

Ejemplo comparación de resultados predictores in silico

Cambio de estudio CHL1 c.197+1G>A (chr 3:326065 G/A, COSV56587931 o NM_006614.4: c.197+1G>A)

Exón 2 e intrones adyacentes:

```
aataatatatatttttaaatttgtgatgacatagcaacttttcttcgtctataaaaatcat
taaattattctgggtttaataaaatgaacatatattaaaacatgaaatgatggcagtgga
agtgtattccttagaagtaagtacatgttgcaaagagagagtgaagatttatcttgatttt
cacttttgatcatcctttcaactatcagtatacaaaagggtttaaagtggttttcgcta
tagattaaccttgcctgctggttgaatagtgtgtttttaagacatatatttaattatttag
TTCAACAGGTTCCAACAATCATAAAACAGTCAAAGTCCAAGTTGCCTTCCCTTCGATG
AGTATTTTCAAATTGAATGTGAAGCTAAGGAAATCCAGAACCAAC
gtgagtattgtttcaaacgaagtgtgttcttttaaatgcgtgctgttctttatgctgctctt
tactcttaacatcacaaagcctgttggaactatgaatccacgcatgattaaagctaagggtt
tacaaaaaatcatatccatgcaaaacaaaaaaaaaatttggttatttcacatatgtttttc
tgccacttgaatttttttgaccataaactaagggtgtgttataaaatacttagagtagt
ctttggtaacagcagaaaactccaacatcccttagtctacaaataaataaggtagctgaa
```

El cambio se encuentra en la primera posición justo después 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

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
137	2	+	0.58	TTCCTTAGAA	^GTAAGTACAT		
407	1	+	0.93	CAGAACCAAC	^GTGAGTATTG		
698	-	+	0.00	AATAATAAG	^GTAGCTGAAN		

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
475	232	1	-	0.86	GCTTGTGATG	^GTAAGAGTAA		

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
308	0	+	0.82	AGTTCAACAG	^GTTCCAACAA		
329	0	+	0.17	CATAAACAG	^TCAAAAGTCC		
336	1	+	0.17	CAGTCAAAAG	^TCCAAGTTGC		
342	1	+	0.07	AAAGTCCAAG	^TTGCCTTTCC		
362	2	+	0.07	CTTCGATGAG	^TATTTTCAAA		

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
518	189	1	-	0.17	AAACCCTTAG	^CTTTAATCAT		
385	322	2	-	0.14	TTTCCTTAG	^CTTCACATTC		

Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
137	2	+	0.58	TTCCTTAGAA	^GTAAGTACAT		
698	-	+	0.00	AATAATAAG	^GTAGCTGAAN		

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
475	232	1	-	0.86	GCTTGTGATG	^GTAAGAGTAA		

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
308	0	+	0.82	AGTTCAACAG	^GTTCCAACAA		
329	0	+	0.07	CATAAACAG	^TCAAAAGTCC		
336	1	+	0.07	CAGTCAAAAG	^TCCAAGTTGC		
342	1	+	0.07	AAAGTCCAAG	^TTGCCTTTCC		

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
518	189	1	-	0.17	AAACCCTTAG	^CTTTAATCAT		

CUTOFF values used for confidence:

Se pierde uno de los sitios *donor* (en azul) de la secuencia WT a la secuencia mutante. Este sitio *donor* coincide con el correspondiente al exón, por lo que estará afectando al *splicing*, bien perdiendo el exón completo, o en caso de que se empleara el siguiente sitio *donor* detectado, se produciría una inclusión de 291 pb al mRNA.

También desaparece un sitio *acceptor* (en rojo) de la secuencia WT a la mutante, pero este se encuentra dentro del exón y no participa en el *splicing* normal, por lo tanto no se tendrá en cuenta.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.3.123.581458.0 :

Start	End	Score	Exon	Intron
130	144	0.97	cttagaa	gt aagtac
400	414	0.99	aaccaac	gt gagtat
691	705	0.75	aaataag	gt agctga

Acceptor site predictions for 10.42.3.123.581458.0 :

Start	End	Score	Intron	Exon
189	229	0.65	tatcatcctttcacttatc	agtatacaaaagaggtttaaag
223	263	0.92	tttaaagtgttttcgctat	agattaaccttgctgctgttt
288	328	0.83	ttttaatatttagttcaac	aggttccaacaatcataaaaca

Donor site predictions for 10.42.1.119.581470.0 :

Start	End	Score	Exon	Intron
130	144	0.97	cttagaa	gt aagtac
691	705	0.75	aaataag	gt agctga

Acceptor site predictions for 10.42.1.119.581470.0 :


Start	End	Score	Intron	Exon
189	229	0.65	tatcatcctttcacttatc	agtatacaaaagaggtttaaag
223	263	0.92	tttaaagtgttttcgctat	agattaaccttgctgctgttt
288	328	0.83	ttttaatatttagttcaac	aggttccaacaatcataaaaca

Se pierde un sitio *donor* en la secuencia mutante que coincide con el sitio *donor* normal del exón, por lo que el splicing se verá afectado, bien perdiéndose el exón completo o bien, en el caso de que se empleara el siguiente sitio *donor* detectado, se produciría la inclusión de 291 pb en el mRNA.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
ccaac(g/a)tgagt	aacgtg	aacatg	28523	65%

Human Splicing Finder

 Broken WT Donor Site	Alteration of the WT Donor site, most probably affecting splicing		
Algorithm/Matix	position	sequences	variation
HSF Donor site (matrix GT)	chr3:326062	- REF : AAC G TGAGT - ALT : AAC A TGAGT	85.86 > 58.72 => -31.61%
MaxEnt Donor site	chr3:326062	- REF : AAC G TGAGT - ALT : AAC A TGAGT	9.54 > 1.36 => -85.74%

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr	
wt	17	68	cgatgagta		-1.21273751968	0.428571428571	2	7	15	0.078490847
wt	17	59	ttttcaaat		-2.07455508372	0.388888888889	44	7	15	-2.9302973
wt	17	53	aattgaatg		-0.622181008521	0.395833333333	38	7	15	-1.9795914
wt	17	47	atgtgaagc		-0.960181259849	0.404761904762	32	7	15	-1.7292608
wt	17	41	agctaaagg		0.0466591828497	0.416666666667	26	7	15	-0.95140041
wt	17	19	acgtgagta		-2.05876143174	0.571428571429	4	7	15	-0.33322543
mut	17	68	cgatgagta		-1.21273751968	0.428571428571	2	7	15	0.078490847
mut	17	59	ttttcaaat		-2.07455508372	0.388888888889	44	7	15	-2.9302973
mut	17	53	aattgaatg		-0.622181008521	0.395833333333	38	7	15	-1.9795914
mut	17	47	atgtgaagc		-0.960181259849	0.404761904762	32	7	15	-1.7292608
mut	17	41	agctaaagg		0.0466591828497	0.416666666667	26	7	15	-0.95140041
mut	17	19	acatgagta		-2.66149115316	0.571428571429	4	7	15	-0.56922263

Hay un único cambio en las predicciones por la presencia de la mutación (tercera posición del fragmento), pero las puntuaciones se hacen más negativas, por lo que no se tendrá en cuenta.

Variant Effect Predictor tool

ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 splice_donor_variant	CHL1	ENSG00000134121	Transcript	ENST00000256509.7	protein_coding	-	-	-	-	-	-	COSV56587931
ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 splice_donor_variant	CHL1	ENSG00000134121	Transcript	ENST00000397491.6	protein_coding	-	-	-	-	-	-	COSV56587931
ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 splice_donor_variant	CHL1	ENSG00000134121	Transcript	ENST00000421198.5	protein_coding	-	-	-	-	-	-	COSV56587931
ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 downstream_gene_variant	CHL1	ENSG00000134121	Transcript	ENST00000427688.5	protein_coding	-	-	-	-	-	-	COSV56587931
ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 splice_donor_variant	CHL1	ENSG00000134121	Transcript	ENST00000435603.5	protein_coding	-	-	-	-	-	-	COSV56587931
ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 splice_donor_variant	CHL1	ENSG00000134121	Transcript	ENST00000449294.6	protein_coding	-	-	-	-	-	-	COSV56587931
ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 splice_donor_variant, NMD_transcript_variant	CHL1	ENSG00000134121	Transcript	ENST00000453040.5	nonsense_mediated_decay	-	-	-	-	-	-	COSV56587931
ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 splice_donor_variant, non_coding_transcript_variant	CHL1	ENSG00000134121	Transcript	ENST00000461289.2	processed_transcript	-	-	-	-	-	-	COSV56587931
ENST00000620033.4:c.197+1G>A	3-326065-326065	A	 splice_donor_variant	CHL1	ENSG00000134121	Transcript	ENST00000620033.4	protein_coding	-	-	-	-	-	-	COSV56587931

ESEfinder

Solo se encuentra una predicción con puntuación positiva en las matrices 5'SS. Si se compara con la región equivalente en la secuencia mutante, se observa que la puntuación pasa a ser negativa. Por lo tanto, se puede decir que se está perdiendo un sitio *donor*, afectando al *splicing*.

392 (-315)	AAATCCAGAACCAACgtgagtattgtttca	7.94180	392 (-315)	AAATCCAGAACCAACgtgagtattgtttca	-15.31660	392 (-315)	AAATCCAGAACCAACgtgagtattgtttca	7.61730	392 (-315)	AAATCCAGAACCAACgtgagtattgtttca	-15.74540
392 (-315)	AAATCCAGAACCAACatgagtattgtttca	-3.04070	392 (-315)	AAATCCAGAACCAACatgagtattgtttca	-16.64660	392 (-315)	AAATCCAGAACCAACatgagtattgtttca	-3.51100	392 (-315)	AAATCCAGAACCAACatgagtattgtttca	-16.99580