

Identifying Genes on 2p13 as Potential Genetic Factors of Parkinson's Disease and PARK3 Locus Candidate

Jingyu Cai*

The Hill School, 301 W 53rd St., New York, NY, 10019, USA.

* Corresponding author. Tel.: +1 (917) 654-7851; email: caijingyududu@gmail.com

Manuscript submitted October 22, 2017; accepted December 1, 2017.

doi: 10.17706/ijbbb.2018.8.4.245-251

Abstract: Parkinson's Disease(PD) is a common progressive neurodegenerative disorder caused by both genetic and environmental risk factors, at least 25 genetic loci and 20 genes have been identified in monogenic PD. In 1998, Thomas Gasser described a genetic locus on chromosomal location 2p13, called PARK3. The monogenic gene of PARK3 is still unknown. To explore the genetic cause of PD and potential candidates for the PARK3 locus, gene expressions and biological pathways were closely analyzed between 6 known PD-associated genes and the genes on 2p13. By excluding non-pathogenic variants from the current research, this study proposes 3 genes - *AAK1*, *EXOC6B*, and *SEMA4F* - should be extensively researched as potential genetic factors of PD and probable PARK3 locus, and more clinical data are necessary for predicting new genetic variants.

Key words: Biological pathway, PARK3, parkinson's disease, RPKM.

1. Background

Parkinson's Disease(PD) is the second most common neurodegenerative disorder. Widely affecting movement and motor functions, its clinical features include bradykinesia, tremors and rigidity in muscles, and postural instability [1]. These debilitating symptoms are known as parkinsonism as a whole. Patients are often observed with a progressive loss of dopamine and impairment of dopaminergic neurons in the substantia nigra - a basal ganglia structure in the mid brain that plays an important role in motor control. Although there are no current cures, medications such as L-dopa attempt to increase dopamine levels in the brain to help alleviate the symptoms and facilitate movement.

The etiology of Parkinson's Disease is still unknown, but it is believed to be caused by a combination of genetic and environmental factors. With the success of the Human's Genome Project in the last decade, researchers have identified "about 10 to 15 percent of all cases of Parkinson's are...to be genetic forms of the disease", while the "other 85 to 90 percent of cases are classified as idiopathic" [2].

In 1998, Thomas Gasser described a genetic locus on 2p13, called PARK3, with clinical features resembling those of sporadic PD, including a similar mean age of onset of 59.7 years [3]. Similar to six other PD genetic loci - PARK 1, PARK2, PARK5, PARK6, PARK7 and PARK8, the PARK3 locus is also monogenic. Many sequence analysis have been done since in attempt to locate the locus. The main goal of this study is to exclude the non-pathogenic variants, and identify the potential monogenic PARK3 locus on 2p13.

2. Methods

Six known genes whose mutations are associated with an increased chance of developing PD were used as the train set for this paper, three of these genes (and their respective loci) are *SNCA*(PARK1), *UCHL1*(PARK5), and *LRRK2*(PARK8) which express dominant inheritance, and the other three (and their loci) are *PRKN* (PARK2), *PINK1*(PARK6), and *DJ1*(PARK7) which express recessive inheritance [4].

The UCSC Genome Browser (<http://genome.ucsc.edu/>) mapped 84 genes on 2p13, these genes were the test set. We were particularly interested in their gene expressions in comparison to the 6 known genes associated with PD, for the more similar their expression, the more likely their protein would function in the same region, and the more likely their mutation may affect the same region.

The gene expressions of the 2p13 genes and PD-associated genes were examined using the online portal of the Genotype Tissue Expression (GTEx) project (<https://gtexportal.org/>). The data are shown in RPKM, or Reads Per Kilobase of transcript per Million mapped reads, which is a quantitative measure of gene expression by RNA sequencing (RNA-Seq), revealing the quantity of RNA and in turn its protein expression in a biological sample. The RPKM values of the genes in 31 target tissues were extracted.

The functional investigation of the genes, including their functions, mutations and relation to Parkinson's Disease, was primarily based on online resources such as *PubMed* and *Nature*.

3. Results

3.1. Mutation Patterns of the 6 PD-Associated Genes

The biological pathways of the 6 PD-associated genes during mutation show the known pathogenic mechanisms in disrupting the normal function of neurons in the brain, and affecting the regulation of dopamine. The genetic functions and mutations of the 3 autosomal dominant PD-associated genes are shown in Table 1, and the genetic functions and mutations of the 3 autosomal recessive PD-associated genes are shown in Table 2.

These known neuropathology observed in PD patients, such as Lewy bodies, ubiquitination dysfunction, mitochondrial dysfunction, synaptic disruption and oxidative stress may give insights to identifying potential genetic factors of the disease.

3.2. RPKM Heat Map of the 84 2p13 Genes and the 6 PD-Associated Genes

Fig. 1 shows the RPKM heat map of the 84 genes on 2p13 and the 6 PD-associated genes in 31 tissues. The "candidate genes" are the ones in the same cluster (indicated by the red node on the top left) as 5 of the 6 PD-associated genes (indicated by the red arrows on the bottom left).

These 11 "candidate genes" - *ADD2*, *AAK1*, *EXOC6B*, *FBXO41*, *SEMA4F*, *DCTN1*, *EGR4*, *VAX2*, *PAIP2B*, *SFXN5* and *BOLA3*, similar to *SNCA*, *PRKN*, *UCHL1*, *DJ1*, and *PINK1*, are highly expressed in many brain tissues. The mutations of these genes are more likely to directly affect the central nervous system, contribute to the degeneration of neurons and diminish motor control. It is not necessary to look at other genes with low expressions in the brain, because they are less likely to cause effects with low abundance in these regions.

3.3. Potential Genetic Factors of PD and the PARK3 Locus

Genes that did not correspond to PD or the PARK3 locus were excluded from the search. From the 11 "candidate genes" with similar expression to 5 of the 6 PD-associated genes, 8 genes were excluded, 3 genes - *AAK1*, *EXOC6B*, and *SEMA4F*, remained. Studies of these genes have not excluded them from the PARK3 locus search.

Table 3 shows the genetic function and mutation of the three remaining genes as probable genetic factors of PD and the PARK3 locus. Although using exclusion didn't indicate pathological confirmation, they did show promises. The three result candidates for the PARK3 locus - *AAK1*, *EXOC6B*, and *SEMA4F* - exhibit similar mutation patterns to the 6 PD-associated genes in Table 1 & Table 2: proteasome dysfunction,

α -synuclein buildup, Lewy bodies, molecular stress, synaptic vesicular disruptions, etc., all of these pathways may eventually cause dopamine deficiency, and likely result in PD.

Table 1. Genetic Function and Mutation of Three Autosomal Dominant PD-Associated Genes

Gene	Protein & Function	Mutation
SNCA	Alpha-synuclein: Found in presynaptic terminals, it maintains an adequate supply of synaptic vesicles, and regulates the release of dopamine. [5]	Single Nucleotide Polymorphisms lead to mis-folded or excess alpha-synuclein protein, which aggregate to form clumps called Lewy bodies. Present in many PD patients, it's believed that they "clog up" and impair neural functions, disrupting the regulation of dopamine. [5]
UCHL1	Ubiquitin C-terminal hydrolase: Its enzymes act in hydrolase activity which removes and recycles ubiquitin, sustaining the degradation process; ligase activity links ubiquitin for use in tagging proteins for disposal. [6]	The I93M mutation leads to decreased hydrolase activity, which may disrupt the ubiquitin-proteasome system. Instead of being degraded, unneeded proteins could accumulate to toxic levels that impair or kill nerve cells in the brain. [6]
LRRK2	Dardarin: Play a role in protein interactions, including transmitting signals and assembling cellular cytoskeleton. [7]	These mutations replace single amino acids in the Dardarin protein, which affects the protein's structure and function. It is unclear how <i>LRRK2</i> gene mutations lead to the movement and balance problems characteristic of Parkinson disease. [7]

Table 2. Genetic Function and Mutation of Three Autosomal Recessive PD-Associated Genes

Gene	Protein & Function	Mutation
PRKN	Parkin: Tag damaged/excess protein with ubiquitin, which then sends unneeded proteins into cell structures called proteasomes to be degraded. [8]	The loss of Parkin activity disturbs the ubiquitin-proteasome system, which allows unneeded proteins to accumulate. A buildup of these proteins could disrupt the supply and release of synaptic vesicles, particularly dopamine. [8]
PINK1	PTEN induced putative kinase 1: Found in the mitochondria of cells, it protects mitochondria during cellular stress. [9]	A disruption in PTEN-induced putative kinase 1 activity may cause malfunction in the mitochondria and an increase in cell stress. The loss of motor neurons due to insufficient energy weakens communication between the brain and muscles, leading to the loss of control of muscle movement. [9]
DJ1	DJ-1 protein: Protect brain cells from oxidative stress, fold proteins into tertiary structures, and deliver proteins to proteasome. [10]	Its mutations may lead to a toxic buildup of misfolded or damaged proteins and eventually to cell death from the oxidative stress. [10]

Table 3. Genetic Function and Mutation of the Three Potential Candidate for the PARK3 Locus

Gene	Protein & Function	Mutation
AAK1	AP2-associated protein kinase 1: A serine/threonine kinase that carries out clathrin-mediated vesicular transport. Its functions in clathrin-mediated vesicular transport and the serine/threonine domain are closely related to that of gene <i>GAK</i> [11].	Identified as a PD susceptibility, <i>GAK</i> 's depletion during mutation "impairs the lysosomal sorting of proteasome cathepsin D" [11], which serves as the lysosomal enzyme for degrading the α -synuclein protein. Disrupting its regulation may lead to a buildup of unneeded α -synuclein, leading to the formation of Lewy bodies.
EXOC6B	Exocyst complex component 6B: Encodes for part of a multi-protein complex necessary for exocytosis. This complex is highly expressed in the central nervous system and is crucial for cell growth, polarity and migration [12].	Variants of <i>EXOC6B</i> have been associated with cerebral disorders, such as intellectual disability [12]. A discovery in 2015 indicated molecular deficits in "synaptic exo- and endocytosis [and] vesicular trafficking" are "genetically linked to the phenotypic spectrum of parkinsonism associated with Lewy body pathology", which may contribute to neurodegeneration in the brain [13].
SEMA4F	Semaphorin-4F: A transmembrane semaphorin	Because of its synaptic plasticity, alterations in its signal

Gene	Protein & Function	Mutation
	found across glutamatergic synapses[14], where the synaptic plasticity of the neurotransmitter glutamate allows long-term potentiation(LTP) to take place.	transmission may “contribute to a number of neurological diseases, including depression, Parkinson's disease, [and] epilepsy” [15]. Mutant semaphorins may alter the physiological process of LTP, causing PD.

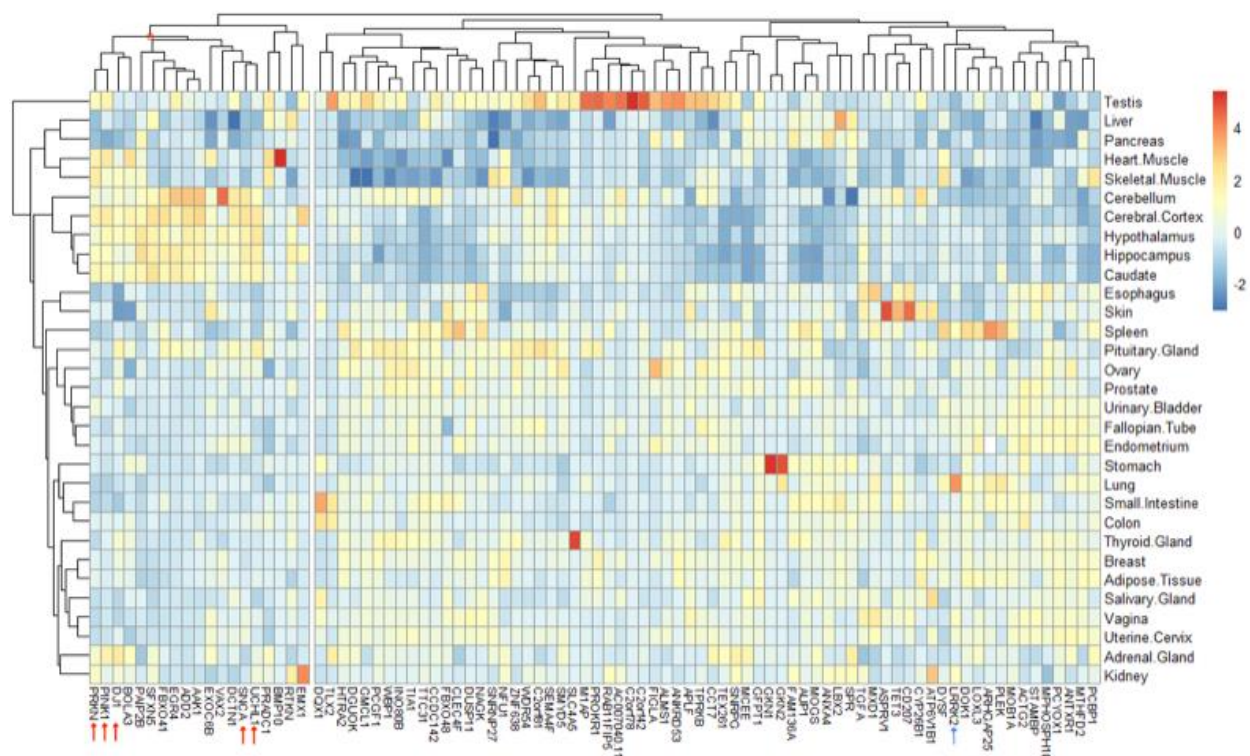


Fig. 1. The RPKM heat map of genes on 2p13 and the 6 PD-associated genes for 31 tissues.

4. Discussion

4.1. Exclusion of the 8 Genes

Genes *VAX2*, *PAIP2B*, *ADD2*, *FBXO41*, *DCTN1*, *EGR4*, *SFXN5*, and *BOLA3* were first excluded from the PARK3 candidates for this study, but it doesn't mean that they might not correspond to PD in future research. More clinical features and genetic data for these genes are needed for making a certain exclusion, the reasoning for this paper is primarily based on current online researches and databases. Studies have not suggested any variants in genes *VAX2*, *PAIP2B*, and *ADD2* may show correspondence to PD or parkinsonism. Although they are highly expressed in the brain, their functions may be more concentrated in other regions and neural pathways. Therefore these 3 genes were excluded.

5 more genes - *FBXO41*, *DCTN1*, *EGR4*, *SFXN5*, and *BOLA3* were also excluded. Shown in Table 4, while certain variants in some of these genes cause parkinsonism symptoms in other neurological disorders, they show no correspondence to the mutation patterns of PD; while others may show relation to PD, their mutation patterns have been excluded from the PARK3 locus.

4.2. The Exception of PD-Associated Gene *LRRK2*

Fig. 1 shows the cluster of 16 relatively related genes (11 on 2p13, 5 PD-associated) with high expressions in brain tissues. Another PD-associated gene *LRRK2* (as indicated by the blue arrow on the bottom right in Fig. 1) has relatively low expression in the brain, even though 6 of its mutations have been

demonstrated to be pathogenic to PD, including pathogenic features of the presence of Lewy bodies and ubiquitin dysfunction, and phenotypic conditions of progressive impairment in motor functions [22]. This suggests that more clinical data of the pathogenic mechanisms of PD are necessary for establishing not only intricate relationships of gene expressions, but also gene variants.

Table 4. Current Identification on the Correspondence Genetic Variants to *FBXO41*, *DCTN1*, *EGR4*, *SFXN5*, and *BOLA3* to PD and PARK3

Gene	Mutation & Correspondence
<i>FBXO41</i>	A study in 2013 of Chinese Han patients identified 5 coding variants rs151107579, rs526106, rs61733550, rs187317553, and rs34363861 of <i>FBXO41</i> were not of statistical significance to influence PD genetically. The variant rs61733550 was identified in a proband of a large PD family of autosomal dominant inheritance. However, further analysis of all members of the PD family showed this variant “did not co-segregate with disease”, suggesting it may not be a pathogenic variant. [16]
<i>DCTN1</i>	“DCTN1 mutations may contribute to disparate neurodegenerative diagnoses...but further correspondence to PARK3 is not specified.” [17]
<i>EGR4</i>	“...[<i>EGR4</i> was] sequenced and no potentially pathogenic mutations were detected, suggesting [<i>EGR4</i> does not] correspond to PARK3.” [18]
<i>SFXN5</i>	“Sequence analysis of 2p13 linked individuals affected with PD did not reveal any potentially pathogenic mutations within <i>SFXN5</i> , suggesting <i>SFXN5</i> does not correspond to PARK3.” [19]
<i>BOLA3</i>	Three mutations in <i>BOLA3</i> have been found to cause mitochondrial dysfunctions syndrome, a disease characterized by the impairment of mitochondrial functions, which may lead to the disruption of the Fe-S cluster, and affect energy production and glycine breakdown. Consequently, an excess glycine builds up in the brain, which may lead to glycine encephalopathy [20]. Also mapped on 2p13, the mutation of gene <i>HTRA2</i> leads to <i>PARK13</i> which is also characterized by encephalopathy and dystonia. However, <i>HTRA2</i> has been excluded from the finding of the PARK3 locus [21], suggesting that though <i>BOLA3</i> may show PD-related mutation patterns, it may not be a potential candidate for PARK3.

4.3. Prediction for Potential Genetic Factors of PD Needs a Larger Train Set

For this study, only 6 PD-associated genes were used as the train set for predicting potential genetic factors of PD and the PARK3 locus. The size of the this train set was not enough to assess and confirm the relationships with the test set - the 84 genes mapped on 2p13. Again, more clinical data and more pathogenic variants need to be identified for the validity of the prediction.

5. Conclusion

From the 11 “candidate genes”, this study has identified 3 genes - *AAK1*, *EXOC6B*, and *SEMA4F* as probable genetic factors of PD and candidates for the PARK3 locus, and more studies are necessary for showing their correspondence. This study highlights the potential for using biological pathways to identify new genetic factors, as genes displaying the similar pathways during mutation may contribute to similar diseases.

Acknowledgement

The author thanks his mentor for helping him with data analysis and paper revision.

References

- [1] Jankovic, J. (2008). Parkinson's disease: Clinical features and diagnosis. *Journal of Neurology, Neurosurgery, and Psychiatry*, 79(4), 368-376.
- [2] Believe in Better. Retrieved from the website: <http://www.parkinson.org/understanding-parkinsons/what-is-parkinsons/Genetics-and-Parkinsons->

Disease

- [3] Gasser, T., Müller-Myhsok, B., Wszolek, Z. K., Oehlmann, R., Calne, D. B., Bonifati, V., Horstmann, R. D. (1998). A susceptibility locus for Parkinsons disease maps to chromosome 2p13. *Nature Genetics*, 18(3), 262-265.
- [4] Learning About Parkinson's Disease. Retrieved from the website: <https://www.genome.gov/10001217/learning-about-parkinsons-disease/>
- [5] SNCA gene - Genetics Home Reference. (2017). Retrieved from the website: <https://ghr.nlm.nih.gov/gene/SNCA>
- [6] UCHL1 gene - Genetics Home Reference. (2017). Retrieved from the website: <https://ghr.nlm.nih.gov/gene/UCHL1>
- [7] LRRK2 gene - Genetics Home Reference. (2017). Retrieved from the website: <https://ghr.nlm.nih.gov/gene/LRRK2>
- [8] PARK2 gene - Genetics Home Reference. (2017). Retrieved from the website: <https://ghr.nlm.nih.gov/gene/PARK2>
- [9] PINK1 gene - Genetics Home Reference. (2017). Retrieved from the website: <https://ghr.nlm.nih.gov/gene/PINK1>
- [10] PARK7 gene - Genetics Home Reference. (2017). Retrieved from the website: <https://ghr.nlm.nih.gov/gene/PARK7>
- [11] Latourelle, J. C., Pankratz, N., Dumitriu, A., Wilk, J. B., Goldwurm, S., Pezzoli, G., Mariani, C. B., DeStefano, A. L., Halter, C., & Gusella, J. F. (2009). Genomewide association study for onset age in Parkinson disease. *BMC Medical Genetics*, 10(98).
- [12] Kloek, A. T., Setten, J. V., Ende, A. V., Bots, M. L., Asselbergs, F. W., Serón, M. V., & Ferwerda, B. (2016). Exome array analysis of susceptibility to pneumococcal meningitis. *Scientific Reports*, 6, 29351.
- [13] Vilariño-Güell, C., Rajput, A., Milnerwood, A. J., Shah, B., Szu-Tu, C., Trinh, J., Yu, I., & Farrer, M. J. (2014). DNAJC13 mutations in Parkinson disease. *Human Molecular Genetics*, 23(7), 1794-1801.
- [14] Armendáriz, B. G., Bribian, A., Pérez-Martínez, E., Martínez, A., Castro, F. D., Soriano, E., & Burgaya, F. (2012). Expression of Semaphorin 4F in neurons and brain oligodendrocytes and the regulation of oligodendrocyte precursor migration in the optic nerve. *Molecular and Cellular Neuroscience*, 49(1), 54-67.
- [15] Yazdani, U., & Terman, J. R. (2006). The Semaphorins. *Genome Biology*, 7(3), 211.
- [16] Liang, H., Song, Z., Lei, J., Tang, J., Xu, H., Deng, X., Qi, Y., Zheng, W., & Deng, H. (2013). Genetic analysis of the F-box only protein 41 gene in Chinese Han patients with Parkinson's disease. *Parkinsonism & Related Disorders*, 19(10), 918-920.
- [17] Araki, E., Tsuboi, Y., Daechsel, J., Milnerwood, A., Vilarino-Guell, C., Fujii, N., Mishima, T., Oka, T., Hara, H., Fukae, J., & Farrer, M. J. (2014). A novel DCTN1 mutation with late-onset parkinsonism and frontotemporal atrophy. *Movement Disorders*, 29(9), 1201-1204.
- [18] West, A. B., Zimprich, A., Lockhart, P. J., Farrer, M., Singleton, A., Holtom, B., Gasser, T. (2001). Refinement of the PARK3 locus on chromosome 2p13 and the analysis of 14 candidate genes. *European Journal of Human Genetics*, 9(9), 659-666.
- [19] Lockhart, P. J., Holtom, B., Lincoln, S., Hussey, J., Zimprich, A., Gasser, T., & Farrer, M. J. (2002). The human sideroflexin 5 (SFXN5) gene: Sequence, expression analysis and exclusion as a candidate for PARK3. *Gene*, 285(1-2), 229-237.
- [20] BOLA3 gene - Genetics Home Reference. (2017). Retrieved from the website: <https://ghr.nlm.nih.gov/gene/BOLA3>
- [21] Strauss, K. M., Martins, L. M., Plun-Favreau, H., Marx, F. P., Kautzmann, S., Berg, D., & Krüger, R. (2005).

Loss of function mutations in the gene encoding Omi/HtrA2 in Parkinson's disease. *Human Molecular Genetics*, 14(15), 2099-2111.

[22] Li, J., Tan, L., & Yu, J. (2014). The role of the LRRK2 gene in Parkinsonism. *BioMed Central*, 9, 47.



Jingyu Cai was born in Beijing, China on Jan. 12th, 2001. Currently he is a high school student at the Hill School, in Pottstown, PA, USA.