Mitochondrial DNA heteroplasmy distinguishes disease manifestation in *PINK1*/*PRKN*-linked Parkinson's disease

Joanne Trinh,[1] Andrew A. Hicks,[2] Inke R. König,[3] Sylvie Delcambre,[4] Theresa Lüth,[1] Susen Schaake,[1] Kobi Wasner,[4] Jenny Ghelfi,[4] Max Borsche,[1] Carles Vilariño-Güell,[5] Faycel Hentati,[6] Elisabeth L. Germer[1] Peter Bauer,[7] Masashi Takanashi,[8] Vladimir Kostić,[9] Anthony E. Lang,[10] Norbert Brüggemann,[1,11] Peter P. Pramstaller,[2] Irene Pichler,[2] Alex Rajput,[12] Nobutaka Hattori,[8] Matthew J. Farrer,[5] Katja Lohmann,[1] Hansi Weissensteiner,[13] Patrick May,[4] Christine Klein[1] and Anne Grünewald[1,4]

Abstract

Biallelic mutations in *PINK1*/*PRKN* cause recessive Parkinson's disease. Given the established role of PINK1/Parkin in regulating mitochondrial dynamics, we explored mitochondrial DNA (mtDNA) integrity and inflammation as disease modifiers in carriers of mutations in these genes. MtDNA integrity was investigated in a large collection of biallelic (*n*=84) and monoallelic (*n*=170) carriers of *PINK1/PRKN* mutations, idiopathic Parkinson's disease patients (*n*=67) and controls (*n*=90). In addition, we studied global gene expression and serum cytokine levels in a subset. Affected and unaffected *PINK1*/*PRKN* monoallelic mutation carriers can be distinguished by heteroplasmic mtDNA variant load (AUC=0.83, CI:0.74-0.93). Biallelic *PINK1*/*PRKN* mutation carriers harbor more heteroplasmic mtDNA variants in blood (*p*=0.0006, *Z*=3.63) compared to monoallelic mutation carriers. This enrichment was confirmed in iPSC-derived (controls, *n*=3; biallelic PRKN mutation carriers, *n*=4) and *postmortem* (control, *n*=1; biallelic PRKN mutation carrier, *n*=1) midbrain neurons. Lastly, the heteroplasmic mtDNA variant load correlated with IL6 levels in *PINK1*/*PRKN* mutation carriers (*r*=0.57, *p*=0.0074). *PINK1*/*PRKN* mutations predispose individuals to mtDNA variant accumulation in a dose- and disease-dependent manner.

Author affiliations:

1 Institute of Neurogenetics, University of Lübeck, Lübeck, Germany, 23562

2 Institute for Biomedicine, EURAC, Bolzano, Italy, 39100

3 Institute of Medical Biometry and Statistics, University of Lübeck, Lübeck, Germany, 23562

4 Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur Alzette, Luxembourg, L-4362

5 Department of Medical Genetics, University of British Columbia, Canada, V6T 1Z3

6 Service de Neurologie, Institut National de Neurologie, La Rabta, Tunis, Tunisia, 1007

7 Centogene GmbH, Rostock, Germany, 18055

8 Department of Neurology, Juntendo University, Tokyo, Japan, 113-8421

9 Institute of Neurology, University of Belgrade, Belgrade, Serbia, 11000

10 Toronto Western Hospital, University of Toronto, Toronto, Canada, M5T 2S8

11 Department of Neurology, University of Lübeck, Lübeck, Germany, 23562

12 Division of Neuropathology, University of Saskatchewan and Saskatoon Health Region, Saskatoon, Canada, S7N 5A2

13 Institute of Genetic Epidemiology, Medical University of Innsbruck, Innsbruck, Austria, 6020

Correspondence to: Anne Grünewald, PhD

Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Campus Belval, 6, avenue du Swing, L-4367 Belvaux, Luxembourg

E-mail: anne.gruenewald@uni.lu

Running title: MtDNA heteroplasmy in PRKN/PINK1 parkinsonism

Keywords: mtDNA heteroplasmy; penetrance; modifiers; PINK1; PRKN

Abbreviations: CD38 = ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase; DAMPs = damage-associated molecular patterns; DEGs = differentially expressed genes; GBA = glucocerebrosidase; GSDMA = Gasdermin A; HF = heteroplasmy frequency; IL6 = interleukin 6; L2F = Log2fold; LCM = laser capture microdissection; MDSGene = Movement Disorder

Society Genetic mutation database; MLPA = multiplex ligation probe-dependent amplification; mtDNA = mitochondrial DNA; ND4:ND1 = NADH:ubiquinone oxidoreductase core subunit 4 and NADH-dehydrogenase 1; NUMTs = Nuclear Mitochondrial DNA segments; OPTN = optineurin; PINK1/PRKN/all = biallelic and monoallelic PINK1/ PRKN mutation carriers regardless of affection status; PINK1+/PRKN+/PD- = asymptomatic carriers of PINK1 and PRKN monoallelic mutation carriers; PINK1+/PRKN+/PD+ = symptomatic carriers of PINK1 and PRKN monoallelic mutation carriers; PINK1++/PRKN++/PD+ = Patients with biallelic PINK1 or PRKN mutations; POLG = mitochondrial polymerase gamma; PRKN/all = PRKN biallelic and monoallelic mutation carriers; q-score = quality scores; RNASE1 = Ribonuclease A Family Member 1; ROC = receiving operating characteristic; smNPCs = small molecule neural precursor cells; VUS = variants of uncertain significance

Introduction

Parkinson's disease is a neurodegenerative disorder affecting over seven million patients world-wide,[1] of which ~10% can be genetically explained.[1,2] Multiple genes have been found to be implicated in familial and sporadic Parkinson's disease, including pathogenic mutations and strong risk factors. Biallelic *PINK1* or *PRKN* mutations are a well-known cause of early-onset Parkinson's disease.[3,4] Interestingly, monoallelic *PINK1/PRKN* mutations have been considered a Parkinson's disease risk factor.[5] Based on a conservative estimate of a frequency of 2% of monoallelic *PINK1* and *PRKN* mutations among all Parkinson's disease patients, one can expect ~140,000 carriers worldwide, this warrants a deeper understanding of the possible functional consequence of *PINK1* or *PRKN* monoallelic mutations. Affected monoallelic *PINK1* mutation carriers have a ~10 year later age at onset (AAO of 43.2 years) than carriers of biallelic mutations (AAO 32.6 years), but were on average still younger than idiopathic Parkinson's disease patients (AAO 70 years) used for comparison in the meta-analysis.[6] Furthermore, neuroimaging studies have demonstrated Parkinson's disease-like changes in humans with monoallelic *PINK1/PRKN* mutations.[7,8] However, the possible role of monoallelic pathogenic *PINK1* and *PRKN* mutations in Parkinson's disease remains a matter of vivid debate: a recent study found 0.089% of monoallelic *PINK1* mutation carriers in cases compared to 0.071% in controls in a large population screen (*n*=376,558).[9] Another study did not find any associations

between *PRKN* variants and risk of Parkinson's disease. Pathogenic and likely-pathogenic heterozygous single nucleotide variants and copy-number variations were less frequent in Parkinson's disease patients (1.52%) compared to controls (1.8%).[10] The frequency of *PRKN* variants in 500,000 individuals was evaluated and there were no differences in Parkinson's disease risk for carriers of heterozygous variants.[11] These studies suggest that *PINK1/PRKN* monoallelic mutations may not contribute to risk, albeit with limited power and no possibility to assess (subtle) signs of parkinsonism in the seemingly healthy controls. Still, when deep clinical phenotyping was applied, patients with Parkinson's disease harboring *PINK1* or *PRKN* monoallelic pathogenic variants were at a significantly higher rate than healthy controls.[5] There are further examples of recessive neurodegenerative diseases where carriers of pathogenic variants have low disease risk such as in the optineurin (*OPTN*) gene in ALS/FTD.[12] Even within Parkinson's disease, while biallelic mutations in glucocerebrosidase (*GBA)* cause Gaucher's disease, heterozygous variants in GBA lead to a higher Parkinson's disease risk in later life.[13]

Biologically, PINK1 and Parkin are important in regulating mitochondrial quality control. Activation of the PINK1/Parkin mitophagy pathway has been shown to result in the selective removal of dysfunctional mitochondria and their damaged DNA (mtDNA) molecules.[14-16] In a recent study, Parkin-deficient mice were crossed with "mutator mice", which harbor a proof-reading defective mitochondrial polymerase gamma. In the resulting animals, mtDNA disintegration and impaired mitophagy led to the release of mtDNA molecules into the extracellular space, where they act as damage-associated molecular patterns (DAMPs) triggering inflammatory pathways.[17] However, whether accumulation of mtDNA damage has an influence on the disease manifestation of *PINK1* and *PRKN*-associated Parkinson's disease still needs to be elucidated.

Herein, we present an in-depth characterization of the mitochondrial genome in blood cells as a potential disease marker for genetic Parkinson's disease. We investigated the largest collection of *PINK1*/*PRKN* mutation carriers studied thus far, including blood from two homogeneous founder populations (South Tyrolean and Tunisian Arab Berber), neuronal models and postmortem brain tissue to assess mitochondrial dynamics as a marker of disease manifestation.

Materials and methods

Study individuals and genetic analysis

Deep mitochondrial sequencing and real-time PCR assays using TaqMan technology were performed on a total of 411 individuals (**Table 1, 2 and S1**). Patients were recruited from Germany, Italy, Tunisia and Serbia fulfilling the UK Brain Bank Criteria for Parkinson's disease. Demographic and basic clinical data from *PINK1* and *PRKN* mutation carriers (biallelic and monoallelic, *n*=254) are shown in **Tables 1 and 2**. The AAO of disease for individuals with *PINK1* mutations, *PRKN* mutations, and idiopathic Parkinson's disease is reported in **Tables 1, 2 and S1**. MtDNA was extracted from whole blood using standard protocols for extraction of genomic DNA. *PINK1* and *PRKN* genetic screening was performed with Sanger sequencing and MLPA (multiplex ligation probe-dependent amplification) as previously described.[18] All individuals with Parkinson's disease were negatively screened for other pathogenic variants in known Parkinson's disease genes (*SNCA, LRRK2, VPS35,* and *DJ1*) with gene panel sequencing using either the Illumina NextSeq500 or HiSeq4000 sequencers and MLPA analysis. Genes were selected based on typical Parkinson's disease forms in MDSGene. *PINK1* and *PRKN* mutations refer to 'possibly, probably or definitely' pathogenic variants and variants of uncertain significance (VUS, e.g. *PINK1* p.G411S, p.N451T) were not included based on the Movement Disorder Society Genetic mutation database (MDSGene).[19] *PINK1* and *PRKN* variants were labeled as "pathogenic mutations" based on the criteria of 'possibly, probably or definitely' pathogenic with clear guidelines from the MDSGene consortium.[19,20] All individuals provided informed consent prior to their participation. Specific approvals were obtained from the local ethics committees at the Research Ethics Boards of the Universities of Lübeck and Luxembourg and the respective local ethics committees from the cohort centers. Clinical assessment of all patients was performed by movement disorders specialists. Population-matched healthy controls were used in the study. The AAO was self-reported by patients and based on the occurrence of the first motor symptoms. Familial relationships were defined based on self-report. An overview of individuals in the study is depicted in **Figure 1**.

Deep mitochondrial sequencing

Illumina NextSeq short-read sequencing

Deep mitochondrial sequencing was performed with bait enrichment with IDT lockdown probes specific for the entire mitochondrial genome to generate libraries of 6-plex or long-range PCR. The two primer sets used for the long-range PCR are described in **Table S2**. Subsequent next-generation sequencing was performed on Illumina, using the NextSeq500 (Illumina, Inc.) to produce 2 × 150bp reads. Raw sequencing reads were converted to standard FASTQ format using bcl2fastq software 2.17.1.14 (Illumina, Inc.). Raw FASTQ files were analyzed with FastQC (v011.8) and aligned with BWA MEM[21] (v07.17) to the mitochondrial reference genome (rCRS).[22] The resulting SAM files were processed with SAMtools (v1.3.1) to sorted BAM files.[23] Subsequently, the BAM files were processed with GATK[24] (v3.8) and duplicates removed with MarkDuplicates. Additional quality control was performed with QualiMap v2.2.2d, [25] and subsequently MultiQC[26] (v1.6) was applied on the resulting reports, including the FastQC reports. The mtDNA-Server[27] successor Mutserve v2.0.0[28] was used for variant calling on the mitochondrial genome to detect homoplasmic and heteroplasmic variants in the sequence data. After additional quality control with Haplocheck[28] (v1.2.1), we set the minimum heteroplasmy frequency (HF) level to 0.05[27] and kept all other parameters with the default parameters (per base quality 20, mapping quality 30, alignment score 30), as some samples showed contamination either from different samples or from Nuclear Mitochondrial DNA segments (NUMTs) below that threshold. Mitochondrial genomes with a mean depth of 5,538X (SD:4,919X) coverage were obtained. A cut-off of >1,000X coverage was applied and samples that were contaminated or had an unusually high number of heteroplasmic variants that may be related to NUMTs (>40) were removed (n=86). We defined variants as homoplasmic (>95%) or heteroplasmic (<95% and >5%) based on allele frequency.

Oxford Nanopore long-read single-molecule sequencing

To obtain more even coverage across the mtDNA genome and eliminate NUMTs, a long-read sequencing approach was additionally used to detect mtDNA heteroplasmy in blood-derived (n=16) and neuronal-derived DNA (n=10) of *PRKN* mutation carriers and controls. The same two primer sets from Illumina sequencing were used for the long-range PCR (**Table S2** and

Figure S1). Samples were barcoded and multiplexed, and subsequent sequencing was performed on an R9 flow cell. Mitochondrial genomes with a mean depth of 3433.2X (SD:±2299.9X) coverage were obtained. Base-calling was performed with Guppy v5.0.11 with the super-accurate model (https://community.nanoporetech.com/). The base-called reads were filtered with Filtlong (v0.2.0) to only include the best 50% of the reads, based on Phred quality scores (q-score), with a minimum read length of 9kb (https://github.com/rrwick/Filtlong). Then, the reads were trimmed with NanoFilt.[29] Subsequently, the Nanopore reads were aligned against the mitochondrial genome reference sequence using Minimap2 (v2.17).[30] The alignments were sorted and indexed with SAMtools[23] (v1.3.1). Mutserve2[27] (v2.0.0) and Haplocheck (v1.2.1)[28] was used to call variants and the same default parameters were applied in the same fashion as with our short-read sequencing. However, different thresholds for HF levels were set at 0.05, 0.02 and down to 0.01. The per base quality scores allowed for more accurate identification of heteroplasmic variants. All variants from third and second generation sequencing were annotated with a predefined file provided by Mutserve2, with allele frequencies from the 1000 Genomes Project[31], HelixMtDB[32], NUMTs defining variants[33] as well as pathogenicity scores.[34,35]

MtDNA deletion and 7S DNA analysis

Real-time PCR assays were performed using TaqMan technology. MtDNA deletion rates were derived from the ratio of the mtDNA-encoded genes *NADH:ubiquinone oxidoreductase core subunit 4* and *NADH-dehydrogenase 1* (*ND4:ND1*). *ND1* is a gene located in the minor arc of the mitochondrial genome, in an area spared from deletions. *ND4* is located in the major arc, in a region commonly affected by somatic deletions. The deletion assay has been established previously and has been used quite frequently.[36-39] For instance, in a study where a specific accumulation of mtDNA major arc deletions in neuromelanin-containing dopaminergic neurons isolated from postmortem nigral tissue of Parkinson's disease patients was detected.[36] In addition, they characterized the size of the deletions by long-range PCR, providing evidence that the real-time PCR assay is in fact specific. MtDNA transcription initiation was assessed by calculating the 7S DNA:*ND1* ratio, as the relative abundance of 7S DNA in the D-Loop region allows for an estimation of the mitochondrial transcription initiation rate.[40] Forward and reverse primers as well as probes coupled with a non-fluorescent quencher described in **Table S2** were used to quantify the concentrations of *ND1*, *ND4*, and D-loop, where 7S DNA is integrated

during the initial phase of mtDNA transcription. The reactions were performed in 384-well plates in a final reaction volume of 5 µL using a LightCycler480 (Roche) as previously described.[41] To minimize inter-experimental variability, reagents for each real-time PCR analysis were pipetted using an automated liquid handler (BioMek FX). A serial dilution of the plasmid p7D1 encoding the target sequence of *ND1, ND4* and the D-Loop region was used to quantify the samples. All samples were measured in triplicates. A mixture of all samples was used to normalize across plates. The average and standard deviation of the normalized samples were used for analysis.

Generation of iPSC-derived neurons from PRKN-mutant Parkinson's disease patients and controls

iPSCs were generated from fibroblasts of four Parkinson's disease patients (mean age±SD: 57.75±12.03 years; three females and one male) with biallelic *PRKN* mutations (homozygous c.1072Tdel; compound-heterozygous c.1072Tdel+delEx7; compound-heterozygous delEx4+c.924C>T; compound-heterozygous delEx1+c.924C>T) and three age-matched controls (mean age±SD: 45.33±9.74 years, one female and two male) as previously described[42] and cultured in mTeSR1™ complete medium (StemCell Technologies). First, iPSCs were differentiated into small molecule neural precursor cells (smNPCs) using an established protocol.[43] In short, smNPCs were cultured in N2B27 medium: Neurobasal (Gibco)/DMEM-F12 (Gibco) 50:50 and supplemented with 1% B27 without vitamin A (ThermoFisher), 0.5% N2 (Life Technologies), 1% penicillin-streptomycin (ThermoFisher) and 1% 200 mM glutamine (Westburg). During smNPC cultivation, N2/B27 medium was additionally supplemented with 3 µM CHIR 99021 (Sigma), 150 µM ascorbic acid (Sigma) and 0.5 µM purmorphamine (PMA) (Sigma). The medium was changed every other day.

The induction of midbrain neurons began after smNPCs reached the 15th passage. Once ~75% confluent, cells were detached by subjection to Accutase (Merck Millipore) for 5 min at 37°C and then centrifuged for 3 min at 300 x g at room temperature. Cells were then resuspended and counted via the Countess II FL Automated Cell Counter (ThermoFisher). 750,000 smNPCs per well were seeded onto Matrigel-coated 6-well plates in N2/B27 medium with 1 µM PMA, 200 µM ascorbic acid and 100 ng FGF8 (PeproTech) for 8 days. Next, cells were cultured in N2/B27 medium with 0.5 µM PMA and 200 µM ascorbic acid for two days. For the next 22 days, cells were cultured in N2/B27 medium with 200 µM ascorbic acid, 500 µM dibutyryl-cAMP

(Applichem), 1 ng/mL TGF-β3 (Peprotech), 10 ng/mL GDNF (Peprotech) and 20 ng/mL BDNF (PeproTech). The medium was changed every second day. Finally, DNA was prepared from previously characterized iPSC-derived neurons (*n*=5 million)[42] with the QIAAmp Minikit (Qiagen) and was quantified using standard methods.

Laser-microdissected midbrain tissue from PRKN-mutant Parkinson's disease patients and controls

Frozen human *postmortem* midbrain tissue sections and associated clinical and neuropathological data were supplied by Juntendo University and the Parkinson's UK brain bank at Imperial College London. One Parkinson's disease patient (age at death: 71 years old, male, postmortem interval: 14h, compound-heterozygous *PRKN* delEx2 + delEx3) and one control (age at death: 81 years old, male, postmortem interval: 34h) was used for this analysis. Frozen midbrain blocks were cryosectioned at ~15 µm thickness in the transverse plane. For the isolation of neuromelanin-positive neurons from the *substantia nigra*, sections were gradually thawed and fixed with 4% paraformaldehyde for 10 min. After three brief washes in TBST buffer, sections were left to dry on room temperature. Laser capture microdissection (LCM) was performed with the PALM MicroBeam (Zeiss) and isolated neurons were lysed in a buffer (50 mM Tris-HCl, pH 8.5, 1 mM EDTA, 0.5% Tween-20, 200ng/mL proteinase K) for 3hrs at 55°C and 10min at 90°C. For each person, we cut out 200 individual neurons, which were identified by the presence of neuromelanin deposits, from the *substantia nigra*. The total area of these separately laser-microdissected 200 cells was for the control: 445470 µm² and the PRKN-Parkinson's disease patient: 249728 μm^2, respectively. In addition, we isolated a region each from the center of the crus cerebri (without visible neuromelanin deposits) of the control (total area: 510609 μm^2) and the PRKN-Parkinson's disease (total area: 466368 μm^2) tissue. These dissected tissue areas were lysed in the same manner as the individual neurons. The resulting DNA lysates were used for Nanopore sequencing without further purification. The resulting DNA lysates were used in a long-range PCR (prior to Nanopore sequencing) without further purification. With our approach, we aimed to compare a dopaminergic neuron-rich region with a midbrain region that is free of dopaminergic neurons

RNA Ampliseq Transcriptome

RNA was extracted from blood of *PRKN* biallelic and monoallelic mutation carriers (*PRKN*/all) (*n*=23, mean age±SD: 56.2±13.2, mean AAO±SD: 38.8±18.2) and controls (*n*=6, mean age±SD: 71.7±9.1) with RNeasy Qiagen Minikit and DNase I digested prior to assessing the concentration, quality and RNA integrity with Agilent 2100 bioanalyzer, RNA LabChip kit and associated software (Agilent). Ampliseq™ whole transcriptome analysis was performed with the Ion Proton (Life Technologies, Inc) with average of 9.96M±1.14M reads. Transcriptome reads were aligned to the GRCh37/hg19 reference genome and analyzed by RNAseq Analysis plugin (V5.0.0.2) using default analysis settings. DESeq2 R[44] package (V1.18.1) within Bioconductor was used to test for differential expression by use of negative binomial generalized linear models; the estimates of dispersion and logarithmic fold changes incorporate data-driven prior distributions. The non-normalized counts of sequencing reads/fragments were used in DESeq2's statistical model that accounts for library size differences internally. Differentially expressed genes (DEGs) with nominally-significant p-values (<0.05) and Log2fold (L2F) changes (>|0.2|) were mapped on KEGG pathways with pathview.[44]

IL6 measurements in serum

To investigate interleukin 6 (IL6) levels in *PINK1/PRKN* biallelic and monoallelic mutation carriers (*n*=20, mean age±SD: 52.9 years ±14.3 years, 9 females), venous blood was collected in serum tubes and processed within one hour. Serum samples were centrifuged at 4°C for 10 min at 3000g, immediately frozen at -80°C and transferred to a certified diagnostic laboratory, where IL6 levels were measured using the Elecsys IL6 assay (Cobax).

Statistical Analyses

The affection status (whether the individual was diagnosed with Parkinson's disease), *PINK1/PRKN* mutational dosage or heteroplasmic variant load were used as outcome variables in this study. JMP software, Version 10 (SAS Institute Inc, Cary, NC USA) and R,[45] Version 4.1 were used to compare the clinical observations across different groups (i.e. Parkinson's disease vs asymptomatic; AAO; genotypes) and number of heteroplasmic mtDNA variants. For analyses in the blood-derived samples from the cohorts, Mann-Whitney U-tests were applied to compare two groups and Kruskal-Wallis tests and subsequent post-hoc pairwise comparisons were used

for more than two groups if $p<0.05$. With this approach, we corrected for multiple comparisons within each data set for a given research question. More specifically, six comparisons were performed post-hoc when investigating heteroplasmic mtDNA variants in control subjects, idiopathic Parkinson's disease, monoallelic and biallelic PINK1/PRKN mutation carriers. When looking at ND4:ND1 ratios across idiopathic Parkinson's disease, and patients with monoallelic and biallelic PINK1/PRKN mutations, we performed three post-hoc tests. Lastly, when comparing control subjects, idiopathic Parkinson's disease, monoallelic and biallelic PINK1/PRKN mutation carriers for ND4:ND1 and 7SDNA:ND1 ratios we performed six post-hoc tests. By contrast, we did not account for multiplicity between different data sets (e.g. mtDNA deletions vs mtDNA somatic point mutational load). To distinguish affection status with heteroplasmic mtDNA variants in monoallelic *PINK1*/*PRKN* mutation carriers, multivariable logistic regression models were used to predict disease status by mtDNA variants and age at examination. Then, the corresponding receiving operating characteristic (ROC) curves were reported. DESeq2 R package within Bioconductor was used to describe differences in gene expression using negative binomial distribution based on estimate variance-mean dependence in count data from high-throughput sequencing assays. The p-values were attained by the Wald test for the gene expression differences. Pearson correlation coefficient was reported for the correlation of IL6 and heteroplasmic variants. For analyses in iPSC-derived neurons, t-tests were performed.

Graphical output

All graphics and figures were created with GraphPad Prism and BioRender.com.

Data availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Results

Deep mitochondrial DNA sequencing with Illumina technology

After stringent filtering for uncontaminated samples, >1,000X coverage, <40 heteroplasmic mtDNA variants, *n*=325 individuals were included for further analyses. We identified a total

9,799 variants in 325 individuals. The mean number of overall homoplasmic variants was 28.15 (SD: ±15.1) per person and 2.3 (SD: ±2.67) per person for overall heteroplasmic variants.

Heteroplasmic mtDNA variants are associated with disease manifestation in PINK1/PRKN mutation carriers

We explored whether an association between mtDNA variant burden and Parkinson's disease status exists by comparing the cumulative mtDNA variant load between symptomatic (PD+) and asymptomatic carriers (PD-) of *PINK1* and *PRKN* monoallelic mutation carriers.

This analysis revealed a higher number of heteroplasmic variants in symptomatic (PINK1+/PRKN+/PD+, *n*=29), compared to asymptomatic monoallelic carriers (PINK1+/PRKN+/PD-, *n*=109) (Mann-Whitney U-test, Z=2.93, p=0.0033, *n*=138)**.** The resulting area under the curve indicated good discrimination based on low-level mtDNA variants and age at blood drawn for comparison between symptomatic (PD+) and asymptomatic carriers (PD-) of *PINK1* and *PRKN* monoallelic mutation carriers (AUC=0.83, CI:0.74-0.93) **(Figure 2a and b)**.

By contrast, a comparison of homoplasmic mtDNA variants did not show differences between symptomatic (PINK1+/PRKN+/PD+, *n*=29) and asymptomatic *PINK1* or *PRKN* monoallelic mutation carriers (PINK1+/PRKN+/PD-, *n*=109) (Mann-Whitney U-test, *Z*=0.43, *p*=0.66, **Figure S2**).

Heteroplasmic mtDNA variants increase with number of mutant PINK1/PRKN alleles

To assess whether the number of *PINK1* or *PRKN* mutations has an impact on the mtDNA variant burden, we compared the heteroplasmic variant load between affected and unaffected individuals with biallelic or monoallelic mutations, idiopathic Parkinson's disease patients and controls.

The number of heteroplasmic mtDNA variant load differed between patients with idiopathic Parkinson's disease (*n*=54) and patients with *PINK1/PRKN* biallelic or monoallelic mutations (PINK1/PRKN/PD+, *n*=78) (Mann-Whitney U-test, *p*=0.0441) **(Figure 2c)**. Separating patients with monoallelic (PINK1+/PRKN+/PD+, *n*=29) and biallelic *PINK1/PRKN* mutations

(PINK1++/PRKN++/PD+, *n*=49), no overall difference was seen to idiopathic Parkinson's disease patients (Kruskal-Wallis test, *p*=0.104) **(Figure 2d).**

Heteroplasmic mtDNA variant load differed between biallelic and monoallelic *PINK1/ PRKN* mutation carriers regardless of affection status (PINK1/PRKN/all) and controls (*n*=67), higher mtDNA variant load was observed for biallelic compared to monoallelic mutation carriers (Kruskal-Wallis test, *p*=0.0006) **(Figure 2e)**. Post-hoc pairwise analyses revealed higher heteroplasmic mtDNA variant load for biallelic compared to monoallelic *PINK1/ PRKN* mutation carriers (*p*=0.002) and higher variant load for biallelic PINK1/PRKN/ mutation carriers compared to controls (*p*=0.001).

The number of homoplasmic variants also increased with the number of mutant *PINK1/PRKN* alleles: mtDNA variants burden differed between patients with idiopathic Parkinson's disease (*n*=54) and *PINK1* or *PRKN* carriers regardless of affection status (*n*=78) **(Figure S3**, Kruskal-Wallis test, *p*<0.001**)**. As the homoplasmic mtDNA variants burden should not be influenced by age, we did not perform the age adjustment. These results suggest that heteroplasmic and homoplasmic mtDNA variant load are associated with *PINK1/PRKN* genotype. Of note, the samples with a higher number of heteroplasmic variants (**Figure 2e**) are not the same as those with a high number of homoplasmic variants (**Figure S3**).

MtDNA variants were also compared between idiopathic Parkinson's disease patients and controls in the same manner. The median values for the number of heteroplasmic variants and homoplasmic variants was not found to differ between idiopathic Parkinson's disease patients and controls (Mann-Whitney U-test *p*=0.0846 and *p*=0.4668, respectively).

Quantitative analyses of mtDNA-associated 7S DNA and major arc deletions

To extend our mtDNA variant burden analysis to include large-scale deletions, we investigated ND4:ND1 ratios. After filtering for quality control and removing outliers, *n*=302 individuals were included for further analyses.

Patients with biallelic *PINK1* or *PRKN* mutations (PINK1++/PRKN++/PD+; *n*=57) had more mtDNA deletions than patients with monoallelic *PINK1* or *PRKN* mutations (PINK1+/PRKN+/PD+; *n*=25) or idiopathic Parkinson's disease cases (*n*=50) (Kruskal-Wallis

test, p=0.009) **(Figure 3a)**. Disregarding the disease status, mtDNA major arc deletions were more abundant in biallelic mutation carriers (n=57) than in monoallelic mutation carriers (n=130) or controls (n=65) (Kruskal-Wallis test, p=0.003) **(Figure 3b)**.

MtDNA maintenance impairments may also extend to changes in transcription and replication of the mitochondrial genome. Events of mtDNA transcription/replication initiation were measured by determining 7S DNA:ND1 ratios **(Figure 3c and 3d)**. Biallelic mutation carriers (PINK1++/PRKN++/all; n=57) showed lower 7S DNA:ND1 ratios than controls (n=65); a smaller reduction in 7S DNA:ND1 ratios was seen for monoallelic mutation carriers (PINK1+/PRKN+/all; n=131) compared to controls, and the 7S DNA:ND1 ratios were different between biallelic (PINK1++/PRKN++/all) and monoallelic (PINK1+/PRKN+/all) *PINK1* or *PRKN* mutation carriers; lastly, compared to controls, idiopathic Parkinson's disease patients (n=49) also showed a similar reduction in 7S DNA:ND1 ratios (Kruskal-Wallis test p<0.001) like biallelic *PINK1* or *PRKN* mutation carriers (PINK1++/PRKN++/all) **(Figure 3d)**.

MtDNA variants in iPSC-derived dopaminergic neurons and midbrain tissue with Nanopore sequencing

Next, we tested whether alterations in mtDNA variant load in *PINK1/PRKN*-linked Parkinson's disease are a phenomenon restricted to peripheral tissues or whether they can also be observed in patient-derived neurons and midbrain tissue. To assess whether heteroplasmic mtDNA variants are present in neurons, we generated iPSC-derived midbrain neurons from four Parkinson's disease patients with biallelic *PRKN* mutations and three age-matched controls. MtDNA Illumina short-read deep-sequencing of these neuronal cultures revealed an increase in heteroplasmic variants in patient cells (p=0.0389) **(Figure 4a)**. These findings suggest that somatic mtDNA mutation accumulation not only occurs in peripheral tissues but also in iPSC-derived midbrain neurons from Parkinson's disease patients with biallelic *PRKN* mutations.

Due to limitations in short-read sequencing, we utilized Nanopore long-read sequencing for a more even coverage within amplicons, better mapping quality, and single-molecule technology (**Figure S4**). We first tested the technology in blood-derived *PRKN* monoallelic mutation carriers (n=16 select samples). Nanopore sequencing at >1,000X coverage confirmed 28 out of 32 (87.5%) variants (HF>0.05) in the Illumina short-read dataset. Besides data from our own

study, we followed the benchmarking protocol established for Nanopore long-read sequencing in low-frequency variants.[46] When applying 5%, 2% and 1% heteroplasmy thresholds on long-read sequencing data in the neurons, we found that the number of heteroplasmic variants was higher in Parkinson's disease patients with biallelic *PRKN* mutations compared to controls (HF>5%, p=0.0633, HF>2%, p=0.0024, HF>1%, *p*=0.0466) **(Figure 4b-d)**.

Finally, in light of recently reported limitations of mtDNA studies in iPSC-derived models,[47] we compared mtDNA variants in laser-microdissected neuromelanin-positive neurons isolated from postmortem *substantia nigra* and non-nigral midbrain tissue of a Parkinson's disease patient with biallelic *PRKN* mutations and one control. As we investigated lower HF levels, the difference between the nigral neurons (HF>5%, *n*=4; HF>2%, *n*=21; HF>1%, *n*=33) and the non-nigral midbrain tissue (HF>5%, *n*=5; HF>2%, *n*=7; HF>1%, *n*=13) increases in the *PRKN* mutation carrier. This phenomenon was not as pronounced in nigral neurons (HF>5%, *n*=7; HF>2%, *n*=10; HF>1%, *n*=18) compared to non-nigral midbrain tissue (HF>5%, *n*=5; HF>2%, *n*=6; HF>1%, *n*=14) in the control **(Figure 4e)**.

MtDNA variant load is associated with differentially-expressed genes (DEGs) in PRKN Parkinson's disease

Recently, mtDNA disintegration in *PRKN* KO mice and patient neurons with an error-prone mitochondrial polymerase gamma (POLG) has been linked to mtDNA release and the activation of pro-inflammatory pathways.[17,42] Thus, we first studied the global expression profile in *PRKN* mutation carriers (*PRKN*/all) and controls from the South Tyrolean cohort with RNA-seq data from available blood samples **(Table S3)**. This analysis revealed 1115 genes with nominally-significant differential expression L2F change >|0.2| (*p*<0.05). In this set, there was an enrichment for genes in the olfactory transduction pathway (*n*=22 genes), cytokine-cytokine receptor interaction (*n*=17 genes), and PI3K-Akt signaling pathway (*n*=14 genes) **(Table S4)**.

Second, to test whether mtDNA variant load is associated with changes in gene expression, we analyzed the RNA profiles in *PRKN* biallelic and monoallelic patients with more (*n*=5, mean±SD: 41.1±1.9) or fewer (*n*=6, mean±SD: 13.3±1.3) homoplasmic mtDNA variants. Still, these individuals were all from the South Tyrolean cohort. This analysis revealed 936 genes with nominally significant differential expression between high and low mtDNA variant load L2F

change >|0.2| (p<0.05 before adjustment), 40 genes overlapped with 1115 DEGs comparing *PRKN* mutation carriers to controls. In this set, there was also an enrichment for genes in the olfactory transduction pathway (n=70 genes), the PI3K-Akt signaling pathway (n=17 genes) and the cytokine-cytokine receptor interaction (n=15 genes) **(Table S4b)**. Three genes showed the strongest effects: *CD38* (ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase, L2F=-1.13, p=2.64x10^{-6}), *GSDMA* (Gasdermin A, L2F=-1.16, p=4.44x10^{-6}), and *RNASE1* (Ribonuclease A Family Member 1, L2F=-1.10, p=2.31x10^{-5}) **(Figure S5a)**.

IL6 levels and mtDNA heteroplasmy

Finally, prompted by our finding of altered gene expression in cytokine-cytokine receptor interaction pathways and by the recent literature suggesting a role for proinflammatory signaling in *PINK1*- and *PRKN*-associated Parkinson's disease,[17,48] we next assessed whether the heteroplasmic mtDNA variant load correlated with serum IL6 levels in a subset of *PINK1* and *PRKN* mutation carriers (PINK1/PRKN/all, n=20), where serum was available. In these individuals, heteroplasmic mtDNA variant load was found to correlate with serum IL6 levels (r=0.57, p=0.0074) **(Figure 4f)**. Similarly, the heteroplasmic mtDNA variant load correlated with serum IL6 in *PINK1* and *PRKN* monoallelic mutation carriers (PINK1+/PRKN+/all, n=14) (r=0.86, p=8.1x10^{-5}) **(Figure S5b)**. Homoplasmic mtDNA variant load did not correlate with serum IL6 levels (p=0.1812).

Discussion

In this study, we investigated the mtDNA variant burden as a marker of disease state in patients with *PINK1* and *PRKN* mutations. Deep sequencing of blood-derived mtDNA revealed an accumulation of heteroplasmic variants, which are likely to be of somatic origin[49] that could influence the clinical manifestation of *PINK1* or *PRKN* pathogenic mutations. Although the effect size of mtDNA heteroplasmy as a disease marker is relatively low, the ROC curve analyses showed predictive potential after including age (AUC>0.83). Unfortunately, only for a subset of individuals the exact age at blood-draw was available (n=231/325 total individuals). Moreover, the number of affected carriers was too sparse for a continuous AAO trait analysis, and may have been further hampered by the linear association of mtDNA damage and age.[50] However, being mindful of these limitations, the fact that mtDNA heteroplasmy could

distinguish between symptomatic and asymptomatic monoallelic carriers of *PINK1* or *PRKN* mutations but not between idiopathic Parkinson's disease and controls, still suggests that the effect is specific for individuals with *PINK1 or PRKN* mutations. As we found that mtDNA heteroplasmy is associated with disease manifestation, we may extrapolate mtDNA heteroplasmic variant load as a penetrance modifier. Interestingly, the analysis of heteroplasmic mtDNA variant load revealed a dosage effect for patients with *PINK1* or *PRKN* mutations. A similar dosage effect was also observed for major arc deletions and 7S DNA. Together, these findings suggest a mechanistic link between impaired mtDNA maintenance and increased mtDNA variant load. Parkin has been shown to modulate mtDNA transcription by ubiquitination of PARIS, which represses the mitochondrial biogenesis master regulator PGC1-alpha.[51] In addition, our own recent work revealed that a metabolic shift in Parkin-deficient neurons is sensed by SIRT1, which in turn causes a depletion of PGC1-alpha and mtDNA dyshomeostasis.[42] Accordingly, Parkin protects mtDNA from damage and stimulates mtDNA repair[52] (see **Figure S6** for an overview scheme).

An increase in somatic mutational load likely indicates a development during an individual's lifetime that can influence phenoconversion. While low-frequency heteroplasmic mtDNA changes are possibly a direct result of a disturbance in mitochondrial quality control in carriers of *PINK1* or *PRKN* mutations, homoplasmic mtDNA variants are likely to be inherited. The accumulation of homoplasmic variants could be explained by transmission of mtDNA variants across multiple generations over decades in families, a generational conversion of heteroplasmic into homoplasmic variants in populations with *PINK1* or *PRKN* mutations, though this can be influenced by specific mtDNA haplotypes. Individuals with a higher heteroplasmic variant load (n>14 heteroplasmic variants) are detailed in **Table S5**. These individuals do not overlap with individuals who have higher homoplasmic load, (n>78 homoplasmic variants) which originate mainly from the Tunisian population. These samples were also assigned to the "L" haplogroup (i.e., L0a1a3, L0a1a2, L1b1a, L1b1a16, L2a1c3a, L2a1b+143, L2a1+143 or L3h1b2), which are of African origin (**Table S6**).

To investigate whether the observed mtDNA phenotypes in carriers of *PINK1* or *PRKN* mutations are only a feature detectable in peripheral tissue or whether mtDNA integrity contributes to neurodegeneration in these individuals, we additionally studied neuronal samples. Comparing iPSC-derived midbrain neurons from four Parkinson's disease patients deficient of

Parkin and three controls, we confirmed an increase of somatic mtDNA mutations in the patient cells. Interestingly, our previous work in the *PRKN*-mutant neuronal cultures revealed impairments in the mtDNA maintenance and mitochondrial biogenesis pathways, which resulted in respiratory chain disturbance.[42,53] Together, these findings suggest a causal link between mtDNA dyshomeostasis, somatic mutation accumulation and mitochondrial dysfunction. By contrast, recent sequencing analyses in iPSC lines implicated a high mtDNA mutation rate during reprogramming.[47] While this technical limitation alone would not explain the genotype-specific differences in mtDNA mutational load observed here, it motivated us to additionally study *postmortem* nigral neurons from a biallelic *PRKN*- Parkinson's disease patient and a matched control. Indeed, in laser-microdissected dopaminergic neurons (which were identified based on their neuromelanin deposits) from the *PRKN*-mutant patient, we detected elevated numbers of somatic mutations when considering heteroplasmic frequencies >1% or >2%. Given that previous deep-sequencing analyses in single *postmortem* dopaminergic neurons from nigral tissue of idiopathic Parkinson's disease patients and controls did not reveal differences in terms of mtDNA somatic mutational load,[54] these findings further strengthen the role of Parkin in regulating mtDNA homeostasis. In addition, in line with previous results from idiopathic Parkinson's disease patient and control postmortem sections,[36,54] our comparison of the mtDNA status in neuromelanin-positive nigral neurons versus non-nigral midbrain tissue revealed that dopaminergic neurons are particularly vulnerable to this kind of damage. This phenomenon may be explained by the autoxidation properties of dopamine,[55] which facilitate the formation of free radicals. ROS, in turn, can cause single- and double-strand breaks, create abasic sites and oxidize purines and pyrimidines in the mitochondrial genome.[56] However, keeping in mind that due to the rarity of *PRKN*-Parkinson's disease brain tissue, we could only explore mtDNA somatic alterations in a single case, our *postmortem* results warrant further investigation in additional samples.

Apart from its role in mtDNA maintenance, Parkin deficiency has been shown to trigger mtDNA disintegration and dyshomeostasis as a result of impaired mitochondrial clearance.[51] Upon depolarization, oxidative stress or protein misfolding, PINK1 recruits Parkin to the mitochondria to initiate mitophagy.[51,57] If mitochondrial clearance fails in the absence of PINK1 or Parkin, pro-inflammatory signaling is induced. *PRKN* KO mice and iPSC-derived neurons with proofreading-deficient POLG as well as Parkinson's disease patients with *PINK1* or *PRKN*

mutations showed elevated circulating cell-free mtDNA levels,[58] which escapes from cells as consequence of impaired mitochondrial clearance and DNA maintenance. In the extracellular space, mtDNA can act as a damage-associated molecular pattern that is recognized by cGAS/STING and leads to inflammasome activation. Accordingly, *PRKN*-knockout mutator mice and *PINK1* or *PRKN*-mutant Parkinson's disease patients showed an upregulation of the cytokine IL6 in serum.[17,58] To further investigate the relationship between mtDNA integrity and inflammation, we quantified serum IL6 levels in a small subset of our *PINK1*- and *PRKN* mutation carriers. Remarkably, IL6 concentrations correlated with mtDNA heteroplasmic variant load in these samples. However, the sample size does not allow us to directly compare correlative differences between biallelic and monoallelic *PINK1/PRKN* mutation carriers. Whether IL6 levels differ between monoallelic and biallelic PINK1/PRKN mutation carriers and correlate with heteroplasmic mtDNA load needs to be further elucidated with larger sample sizes. In addition, we observed DEGs enriched in inflammatory pathways when comparing the gene expression profiles of *PRKN* mutation carriers with differential mtDNA variant load. These results concur with the above-mentioned findings in mice and human biosamples.[17,58] However, cellular studies will be needed to conclusively clarify if inflammation is a downstream effect triggered by accumulation of mtDNA variants and subsequent mtDNA release from damaged mitochondria.

The main strength lies in the in-depth analysis of mtDNA samples from the largest collection of *PINK1/PRKN* mutation carriers in one comprehensive investigation. Thus, while we would ideally aim to replicate our findings in an independent cohort, the overall rarity of individuals with *PINK1/PRKN* mutations currently precludes such a validation approach. The majority of samples included two founder populations: South Tyroleans and North Africa Arab Berbers, where the frequency of biallelic *PINK1* and *PRKN* variants are higher than in other ethnic groups. These populations provide genetic and environmental homogeneity to increase power for discovery and comprise the majority of our study cohort. Furthermore, our analysis compared specific controls from these founder populations, which also avoids genetic heterogeneity and background bias. However, the disease severity of patients and affection status of currently unaffected carriers may change and longitudinal follow up of our study cohort is warranted. We only had access to a small number of the corresponding mothers of investigated individuals to unequivocally exclude inherited mtDNA variants. In our study, heteroplasmic variants from deep

sequencing served as a surrogate marker for somatic mtDNA variation.[49] Using this approach, we provided functional evidence for a role of monoallelic mutations in *PINK1/PRKN* as Parkinson's disease risk factors or pathogenic variants of highly reduced penetrance. *PINK1* and *PRKN* variants were labelled as "pathogenic mutations" based on the MDSGene consortium guidelines criteria of 'possibly, probably or definitely' pathogenic. Although it is becoming common practice to use ACMG, there is much room for interpretation in the criteria, especially for age-related neurodegenerative diseases that have reduced penetrance. Instead MDSGene scoring of pathogenicity has been used as it has been reviewed thoroughly, published and used effectively in scoring pathogenicity.

In addition, we applied single-molecule sequencing to *postmortem* midbrain tissue and confirmed an increase of somatic mutations in nigral dopaminergic neurons from a *PRKN* Parkinson's disease case. Finally, our data provides further evidence for the recently discovered link between the PINK1/Parkin mitophagy pathway and inflammation in the context of increased mtDNA heteroplasmy in human samples. In the future, mitochondrial variant load should also be investigated in patients and control subjects with pathogenic variants in Parkinson's disease genes that play a role in mitochondrial metabolism (e.g. *VPS13C, DNAJC6*).

In summary, our findings strengthen the relevance of the PINK1/Parkin pathway in maintaining a healthy mitochondrial pool. In addition, our study highlights mtDNA heteroplasmy as a potential disease manifestation marker for *PINK1*/*PRKN* Parkinson's disease. Thus, patient blood-based deep mtDNA sequencing may be considered for genetic counseling and future clinical trials. This is important for monitoring disease onset in asymptomatic carriers, as mtDNA heteroplasmy is reliably measurable.

Acknowledgements

The authors wish to thank the many patients and their families who volunteered, and the efforts of the many clinical teams involved. Initial studies in Tunisia on familial parkinsonism were in collaboration with Lefkos Middleton, Rachel Gibson, and the GlaxoSmithKline PD Programme Team (2002-2005). We would like to thank Dr Helen Tuppen from the Welcome Trust Centre for Mitochondrial Research, Newcastle University, UK for providing us with the plasmid p7D1.

Moreover, this project was supported by the high throughput/high content screening platform and HPC facility at the Luxembourg Centre for Systems Biomedicine, and the University of Luxembourg.

Funding

Funding has been obtained from the German Research Foundation ("ProtectMove"; FOR 2488, GR 3731/5-1; SE 2608/2-1; KO 2250/7-1), the Luxembourg National Research Fund in the ATTRACT ("Model-IPD", FNR9631103), NCER-PD (FNR11264123) and INTER programmes ("ProtectMove", FNR11250962; "MiRisk-PD", C17/BM/11676395, NB 4328/2-1), the BMBF (MitoPD), the Hermann and Lilly Schilling Foundation, the European Community (SysMedPD), the Canadian Institutes of Health Research (CIHR), Peter and Traudl Engelhorn Foundation.

Competing interests

Authors report no competing interests relevant to the manuscript.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Dorsey ER, Sherer T, Okun MS, Bloem BR. The Emerging Evidence of the Parkinson Pandemic. *J Parkinsons Dis*. 2018;8(s1):S3-S8. doi:10.3233/JPD-181474

JPD181474 [pii]

2. Bloem BR, Okun MS, Klein C. Parkinson's disease. *Lancet*. Jun 12 2021;397(10291):2284-2303. doi:10.1016/S0140-6736(21)00218-X

3. Hedrich K, Hagenah J, Djarmati A, et al. Clinical spectrum of homozygous and heterozygous PINK1 mutations in a large German family with Parkinson disease: role of a single hit? *Arch Neurol*. Jun 2006;63(6):833-8. doi:63/6/833 [pii]

10.1001/archneur.63.6.833

4. Pramstaller PP, Schlossmacher MG, Jacques TS, et al. Lewy body Parkinson's disease in a large pedigree with 77 Parkin mutation carriers. *Ann Neurol*. Sep 2005;58(3):411-22. doi:10.1002/ana.20587

5. Klein C, Lohmann-Hedrich K, Rogaeva E, Schlossmacher MG, Lang AE. Deciphering the role of heterozygous mutations in genes associated with parkinsonism. *Lancet Neurol*. Jul 2007;6(7):652-62. doi:S1474-4422(07)70174-6 [pii]

10.1016/S1474-4422(07)70174-6

6. Kasten M, Weichert C, Lohmann K, Klein C. Clinical and demographic characteristics of PINK1 mutation carriers--a meta-analysis. *Mov Disord*. May 15 2010;25(7):952-4. doi:10.1002/mds.23031

7. Weissbach A, Konig IR, Huckelheim K, et al. Influence of L-dopa on subtle motor signs in heterozygous Parkin- and PINK1 mutation carriers. *Parkinsonism Relat Disord*. Sep 2017;42:95-99. doi:S1353-8020(17)30236-5 [pii]

10.1016/j.parkreldis.2017.07.003

8. Hilker R, Klein C, Ghaemi M, et al. Positron emission tomographic analysis of the nigrostriatal dopaminergic system in familial parkinsonism associated with mutations in the parkin gene. *Ann Neurol*. Mar 2001;49(3):367-76.

9. Krohn L, Grenn FP, Makarious MB, et al. Comprehensive assessment of PINK1 variants in Parkinson's disease. *medRxiv*. 2020:2020.01.21.20018101. doi:10.1101/2020.01.21.20018101

10. Yu E, Rudakou U, Krohn L, et al. Analysis of Heterozygous PRKN Variants and Copy-Number Variations in Parkinson's Disease. *Mov Disord*. Jan 2021;36(1):178-187. doi:10.1002/mds.28299

11. Zhu W, Huang X, Yoon E, et al. Heterozygous PRKN mutations are common but do not increase the risk of Parkinson's disease. *Brain*. Jun 30 2022;145(6):2077-2091. doi:10.1093/brain/awab456

12. Pottier C, Rampersaud E, Baker M, et al. Identification of compound heterozygous variants in OPTN in an ALS-FTD patient from the CReATe consortium: a case report. *Amyotroph Lateral Scler Frontotemporal Degener*. Aug 2018;19(5-6):469-471. doi:10.1080/21678421.2018.1452947

13. Greuel A, Trezzi JP, Glaab E, et al. GBA Variants in Parkinson's Disease: Clinical, Metabolomic, and Multimodal Neuroimaging Phenotypes. *Mov Disord*. Dec 2020;35(12):2201-2210. doi:10.1002/mds.28225

14. Kandul NP, Zhang T, Hay BA, Guo M. Selective removal of deletion-bearing mitochondrial DNA in heteroplasmic Drosophila. *Nat Commun*. Nov 14 2016;7:13100. doi:10.1038/ncomms13100

ncomms13100 [pii]

15. Gilkerson RW, De Vries RL, Lebot P, et al. Mitochondrial autophagy in cells with mtDNA mutations results from synergistic loss of transmembrane potential and mTORC1 inhibition. *Hum Mol Genet*. Mar 1 2012;21(5):978-90. doi:10.1093/hmg/ddr529

ddr529 [pii]

16. Suen DF, Narendra DP, Tanaka A, Manfredi G, Youle RJ. Parkin overexpression selects against a deleterious mtDNA mutation in heteroplasmic cybrid cells. *Proc Natl Acad Sci U S A*. Jun 29 2010;107(26):11835-40. doi:10.1073/pnas.0914569107

0914569107 [pii]

17. Sliter DA, Martinez J, Hao L, et al. Parkin and PINK1 mitigate STING-induced inflammation. *Nature*. Aug 22 2018;doi:10.1038/s41586-018-0448-9

10.1038/s41586-018-0448-9 [pii]

18. Trinh J, Lohmann K, Baumann H, et al. Utility and implications of exome sequencing in early-onset Parkinson's disease. *Mov Disord*. Jan 2019;34(1):133-137. doi:10.1002/mds.27559

19. Kasten M, Hartmann C, Hampf J, et al. Genotype-Phenotype Relations for the Parkinson's Disease Genes Parkin, PINK1, DJ1: MDSGene Systematic Review. *Mov Disord*. May 2018;33(5):730-741. doi:10.1002/mds.27352

20. Lill CM, Mashychev A, Hartmann C, et al. Launching the movement disorders society genetic mutation database (MDSGene). *Mov Disord*. May 2016;31(5):607-9. doi:10.1002/mds.26651

21. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. Jul 15 2009;25(14):1754-60. doi:10.1093/bioinformatics/btp324

22. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. Oct 1999;23(2):147. doi:10.1038/13779

23. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. Aug 15 2009;25(16):2078-9. doi:10.1093/bioinformatics/btp352

24. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. Sep 2010;20(9):1297-303. doi:10.1101/gr.107524.110

25. Okonechnikov K, Conesa A, Garcia-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*. Jan 15 2016;32(2):292-4. doi:10.1093/bioinformatics/btv566

26. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. Oct 1 2016;32(19):3047-8. doi:10.1093/bioinformatics/btw354

27. Weissensteiner H, Forer L, Fuchsberger C, et al. mtDNA-Server: next-generation sequencing data analysis of human mitochondrial DNA in the cloud. *Nucleic Acids Res*. Jul 8 2016;44(W1):W64-9. doi:10.1093/nar/gkw247

28. Weissensteiner H, Forer L, Fendt L, et al. Contamination detection in sequencing studies using the mitochondrial phylogeny. *Genome Res*. Jan 15 2021;doi:10.1101/gr.256545.119

29. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*. Aug 1 2018;34(15):2666-2669. doi:10.1093/bioinformatics/bty149

30. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. Sep 15 2018;34(18):3094-3100. doi:10.1093/bioinformatics/bty191

31. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. Oct 1 2015;526(7571):68-74. doi:10.1038/nature15393

32. Bolze A, Mendez F, White S, et al. Selective constraints and pathogenicity of mitochondrial DNA variants inferred from a novel database of 196,554 unrelated individuals. *bioRxiv*. 2019:798264. doi:10.1101/798264

33. Dayama G, Emery SB, Kidd JM, Mills RE. The genomic landscape of polymorphic human nuclear mitochondrial insertions. *Nucleic Acids Res*. Nov 10 2014;42(20):12640-9. doi:10.1093/nar/gku1038

34. Castellana S, Biagini T, Petrizzelli F, et al. MitImpact 3: modeling the residue interaction network of the Respiratory Chain subunits. *Nucleic Acids Res*. Jan 8 2021;49(D1):D1282-D1288. doi:10.1093/nar/gkaa1032

35. Pereira L, Soares P, Triska P, et al. Global human frequencies of predicted nuclear pathogenic variants and the role played by protein hydrophobicity in pathogenicity potential. *Sci Rep*. Nov 21 2014;4:7155. doi:10.1038/srep07155

36. Bender A, Krishnan KJ, Morris CM, et al. High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. *Nat Genet*. May 2006;38(5):515-7. doi:ng1769 [pii]

10.1038/ng1769

37. Rocha MC, Rosa HS, Grady JP, et al. Pathological mechanisms underlying single large-scale mitochondrial DNA deletions. *Ann Neurol*. Jan 2018;83(1):115-130. doi:10.1002/ana.25127

38. Grady JP, Murphy JL, Blakely EL, et al. Accurate measurement of mitochondrial DNA deletion level and copy number differences in human skeletal muscle. *PLoS One*. 2014;9(12):e114462. doi:10.1371/journal.pone.0114462

39. Yu-Wai-Man P, Lai-Cheong J, Borthwick GM, et al. Somatic mitochondrial DNA deletions accumulate to high levels in aging human extraocular muscles. *Invest Ophthalmol Vis Sci*. Jul 2010;51(7):3347-53. doi:10.1167/iovs.09-4660

40. Nicholls TJ, Zsurka G, Peeva V, et al. Linear mtDNA fragments and unusual mtDNA rearrangements associated with pathological deficiency of MGME1 exonuclease. *Hum Mol Genet*. Dec 1 2014;23(23):6147-62. doi:10.1093/hmg/ddu336

ddu336 [pii]

41. Grunewald A, Rygiel KA, Hepplewhite PD, Morris CM, Picard M, Turnbull DM. Mitochondrial DNA Depletion in Respiratory Chain-Deficient Parkinson Disease Neurons. *Ann Neurol*. Mar 2016;79(3):366-78. doi:10.1002/ana.24571

42. Wasner K, Smajic S, Ghelfi J, et al. Parkin Deficiency Impairs Mitochondrial DNA Dynamics and Propagates Inflammation. *Mov Disord*. Jul 2022;37(7):1405-1415. doi:10.1002/mds.29025

43. Reinhardt P, Glatza M, Hemmer K, et al. Derivation and expansion using only small molecules of human neural progenitors for neurodegenerative disease modeling. *PLoS One*. 2013;8(3):e59252. doi:10.1371/journal.pone.0059252

44. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:s13059-014-0550-8 [pii]

10.1186/s13059-014-0550-8

45. *A language and environment for statistical computing. R Foundation for Statistical Computing*. 2020. https://www.R-project.org/

46. Luth T, Schaake S, Grunewald A, May P, Trinh J, Weissensteiner H. Benchmarking Low-Frequency Variant Calling With Long-Read Data on Mitochondrial DNA. *Front Genet*. 2022;13:887644. doi:10.3389/fgene.2022.887644

47. Wei W, Gaffney DJ, Chinnery PF. Cell reprogramming shapes the mitochondrial DNA landscape. *Nat Commun*. Sep 2 2021;12(1):5241. doi:10.1038/s41467-021-25482-x

48. Matheoud D, Cannon T, Voisin A, et al. Intestinal infection triggers Parkinson's disease-like symptoms in Pink1(-/-) mice. *Nature*. Jul 2019;571(7766):565-569. doi:10.1038/s41586-019-1405-y

10.1038/s41586-019-1405-y [pii]

49. Payne BA, Wilson IJ, Yu-Wai-Man P, et al. Universal heteroplasmy of human mitochondrial DNA. *Hum Mol Genet*. Jan 15 2013;22(2):384-90. doi:10.1093/hmg/dds435

dds435 [pii]

50. Kraytsberg Y, Kudryavtseva E, McKee AC, Geula C, Kowall NW, Khrapko K. Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nat Genet*. May 2006;38(5):518-20. doi:ng1778 [pii]

10.1038/ng1778

51. Grunewald A, Kumar KR, Sue CM. New insights into the complex role of mitochondria in Parkinson's disease. *Prog Neurobiol*. Jun 2019;177:73-93. doi:S0301-0082(18)30065-0 [pii]

10.1016/j.pneurobio.2018.09.003

52. Rothfuss O, Fischer H, Hasegawa T, et al. Parkin protects mitochondrial genome integrity and supports mitochondrial DNA repair. *Hum Mol Genet*. Oct 15 2009;18(20):3832-50. doi:10.1093/hmg/ddp327

ddp327 [pii]

53. Zanon A, Kalvakuri S, Rakovic A, et al. SLP-2 interacts with Parkin in mitochondria and prevents mitochondrial dysfunction in Parkin-deficient human iPSC-derived neurons and Drosophila. *Hum Mol Genet*. Jul 1 2017;26(13):2412-2425. doi:10.1093/hmg/ddx132

3098491 [pii]

54. Dolle C, Flones I, Nido GS, et al. Defective mitochondrial DNA homeostasis in the substantia nigra in Parkinson disease. *Nat Commun*. Nov 22 2016;7:13548. doi:10.1038/ncomms13548

ncomms13548 [pii]

55. LaVoie MJ, Hastings TG. Dopamine quinone formation and protein modification associated with the striatal neurotoxicity of methamphetamine: evidence against a role for extracellular dopamine. *J Neurosci*. Feb 15 1999;19(4):1484-91.

56. Yakes FM, Van Houten B. Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc Natl Acad Sci U S A*. Jan 21 1997;94(2):514-9. doi:10.1073/pnas.94.2.514

57. Pickrell AM, Huang CH, Kennedy SR, et al. Endogenous Parkin Preserves Dopaminergic Substantia Nigral Neurons following Mitochondrial DNA Mutagenic Stress. *Neuron*. Jul 15 2015;87(2):371-81. doi:10.1016/j.neuron.2015.06.034

S0896-6273(15)00597-8 [pii]

58. Borsche M, Konig IR, Delcambre S, et al. Mitochondrial damage-associated inflammation highlights biomarkers in PRKN/PINK1 parkinsonism. *Brain*. Oct 1 2020;143(10):3041-3051. doi:10.1093/brain/awaa246

Figure legends

Figure 1 Overview of study. Flow chart showing individuals incorporated in this study, biospecimen, rationale and techniques. mtDNA: mitochondrial DNA, NGS: next-generation sequencing, NM: neuromelanin, TGS: third-generation sequencing.

Figure 2 Disease manifestation and *PRKN*/*PINK1* status are associated with number of heteroplasmic mtDNA variants in blood-derived DNA. **(a)** The sensitivity and specificity of the heteroplasmic mtDNA variants (HF>5%) as a biological marker plus age at examination were assessed with ROC analysis. The area under the curve (AUC) indicated good discrimination for affected (n=29) vs. unaffected (n=109) monoallelic *PRKN/PINK1* mutation carriers (AUC=0.83, CI:0.74-0.93). **(b)** Scatter plot showing the number of heteroplasmic mtDNA variants for affected (n=29) vs. unaffected (n=109) monoallelic *PRKN/PINK1* mutation carriers. Mann-Whitney U-test was performed. **(c-d)** Scatter plots showing number heteroplasmic mtDNA variants (HF>5%) for **(c)** idiopathic PD (iPD, n=54) vs PINK1/PRKN/PD+ (n=78); **(d)** iPD (n=54) vs PINK1+/PRKN+/PD+ (n=29) vs PINK1++/PRKN++/PD+ (n=49); **(e)** controls (n=67) vs iPD (n=54) vs all monoallelic mutation carriers (n=150) vs all biallelic mutation carriers (n=52). Mann-Whitney U-test was performed for (c) and Kruskal-Wallis tests for (d) and (e) shown in bold. Post-hoc analyses are not bolded. Bars indicate means and 95%CI. PINK1: monoallelic and biallelic *PINK1* mutations; PRKN: monoallelic and biallelic *PRKN* mutations; PINK1+:*PINK1* monoallelic, PINK1++: *PINK1* biallelic mutations; PRKN+=*PRKN* monoallelic, PRKN++:*PRKN* biallelic mutations; PD+:patient with Parkinson's disease; all=individuals in study with or without Parkinson's disease.

Figure 3 Analysis of mtDNA major arc deletions and mtDNA transcription-associated 7S DNA by real-time PCR in blood-derived DNA. Simultaneous real-time PCR quantification of the mtDNA genes *ND4* relative to *ND1* or detection of 7S DNA relative to *ND1*. Scatter plots showing ND4:ND1 ratios for **(a)** patients with idiopathic PD (iPD, n=50) vs patients with PD due to monoallelic (n=25) or biallelic (n=57) mutations in *PINK1* or *PRKN*. **(b)** Controls (n=65)

vs iPD (n=50) vs all *PINK1* or *PRKN* monoallelic mutation carriers (n=130) vs all *PINK1* or *PRKN* biallelic mutation carriers (n=57). Scatter plot showing 7S DNA:ND1 ratios for **(c)** patients with iPD (n=49) vs patients with PD due to monoallelic (n=25) or biallelic (n=57) mutations in *PINK1* or *PRKN*. **(d)** Controls (n=65) vs iPD patients (n=49) vs all monoallelic *PINK1*/*PRKN* mutation carriers (n=131) vs all *PINK1* or *PRKN* biallelic mutation carriers (n=57). Kruskal-Wallis tests were performed and in bold. Post-hoc analyses are not bolded. Bars indicate means and 95%CI; PINK1+:PINK1 monoallelic, PINK1++: *PINK1* biallelic mutations; PRKN+: *PRKN* monoallelic, PRKN++: *PRKN* biallelic mutations; PD+: patient with Parkinson's disease.

Figure 4 Heteroplasmic mtDNA variants are associated with *PRKN* mutation carriers in iPSC-derived and postmortem neurons and inflammation. Scatter plot showing number of heteroplasmic mtDNA variants in iPSC-derived midbrain neurons from 3-4 differentiations of controls (n=4) and biallelic *PRKN*-PD patients (n=4) performed with **(a)** Illumina short-read sequencing at >5%, **(b)** Nanopore long-read sequencing at >5%, **(c)** Nanopore long-read sequencing at >2%, and **(d)** Nanopore long-read sequencing at >1%. T-tests were performed. PRKN++: *PRKN* biallelic mutations; PD+: patients with Parkinson's disease. **(e)** Line plot showing number heteroplasmic mtDNA variants for control- vs PRKN++/PD+-derived nigral laser-microdissected neuromelonin (NM)-positive neurons and non-nigral midbrain tissue at >5%, >2%, and >1%. **(f)** Correlation of serum IL6 and total mtDNA variant load in *PINK1* or *PRKN* monoallelic (n=14) and biallelic (n=6) mutation carriers. Bars indicate means and 95% CI.

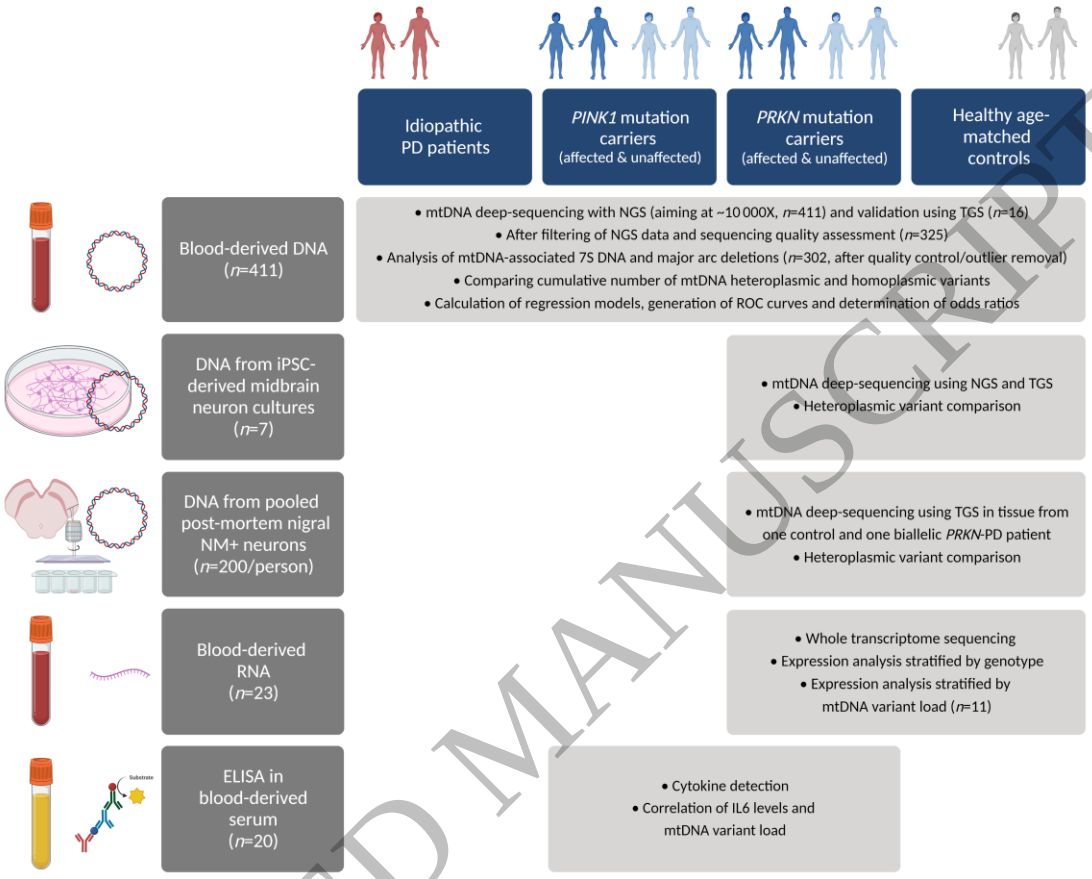

Figure 1
159x124 mm (3.8 x DPI)

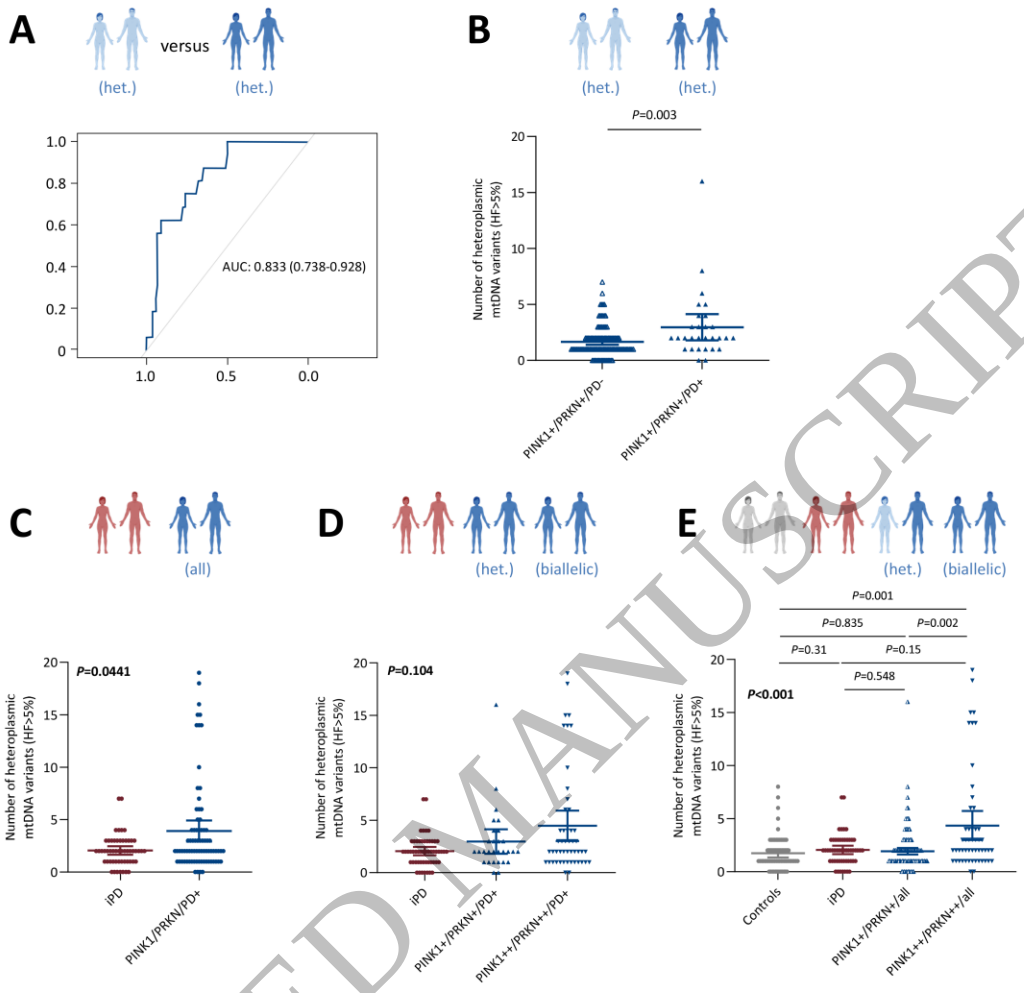

Figure 2
142x132 mm (3.8 x DPI)

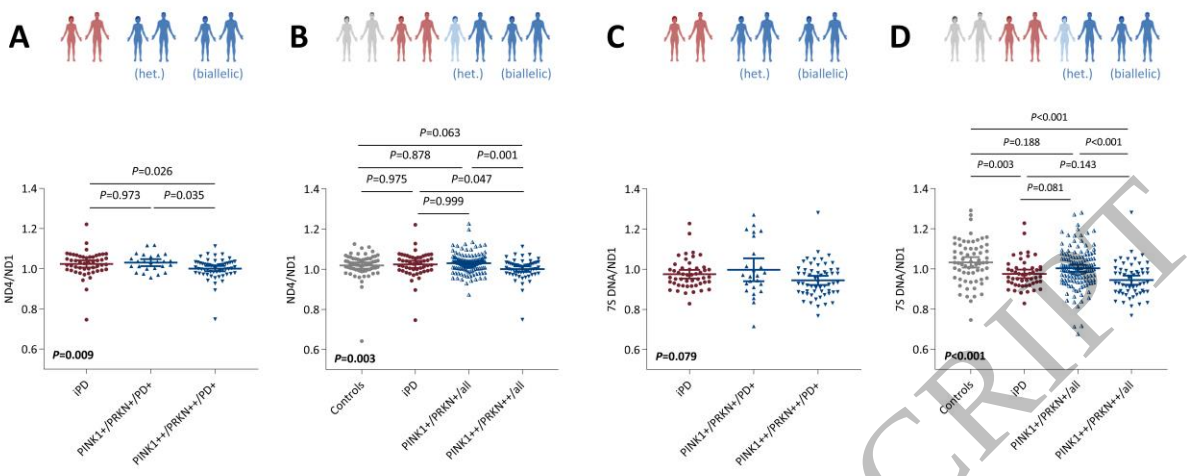

Figure 3
159x63 mm (3.8 x DPI)

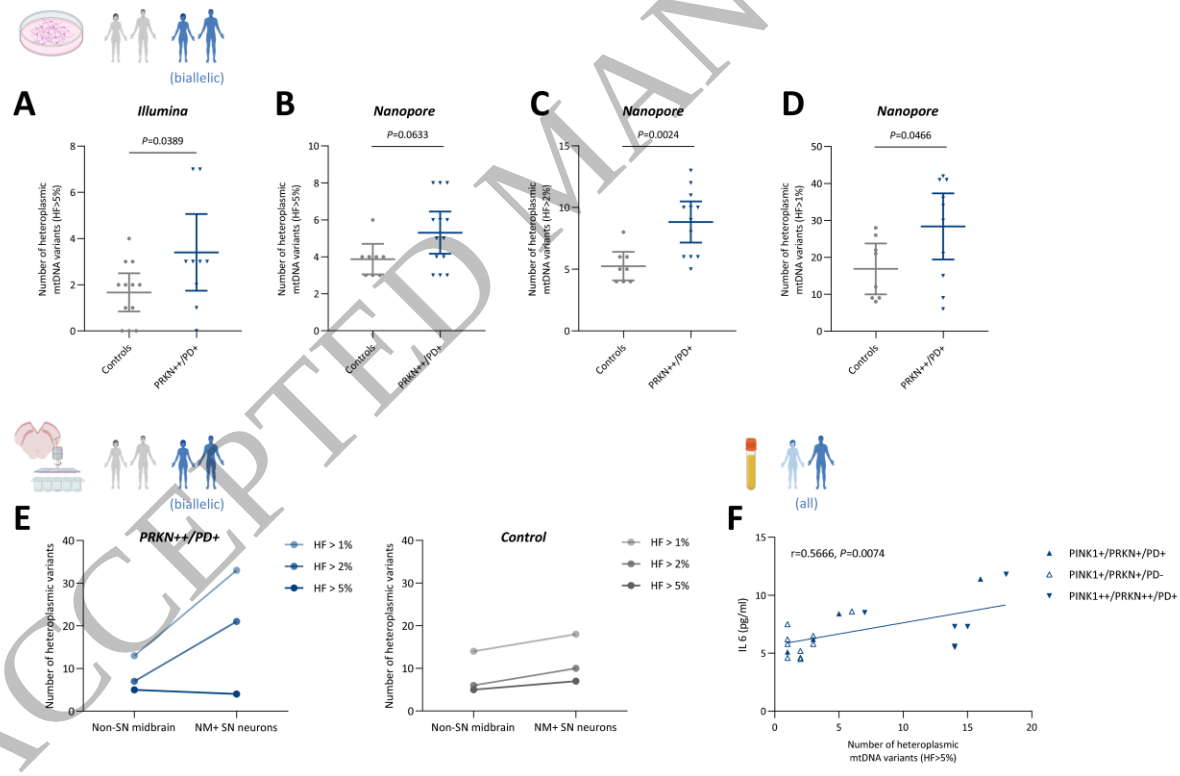

Figure 4
159x104 mm (3.8 x DPI)

Table 1 Demographics of all cohorts: individuals with *PINK1*/*PRKN* mutations, idiopathic Parkinson's disease and controls

	PINK1*/*all			***PRKN*/*all***			**iPD**	**Control subjects**
	Affected (PD+)	**Unaffected (PD−)**	**NA**	**Affected (PD+)**	**Unaffected (PD−)**	**NA**		
N	**59**	**47**	**6**	**54**	**79**	**9**	**67**	**90**
Number of men (%)	31 (52.5%)	25 (53%)	6 (100%)	25 (46%)	38 (48%)	1 (11%)	39 (60%)	38 (43%)
Mean age (SD) years	64.2 (15.2)	62.1 (11.6)	61.6 (17.7)	59.7 (16.9)	47.9 (16.9)	59.3(16.1)	74.3 (10.2)	61.7 (18.3)
Median age (IQR)	64.5 (50.75–79)	60(54–71)	66 (48–73)	59.5 (47.75–73)	51 (33–61)	57 (50–74.25)	77 (69–82)	64 (53.5–75.25)
Mean AAO (SD)	38.8 (14.2) (*n* = 53)	–	–	36.9 (16.4) (*n* = 50)	–	–		–
Median AAO (IQR)	36 (29.25–50) (*n* = 53)	–	–	34 (26–46) (*n* = 50)	–	–		–

PINK1/all: Monoallelic and biallelic *PINK1* mutation carriers regardless of affection status, *PRKN*/all: Monoallelic and biallelic *PRKN* mutation carriers regardless of affection status, IPD: idiopathic PD, SD: standard deviation, IQR: interquartile range, NA: not available, PD+: affected, PD-: unaffected, AAO: age at onset.

Table 2 Demographics of cohorts stratified by genotype

	PINK1*/*all		***PRKN*/*all***	
	Biallelic (PINK1++/all)	**Monoallelic (PINK1+/all)**	**Biallelic (PRKN++/all)**	**Monoallelic (PRKN+/all)**
N	**57**	**55**	**27**	**115**
Number of men (%)	29 (51%)	32 (58%)	10 (37%)	54 (47%)
Mean age (SD) years	63.4 (14.6)	63 (13.1)	57.3 (17)	52 (17.9)
Median age (IQR)	63 (51–75)	61 (54–75)	58 (48–69)	54 (36–65.3)
Mean AAO (SD)	37.5 (12.9) (*n* = 51)	71 (1.4) (*n* = 2)	35.9 (16.9) (*n* = 28)	40.2 (15.6) (*n* = 22)
Median AAO (IQR)	36 (29–49.3) (*n* = 51)	71 (70–71) (*n* = 2)	32.5 (22.8–46) (*n* = 28)	39.5 (31.3–50) (*n* = 22)

PINK1/all: Monoallelic and biallelic *PINK1* mutation carriers regardless of affection status, *PRKN*/all: Monoallelic and biallelic *PRKN* mutation carriers regardless of affection status, PINK1++: *PINK1* biallelic, *PRKN*++: *PRKN* biallelic, PINK1+: *PINK1* Monoallelic, *PRKN*+: *PRKN* Monoallelic, SD: standard deviation, IQR: interquartile range, NA: not available, PD+: affected, PD-: unaffected, AAO: age at onset