

Ejemplo comparación de resultados predictores in silico

Cambio de estudio CLASRP c.100-1G>A (chr 19:45052070 G/A, COSV55516987 o NM_007056.3: c.100-1G>A)

Exón 2 e intrones adyacentes:

```
ctgggaagcagagccttgacgtgagccgagacagcaccattgcactccaggcctgggcgaca  
gagtgaggctccatctcaaaaaaaaaaagaagtgagcttgetgctaaggctcggttgtgtg  
tgaggsggtggggtttggagcctctgttgggtggtgacctcagtcctgtttacccctgcag  
AAGAAGGACCCAGCCAGTTCCTCAGGTACATGGCCGAGCTTCAAGGTGCACCTGGAT  
TCTGCAGTCGCCCTGGCCGCTGAGAGCCCTGTTAATA  
gtaagactgatatggaaggcaggggagtgtctgaagtctgggagtcacaaacggacctggg  
gtggtgttagagtgtggagcagagactgaagtatggacagacccttggggactcagaggca  
gggaagggtgtgacccgcatatgaagggtgatagtcacagggtgttaagaaacatgagcaggg
```

El cambio se encuentra en la posición justo antes del exón 2 (la **g** 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

No donor site predictions above threshold.

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
172	287	2	-	0.39	TTCTGCAGGG		GTAAACAGGA	

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
162	0	+	0.14	GTGACCTCAG		TCCTGTTTAC	
180	0	+	0.93	ACCCCTGCAG		AAGAAGGACC	
183	0	+	0.20	CCTGCAGAAG		AAGGACCCAG	
186	0	+	0.25	GCAGAAGAAG		GACCCAGCCC	
193	1	+	0.20	AAGGACCCAG		CCCAGTTTCT	
198	0	+	0.19	CCCAGCCCAG		TTCTGCAGG	
207	0	+	0.79	GTTCTGCAG		GTACATGGCC	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
285	174	1	-	0.20	TCCATATCAG		TCTTACATAT	
269	190	2	-	0.96	ATATTAACAG		GGCTCTCAGC	
260	199	2	-	0.19	GGGCTCTCAG		CGGCCAGGGC	
253	206	0	-	0.18	CAGCGGCCAG		GGCGACTGCA	
175	284	1	-	0.28	TCTTCTGCAG		GGGTAAACAG	

Donor splice sites, direct strand

No donor site predictions above threshold.

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
172	287	2	-	0.32	TTTTCAGGG		GTAAACAGGA	

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
183	0	+	0.19	CCTGCAAAAG		AAGGACCCAG	
186	0	+	0.19	GCAAAAGAAG		GACCCAGCCC	
193	1	+	0.19	AAGGACCCAG		CCCAGTTTCT	
198	0	+	0.18	CCCAGCCCAG		TTCTGCAGG	
207	0	+	0.72	GTTCTGCAG		GTACATGGCC	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
285	174	1	-	0.19	TCCATATCAG		TCTTACATAT	
269	190	2	-	0.95	ATATTAACAG		GGCTCTCAGC	
260	199	2	-	0.18	GGGCTCTCAG		CGGCCAGGGC	
253	206	0	-	0.17	CAGCGGCCAG		GGCGACTGCA	
175	284	1	-	0.30	TCTTTTGCAG		GGGTAAACAG	

Desaparecen dos sitios *aceptores* en la secuencia mutante. El primero (en rojo) se encuentra antes del exón, por lo que no lo tendremos en cuenta, mientras que el segundo (en verde) es el sitio *acceptor* del exón 2. Si desaparece pueden ocurrir dos cosas: bien que se pierda por completo el exón o que se esté utilizando otro sitio *acceptor*, lo que provocaría (por orden de aparición y siempre que se emplee el sitio *donor* natural del exón) la pérdida de 3, 6, 10, 15 o 27 nucleótidos del exón, respectivamente.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.1.119.425895.0 :

Start	End	Score	Exon	Intron
201	215	0.90	cctgcag	gt acatgg
272	286	0.42	ttaatat	gt aagact
398	412	0.41	agggaag	gt gtgacc
433	447	0.58	acagggt	gt aagaac

Donor site predictions for 10.42.2.148.425907.0 :

Start	End	Score	Exon	Intron
201	215	0.90	cctgcag	gt acatgg
272	286	0.42	ttaatat	gt aagact
398	412	0.41	agggaag	gt gtgacc
433	447	0.58	acagggt	gt aagaac

Acceptor site predictions for 10.42.1.119.425895.0 :

Start	End	Score	Intron	Exon
160	200	0.82	cagtcctggtttacccctgc	ag aagaaggaccagcccagtt
187	227	0.47	gaccagcccagttcctgc	ag gtacatggccgagcttgcaa

Acceptor site predictions for 10.42.2.148.425907.0 :


Start	End	Score	Intron	Exon
187	227	0.47	gaccagcccagttcctgc	ag gtacatggccgagcttgcaa

Se produce la pérdida de un sitio *acceptor* (en rojo) en la secuencia mutante. Este coincide con el sitio *acceptor* del exón, por lo que se va a ver alterado el *splicing*, bien con la pérdida de exón completo o, en caso de que se emplee el otro sitio *acceptor*, se produciría la pérdida de los primeros 27 nt del exón.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
ctgca(g/a)aagaa	ctgcag	ctgcaa	30626	78%

Human Splicing Finder

 Broken WT Acceptor Site	Alteration of the WT Acceptor site, most probably affecting splicing		
Algorithm/Matix	position	sequences	variation
MaxEnt Acceptor site	chr19:45052051	- REF : AGTCCTGTTTACCCCTGCA G AAG - ALT : AGTCCTGTTTACCCCTGCA A AAG	7.77 > -0.98 => -112.61%
HSF Acceptor site (matrix AG)	chr19:45052059	- REF : TTACCCCTGCA G AA - ALT : TTACCCCTGCA A AA	86.93 > 59.06 => -32.06%

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr			
wt	18	152	aagtgaagct		-0.657793522583	0.496598639456	2	9	14		0.30843261	
wt	18	140	tgctaagct		1.62749920925	0.488888888889	7	8	9		0.83767188	
wt	18	124	gtgtgaggg		-1.52926676534	0.487394957983	34	22	31		-1.9029398	
wt	18	91	tggtgacct		2.63890969484	0.53488372093	1	22	31		1.8332803	
wt	18	86	acctcagtc		-0.101878544138	0.518518518519	2	16	26		0.64496921	
wt	18	76	tgtttaccc		-0.394340216713	0.492957746479	23	12	15		-0.90953777	
mut	18	152	aagtgaagct		-0.657793522583	0.496598639456	2	9	14		0.30843261	
mut	18	140	tgctaagct		1.62749920925	0.488888888889	7	8	9		0.83767188	
mut	18	124	gtgtgaggg		-1.52926676534	0.487394957983	34	22	31		-1.9029398	
mut	18	91	tggtgacct		2.63890969484	0.53488372093	1	22	31		1.8332803	
mut	18	86	acctcagtc		-0.101878544138	0.518518518519	2	16	26		0.64496921	
mut	18	76	tgtttaccc		-0.394340216713	0.492957746479	23	12	15		-0.90953777	

Variant Effect Predictor tool

ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	splice_acceptor_variant	CLASRP	ENSG00000104859	Transcript	ENST00000221455.8	protein_coding	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	downstream_gene_variant	RNU6-611P	ENSG00000207003	Transcript	ENST00000384276.1	snRNA	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	splice_acceptor_variant, NMD_transcript_variant	CLASRP	ENSG00000104859	Transcript	ENST00000391952.7	nonsense_mediated_decay	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	splice_acceptor_variant	CLASRP	ENSG00000104859	Transcript	ENST00000391953.8	protein_coding	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	splice_acceptor_variant	CLASRP	ENSG00000104859	Transcript	ENST00000544944.6	protein_coding	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	splice_acceptor_variant, NMD_transcript_variant	CLASRP	ENSG00000104859	Transcript	ENST00000587112.1	nonsense_mediated_decay	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	upstream_gene_variant	CLASRP	ENSG00000104859	Transcript	ENST00000588016.1	retained_intron	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	splice_acceptor_variant	CLASRP	ENSG00000104859	Transcript	ENST00000588936.5	protein_coding	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	intron_variant, NMD_transcript_variant	CLASRP	ENSG00000104859	Transcript	ENST00000591410.5	nonsense_mediated_decay	-	-	COSV55516987 , COSV55517115
ENST00000544944.6:c.100-1G>A	19:45052070-45052070	A	splice_acceptor_variant, non_coding_transcript_variant	CLASRP	ENSG00000104859	Transcript	ENST00000592056.5	retained_intron	-	-	COSV55516987 , COSV55517115

ESEfinder

Se obtienen 2 resultados con puntuaciones positivas para la matriz 3 prima en la secuencia *wild type* 166 y 169. 172 lo tiene puntuación positiva para una de las matrices 3' por lo que no se tendrá en cuenta:

166 (-293)	tgtttaccctgcagAAGAAGGACCCAGCC	-19.49950	166 (-293)	tgtttaccctgcagAAGAAGGACCCAGCC	10.88720	166 (-293)	tgtttaccctgcagAAGAAGGACCCAGCC	-15.84870	166 (-293)	tgtttaccctgcagAAGAAGGACCCAGCC	10.83330
167 (-292)	gtttaccctgcagAAGAAGGACCCAGCCC	-20.27060	167 (-292)	gtttaccctgcagAAGAAGGACCCAGCCC	-26.39500	167 (-292)	gtttaccctgcagAAGAAGGACCCAGCCC	-18.07390	167 (-292)	gtttaccctgcagAAGAAGGACCCAGCCC	-27.85190
168 (-291)	tttaccctgcagAAGAAGGACCCAGCCCA	-10.76270	168 (-291)	tttaccctgcagAAGAAGGACCCAGCCCA	-12.83510	168 (-291)	tttaccctgcagAAGAAGGACCCAGCCCA	-7.09130	168 (-291)	tttaccctgcagAAGAAGGACCCAGCCCA	-16.80940
169 (-290)	ttaccctgcagAAGAAGGACCCAGCCAG	-21.55670	169 (-290)	ttaccctgcagAAGAAGGACCCAGCCAG	3.06410	169 (-290)	ttaccctgcagAAGAAGGACCCAGCCAG	-17.98410	169 (-290)	ttaccctgcagAAGAAGGACCCAGCCAG	2.74320

Cuando comparamos con las puntuaciones que se obtienen para la secuencia mutante para 169 descienden muy ligeramente, mientras que para 166 pasan a ser negativas:

166 (-293)	tgtttaccctgcaaAAGAAGGACCCAGCC	-22.88690	166 (-293)	tgtttaccctgcaaAAGAAGGACCCAGCC	-3.09770	166 (-293)	tgtttaccctgcaaAAGAAGGACCCAGCC	-19.13760	166 (-293)	tgtttaccctgcaaAAGAAGGACCCAGCC	-6.51470
167 (-292)	gtttaccctgcaaAAGAAGGACCCAGCCC	-17.96510	167 (-292)	gtttaccctgcaaAAGAAGGACCCAGCCC	-9.81570	167 (-292)	gtttaccctgcaaAAGAAGGACCCAGCCC	-15.71130	167 (-292)	gtttaccctgcaaAAGAAGGACCCAGCCC	-13.31460
168 (-291)	tttaccctgcaaAAGAAGGACCCAGCCCA	-10.21450	168 (-291)	tttaccctgcaaAAGAAGGACCCAGCCCA	-8.69570	168 (-291)	tttaccctgcaaAAGAAGGACCCAGCCCA	-6.45050	168 (-291)	tttaccctgcaaAAGAAGGACCCAGCCCA	-12.20110
169 (-290)	ttaccctgcaaAAGAAGGACCCAGCCAG	-21.37090	169 (-290)	ttaccctgcaaAAGAAGGACCCAGCCAG	2.99290	169 (-290)	ttaccctgcaaAAGAAGGACCCAGCCAG	-17.84430	169 (-290)	ttaccctgcaaAAGAAGGACCCAGCCAG	2.74460

Por lo tanto, se está perdiendo un sitio *acceptor* en la secuencia mutante.