointestinal tract; DBS = deep brain stimulation; STN = subthalamic nucleus; GPi = globus pallidus internus.

Table 3: Heterozygous variants detected in *PARK2*

For header and abbreviations used refer to table 1. HGMD, MDSGene, PDmutDB = classification of variant in the respective databases. DM = variant reported by HGMD as disease-causing. DM? = variant reported by HGMD as likely disease-causing. [d]pathogenic refers to classification of variant if seen in compound heterozygous or homozygous *PARK2* variant carriers.

Table 1

Gene	Genomic position (hg19)	Variation nucleotide[a]	Variation amino acid[a] / deleted exons	Mutation type	Number of cases	zygosity	dbSNP142	Frequency ExAC European	Frequency gnomAD European	Frequency in-house exomes[b]	ClinVar classification[c]	study-based classification
	a) Mendelian PD genes											
PINK1	chr1:20975710	c.1474C>T	p.Arg492Ter	nonsense	1	hom	rs34208370	not found	8/110868	3/16986	not found	pathogenic
PARK2	chr6:162622163-162683796	c.(171+1_172-1)_(534+1_535-1)del	EX3-4DEL	deletion	1	hom	not found	na	na	1/16986	not found	pathogenic
LRRK2	chr12:40734202	c.6055G>A	p.Gly2019Ser	missense	2	het	rs34637584	42/66646	31/126412	5/16986	pathogenic	pathogenic
VPS13C	chr15:62274656	c.2029+2T>G	p.Lys639AspfsTer14 p.Leu678GlufsTer26	splice site	1	het	not found	not found	not found	not found	not found	pathogenic
	chr15:62256151	c.3215-1G>T	p.Ser1076ArgfsTer4 p.Ala1072GlyfsTer91 p.Ala1072GlufsTer14	splice site		het	rs374621625	1/66406	1/109162	1/16986	not found	pathogenic
LRRK2	chr12:40643706	c.917C>T	p.Ala306Val	missense	1	het	not found	not found	not found	not found	not found	VUS
LRRK2	chr12:40702910	c.4192C>T	p.Arg1398Cys	missense	1	het	rs373268136	2/66584	2/111322	1/16986	not found	VUS
LRRK2	chr12:40757242	c.7067C>T	p.Thr2356Ile	missense	1	het	rs113511708	17/66438	39/126416	10/16986	uncertain significance	benign
VPS35	chr16:46694455	c.2320C>A	p.Leu774Met	missense	1	het	rs192419029	5/66728	13/126634	9/16986	not found	VUS
	b) GBA											
GBA	chr1:155205634	c.1226A>G	p.Asn409Ser	missense	2	het	rs76763715	242/66184 (1 hom)	255/126662 (1 hom)	22/16986	pathogenic / risk factor	risk factor
GBA	chr1:155205518	c.1342G>C	p.Asp448His	missense	1	het	rs1064651	8/56216	20/111186	not found	pathogenic	risk factor
GBA	chr1:155205564	c.1296G>A	p.Trp432Ter	nonsense	1	het	not found	not found	1/111698	not found	not found	risk factor
GBA	chr1:155206037	c.1093C>T	p.Glu365Lys	missense	3	het	rs2230288	798/66700 (7 hom)	1564/126692 (10 hom)	161/16986	risk factor	risk factor (mild)
GBA	chr1:155206167	c.1223G>A	p.Thr408Met	missense	10	het	rs75548401	627/64196 (5 hom)	1190/126534 (5 hom)	158/16986	VUS	risk factor (mild)

Table 2

PD case	Mendelian PD gene pathogenic variant	Sex	Ancestry	Age at onset of PD (y)	Predominant symptoms	Motor symptoms	Psychiatric symptoms / cognition / memory	Further symptoms	Dyskinesias	DBS treatment
1	*PINK1* p.Arg492Ter	female	Caucasian	20	falls, dysarthria,	tremor	depression, anxiety, mild attention deficit	dysphagia	moderate	10y STN, good effect on rigitiy and akinesiia
2	*PARK2* EX3-4DEL	male	Caucasian	39	dysarthrophonia	rigidity	depressive symptoms	sialorrhea	severe peak-dose dyskinesias, biphasic dyskinesias	none
3	*LRRK2* p.Gly2019Ser	female	Caucasian	40	right-sided hypokinesia, rigidity	akinesia, rigidity	minor depression	orthostatic hypotension, synkope, artrial fibrillation, hypakusis	none	none
4	*LRRK2* p.Gly2019Ser	male	Caucasian	31	fluctuations, freezing of gait	rigidity, micrographia	minor depression	sialorrhea	severe	1y GPi (infection), 5y STN, after 3y major fluctuations
5	*VPS13C* p.Lys639AspfsTer14 p.Leu678GlufsTer26 p.Ala1072GlufsTer14 p.Ala1072_Gln1110del p.Ser1076ArgfsTer	female	Caucasian	39	dysarthria	rigidity, micrographia, tremor	minor delusions, hallucinations, mild dementia	mild dystonia, moderate gait problems; severe GIT problems under dopaminergic treatment	severe under dopaminergic treatment	4y STN, initial improvement, after 2.5y severe dysarthria and mild aphasia

Table 3

Gene	Variation nucleotide[a]	Variation amino acid[a] /deleted exons	dbSNP142	Frequency ExAC European	Frequency gnomAD European	Frequency in-house exomes[b]	HGMD	ClinVar classification[c]	MDSGene classification	PDMutDB classification	study-based classification[d]
PARK2	c.245C>A	p.Ala82Glu	rs55774500	471/66736 (2 hom)	648/126574 (4 hom)	61/16986	CM012632 DM ?	not found	not found	pathogenic nature unclear	VUS
PARK2	c.1000C>T	p.Arg334Cys	rs199657839	1/66126	9/126462	9/16986	CM003865 DM ?	not found	not found	pathogenic	VUS
PARK2	c.1204C>T	p.Arg402Cys	rs55830907	170/66740	280/126724 (2 hom)	61/16986	CM056983 DM	not found	not found	pathogenic nature unclear	VUS
PARK2	c.823C>T	p.Arg275Trp	rs34424986	204/65874	410/126516	21/16986	CM991007 DM	pathogenic	definitely pathogenic	pathogenic nature unclear	pathogenic
PARK2	c.247A>G	p.Thr83Ala	rs141825163	10/66736	18/126508	2/16986	not found	not found	not found	pathogenic nature unclear	VUS
PARK2	c.(?-103)_ (171+1_172-1)del	EX1-2DEL	na	na	na	1/16986	not found	not found	not found	not found	pathogenic
PARK2	c.(7+1_8-1)_ (171+1_172-1)del	EX2DEL	na	na	na	1/16986	CG078014 DM	pathogenic	definitely pathogenic	pathogenic	pathogenic
PARK2	c.(7+1_8-1)_ (618+1_619-1)del	EX2-5DEL	na	na	na	not found	CG044415 DM	not found	probably pathogenic	pathogenic	pathogenic