

Research Article

In Silico Analysis of SNPs in *PARK2* and *PINK1* Genes That Potentially Cause Autosomal Recessive Parkinson Disease

Yousuf Hasan Yousuf Bakhit,¹ Mohamed Osama Mirghani Ibrahim,² Mutaz Amin,³
Yousra Abdelazim Mirghani,³ and Mohamed Ahmed Salih Hassan⁴

¹Department of Basic Medical Sciences, Faculty of Dentistry, University of Khartoum, Khartoum, Sudan

²Department of Medical Biochemistry, Faculty of Medicine, Elrazi University, Khartoum, Sudan

³Department of Medical Biochemistry, Faculty of Medicine, University of Khartoum, Khartoum, Sudan

⁴Department of Biotechnology, Africa City of Technology, Khartoum, Sudan

Correspondence should be addressed to Yousuf Hasan Yousuf Bakhit; npgy1@yahoo.com

Received 6 October 2016; Accepted 24 November 2016

Academic Editor: Huixiao Hong

Copyright © 2016 Yousuf Hasan Yousuf Bakhit et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction. Parkinson's disease (PD) is a common neurodegenerative disorder. Mutations in *PINK1* are the second most common agents causing autosomal recessive, early onset PD. We aimed to identify the pathogenic SNPs in *PARK2* and *PINK1* using in silico prediction software and their effect on the structure, function, and regulation of the proteins. **Materials and Methods.** We carried out in silico prediction of structural effect of each SNP using different bioinformatics tools to predict substitution influence on protein structure and function. **Result.** Twenty-one SNPs in *PARK2* gene were found to affect transcription factor binding activity. 185 SNPs were found to affect splicing. Ten SNPs were found to affect the miRNA binding site. Two SNPs rs55961220 and rs56092260 affected the structure, function, and stability of Parkin protein. In *PINK1* gene only one SNP (rs7349186) was found to affect the structure, function, and stability of the PINK1 protein. Ten SNPs were found to affect the microRNA binding site. **Conclusion.** Better understanding of Parkinson's disease caused by mutations in *PARK2* and *PINK1* genes was achieved using in silico prediction. Further studies should be conducted with a special consideration of the ethnic diversity of the different populations.

1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder after Alzheimer's disease and was originally described by James Parkinson, in 1817, who discovered and described what he called "Essay on the Shaking Palsy." The cardinal signs of PD are resting tremor, bradykinesia, rigidity, and postural reflex impairment. Other manifestations include psychiatric symptoms such as anxiety and depression and dysautonomic symptoms such as hypotension and constipation, paresthesias, cramps, olfactory dysfunction, and seborrheic dermatitis. As the disease progresses, decreased cognitive ability may appear. On histopathology, PD is manifested as degeneration of the dopaminergic neurons of the pars compacta of the substantia nigra resulting in loss of dopamine in the striatum. This may occur only when 50–60% of the nigral neurons are lost and about 80–85%

of the dopamine content of the striatum is depleted [1]. Pathological features including progressive loss of striatal dopamine are those of bradykinesia and rigidity, which relate to striatal dopamine deficiency and loss of SNc dopaminergic neurons [1].

Studies on genetics of Parkinson's disease are worldwide and mutations are increasingly reported. The most frequent genes involved are *PINK1* and *PARK2* both of which are associated with autosomal recessive Parkinson's disease [2].

Linkage analysis of families with autosomal recessive juvenile Parkinsonism mapped the *PARK2* locus to chromosome 6q26, near the *sod2* locus. It encodes parkin, a 465-amino-acid protein that belongs to the "ring between ring fingers" (RBR) family of E3 ubiquitin ligases. The RBR domain interacts with ubiquitin-conjugating enzymes (E2s) to catalyze attachment of ubiquitin to protein targets, thus tagging these proteins for destruction by the proteasome [3, 4].

Reported mutations in *PARK2* exceed 100 including missense and nonsense mutations, as well as exonic deletions, rearrangements, and duplications [5, 6].

PARK6 was first mapped to chromosome 1p35-p36 in a large consanguineous Italian family with autosomal recessive, early onset PD. Subsequently, phosphatase and tensin homolog- (PTEN-) induced putative kinase 1 (*PINK1*) was determined to be the disease-causing gene. *PINK1* mutations have been reported to account for approximately 1% to 3% of early onset PD in populations of European ancestry [7, 8], 8.9% of autosomal recessive PD in a sample of Japanese families [9], and 2.5% of early onset PD in a sample of ethnic Chinese, Malays, and Indians.

The usage of in silico studies has strong impact on the identification of candidate SNPs since they are easy and less costly and can facilitate future genetic studies. The aim of this study was to identify the pathogenic SNPs in *PARK2* and *PINK1* using in silico prediction software and to determine the effect of these SNPs on the structure, function, and regulation of their respective proteins.

2. Materials and Methods

SNPs in *PARK2* and *PINK1* were obtained from the national center for biotechnology information (<https://www.ncbi.nlm.nih.gov/mapview/>).

2.1. Analysis of Single Nucleotide Polymorphism. We carried out in silico prediction of structural effect of each SNP using *PolyPhen* tool (pph2, Polymorphism Phenotyping v2, version 2.1.0) [10]. SNPs are also appraised quantitatively as benign, possibly damaging, and probably damaging. The program calculates position-specific independent count (PSIC) scores for every variant and estimates the differences between the variant scores. The software is found on the following link: <http://genetics.bwh.harvard.edu/pph2/>. Potentially damaging SNPs were further tested using *SIFT* [11] online software. It classifies SNPs into neutral and deleterious and it calculates a probability based on that prediction. A tolerance index of <0.05 was considered deleterious. *I-mutant* 2.0 [12] was used to predict the effect of coding SNPs on the stability of the protein. The resultant variants were visualized using UCSF Chimera 2.0 software using phylogenetic and structural information to predict possible impact of an amino acid substitution on the structure and function of a human protein.

We carried out further prediction of noncoding SNPs using two more online software types: Regulomedb [13] which integrates a large collection of regulatory information and an approach that enables the functional assignment of regulatory information onto any set of variants derived from genomic sequencing or GWAS studies. It functions through aligning SNPs with regulatory information and then comparing them to different types of transcription factor data set. The scoring system of it is based on increasing confidence that a variant lies in a functional location and likely results in a functional consequence. Noncoding SNPs were analyzed using Regulomedb, MIRSNP, and *SNP Function Prediction* [14] to determine their effects on transcription factor binding,

miRNA binding, and splicing database, which merge data set from computational, experimental, and epidemiological studies with genomewide association studies result and linkage disequilibrium information to prioritize SNPs for further genetic mapping studies. This software can predict SNPs with transcription factor binding sites (TFBS) activity; premature termination of amino acid sequence (stop codons); changing of splicing pattern or efficiency by disrupting splice site, exonic splicing enhancers (ESE), or silencers (ESS); alteration of protein structures or properties by changing single amino acids; or regulation of protein translation by affecting microRNA (miRNA) binding sites activity.

For SNPs within the 3'UTR region we used *MIRSNTs* [15] online software, which contains human SNPs in predicted miRNA-mRNA binding sites based on information from dbSNP135 and mirBASE18. It performs sequence alignment between 20 bp DNA sequences surrounding 3'UTR SNPs and the corresponding mRNA sequences. SNPs were classified into four groups, labeled as create, enhance, decrease, or break.

SNPs Stability Prediction

3D Modeling and Visualization. UCSF Chimera [16] is a program used to visualize the modeled PDP files and take images. It is used to visualize and analyze molecular structural data like supramolecular assemblies, sequence alignment, docking result, and conformational assemblies.

3. Result

PARK2. Twenty-nine thousand and nine hundred and sixty-four SNPs of *PARK2* gene were retrieved from dbSNP of which twenty-nine thousand, seven hundred and fifty-two SNPs (99.3%) were found intronic. 69 SNPs (0.2%) were found to be missense, twenty-three (0.08%) were synonymous, forty-seven (0.2%) were found at 3'UTR, and the remaining 120 (0.22%) were found as 5'UTR, upstream and downstream variants, frame shift variant, and also unknown variants.

One hundred and fifty-three SNPs were considered deleterious according to SIFT with score less than 0.05; the effects of the structural alteration were supported by PolyPhen in only two of them. All amino acid substitutions of these nsSNPs reduced the stability of Parkin protein as predicted by I-mutant 2.0 tool, Table 1.

The 3D structure of Parkin protein was modeled and visualized using Chimera 1.8 software, Figure 1.

Twenty-one SNPs were found to affect transcription factor binding activity (predicted by Regulomedb database) as in Table 2.

One hundred and eighty-five SNPs were found to be exonic splicing enhancer (ESE), exonic splicing silencer (ESS), or abolishing domains as listed below.

PINK1. One thousand four hundred and seventy-six SNPs were retrieved from dbSNP of which thirty-three SNPs were found deleterious according to SIFT with damaging scores of less than 0.05 of which one matching SNP was found to

TABLE 1: SNPs within PARK2 gene that affect the structure and function and stability of Parkin protein predicted by SIFT, PolyPhen, and I-mutant tools, respectively.

SNP ID	AA change	SIFT result	SIFT score	PolyPhen result	PolyPhen score	Protein stability	I-mutant score
rs55961220	C289G	Intolerable	0	Damaging	0.999	Decreased	9
rs56092260	R366W	Intolerable	0.001	Damaging	0.999	Decreased	5

TABLE 2: SNPs within PARK2 gene that affect transcription factor binding activity.

SNP	Score*
rs58457024	6
rs372637609	6
srs372626811	6
rs372664169	6
rs2358	6
rs372652246	6
rs372666477	6
rs372648356	6
rs372650102	6
rs372666184	6
rs25622	6
rs477943	6
rs480557	6
rs475158	6
rs75730600	6
rs372632202	5
rs372678887	5
rs372649620	5
rs372691334	5
rs372685162	5
rs372677454	4

* indicates evidence for a variant to be located in a functional region. 6: motive hit, 5: TF binding or DNase peak. 4: TF binding + DNase peak.

have structural level of alteration in PolyPhen. Amino acid substitution of this nsSNP reduced the stability of the protein as predicted by I-mutant 2.0 tool, Table 3.

The 3D structure of PINK1 protein was modeled and visualized using Chimera 1.8 software, Figure 2.

Forty-seven variants were found to affect transcription factor binding activity (predicted by Regulomedb database) as in Table 4 (see Supplementary Material available online at <http://dx.doi.org/10.1155/2016/9313746>).

No intronic SNPs were found to affect the splicing site.

Ten SNPs were found to affect the microRNA binding site using SNP Function Prediction website: rs10493377, rs11543262, rs12093960, rs1801792, rs41296144, rs41301082, rs41311152, rs9436293, rs9436732, and rs9436733.

4. Discussion

Analysis of in silico single nucleotide polymorphisms (SNPs) has become a very valuable tool recently in order to predict

variants most likely associated with disease which facilitates future association studies. This approach has been done for many disorders especially for cancer related genes [17–20]. In an effort to elucidate the identity of the pathogenic SNPs implicated in autosomal recessive nonsyndromic Parkinson's disease, we performed an in silico analysis of candidate SNPs in *PARK2* and *PINK1* gene and analyzed twenty-nine thousand, nine hundred and sixty-four SNPs in *PARK2* and one thousand four hundred and seventy-six SNPs in *PINK1*. In this study, two variants were predicted to be damaging in *PARK2*; the first (rs55961220) is A to C missense mutation which caused amino acid substitution C289G and the second (rs56092260) is G to A missense mutation that alters the amino acid R366W. Cysteine is sulfur containing hydrophilic amino acid while glycine is nonpolar hydrophobic one. Also arginine is a basic, hydrophobic amino acid containing amino group while tryptophan is nonpolar hydrophobic, ring containing amino acid. The later mutation was detected in two previous studies associated with a sporadic case of Parkinson's and the other with Juvenile Parkinson's disease. [21–23]

One hundred eighty-five alternative splice variants were revealed, which is much more than what was found by Scuderi et al. in 2014 who described 21 unique *PARK2* alternative splice variants through 26 different exon combinations [24]. Despite that majority of SNPs in *PARK2* gene are intronic, few are lying at exon/intron junctions and hence may cause splicing dysfunction.

On the other hand, one variant was found to be damaging in *PINK1*: a missense C to T mutation which alters the amino acid substitution P305L. Proline is an amino acid with a nonpolar ring usually responsible for the articulation within a protein while leucine is a nonpolar aliphatic one. This variant is not currently associated with any human disease up to date. Interestingly, although a 23 bp deletion of splicing regulation of exon 7 was found [25], no apparent variants with splicing regulation were found in our study. These results suggest that common variations in *PINK1* protein between individuals are rather quantitative which correlates with the spectrum of phenotypes seen in Parkinson's disease.

In brain samples of PD patients, miR-34b and miR-34c are significantly downregulated (almost by half) in the early stages of the disease [26] especially in the amygdala, substantia nigra, and cerebral cortex. On the other hand, deficiency of the same miRNAs results in apoptosis of differentiated neuroblastoma cells [27]. In our study variants that upregulate or downregulate the miRNA binding site in both genes might consequently alter the function of their respective proteins leading to the development of PD, a promising area to be further investigated for better understanding of its pathogenicity.

TABLE 3: SNPs within PINK1 gene that affect the structure and function and stability of PINK1 protein predicted by SIFT, PolyPhen, and I-mutant tools, respectively.

SNP ID	AA change	SIFT result	SIFT score	PolyPhen result	PolyPhen score	Stability	I-mutant score
rs7349186	P305L	Deleterious	0.004	Damaging	0.99	Decreased	7

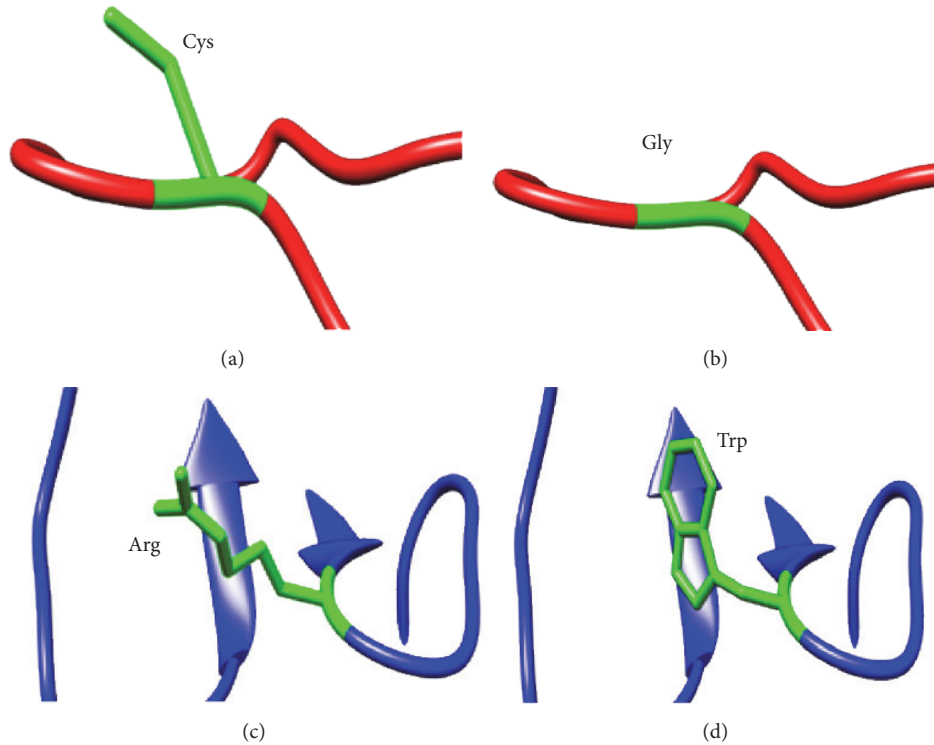


FIGURE 1: (a) Segment of reference Parkin protein highlighting the amino acid cysteine in position 239. (b) Mutant Parkin protein showing cysteine to glycine amino acid change in position 239. (c) Segment of native Parkin protein highlighting the amino acid arginine in position 366. (d) Mutant Parkin protein showing arginine to tryptophan in position 366.

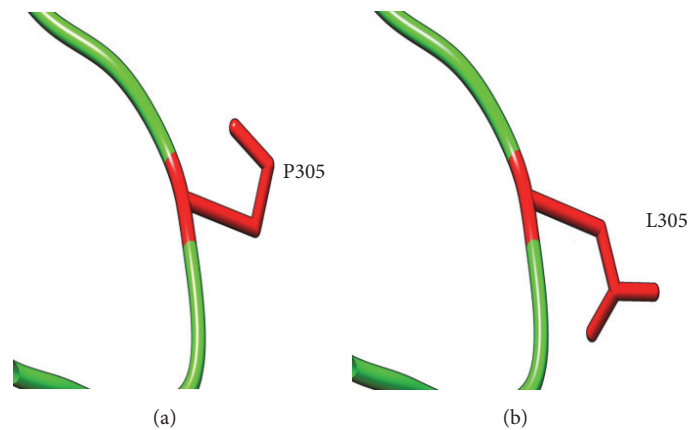


FIGURE 2: (a) The site of reference PTEN protein showing amino acid residue proline at position 305 which has changed to lysine due to point mutation, as shown in (b).

5. Conclusion

In silico analysis of SNPs in genes known to be associated with PD is a great aid for future candidate gene studies. All of these mutations constitute possible candidates for further genetic epidemiological studies with a special consideration of the large heterogeneity of PD among the different populations.

Competing Interests

The authors declare that they have no competing interests.

References

- [1] C. D. Marsden, D. J. Weatherall, J. G. G. Ledingham, and D. A. Warrell, *Movement Disorders. Oxford Textbook of Medicine*, vol. 3, Oxford University Press, New York, NY, USA, 1996.
- [2] S. Lesage and A. Brice, "Parkinson's disease: from monogenic forms to genetic susceptibility factors," *Human Molecular Genetics*, vol. 18, no. 1, pp. R48–R59, 2009.
- [3] M. Farrer, P. Chan, R. Chen et al., "Lewy bodies and parkinsonism in families with parkin mutations," *Annals of Neurology*, vol. 50, no. 3, pp. 293–300, 2001.
- [4] P. P. Pramstaller, M. G. Schlossmacher, T. S. Jacques et al., "Lewy body Parkinson's disease in a large pedigree with 77 Parkin mutation carriers," *Annals of Neurology*, vol. 58, no. 3, pp. 411–422, 2005.
- [5] N. Hattori, H. Matsumine, S. Asakawa et al., "Point mutations (Thr240Arg and Gln311Stop) [correction of Thr240Arg and Ala311Stop] in the Parkin gene," *Biochemical and Biophysical Research Communications*, vol. 249, no. 3, pp. 754–758, 1998.
- [6] C. Klein, K. Hedrich, C. Wellenbrock et al., "Frequency of parkin mutations in late-onset Parkinson's disease," *Annals of Neurology*, vol. 54, no. 3, pp. 415–416, 2003.
- [7] D. G. Healy, P. M. Abou-Sleiman, J. M. Gibson et al., "PINK1 (PARK6) associated Parkinson disease in Ireland," *Neurology*, vol. 63, no. 8, pp. 1486–1488, 2004.
- [8] V. Bonifati, C. F. Rohé, G. J. Breedveld et al., "Early-onset parkinsonism associated with PINK1 mutations: frequency, genotypes, and phenotypes," *Neurology*, vol. 65, no. 1, pp. 87–95, 2005.
- [9] Y. Li, H. Tomiyama, K. Sato et al., "Clinicogenetic study of PINK1 mutations in autosomal recessive early-onset parkinsonism," *Neurology*, vol. 64, no. 11, pp. 1955–1957, 2005.
- [10] I. A. Adzhubei, S. Schmidt, L. Peshkin et al., "A method and server for predicting damaging missense mutations," *Nature Methods*, vol. 7, no. 4, pp. 248–249, 2010.
- [11] P. C. Ng and S. Henikoff, "SIFT: predicting amino acid changes that affect protein function," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3812–3814, 2003.
- [12] E. Capriotti, P. Fariselli, and R. Casadio, "I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure," *Nucleic Acids Research*, vol. 33, no. 2, pp. W306–W310, 2005.
- [13] A. P. Boyle, E. L. Hong, M. Hariharan et al., "Annotation of functional variation in personal genomes using RegulomeDB," *Genome Research*, vol. 22, no. 9, pp. 1790–1797, 2012.
- [14] Z. Xu and J. A. Taylor, "SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies," *Nucleic Acids Research*, vol. 37, no. 2, pp. W600–W605, 2009.
- [15] C. Liu, F. Zhang, T. Li et al., "MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs," *BMC Genomics*, vol. 13, article 661, 2012.
- [16] E. F. Pettersen, T. D. Goddard, C. C. Huang et al., "UCSF Chimera—a visualization system for exploratory research and analysis," *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1605–1612, 2004.
- [17] R. Rajasekaran, C. G. Priya Doss, C. Sudandiradoss, K. Ramanathan, and R. Sethumadhavan, "In silico analysis of structural and functional consequences in p16INK4A by deleterious nsSNPs associated CDKN2A gene in malignant melanoma," *Biochimie*, vol. 90, no. 10, pp. 1523–1529, 2008.
- [18] R. Rajasekaran and R. Sethumadhavan, "Exploring the cause of drug resistance by the detrimental missense mutations in KIT receptor: computational approach," *Amino Acids*, vol. 39, no. 3, pp. 651–660, 2010.
- [19] R. Rajasekaran, C. Sudandiradoss, C. G. P. Doss, and R. Sethumadhavan, "Identification and in silico analysis of functional SNPs of the BRCA1 gene," *Genomics*, vol. 90, no. 4, pp. 447–452, 2007.
- [20] R. Rajasekaran, C. George Priya Doss, C. Sudandiradoss, K. Ramanathan, R. Purohit, and R. Sethumadhavan, "Effect of deleterious nsSNP on the HER2 receptor based on stability and binding affinity with herceptin: a computational approach," *Comptes Rendus—Biologies*, vol. 331, no. 6, pp. 409–417, 2008.
- [21] A. T. Portman, N. Giladi, K. L. Leenders et al., "The nigrostriatal dopaminergic system in familial early onset parkinsonism with parkin mutations," *Neurology*, vol. 56, no. 12, pp. 1759–1762, 2001.
- [22] C.-B. Lücking, V. Chesneau, E. Lohmann et al., "Coding polymorphisms in the parkin gene and susceptibility to Parkinson disease," *Archives of Neurology*, vol. 60, no. 9, pp. 1253–1256, 2003.
- [23] M. Wang, N. Hattori, H. Matsumine et al., "Polymorphism in the parkin gene in sporadic Parkinson's disease," *Annals of Neurology*, vol. 45, no. 5, pp. 655–658, 1999.
- [24] S. Scuderi, V. La Cognata, F. Drago, S. Cavallaro, V. D'Agata, and S. Lesage, "Alternative splicing generates different parkin protein isoforms: evidences in human, rat, and mouse brain," *Bio-Med Research International*, vol. 2014, Article ID 690796, 14 pages, 2014.
- [25] R. Marongiu, F. Brancati, A. Antonini et al., "Whole gene deletion and splicing mutations expand the PINK1 genotypic spectrum," *Human Mutation*, vol. 28, no. 1, article 98, 2007.
- [26] E. Miñones-Moyano, S. Porta, G. Escaramís et al., "MicroRNA profiling of Parkinson's disease brains identifies early downregulation of miR-34b/c which modulate mitochondrial function," *Human Molecular Genetics*, vol. 20, no. 15, pp. 3067–3078, 2011.
- [27] P. Shapshak, "Molecule of the month: miRNA and Parkinson's disease protein PARK2," *Bioinformatics*, vol. 9, no. 8, pp. 381–382, 2013.