

Ejemplo comparación de resultados predictores in silico

Cambio de estudio VAV3 c.321+1896C>T (chr1:107873005 C/T, COSV64067143 o NM_001079874.2: c.321+1896C>T)

Exón 2 e intrones adyacentes:

```
gc a a c t t c a a a g g c t g c t t g a t g c t t t c t t g t c a t t t t t a g a g c a g t c t t g a g t g t t a c a
g a g g g t g g a t a g c a t t c a g a a t a a t t a a c t t t a t g c t a a c a g c t c t t t t c c a t t c c t c a g
T T T C T C T G T T T G A A G A C A T A A G G A C A T T T C T C A C G G C C T G T T G T G A G A C G T T G G A A T G
A G G A A A A G T G A A C T T T T C G A G G C A T T T G A C T T G T T T G A T G T T C G T G A C T T T G G A A G
g t a a t t a t a c t t c a t t t t t t a a c a c t a c t t t t t t a a c t g g a a a g c a t t t a a g c a a a t t g t t t
c a a a t c t t g a a g t a g g g c a t g t g a c t t a a g t g g t t t a t g t a t c t g c a t t a t a t a g t t a a a
t t t c a c a t g t a g a t g g c t t a g c a a a a c c a a a a a c g a a a a a c a a c a c a a a a a a g a a a c c
c c a c a c a t a c a a c a c a a c a c c a a a a a a t t a g c a t c a a c t t t t c c a a t g a g a t t a g a a t t
a t t t t g t a a t t a a a a t a g t a t a c a g g g a g g g a a t c t t t t c t c a g a g a a a t t t t t t t t t c
t g c t g g t c a c c t c a g a g g t t a g g g t t a c t a t t a t c t g c c c t t t a g g t t g a t c a c t g g g c a
c t c a g g c c t g a g g a c c a g t a t t c t t t t c c t c a g t c t t c t c t a t a t c t g a a a g c t g c t a a
t t t t g g a a g t t g a g a a t g t t c t g t t g t c c t t t t c t c t a t t t t g a t a c t t g g t t g g t g a
c t c a g a g a a g a c a t g a a t c a c a a g t g t t c a a g a t t g t t g a a t a a t t t c t g t t g c t t g a a t
t t c a g c a g a t a t t g t t g g a t t t c a a g g a c a t t t t t t c a c a t g g a a a t t g t t t t a a t a t t t
a c t g t t a t c c g t a t g c a c a c t a a a c a c a t g t a t a a t a a g t g a a c a a a t c a t c a t t t c a a
a a a t t a t g g g c a t t t t a t t a t g a a g t g a t g g a c a t a a a c c t c t c a c a a g t a a c t a g c a t
a a a a g a g t c a g g t a a t t a a g g t g a c a g g a c t t a g c t t a t g c t g a t g t a a t t g c a t t a t g
a a a a t a a g c a a a t a a g g a g g a a a a a c t a a t g c t a t t g a g a a g t t g t c c t t a t t c t t g t
a a t t c a g g t g c t a g t a t c t t a t c a a a a g g c t g t t t g a g a g t a t a c a t t g t t t t a a c t c t g
a g a t a t a a t a t t t c a t t a g a a a a t a a c c a t g g a t t g c t c t t t g c t a a c a g c a t a a a c
t t g a g a c a g a a a t a t a g g a a a t t g a t g c t g a a a t a c a g c g c a t a t g g g a a g c a d c t c t
a c c t g g g g t g g g g a a g g c c t g g t t c t c t c t c c g g t c t t a a c c a t g g t c a t t t t g t g a
t c t a g a a g g c a c c a g c c t t c t c a g g c a t c a g c c t g c c t g c c a g c t t g a a a g c t g t a c a a
a a g g a a g g g g g a t a a g a a g g t g g t t c t a a g a t t c c t t t a a t c t c t g g a a a g c t g t g a t
t t a a c t g t c t g a g g t c c a a a a a g t g a t g g a g g a t a a a a g a t g a t a a t t t g t a c a a c g a a
c t g a a a g c a a g c c c t a g t a a g g a a t g g g a g c g c a t t t g t c t g t g g c c a t g a a t c a c a
g g t g a g t g g a c a c t t c t c t a a g c t g t t g g c c t g a a a a c a c c a g a g a t a t t t c t g g t t
t t a t t t t t c t t t a t c c t t c c g c t g t t c t c t g c a c c t t t a a a a a a c t t c c t t g a t
t t t t c t t t a t t t t c a c c c t c t c t t t g g g g a a a t g a a a a t g a a a t g a a t t g g g g a a a t
g a a a a t g a a a a g t c t t g a c t t t t t a a a g c t c a t t a a g g a a c t a t t t a t t t t t c a t t t a
a g t t t g c a a g a a a t g t g t t g a t a a a t g t t g g a a a a t a t a t a a a t a c t t a a a a t g c a t
g g t g g t g t t c t a c t g g a a t g a g a t t g c a t g c t a g a t t g a t t a g a c a a c g g a a g c t t t g t a
g g a g a g t t g a t a c t a c a a a a a a t t g a a t g t g a a a g c a t a a a a c t g t a t g c t t c t c a a g g a t
a a a t g t a t t t t t g a c t g c t t a a a a t a t t t t a g c a g t g g t a g a t c t g a a a c a g a a t t a a a c
a t g a g a t t g t g g t t t a g a a c t t g a a c c t a c a t t a t a g t c a c t t a c a t t t t t g c t a c t g c c a
a g a t t g g a g t a g c t a t c t c t t t a g a a c t g t a t g a t c a a a g a a a a c a c t g a a g t t a a a g
a a t t t g t g a t a g c t t t c a t a t t t a c c a g t a c t a a t a g t a t t a a g c a t t a t a t a c t c t a
```

El cambio se encuentra en la penúltima línea del intrón 2 (la **c** 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'
238		0	+	0.83	CTTTGGAAG	^	GTAATTATAC	
555		0	+	0.34	CACCTCAGAG	^	GTTAGGGTTA	
921		0	+	0.37	TTATTATGAA	^	GTGAGTGGAC	
969		0	+	0.37	AAAGAGTCAG	^	GTAATTAAGG	
1559		1	+	0.58	TGAATCACAG	^	GTGAGGTGGA	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
2082		136		1	-	0.58	GTAGCAAAAT	^	GTAAGTGACT	
1259		959		0	-	0.41	CCCACCCAG	^	GTAGAGTGGC	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
120		1	+	0.14	CATTCTCAG	^	TTTCTCTGTT	
135		1	+	0.18	CTGTTTGAAG	^	AACATAAGGA	
143		2	+	0.17	AGAACATAAG	^	GACATTTCTC	
582		2	+	0.56	TGCCCTTTAG	^	GTTGATCACT	
631		1	+	0.25	TTTCTCCAG	^	TCTTCTCTAT	
1322		0	+	0.26	TGTGATCTAG	^	AAGGCACCAAG	

Acceptor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
2021		197		2	-	0.33	TCTGTTTCAG	^	ATCTACCACT	
1590		628		2	-	0.33	TGTTTTTCAG	^	GGCCAACAGC	
1370		848		0	-	0.97	TTTTGTACAG	^	CTTTCCAAGC	
1361		857		0	-	0.18	GCTTTCCAAG	^	CTGGCAGGCA	
1354		864		1	-	0.22	AAGCTGGCAG	^	GCAGCCTGAT	

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
238		0	+	0.83	CTTTGGAAG	^	GTAATTATAC	
555		0	+	0.34	CACCTCAGAG	^	GTTAGGGTTA	
921		0	+	0.37	TTATTATGAA	^	GTGAGTGGAC	
969		0	+	0.37	AAAGAGTCAG	^	GTAATTAAGG	
1559		1	+	0.58	TGAATCACAG	^	GTGAGGTGGA	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
2082		136		1	-	0.54	GTAGCAAAAT	^	GTAAGTGACT	
1259		959		0	-	0.41	CCCACCCAG	^	GTAGAGTGGC	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
120		1	+	0.14	CATTCTCAG	^	TTTCTCTGTT	
135		1	+	0.18	CTGTTTGAAG	^	AACATAAGGA	
143		2	+	0.17	AGAACATAAG	^	GACATTTCTC	
582		2	+	0.56	TGCCCTTTAG	^	GTTGATCACT	
631		1	+	0.25	TTTCTCCAG	^	TCTTCTCTAT	
1322		0	+	0.26	TGTGATCTAG	^	AAGGCACCAAG	

Acceptor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
2021		197		2	-	0.33	TCTGTTTCAG	^	ATCTACCACT	
1590		628		2	-	0.33	TGTTTTTCAG	^	GGCCAACAGC	
1370		848		0	-	0.97	TTTTGTACAG	^	CTTTCCAAGC	
1361		857		0	-	0.18	GCTTTCCAAG	^	CTGGCAGGCA	
1354		864		1	-	0.22	AAGCTGGCAG	^	GCAGCCTGAT	

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
231	245	0.99	tggaaag	gt aattat
548	562	0.82	ctcagag	gt taggg
914	928	0.93	ttatgaa	gt gagtg
962	976	0.99	gagtcag	gt aattaa
971	985	0.53	aattaag	gt gacagg
1078	1092	0.60	aattcag	gt gctagt
1509	1523	0.41	gccccta	gt aaggaa
1552	1566	0.94	atcacag	gt gaggtg

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
21	61	0.84	atgctttcttgtcattttt	ag agcagctcttgagtgttacag
100	140	0.97	cagctctttccattcctc	ag tttctctgtttgaagaacat
562	602	0.97	ttactattatctgcccttt	ag gttgatcactgggcactcag
611	651	0.97	accagtattcttttctctcc	ag tcttctctatatctgaaagc
1064	1104	0.81	gtccttattcttgtaattc	ag gtgctagtatcttatcaaaa
1321	1361	0.57	agaaggcaccagccttctc	ag gcatcaggctgcctgccagc
1779	1819	0.89	actattttattttcattta	ag tttgcaagaaatggtgttga

Donor site predictions for mut :

Start	End	Score	Exon	Intron
231	245	0.99	tggaaag	gt aattat
548	562	0.82	ctcagag	gt taggg
914	928	0.93	ttatgaa	gt gagtg
962	976	0.99	gagtcag	gt aattaa
971	985	0.53	aattaag	gt gacagg
1078	1092	0.60	aattcag	gt gctagt
1509	1523	0.41	gccccta	gt aaggaa
1552	1566	0.94	atcacag	gt gaggtg


Acceptor site predictions for mut :

Start	End	Score	Intron	Exon
21	61	0.84	atgctttcttgtcattttt	ag agcagctcttgagtgttacag
100	140	0.97	cagctctttccattcctc	ag tttctctgtttgaagaacat
562	602	0.97	ttactattatctgcccttt	ag gttgatcactgggcactcag
611	651	0.97	accagtattcttttctctcc	ag tcttctctatatctgaaagc
1064	1104	0.81	gtccttattcttgtaattc	ag gtgctagtatcttatcaaaa
1321	1361	0.57	agaaggcaccagccttctc	ag gcatcaggctgcctgccagc
1779	1819	0.89	actattttattttcattta	ag tttgcaagaaatggtgttga

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
atgat(c't)taaag	tctaaa	tttaaa	33183	94%

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	19	38	tcttttagaa		-5.13669095267	0.363636363636	33	0	0	-3.5808802		
wt	19	27	gtatgatct		-0.52129078373	0.363636363636	22	0	0	-1.0774513		
wt	19	22	atctaaaga		-0.438718495174	0.294117647059	17	0	0	-0.75108122		
mut	19	38	tcttttagaa		-5.13669095267	0.363636363636	33	0	0	-3.5808802		
mut	19	27	gtatgattt		-1.35364105627	0.363636363636	22	0	0	-1.4033557		
mut	19	23	gatttaaag		-3.64153004199	0.277777777778	18	0	0	-2.0737089		
mut	19	22	atttaaaga		-2.48917522367	0.294117647059	17	0	0	-1.553932		

Se producen cambios en dos de las predicciones por la presencia de la mutación, además de que aparece una nueva predicción de BP para la secuencia mutante. Sin embargo, todas están tienen puntuaciones negativas por lo que no se tendrán en cuenta.

Variant Effect Predictor tool

ENST00000370056.8:c.321+1896C>T	1:107873005-107873005	A	intron_variant	VAV3	ENSG00000134215	Transcript	ENST00000370056.9	protein_coding	-	-	-	-	-	-	COSV64067143
ENST00000370056.8:c.321+1896C>T	1:107873005-107873005	A	intron_variant, non_coding_transcript_variant	VAV3	ENSG00000134215	Transcript	ENST00000469325.5	processed_transcript	-	-	-	-	-	-	COSV64067143
ENST00000370056.8:c.321+1896C>T	1:107873005-107873005	A	intron_variant	VAV3	ENSG00000134215	Transcript	ENST00000490388.2	protein_coding	-	-	-	-	-	-	COSV64067143
ENST00000370056.8:c.321+1896C>T	1:107873005-107873005	A	intron_variant	VAV3	ENSG00000134215	Transcript	ENST00000527011.5	protein_coding	-	-	-	-	-	-	COSV64067143
ENST00000370056.8:c.321+1896C>T	1:107873005-107873005	A	intron_variant, non_coding_transcript_variant	VAV3	ENSG00000134215	Transcript	ENST00000530671.1	processed_transcript	-	-	-	-	-	-	COSV64067143

ESEfinder

Se obtienen dos resultados con puntuaciones positivas para las matrices 3':

2107 (-111)	tagctatctctttagaactgtatgatctaa	-25.13580	2107 (-111)	tagctatctctttagaactgtatgatctaa	4.64430	2107 (-111)	tagctatctctttagaactgtatgatctaa	-21.85170	2107 (-111)	tagctatctctttagaactgtatgatctaa	4.47290
2124 (-94)	ctgtatgatctaaagaaaacactgaaagtt	-21.10170	2124 (-94)	ctgtatgatctaaagaaaacactgaaagtt	1.31240	2124 (-94)	ctgtatgatctaaagaaaacactgaaagtt	-17.59660	2124 (-94)	ctgtatgatctaaagaaaacactgaaagtt	1.33730

Cuando comparamos estas puntuaciones con las de la secuencia mutante se observa que una ha descendido y la otra aumentado, lo que es un poco contradictorio y no podemos dar ninguna resolución concreta aunque lo más probable por la diferencia de puntuaciones es que se estuviera debilitando un sitio *acceptor*:

2107 (-111)	tagctatctctttagaactgtatgatttaa	-25.24110	2107 (-111)	tagctatctctttagaactgtatgatttaa	4.08750	2107 (-111)	tagctatctctttagaactgtatgatttaa	-22.08030	2107 (-111)	tagctatctctttagaactgtatgatttaa	3.95230
2124 (-94)	ctgtatgatttaaagaaaacactgaaagtt	-21.61190	2124 (-94)	ctgtatgatttaaagaaaacactgaaagtt	1.42130	2124 (-94)	ctgtatgatttaaagaaaacactgaaagtt	-18.04890	2124 (-94)	ctgtatgatttaaagaaaacactgaaagtt	1.36780