

Ejemplo comparación de resultados predictores in silico

Cambio de estudio GFPT2 c.2004+4C>T (chr5:180302419 C/T, rs373053816 o NM_005110.4: c.2004+4C>T)

Exón 18 e intrones adyacentes:

```
agcccgctagagttaggagtgagtctgcccgcctctcggaaggcacaatgggtgcc  
aatgatgcagcaggatggcaatctccttaacaggaaggggaagaaccccaaaagca  
tcagacactttatcctttccaagcatttgtgtcccacacataccaactcacccatga  
cagcgaatgcaaaagaggactgttctcatcacagtaatgtagcttcagatactctgtat  
gcataagggatagctttctttcagaggggtctattttatgatgctatttctctgctctag  
GGTCGCCCAATTATACTGTGCTCCAAGGACGATAC TGAAAGTTCCAAGTTTTCGTATAAG  
ACAATTGAGCTGCCCAACTGTGAC TGCCTCCAGGGCATCTGAGCGTGATTCCGCTG  
CAGCTGCTGTCCTTCCACCTGGCTGTCTCCGAGGATATGAC  
gtgcgtctaaaaactgcacctaaccagcagagaggagagacataacccagagttccttc  
tgttctttaataacttgagtaaatagtcatagagtaataattttattttaagtagctt  
tcttttttttttttttgagattgagtctcgctctgtcgccaggtcgagggcagtggca  
caaccgcggctcactgaagctcacctcctgggttcaagcgattctcctgcctcagcct  
cccgagttagctgggaactacagcacccgcaacatgctggctaattttttgtaattttt
```

El cambio se encuentra en la cuarta posición justo después del exón 18 (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'
20	-	+		0.00	AGAGTTAGGA	^	GTGAGTCTGC	
463	0	+		0.82	AGGATATGAC	^	GTGCGTCTAA	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
171		592		0	-	0.34	GGAGTCATGG	^	GTGAGGTGGT	
163		600		2	-	0.37	GGGTGAGGTG	^	GTATGTGTGG	
90		673		1	-	0.00	CCCTTCCCTG	^	GTAAGGAGAT	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
300	0	+		0.96	TCTGCTCTAG	^	GGTCGCCCCA	
327	0	+		0.19	GTGCTCCAAG	^	GACGATACTG	
341	2	+		0.17	ATACTGAAAG	^	TTCCAAGTTT	
348	0	+		0.07	AAGTTCCAAG	^	TTTGCGTATA	

Acceptor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
482		281		2	-	0.76	CTGCTGTTAG	^	GTGCAGTTTT	
476		287		2	-	0.18	TTAGGTGCAG	^	TTTTTAGACG	
469		294		0	-	0.19	CAGTTTTT	^	ACGCACGTCA	
448		315		0	-	0.25	ATCCTCGGAG	^	AACAGCCAGG	
443		320		2	-	0.25	CGGAGAACAG	^	CCAGGTGGAA	
439		324		0	-	0.19	GAACAGCCAG	^	GTGGAAGGAC	
432		331		1	-	0.18	CAGGTGGAAG	^	GACAGCAGCT	
427		336		0	-	0.07	GGAAGGACAG	^	CAGCTGCAGC	

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
20	-	+		0.00	AGAGTTAGGA	^	GTGAGTCTGC	
463	0	+		0.38	AGGATATGAC	^	GTGTGTCTAA	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
171		592		0	-	0.34	GGAGTCATGG	^	GTGAGGTGGT	
163		600		2	-	0.37	GGGTGAGGTG	^	GTATGTGTGG	
90		673		1	-	0.00	CCCTTCCCTG	^	GTAAGGAGAT	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
300	0	+		0.96	TCTGCTCTAG	^	GGTCGCCCCA	
327	0	+		0.19	GTGCTCCAAG	^	GACGATACTG	
341	2	+		0.17	ATACTGAAAG	^	TTCCAAGTTT	
348	0	+		0.07	AAGTTCCAAG	^	TTTGCGTATA	

Acceptor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
482		281		0	-	0.53	CTGCTGTTAG	^	GTGCAGTTTT	
476		287		0	-	0.17	TTAGGTGCAG	^	TTTTTAGACA	
469		294		0	-	0.18	CAGTTTTT	^	ACACACGTCA	
448		315		0	-	0.20	ATCCTCGGAG	^	AACAGCCAGG	
443		320		2	-	0.20	CGGAGAACAG	^	CCAGGTGGAA	
439		324		0	-	0.19	GAACAGCCAG	^	GTGGAAGGAC	
432		331		1	-	0.18	CAGGTGGAAG	^	GACAGCAGCT	
427		336		0	-	0.07	GGAAGGACAG	^	CAGCTGCAGC	

La mutación provoca un cambio en el sitio *donor* (en azul), que corresponde con el *donor* del exón, el cual se sigue prediciendo, pero su confianza ha disminuido, lo que puede significar que este sitio se esté debilitando. Un debilitamiento del sitio de *donor* podría provocar que el complejo del *splicing* no lo reconociera, llevando un cambio en el mensajero.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
13	27	0.67	gtagga	gtgagtct
456	470	0.80	atatgac	gtggtgtct
550	564	0.50	tcataga	gtaaataa

Donor site predictions for mut :

Start	End	Score	Exon	Intron
13	27	0.67	gtagga	gtgagtct
456	470	0.54	atatgac	gtggtgtct
550	564	0.50	tcataga	gtaaataa

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
280	320	0.92	gatgctatttctctgctct	agggtcgccccattatactgtg
376	416	0.45	cacactgtggactgcctcc	aggcatcctgagcgtgattcc
581	621	0.97	tttctttttttttttttg	agattgagtctcgctctgtcgc
605	645	0.54	gagtctcgctctgtcgccc	aggctggagggcagtggcacaa
679	719	0.55	caagcgattctcctgcctc	aggctcccgagtagctgggact

Acceptor site predictions for mut :


Start	End	Score	Intron	Exon
280	320	0.92	gatgctatttctctgctct	agggtcgccccattatactgtg
376	416	0.45	cacactgtggactgcctcc	aggcatcctgagcgtgattcc
581	621	0.97	tttctttttttttttttg	agattgagtctcgctctgtcgc
605	645	0.54	gagtctcgctctgtcgccc	aggctggagggcagtggcacaa
679	719	0.55	caagcgattctcctgcctc	aggctcccgagtagctgggact

La mutación provoca un cambio en el sitio *donor* (en azul), que corresponde con el *donor* del exón, el cual se sigue prediciendo, pero su confianza ha disminuido, lo que puede significar que este sitio se esté debilitando. Un debilitamiento del sitio de *donor* podría provocar que el complejo del *splicing* no lo reconociera, llevando un cambio en el mensajero.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
acgtg(c/t)gtcta	gtgcgt	gtgtgt	31058	81%

Human Splicing Finder

 Broken WT Donor Site	Alteration of the WT Donor site, most probably affecting splicing		
Algorithm/Matix	position	sequences	variation
MaxEnt Donor site	chr5:180302425	- REF : GACGTG C GT - ALT : GACGTG T GT	6.63 > 3.49 => -47.36%

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr	
wt	37	31	atatgacgt		0.608627490222	0.5	26	0	0	-0.8441867
wt	37	20	gtctaaaaa		0.0230418106475	0.466666666667	15	0	0	-0.38795533
mut	37	31	atatgacgt		0.608627490222	0.5	26	0	0	-0.8441867
mut	37	20	gtctaaaaa		0.0230418106475	0.466666666667	15	0	0	-0.38795533

Variant Effect Predictor tool

ENST00000253778.12:c.2004+4C>T	5:180302419-180302419	A	splice_region_variant, intron_variant	GFPT2	ENSG00000131459	Transcript	ENST00000253778.13	protein_coding	-	-	-	-	-	-	rs373053816
--------------------------------	---------------------------------------	---	---------------------------------------	-------	---------------------------------	------------	------------------------------------	----------------	---	---	---	---	---	---	-----------------------------

ESEfinder

Se han encontrado dos predicciones con puntuaciones positivas: una para las matrices 3' (440) y otra para las matrices 5' (448).

440 (-323)	TGGCTGTTCTCCGAGGATATGACgtggtc	-13.24290	440 (-323)	TGGCTGTTCTCCGAGGATATGACgtggtc	2.06600	440 (-323)	TGGCTGTTCTCCGAGGATATGACgtggtc	-9.38710	440 (-323)	TGGCTGTTCTCCGAGGATATGACgtggtc	1.58660
448 (-315)	CTCCGAGGATATGACgtggtctaaaaact	2.48320	448 (-315)	CTCCGAGGATATGACgtggtctaaaaact	-16.19300	448 (-315)	CTCCGAGGATATGACgtggtctaaaaact	2.06720	448 (-315)	CTCCGAGGATATGACgtggtctaaaaact	-17.08050

Si comparamos estas puntuaciones con las equivalentes en la secuencia mutante, se observa que para 440 la puntuación ha disminuido, mientras que para el 448 aumentado ligeramente.

440 (-323)	TGGCTGTTCTCCGAGGATATGACgtgtgtc	-13.34820	440 (-323)	TGGCTGTTCTCCGAGGATATGACgtgtgtc	1.50920	440 (-323)	TGGCTGTTCTCCGAGGATATGACgtgtgtc	-9.61570	440 (-323)	TGGCTGTTCTCCGAGGATATGACgtgtgtc	1.06600
448 (-315)	CTCCGAGGATATGACgtgtgtctaaaaact	2.68680	448 (-315)	CTCCGAGGATATGACgtgtgtctaaaaact	-16.65710	448 (-315)	CTCCGAGGATATGACgtgtgtctaaaaact	2.18630	448 (-315)	CTCCGAGGATATGACgtgtgtctaaaaact	-17.52890

Por lo tanto lo más probable es que se estuviera debilitando un sitio *acceptor*.