

## Ejemplo comparación de resultados predictores in silico

Cambio de estudio TREML4 c.507-1699A>T (chr6:41234787 A/T, COSV58385460 o NM\_198153.3:c.507-1699A>T)

Exón 5 e intrones adyacentes:

```
t c t a g a a g c a c a c a a t t t c t a a c a a t a t g a a c t g c t a a g a t t a a c a c a a g a a a c a
g a t t t c a a t a g a t t a a t a a a t a g g a a a a c t g a a a c a g i t t g a t a a g a c c c a t c c t c a c a g a
a a a t t c a a g a c t c a t g t g g t t t t c a g g t a a c c c a a t t t t t a a g a a a c a t t t a c a t t t c a
t g a t t t a t a t t t c a t a c t a g a t g t t g c a g a a a t a g a a a a g a a g c a a a a c a g c c c a g g t
c a t t t g a a c a g g t t a g t g t c a t c a a a a t c t c a g a a c t a a t t a a g g a g a t a c a a a a a a a t
t a a a t t a t a g g t t c a t t t c t c t g a t g c a c a t g g g g t c c a t t c c c t a a t t a c c a t t a g e t a
a c t g a a t t c a a g a g t g c a t t g a a a a g a a g a c t g a t t a t a a t c g a g g a g a g t c t t a t c a c a
a a a c a a a a a g g a t t t t t c a a c t t t a g a a a a t c t a t t a a t a t a a g t a a a c c a a a t t c a t a a
a c t a c a g g a a a a a a g t c a a c t g c t g c a c a a a a a g c a t t t g g g g a c c t t g a g c a c a t t t a t
g t t a t c t g a a a a a g c t c t t a a t a a a a t g g a a g t a a g a a c a g a t a a c t t g g t a a a g a t t g t
a c a c c a a t g a t t a c c a t a a a g a t t a c t t a t t g g a g g t a t a a a c t c a a a t t a a g g a t t t t c
a t a g c t a c c a t a a a t g t c t t c t a a a a t a t t a g a g a t t c t g a t c a a t g c t t g a a g g a a a g g
a a a a c a g g t t a c a t g a a a g a a g a t t a g a c a a g a a g a g g a a a a a c t g t c a t g a c t t g c a a t
t g a t g g g a t c a t c t a c c c a g a a a a a g a a t g a g a a c a c a a a c t a t t a g a t c t c a c a a g a
g a g t t t a g c a a g g c t g c a g a t a c a a a a t c a a c t g a c a t a c a t t a c a a a t a g a t t t c t a t
g a g g a a t a a t c a t t t a g a a a a t a a g a t c g t t t t c a t t a c a g t c a c a a a g t c t a t a a a g t
t t c t a g a a a g a t t a a t t t a a c c a a g a c c t g t t t g g a g a a a g t a t a a a a t g c t a a a a a
t g a c a g g t g a t c t g a a t a a a t a g a a a c a a t c c a t g a c t t t g g a a g a g g c a a t a a c a t t g t
a a a t a t g t c a a t t t c t c c a a a t t a a a a g t c t t t a a a t g t a a t c c c a a c a c a a c t t c t
t t a c t t t t g t a a a g a a c t t g a a a a a t t t a c t c t a a a a t t a t a t g a a a a a g g a a g a t a c a t
g a g t a g c t a a g t c a a c c a t g a a a a a g t g a a a a t a a a a g a g a c t c a c t g a c t t t a t c a t a t
a t t a t g t a a t a c t g a g g c a g a a a g t t a c c a a g g c c a c a a g a a a a g c a g a t t t g c t g g g
a c a g g g t a g t c c t a g a t c a g a g c t a g g c a g g c c a g g a g t g g g g g c t a c a c a g a g g g g a t g
a g g g a g g a c c c c t g g g a c c a a c a g a c a g g a g g t c t a g g t c t g t g c t g t c a c a c t c c t a a
g g g t c t g g t c a g g t c c t g t g c a g t c t c t a g g c c t c a t t t c t c a g a t g t g a t a t c t t g g c
t g g t c t t g a a g a t g c a c t g g g c t c t g a t t c c t a a g g t c g g a t g g g t c c c a t c c g a c c t g a c
c t g g t t a g g g a g a a t t c d t g d t c c c c a g a c t c t a t d c t c c c c t t c c c t g a a c t c c a a t c
c c t t g t t c t t c c c t g g c c a g c d t t t t t c t c c g t t c a a g a c t t t g a g g t t c c t t c t g a d t
a g g t c c c c t c t c a g t c c c a g a g a c c t c t g t g g t c c t g c c c c a c a c c c c t c t g t c a t c g t
t c a g g g t g g g c c c t t a c g c a a a c a g c t c a g c t a c a a g g g c c t g c c t g g c t t a g g g a c a
c d g c a g g a g g t g g t g a c a t c c t g g c c c a a g t c a t c c t c c t t t t c t c t t c c c t g c a g
G A A A T C A A G A G C C C C T G C C T G C C T T G G C T C A G G T G G C C C C A G A T T C C T G G T C T T G G T G C T
A T G T G G A C T C C T C C T G G C C A A G G C C T G A T G T G T C A G T C T G T A G T G C T C C T G A T C T G
C A G G T G C C A C A G
g t g a g g g g c c t g g a t t t c a c c t g g g g g c a a t g g a g c t g g g g c a g a t t c t g t t c c t a g a
```

El cambio se encuentra en la tercera línea del intrón 4 (la **a** 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'
252	0	+	0.67	ATTTGAACAG	^GTTAGTGTCA		
572	0	+	0.44	TAAAATGGAA	^GTAAGAACAG		
1993	0	+	0.00	GGTGCCACAG	^GTGAGGGGGC		

### Donor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5'	exon	intron	3'
1757	296	2	-	0.32	AGCTGTTTGC	^GTAAGGGCCC		

### Acceptor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5'	intron	exon	3'
146	1	+	0.17	TGGTTTTTCAG	^GTAACCCAAT		
200	0	+	0.56	TTCATACTAG	^ATGTTGCAGA		
1485	1	+	0.23	ATTTCTCAG	^ATGTGATATC		
1641	1	+	0.43	CCCTGGCCAG	^CTTTTTCCTC		
1700	1	+	0.30	TCAGTCCCAG	^AGACCTCTGT		
1744	2	+	0.63	CATCGTTCAG	^GGTGGGCCCT		
1860	1	+	0.77	TTCCCTGCAG	^GAAATCAAGA		
1892	1	+	0.25	CTTGGCTCAG	^GTGGCCCCAG		
1983	0	+	0.00	TGATCTGCAG	^GTGCCACAGG		

### Acceptor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5'	intron	exon	3'
2014	39	1	-	0.00	TTGCCCCCAG	^GTGAAATCCA		
1978	75	1	-	0.00	GCACCTGCAG	^ATCAGGAGCA		
1855	198	2	-	0.43	TTTCCTGCAG	^GGAAGAAGGA		
1392	661	0	-	0.33	TTGGTCCCAG	^GGGTCCTCCC		
989	1064	0	-	0.51	CTCCAAACAG	^GTCTTGGTTA		
897	1156	0	-	0.16	TTCTCATAG	^AAATCTATTT		
708	1345	1	-	0.25	TTCTTCAAG	^CATTGATCAG		

### Donor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5'	exon	intron	3'
252	0	+	0.66	ATTTGAACAG	^GTTAGTGTCA		
572	0	+	0.44	TAAAATGGAA	^GTAAGAACAG		
1993	0	+	0.00	GGTGCCACAG	^GTGAGGGGGC		

### Donor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5'	exon	intron	3'
1757	296	2	-	0.32	AGCTGTTTGC	^GTAAGGGCCC		

### Acceptor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5'	intron	exon	3'
146	1	+	0.16	TGGTTTTTCAG	^GTAACCCAAT		
200	0	+	0.53	TTCATACTAG	^ATGTTGCAGA		
1485	1	+	0.23	ATTTCTCAG	^ATGTGATATC		
1641	1	+	0.43	CCCTGGCCAG	^CTTTTTCCTC		
1700	1	+	0.30	TCAGTCCCAG	^AGACCTCTGT		
1744	2	+	0.63	CATCGTTCAG	^GGTGGGCCCT		
1860	1	+	0.77	TTCCCTGCAG	^GAAATCAAGA		
1892	1	+	0.25	CTTGGCTCAG	^GTGGCCCCAG		
1983	0	+	0.00	TGATCTGCAG	^GTGCCACAGG		

### Acceptor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5'	intron	exon	3'
2014	39	1	-	0.00	TTGCCCCCAG	^GTGAAATCCA		
1978	75	1	-	0.00	GCACCTGCAG	^ATCAGGAGCA		
1855	198	2	-	0.43	TTTCCTGCAG	^GGAAGAAGGA		
1392	661	0	-	0.33	TTGGTCCCAG	^GGGTCCTCCC		
989	1064	0	-	0.51	CTCCAAACAG	^GTCTTGGTTA		
897	1156	0	-	0.16	TTCTCATAG	^AAATCTATTT		
708	1345	1	-	0.25	TTCTTCAAG	^CATTGATCAG		

## Splice Site Prediction by Neural Network (NNSplice)

### Donor site predictions for wt :

Start	End	Score	Exon	Intron
140	154	0.92	ttttcag	gt <sup>a</sup> aaccca
245	259	0.97	tgaacag	gt <sup>a</sup> tagtgt
565	579	0.94	aatggaa	gt <sup>a</sup> aagaac
583	597	0.53	taacttg	gt <sup>a</sup> aaagat
721	735	0.45	aaaacag	gt <sup>a</sup> tacatg
1947	1961	0.70	tgatgtt	gt <sup>a</sup> gagtcc
1986	2000	0.98	gccacag	gt <sup>a</sup> gagggg

### Donor site predictions for mut :

Start	End	Score	Exon	Intron
140	154	0.92	ttttcag	gt <sup>a</sup> aaccca
245	259	0.97	tgaacag	gt <sup>a</sup> tagtgt
565	579	0.94	aatggaa	gt <sup>a</sup> aagaac
583	597	0.53	taacttg	gt <sup>a</sup> aaagat
721	735	0.45	aaaacag	gt <sup>a</sup> tacatg
1947	1961	0.70	tgatgtt	gt <sup>a</sup> gagtcc
1986	2000	0.98	gccacag	gt <sup>a</sup> gagggg

### Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
126	166	0.92	caagactcatgtggttttc	ag <sup>a</sup> gtaaccaatttttaaagaa
180	220	0.54	atgatttatatttcatact	ag <sup>a</sup> atgttcagaaaaatagaaaa
922	962	0.79	ataagatcgttttcattac	ag <sup>a</sup> tcacaaagtcataaagttt
1465	1505	0.89	ctctaggcctcatttcctc	ag <sup>a</sup> atgtgatatcttggtggtc
1515	1555	0.53	cactgggctctgattccta	ag <sup>a</sup> gtcggatggtcccatccgac
1621	1661	0.70	ccttgttcttccctggcc	ag <sup>a</sup> cttttctcctcgttcaagac
1639	1679	0.76	cagcttttctcctcggtca	ag <sup>a</sup> actttgaggttccttctgac
1662	1702	0.96	tttgaggttccttctgact	ag <sup>a</sup> gtcccctctcagtcacagag
1680	1720	0.69	taggtcccctctcagtcac	ag <sup>a</sup> agacctctgtggctcctgcc
1724	1764	0.74	caccctctgtcatcggtc	ag <sup>a</sup> ggtagggcccttacgcaaaaca
1840	1880	0.99	ctttctccttcttccctgc	ag <sup>a</sup> gaaatcaagagccctgcct
1872	1912	0.91	cccctgcctgccttggtc	ag <sup>a</sup> gtggcccagattcctggtc
1963	2003	0.88	gttagtgctcctgatctgc	ag <sup>a</sup> gtgccacaggtgagggggcc

### Acceptor site predictions for mut :


Start	End	Score	Intron	Exon
126	166	0.90	caagactcatgtggttttc	ag <sup>a</sup> gtaaccaatttttatagaa
180	220	0.54	atgatttatatttcatact	ag <sup>a</sup> atgttcagaaaaatagaaaa
922	962	0.79	ataagatcgttttcattac	ag <sup>a</sup> tcacaaagtcataaagttt
1465	1505	0.89	ctctaggcctcatttcctc	ag <sup>a</sup> atgtgatatcttggtggtc
1515	1555	0.53	cactgggctctgattccta	ag <sup>a</sup> gtcggatggtcccatccgac
1621	1661	0.70	ccttgttcttccctggcc	ag <sup>a</sup> cttttctcctcgttcaagac
1639	1679	0.76	cagcttttctcctcggtca	ag <sup>a</sup> actttgaggttccttctgac
1662	1702	0.96	tttgaggttccttctgact	ag <sup>a</sup> gtcccctctcagtcacagag
1680	1720	0.69	taggtcccctctcagtcac	ag <sup>a</sup> agacctctgtggctcctgcc
1724	1764	0.74	caccctctgtcatcggtc	ag <sup>a</sup> ggtagggcccttacgcaaaaca
1840	1880	0.99	ctttctccttcttccctgc	ag <sup>a</sup> gaaatcaagagccctgcct
1872	1912	0.91	cccctgcctgccttggtc	ag <sup>a</sup> gtggcccagattcctggtc
1963	2003	0.88	gttagtgctcctgatctgc	ag <sup>a</sup> gtgccacaggtgagggggcc

El cambio entre secuencias es la presencia de la mutación que hace el sitio *aceptor* (en azul) predicho más débil, pero este no se emplea en el *splicing*, por lo que no tendrá efecto.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
tttta(a/t)agaaa	tttaaa	tttata	31772	85%

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	15	79	agttgataa		-1.07570997305	0.5	6	11	16	-0.088664473
wt	15	76	tgataagac		-1.19553354418	0.507042253521	3	11	16	0.056588331
wt	15	64	tcctcacag		1.34063664272	0.474576271186	35	10	17	-0.97709718
wt	15	53	aattcaaga		-2.38911935059	0.5	24	10	17	-1.7329803
wt	15	46	gactcatgt		1.41608889792	0.512195121951	17	10	17	0.20396553
wt	15	35	ttttcaggt		-2.23210782915	0.466666666667	6	10	17	-0.54289876
wt	15	30	aggtaaccc		2.17647894984	0.52	1	10	17	1.516988
wt	15	19	tttttaag		-4.30072364603	0.357142857143	14	0	0	-2.0529885
wt	15	18	ttttaaga		-2.02572836426	0.384615384615	13	0	0	-1.0900492
mut	15	79	agttgataa		-1.07570997305	0.513513513514	6	11	16	-0.084300007
mut	15	76	tgataagac		-1.19553354418	0.521126760563	3	11	16	0.061137217
mut	15	64	tcctcacag		1.34063664272	0.491525423729	35	12	18	-0.96230726
mut	15	53	aattcaaga		-2.38911935059	0.520833333333	24	12	18	-1.7169359
mut	15	46	gactcatgt		1.41608889792	0.536585365854	17	12	18	0.22115872
mut	15	35	ttttcaggt		-2.23210782915	0.5	6	12	18	-0.52281723
mut	15	30	aggtaaccc		2.17647894984	0.56	1	12	18	1.5392227
mut	15	19	tttttatag		-3.89627971916	0.428571428571	14	0	0	-1.8715603

Se produce cambio en uno de los BP por la presencia de la mutación y desaparece otro en la secuencia mutante. Como estos tienen puntuación negativa, no se tendrán en cuenta.

Variant Effect Predictor tool

ENST00000341495.6:c.507-1699A>T	<a href="#">6:41234787-41234787</a>	T	intron_variant	TREML4	<a href="#">ENSG00000188056</a>	Transcript	<a href="#">ENST00000341495.7</a>	protein_coding	-	-	-	-	-	-	<a href="#">COSV58385460</a>
ENST00000341495.6:c.507-1699A>T	<a href="#">6:41234787-41234787</a>	T	downstream_gene_variant	RNA5SP207	<a href="#">ENSG00000212176</a>	Transcript	<a href="#">ENST00000390874.1</a>	rRNA_pseudogene	-	-	-	-	-	-	<a href="#">COSV58385460</a>
ENST00000341495.6:c.507-1699A>T	<a href="#">6:41234787-41234787</a>	T	intron_variant	TREML4	<a href="#">ENSG00000188056</a>	Transcript	<a href="#">ENST00000448827.6</a>	protein_coding	-	-	-	-	-	-	<a href="#">COSV58385460</a>
ENST00000341495.6:c.507-1699A>T	<a href="#">6:41234787-41234787</a>	T	intron_variant, NMD_transcript_variant	TREML4	<a href="#">ENSG00000188056</a>	Transcript	<a href="#">ENST00000461240.1</a>	nonsense_mediated_decay	-	-	-	-	-	-	<a href="#">COSV58385460</a>

ESEfinder

Se observa una predicción con puntuación positiva en la secuencia WT:

150 (-1903)	acccaattttttaaagaacatttacatttc	-21.91030	150 (-1903)	acccaattttttaaagaacatttacatttc	0.27320	150 (-1903)	acccaattttttaaagaacatttacatttc	-18.74680	150 (-1903)	acccaattttttaaagaacatttacatttc	0.11610
----------------	--------------------------------	-----------	----------------	--------------------------------	---------	----------------	--------------------------------	-----------	----------------	--------------------------------	---------

En la secuencia mutante, la puntuación para esta región aumenta:

150 (-1903)	acccaattttttatagaacatttacatttc	-23.55580	150 (-1903)	acccaattttttatagaacatttacatttc	2.55330	150 (-1903)	acccaattttttatagaacatttacatttc	-20.41450	150 (-1903)	acccaattttttatagaacatttacatttc	2.44280
----------------	--------------------------------	-----------	----------------	--------------------------------	---------	----------------	--------------------------------	-----------	----------------	--------------------------------	---------

Por lo tanto, lo que podría estar pasando es que se hiciera más fuerte un sitio *acceptor*, afectando al *splicing*.