Ejemplo comparación de resultados predictores in sillico

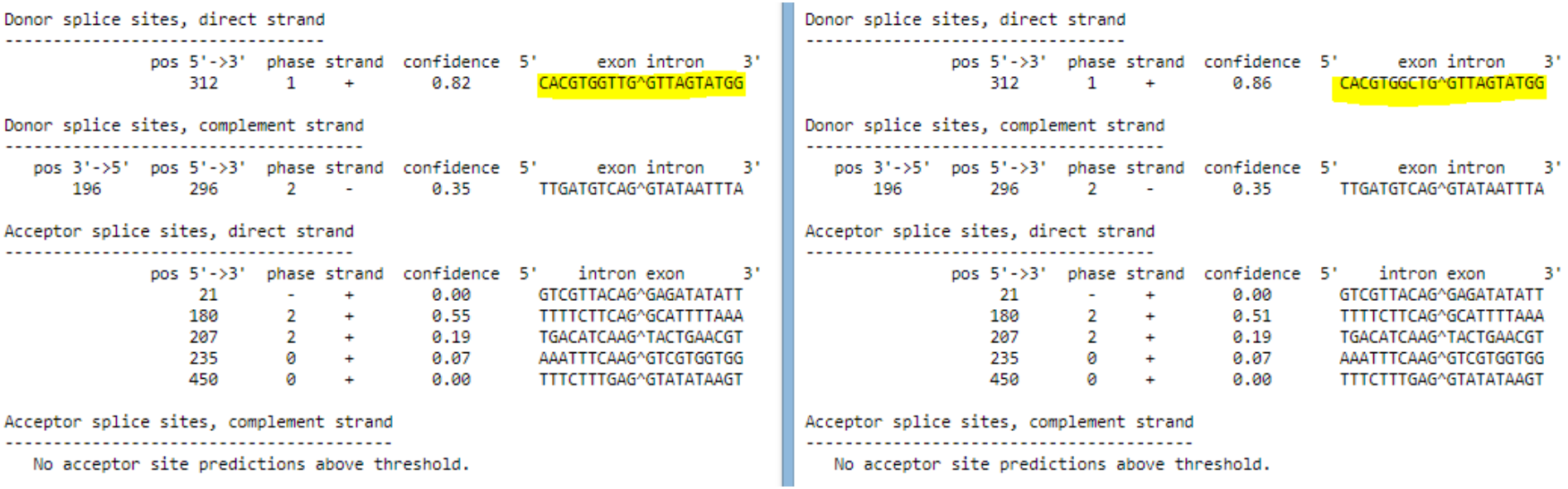
Cambio de estudio ADAM17 c.359T>C (chr2:9536700 T/C, COSV60397784 o NM_003183.6: c.359T>C)

Exón 3 e intrones adyacentes:

El cambio se encuentra en la última línea del exón 3 (la **t** 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'
                   312       1     +        0.82       CACGTGGTTG^GTTAGTATGG

Donor splice sites, complement strand
-------------------------------------
   pos 3'->5'  pos 5'->3'  phase strand  confidence  5'      exon intron    3'
       196         296       2     -        0.35       TTGATGTCAG^GTATAATTTA

Acceptor splice sites, direct strand
------------------------------------
               pos 5'->3'  phase strand  confidence  5'    intron exon      3'
                    21       -     +        0.00       GTCGTTACAG^GAGATATATT
                   180       2     +        0.55       TTTTCTTCAG^GCATTTTAAA
                   207       2     +        0.19       TGACATCAAG^TACTGAACGT
                   235       0     +        0.07       AAATTTCAAG^GTCGTGGTGG
                   450       0     +        0.00       TTTCTTTGAG^GTATATAAGT

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'
                   312       1     +        0.86       CACGTGGCTG^GTTAGTATGG

Donor splice sites, complement strand
-------------------------------------
   pos 3'->5'  pos 5'->3'  phase strand  confidence  5'      exon intron    3'
       196         296       2     -        0.35       TTGATGTCAG^GTATAATTTA

Acceptor splice sites, direct strand
------------------------------------
               pos 5'->3'  phase strand  confidence  5'    intron exon      3'
                    21       -     +        0.00       GTCGTTACAG^GAGATATATT
                   180       2     +        0.51       TTTTCTTCAG^GCATTTTAAA
                   207       2     +        0.19       TGACATCAAG^TACTGAACGT
                   235       0     +        0.07       AAATTTCAAG^GTCGTGGTGG
                   450       0     +        0.00       TTTCTTTGAG^GTATATAAGT

Acceptor splice sites, complement strand
----------------------------------------
   No acceptor site predictions above threshold.
```

La presencia de la mutación produce una alteración en el sitio *donor* predicho (en amarillo) para la secuencia WT, teniendo mayor confianza para la secuencia mutante. Sin embargo como este sitio ya se estaba utilizando en el *splicing*, la mutación no tendrá un efecto significativo en el mensajero resultante.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
305	319	0.63	gtggttg	gttagtat
444	458	0.85	ctttgag	gtatataa

Donor site predictions for mut :

Start	End	Score	Exon	Intron
305	319	0.86	gtggctg	gttagtat
444	458	0.85	ctttgag	gtatataa

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
1	41	0.95	ttagttcatttgtcgttac	aggagatatattgctctaagta
160	200	0.85	tatgaatgcatttttcttc	aggcattttaaattatacctga
336	376	0.49	cccctgtgtctggtgtatt	aggtctccaaatctaacttgat
378	418	0.86	gcatattcttttctggtgc	aggatagtggggacggggtgtc

Acceptor site predictions for mut :

Start	End	Score	Intron	Exon
1	41	0.95	ttagttcatttgtcgttac	aggagatatattgctctaagta
160	200	0.85	tatgaatgcatttttcttc	aggcattttaaattatacctga
336	376	0.49	cccctgtgtctggtgtatt	aggtctccaaatctaacttgat
378	418	0.86	gcatattcttttctggtgc	aggatagtggggacggggtgtc

La presencia de la mutación produce una alteración en el sitio *donor* predicho (en amarillo) para la secuencia WT, teniendo mayor *score* para la secuencia mutante. Sin embargo como este sitio ya se estaba utilizando en el *splicing*, la mutación no tendrá un efecto significativo en el mensajero resultante.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
cgtgg(t/c)tggtt	ggttgg	ggctgg	31388	83%

Human Splicing Finder

New Donor splice site	Activation of a cryptic Donor site. Potential alteration of splicing

Algorithm/Matix	position	sequences	variation
MaxEnt Donor site	chr2:9536700	- REF : TTGGTTAGT - ALT : CTGGTTAGT	4.82 > 6.32 => 31.12%

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr
wt	40	49	ctgtaaaat		-0.472832092723	0.454545454545	11	10	21	-0.13720216
wt	40	31	tcttcactg		0.165684009896	0.423076923077	26	0	0	-1.0424639
mut	40	49	ctgtaaaat		-0.472832092723	0.454545454545	11	10	21	-0.13720216
mut	40	31	tcttcactg		0.165684009896	0.423076923077	26	0	0	-1.0424639

Variant Effect Predictor tool

ENST00000310823.8:c.359T>C	2:9536700-9536700	G	missense_variant, splice_region_variant	ADAM17	ENSG00000151694	Transcript	ENST00000310823.8	protein_coding	3/19	584	359	120	V/A	GTT/GCT	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, non_coding_transcript_exon_variant	ADAM17	ENSG00000151694	Transcript	ENST00000478059.1	retained_intron	3/5	528	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	missense_variant, splice_region_variant, NMD_transcript_variant	ADAM17	ENSG00000151694	Transcript	ENST00000618923.2	nonsense_mediated_decay	3/8	546	359	120	V/A	GTT/GCT	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	intron_variant, NMD_transcript_variant	ADAM17	ENSG00000151694	Transcript	ENST00000647610.1	nonsense_mediated_decay	-	-	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, non_coding_transcript_exon_variant	ADAM17	ENSG00000151694	Transcript	ENST00000647622.1	retained_intron	3/12	538	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	intron_variant, NMD_transcript_variant	ADAM17	ENSG00000151694	Transcript	ENST00000647979.1	nonsense_mediated_decay	-	-	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	missense_variant, splice_region_variant	ADAM17	ENSG00000151694	Transcript	ENST00000648002.1	protein_coding	2/4	248	248	83	V/A	GTT/GCT	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, non_coding_transcript_exon_variant	ADAM17	ENSG00000151694	Transcript	ENST00000648548.1	processed_transcript	4/20	566	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, non_coding_transcript_exon_variant	ADAM17	ENSG00000151694	Transcript	ENST00000648857.1	processed_transcript	3/19	404	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, 3_prime_UTR_variant, NMD_transcript_variant	ADAM17	ENSG00000151694	Transcript	ENST00000649227.1	nonsense_mediated_decay	3/19	483	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, non_coding_transcript_exon_variant	ADAM17	ENSG00000151694	Transcript	ENST00000649686.1	processed_transcript	2/15	149	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, non_coding_transcript_exon_variant	ADAM17	ENSG00000151694	Transcript	ENST00000649972.1	retained_intron	4/13	597	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, 3_prime_UTR_variant, NMD_transcript_variant	ADAM17	ENSG00000151694	Transcript	ENST00000650116.1	nonsense_mediated_decay	3/19	357	-	-	-	-	COSV60397784
ENST00000310823.8:c.359T>C	2:9536700-9536700	G	splice_region_variant, non_coding_transcript_exon_variant	ADAM17	ENSG00000151694	Transcript	ENST00000650241.1	retained_intron	3/14	302	-	-	-	-	COSV60397784

ESEfinder

Se obtienen 4 resultados con puntuaciones positivas para las matrices 5’ (290, 293, 297 y 301):

290 (-202)	TTCTTCACTGGACACGTGGTTGgttagtat	0.67910	290 (-202)	TTCTTCACTGGACACGTGGTTGgttagtat	-4.18170	290 (-202)	TTCTTCACTGGACACGTGGTTGgttagtat	0.13170	290 (-202)	TTCTTCACTGGACACGTGGTTGgttagtat	-4.32780
293 (-199)	TTCACTGGACACGTGGTTGgttagtatgga	1.99570	293 (-199)	TTCACTGGACACGTGGTTGgttagtatgga	-16.15810	293 (-199)	TTCACTGGACACGTGGTTGgttagtatgga	1.57140	293 (-199)	TTCACTGGACACGTGGTTGgttagtatgga	-17.85770
297 (-195)	CTGGACACGTGGTTGgttagtatggagctt	4.98040	297 (-195)	CTGGACACGTGGTTGgttagtatggagctt	-11.96190	297 (-195)	CTGGACACGTGGTTGgttagtatggagctt	5.08900	297 (-195)	CTGGACACGTGGTTGgttagtatggagctt	-13.18810
301 (-191)	ACACGTGGTTGgttagtatggagcttgttg	1.93890	301 (-191)	ACACGTGGTTGgttagtatggagcttgttg	-25.41470	301 (-191)	ACACGTGGTTGgttagtatggagcttgttg	1.83790	301 (-191)	ACACGTGGTTGgttagtatggagcttgttg	-29.85810

Cuando se comparan estas puntuaciones con las equivalentes en la secuencia mutante se observa que 290, 297 y 301 han aumentado sus puntuaciones mientras que 293 ha pasado tener puntuación negativa esto supone resultados contradictorios porque por un lado se estaría fortaleciendo un sitio *donor* mientras que por otro se estaría perdiendo:

290 (-202)	TTCTTCACTGGACACGTGGCTGgttagtat	0.62500	290 (-202)	TTCTTCACTGGACACGTGGCTGgttagtat	-3.81600	290 (-202)	TTCTTCACTGGACACGTGGCTGgttagtat	0.29890	290 (-202)	TTCTTCACTGGACACGTGGCTGgttagtat	-3.99580
293 (-199)	TTCACTGGACACGTGGCTGgttagtatgga	-4.94040	293 (-199)	TTCACTGGACACGTGGCTGgttagtatgga	-16.62150	293 (-199)	TTCACTGGACACGTGGCTGgttagtatgga	-5.53870	293 (-199)	TTCACTGGACACGTGGCTGgttagtatgga	-18.32320
297 (-195)	CTGGACACGTGGCTGgttagtatggagctt	7.13120	297 (-195)	CTGGACACGTGGCTGgttagtatggagctt	-10.37360	297 (-195)	CTGGACACGTGGCTGgttagtatggagctt	7.19090	297 (-195)	CTGGACACGTGGCTGgttagtatggagctt	-11.57440
301 (-191)	ACACGTGGCTGgttagtatggagcttgttg	2.42120	301 (-191)	ACACGTGGCTGgttagtatggagcttgttg	-25.40410	301 (-191)	ACACGTGGCTGgttagtatggagcttgttg	2.34010	301 (-191)	ACACGTGGCTGgttagtatggagcttgttg	-29.78970

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

303 (-189)	ACGTGGT	-1.29601	303 (-189)	ACGTGGT	-0.89232	303 (-189)	ACGTGGTT	-3.94807	303 (-189)	ACGTGGT	-4.65494
304 (-188)	CGTGGTT	-2.15411	304 (-188)	CGTGGTT	-0.06841	304 (-188)	CGTGGTTG	-1.47062	304 (-188)	CGTGGTT	-4.58428
305 (-187)	GTGGTTG	-4.63609	305 (-187)	GTGGTTG	-3.82003	305 (-187)	GTGGTTGg	-4.00999	305 (-187)	GTGGTTG	-3.09338
306 (-186)	TGGTTGg	-2.75623	306 (-186)	TGGTTGg	-1.69491	306 (-186)	TGGTTGgt	-4.39112	306 (-186)	TGGTTGg	0.26370
307 (-185)	GGTTGgt	-1.15368	307 (-185)	GGTTGgt	-1.36411	307 (-185)	GGTTGgtt	1.57873	307 (-185)	GGTTGgt	-5.36969
308 (-184)	GTTGgtt	-4.40672	308 (-184)	GTTGgtt	-3.69608	308 (-184)	GTTGgtta	-0.53968	308 (-184)	GTTGgtt	-6.22243
309 (-183)	TTGgtta	-5.28249	309 (-183)	TTGgtta	-3.59151	309 (-183)	TTGgttag	-4.76274	309 (-183)	TTGgtta	-2.97078

303 (-189)	ACGTGGC	-3.14941	303 (-189)	ACGTGGC	-2.08902	303 (-189)	ACGTGGCT	-4.59251	303 (-189)	ACGTGGC	-3.11584
304 (-188)	CGTGGCT	-2.61182	304 (-188)	CGTGGCT	-0.17241	304 (-188)	CGTGGCTG	0.25839	304 (-188)	CGTGGCT	-2.86735
305 (-187)	GTGGCTG	-2.11236	305 (-187)	GTGGCTG	-2.00823	305 (-187)	GTGGCTGg	-3.25075	305 (-187)	GTGGCTG	-4.58464
306 (-186)	TGGCTGg	-1.29891	306 (-186)	TGGCTGg	-0.44015	306 (-186)	TGGCTGgt	-4.16028	306 (-186)	TGGCTGg	2.64833
307 (-185)	GGCTGgt	1.15633	307 (-185)	GGCTGgt	0.70475	307 (-185)	GGCTGgtt	2.14766	307 (-185)	GGCTGgt	-6.71796
308 (-184)	GCTGgtt	-5.00965	308 (-184)	GCTGgtt	-3.94487	308 (-184)	GCTGgtta	-1.92809	308 (-184)	GCTGgtt	-5.91038
309 (-183)	CTGgtta	-2.33200	309 (-183)	CTGgtta	-0.48646	309 (-183)	CTGgttag	-4.74162	309 (-183)	CTGgtta	-3.32748

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	5	5	3	23	367.5158	28	-35.0508	8	10	45	493.3854	40	50.7204	64	103	0.62
mut	5	4	3	23	350.5170	26	-33.1891	8	10	45	493.3854	41	52.2456	61	104	0.59

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

HOT-SKIP

ttagttcatttgtcgttacaggagatatattgctctaagtacctaataagactttagttgtaacttgcatagtgttttat
taatgtcaatgcttctggctttttaagaaagtttaattatgtgaaaatgagacttaaaatttgtgtgcttagagtaaaat
atgaatgcatttttcttcagGCATTTTAAATTATACCTGACATCAAGTACTGAACGTTTTTCACAAAATTTCAAGGTCGT
GGTGGTGGATGGTAAAAACGAAAGCGAGTACACTGTAAAATGGCAGGACTTCTTCACTGGACACGTGGTTGgttagtatg
gagcttgttggtttgcccctgtgtctggtgtattaggtctccaaatctaacttgattgcatattcttttctggtgcagga
tagtggggacggggtgtcatagaagtgctattctttgacttttctttgaggtatataagtttagaggcaggaacattcta
atagaaggaaa