

Ejemplo comparación de resultados predictores in silico

Cambio de estudio GPX5 c.242-3C>T (chr6:28532318 C/T, rs1454730768 o NM_001509.3: c.242-3C>T)

Exón 3 e intrones adyacentes:

```
agtttaacttttggcctaccggtgatgagcggtctggcagtttccagactctctctccatccc  
tggctgagaatttgatggactatggaggacagaagggatggaggaaaagagacaagtc  
gggtagagagaaataaagtgagtcagagagccaaagtcataatgcataattcctc  
cccacccacaaacgatacttccctctttctccctttctcctgacttcaaatacctagtgtc  
ccattataatctttacttgcatatcctggagcttcctgggaacattatggcttgttgcag  
GTATGTCCTCCAGGGGGAGGATTGTACCTAGTTTCCAGCTTTTGGAGAAAGGGGATGT  
GAATGGTGA AAAAGAACAGAAAGTCTCAGTTTCTTG AAG  
gtgagtataattcctggcataagcctccctcctcttttctaataattgctaacaatagagttgg  
tgtggcagaaaatggatccaagggtctatgcccagaggtgggattggctttgtggagaaa  
aatctccagatcagcattccacattgcacatgctgatctggaattaatctgtgagtatc  
tagggggtctctaaactctgtgtgtggaagaaagtgtgtggaatggcaacttttactagga  
ccacccctcttttgctcattgacattcattcattcaggtacacttgatgaat
```

El cambio se encuentra en la tercera posición justo antes del exón 3 (la **c** en color naranja).

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'
139		0	+	0.35	GAGAAATAAA	GTGAGTCAGA		
301		2	+	0.34	CTTGTTGCAG	GTATGTCCGT		
401		0	+	0.89	TTTCTTGAAG	GTGAGTAAAT		
686		-	+	0.00	TTCATTCCAG	GTACACTTGA		

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
18	683	-	-	0.00	GCTCATCACG	GTAGGCCAAA		

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
300		2	+	0.26	CTTGTTGCAG	GTATGTCCGT		

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
596	105	1	-	0.25	CACCACACAG	AGTTAGAGCC		
80	621	2	-	0.00	TCCTCCATAG	TCCATCCAAG		

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
139		0	+	0.35	GAGAAATAAA	GTGAGTCAGA		
301		2	+	0.41	CTTGTTGTAG	GTATGTCCGT		
401		0	+	0.89	TTTCTTGAAG	GTGAGTAAAT		
686		-	+	0.00	TTCATTCCAG	GTACACTTGA		

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
329	372	2	-	0.31	TGGAACTAG	GTACAAATCC		
18	683	-	-	0.00	GCTCATCACG	GTAGGCCAAA		

Acceptor splice sites, direct strand

No acceptor site predictions above threshold.

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
596	105	1	-	0.25	CACCACACAG	AGTTAGAGCC		
80	621	2	-	0.00	TCCTCCATAG	TCCATCCAAG		

Se está viendo alterado uno de los sitios *donor* (en azul) predichos de la secuencia WT a la mutante. Sin embargo este sitio no corresponde al *donor* del exón, ya que corresponde al sitio *acceptor*, por lo tanto, no se va a tener en cuenta.

Por otro lado se pierde el sitio *acceptor* (en rojo) predicho de la secuencia WT en la secuencia mutante y este sitio corresponde al *acceptor* del exón, por lo tanto, va a haber un cambio en el *splicing*.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
132	146	0.99	aaataaa	gtgagtca
294	308	1.00	gttcag	gtatgtcc
394	408	0.99	cttgaag	gtgagtaa
490	504	0.63	cccagag	gtgggatt
565	579	0.48	ttaatct	gtgagtat
679	693	0.92	attccag	gtacactt

Donor site predictions for mut :

Start	End	Score	Exon	Intron
132	146	0.99	aaataaa	gtgagtca
294	308	0.99	gtttag	gtatgtcc
394	408	0.99	cttgaag	gtgagtaa
490	504	0.63	cccagag	gtgggatt
565	579	0.48	ttaatct	gtgagtat
679	693	0.92	attccag	gtacactt

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
280	320	0.83	gaacattatggccttggtgc	aggtatgtccgtccagggggag

Acceptor site predictions for mut :

Start	End	Score	Intron	Exon
280	320	0.83	gaacattatggccttggtgta	aggtatgtccgtccagggggag


Se ve alterado uno de los sitios *donor* (en azul) de la secuencia WT a la secuencia mutante. Sin embargo este no es un sitio *donor* en la realidad, por lo que no se te va a tener en cuenta.

Por otro lado, se ve modificado uno de los sitios *acceptor* (en rojo) de la secuencia WT a la secuencia mutante. Sin embargo, su *score* sigue siendo el mismo, por lo que es probable que, a pesar de la presencia de la mutación, no se vea alterado el *splicing*.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
tgttg(c/t)aggtta	tgcagg	tgtagg	29334	70%

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	62	acattatgg		-2.59233579663	0.543859649123	17	9	13	-1.3925588		
mut	15	62	acattatgg		-2.59233579663	0.543859649123	17	9	13	-1.3925588		

Variant Effect Predictor tool

ENST00000469384.1:c.242-3C>T	6:28532318-28532318	T	splice_region_variant, intron_variant	GPX5	ENSG00000224586	Transcript	ENST00000412168.7	protein_coding	-	-	-	-	-	-	rs1454730768 , COSV101297260
ENST00000469384.1:c.242-3C>T	6:28532318-28532318	T	splice_region_variant, intron_variant, non_coding_transcript_variant	GPX5	ENSG00000224586	Transcript	ENST00000442674.6	processed_transcript	-	-	-	-	-	-	rs1454730768 , COSV101297260
ENST00000469384.1:c.242-3C>T	6:28532318-28532318	T	splice_region_variant, intron_variant	GPX5	ENSG00000224586	Transcript	ENST00000469384.1	protein_coding	-	-	-	-	-	-	rs1454730768 , COSV101297260
ENST00000469384.1:c.242-3C>T	6:28532318-28532318	T	upstream_gene_variant	GPX6	ENSG00000198704	Transcript	ENST00000483058.1	processed_transcript	-	-	-	-	-	-	rs1454730768 , COSV101297260
ENST00000469384.1:c.242-3C>T	6:28532318-28532318	T	splice_region_variant, intron_variant, non_coding_transcript_variant	GPX5	ENSG00000224586	Transcript	ENST00000483784.1	processed_transcript	-	-	-	-	-	-	rs1454730768 , COSV101297260

ESEfinder

Se encuentra una única predicción con puntuaciones positivas para todas las matrices. Cuando se comparan estas puntuaciones con las equivalentes en la secuencia mutante se observa que todas han disminuido, por lo que esto debería significar que se debilita un sitio *acceptor* y *donor* a la vez y no tiene mucho sentido.

286 (-415)	tatggcttggtgcagGTATGTCGTCAGG	9.30870	286 (-415)	tatggcttggtgcagGTATGTCGTCAGG	7.69050	286 (-415)	tatggcttggtgcagGTATGTCGTCAGG	9.55540	286 (-415)	tatggcttggtgcagGTATGTCGTCAGG	7.39660
286 (-415)	tatggcttggttagGTATGTCGTCAGG	7.15790	286 (-415)	tatggcttggttagGTATGTCGTCAGG	6.10220	286 (-415)	tatggcttggttagGTATGTCGTCAGG	7.45350	286 (-415)	tatggcttggttagGTATGTCGTCAGG	5.78290

Por otro lado, se encuentra una predicción que tiene puntuaciones negativas para la secuencia *wild type*, pero positivas para las matrices 5' en la secuencia mutante, por lo que podría estar activándose un sitio de *donor*.

282 (-419)	acattatggcttggtgcagGTATGTCGTC	-4.12700	282 (-419)	acattatggcttggtgcagGTATGTCGTC	-27.54660	282 (-419)	acattatggcttggtgcagGTATGTCGTC	-4.16240	282 (-419)	acattatggcttggtgcagGTATGTCGTC	-30.24120
282 (-419)	acattatggcttggttagGTATGTCGTC	2.80910	282 (-419)	acattatggcttggttagGTATGTCGTC	-27.08320	282 (-419)	acattatggcttggttagGTATGTCGTC	2.94770	282 (-419)	acattatggcttggttagGTATGTCGTC	-29.77570