

Ejemplo comparación de resultados predictores in silico

Cambio de estudio PIR c.273+1G>A (chr23:15459656 G/A, COSV62916238 o NM_001018109.3:c.273+1G>A)

Exón 4 e intrones adyacentes:

```
catgtagctttcatttgcatttttctgacaatggttgactaccttttcataatacctggtg
gccatttgtatgtcttccttgggaaagtgtctattcaagtcctttgcttatttttttttt
ttttttaatgaaaaggcaactcaggaatgggagaaaatgtttgcaaaccatattatcca
ataagggattaagttccaaaatatacaaggaaactcatacaagtcagtagggataattaaa
taataacccaattttaaaaaataggcaaaaggatctactttttttcctgtttgtttttcaag
GTATCCTACCTCCTGGAA GGGGCGAGCATGGCCCATGAAGACTCTGTGGACACACTGCT
AAAATGAACCCAGGAGATTTGCAG
gtatggccaaggacaagggaattgctgggtggttttagtagtttttgggtgdcggtgcttgat
cctgcccacttccagtagattcaccctttccacactctctctctattatattcttgactttt
ccttttctacagattcttttactcagcctggaaaccacacaggtctcttatattcttctt
tcacactttcttgacccctcaatcaacttttagctcccacctaagcttttttcttacatt
tctcaaaagagtaatcagggaatgatatcccaacttccttatcttcaattcgttttgtact
```

El cambio se encuentra en primera posición justo después del exón 4 (la **g** 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'
359		1	+	0.55	GGACACACTG	^	GTAAAATGAA	
385		0	+	0.83	AGATTTGCAG	^	GTATGGCCAA	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
309		376		0	-	0.47	CTTCCAGGAG	^	GTAGGATACC	
144		541		1	-	0.37	CTCCATTCC	^	GTGAGTTGCC	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
300		0	+	0.97	TTTTTCAAAG	^	GTATCCTACC	
319		1	+	0.17	CTCCTGGAAG	^	GGGGCAGCAT	
516		1	+	0.33	TTTCTACAG	^	ATTCTTTTAC	

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
359		1	+	0.46	GGACACACTG	^	GTAAAATGAA	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
309		376		0	-	0.46	CTTCCAGGAG	^	GTAGGATACC	
144		541		1	-	0.37	CTCCATTCC	^	GTGAGTTGCC	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
300		0	+	0.97	TTTTTCAAAG	^	GTATCCTACC	
516		1	+	0.33	TTTCTACAG	^	ATTCTTTTAC	

Desaparecen en la secuencia mutante tanto un sitio *donor* (en rojo) como un *acceptor* (en verde). El sitio *donor* coincide con el sitio *donor* del exón 4, mientras que el *acceptor* se encuentra dentro del exón 4, por lo que no se tendrá en cuenta. En conclusión, se pierde el sitio *donor* del exón 4, lo que llevará a cambios en el *splicing*.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 85.53.15.54.10054.0 :

Start	End	Score	Exon	Intron
294	308	0.42	ttcaaag	gtatccta
352	366	0.83	cacactg	gtaaaatg
378	392	0.96	tttgcag	gtatggcc

Donor site predictions for 85.53.15.54.10085.0 :

Start	End	Score	Exon	Intron
294	308	0.42	ttcaaag	gtatccta
352	366	0.83	cacactg	gtaaaatg

Acceptor site predictions for 85.53.15.54.10054.0 :

Start	End	Score	Intron	Exon
280	320	0.98	tttcctgtttgtttttcaa	aggtatcctacctcctggaagg
496	536	0.99	ttgacttttccttttctac	agattcttttactcagcctgga
510	550	0.42	tctacagattcttttactc	agcctggaaaccacacaggtct
578	618	0.77	gacccctcaatcacttttt	agctcccacctagtctttttct

Acceptor site predictions for 85.53.15.54.10085.0 :

Start	End	Score	Intron	Exon
280	320	0.98	tttcctgtttgtttttcaa	aggtatcctacctcctggaagg
496	536	0.99	ttgacttttccttttctac	agattcttttactcagcctgga
510	550	0.42	tctacagattcttttactc	agcctggaaaccacacaggtct
578	618	0.77	gacccctcaatcacttttt	agctcccacctagtctttttct

Se pierde un sitio *donor* en la secuencia mutante, el cual coincide con el sitio *donor* del exón 4. Por lo tanto, o bien se deberá usar el sitio *donor* del exón 5, produciéndose el *skipping* del exón 4 o bien se activa el *donor* anterior (cacactggtaaaatg) y se pierden los 18 últimos nucleótidos del exón 4.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
tcag(g/a)tatgg	gcaggt	gcagat	26697	53%

Broken WT Donor Site		Alteration of the WT Donor site, most probably affecting splicing	
Algorithm/Matix	position	sequences	variation
HSF Donor site (matrix GT)	chrX:15459659	- REF : CAGGTATGG - ALT : CAGATATGG	90.21 > 63.07 => -30.09%
MaxEnt Donor site	chrX:15459659	- REF : CAGGTATGG - ALT : CAGATATGG	9.99 > 1.8 => -81.98%

wt	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr	seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr		
wt	14	481	aactcacgg		1.67756170147	0.552521008403	128	26	58	-6.3247954	mut	14	481	aactcacgg		1.67756170147	0.552521008403	128	26	58	-6.3247954
wt	14	448	atattatcc		-2.79093021925	0.566591422122	95	26	58	-5.9810323	mut	14	448	atattatcc		-2.79093021925	0.566591422122	95	26	58	-5.9810323
wt	14	440	caataagggg		-1.14639175413	0.567816091954	87	26	58	-4.8303357	mut	14	440	caataagggg		-1.14639175413	0.567816091954	87	26	58	-4.8303357
wt	14	433	ggatataagt		-3.39798143043	0.572429906542	80	26	58	-5.2673611	mut	14	433	ggatataagt		-3.39798143043	0.572429906542	80	26	58	-5.2673611
wt	14	432	gattaaagtt		-0.90787848085	0.573778491803	79	26	58	-4.2286368	mut	14	432	gattaaagtt		-0.90787848085	0.573778491803	79	26	58	-4.2286368
wt	14	408	aactcatac		0.540765119857	0.583126550868	55	26	58	-2.1392429	mut	14	408	aactcatac		0.540765119857	0.583126550868	55	26	58	-2.1392429
wt	14	399	aagtacgta		-2.77579708249	0.586294416244	46	26	58	-2.8671258	mut	14	399	aagtacgta		-2.77579708249	0.586294416244	46	26	58	-2.8671258
wt	14	389	ggataaatta		-1.60480423366	0.596354166667	36	26	58	-1.7723945	mut	14	389	ggataaatta		-1.60480423366	0.596354166667	36	26	58	-1.7723945
wt	14	386	taattaaat		-2.62817416229	0.595800524934	33	26	58	-1.983376	mut	14	386	taattaaat		-2.62817416229	0.595800524934	33	26	58	-1.983376
wt	14	385	aattaaata		-1.5801248805	0.597368421053	32	26	58	-1.5092105	mut	14	385	aattaaata		-1.5801248805	0.597368421053	32	26	58	-1.5092105
wt	14	381	aaataataa		-1.03280410638	0.601063829787	28	26	58	-1.0405218	mut	14	381	aaataataa		-1.03280410638	0.601063829787	28	26	58	-1.0405218
wt	14	378	taataacccc		2.22335077733	0.603217158177	25	26	58	0.42500721	mut	14	378	taataacccc		2.22335077733	0.603217158177	25	26	58	0.42500721
wt	14	369	aatttaaaaa		-4.65895798149	0.601648351648	16	26	58	-1.7005637	mut	14	369	aatttaaaaa		-4.65895798149	0.601648351648	16	26	58	-1.7005637
wt	14	368	atttaaaaaa		-2.44735197076	0.603305785124	15	26	58	-0.77078183	mut	14	368	atttaaaaaa		-2.44735197076	0.603305785124	15	26	58	-0.77078183
wt	14	327	ttttcaaga		-1.98906176222	0.60248447205	5	13	23	-0.28467615	mut	14	327	ttttcaaga		-1.98906176222	0.60248447205	5	13	23	-0.28467615
wt	14	286	ccatgaaga		-1.58946819573	0.626334519573	85	5	15	-5.258906	mut	14	286	ccatgaaga		-1.58946819573	0.626334519573	85	5	15	-5.258906
wt	14	262	tggtaaaat		-0.155268706687	0.642023346304	61	5	15	-3.1731224	mut	14	262	tggtaaaat		-0.155268706687	0.642023346304	61	5	15	-3.1731224
wt	14	257	aaatgaacc		-0.459301866588	0.650793650794	56	5	15	-2.9728416	mut	14	257	aaatgaacc		-0.459301866588	0.650793650794	56	5	15	-2.9728416
wt	14	206	ggtttagta		-4.23225815046	0.716417910448															

No hay cambios entre ambas secuencias, por lo que la mutación no tendrá efecto en los *BP*.

Variant Effect Predictor tool

ENST00000380420.9:c.273+1G>A	X-15459656-15459656	T	splice_donor_variant	PIR	ENSG00000087842	Transcript	ENST00000380420.10	protein_coding	COSV62916238
ENST00000380420.9:c.273+1G>A	X-15459656-15459656	T	splice_donor_variant	PIR	ENSG00000087842	Transcript	ENST00000380421.3	protein_coding	COSV62916238
ENST00000380420.9:c.273+1G>A	X-15459656-15459656	T	splice_donor_variant non coding transcript variant	PIR	ENSG00000087842	Transcript	ENST00000476381.5	processed_transcript	COSV62916238

Se trata de una mutación que afecta al *splicing*, concretamente al *donor* del exón.

ESEfinder

Se obtiene un único resultado donde se encuentra la posición de interés (g detrás de las mayúsculas que indican el exón) con puntuaciones positivas, tanto para las matrices 5'SS como 3'SS:

370 (-315)	CCAGGAGATTTCAGGgtatggccaaggaca	9.11800	370 (-315)	CCAGGAGATTTCAGGgtatggccaaggaca	3.69350	370 (-315)	CCAGGAGATTTCAGGgtatggccaaggaca	8.83710	370 (-315)	CCAGGAGATTTCAGGgtatggccaaggaca	3.43960
---------------	--------------------------------	---------	---------------	--------------------------------	---------	---------------	--------------------------------	---------	---------------	--------------------------------	---------

Si buscamos la resultado equivalente en la secuencia mutante, se observa que las puntuaciones para las matrices 5'SS pasan a ser negativas, mientras que para las matrices de 3'SS descienden un punto.

370 (-315)	CCAGGAGATTTCAGGatatggccaaggaca	-1.86450	370 (-315)	CCAGGAGATTTCAGGatatggccaaggaca	2.36350	370 (-315)	CCAGGAGATTTCAGGatatggccaaggaca	-2.29120	370 (-315)	CCAGGAGATTTCAGGatatggccaaggaca	2.18920
---------------	--------------------------------	----------	---------------	--------------------------------	---------	---------------	--------------------------------	----------	---------------	--------------------------------	---------

Por lo tanto, lo que nos indican estos resultados es que se está perdiendo un sitio 5'SS (*donor*) en la secuencia mutante, lo que muy probablemente esté afectando al *splicing*.