All Sciences Proceedings

3rd International Conference on Innovative Academic Studies

September 26-28, 2023 : Konya, Turkey

ICIAS

All Sciences Proceedings
http://as-proceeding.com/

https://www.icias.net/

Comparative assay of LRRK2 gene mutations via bioinformatics approach

Ledia Vasjari[1*], Erisa Miçuli[1] and Gledjan Caka[2]

[1]*Department of Biology, University of Tirana, Albania*
[2]*Department of Biotechnology, University of Tirana, Albania*

[*]*(lediavasjari@gmail.com) Email of the corresponding author*

Abstract **–** Parkinson's disease (PD) is a neurodegenerative disease known to cause irreversible brain damage to afflicted individuals. It has been characterized as a chronic brain deterioration illness with an increased age-related risk factor. Being the second most common neural degenerative affliction after Alzheimer's its most prominent symptoms are known to be slow movement, tremors, loss of balance and muscle atrophy. This progressive disorder affects the neural cells embedded deep in the brain called basal ganglions and the substantia nigra. The substantia nigra has a key role in the production of the neurotransmitter dopamine of the human body responsible for the reward center and muscle movement. Since there is no known definitive diagnosis for Parkinson's, different approaches have identified a high number of mutations, especially in the LRRK2 gene (over 40%) in PD cases. This has made this gene a target for analysis. The mutations that affect this gene can create a series of damaging effects in the human body, ranging from abnormal signalling up to onset of Parkinson's disease and/or cancer. To assess these deleterious effects, we analysed 30-point mutations by predictive algorithms. Since our previous study and this one correlate on most effects and predictions, we can safely assume that the predictors we additionally chose were up to standard, up to date, with high sensitivity, specificity and accuracy.

Keywords – Parkinson\S Disease, LRRK2, Point Mutation, Prediction, Algorithm

I. INTRODUCTION

Parkinson's disease (PD) is a disease known to cause irreversible brain damage. It has an increased age-related risk factor. It is the second most common neurodegenerative disease after Alzheimer's and it is characterized by slow movement, shaking, balance problems and muscle rigidity [1]. This disease has a late onset, with individuals close to 60 years old, but nonetheless cases younger than 40 have also been identified.

This neurodegenerative chronic progressive disorder affects the neural cells embedded deep in the brain called basal ganglions and the substantia nigra. The substantia nigra plays an important role in the production of dopamine, a major neurotransmitter in the human body responsible for reward and motor movement. In the brain, dopamine hormone is attached to both the postsynaptic or presynaptic receptor or each of them individually. In cases when it binds to the presynaptic receptor it is then transmitted to be expressed in the postsynaptic neuron. The later is switched on and the signal is then propagated to the next neuron. When a patient is affected by PD, this normal transition is lacking, leading to an arrest of the dopamine signal to other cells [2]. There are no known definitive factors that can be attributed to the cause of PD cases, but studies have shown an increasing number of mutations affecting PD patients, occurring mostly in Leucine-Rich Repeat Kinase-2 (LRRK2) [3].

Up to date there is still a lack of definitive diagnosis for Parkinson's, however the high number of mutations revealed in the LRRK2 gene (over 40%) in PD cases [4], have made it a target for analysis.

Previous research from [5] have shown a pool of 30 mutations with possible deleterious effects, protein function and structural changes of the LRRK2 protein. In this study we try to elucidate these effects even further by adding additional predictive algorithms in order to assess the single nucleotide polymorphisms, potentially responsible for PD.

II. Materials and Method

To obtain the necessary sequences of the LRRK2 gene, which is believed to be responsible for Parkinson's disease, the National Center for Biotechnology Information (NCBI) databases were used [6].

Table 1: Protein stability comparison between IMutant and iStable

Mutation	iMutant	iStable
V21R	DEC	D
A211V	IN	IN
H230R	DEC	IN
A397T	DEC	D
G472R	DEC	D
Q501W	DEC	IN
L550W	DEC	D
N551K	DEC	D
T1343V	DEC	IN
K1347A	IN	IN
T1348N	DEC	D
R1398H	DEC	D
R1398L	DEC	IN
K1423E	DEC	D
N1437H	DEC	D
A1440P	IN	D
R1441C	DEC	D
R1441G	DEC	D
R1441H	DEC	D
R1628P	DEC	D
Y1699C	DEC	D
S1761R	DEC	IN
K1906M	DEC	IN
D1994A	DEC	D
D1994N	DEC	D
D1994S	DEC	D
G2019S	DEC	D
I2020T	DEC	D
Y2064K	DEC	D
G2385R	DEC	D

In order to do a comparison with our previous work to understand the effect of deleterious mutations affecting the gene we used:

The deleterious SNP mutations that affect the LRRK2 gene were extracted for the NCBI Variation Viewer [7].

The BLAST job for human LRRK2 gene regarding its protein sequence was gained in FASTA format from Uniprot [8], the biggest database for the known sequence and function of proteins.

FATHMM (Functional Analysis through Hidden Markov Models), a computer algorıthm able to predict effects in the function of coding variants [9] MutPred2 which is a method to determine via probability the effects of pathogenicity of amino acids rearrangement [10].

Table 2: Deleterious effects of mutations of the LRRK2 gene

Mutation	Mutpred		Fathmm	
V21R	disease	0.6	neutral	0.85
A211V	neutral	0.3	neutral	1.36
H230R	neutral	0.3	neutral	1.44
A397T	neutral	0.4	neutral	1.26
G472R	disease	0.51	neutral	1.33
Q501W	disease	0.8	neutral	1.18
L550W	disease	0.8	neutral	0.32
N551K	disease	0.6	neutral	0.35
T1343V	neutral	0.3	neutral	-0.26
K1347A	disease	0.66	disease	-4.08
T1348N	neutral	0.4	disease	-2.33
R1398H	neutral	0.24	neutral	-1.24
R1398L	neutral	0.5	neutral	-1.21
K1423K	disease	0.6	neutral	-1.3
N1437H	neutral	0.5	neutral	-0.24
A1440P	disease	0.8	neutral	-0.25
R1441C	neutral	0.5	neutral	-1.38
R1441G	disease	0.6	neutral	-1.37
R1441H	neutral	0.3	neutral	-1.3
R1628P	disease	0.8	neutral	-0.68
Y1699C	disease	0.82	neutral	-1.06
S1761R	neutral	0.4	neutral	-0.67
K1906M	disease	0.6	disease	-5.41
D1994A	disease	0.8	disease	-3.15
D1994N	disease	0.7	disease	-3.14
D1994S	disease	0.8	disease	-3.15
G2019S	disease	0.8	disease	-6.18
I2020T	disease	0.7	disease	-3.26
Y2064K	disease	0.9	disease	-3.45
G2385R	disease	0.6	neutral	1.38

One key aspect affecting the normal function and structure of proteins is their stability. Deleterious effects that affect said protein have a tendency to alter this feature. İn order to assess and then compare this aspect we used I-Mutant2.0. this software is used to predict alterations of the stability of the protein affected by mutations [11]

Another predicting tool to assess a multitude of protein characteristics was used as well. We used NetSurfP3.0 to evaluate the secondary structure of the protein, the changes in the dihedral angles, structural disorders and solvent accessibility of the corresponding amino acids in the protein sequence [12].

All the aforementioned effects lead to changes and/or damage to the activity and metabolism of the cell.

III. Results and Discussion

The LRRK2 gene is a very important gene in the human body, responsible for coding the dardarin protein and also helps to coordinate and adjust problems in the endomembrane. The multiple connections of this gene with other surrounding proteins helps to maintain, develop and proliferate brain cells. The mutations that affect this gene can create a series of deleterious effects in the human body, from aberrant signalling up to onset of Parkinson's disease and/or cancer. To evaluate these damages, we analysed these mutations by predictive algorithms:

30-point mutations were considered and then assessed by different computational algorithms. Our data (table 2) show that according to FATHMM, 21 out of 30 mutations were considered as neutral or as not having a deleterious effect on the protein. Only 9 were considered as deleterious (K1347A, T1348N, K1906M, D1994A, D1994N, D1994S, G2019S, I2020T, Y2064K). The cut-off value for the mutations to be ascertained as damaging must be >0.5.
Mutpred2 algorithm showed that 19 mutations were predicted as damaging (V21R, G472R, Q501W, L550W, N551K, K1347A, K1423K, A1440P, R1441G, R1628P, Y1699C, K1906M, D1994A/N/S, G2019S, I2020T, Y2064K, G2385R). This software uses the same cut-off value as FATHMM (>0.5). when compared to each other only 8 mutations were assessed as damaging by both programs (K1347A, K1906M, D1994A/N/S, G2019S, I2020T and Y2064K). When compared to the data from [5] it is shown that only I2020T is considered as damaging by all previous and current predictions. The same goes for A221V that is the only mutation considered as neutral/benign by all predictors with relatively high scores. The G2385R mutation is also to be viewed closely because with the exception of MutPred that states it as damaging, the other tools regard it

as neutral or without an effect. The reason this mutation is important to mention is the fact that this mutation has been encountered and associated with late onset PD in Asian populations [13]. Regarding the other mutations, in order to analyse and come to the conclusion which of them must be further evaluated, the rule of the majority is applied. İf a higher number of logarithms showed that the effect is either deleterious and/or neutral, that is the final result.

Another aspect to take into account is the proper folding of the protein in order to make it specific and to have a proper function [3]. To analyse this important feature which leads to understanding if the protein is functioning correctly or not we used I-Mutant2.0 and compared the results with our previous study [5]. The parameters were kept the same for both programs: sequence predictions with a temperature of 36 Celsius degrees (to simulate body temperature) and a pH of 7. The predictions shown in table 1 resulted in 23 mutations with the same results from both algorithm tools and just 7 (23%) that differ. Out of these seven, only 1 showed a decrease in stability compared to iStable, whereas the other 6 showed an increase in stability. This could be attributed to either the specificity of iStable and I-Mutant prediction (0.942>0.909) or it's accuracy (0.914>0.809) respectively. Compared to iStable, I-Mutant showed that only 9 of the ten most risk factor amino acids substitution (N1437H, R1441C/G/H, R1628P, Y1699C, G2019S, I2020T and G2385R) lead to a decrease in protein stability. The last one, S1761R, differed and showed an increase in protein stability. This could be attributed to a lower sensitivity of I-Mutant regarding this residue.

The last analyses performed was to assess the protein structure and whether its amino acids residues were in an exposed or buried state. Their location shows the role the residue has in cell, be it functional or structural. NetSurfP 3.0 was the tool used to understand this role. Compared to [5] where we used NetSurfP2.0, the 3.0 version was ultimately faster, but had the same prediction performance. What we were able to conclude, is that both variants are very optimal to use for these predictions.

IV. CONCLUSION

Our study tried to focus and to clarify on the role, function, properties and possible effects that point mutations have on the LRRK2 protein. We also tried to do a comparison of our previous study with additional predictive algorithms to understand and check if our data is compatible with future studies. Since our previous study and this one correlate on most effects and predictions, we can safely assume that the predictors we additionally chose were up to standard, up to date, with high sensitivity, specificity and accuracy. This could lead to further assessing the mutations that affect the LRRK2 gene and come to final conclusions with high confidence about these SNPs role and possible deleterious effects.

ACKNOWLEDGMENT

REFERENCES

[1] Dauer W, Przedborski S. Parkinson's disease: mechanisms and models. Neuron. 2003 Sep 11;39(6):889-909. doi: 10.1016/s0896-6273(03)00568-3. PMID: 12971891.

[2] Rivero-Ríos P, Romo-Lozano M, Fasiczka R, Naaldijk Y and Hilfiker S (2020) LRRK2-Related Parkinson's Disease Due to Altered Endolysosomal Biology With Variable Lewy Body Pathology: A Hypothesis. Front. Neurosci. 14:556.

[3] Joshi N, Raveendran A, Nagotu S. Chaperones and Proteostasis: Role in Parkinson's Disease. Diseases. 2020 Jun 22;8(2):24. doi: 10.3390/diseases8020024. PMID: 32580484; PMCID: PMC7349525

[4] Zhao Y, Vavouraki N, Lovering RC, Escott-Price V, Harvey K, Lewis PA, et al. (2023) Tissue specific LRRK2 interactomes reveal a distinct striatal functional unit. PLoS Comput Biol 19(1): e1010847. https://doi.org/10.1371/journal.pcbi.1010847.

[5] Vasjari L. et al., 2023, Alterations affecting LRRK2 gene signaling in Parkinson disease, Hristina Rucheva Tasev, IAI Academic Conference Proceedings Florence, 38-43, IAI ACADEMIC CONFERENCE PROCEEDINGS, ISSN 2671-3179, https://ia-institute.com/iai-academic-conference-proceedings-florence-20-june-2023/

[6] https://www.ncbi.nlm.nih.gov/gene

[7] https://www.ncbi.nlm.nih.gov/variation/view/

[8] Coudert E, Gehant S, de Castro E, Pozzato M, Baratin D, Neto T, Sigrist CJA, Redaschi N, Bridge A; UniProt Consortium. Annotation of biologically relevant ligands in UniProtKB using ChEBI. Bioinformatics. 2023 Jan 1;39(1):btac793. doi: 10.1093/bioinformatics/btac793. PMID: 36484697; PMCID: PMC9825770.

[9] Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GLA, Edwards KJ, Day INM, Gaunt, TR. (2013). Predicting the Functional, Molecular and Phenotypic Consequences of Amino Acid Substitutions using Hidden Markov Models. *Hum. Mutat.*, **34**:57-65

[10] Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD, Radivojac P. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nat. Commun.* **11**, 5918 (2020)

[11] Capriotti E, Fariselli P and Casadio R (2005). I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucl. Acids Res. 33:W306-W310.
http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant2.0/I-Mutant2.0.cgi

[12] Magnus Haraldson Høie, Erik Nikolas Kiehl, Bent Petersen, Henrik Nielsen, Ole Winther, Jeppe Hallgren, Paolo Marcatili. *Nucleic Acid Research (June 2022). doi: 10.1093/nar/gkac439*

[13] Qian Sun, Tian Wang, Tian-Fang Jiang , Pei Huang , Dun-Hui Li , Ying Wang , Qin Xiao , Jun Liu , Sheng-Di Chen. Effect of a Leucine-rich Repeat Kinase 2 Variant on Motor and Non-motor Symptoms in Chinese Parkinson's Disease Patients. *Aging and disease*. 2016, 7(3): 230-236 https://doi.org/10.14336/AD.2015.1026