

Ejemplo comparación de resultados predictores in silico

Cambio de estudio PTPRT c.3771+1049T>G (chr20:42101075 T/G, COSV61967072 o NM_007050.6: c.3771+1049T>G)

Exón 27 e intrones adyacentes:

```
cgcttagaggggtggaatataggcgagtttaggctgagttggaacattactgtctcttttgct
caggggcagcctaaatgagcagggcctatctatttcacctttgtgccccacag
ACCTCAACATTGTGACACCCCGTGTGCGGCCCGGACTGCAGCATGGGCTCCTGCC
CGGAACCATGATAAGAATCGAAGTATGGACGCTGCTGCCTCTGGACCGCTGCCCTTC
CTTATCTCAGTGGACGGAGAATCCAGCAATTACATCAACGCAGCACTGATGGAT
gtaagccgagcccgggcagaaagctccctagctggcattctagggaatattcttctctga
ggtggaggtggctggacctcagggccctgacccaccagccagccccccctgcttctgat
ctactagtcccaaggacctttgtccctgcaggctctgtgtccaaacctagacadaagg
ggcctcttctgctccctttctcacaagagcctggaaagaatagaccagaagcgttaagg
tggctcccggtgtcaccatggcaagtctacaggcctgcctagcaaaccctctccattccca
cagagccgactctcagcctgaagttagtaaaagcaggtattttctgtgctgggcttggc
ctgagcgtgggaacaacaagacgggtgcagacacagggttctgpcctctggtagctccc
atgaaggttccagataactcctgagcagagacaaaaatgaggatggcgacagtaggggctcag
taaatgggtatttgggaaagttagtggaagcagtaaaacaagcatggggtggaagcgggct
gtgttccctggtgtcctgtgaggccactgccccccacttcttgtcatcaggtccagga
aggggcttccacctaaccctggtttgggattagggtcatgttccctggtccagcgagcctaa
gcttctcccttgcggtttttcgatacttctgattatcatcacttccccctccagggtgtac
aggtgtttgcagtaggaaggaaactatggaacatctactaccagccatggcacttttcagca
gagaaaacagaggccagacagagaaagagacttgctcaagggtaccagggagctagtgg
cacagccacacacagagctcctgggtcacactctggaattcatttcttgagcctccctg
caatttgggtggatccaataagaggcaaacatctcacctcctgagggtcaaaaccaagaac
accattgctgatataatgagcccttccagcttccctccccagcccatctgaggttcaca
tccacctagcccatgtggccctcctatctcactcagaatagacagaaacatgggagtgtt
totcctggggtcaaggccaacaactctctctgacctctaaccctctgaagtttctccaa
gggtctgggtgttagaggcaggaggtgatgatggaaggtcaatgctatccagctctc
```

El cambio se encuentra en la antepenúltima línea del intrón 27 (la **t** en color granate).

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'
286		0	+	0.71	ACTGATGGAT	^	GTAAGCCGAG	
464		1	+	0.34	TAGACACAAG	^	GTGGGCTCTT	
520		2	+	0.45	GACCAGAAGC	^	GTAAGGTGGC	
622		1	+	0.44	AGTAAAGCAG	^	GTATTTTCTG	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
1221		265		0	-	0.41	GCCTCAGGAG	^	GTGAGATGTT	
901		585		2	-	0.34	CCAAACCAGG	^	GTAGGGTGAA	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
63		0	+	0.00	TTTTGCTCAG	^	GGGCAGCCAA	
111		0	+	1.00	TGCCCCACAG	^	ACCCTCAACA	H
147		0	+	0.18	GCGGCCCGAG	^	GACTGCAGCA	
241		1	+	0.27	CTTATCTCAG	^	TGGACGGAGA	
438		2	+	0.43	TCCCCTGCAG	^	GCTCTGGTCC	
588		1	+	0.56	TTCCCCACAG	^	AGCCGACTCT	
877		1	+	0.56	TTGTCATCAG	^	GTCCAGGAAG	
999		1	+	0.33	CCCCTCCAAG	^	GTGTACAGGT	
1066		0	+	0.25	TTTTCAGCAG	^	AGAAAACAGA	
1341		0	+	0.23	TCTCACTCAG	^	AATAGACAGA	

Acceptor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
722		764		0	-	0.33	CTCTGCTCAG	^	GAGTTATCTG	
646		840		0	-	0.31	CCACGCTCAG	^	GCCAAGGCC	
498		988		1	-	0.48	TTCTTTCCAG	^	GCTCTTTGTA	

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
286		0	+	0.71	ACTGATGGAT	^	GTAAGCCGAG	
464		1	+	0.34	TAGACACAAG	^	GTGGGCTCTT	
520		2	+	0.45	GACCAGAAGC	^	GTAAGGTGGC	
622		1	+	0.44	AGTAAAGCAG	^	GTATTTTCTG	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
1221		265		0	-	0.41	GCCTCAGGAG	^	GTGAGATGTT	
901		585		2	-	0.34	CCAAACCAGG	^	GTAGGGTGAA	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
63		0	+	0.00	TTTTGCTCAG	^	GGGCAGCCAA	
111		0	+	1.00	TGCCCCACAG	^	ACCCTCAACA	H
147		0	+	0.18	GCGGCCCGAG	^	GACTGCAGCA	
241		1	+	0.27	CTTATCTCAG	^	TGGACGGAGA	
438		2	+	0.43	TCCCCTGCAG	^	GCTCTGGTCC	
588		1	+	0.56	TTCCCCACAG	^	AGCCGACTCT	
877		1	+	0.56	TTGTCATCAG	^	GTCCAGGAAG	
999		1	+	0.33	CCCCTCCAAG	^	GTGTACAGGT	
1066		0	+	0.25	TTTTCAGCAG	^	AGAAAACAGA	
1289		2	+	0.16	CCCTCCCGAG	^	CCCCTCTGAG	

Acceptor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
722		764		0	-	0.33	CTCTGCTCAG	^	GAGTTATCTG	
646		840		0	-	0.31	CCACGCTCAG	^	GCCAAGGCC	
498		988		1	-	0.48	TTCTTTCCAG	^	GCTCTTTGTA	

Desaparece un sitio *acceptor* (en azul) de la secuencia WT (que no participa en el *splicing*, por lo que no lo tendremos en cuenta) y aparece uno nuevo (en rojo) en la secuencia mutante. En caso de que fuera identificado como tal por el *spliceosome* y hubiera un sitio *donor* después que se pudiera emplear, podría producirse la inclusión de un exón críptico.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.0.139.574296.0 :

Start	End	Score	Exon	Intron
187	201	0.52	aatcgaa	gtatggac
279	293	0.85	gatggat	gt aagccg
457	471	0.88	acacaag	gtgggctc
513	527	0.84	cagaagc	gt aaggtg
615	629	0.81	aaagcag	gtattttc

Acceptor site predictions for 10.42.0.139.574296.0 :

Start	End	Score	Intron	Exon
43	83	0.94	attactgtctcttttgctc	aggggcagccaaatgagcaggg
91	131	0.96	attcacctttgtgccccac	agaccctcaacattgtgacacc
221	261	0.70	gcctgcccttccttatctc	agtggacggagaatccagcaat
418	458	0.95	aggacctttgttccccctgc	aggtctgtgtccaaacctagac
568	608	0.90	aacctcttccattccccac	agagccgactctcagcctgaag
676	716	0.59	acaggttctgccctctggt	agctcccatgaagggtccagat
857	897	0.98	ccccactttctgtcatc	aggtccaggaaggggcttcacc
979	1019	0.88	tatcatcacttccccctcca	aggtgtacaggtgtttgcagta
1269	1309	0.80	cttcagcttccccctcccc	agcccatctgagttcacatcca
1321	1361	0.50	tggccctcctatctcactc	agaatagacagaaacatgggat
1395	1435	0.48	tgaccttaatcctcttga	agtttctccaaagggtctgggt
1406	1446	0.58	cctcttgaagtttctccca	aggtctgggtgttagaggcag

Donor site predictions for 10.42.1.119.574308.0 :

Start	End	Score	Exon	Intron
187	201	0.52	aatcgaa	gtatggac
279	293	0.85	gatggat	gt aagccg
457	471	0.88	acacaag	gtgggctc
513	527	0.84	cagaagc	gt aaggtg
615	629	0.81	aaagcag	gtattttc

Acceptor site predictions for 10.42.1.119.574308.0 :

Start	End	Score	Intron	Exon
43	83	0.94	attactgtctcttttgctc	aggggcagccaaatgagcaggg
91	131	0.96	attcacctttgtgccccac	agaccctcaacattgtgacacc
221	261	0.70	gcctgcccttccttatctc	agtggacggagaatccagcaat
418	458	0.95	aggacctttgttccccctgc	aggtctgtgtccaaacctagac
568	608	0.90	aacctcttccattccccac	agagccgactctcagcctgaag
676	716	0.59	acaggttctgccctctggt	agctcccatgaagggtccagat
857	897	0.98	ccccactttctgtcatc	aggtccaggaaggggcttcacc
979	1019	0.88	tatcatcacttccccctcca	aggtgtacaggtgtttgcagta
1269	1309	0.80	cttcagcttccccctcccc	agcccatctgagttcacatcca
1395	1435	0.48	tgaccttaatcctcttga	agtttctccaaagggtctgggt
1406	1446	0.58	cctcttgaagtttctccca	aggtctgggtgttagaggcag

Desaparece un sitio *acceptor* en la secuencia mutante. Este está dentro del intrón y no participa en el *splicing*, por lo que no se tendrá en cuenta.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
ctatc(t/g)cactc	ctcact	cgcact	29856	73%

Human Splicing Finder

 Alteration of auxiliary sequences	Significant alteration of ESE / ESS motifs ratio (2)	
Algorithm/Matix	position	sequence
ESE_SRp40 (ESE Site Broken)	chr20:42101075	TCACTCA
Sironi_motif3 (ESS Site Broken)	chr20:42101077	TCTCACTC
EIE (New ESE Site)	chr20:42101080	CTATCG
ESE_SRp40 (New ESE Site)	chr20:42101081	CCTATCG

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr	
wt	33	86	atctgagtt		0.0466669245856	0.555555555556	2	13	19	0.64988276
wt	33	81	agttcacat		-0.512706263568	0.552631578947	1	9	13	0.43732059
wt	33	48	atctcactc		1.56689513371	0.441860465116	25	11	24	-0.2008784
wt	33	44	cactcagaa		-0.414550395741	0.410256410256	21	11	24	-0.73372187
mut	32	85	atctgagtt		0.0466669245856	0.55	2	13	19	0.64808847
mut	32	80	agttcacat		-0.512706263568	0.546666666667	1	9	13	0.43539411
mut	32	43	cactcagaa		-0.414550395741	0.421052631579	21	11	24	-0.730235

Desaparece un sitio BP en la secuencia mutante, pero como tenía puntuación negativa, no se tendrá en cuenta.

Variant Effect Predictor tool

ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000356100.6	protein_coding	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000373184.5	protein_coding	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000373187.5	protein_coding	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000373190.5	protein_coding	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000373193.7	protein_coding	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000373198.8	protein_coding	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000373201.5	protein_coding	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000612229.4	protein_coding	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant, NMD_transcript_variant	PTPRT	ENSG00000196090	Transcript	ENST00000617474.1	nonsense_mediated_decay	COSV61967072
ENST00000373198.8:c.3771+1049T>G	20:42101075-42101075	C	intron_variant	PTPRT	ENSG00000196090	Transcript	ENST00000618610.1	protein_coding	COSV61967072

ESEfinder

Se encuentran dos resultados con puntuaciones positivas en las matrices 3'SS (1327 y 1332) y uno con puntuaciones positivas en las matrices 5'SS (1305):

1305 (-181)	atccatctagcccatgtggccctcctatct	1.99310	1305 (-181)	atccatctagcccatgtggccctcctatct	-6.27660	1305 (-181)	atccatctagcccatgtggccctcctatct	1.84160	1305 (-181)	atccatctagcccatgtggccctcctatct	-7.61550
1327 (-159)	tcctatctcactcagaatagacagaaacat	-21.96750	1327 (-159)	tcctatctcactcagaatagacagaaacat	7.31010	1327 (-159)	tcctatctcactcagaatagacagaaacat	-18.54730	1327 (-159)	tcctatctcactcagaatagacagaaacat	7.28890
1332 (-154)	tctcactcagaatagacagaaacatgggat	-13.45840	1332 (-154)	tctcactcagaatagacagaaacatgggat	1.17670	1332 (-154)	tctcactcagaatagacagaaacatgggat	-14.31230	1332 (-154)	tctcactcagaatagacagaaacatgggat	0.96140

1305 aumento muy poco para tenerse cuenta, pero las otras dos bajan las puntuaciones, hasta 1332 pasar a negativa.

1305 (-181)	atccatctagcccatgtggccctcctatcg	2.09580	1305 (-181)	atccatctagcccatgtggccctcctatcg	-5.64190	1305 (-181)	atccatctagcccatgtggccctcctatcg	2.09070	1305 (-181)	atccatctagcccatgtggccctcctatcg	-6.99730
1327 (-159)	tcctatcgcaactcagaatagacagaaacat	-21.74810	1327 (-159)	tcctatcgcaactcagaatagacagaaacat	5.64690	1327 (-159)	tcctatcgcaactcagaatagacagaaacat	-18.33580	1327 (-159)	tcctatcgcaactcagaatagacagaaacat	5.66590
1332 (-154)	tcgcactcagaatagacagaaacatgggat	-12.83850	1332 (-154)	tcgcactcagaatagacagaaacatgggat	-0.48500	1332 (-154)	tcgcactcagaatagacagaaacatgggat	-13.67240	1332 (-154)	tcgcactcagaatagacagaaacatgggat	-0.60730

Por lo tanto, lo más probable es que se estuviera debilitando y/o perdiendo un sitio *acceptor* en la secuencia mutante.