

Ejemplo comparación de resultados predictores in silico

Cambio de estudio TYW1 c.270G>A (chr7:66998951 G/A, rs11547292 o NM_018264.4: c.270G>A)

Exón 3 e intrones adyacentes:

```
accttattatccaaggggctaataatttgagggaatacaaccagaaagattaagagaaaaa  
tacagtgtgtcagattaagagaaaaataaatagtgtcagtaactgttgctaattttcatc  
ctgctgaagtgtgggaaacgatcttagacttgatggattttgtgttaattatttgtcttcaag  
GGCAAGAACTTACAGGAAAAATCTGTTCCAAGAGCAGCTCAGGATTGATGACAAATGGT  
TATGCTCCCTTCAAGAGAAAGACATCTTTGTGCTTGGAGTGAAGATTTTTTATGGTTCT  
CAGACTGGAAACAGCGAAG  
gtaagaaatttatatgatgcttttcctttgggtttcctgacattttaagaaaatcccttta  
gcagtgaacttttagcagtgaactgtccttttagcagtgaactgaattggtcatctaaccctt  
ctttattctttctcccagccttttcccccaaaatatttttcattctgtgacggtaggca
```

El cambio se encuentra en la última línea de exón 3 (la g en color verde).

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'
319	0	+		0.99	AACAGCGAAG	^GTAAGAAATT	H	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
435	64	2	-	0.00	AATAAGAAG	^GTTAGATGAC		

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
180	0	+		0.28	TGTCTTCAAG	^GGCAAGAACT		
186	0	+		0.18	CAAGGGCAAG	^AACTTACAGG		
195	0	+		0.18	GAACTTACAG	^GAAAAATCTG		
456	2	+		0.00	TTTCTCCAG	^CCTTTTCCCC		

Acceptor splice sites, complement strand

No acceptor site predictions above threshold.

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
319	0	+		0.94	AACAGCAAAG	^GTAAGAAATT	H	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
435	64	2	-	0.00	AATAAGAAG	^GTTAGATGAC		

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
180	0	+		0.27	TGTCTTCAAG	^GGCAAGAACT		
186	0	+		0.18	CAAGGGCAAG	^AACTTACAGG		
195	0	+		0.17	GAACTTACAG	^GAAAAATCTG		
456	2	+		0.00	TTTCTCCAG	^CCTTTTCCCC		

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
305	194	2	-	0.25	GCTGTCCAG	^TCTGAGAACC		

El sitio *donor* (en azul) predicho es el correspondiente al exón y, por la presencia de la mutación, su nivel de confianza disminuye. Esto podría producir que se debilitara el sitio, haciendo que el *spliceosome* no lo reconociera, lo que provocaría cambios en el *splicing*. Esto es poco probable, dado que su nivel de confianza sigue siendo alto.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
312	326	1.00	agcgaag	gt aagaaa

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
283	323	0.89	aagatTTTTTatggttctc	ag actggaacagcgaagtaag
436	476	0.92	cttctttattctttctccc	ag ccttttccccctcaaataatt

Donor site predictions for mut :

Start	End	Score	Exon	Intron
312	326	1.00	agcaaag	gt aagaaa

Acceptor site predictions for mut :


Start	End	Score	Intron	Exon
283	323	0.84	aagatTTTTTatggttctc	ag actggaacagcaaagtaag
436	476	0.92	cttctttattctttctccc	ag ccttttccccctcaaataatt

Se producen dos cambios entre las predicciones para las secuencias WT y la mutante. El primero será en el sitio *donor* predicho, que se ve alterado debido a la presencia de la mutación. Pero como el *score* sigue siendo el mismo y es este el sitio que corresponde con el *donor* del exón, el *donor* no se verá alterado. El otro cambio se produce en el primer sitio *acceptor* predicho que cambia debido a la presencia de la mutación, pero este no coincide con el sitio *acceptor* normal del *splicing* por lo que no se tendrá en cuenta.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
acagc(g/a)aaggt	agcgaa	agcaaa	29734	72%

Human Splicing Finder

 No significant impact on splicing signals.	No significant impact on splicing signals.
--	--

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr	
wt	17	130	gagtgaaga		-1.43700571908	0.536	4	8	19	-0.063957738
wt	17	120	tttttatgg		-3.09494031759	0.530434782609	43	20	40	-2.9879167
wt	17	111	ttctcagac		0.235315307719	0.518867924528	34	20	40	-1.1180151
wt	17	91	aggtaagaa		-1.949418832	0.581395348837	14	20	40	-0.68728078
wt	17	83	aatttatat		-3.75777453105	0.602564102564	6	20	40	-0.88211408
wt	17	78	atatgatgc		-0.0802349284277	0.616438356164	1	20	40	0.87878908
wt	17	56	tcctgacat		2.41746425666	0.549019607843	43	9	19	-1.0191801
wt	17	49	attttaaga		-5.38822660344	0.522727272727	36	9	19	-3.6408804
wt	17	48	ttttaagaa		-2.42879329846	0.53488372093	35	9	19	-2.4148979
wt	17	35	cctttagca		-3.3222356194	0.533333333333	22	9	19	-1.9423407
wt	17	28	cagtgaact		0.363286991875	0.608695652174	15	9	19	-0.031861019
wt	17	22	actttagca		-4.54172837212	0.588235294118	9	9	19	-1.5792246
wt	17	15	cagtgaact		0.363286991875	0.8	2	9	19	0.85280286
mut	17	130	gagtgaaga		-1.43700571908	0.536	4	8	19	-0.063957738
mut	17	120	tttttatgg		-3.09494031759	0.530434782609	43	20	40	-2.9879167
mut	17	111	ttctcagac		0.235315307719	0.518867924528	34	20	40	-1.1180151
mut	17	91	aggtaagaa		-1.949418832	0.581395348837	14	20	40	-0.68728078
mut	17	83	aatttatat		-3.75777453105	0.602564102564	6	20	40	-0.88211408
mut	17	78	atatgatgc		-0.0802349284277	0.616438356164	1	20	40	0.87878908
mut	17	56	tcctgacat		2.41746425666	0.549019607843	43	9	19	-1.0191801
mut	17	49	attttaaga		-5.38822660344	0.522727272727	36	9	19	-3.6408804
mut	17	48	ttttaagaa		-2.42879329846	0.53488372093	35	9	19	-2.4148979
mut	17	35	cctttagca		-3.3222356194	0.533333333333	22	9	19	-1.9423407
mut	17	28	cagtgaact		0.363286991875	0.608695652174	15	9	19	-0.031861019
mut	17	22	actttagca		-4.54172837212	0.588235294118	9	9	19	-1.5792246
mut	17	15	cagtgaact		0.363286991875	0.8	2	9	19	0.85280286

Variant Effect Predictor tool

ENST00000359626.9:c.270G>A	7-66998951-66998951	A	synonymous_variant	TYW1	ENSG00000198874	Transcript	ENST00000359626.10	protein_coding	3/16	416	270	90	A	GCG/GCA	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	synonymous_variant, NMD_transcript_variant	TYW1	ENSG00000198874	Transcript	ENST00000361660.8	nonsense_mediated_decay	3/15	396	270	90	A	GCG/GCA	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	upstream_gene_variant	SBDS	ENSG00000126524	Transcript	ENST00000414306.5	nonsense_mediated_decay	-	-	-	-	-	-	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	synonymous_variant	TYW1	ENSG00000198874	Transcript	ENST00000442959.5	protein_coding	3/5	429	270	90	A	GCG/GCA	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	upstream_gene_variant	SBDS	ENSG00000126524	Transcript	ENST00000463579.1	retained_intron	-	-	-	-	-	-	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	non_coding_transcript_exon_variant	TYW1	ENSG00000198874	Transcript	ENST00000475392.1	retained_intron	3/4	404	-	-	-	-	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	upstream_gene_variant	SBDS	ENSG00000126524	Transcript	ENST00000490953.5	retained_intron	-	-	-	-	-	-	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	downstream_gene_variant	TYW1	ENSG00000198874	Transcript	ENST00000491969.5	processed_transcript	-	-	-	-	-	-	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	intron_variant	TYW1	ENSG00000198874	Transcript	ENST00000615572.4	protein_coding	-	-	-	-	-	-	rs11547292
ENST00000359626.9:c.270G>A	7-66998951-66998951	A	upstream_gene_variant	SBDS	ENSG00000126524	Transcript	ENST00000617799.1	protein_coding	-	-	-	-	-	-	rs11547292

ESEfinder

Solo se observa un resultado con puntuación positiva para las matrices 3'. Sin embargo, cuando se compara esta puntuación con la equivalente en la secuencia mutante se observa que ésta disminuido muy poco, por lo que el efecto más probable que sería que se estuviera debilitando un sitio *acceptor* es poco probable y tampoco tendrá mucho efecto sobre el *splicing* porque nos encontramos en una región *donor*:

289 (-210)	TTTTATGGTTCTCAGACTGGAACAGCGAAG	-15.65210	289 (-210)	TTTTATGGTTCTCAGACTGGAACAGCGAAG	8.12890	289 (-210)	TTTTATGGTTCTCAGACTGGAACAGCGAAG	-16.23750	289 (-210)	TTTTATGGTTCTCAGACTGGAACAGCGAAG	8.02270
289 (-210)	TTTTATGGTTCTCAGACTGGAACAGCAAAG	-16.06320	289 (-210)	TTTTATGGTTCTCAGACTGGAACAGCAAAG	7.79570	289 (-210)	TTTTATGGTTCTCAGACTGGAACAGCAAAG	-16.77230	289 (-210)	TTTTATGGTTCTCAGACTGGAACAGCAAAG	7.68630

En cuanto a los ESE, se producen algunas alteraciones que pueden estar afectando al *splicing*:

309 (-190)	AACAGCG	0.17527	309 (-190)	AACAGCG	-0.02544	309 (-190)	AACAGCGA	-2.02699	309 (-190)	AACAGCG	-3.81669
310 (-189)	ACAGCGA	-2.86015	310 (-189)	ACAGCGA	-1.89965	310 (-189)	ACAGCGAA	-3.43363	310 (-189)	ACAGCGA	-0.56825
311 (-188)	CAGCGAA	2.16746	311 (-188)	CAGCGAA	2.62625	311 (-188)	CAGCGAAG	-1.44287	311 (-188)	CAGCGAA	-4.41193
312 (-187)	AGCGAAG	-5.10014	312 (-187)	AGCGAAG	-3.28570	312 (-187)	AGCGAAGg	-2.06777	312 (-187)	AGCGAAG	-0.23447
313 (-186)	GCGAAGg	-0.21041	313 (-186)	GCGAAGg	-0.90828	313 (-186)	GCGAAGgt	-6.55803	313 (-186)	GCGAAGg	-0.35776
314 (-185)	CGAAGgt	2.87450	314 (-185)	CGAAGgt	3.49849	314 (-185)	CGAAGgta	-1.54856	314 (-185)	CGAAGgt	-1.35916
315 (-184)	GAAGgta	-2.93096	315 (-184)	GAAGgta	-2.73329	315 (-184)	GAAGgtaa	-1.71772	315 (-184)	GAAGgta	-7.12982

309 (-190)	AACAGCA	0.90205	309 (-190)	AACAGCA	0.40993	309 (-190)	AACAGCAA	-0.73009	309 (-190)	AACAGCA	-6.19842
310 (-189)	ACAGCAA	-5.43788	310 (-189)	ACAGCAA	-3.67930	310 (-189)	ACAGCAAA	-3.79512	310 (-189)	ACAGCAA	-1.13608
311 (-188)	CAGCAAA	0.24785	311 (-188)	CAGCAAA	0.91987	311 (-188)	CAGCAAAG	-1.01760	311 (-188)	CAGCAAA	-1.85737
312 (-187)	AGCAAAG	-2.19157	312 (-187)	AGCAAAG	-1.56595	312 (-187)	AGCAAAGg	-2.06777	312 (-187)	AGCAAAG	-0.07917
313 (-186)	GCAAAGg	-2.27420	313 (-186)	GCAAAGg	-2.29203	313 (-186)	GCAAAGgt	-5.26114	313 (-186)	GCAAAGg	2.25104
314 (-185)	CAAAGgt	3.32172	314 (-185)	CAAAGgt	3.20026	314 (-185)	CAAAGgta	-1.90942	314 (-185)	CAAAGgt	-2.80433
315 (-184)	AAAGgta	-3.85883	315 (-184)	AAAGgta	-2.97084	315 (-184)	AAAGgtaa	-3.46979	315 (-184)	AAAGgta	-5.67510

EX-SKIP

Seq	PESS (count)	FAS-ESS hex2 (count)	FAS-ESS hex3 (count)	IIE (count)	IIE (sum)	NI-ESS trusted (count)	NI-ESS all (sum)	PESE (count)	RESCUE -ESE (count)	EIE (count)	EIE (sum)	NI-ESE trusted (count)	NI-ESE all (sum)	ESS (total)	ESE (total)	ESS/ESE (ratio)
wt	12	3	1	38	722.8436	24	-32.9953	13	31	57	799.5271	61	68.6502	78	162	0.48
mut	12	3	1	38	722.8436	24	-32.9953	14	32	58	803.9295	62	69.3027	78	166	0.47

Allele wt has a higher chance of exon skipping than allele mut.

HOT-SKIP

accttattataccaaggggctaataatatttgaggaatacaaccagaaagattaagagaaaaatacagtggtgtcagattaaga
gaaaaataaaatgtgtcagtatctgtgctaattttcatccctgctgaagttgggaaacgatttagacttgatggattttg
tgtaattatttgtcttcaagGGCAAGAACTTACAGGAAAAATCTGTTCCAAAAGCAGCTCAGGATTTGATGACAAATGGT
TATGTCTCCCTTCAAGAGAAAGACATCTTTGTGTCTGGAGTGAAGATTTTATGGTTCTCAGACTGGAACAGCGAAGgt
aagaaatttatatgatgcttttcctttggtttcctgacattttaagaaaatcccttttagcagtgaaacttttagcagtgaaac
tgtccttttagcagtgaaactgaattggtcatcctaactctttattctttctcccagccttttccctcaaatatttttcc
attctgtgacggtaggca