

Research Article

Finding a possible biomarker to tackle Parkinson's disease by splice analysis and random point mutations to counter the expression of genes involved in Parkinson's disease

Piyusha Kalwad¹, Ranjitha Guttapadu¹ and Pradeep S^{1*}

¹Department of Biotechnology, BMS College of Engineering, Bangalore, India

Abstract

The genes showing aberrant alternative splicing in Parkinson's disease namely SNCA, SNCAIP, LRRK2, SRRM2, MAPT and PARK2 were analysed. Two of the genes, namely SNCAIP and SRRM2 that showed high effect were taken and splice site analysis was carried out. Random mutations were carried out on these two genes using Human Splicing Finder tool and the mutations showing the most promising results (i.e., mutations that can restore natural gene expression) were appropriately chosen to tackle Parkinson's disease.

Keywords: Parkinson's disease; phylogenetic analysis; splice site analysis; mutations, Human Splicing Finder; biomarker discovery

Introduction

There are many diseases which are known to affect the nervous system in humans. One such disease is Parkinson's disease. It is a disorder caused due to aberrant alternative splicing events. (Fu *et al.*, 2013).

The alternative splicing process involves removal of the intronic regions from the RNA primary transcript and simultaneous assembly of the exonic regions in different combinations to form a mature mRNA, which then undergoes post transcriptional modification, and gets translated into protein. (La Cognata *et al.*, 2015).

Parkinson's disease (PD) mainly affects the motor neuron system. The most common symptoms which are observed are shaking, slowness in movement, difficulty in walking etc. Other symptoms also include sensory, sleep and emotional problems. Several studies show the involvement of alternative splicing in nervous system diseases. Alternative splicing events in the nervous system play an important role in ion transportation, neurotransportation, memory, and learning. (Fu *et al.*, 2013).

There are six genes which are known to be involved in aberrant alternative splicing in Parkinson's disease. Two of

Article may be cited as:

P. Kalwad et al. (2017) Int. J. Appl. Sci. Biotechnol. Vol 5(3): 336-344. DOI: http://dx.doi.org/10.3126/ijasbt.v5i3.18290

¹*Corresponding author
Pradeep S,
Department of Biotechnology, BMS College of Engineering, Bangalore, India
Email: padees@gmail.com
Peer reviewed under authority of IJASBT
© 2017 International Journal of Applied Sciences and Biotechnology
Image: State S

This paper can be downloaded online at http://ijasbt.org&http://nepjol.info/index.php/IJASBT

these genes, SRRM2 and SNCAIP (codes for synphilin-1), show the most promising effect.

Eyal et al. (2006) and Humbert *et al.* (2007) have shown that synphilin-1A (lacking exons 3 and 4 and containing exon 9A) is present in PD. Overexpression of synphilin-1A causes proteasome saturation; is aggregation enhancing; and directly promotes the inclusion, formation, and neurotoxicity of proteins, which indicates that this isoform may contribute to neuronal degeneration.

Shehadeh *et al.* (2010) performed exon microarray analyses from the peripheral blood of 17 PD patients. They found a noteworthy upregulation of the upstream (5') exons of SRRM2 and a downregulation of the downstream exons, which caused downregulation of the long isoform. (Fu *et al.*, 2013).

Thus, we targeted these genes and tried to counter their regulation to restore original gene expression.

The sequences of these genes were obtained and splice site analysis was performed. Random mutations were carried out in these sequences using online software tools. The effect of these mutations was analysed to check the possibility of upregulation and downregulation.

The main aim of this study was to counter-regulate the gene expression in the alternatively spliced genes.

Tools and Databases

1. The website 'HUMAN SPLICING FINDER' was used to carry out the splice site analysis and mutations. This website allowed us to carry out the splice site analysis by giving the input as the gene name and the specific exon number being targeted.

- 2. NCBI was used as a resource to find out the sequences of the respective genes. The gene name was entered and the specific sequence was downloaded in FASTA format, both for nucleotide and protein.
- 3. EBI tools like CLUSTALW were used to carry phylogenetic analysis. CLUSTALW took all the sequences (of all six genes) and generated a phylogram (with evolutionary distances) and cladogram (without evolutionary distances).

Results

SNCAIP

Synphilin-1A (lacking exons 3 and 4 and containing exon 9A) is over expressed in PD. We chose the transcript of Exon 9 containing 11 exons.

Over expression of synphilin-1A causes proteosome saturation. (Fu et al., 2013) Thus, mutations were carried out on Exon 9 of SNCAIP gene to decrease the expression of exon 9. Various mutations were carried out (insertions, deletions, substitutions, indels and duplications) with an aim to find silencer motifs which could silence the expression of SNCAIP. However, upon analysis in Human Splicing Finder, we could not obtain any silencer motifs. Although we couldn't find silencer motifs, the presence of breakage of the potential splice sites and enhancer motifs indicates that there would be a negative expression and these sites would no longer be a part of the exon. This would make the site unrecognisable for alternate splicing, or would decrease the possibility of alternate splicing. Additionally, the enhancer motifs also showed a negative variation indicating a decrease in the expression of the exon (Fig. 1 to 4).

- HDF ministram									
Position	Splice site type	Motit	New polyceful splice site	Conserves value (0.100)					
-64	Lateria	obtasttratagas	dotastitétag72	74.44					
-54	Ennepthy.	asttratagoagta	aanttataquagTA	79.65					
-61	Scorephon .	Agradiascoepts.	agtagtascisg72	74.04					
-62	Danis	gragtaapp.	SCACTERIC	48.72					
+63	Acceptor.	Astongtatonges	sattagtatrad&	- TT.32					
-61	Scorphon	nitegtahdagenn	constantage []	42.22					
125	Denit	inabgeater.	CADovenne	64.22					
117	benite	110010251	TERRINANA	、「南京の泉泉」					
-12	Acceptor.	3455TEODOORAGE	TADALACONTROL	16.44.017)					
-4	assegner.	O'MORCARCASTA	ocagacaaacagTh	78.41					
-4	Anneptic	gerakkenstansk	gacassingters.	48.04					
14	Duniz	CHTOTERUS	Cillatosta	0.44134					
149	Acceptor	UTCACOCTUCAUAA	gecacgougeagaa	45.77					
- 29	Acceptor	AACAATTICTASAA	esciettorieghA	13.55					
- 52	Acceptate	AATITCIASAASCO	asttbdtegasgCC	74.49					
127	Acceptor	CERGAROCCCASAA	stepsepoteda						
2.9	Addeptor	COURSERATOASES	W0000088850bg\$2	78,65					
	Advention.	CARALLYCANE (MIN	wepakat/regegra	75.88					
44	tints	ABOULAART	Anterest.						
- 19	Lawy r	caaloract	Talgitait	12.10					
13	Acceptor	TTUTTONOCCASET	Tigarriarriar	84.27					
\$1	Dokar	Chigtosta	ChSgtasts	85.35 (WT)					
42	Acceptat	bigtesterrepre	applaatecceptA						
929	Aristitt	ettgtogtttepre	structure and the						
.424	Aristtii	AUCOTOCTORY IN CONTRACTOR	AppailtritingTk	72,50					
741	ANNESS	******	SAMOUTINE SAME	30,2:					
	Assegues	#X#C#1/(1+4g#g	由于由于非过大大年的自己的	78					

Fig. 1: Potential splice sites. The picture above shows the potential splice sites of exon 9 of SNCAIP before carrying out mutations. This gives probable locations where the splicing could occur.



Fig. 2: The picture above shows an INDEL: Insertion of the nucleotide sequence 'ttcgcc' and deletion of 8 nucleotides at position 100. Many different combinations of indels, duplications and substitutions were carried out to study the effect of mutations. The mutation showing the best result (as we could not find any silencer motif) was the above mutation.

Sequence	Canal C	1	- % Modf				7 Mod				
Position	dDNA Position	Ref Model	Ref	Mut Not?	Mag Score	Valiation	Ref Most	Red Score	Mut Moth	Mat	Variation (1%)
14	-67						11station/internitionary	3.83	Concentration of the second	2.83	1 +2
37	-68						terecottexthatsprogtes	4.84	tacabletastctatagesgTAR	0.54	+0
81	-20						gettteggttgsttdstecag&SA	1.40	gtstriggingtbircobis.TO2	-0.97	-In Li
96	46						2VILLOUCOBREALASTAGA	25,75	gestiseccostiogroptegAAC	0.03	100
999							TAGRACIET ET CECECTECASIAC	4.02	CARGAUGINGUADITEISAGAN	4.02	140
331	31						CORACTOCRACEMENT CTAGRAGE	2.44	reserveneesstructep585	3.44	+0
106	- 66	ASSIGCABLET	32	ASSociate's	22	:+3				and and a second	
- 374	78						TOACTORS TO TO ACCOMPANY	\$.76	CARDLEGETERCORDUNGTER	9.78	.+6
291	.56	CAUCTRACE	8.55	CROSTRATA	8.55	- 165					
199	46.						accaposch/bgt/cdtthagost	2.66.	accageseattettetttagCAT	2.66	1.40
223	+20	1000038400	1.27	COOLSANDS	1.27	- 140					
258	445	CLOLASSOC	0.22	TTONTAROO	0.22	30					
258	495						cttocostgetatettotéageg	1.41	ostdottogatetetttaaGAG	1.41	+0.
209	+66						COLUMN TRANSPORTER TO A COLUMN TRANSPORT	1.46	CONCEPTION AT A CONSIGNATION OF A DESCRIPTION OF A DESCRI	3.46	1481

Fig 3: The figure above shows the Maximum entropy values after splicing. Entropy values indicate the stability of the sequence after mutation. The variation in the entropy values after splicing has been shown in the labelled rows.

stantial splice site	Potential	Branch Points	Enhancermotifs	Silencer motifs	Other	plicing	motifs	
HSF Makkas								
Sequence Position	cDNA Position	Splice site type	Most	New splice site	Wild Type	Motant	If cryptic site use, exen length variation	Variation (%3
42.	-59	Acceptor	00CAALICALABOA	sitestcostaptA	76.84	76.84	+47	0
-45	-56	Acceptor	eaccoscagrappa	AstroategragTA	79.65	79.65	+44	- 0.
Ω.	-48	Acceptor	appagraaccaagta	agreqteaccegTA	74.66	74.06	9.27	- 0
. 55	-40	Donor	016254800	0CAdtaxos	36.72	68,72	+36	
58	-43	Acceptor	escrepterrepsp	-ascreigtatougAG	77.42	77.02	+35 -	-0-
65	-41	Acceptor	ocagtatiagagto	disgustcagagTC	-10.88	69.88	>29	3
75	- 25	Donor	085252555	CREcterse	66.85	96.85	+13	-Q.
84	-17	Donor	ctggtogtt	TYGgeoget	67.55	67.59	-15	0
49	-12	Acceptor	TgiffExeccagAC.	TgHTtecscalT/C	.96.44.	60.33	¢	
30	4	kongena	TTS OR DALLARS AND A	southagenerate	25,95	19.74	- 48 - I	
38		Acceptac	нараспалскотал	Atterrepreparate	19:00	19:11		The state
20E	41	Anteptte	decessions and	anapingengandia.	44.38	48,118	211	
114	-14	Donor	COPUTIONCO	CSTotcaby	65.31	65.31	-35	- 6
- 1937	EF	Acceptor	OTCL/OCTSCASSA	guesogoupesgak.	85.77	35.77	-28	.0
179	39	Acceptor	ARCANTTICIASAA	sapastitttagAA	73.55	71.65	-50.	6
142	42	Acceptor	AATITCTAGAAGIC	ANTICOTADAGOC	72.69	72.85	-63	4
147	47	Acceptor	CTASSAGCUCASSA	CERGERDOCCERGRA.	79	79	-94	9
154	54	Acceptor	COUNCERAGE AND CAUSAS	crospanatesg%5	76.4B	76.49	括	-0
156	56	Acceptor	CASSAATCASA3999	cagasatosgagSD.	76.36	28.36	-67	¢
162	62	Acceptor	TCARRENGCARTO	DesgagggcasgTC	88	65	-73	.0
166	66	Donor	AGOGCAAGT	305prsagt	67.35	67.35	-17	3
	70	Press	2018/0922019	1222005AT	23.16	73.98	-261-	

Fig. 4: The figure above shows the potential splice sites after the above-mentioned mutation was carried out. 3 sites were broken and 1 new site was formed

The best result that we obtained after performing mutations showed 3 potential splice sites broken and 5 enhancer motifs sites broken. As seen above, the location of the broken potential splice sites can be easily observed. In the following picture, the broken enhancer motif sites can be observed.

Along with this, three new sites are also formed.

The breakage of the potential splice sites and the enhancer motifs indicate the loss of function of those motifs which enhance the gene regulation. Thus, the gene now is downregulated.

This is the desired level of SNCAIP gene under normal condition as opposed to the overexpressed level observed in Parkinson's disease.

SRRM2

Exon 1 – One of the upstream exons which is upregulated in PD. Transcript containing 15 exons was chosen (Fig. 5 to 11).

Random mutations were carried out to find silencer motifs which would bind to these upregulated sequences and silence their expression. But we were unable to find such silencer motifs. So, in the place of silencer motifs we went on to find if any sites in the potential splice sites were broken. Once the sites are broken, they would not be recognised and hence the alternative splicing at these sites would either reduce or be completely prevented.

On carrying out many random mutations, we found that on the insertion of 4 nucleotides (ATTT) at position 2, 4 potential splice sites were broken. This could be a possible strategy to reduce the alternative splicing of Exon1.

Insertion of 'attt' was carried out at position 2:

tential splice s	ides Pob	initial Branch Points	Enhancer mutits	illencer motifs Other	splicing motifs	
SE From mayer	Ne toy SRovin SC	16 SF2 ASF and SHp55 pro	-			
wshold values						
		SF (IgNI-ERCA1): 70.51				
saudos expressie	a prot carrierone	30. DOWNOOD FENETICISE AFRO	recom values, unit ryp	e value is taken as referenz	#.	
Sequence Position	cONA Position	Linked SR protein	Reference Motif (value 100)	ranked an broken	Mutarri Motif (value G-100)	Variation
3	-96	\$Rp55	550840 (75.52)	SRp65	101460 (75.52)	
9	/82	SRp40	rtaczęg (96.65)	8Rp4)	ttestgg (96.65)	0%
-36	-45	SRp40	atecade (84.91)	5Rp40	acaosos (84.91)	3%
39	-62	SF2(ASF (igM-BRCA1)	08/00/54 (\$2.99)	SF2IASF (IgM-BRCA1)	GADGITA (82.49)	4%
19 39	42	SF2(ASF (igM-BRCA1) SF2(ASF	calcitota (82.85)	SF2/ASF SF2/ASF-0aM-BRCA15	carcota (81.85)	0.97%
38	-16 -62	SF2IASF	capacity (81,89)	SF2/ASF	caccata (82.89)	8 %
48	-62	SR640	cectoria (81.85) ttataor (76.88)	SPatier SPa40	0800004 (81,89) ttatapt (78,80)	3%
40	-63	SC35	excoupts (75.70)	9035	waucegue (76.76)	4 %
73	-45	SRo40		SRp40	eccests (79.46)	0.5
74	- 27	SF2(ASF (igM-BRCA1)	etcatgt (78.77)	SF2(ASF ()gM-BRCA1)	ticator (78.77)	45
74	-27	SP2/ASF (gM-BRCA1)	geowegt (78.77)	SF2IASE	atostat (78.26)	3.55%
74	-27	SF2IASF	etcatgt (76.76)	SF2(ASF ()(MBRCA1)	ctioatus (78.77)	+2.62%
74	-27	SF3GSF	steatet (76.76)	SF2040F	otratgt (78.76)	
	-8	8078	Secondary (TT.23)			
	: =0.1	##2/\$55 (140- 580532)	00000456 (14.49)			
41	-8	- Advertiserer		34241	30877C8 (30.34)	tion, pillip
00	(18)	REFLEXANT LAND	9828CAA (51.13)			
(QuirkNAmenther	ац., —	ALT.	(mp1738 (01.14)			
124	1.1			M2/AM (THE-REAL)	COLUMN (191-191)	100.011
304				STI/ANT	COCCUTA (15,11)	
122		84240	ADADTAG ITS ARE			
114	- 14	SRp55	CREWER (80.84)	SRp65	(1019年12月1日)	0 %
116	16	SRp40	1010200 (86.95)	SRp40	TOTCACO (BE 56)	0%
-121	28	SF2(ASP (igM-8RCA1)	08039CA (72.92)	SF2IASF (igM-BRCAT)	COCTOCA (72.92)	0%
123	29	SF2/ASF (IgM-BRCA1)	CESCROR (73.69)	BF2/ASF (ig4/BRCA1)	CESCAGA (73.69)	0.%
- 124	- 24	SRp65	292838 (84.82)	SRp85	TOCASA (84.62)	0%
333	-23	\$035	AACTOCAN (76.88)	50,8	AACTICERA (78.58)	.0.%
144	44	SRp48	TITCIAS (86.95)	SRp45	1110(Ta# (86.95)	0%
547	47	SRb45	CTAGRAS (82.28)	SRp40	CTNBAAD (82.28)	.0%
154	- 54	SF2(ASF (IgM-BRCA1)	CCICA (844 (77, 23)	SF2(ASF (IgA-BRCA1)	OCCASAA (77,23)	0.5
162	- 62	SRp40	TCAGAGG (89.70)	SRpla	TCAGA00 (89.70)	0%
- 963	43	SF2/ASF (IgM-BRCA1) SF2/ASF (IgM-BRCA3)	02492309 (89.52)	SF2/ASF (IgM-BRCA3) SF2/ASF	CA3A832 (89.92)	42.31 %
163	- 6) - 6)	SF2(ASF (IgN-DRCA3) SF2(ASF	Cassaurus (85.92)		Calaboratio (99,74)	-0.9%
163	- 63	SF2ASF	Chickologic (90,74) Chicklesis (90,74)	SF2IASF (IgH-BRCA1) SF2IASF	CA3A030 (89.52)	0%
165	74	SP2045F SRp40	CASASSS (80.74) 20A0200 (91.66)	SF2ASF SFp40	C3335999 (98:74)	0%
104		SF2/ASF JgM-8RCA1)	CTCCCTT (77.62)	SF2/ASF (IgM-BRCA1)	TCACICC (91.56) CTCCCTT (77.62)	0.5
		GRANDE ISMONIATI	163 6 163 2 (P.1.963)	anaryon (ige-prove))	048-917-64 \$72,047	10 100
177	17	SF2/ASF IgM-BIRCA31	CICCCCIT (77.42)	58/2/458	ettecht (73.33)	6.53 5

Fig 5: The figure shows the breakage of sites in Enhancer motifs that indicates the loss of function of the enhancer motifs that help in the up-regulation of the gene. Thus, the gene is successfully downregulated.

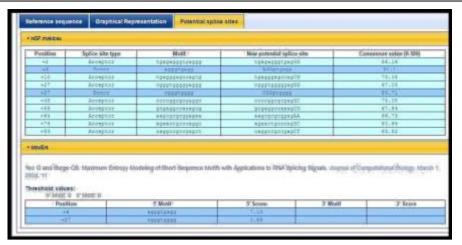


Fig. 6: The figure above shows the potential splice sites in the exon sequence before mutating.

HGF Methon								
Sequence Position	CENA Position	Splice site type	Mosit	New splice site	Weld Type	Mutant	E cryptic site use, exon length variation	Variation (%)
2.	- 4	heapport	121210000000000000000000000000000000000	attettungagog10	46.114	41.,62	-4	
3K		Anti-mphone (ANTAL CONTRACTOR OF T	ALTERNIE STREET	Neith.	18,25	18	
	1	Acceptor	124242312428	11100000001080	66.26	43.04	-	
<u>s.</u>	4	Rempted	*241*277*1*071	Trooproprint	82.58	14.65		
6	6	Acceptor	10104000000000000	SpegagggtgagGG	98,16	. 68.10	-6	
-10	98	Donter	929273A93	W09gtgaagg	95.4	.91.1	-10	10
:14	54	Accessar	topeopgagocagog	tpagggagcobgCu	73.98	73.08	-54	9
14	31	Acceptor	1000010000000000	100010000000000	97.05	67,05	- 31	.0
31	31	Donor	020212022	000000000	83.71	83.71	- Mi	
52	52	Acceptor	2004Epogogiagos	conceptor degR?	75.00	76.08	-42	0
- 50	59	Accessor	pogagginasagig	grightgorenegtD	67.94	67.94	-59	2
38	-86	Acceptor	angogogogogoka	*septerconspagala	98.73	65.73	-86	.0
78	. 78	Acceptor	sgaactgoooaggo	3088050000000000	80.94	85.94	78	0
47	187	Acceptor	ing pring birg sigdt	Chapterspectare.	89.82	69.62	-337	9
Mediat venshold values: 5 Molf. 0-3	Modf: 0							
Requesce Position	cDNA Position	Rut Mont Hat so	5' Mote one Mat Motel Mi	It Score Verlation (*	NJ Bet Mot	Rot Sci	3' Hotf In Mat Motif Mot Score	Variation
35	10	A20272802 7.1		7.13. +0				1
35	31	opertapps 3.0	9 000010000	3.69 +0				1

Fig. 7: The figure above shows the Potential splice sites after the mutation was carried out. Three potential splice sites were broken.

The effect of mutation is seen to have brought the breakage of 3 potential splice sites, thus reducing the probability of alternative splicing occurring in those regions.

Exon 15- One of the downstream exons that is down regulated in PD. Transcript containing 15 exons was chosen.

Since this is one of the exons in the gene that is down regulated during Parkinson's disease, we tried to find enhancer motifs which would cause the up-regulation of the exon. From many random mutations, it was found that many point mutations led to the formation of enhancer motifs.

Some of the sites in these enhancer motifs were broken, while some had new sites formed. This caused a variation in the expression. One of the mutations, which was a duplication of 8 nucleotides at the 8th position gave the maximum positive variation and hence was chosen as the most efficient strategy for the up-regulation of exon15, SRRM2 gene expression.

Duplication of 8 nucleotides at position 8

erence sequ	ence Graphical Repre	sectation Patential a	place attes	
Friettee		1.156.00		11.00 million
testine -	Splice site type	Motif	New potential splice sits	Consenses value (0-100)
-40	2001	1111111	arigt syng	11
244	(2000)	274274297		-84,48
一种	(1000)	Alleherre .	131414300	11.72
	Amosphon	contrationario	UIQuipping II	72,58
- 250	Amospher	00x01200004046	agegt packageA.	11.18
-36	Acceptus	401000000000000000000000000000000000000	#/tgpsttd#gtD	27,65
-12	Timit:	83357,0004	12731-0364	81,498
-12	Acceptor	05555555959977	ettotettgaagdT	-\$6.67 (W(T))
18	Timit.	PRATICITY	TAIgt (11)	112-18
eda n				

Position	Spillov elle type	Modi	New potential splice site	Consenses value (0-100)
-10-11-1	Desta	878878878	of Generative -	69.74
-24	State 1	11111111111	GEGER.pogr.:	11.00
-10	Dutter	10001000	Tabyyerra	19.72
-78	Anieptus	mehtitiggoogna	contit tigupadid.	142-88
-19	Anomiter	228222220018289	sysphydronaxiXA	78.28
-14		art part is a part	sonpprovingsoff.	77,43
-10	Tenir	000070000	000014004	教育学教教主义
-12	Acceptor	191402091004097	1011050575998077	- 45.67 (MT)
-8- 17	Damag	ALC: NO. OF THE OWNER.	CLOTHER	TL.M.
NEY				
iciti Ishold value		eeing at Stein Sequence Monts	with Applications to RNA Specing Signals, a	sumer of Computational Beauty Matrice
G and Barge L 11 shold value	161	nemy of Short Sequence Mottle	vith Applications 15: RNA Splitting Signalis, J 27 Madd	
G and Barga L 11 Shold value S' Mot?	HAT 0. 31 MARIE (5			sumail of Computational Basings March 3' Some

Fig. 8: This figure above shows the Potential splice sites and entropy modelling that represents the stability of the gene sequence before carrying out mutations.

Potential splice s	dies Poten	ial Branch Points	Enhancer multis	cer moths Othe	splicing motils						
ESEFECter manage	af fictor manuses for Spon State Spon State and Spon Sponsores										
SF2/ASF.	reshalid values: SF2IASF: 72.98: SF2IASF (IgNABRCA1): 70.61: SRp40: 78.08: SC36: 75.05: SRp56: 73.86 Istion expresses the difference between reference and mutant values. Wild Type value is tailen as reference.										
lequence Position	CONA Position	Linked SR protein	Reference Mutil (value 0-100)	Linked SR protein	Mutant Motif (value 0.100)	Variation					
3	:3	\$C35	tgassestg (84.63)	\$035	00008030(0443)	6%					
5	5	SRp45	coactgo (95.26)	SRp40	coacter (9126)	非所					
24	H. 1	1011	accesso (75.27)	1							
11	12			2019	manage (19.811.	204 0110					
11	12	38940	(recense (fil.34)								
11	CE .	33(24)	utaduna: (78,32)								
- 18	28			2215	20111054 (TE.27)	Des gala					
20	15			與相	ttrates [E1.54]						
12	1 2t		1	(法)	TTREAM (11.31)	line with					

Fig. 9: The figure above shows the Enhancer motifs after the above-mentioned mutation. Four new sites were formed while three sites were broken. The enhancer motifs variation represents the difference between the reference and the mutant values. The wild type value was taken as reference.

lotential splice site	Potential E	Branch Foints	Enhancer motifs	Silencer motifs	Others	splicing	notils	
HEF Nations					-			
Sequence Position	CONA Position	Splice site type	Math	New splice site	Wild Type	Matant	If cryptic site use, exon length variation	Verietion (%)
11	::	Acceptor	статоствоннорн	CitizeogettaC	10.17	25.76	-44	
3	ii	Acceptor	titacasagaagaa	teetgetttetsG	11.47	31.77	-14	
111	в	Acceptor	taceesgeegeeg	ctttctacaaagiik	16.41	10.17	-13	500.010 +626.01
22	22	Inceptor	TROBARDSROAMED	triataagaagla	35.41	12.47	-22	Ser att

Fig. 10: The above figure shows the potential splice sites after the mutation. Two sites are observed to be broken while two new sites are formed. The broken splice sites reduce the possibility of alternative splicing occurring at that region.

	3' Motif		
	Mut Motif	Mut Score	Variation (%)
	egecactgettteetgetttCTA	-10.96	-637.25
	actgettteetgetttetacAAA	1.87	+108.1
	tgettteetgetttetacaaAGA	0.61	+103.16
	otttoctgotttotacaaagAAG	5.4	+117.9
	atgtctgtcctgtgttgcagCAG	10.09	÷0
	tetgteetgtgttgeageagCAG	4.63	+0
	atgagacaccgetectecagGIG	5.77	÷0
10 m 10	ccaggtgcgtgtcctggaagGCT	2.08	ф.

Fig. 11: The figure above shows the Entropy variation after mutating. A large positive variation is observed on an average indicating good stability of the sequence after mutating.

Gene	Regulation To Restore Original Gene Functionality	Potential Splice Sites	Enhancer Motifs	Result
Sncaip (Exon 9)	Downregulation	Broken- 3 Formed – 1	Broken – 5 Formed - 3 (Negative Variation)	Down bregulated
Srrm2 (Exon 1)	Downregulation	Broken – 4	_	Down regulated
Srrm2 (Exon 15)	Upregulation	Broken – 2 Formed – 2	Broken – 3 Formed – 4 (Max Positive Variation)	Up regulated

 Table 1: Consolidated Results

The Table 1 summarizes our study on these two genes. It depicts the result of random mutations on the expression of these genes.

The future scope of study of this project involves the study of the extent to which the expression of the genes alters to develop an effect therapy to tackle Parkinson's disease. Probable wet lab studies including site directed mutagenesis along with techniques such as FISH can help to validate the results.

Discussion

Alternative splicing is a key element in eukaryotic gene expression that increases the coding capacity of the human genome and an increasing number of examples illustrate that the selection of wrong splice sites cause various human diseases (Tazi *et al.*, 2009). Improper selection of splice sites leads to differential gene expression. We thus targeted the differentially expressed genes with an aim to restore their native gene expression.

The expression of genes is controlled by various external and internal factors. Some of them include enhancer motifs and silencer motifs. Enhancer motifs are sequences that direct accurate slicing of heterogeneous nuclear RNA to pre mRNA or mRNA. Thus, their presence leads to up regulation of genes. While silencer motifs are sequences that bind to enhancer motifs and prevent up regulation of genes. Enhancer motifs have been shown to bind negative regulators belonging to the heterogeneous nuclear ribonucleoprotein (hnRNP) family (Sironi *et al.*, 2004).

Thus, to restore original gene expression of the alternatively spliced genes (whose differential expression led to Parkinson's disease), we looked for silencer motifs for differentially upregulated genes and enhancer motifs for differentially down regulated genes.

We also tried out random point mutations in the sequences of the alternatively spliced genes to see if we can obtain the enhancer or silencer motifs required. Insertions and deletions were also carried out to see if any of the sites break or new sites is formed as origins of replication could conceivably be made non-functional by mutations that change, delete or disrupt sequences recognized by the relevant binding proteins (Genomes 2, Chapter 13).

Based on the number of sites broken and number of sites formed the regulation of expression can be manipulated. For example, by mutating Exon 9 of SNCAIP we found that 3 potential splice sites are broken. This implies that the original binding sites for enhancer motifs or inducers of splicing are unavailable. Thus, the up regulation of this gene will be disrupted. Similar analysis was carried out for the other genes.

The future scope of study of this project involves the study of the extent to which the expression of the genes alters to develop an effect therapy to tackle Parkinson's disease. Probable wet lab studies including site directed mutagenesis along with techniques such as FISH to validate the results can be carried out.

Conclusion

The expression level of genes plays an important role in various cell signalling pathways in the body. Up-regulation of genes leads to their over expression while downregulation leads to their under expression. Both the conditions lead to the alteration of homeostasis and may lead to a probable disorder. Two such genes involved in the development of Parkinson's disease were targeted in this study.

Genes that have undergone aberrant alternative splicing leading to the development of the disease were randomly mutated in such a way that the expression of up-regulated genes is decreased and expression of down-regulated genes is increased thus restoring their normal gene expression and indirectly their function.

Acknowledgements

We would like to express our heartfelt gratitude to our HOD, Dr. Savithri Bhat for her constant support. We would also like to thank the lab assistants for helping us complete this project. We would also like to extend our heartfelt gratitude to our institution, BMS College of Engineering for the encouragement and support they provide us in all our endeavours.

References

- Eyal A, Szargel R, Avraham E, Liani E, Haskin J, Rott R and Engelender S (2006) Synphilin-1A: An aggregation prone isoform of synphilin-1 that causes neuronal death and is present in aggregates from □-synucleinopathy patients. *Proc. Natl. Acad. Sci.* USA **103**: 5917–5922; DOI: <u>10.1073/pnas.0509707103</u>
- Fu RH, Liu SP, Huang SJ, Chen HJ, Chen PR, Lin YH, Ho YC, Chang WL, Tsai CH, Shyu WC and Lin SZ. (2013) Aberrant Alternative Splicing Events in Parkinson's Disease. *Cell Transplantation* 22: 653–661 DOI: <u>10.3727/096368912X655154</u>
- Humbert J, Beyer K, Carrato C, Mate JL, Ferrer I, and Ariza A (2007) Parkin and synphilin-1 isoform expression changes in Lewy body diseases. *Neurobiol Dis* 26: 681–687. DOI: <u>10.1007/s10048-008-0124-6</u>
- La-Cognata V, D'Agata V, Cavalcanti F and Cavallaro S (2015) Splicing: is there an alternative contribution to Parkinson's disease?. neurogenetics. 16(4):245-263. DOI: <u>10.1007/s10048-015-0449-x</u>
- Shehadeh LA, Yu K, Wang L, Guevara A, Singer C, Vance J and Papapetropoulos S (2010) SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's disease. *PLoS One* 5: e9104. DOI: <u>10.1371/journal.pone.0009104</u>

Sironi M, Menozzi G, Riva L, Cagliani R, Comi GP, Bresolin N, Giorda R, Pozzoli U (2004) Silencer elements as possible inhibitors of pseudoexon splicing. *Nucleic acids research* 32(5): 1783-1791. DOI: <u>10.1093/nar/gkh341</u> Tazi J, Bakkour N, Stamm S (2009) Alternative splicing and disease. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease. **1792**(1): 14-26. DOI: <u>10.1016/j.bbadis.2008.09.017</u>