

Ejemplo comparación de resultados predictores in silico

Cambio de estudio GABRB2 c.353-1057G>T (chr5:161337826 G/T, o NM_000813.3: c.353-1057G>T)

Exón 6 e intrones adyacentes:

```
atttaagcagtccttttaaatatttgggttggttttggaaatggcagagcaatgtccacC
ttgocctgtctccagacattgtgaacatcaactatccagatgcaaaagtctttgaaaaga
taaattgtgatatctcttcacatgaatgtgacatatgcttttatttgggttttttttttgg
gagggtgggtgaattttgtaagatatatttaaatattcactggattttatggcatacttccc
tgtacagcactgtttacaatctaataatgtgcttattcatgtgctgttatatgtttgcttac
tagctaggattatatgtgtgctcacttcttaagattgtactccacCctgtgtattgagag
tatoaataataacagagaatatgtcttggttatatagactccatccccagcaactgtgcC
cctgatatgcatagtaggtgtttaataaatagttgtggaatgaatggaagcatgCctcca
aaatgcactgtgcatgtccagtgtaactatgcatttatatctactcctcaaagtgattg
gagttaatgttgttttgagtttttctgtggaagattttctttgaaactaataatgctaca
aagtcaaatttgaaaatcaactacatgaatattgaacagatgattcttcttcgtaatcac
ttctcttctgacttgaaattcaaatgtattgtaattcataatattttataattgggggagaa
tttgataatcagcattataactaatagtttaattattgttctcttatttttatcttcttat
aaagtaaaaaactagtgttatgtgataatgttatatgataatgtgataatgtttgatagaa
tggtattcctaatacaggtaaactgtggtttaagtcctcaactacCctattcctattttat
tattaaggtcccaactatactattcctagtgtgcaagagaatctatgctgttaatttctcta
agcttttgtaagtctaataatgggcctaattatactaccctaatagagatcgctcatgaggat
taaatacagtaatagacttgaaaatgttagtatgataccaaacacataacagccttttagc
tataatctttatattatttatataggtcatcttctaaacaggtattttttctgttttgtc
tgttttttttttttttgttttctgtattgtgtgtgtgtgtgtgtgtgtatgtttctttcaCag
ATGGATACACAACTGATGACATTGAGTTTTACTGGCGTGGCGATGATAATGCAAGTAACAG
GAGTAACGAAATTTGAACCTCCACAGTCTCTATTGTAGATTACAAACTTATCACCAAGA
AGTTTGTTTTTTCACAG
gtttgtatcatggaaaaagtgttaaggaaaaataatcttcaacgtattttgtttaaacatgt
toacttatattagagtttctgtcccaagggaagcaccatgacattacaagacaatccactg
```

El cambio se encuentra en la tercera línea del intrón 5 (la **g** en color rojo subrayada en azul).

Se va a obtener los resultados que produce analizar esta variable con los diferente predictores y ver cuál de ellos es más preciso.

NetGene2

Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
857	2	+	0.35	TCCTAATCAG	^	GTAACCTGG	
1121	1	+	0.34	TTCTAAACAG	^	GTATTTTTTT	
1339	1	+	0.71	TTTTCCACAG	^	GTTTGTATCA	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
325	1134	0	-	0.32	ATCTTAAGAA	^	GTGAGCACAC	

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
603	0	+	0.07	TGCTACAAAG	^	TCAAATTTGA	
1104	2	+	0.33	ATTTATATAG	^	GTGATCTTCT	
1200	1	+	1.00	CCTTTCACAG	^	ATGGATACAC	H
1226	0	+	0.07	TGACATTGAG	^	TTTTACTGGC	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
1308	151	0	-	0.17	TGGTGATAAG	^	TTTGTAATCT	
422	1037	0	-	0.15	TGCATATCAG	^	GGCACAGTTG	

Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
857	2	+	0.35	TCCTAATCAG	^	GTAACCTGG	
1121	1	+	0.34	TTCTAAACAG	^	GTATTTTTTT	
1339	1	+	0.71	TTTTCCACAG	^	GTTTGTATCA	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
325	1134	0	-	0.32	ATCTTAAGAA	^	GTGAGCACAC	

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
603	0	+	0.07	TGCTACAAAG	^	TCAAATTTGA	
1104	2	+	0.33	ATTTATATAG	^	GTGATCTTCT	
1200	1	+	1.00	CCTTTCACAG	^	ATGGATACAC	H
1226	0	+	0.07	TGACATTGAG	^	TTTTACTGGC	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
1308	151	0	-	0.17	TGGTGATAAG	^	TTTGTAATCT	
422	1037	0	-	0.15	TGCATATCAG	^	GGCACAGTTG	

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.2.148.574485.0 :

Start	End	Score	Exon	Intron
850	864	0.96	taatcag	gtaaacct
961	975	0.96	agctttt	gt aagtct
1114	1128	0.80	taaacag	gtatTTTT
1332	1346	0.89	tccacag	gtttgtat

Acceptor site predictions for 10.42.2.148.574485.0 :

Start	End	Score	Intron	Exon
55	95	0.80	tccaccttgccctgtctccc	agacttgtgaacatcaactatc
163	203	0.51	tatttgggttttttttgg	aggggtgggtgaatttgaagat
283	323	0.49	tgttatatgtttgcttact	agctaggattatatgtgtgctc
377	417	0.58	gaatatgtctttgttatat	agactctcatccccagcaactg
1084	1124	0.91	atctttatattatttatat	aggtcatcttctaacagggtat
1100	1140	0.51	tataggtcatcttctaaac	aggtattttttctgtttgtc
1180	1220	0.96	tgtgtatgtttcctttcac	agatggatacacaactgatgac
1279	1319	0.81	ttccacagttctctattgt	agattacaaacttatcaccaag
1318	1358	0.98	agaaggttgTTTTTccac	aggtttgtatcatggaaaagtg

Donor site predictions for 10.42.1.119.574473.0 :

Start	End	Score	Exon	Intron
850	864	0.96	taatcag	gtaaacct
961	975	0.96	agctttt	gt aagtct
1114	1128	0.80	taaacag	gtatTTTT
1332	1346	0.89	tccacag	gtttgtat

Acceptor site predictions for 10.42.1.119.574473.0 :

Start	End	Score	Intron	Exon
55	95	0.80	tccaccttgccctgtctccc	agacttgtgaacatcaactatc
163	203	0.51	tatttgggttttttttgg	aggggtgggtgaatttgaagat
283	323	0.49	tgttatatgtttgcttact	agctaggattatatgtgtgctc
377	417	0.58	gaatatgtctttgttatat	agactctcatccccagcaactg
1084	1124	0.91	atctttatattatttatat	aggtcatcttctaacagggtat
1100	1140	0.51	tataggtcatcttctaaac	aggtattttttctgtttgtc
1180	1220	0.96	tgtgtatgtttcctttcac	agatggatacacaactgatgac
1279	1319	0.81	ttccacagttctctattgt	agattacaaacttatcaccaag
1318	1358	0.98	agaaggttgTTTTTccac	aggtttgtatcatggaaaagtg

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
acatg(g/t)aatgt	ggaatg	gtaatg	26828	54%

Human Splicing Finder

 New Donor splice site		Activation of a cryptic Donor site. Potential alteration of splicing	
Algorithm/Matix	position	sequences	variation
HSF Donor site (matrix GT)	chr5:161337830	- REF : CATGGAATG - ALT : CATGTAATG	50.05 > 77.19 => 54.23%

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr			
wt	67	72	ctttgaaaa		-1.69735314047	0.537313432836	15	11	19	-0.86175337		
wt	67	64	agataaatg		-1.62819181977	0.593220338983	7	11	19	-0.31023059		
wt	67	58	atgtgatat		-0.618409855806	0.622641509434	1	11	19	0.47443886		
wt	67	48	tcttcacat		-0.494973932345	0.581395348837	17	12	25	-0.44742951		
wt	67	36	atgtgacat		0.524225877991	0.677419354839	5	12	25	0.74222819		
wt	67	23	tttttat		-4.34020649867	0.666666666667	7	9	27	-1.2738648		
mut	67	72	ctttgaaaa		-1.69735314047	0.55223880597	15	11	19	-0.85693291		
mut	67	64	agataaatg		-1.62819181977	0.610169491525	7	11	19	-0.30475653		
mut	67	58	atgtgatat		-0.618409855806	0.641509433962	1	11	19	0.48053264		
mut	67	48	tcttcacat		-0.494973932345	0.604651162791	17	12	25	-0.43991856		
mut	67	41	atgtaatgt		0.733804510822	0.638888888889	10	12	25	0.49535232		
mut	67	36	atgtgacat		0.524225877991	0.677419354839	5	12	25	0.74222819		
mut	67	23	tttttat		-4.34020649867	0.666666666667	7	9	27	-1.2738648		

Variant Effect Predictor tool

ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000274547.7	protein_coding	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000353437.10	protein_coding	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000393959.6	protein_coding	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000517547.5	protein_coding	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000517901.5	protein_coding	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000520240.5	protein_coding	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant, non_coding_transcript_variant	GABRB2	ENSG00000145864	Transcript	ENST00000522758.1	retained_intron	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant, non_coding_transcript_variant	GABRB2	ENSG00000145864	Transcript	ENST00000612710.1	retained_intron	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant, non_coding_transcript_variant	GABRB2	ENSG00000145864	Transcript	ENST00000674514.1	retained_intron	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant, NMD_transcript_variant	GABRB2	ENSG00000145864	Transcript	ENST00000675081.1	nonsense_mediated_decay	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant, non_coding_transcript_variant	GABRB2	ENSG00000145864	Transcript	ENST00000675245.1	processed_transcript	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000675303.1	protein_coding	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000675381.1	protein_coding	-	-	-
ENST00000393959.6:c.353-1057G>T	5:161460786-161460786	A	intron_variant	GABRB2	ENSG00000145864	Transcript	ENST00000675773.1	protein_coding	-	-	-

ESEfinder

Solo se encuentra una predicción en la secuencia WT que es positiva, pero cuando se comprueba con la correspondiente en la secuencia mutante, la puntuación no sea alterado.

133 (-1326)	ctottoacatggaatgtgacatatgctttt	2.11910	133 (-1326)	ctottoacatggaatgtgacatatgctttt	-9.85440	133 (-1326)	ctottoacatggaatgtgacatatgctttt	1.71670	133 (-1326)	ctottoacatggaatgtgacatatgctttt	-11.03410
133 (-1326)	ctottoacatgtaatgtgacatatgctttt	1.65990	133 (-1326)	ctottoacatgtaatgtgacatatgctttt	-9.78800	133 (-1326)	ctottoacatgtaatgtgacatatgctttt	1.22290	133 (-1326)	ctottoacatgtaatgtgacatatgctttt	-11.02970

Sin embargo, se encuentran una predicción positiva en la secuencia mutante que es negativa en WT para las matrices 5'SS.

$\frac{128}{(-1331)}$	gatatctttcacatggaatgtgacatatg	-13.48100	$\frac{128}{(-1331)}$	gatatctttcacatggaatgtgacatatg	-4.52370	$\frac{128}{(-1331)}$	gatatctttcacatggaatgtgacatatg	-11.94120	$\frac{128}{(-1331)}$	gatatctttcacatggaatgtgacatatg	-5.48180
$\frac{128}{(-1331)}$	gatatctttcacatgtaatgtgacatatg	2.95990	$\frac{128}{(-1331)}$	gatatctttcacatgtaatgtgacatatg	-4.03760	$\frac{128}{(-1331)}$	gatatctttcacatgtaatgtgacatatg	2.45000	$\frac{128}{(-1331)}$	gatatctttcacatgtaatgtgacatatg	-5.01550

Por lo tanto lo más probable es que se esté activando un sitio de *donor* críptico