AWARD NUMBER: W81XWH-17-1-0249

TITLE: Early-Onset Parkinson's Disease Is a Mitochondrial Disease: A Nigral Mitochondrial Cytopathy

PRINCIPAL INVESTIGATOR: Owen A. Ross, PhD

CONTRACTING ORGANIZATION: Mayo Clinic, Jacksonville, FL

REPORT DATE: December 2022

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
December 2022	Final	15Aug2017 - 14Aug2022

4. TITLE AND SUBTITLE

Early-Onset Parkinson's Disease Is a Mitochondrial Disease: A Nigral Mitochondrial Cytopathy

5a. CONTRACT NUMBER
PR160606

5b. GRANT NUMBER
W81XWH-17-1-0249

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Owen A. Ross, PhD

E-Mail:

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Mayo Clinic Jacksonville
4500 San Pablo Road
Jacksonville, Florida 32224

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Loss of function mutations in the genes encoding PINK1 and Parkin results in early-onset forms of Parkinson's disease (EOPD). Both enzymes are functionally linked and together direct a neuroprotective mitochondrial quality control (mitoQC) ensuring elimination of damaged organelles from cells via the autophagy-lysosome system (i.e. mitophagy), which is lost in EOPD. Given the complexity of this pathway and the general missing heritability in EOPD, it is highly likely that additional genes regulating this pathway may also be found mutated in EOPD. The overarching goals of this project are to 1) identify high confidence genetic modifiers of the PINK1/Parkin pathway by a two-tiered functional screening (overlay of genome-wide siRNA and miRNA screens) in cells, 2) to identify the underlying genetic variation and characterize the EOPD genome (whole-genome-sequencing of patients), as well as 3) to determine the pathogenicity of these novel EOPD sequence variants in functional readout studies. Using this combined functional genetics approach we will determine the regulation of mitophagy as well as the genetic architecture of EOPD.

15. SUBJECT TERMS
early-onset Parkinson's disease, mitochondrial quality control, mitophagy, PINK1, Parkin, functional genomic screening, genetic architecture

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRDC
Unclassified	Unclassified	Unclassified	Unclassified	24	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

1. Introduction……....4

2. Keywords……....4

3. Accomplishments……..…….....4

4. Impact……...……11

5.Changes/Problems...….……11

6. Products, Inventions, Patent Applications, and/or Licenses..……..12

7. Participants & Other Collaborating Organizations…….14

8. Special Reporting Requirements…….....16

9. Appendices……....17

1. **INTRODUCTION:** Loss of function mutations in the genes encoding PINK1 and Parkin results in early-onset forms of Parkinson's disease (EOPD). Both enzymes are functionally linked and together direct a neuroprotective mitochondrial quality control (mitoQC) ensuring elimination of damaged organelles from cells via the autophagy-lysosome system (i.e. mitophagy), which is lost in EOPD. Given the complexity of this pathway and the general missing heritability in EOPD, it is highly likely that additional genes regulating this pathway may also be found mutated in EOPD. The overarching goals of this project are to 1) identify high confidence genetic modifiers of the PINK1/Parkin pathway by a two-tiered functional screening (overlay of genome-wide siRNA and miRNA screens) in cells, 2) to identify the underlying genetic variation and characterize the EOPD genome (whole-genome-sequencing of patients), as well as 3) to determine the pathogenicity of these novel EOPD sequence variants in functional readout studies. Using this combined functional genetics approach we will determine the regulation of mitophagy as well as the genetic architecture of EOPD.

2. **KEYWORDS:** early-onset Parkinson's disease, mitochondrial quality control, mitophagy, PINK1, Parkin, functional genomic screening

3. **ACCOMPLISHMENTS:**
   - **What were the major goals of the project?**
     Major Task 1: Nomination of mitoQC candidate genes by an accelerated, two-tiered functional screen and processing through bioinformatics resource/filtering strategy – Month 1-18

     Major Task 2: Whole-Genome sequencing in patients with EOPD and nomination of disease genes/variants – Month 1-36

     Major Task 3: Validation of high-confidence mitoQC/EOPD genes and dysfunctions of sequence variants on molecular, cellular, and organismal level – Month 6-36

   - **What was accomplished under these goals?**
     Towards accomplishing the major goals of the project, we first completed the regulatory review and approval by Mayo Clinic IRB and the HRPO (subtask 1.1); and this has been renewed for all subsequent years (Appendix).

     We have continued to screen newly collected early-onset patients with Parkinson's disease to identify those that are negative for *PINK1* and *PARKIN* mutations, and have prioritized these subjects for whole-genome sequencing. In Year 1, we identified all early-onset patients for which we have collected fibroblast biopsies and grown the cell lines and extracted DNA. In Year 2 we sequenced these cell lines and are analyzing the sequence variants identified. We plan to functionally assess the influence of variants identified within these subjects on the efficiency of the mitophagy pathway working with Dr. Springer (Co-PI). In year 3, we had whole-genome sequencing for 95 patients with early-onset disease and we had a further 55 patient samples being processed and focused on analysis single nucleotide variation. In Years 3 and 4 we have encountered major delays in both processing and bioinformatic analysis of the data given the situation with the COVID-19 pandemic. The Core facilities at the clinic were either closed or major efforts were repurposed on projects directly-related to the pandemic. We also unfortunately lost some staff members as well and with recruitment placed on hold at the clinic during most of Year 3 we have struggled to recruit replacement fellows.

     In year 4 we completed the sequencing of the 150 patients EOPD and have refocused our efforts on copy number/structural variant efforts and on non-coding variation that may alter

gene expression patterns. Another more complex variation is expanded (or simply longer) repeat elements. We are applying specific algorithms to screen for repeat expansions and are assessing the variability across multiple callers (ExpansionHunter, gangSTR and STRetch). The hypothesis is that non-coding repeat expansions could knockout or drastically reduce expression of a given allele and this may be another process by which recessive loss-of-function of a gene could result in early-onset forms of PD. In Year 5 explored the potential of long-read (LR-) WGS data to supplement our short-read and inform into our potential CNV/SV/repeat hits. We have established protocols for high molecular weight DNA isolation and are currently awaiting the results of 25 EOPD patients with disease onset under 40 years of age.

At the beginning of the project we focused on single nucleotide variation and copy number changes across our series of people with early-onset Parkinson's disease (EOPD). In years 3 and 4 given the COVID-19 pandemic research had slowed but we are beginning to implement a further analysis tool to the samples and data and look at non-coding DNA and possible expanded repeat sequences as a cause for loss of expression and phenotypic presentation of disease. In addition, we have access to a whole-genome sequence (WGS) data that is available through a number of consortia efforts including the Accelerating Medicines Partnership (AMP) for Parkinson's (PD; AMP-PD) datasets and Lewy body dementia sequencing cohorts.

Single nucleotide variation in Known Parkinsonism Genes

Variants within the well-known and published parkinsonism genes (*SNCA, PINK1, PRKN, LRRK2, PARK7, VPS13C* and *GBA*) were filtered to identify individuals with variants in these genes. A total of 13,272 variants within those genes were present in the sequences (3'-UTR variants: 94, 5'-UTR variants: 12, intron variants: 13,070, missense variants: 60, splice-region variants: 7, synonymous: 27, stop gained: 1).

From the Mayo sequenced individuals, 37 carry coding or loss of function variants in one or more of the listed genes (Table 1). Two individuals carried homozygous common *PRKN* variants (PRKN p.Ser167Asn and PRKN p.Arg444Asn) and were not considered the pathogenic cause; one individual carries two novel heterozygous mutations in *VPS13C* (VPS13C p.Ala1687Val and VPS13C p.Leu216 Val); one individual carries the published PD risk variant LRRK2 p.Arg1628Pro, and five individuals carry the known PD risk factor GBA p.Leu483Pro and this can influence earlier onset of disease.

Other interesting findings include a patient who is heterozygous to novel variants in the previously published gene *SYNJ1* (p.Arg1328Pro and p.Val1405_Leu1406insAsnThr) who also carries a variant in *PARK7* (p.Val20Ala). In the field, there is a question that remains unanswered in that do heterozygous (known pathogenic) mutations of known recessive genetic forms of disease. A number of studies have reported higher frequencies of heterozygous mutations in patients for certain genes (e.g. *PINK1*) but other studies do not and suggest no influence from heterozygous carriers. This is of course a complex question as there maybe different effects from different types of mutations (e.g. missense, loss-of-function, deletion) and this may be different for different genes. For example, for PARKIN there are three potential pathogenic mechanisms; 1) loss of protein, 2) loss of translocation to mitochondria or 3) loss of ubiquitinase activity, and specific mutations may affect these differently. There is also the potential that we are missing the other pathogenic allele in some patients as it is not detected by current methods. We will try to address some of these questions with long-read whole genome sequencing approaches. For variant information see the table on next page for more details (Table1).

Table 1. Individuals carrying variants in previously published EOPD genes (highlighted ones probably have the disease explained)

Sample ID	Gender	Age of onset (years)	Race	Family History	Known PD genes variants
1487-1	F	40	White	Y	*GBA* p.Leu483Pro het
1622-1	F	35	White	Y	*SYNJ1* p.Arg1328Pro het SYNJ1 p.Val1405_Leu1406insAsnThr het *PARK7* p. Val20ala het
3021-1	M	33	White	Y	*GBA* p.Leu483Pro het
3021-10	F	50	White	Y	*GBA* p.Leu483Pro het
5541-1	F	34	White	Y	*GBA* p.Leu483Pro het
6634-1	M	37	White	N	*PINK1* p.Val184Met het *GBA* p.Asp448His het
6636-1	F	43	White		*PRKN* p.Met192Leu het
6743-1	F	44	White		*VPS13C* p.Ala1687Val het; *VPS13C* p.Leu216Val het
552-1	F	39	White	Y	*PINK1* p.Gln115Leu het
707-3	M	34	White	Y	*PINK1* p.Gln115Leu het
1200-1	F	40	White	Y	*GBA* p.Glu365Lys het; *GBA* p.Asp179His het
1256-1	M	43	White	N	*VPS13C* p.Arg470His het
1263-1	M	40	White	AD	*PRKN* p.Ser167Asn het *VPS13C* p.Ile2789Thr het; *VPS13C* p.Ile1132Val het; *VPS13C* p.Arg153His het
1308-1	M	41	White	Y	*GBA* p.Asn409Ser *PRKN* p.Ser167Asn homozygous
1560-1	F	37	White	Y	*GBA* p.Glu365Lys het
5167-1	F	37	White	N	*PINK1* p.Gln115Leu het
5274-1	F	37	White	Y	*GBA* p.Arg202Ter het *PINK1* p.Gly411Ser het
5311-1	M	36	White	N	*PRKN* p.Arg256Cys het
5400-1	M	34	White	Y	*VPS13C* splicing c.684+7T>C het
5536-1	F	40	White	N	*GBA* p.Ala483Arg het
5779-1	M	45	White	Y	*GBA* p.Ala495Pro het ; *GBA* p.483Pro het
5891-1	F	41	White	N	*GBA* p.Glu365Lys het
5980-1	M	36	White	Y	*PRKN* p.Arg256Cys het
6179-1	M	38	White	Y	*PRKN* p.Arg275Gln het
6375-1	M	45	White	Y	*GBA* p.Glu365Lys het
6560-1	F	36	White	Y	*LRRK2* p.Gly2019Ser het
7200-1	F	47	White	N	*PINK1* p.Gln115Leu het

Novel and rare single nucleotide loss-of-function variants

A total of 3,890 Loss of function (LoF) variants were detected, of which 694 are singletons and 1,461 are rare (<0.01 or absent from GnomAD database). Some of those LoF variants with a CADD score >36 are shown (>20 is predicted to be in the top 1% of deleterious variants) are listed on Table 2.

Table 2. Rare (<0.01 allele frequency in GnomAD) or singleton loss of function variants

ID	Sex	Onset (years)	Gene	aa change	CADD	Freq GnomAD
767-3	M	35	*BCAT2*	p.Tyr200Ter	37	<0.01
1487-1	F	40	*CNTN6*	p.Glu924Ter	49	
1651-2	F	43	*CPNE5*	p.Glu266Ter	45	
1487-1	F	40	*DOCK6*	p.Gln1518Ter	38	
1951-1	M	29	*FMO4*	p.Lys42Ter	38	<0.01
597-1	M	43	*IFIH1*	p.Glu627Ter	38	<0.01
1202-1	M	39	*MMP21*	p.Leu546Ter	39	
5136-1	F	34	*NMNAT3*	p.Trp184Ter	41	<0.01
5704-1	M	40	*PCDHB7*	p.Glu632Ter	42	<0.01
767-3	M	35	*PLSCR5*	p.Arg129Ter	42	<0.01
5948-1	M	41	*SPAG1*	p.Gln672Ter	38	<0.01
1487-1	F	40	*TAOK1*	p.Arg698Ter	39	
1951-1	M	29	*TMEM45A*	p.Arg36Ter	37	<0.01
1695-1	F	38	*TSFM*	p.Leu207Ter	39	<0.01
5619-1	F	37	*ULK4*	p.Arg862Ter	52	<0.01

Novel and rare single nucleotide non-synonymous variants

Lists of variants that were absent or have low allele frequency (<0.01) in the public database GnomAD and CADD scores >25 were generated for each individual. We attempted to identify potential detrimental rare homozygous variants or compound heterozygous. A total of 42,091 missense variants were detected and 5,675 variants have CADD scores > 25 and some examples of those variants are highlighted on Table 3.

Table 3. Examples of rare missense variants (allele freq < 0.01 or singletons) with CADD scores > 30

ID	Sex	Onset (y)	Gene	aa change	CADD	Freq GnomAD
5944-1	M	41	*CSMD3*	p.Pro1233Leu	33	na
6323-1	M	45	*FOXO6*	p.Lys190Thr	32	na
5167-1	F	37	*GMPR2*	p.Gly201Val	35	<0.01
552-1	F	39	*KCNT1*	p.Arg910Gln	35	<0.01
5676-1	M	41	*RPS27A*	p.Ser57Pro	33	<0.01
5255-1	M	40	*TOMM5*	p.Pro9Arg	33	<0.01
5980-1	M	36	*TRPM3*	p.Ser453Leu	33	<0.01

Targeted Pathway Analysis

Having identified and characterized single variant analysis in a genome-wide manner, we then extracted single variant data based on defined molecular pathways that are related to EOPD pathogenesis. We selected a number of key pathways including mitochondrial, lysosomal, and autophagic pathways. We annotated genes that were involved in these pathways and highlight a number of examples below.

Novel or rare single nucleotide variants in mitochondrial quality control genes

We extracted variants within the genes of the mitochondrial quality control pathways to feed into functional studies to assess pathogenicity and clinical relevance. A list with 1,346 genes with the GO term "mitophagy" was obtained from Genecards, and a total of 3,116 are missense and 171 are LoF (includes frameshift, splicing, stop gain and stop loss variants) variants in 713 of those genes were detected. A few examples of variants within genes that may be involved in the mitophagy pathway are summarized on Table 4.

Table 4. Rare variants (<0.01 in GnomAD) within genes that may participate on the mitophagy pathway (CADD scores > 20)

ID	Sex	Onset (yrs)	Gene	aa change	CADD	Freq GnomAD
501-1	M	37	*TRPM2*	p.Leu1034Phe	26.2	<0.01
1308-1	M	41	*TRPM2*	p.Gln953Ter	58	<0.01
6636-1	F	43	*TRPM2*	p.Arg411Trp	26.7	<0.01
1200-1	F	40	*TRAP1*	p.Arg47Ter	35	<0.01
5256-1	F	35	*SLC25A13*	p.Gln312Ter	35	<0.01
5834-1	M	48	*MFN2*	p.Thr362Met	26.4	<0.01
1776-1	F	43	*USP30*	p.Ser171Leu	26.3	<0.01
5430-1	F	43	*ATG13*	p.Asp490Asn	26.8	<0.01
5676-1	M	41	*CLEC16A*	p.Arg860Cys	25.9	<0.01
6179-1	M	38	*TOMM40*	p.Gly24Glu	23.1	<0.01
6026-1	F	42	*VPS13D*	p.Ala171Gly	24.6	<0.01
1064-1	M	39	*VPS13D*	p.Arg3267Trp	27.3	<0.01

Variants in the autophagy-related genes (e.g. *ATG9A)*

We extracted variants within the genes related to the autophagic pathway to feed into functional studies (with Dr. Springer) to assess pathogenicity and clinical relevance. A list of 959 genes with the GO term "autophagy" was obtained from Genecards, and a total of 2,150 are missense and 75 are LoF (includes frameshift, splicing, stop gain and stop loss variants) variants in 914 of those genes were detected. A an example we identified rare predicted damaging variants within the *ATG9A* gene that may be involved in dysfunction of the autophagy pathway and the variants within the gene are summarized on Table 5. The two rare variants that were only observed in EOPD patients are shaded in grey.

Table 5. Rare missense variants (<0.01 in GnomAD) within *ATG9A*, a gene that may participate in the autophagy pathway

											EOPD	LBD	Biobank
Chr	Position	Gene Names	Effect	Exon Number	HGVS c.	HGVS p.	SNP ID	gnomAD Euro Non-Finnish MAF	Minor Allele	Major Allele	Minor Allele Frequency	Minor Allele Frequency	Minor Allele Frequency
2	219220755	*ATG9A*	Missense	15	c.2506G>T	p.Val836Leu	rs371896436	6.49E-05	A	C			0.001
2	219220778	*ATG9A*	Missense	15	c.2483C>T	p.Ser828Leu	rs975693203	7.80E-06	A	G	0.002		
2	219220823	*ATG9A*	Missense	15	c.2438G>A	p.Gly813Glu	rs114840649	6.49E-05	T	C		0.001	
2	219221200	*ATG9A*	Missense	14	c.2248G>C	p.Ala750Pro	rs200858282	6.50E-05	G	C			0.001
2	219221214	*ATG9A*	Missense	14	c.2234A>G	p.Glu745Gly	rs753993322	6.50E-05	C	T		0.001	
2	219221253	*ATG9A*	Missense	14	c.2195G>A	p.Arg732His	rs201843079	0.00026	T	C		0.001	0.001
2	219221257	*ATG9A*	Missense	14	c.2191C>T	p.His731Tyr	rs779637846	0	A	G			0.001
2	219222126	*ATG9A*	Missense	13	c.2069A>G	p.Glu690Gly	N/A	N/A	C	T		0.001	
2	219222318	*ATG9A*	Missense	12	c.1981G>A	p.Gly661Arg	rs775852350	6.48E-05	T	C		0.001	
2	219222341	*ATG9A*	Missense	12	c.1958G>A	p.Arg653His	rs752441032	0.00013	T	C		0.001	
2	219222408	*ATG9A*	Missense	12	c.1891C>T	p.Arg631Trp	rs548518455	4.81E-05	A	G	0.002		
2	219222687	*ATG9A*	Missense	11	c.1806T>G	p.Asn602Lys	rs754362059	8.83E-06	C	A			0.001
2	219222695	*ATG9A*	Missense	11	c.1798C>T	p.Pro600Ser	rs373678234	0.00026	A	G		0.001	
2	219222719	*ATG9A*	Missense	11	c.1774A>G	p.Ser592Gly	rs2276635	0.07612	C	T	0.103	0.091	0.08
2	219222805	*ATG9A*	Missense	11	c.1688C>G	p.Thr563Ser	rs200002062	0.00045	C	G		0.001	
2	219223631	*ATG9A*	Missense	10	c.1553C>A	p.Thr518Asn	N/A	N/A	T	G			0.001
2	219223647	*ATG9A*	Missense	10	c.1537G>A	p.Val513Ile	rs199801466	0.00149	T	C		0.003	0.002
2	219223722	*ATG9A*	Missense	10	c.1462C>T	p.Leu488Phe	rs761747353	0.00013	A	G			0.001
2	219223972	*ATG9A*	Missense	9	c.1316G>A	p.Arg439His	rs199900351	0.00019	T	C		0.001	0.001
2	219224250	*ATG9A*	Missense	8	c.1121C>T	p.Thr374Ile	rs201218710	6.48E-05	A	G		0.002	0.001
2	219224361	*ATG9A*	Missense	8	c.1010G>A	p.Arg337His	rs766730899	0.00026	T	C		0.001	
2	219224494	*ATG9A*	Missense	8	c.877A>G	p.Ile293Val	rs773645781	0.00019	C	T		0.001	
2	219224772	*ATG9A*	Missense	8	c.599A>C	p.Lys200Thr	N/A	N/A	G	T			0.001
2	219225073	*ATG9A*	Missense	7	c.514A>G	p.Met172Val	rs772440768	5.30E-05	C	T		0.001	0.001
2	219225135	*ATG9A*	Missense	7	c.452T>C	p.Ile151Thr	N/A	N/A	G	A		0.001	
2	219225177	*ATG9A*	Missense	7	c.410T>C	p.Val137Ala	rs200773631	0.00084	G	A			0.001
2	219225186	*ATG9A*	Missense	7	c.401C>G	p.Thr134Ser	rs201763658	0.00013	C	G			0.001
2	219225445	*ATG9A*	Missense	6	c.340G>T	p.Asp114Tyr	rs188001238	8.83E-06	A	C			0.001
2	219227797	*ATG9A*	Missense	4	c.120T>G	p.Ile40Met	rs180953382	0.00078	C	A		0.001	0.001

Following on from our studies on coding variation we then wanted to address the possible role of non-coding variants an dhow they could potentially effect gene expression or be perhaps the missing variant in heterozygous recessive gene mutation carriers.

Non-coding variation in known EOPD genes

We also attempted to identify variants that may affect differential gene expression of mitophagy genes such as *PINK1* and *PRKN*. The variants were scored and filtered according to multiple resources. ENCODE was used to determine the genic regions that are predicted to harbor open chromatin in multiple different tissues. Intronic variants located within open chromatin regions are more likely to have a regulatory function in comparison to intronic variants found outside of open chromatin regions. In addition to the ENCODE regions, ATAC-seq data from mice dopaminergic neurons in forebrain (FB) and midbrain (MB) was used to predict regulatory regions specific to dopaminergic neurons by similarity. *PINK1*, for example in figure 1, is predicted to have 2 main regulatory regions in mice dopaminergic neurons, and 10 cis-regulatory regions according to ENCODE.

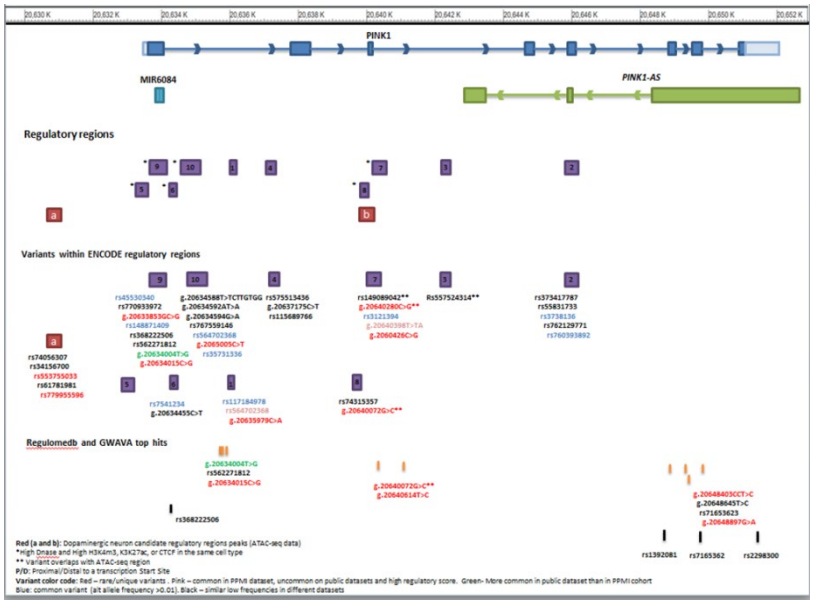
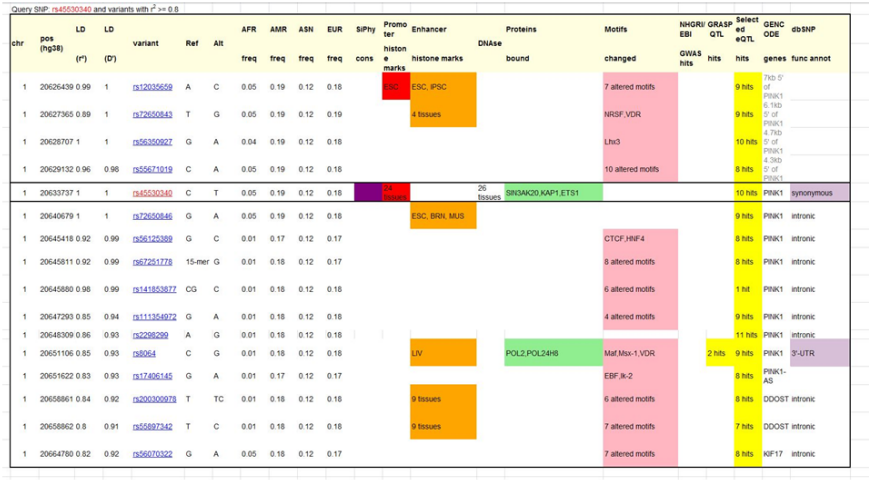


Figure 1. Regulatory regions and variation in the *PINK1*

The scoring algorithm CADD was used to identify variants with high conservation metrics and protein level scores. Variants with a CADD score higher than 20 are more likely to affect protein sequence and function. However, intronic and synonymous variants with CADD scores higher than 10 may indicate a possible regulatory role based on protein binding motif proximity. Regulomedb is a manually curated scoring algorithm that is based on eQTL evidence, chipseq, TF binding motifs and DNAase peaks. Regulomedb scores equal to 1a have many sources of regulatory evidence, whereas scores of 4 and higher are less likely to affects transcription factor binding or other known functional protein binding motifs. Unlike CADD, Regulomedb does not use sequence conservation as a criteria, so the scores may be discrepant and be carefully interpreted.

PINK1

- From the 48 "Regulatory" variants, 21 have HaploReg outputs



Example of Haploreg output

Figure 2. Regulatory variation in the *PINK1*

The GnomAD database was used to access the allele frequency of each variant across the general population.
Other resources such as Haploreg (see Figure 2), GWAVA, BRAINEAC, GTEx, ORegAnno, Enhancer Atlas, SYNAPSE, FANTOM and ConSite were used to identify known eQTLs and fine tune the top candidate variants.

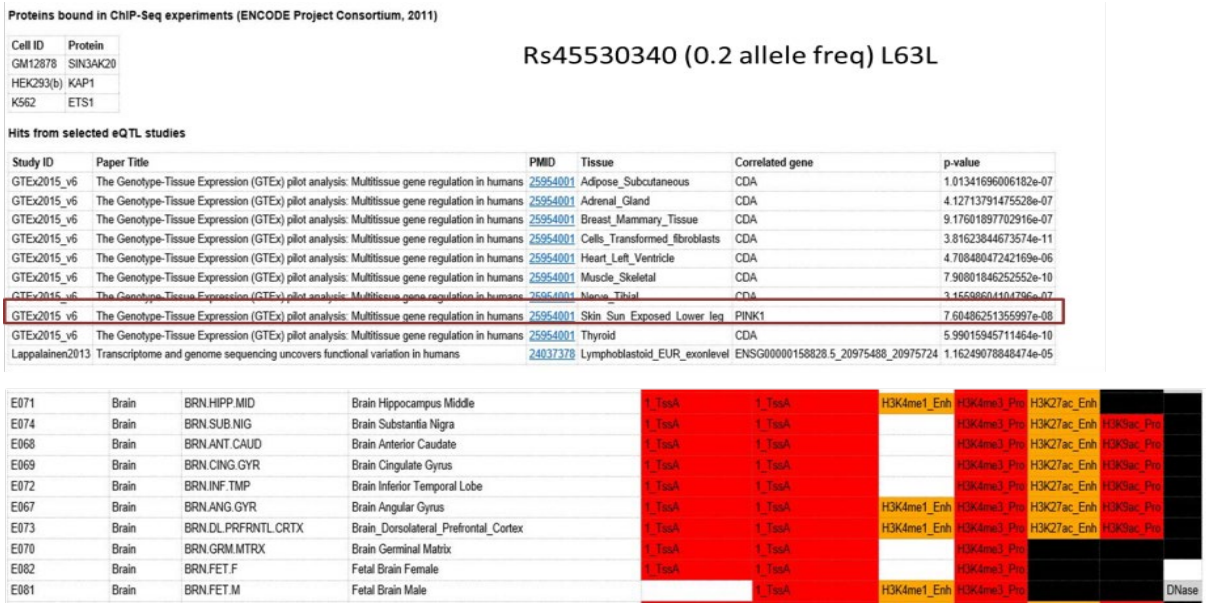


Figure 3. Tissue-specific regulatory variation in the *PINK1* gene

We have been using expression quantitative trait loci (eQTL) databases to help with the annotation across tissues. For example the *PINK1* variant, rs45530340 appears to regulate *PINK1* expression but only in sun exposed skin (Figure 3). We had identified this rare variant in a patient who harbored another single heterozygous *PINK1* mutation.

Another approach we have been using to use programs to predict functional regulatory motifs and this includes transcription factor binding, microRNA binding sites and distant gene expression enhancers. We are using cell-specific derived expression data to identify potentially functional non-coding variation that may act in a cell-specific manner. We have used enhancer atlas (http://www.enhanceratlas.org/) to predict cell-specific enhancers.

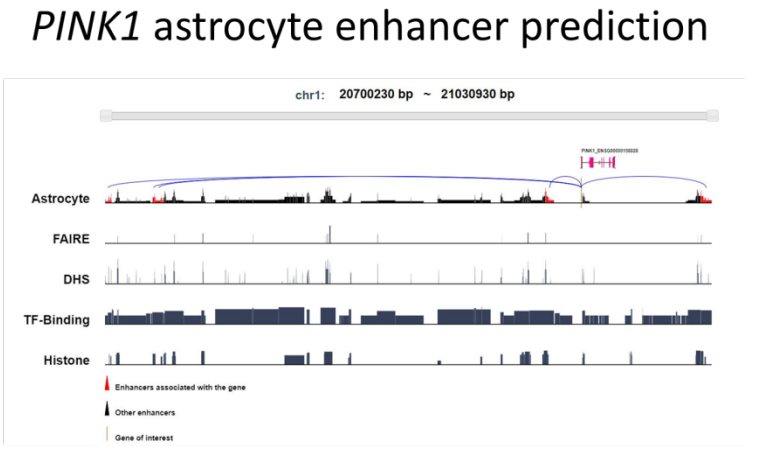


Figure 4. Distant expression enhancer for the *PINK1* gene

The importance of non-coding variation is becoming clearer but is more difficult to functionally test and given the cell-specific nature of the regulation of transcript expression has a higher degree of complexity. In addition to the non-coding

variation we are also performing complementary studies to try to resolve cell-specific transcript expression that may highlight coding variation that is only observed in a specific-cell type.

Investigating Copy Number Variation (CNV)

We have also prioritized assessing copy number variation (CNV) within the early-onset sequence data set and are working with our bioinformatics colleagues (Dr. Ren) to optimize the pipelines and adapt those to work with multiple programs. CNVs were called using the following CNV detection software: Lumpy, CNVnator, PatternCNV and Wandy and are currently being analyzed. Large structural variants and CNVs detection and annotation require a different and more complex approach than SNVs. There were no CNVs disrupting EOPD gene *PINK1*, and all of the rare CNVs disrupting the *PRKN* gene were successfully detected by all algorithms, with slight differences on breakpoint coordinates.

Each rare CNV event was visually inspected using IGV for validation. Efforts to verify the overlap of the common CNVs among the different software, and to calculate their allele frequencies are underway. This task is much more challenging because the breakpoint coordinates from the same CNV event can be variable from sample to sample in addition to algorithm differences. The CNV annotation database SCAN and the software package SVTtyper and intansv for R are being used to assist on the annotation and frequency calls of these inconsistent events.

Table 5 lists examples of CNV calls within genic regions that may disrupt function or increase gene dosage. All CNVs in Table 5 were called using Lumpy. The CNV calls from PatternCNV, CNVnator, Lumpy and Wandy are currently being analyzed.

Table 5. Example of CNVs detected in the Mayo EOPD cohort. CNVs with ID numbers have been previously detected and may be common.

CNV coordinates (GRCh37)	CNV size (bp)	CNV type	Genes	CNV ID
chr1:2886954-3845268	958,314	Deletion	*DFFB MEGF6 TP73 KIAA0562 ARHGEF16 WDR8 KIAA0495 LRRC47 PRDM16 TPRG1L ACTRT2 CCDC27 C1orf174*	NA
chr2:96823988-97075783	251,795	Deletion	*CNNM3 ANKRD39 FAM178B SEMA4C ANKRD23 LOC643445 LOC100128957 LOC100132375*	NA
chr3:14444330-14561629	117,299	Deletion	*GRIP2*	NA
chr6:157691300-164676192	6,984,892	Duplication/ Translocation	*ACAT2 IGF2R LPA MAS1 MAP3K4 PARK2 PLG SLC22A1 SLC22A3 SLC22A2 SOD2 TCP1 DYNLT1 EZR SYNJ2 QKI WTAP MRPL18 SNX9 AGPAT4 TULP4 TMEM181 ZDHHC14 RIPPLY2 LPAL2 RSPH3 FNDC1 SERAC1*	CNVR3123.1 CNVR3126_full CNVR3130.1 CNVR3135.1 CNVR3136_full CNVR3137.1 CNVR3141.1 CNVR3142.1 CNVR3143.1 CNVR3144.1 CNVR3150.1
chr7:72729346-72738350	9,004	Deletion	*DNAJC30*	NA
chr21:14338129-16526867	2,188,738	Duplication	*HSPA13 NRIP1 USP25 VDAC2P RBM11 POLR2CP SAMSN1 LIPI ABCC13 CYCSP42 LOC388813 C21orf126 LOC100128341*	CNVR7955.1 CNVR7956_full

We plan to identify further key variants that are involved in the development of early-onset PD and play a role in the dysfunction of the mitoQC pathway. These target genes/variants will be assessed for functional effects with the co-PI, Dr. Springer. We will identify other dataset that can add to and expand our sample size focusing on early-onset PD. We have already identified and obtained data from the Michael J. Fox Foundation PPMI cohort and will actively seek other data that may be informative (e.g. PDBP, UK biobank). Increasing the sample size will give us more power to genetically identify disease-relevant genes. We will use a cohort of 1000 whole-genome sequences that we have collected from the Mayo Clinic Biobank cohort. In addition, we have obtained whole-genome sequence data from the Parkinson's Progression Markers Initiative (PPMI) cohort supported by the Michael J. Fox Foundation and are analyzing data to use as a replication cohort for our sequence data. We have also run
CNV callers through the early-onset patients within the PPMI cohort to support our CNV studies. The figure shows the CNVs identified at the well-established EOPD locus, PRKN, across the PPMI patients and controls as well as the Mayo Clinic Biobank controls.

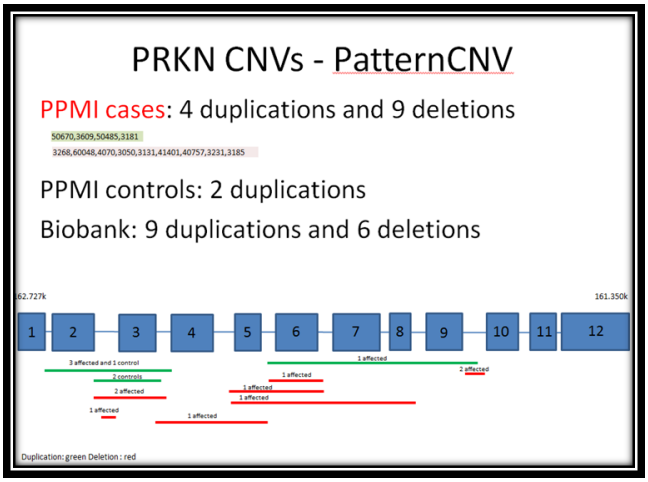


Figure 5. Identification of Copy number variation in the *PARKIN* gene

In 2020, we have applied for access to the recently released Accelerating Medicines Partnership (AMP) for Parkinson's (PD; AMP-PD) datasets that includes the PPMI series mentioned above. This will give us access to over 1800 whole-genome sequences from people with PD and no known genetic cause.

In addition, we have recently participated in the Lewy body dementia sequencing consortium which has given us access to over 1000 whole genome sequences from pathologically-confirmed Lewy body disease from cases collected at the Mayo Clinic brain bank. These resources will allow us to screen genetic variation identified in our early-onset cases to rapidly assess the role of variants/genes in the more frequent late-onset forms of Parkinson's disease and Lewy body disorders that we identify in our studies.

Given the discussion on whether early-onset forms of disease are mitochondrial-centered or if synuclein pathology is still key these additional datasets will allow us to potentially screen novel early-onset genes against related neuropathologies to assess protein aggregation and mitochondrial dysfunction.

Repeat sequence variation in people with early-onset parkinsonism

A number of repeat sequences have been identified to cause neurodegenerative disease for example Huntington disease, the C9orf72 expanded hexanucleotide repeat is the most frequent genetic cause of frontotemporal dementia and amyotrophic lateral sclerosis (and has been implicated in PD), and the group of disorders known as the spinocerebellar ataxias (SCA) are frequently caused by DNA repeat sequences. Indeed, the clinical presentation for a number of SCA genes includes parkinsonism. As part of repeat analysis we screened our EOPD patients for the most common SCA repeats that can results in parkinsonism (SCA1, SCA2 and SCA3) but we are now using this data as an internal validation to apply genome-wide programs that will identify expanded repeats from WGS data. We are testing two programs to begin with, the first is from Illumina termed ExpansionHunter (Dolzhenko et al., 2020) and the second is GangSTR (Mousavi

et al., 2019). We will filter our data against the same control series as we have used for the CNV analysis and the newly acquired AMP-PD cohorts.

This year we completed our whole-genome sequencing of EOPD patients to reach our target of having n=150. We are currently running analysis on the complete series and integrating additional EOPD data from publicly available datasets. For CNV analysis we have been developing criteria to define a true genetic variant across the multiple different algorithms e.g. for CNV/SV we have run; 1) WANDY, 2) Lumpy, 3) Delly, 4) CNVnator, and 5) MANTA. Each of these callers provides a different output and we believe each will have advantages and disadvantages for the different types of genomic variant, e.g. deletion, multiplication, inversion or translocation. We are running a number of programs that assess genomic repeats; 1) gangSTR, 2) hipSTR, 3) expansion hunter and plan to run 4) STRetch. Expanded repeats lengths are known to cause neurodegenerative disease and can be both small, e.g. SCA2 trinucleotide repeats of 33 can cause disease (with early-onset parkinsonism) and normal length can be 31 repeat units; or very large e.g. C9orf72 hexanucleotide repeats can be thousands of repeat units (normal <50 units). This analysis is on-going and requires a major bioinformatic efforts. These analyses will continue through the remainder of the award.

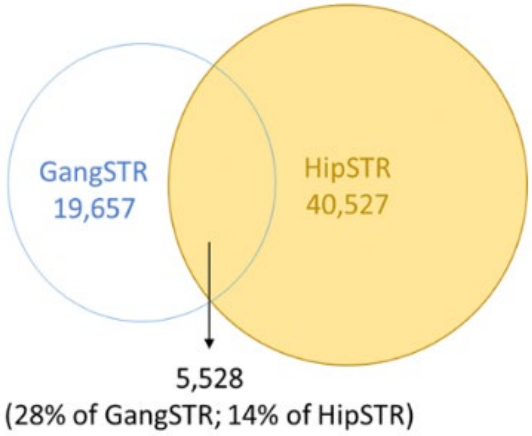


Figure 6. Overlap in calls from two repeat algorithms

Assessing the application of long-read WGS sequencing approaches

This year and in the final year of the grant we hope to assess the utility of long-read WGS approaches in EOPD. We have been working with Pac-Bio to optimize a protocol for LR-WGS with efforts focused on DNA extraction, shearing and fragment length sequencing. To attempt this we have developed a high molecular weight DNA extraction protocol from human brain tissue using New England Biolabs (NEB) Monarch HMW DNA Extraction kit. The protocol used an average of 15-18mg of frozen tissue and we homogenized each sample using a pestle to grind the tissue and a wide-bore tip to ensure a rapid and complete lysis. Samples were incubated in a thermal mixer at 56°C for 2.5 hour with agitation speed of 1500 rpm. The agitation speed is important in fragment size and it's noted on the NEB Monarch HMW DNA Extraction kit protocol that agitation speed between 1400-2000 rpm will produce DNA fragments between 50-250kb and

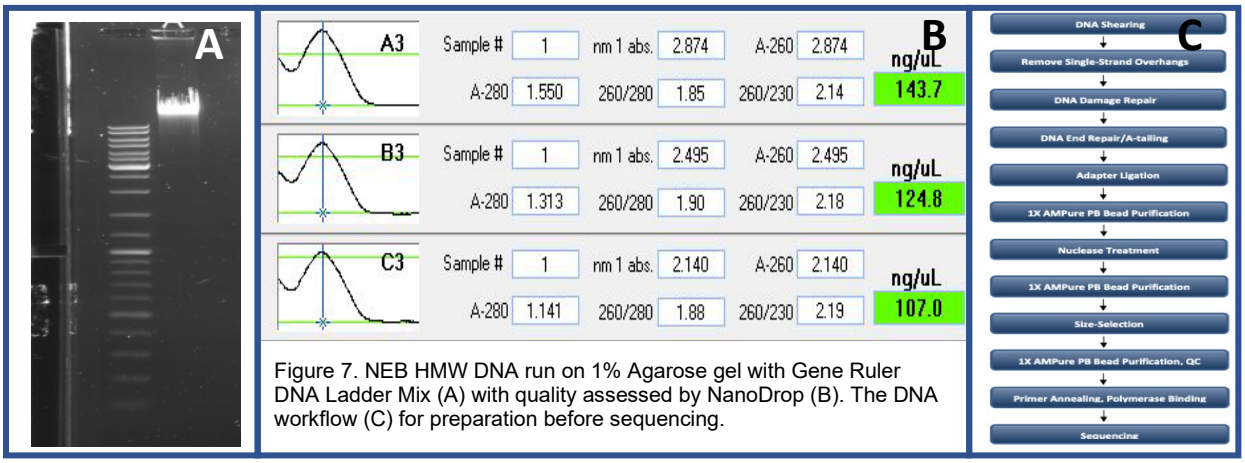


Figure 7. NEB HMW DNA run on 1% Agarose gel with Gene Ruler DNA Ladder Mix (A) with quality assessed by NanoDrop (B). The DNA workflow (C) for preparation before sequencing.

can range up to about 500kb. Th extraction method is gentle so as not to fragment the DNA into short segments.

The DNA quality and quantity were then visualized using the NanoDrop™ 8000 Spectrophotometer; the samples are also run on a 1% agarose gel in order to confirm the DNA was not fragmented during the prep (Figure 7).

Once we have obtained high quality DNA extraction, we optimized the shearing protocol to obtain fragment length of ~15KB (see Figure 8).

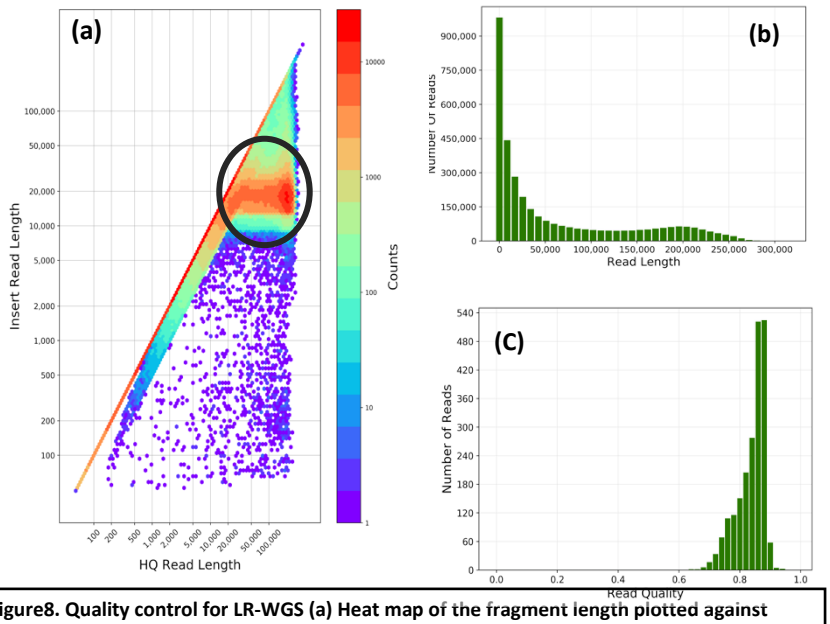


Figure8. Quality control for LR-WGS (a) Heat map of the fragment length plotted against quality read length and (b) number of reads plotted against average polymerase read length and (c) quality. These QC metrics demonstrate that we achieved our greatest CCS/HiFi reads

Given the recent advances in read output, subsequent cost-effectiveness and increased accuracy, the PacBio Sequel II platform provides a great opportunity to begin to characterize the overall architecture of the genome in EOPD. The key advantages of using PacBio SMRT Sequencing are that it provides read lengths in excess of 10 kb, has a high accuracy (>99.999% consensus accuracy) and no biases based on GC-content, producing uniform coverage. It creates high-fidelity (HiFi) reads that are produced by calling consensus from subreads generated by multiple passes of the enzyme around a circularized template.

We used the Sequel II platform to obtain at least 10-fold genome-wide coverage with one SMRT cell (8M) per sample; at this coverage most SVs are captured. This strategy enables the detection of SVs from tens to thousands of base pairs in length, the majority of which would have been missed using SR-WGS (80%). CNV/SVs are called using SMRT Link Structural Variant Calling software powered by pbsv (PacBio).

The new HiFi sequencing and v10.1 software update enables a more streamlined WGS workflow that is scalable and automation-friendly using as little as 5 ug of DNA. These improvements increase the number of HiFi reads at or above 99.9% accuracy for genome assembly and variant detection applications. LR-WGS will also allow us to better define repeat elements (see Figure 9 below) and assess the callers we are using on our short-read WGs (e.g. gangSTR, hipSTR), map in the breakpoints for CNVs/SVs to determine exactly which genes are involved, and phase variants due to the longer reads, which may be critical in understanding genetic influence in recessive forms of disease, such as observed in EOPD.

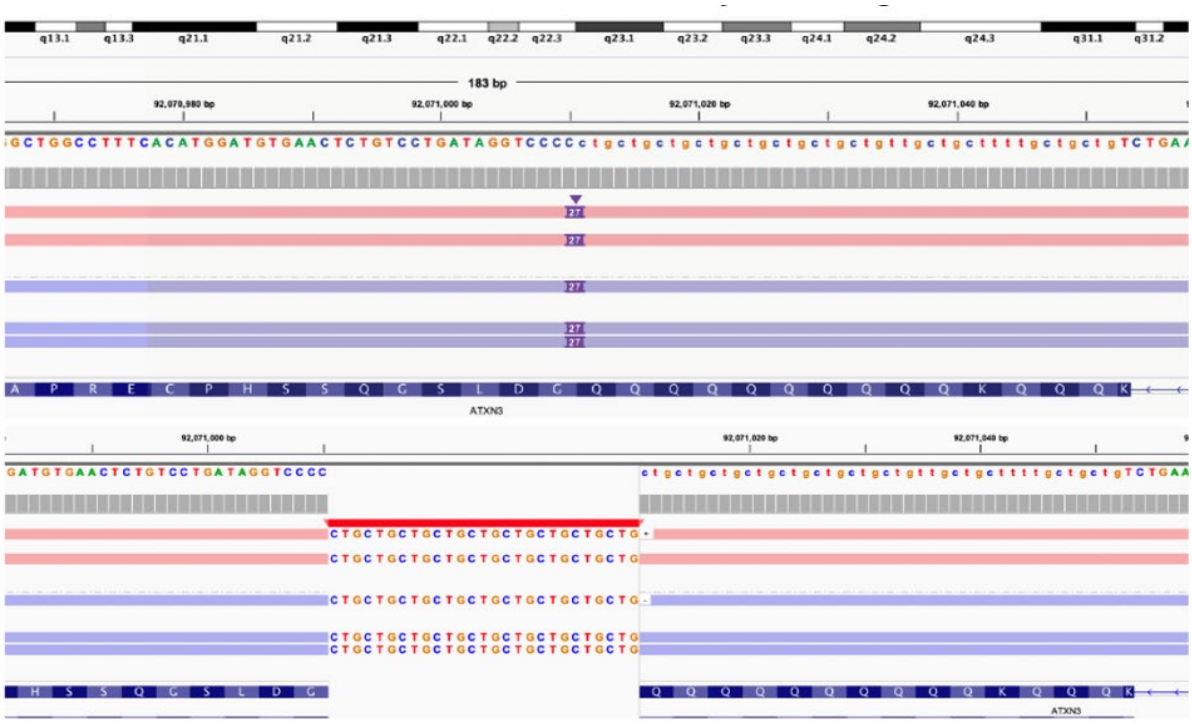
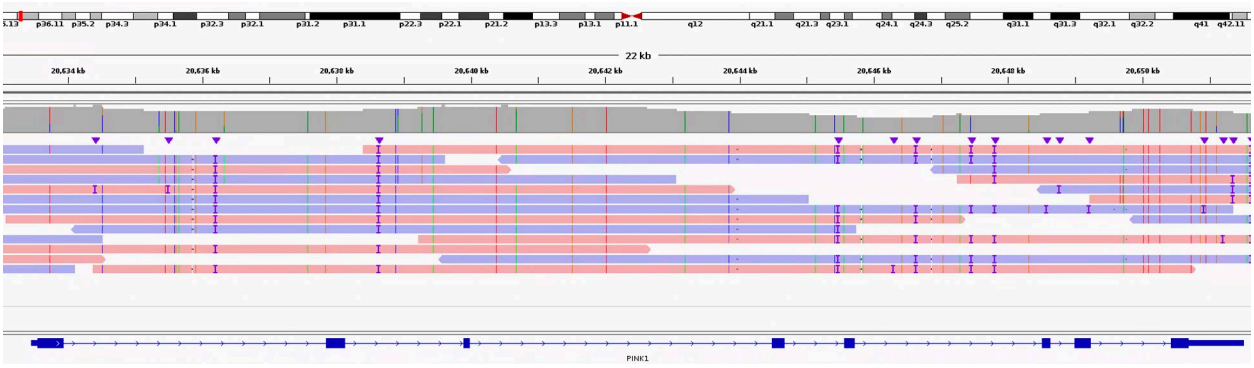


Figure 9. Detection and visualization in IGV viewer of the ataxin 3 repeat using LR-WGS and obtaining CCS/HiFi reads. This subject has a repeat length of 19 units, and that was visualized as a 27bp insertion of a pure CTG insert on the +ve strand (red) and the –ve strand (blue).

We performed LR-WGS on 25 early onset PD patients (age at onset <37 years) using PacBio HiFi sequencing technology and the Sequel II System. Each SMRT cell produces an average 10X coverage, therefore we used two SMRT cells per patient in order to achieve ~20X mean coverage. From the data that has completed thus far, the average coverage is 11X per SMRT cell with a minimum coverage of 9.74X and a maximum of 11.89X. SMRTLink will be used to generate HiFi reads by processing the raw data to generate HiFi fastq files using standard settings (minimum pass 3, minimum predicted RQ 20). Figure 10 below gives a visualization of the length of reads obtained and coverage across *PINK1*, with single reads capturing the entire locus.

Figure 10. Integrated Genome Viewer (IGV) of EOPD gene *PINK1*



Over the course of the project, we have witnessed a revolution in DNA sequencing approaches. AS noted above we have now moved towards integrating the latest long-read approaches with the generated short-read data and we foresee the data currently being generated as helping lay the foundation for future genetic studies in Early-Onset Parkinson's Disease.

- **What opportunities for training and professional development has the project provided?**
  Dr. Ross is now a member of the Movement Disorders Society Early Onset Parkinson's Disease Task Force.

- **How were the results disseminated to communities of interest?**
  Over all the years of the award we published papers relating to genetic factors that may affect early-onset forms of parkinsonism or altering the penetrance of variation and therefore the age-at-onset of clinical PD signs (Tipton et al. (2020) and Wernick et al. (2020)). Deutschlander et al. (2020) investigated genetic factors that may result in an earlier onset of disease. Four other related papers were also published on the role of mtDNA in PSP/CBD, non-MAPT related genes on chr17 haplotype in sporadic PD and overlap in genetic forms of ataxia and MSA (an early-onset parkinsonism disorder). Early in the first quarter of 2020 I also presented some of the EOPD data at the annual Mayo Clinic APDA PD research symposium for patients and at the 2020 Symposium on Alzheimer's Disease and Related Dementias, at the Alvin A. Dubin Alzheimer's Resource Center in Ft. Myers, Florida. In 2021 I presented on the genetics defining PD susceptibility at the XXVI World Congress on Parkinson's Disease and Related Disorders. The lab also published papers on EOPD genetics (Olszewska et al., 2021; Milanowski et al., 2021), and on EOPD as part of the community (Camerucci et al., 2021). We also published two related papers, one is the first large-scale WGS effort in the related disorder Lewy body dementia (Chia et al., 2021) and another identifying genetic modifiers in LRRK2-PD (which can be a cause of EOPD); (Lai et al., 2021). We also recently published a paper defining EOPD as part of the International Parkinson and Movement Disorder Society Task Force on Early Onset Parkinson's Disease (Mehanna et al., 2022) and reviewed the genetic landscape of EOPD (Kolicheski et al., 2022).

- **What do you plan to do during the next reporting period to accomplish the goals?**
  Final report

4. **IMPACT:**
5. **What was the impact on the development of the principal discipline(s) of the project?**
   Nothing to Report.
6. **What was the impact on other disciplines?**
   Nothing to Report.
7. **What was the impact on technology transfer?**
   Nothing to Report.
8. **What was the impact on society beyond science and technology?**
   Nothing to Report.

5. **CHANGES/PROBLEMS:**
   - **Changes in approach and reasons for change**
     Nothing to Report

   - **Actual or anticipated problems or delays and actions or plans to resolve them**

- **Changes that had a significant impact on expenditures**

The COVID-19 pandemic had shut the lab down or reduced productivity for much of Years 3 and 4 and we are again here slowly and carefully ramping back up following Mayo Clinic approved guidelines that include and promote social distancing, mask wearing and protection for staff. In addition, due to the COVID-19 pandemic the Genomics Core has also been closed and was not accepting samples for projects, this has slowed our data generation, plasticware availability has also been issue. Ana Kolicheski, PhD (Research Fellow) on the project has taken a position at Washington University (November 2019) and I will need to replace her with another fellow. I have an active search to find a suitable candidate, however there was a hire freeze at Mayo Clinic due to the COVID-19 pandemic which is preventing recruiting a new Research Fellow at this time. The summer of 2021 had been another difficult challenge again here in Florida with the resurgence of the delta variant. We had implemented a shift system (morning and afternoon) to minimize the number of people in the lab at any time and when work can be performed from home that option is promoted. However, we have now completed the project and all were expenses covered.

- **Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents**

Nothing to Report

6. **PRODUCTS:**

- Tsai PI, Lin CH, Hsieh CH, Papakyrikos AM, Kim MJ, Napolioni V, Schoor C, Couthouis J, Wu RM, Wszolek ZK, Winter D, Greicius MD, **Ross OA**, Wang X. PINK1 Phosphorylates MIC60/Mitofilin to Control Structural Plasticity of Mitochondrial Crista Junctions. Mol Cell. 2018 Mar 1; 69 (5):744-756.e6 PMID: 29456190
- **Ross OA**, Bu G. A novel link between trafficking and Lewy body disorders. Lancet Neurol 2018 Jul; 17 (7):571-573 PMID: 29887160
- McClymont SA, Hook PW, Soto AI, Reed X, Law WD, Kerans SJ, Waite EL, Briceno NJ, Thole JF, Heckman MG, Diehl NN, Wszolek ZK, Moore CD, Zhu H, Akiyama JA, Dickel DE, Visel A, Pennacchio LA, **Ross OA**, Beer MA, McCallion AS. Parkinson-Associated SNCA Enhancer Variants Revealed by Open Chromatin in Mouse Dopamine Neurons. Am J Hum Genet. 2018 Dec 6; 103 (6):874-892 PMID: 30503521 PMCID: 6288322
- Berge-Seidl V, Pihlstrom L, Wszolek ZK, **Ross OA**, Toft M. No evidence for DNM3 as genetic modifier of age at onset in idiopathic Parkinson's disease. Neurobiol Aging. 2019 Feb; 74:236.e1-236.e5 PMID: 30340792
- Ogaki K, Heckman MG, Koga S, Martens YA, Labbe C, Lorenzo-Betancor O, Walton RL, Soto AI, Vargas ER, Fujioka S, Uitti RJ, van Gerpen JA, Cheshire WP, Younkin SG, Wszolek ZK, Low PA, Singer W, Bu G, Dickson DW, **Ross OA**. Association study between multiple system atrophy and TREM2 p.R47H. Neurol Genet. 2018 Aug; 4 (4):e257 PMID: 30109269
- Muller-Nedebock AC, Brennan RR, Venter M, Pienaar IS, van der Westhuizen FH, Elson JL, **Ross OA**, Bardien S. The unresolved role of mitochondrial DNA in Parkinson's disease: An overview of published studies, their limitations, and future prospects. Neurochem Int. 2019 Oct; 129:104495 Epub 2019 June 21 PMID: 31233840 PMCID: 6702091
- Watzlawik JO, Hou X, Truban D, Fricova D, Ramnarine C, Barodia SK, Gendron TF, Heckman MG, DeTure M, Siuda J, Wszolek ZK, Scherzer CR, **Ross OA**, Bu G, Dickson DW, Goldberg MS, Fiesel FC, **Springer W**. Sensitive ELISA-based detection method for the mitophagy marker p-S65-Ub in human cells, autopsy brain, and blood samples. Autophagy. 2020 Oct 28; 1-16 [Epub ahead of print] PMID: 33112198 DOI: 10.1080/15548627.2020.1834712
- Soto-Beasley AI, Walton RL, Valentino RR, Hook PW, Labbe C, Heckman MG, Johnson PW, Goff LA, Uitti RJ, McLean PJ, **Springer W**, McCallion AS, Wszolek ZK, **Ross OA**. Screening non-MAPT genes of the Chr17q21 H1 haplotype in Parkinson's disease. Parkinsonism Relat Disord. 2020 Sep; 78:138-144 Epub 2020 Aug 01 PMID: 32829096 PMCID: 7686230 DOI: 10.1016/j.parkreldis.2020.07.022
- Deutschlander AB, Konno T, Soto-Beasley AI, Walton RL, van Gerpen JA, Uitti RJ, Heckman MG, Wszolek ZK, **Ross OA**. Association of MAPT subhaplotypes with clinical and demographic features in Parkinson's disease. Ann Clin Transl Neurol. 2020 Sep; 7 (9):1557-1563 Epub 2020 Aug 07 PMID: 32767721 PMCID: 7480915 DOI: 10.1002/acn3.51139
- Valentino RR, Tamvaka N, Heckman MG, Johnson PW, Soto-Beasley AI, Walton RL, Koga S, Uitti RJ, Wszolek ZK, Dickson DW, **Ross OA**. Associations of mitochondrial genomic variation with corticobasal degeneration, progressive supranuclear palsy, and neuropathological tau measures. Acta Neuropathol

Commun. 2020 Sep 17; 8 (1):162 Epub 2020 Sept 17 PMID: 32943110 PMCID: 7495714 DOI: 10.1186/s40478-020-01035-z

- Milanowski L, Oshinaike O, Broadway BJ, Lindemann JA, Soto-Beasley AI, Walton RL, Al-Shaikh R, Strongosky A, Fiesel FC, **Ross OA**, **Springer W**, Ogun SA, Wszolek ZK. Early-onset Parkinson disease screening in patients from Nigeria. Front Neurol. 2021 Jan 14;11:594927. doi: 10.3389/fneur.2020.594927.
- Heckman MG, Labbe C, Kolicheski A, Soto-Beasley AI, Walton RL, Valentino RR, Brennan E, Johnson PW, Baheti S, Sarangi V, Ren Y, Uitti RJ, Wszolek ZK, **Ross OA**. Fine-mapping of the non-coding variation driving the Caucasian LRRK2 GWAS signal in Parkinson's disease. Parkinsonism and Related Disorders 81 (2021) 200–204 PMID: 33454605 DOI: 10.1016/j.parkreldis.2020.12.016
- Valentino RR, Heckman MG, Johnson PW, Soto-Beasley AI, Walton RL, Koga S, Uitti RJ, Wszolek ZK, Dickson DW, **Ross OA**. Association of mitochondrial genomic background with risk of Multiple System Atrophy. Parkinsonism and Related Disorders 81 (2020) 200–204 PMID: 33189969 PMCID: PMC7769955
- Olszewska DA, McCarthy A, Soto-Beasley AI, Walton RL, **Ross OA**, Lynch T. PARKIN, PINK1, and DJ1 analysis in early-onset Parkinson's disease in Ireland. Ir J Med Sci. 2021 Mar 22. doi: 10.1007/s11845-021-02563-w.
- Camerucci E, Stang CD, Hajeb M, Turcano P, Mullan AF, Martin P, **Ross OA**, Bower JH, Mielke MM, Savica R. Early-Onset Parkinsonism and Early-Onset Parkinson's Disease: A Population-Based Study (2010-2015). J Parkinsons Dis. 2021 Mar 10. doi: 10.3233/JPD-202464..
- Milanowski ŁM, Lindemann JA, Hoffman-Zacharska D, Soto-Beasley AI, Barcikowska M, Boczarska-Jedynak M, Deutschlander A, Kłodowska G, Dulski J, Fedoryshyn L, Friedman A, Jamrozik Z, Janik P, Karpinsky K, Koziorowski D, Krygowska-Wajs A, Jasińska-Myga B, Opala G, Potulska-Chromik A, Pulyk A, Rektorova I, Sanotsky Y, Siuda J, Sławek J, Śmiłowska K, Szczechowski L, Rudzińska-Bar M, Walton RL, **Ross OA**, Wszolek ZK. Frequency of mutations in PRKN, PINK1, and DJ1 in Patients With Early-Onset Parkinson Disease from neighboring countries in Central Europe. Parkinsonism Relat Disord. 2021 May;86:48-51. PMID: 33845304
- Chia R, Sabir MS. **Ross OA**, Gibbs JR, Dalgard C, Traynor BJ, Scholz SW. Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. Nat Genet. 2021 Mar;53(3):294-303. PMID: 33589841
- Lai D, Alipanahi B, Fontanillas P, Schwantes-An TH, Aasly J, Alcalay RN, Beecham GW, Berg D, Bressman S, Brice A, Brockman K, Clark L, Cookson M, Das S, Van Deerlin V, Follett J, Farrer MJ, Trinh J, Gasser T, Goldwurm S, Gustavsson E, Klein C, Lang AE, Langston JW, Latourelle J, Lynch T, Marder K, Marras C, Martin ER, McLean CY, Mejia-Santana H, Molho E, Myers RH, Nuytemans K, Ozelius L, Payami H, Raymond D, Rogaeva E, Rogers MP, **Ross OA**, Samii A, Saunders-Pullman R, Schüle B, Schulte C, Scott WK, Tanner C, Tolosa E, Tomkins JE, Vilas D, Trojanowski JQ; 23andMe Research Team, Uitti R, Vance JM, Visanji NP, Wszolek ZK, Zabetian CP, Mirelman A, Giladi N, Orr Urtreger A, Cannon P, Fiske B, Foroud T. Genomewide Association Studies of LRRK2 Modifiers of Parkinson's Disease. Ann Neurol. 2021 Jul;90(1):76-88.PMID: 33938021
- Milanowski, L.M.; Hou, X.; Bredenberg, J.M.; Fiesel, F.C.; Cocker, L.T.; Soto-Beasley, A.I.; Walton, R.L.; Strongosky, A.J.; Faroqi, A.H.; Barcikowska, M.; Boczarska-Jedynak M.; Dulski, J.; Fedoryshyn, L.; Janik, P.; Potulska-Chromik, A.; Karpinsky, K.; Krygowska-Wajs, A.; Lynch, T.; Olszewska, D.A.; Opala, G.; Pulyk, A.; Rektorova, I.; Sanotsky, Y.; Siuda, J.; Widlak, M.; Slawek, J.; Rudzinska-Bar, M.; Uitti, R.; Figura, M.; Szlufik, S.; Rzonca-Niewczas, S.; Podgorska, E.; McLean, P.J.; Koziorowski, D.; **Ross, OA**.; Hoffman-Zacharska, D.; **Springer, W**.; Wszolek, Z.K. Cathepsin B p.Gly284Val Variant in Parkinson's Disease Pathogenesis. Int J Mol Sci 2022, 23, doi:10.3390/ijms23137086
- Broadway, B.J.; Boneski, P.K.; Bredenberg, J.M.; Kolicheski, A.; Hou, X.; Soto-Beasley, A.I.; **Ross, O.A**.; **Springer, W**.; Fiesel, F.C. Systematic Functional Analysis of PINK1 and PRKN Coding Variants. Cells 2022, 11, doi:10.3390/cells11152426
- Fiesel, F.C.; Fricova, D.; Hayes, C.S.; Coban, M.A.; Hudec, R.; Bredenberg, J.M.; Broadway, B.J.; Markham, B.N.; Yan, T.; Boneski, P.K.; Fiorino, G.; Watzlawik, J.O.; Hou, X.; McCarty, A.M.; Lewis-Tuffin, L.J.; Zhong, J.; Madden, B.J.; Ordureau, A.; An, H.; Puschmann, A.; Wszolek, Z.K.; **Ross, OA**.; Harper, J.W.; Caulfield, T.R.; **Springer, W**. Substitution of PINK1 Gly411 modulates substrate receptivity and turnover. Autophagy 2022, 1-22, doi:10.1080/15548627.2022.2151294
- Kolicheski, A.; Turcano, P.; Tamvaka, N.; McLean, P.J.; **Springer, W**.; Savica, R.; **Ross, OA**. Early-Onset Parkinson's Disease: Creating the Right Environment for a Genetic Disorder. J Parkinsons Dis 2022, doi:10.3233/JPD-223380
- Markham, B.N.; Ramnarine, C.; Songeum, K.; Bhagwate, A.; Osborne, A.; Wang, C. **Ross, OA**.; Fiesel, F.C.; **Springer, W**. The miRNA family miR-29 inhibits PINK1-PRKN dependent mitophagy via ATG9A. in preparation.

7. **PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS**
   - **What individuals have worked on the project?**

Name:	Owen A. Ross, PhD
Project Role:	PI
Researcher Identifier (e.g. ORCID ID):	0000-0003-4813-756X
Nearest person month worked:	2.0
Contribution to Project:	Together with the co-PI Dr. Springer, Dr. Ross has supervised the project, collected all regulatory material and ensured all necessary steps towards completion of the milestones
Funding Support:	
Name:	Alexandra Beasley, MSc
Project Role:	Laboratory Technician
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	1.0
Contribution to Project:	Ms. Beasley oversaw additional screening of the EOPD patients for atypical forms of early-onset parkinsonism and is establishing the new protocol for analysis of long-read whole genome sequencing.
Funding Support:	
Name:	Yingxue Ren, PhD
Project Role:	Bioinformatician
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	2.0
Contribution to Project:	Dr. Ren is working with Ms. Beasley in establishing the new protocol for analysis of long-read whole genome sequencing.
Funding Support:	
Name:	Ana Kolicheski, PhD
Project Role:	Research Fellow

Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	3.0
Contribution to Project:	Dr. Kolicheski has coordinated the analysis of the next-generation sequence data in consultation with biostatisticians and bioinformaticians.
Funding Support:	
Name:	Ronald Walton, BSc
Project Role:	Laboratory Technician
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	2.0
Contribution to Project:	Mr. Walton processed and organized all the samples from patients with early-onset Parkinson's disease. He completed the gene screening of PINK1 and PARKIN to identify and exclude carriers of known mutations prior to selection for whole-genome sequencing. He prepares and is responsible for shipping samples for sequencing, retrieval of data and facilitating bioinformatics processing.
Funding Support:	

- **Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**

  Owen A. Ross, PhD

  Changes in Active Support:

New

Title: Development and validation of a blood-based biomarker to identify overt and prodromal dementia with Lewy bodies
Grant number: Not applicable
Committed Time: 1.20
Supporting Agency: American Academy of Neurology Institute
Performance Period: 04/01/2022 –12/31/2024
Funding Amount:
Goals & Specific Aims: To identify an early and accurate blood-based assay to determine a diagnosis of Lewy body dementia and that can act as measurable response in drug trials.
Role: PD/PI

Title: Creating a Florida Cerebrovascular Disease Biorepository and Genomics Center
Grant number: 22K01
Committed Time: 1.80
Supporting Agency: James and Esther King Biomedical Research Program (FL State)
Performance Period: 04/01/2022 – 03/31/2025
Funding Amount:
Goals & Specific Aims: To create a biorepository of cerebrovascular patients across the state of Florida.
Role: PD/PI

Ended

Title: Early-Onset Parkinson's Disease Is a Mitochondrial Disease: A Nigral Mitochondrial Cytopathy
Grant number: W81XWH-17-1-0249
Committed Time: 2.40
Supporting Agency: Congressionally Directed Medical Research Programs (CDMRP)
Performance Period: 08/15/2017-**08/14/2022 (no cost extension)**
Funding Amount:
Goals & Specific Aims: The overall goal is to expand the PD-related PINK1/PARK2 pathway through a functional genomics approach through 1) Identification of genetic modifiers of mitoQC by an accelerated, two-tiered functional screen, 2) Whole-Genome sequencing (10X Genomics) in patients with early-onset PD, and 3) Determine pathogenicity of novel EOPD sequence variants in functional readout studies.
Role: PD/PI

- **What other organizations were involved as partners?**

  "Nothing to Report."

8. **SPECIAL REPORTING REQUIREMENTS**
   - **COLLABORATIVE AWARDS:**
     For an update on Major Task 1 see report from the co-PI Dr. Springer.
   - **QUAD CHARTS:**
     See appendix

9. **APPENDICES:** *IRB,Quad Chart.*

From: IRBe
To: Ross, Owen A., Ph.D.
Subject: PR17-008444-05 - Continuing Review Approved by IRB
Date: Wednesday, August 10, 2022 8:27:58 AM

Principal Investigator Notification:

From: Mayo Clinic IRB

To: Owen Ross

CC: Fabienne Fiesel
Owen Ross
Wolfdieter Springer
Zbigniew Wszolek

Re: **Continuing Review #:** PR17-008444-05
Title: Early-Onset Parkinson's Disease Is a Mitochondrial Disease: A Nigral Mitochondrial Cytopathy

IRB Approval Date: 8/10/2022
IRB Expiration Date: 8/9/2023

Continuation of the above referenced study is approved by expedited review procedures (45 CFR 46.110 and 32 CFR 219.110, category 5). This approval is valid for a period of one year. The Reviewer determined the research continues to pose no more than minimal risk to subjects. The Reviewer determined that this research continues to satisfy the requirements of 45 CFR 46.111 and 32 CFR 219.111.

AS THE PRINCIPAL INVESTIGATOR OF THIS PROJECT, YOU ARE RESPONSIBLE FOR THE FOLLOWING RELATING TO THIS STUDY.
1) When applicable, use only IRB approved materials which are located under the documents tab of the IRBe workspace. Materials include consent forms, HIPAA, questionnaires, contact letters, advertisements, etc.
2) Submission to the IRB of any modifications to approved research along with any supporting documents for review and approval prior to initiation of the changes.
3) Submission to the IRB of all Unanticipated Problems Involving Risks to Subjects or Others (UPIRTSO) and major protocol violations/deviations within 5 working days of becoming aware of the occurrence.
4) Compliance with applicable regulations for the protection of human subjects and with Mayo Clinic Institutional Policies.

Mayo Clinic Institutional Reviewer

Early-Onset Parkinson's Disease Is a Mitochondrial Disease: A Nigral Mitochondrial Cytopathy

PR160606P1
W81XWH-17-1-0249

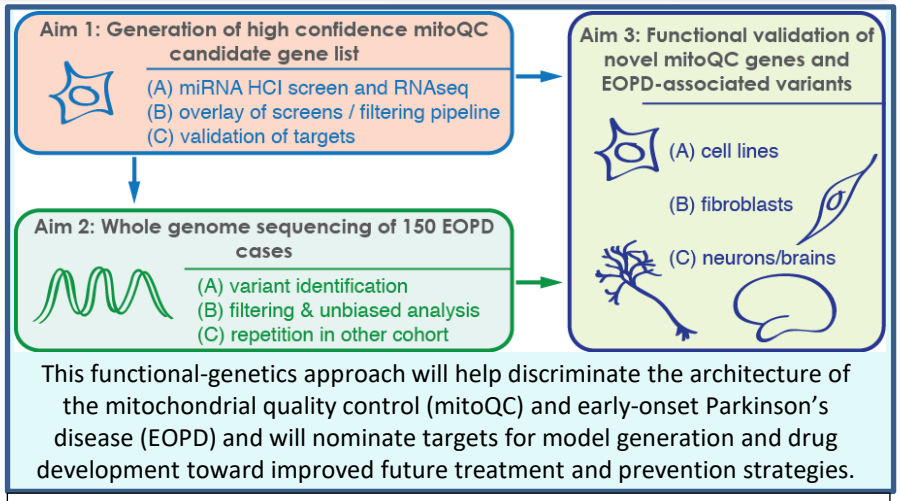
PI: Owen A. Ross, PhD **Org:** Mayo Clinic Jacksonville **Award Amount:** $1,150,197

Study/Product Aim(s)

- Specific Aim 1: To identify high-confidence genetic modifiers of PINK1/PARK2-directed mitochondrial quality control (mitoQC)
- Specific Aim 2: To identify the underlying genetic variation and characterize the early-onset Parkinson's disease (EOPD) genome
- Specific Aim 3: To determine pathogenicity of novel EOPD sequence variants in functional readout studies

Approach

We hypothesize that EOPD is a mitochondrial disease and that its genetic causes cluster around loss of mitoQC functions resulting in failure to safely dispose of damaged organelles. Our overarching goal is to delineate this pathway and the disease relevance of individual key players and their variants towards rationalized biomarker and drug development. This will be achieved through combining whole-genome-sequencing data from EOPD patients with functional genetic screening of genes/variants.



Accomplishment: Complete regulatory review and approval by Mayo Clinic IRB and by HRPO.

Timeline and Cost

Activities / CY	17	18	19	20
Aim 1: Functional screening				
Aim 2: WGS & analysis				
Aim 3: Validation & pathogenicity				
Estimated Budget ($K)	**$196**	**$392**	**$392**	**$196**

Updated: (12/14/2022)

Goals/Milestones

CY17 Goal – Regulatory review and roll-out of study
☑Complete regulatory review and approval by HRPO
☑Whole-genome sequencing of 50 EOPD patients
CY18 Goals – First-tier filtering of genomic variants
☑ Whole-genome sequencing of 100 EOPD patients
☑ Critical variant identification of mitoQC candidates
CY19 Goal – Second-tier filtering of genomic variants
☑Whole-genome sequencing of 150 EOPD patients
☑Critical variant identification with informed and unbiased strategies
CY20 Goal – Third-tier filtering and replication of candidate variants
☑Pathogenic variant identification in EOPD/mitoQC genes
☑Variant replication in additional cases of EOPD, LOPD & controls
☐**Comments/Challenges/Issues/Concerns**

- Dr. Kolicheski (Fellow) left for a position at WashU.
- Pandemic restricted lab and core work.

Budget Expenditure to Date
Projected Expenditure:
Actual Expenditure: $1,150,197