

Ejemplo comparación de resultados predictores in silico

Cambio de estudio JAK2 c.1849G>T (chr 9:5073770 G/T, rs77375493 o NM_004972.4: c.1849G>T)

Exón 14 e intrones adyacentes:

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

El cambio se encuentra en la segunda línea del exón 14 (la primera **g** en color amarillo, empezando por la izquierda).

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					Donor splice sites, direct strand						
-----					-----						
pos 5'->3'	phase	strand	confidence	5' exon intron 3'	pos 5'->3'	phase	strand	confidence	5' exon intron 3'		
269	1	+	0.60	GGAGACGAGA^GTAAGTAAAA	269	1	+	0.57	GGAGACGAGA^GTAAGTAAAA		
Donor splice sites, complement strand					Donor splice sites, complement strand						
-----					-----						
No donor site predictions above threshold.					No donor site predictions above threshold.						
Acceptor splice sites, direct strand					Acceptor splice sites, direct strand						
-----					-----						
pos 5'->3'	phase	strand	confidence	5' intron exon 3'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'		
180	1	+	0.16	TTTTCCTTAG^TCTTTCTTTG	180	1	+	0.17	TTTTCCTTAG^TCTTTCTTTG		
193	1	+	0.17	TTCTTTGAAG^CAGCAAGTAT	193	1	+	0.17	TTCTTTGAAG^CAGCAAGTAT		
Acceptor splice sites, complement strand					Acceptor splice sites, complement strand						
-----					-----						
pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'	pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'
255	194	1	-	0.23	GTCTCCACAG^ACACATACTC	255	194	2	-	0.16	GTCTCCACAG^AAACATACTC

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
262	276	0.97	gacgaga	gt aagtaa

Donor site predictions for mut :

Start	End	Score	Exon	Intron
262	276	0.97	gacgaga	gt aagtaa

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
50	90	0.87	acattgtatcctcatctat	ag tcatgctgaaagtaggagaa
160	200	0.98	tgtacttttttttttcctt	ag tctttctttgaagcagcaag
173	213	0.45	ttccttagtctttctttga	ag cagcaagtatgatgagcaag
285	325	0.63	gctttctaatagcctttctc	ag agcatctgtttttgtttata
308	348	0.93	catctgtttttgtttatat	ag aaaattcagtttcaggatca

Acceptor site predictions for mut :

Start	End	Score	Intron	Exon
50	90	0.87	acattgtatcctcatctat	ag tcatgctgaaagtaggagaa
160	200	0.98	tgtacttttttttttcctt	ag tctttctttgaagcagcaag
173	213	0.45	ttccttagtctttctttga	ag cagcaagtatgatgagcaag
285	325	0.63	gctttctaatagcctttctc	ag agcatctgtttttgtttata
308	348	0.93	catctgtttttgtttatat	ag aaaattcagtttcaggatca

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
tatgt(g t)tctgt	gtctgt	ttctgt	29865	73%

Human Splicing Finder

Alteration of auxiliary sequences		Significant alteration of ESE / ESS motifs ratio (-5)
Algorithm/Matix	position	sequence
EIE (ESE Site Broken)	chr9:5073765	TATGTG
ESE_SRp55 (ESE Site Broken)	chr9:5073767	TGTGTC
Fas ESS (New ESS Site)	chr9:5073767	TGTTTC
ESE_SC35 (New ESE Site)	chr9:5073767	TGTTTCTG
Fas ESS (New ESS Site)	chr9:5073768	GTTTCT
PESS (New ESS Site)	chr9:5073768	GTTTCTGT
Fas ESS (New ESS Site)	chr9:5073770	TTCTGT

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr		
wt	18	71	ttctcaca		0.825333445135	0.409090909091	29	9	11	-0.87611773	
wt	18	55	gttttaaat		-4.42972677405	0.36	13	9	11	-1.9368039	
wt	18	54	ttttaaatt		-1.50013025492	0.367346938776	12	9	11	-0.7240573	
wt	18	50	aaattatgg		-1.7625394338	0.355555555556	8	9	11	-0.57741784	
wt	18	20	gagtaagta		-1.0303404829	0.4	15	0	0	-0.82193564	
wt	18	16	aagtaaaac		-0.527492388703	0.454545454545	11	0	0	-0.354237	
mut	18	71	ttctcaca		0.825333445135	0.424242424242	29	9	16	-0.82464501	
mut	18	55	gttttaaat		-4.42972677405	0.38	13	9	16	-1.8837652	
mut	18	54	ttttaaatt		-1.50013025492	0.387755102041	12	9	16	-0.67088684	
mut	18	50	aaattatgg		-1.7625394338	0.377777777778	8	9	16	-0.52366149	
mut	18	20	gagtaagta		-1.0303404829	0.4	15	0	0	-0.82193564	
mut	18	16	aagtaaaac		-0.527492388703	0.454545454545	11	0	0	-0.354237	

Variant Effect Predictor tool

ENST00000381652.3:c.1849G>T	9:5073770-5073770	T	missense_variant	JAK2	ENSG00000096968	Transcript	ENST00000381652.4	protein_coding	14/25	2316	1849	617	V/F	GTC/TTC	rs77375493 , CM123094 , COSV67569051 , COSV67571909
ENST00000381652.3:c.1849G>T	9:5073770-5073770	T	missense_variant	JAK2	ENSG00000096968	Transcript	ENST000000636127.1	protein_coding	14/16	2333	1849	617	V/F	GTC/TTC	rs77375493 , CM123094 , COSV67569051 , COSV67571909

ESEfinder

Solo se observa un resultado con puntuación positiva para las matrices 5'. Sin embargo, cuando se compara esta puntuación con la equivalente en la secuencia mutante se observa que ésta disminuido muy poco, por lo que el efecto más probable que sería que se estuviera debilitando un sitio *donor* es poco probable y tampoco tendrá mucho efecto sobre el *splicing*:

<div>232 (-217)</div> <div>GTTTAAATTATGGAGTATGTCTCTGTGGA</div> <div>2.16990</div>	<div>232 (-217)</div> <div>GTTTAAATTATGGAGTATGTCTCTGTGGA</div> <div>-34.26990</div>	<div>232 (-217)</div> <div>GTTTAAATTATGGAGTATGTCTCTGTGGA</div> <div>2.08800</div>	<div>232 (-217)</div> <div>GTTTAAATTATGGAGTATGTCTCTGTGGA</div> <div>-36.80710</div>
<div>232 (-217)</div> <div>GTTTAAATTATGGAGTATGTTTCTGTGGA</div> <div>1.44050</div>	<div>232 (-217)</div> <div>GTTTAAATTATGGAGTATGTTTCTGTGGA</div> <div>-34.99850</div>	<div>232 (-217)</div> <div>GTTTAAATTATGGAGTATGTTTCTGTGGA</div> <div>1.23710</div>	<div>232 (-217)</div> <div>GTTTAAATTATGGAGTATGTTTCTGTGGA</div> <div>-37.50050</div>

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

247 (-202)	GTATGTG	-4.32255	247 (-202)	GTATGTG	-3.54072	247 (-202)	GTATGTGT	-2.64441	247 (-202)	GTATGTG	-3.45373
248 (-201)	TATGTGT	-4.45696	248 (-201)	TATGTGT	-3.53224	248 (-201)	TATGTGTC	-3.16063	248 (-201)	TATGTGT	-1.34137
249 (-200)	ATGTGTC	-4.66650	249 (-200)	ATGTGTC	-3.38355	249 (-200)	ATGTGTCT	-4.01781	249 (-200)	ATGTGTC	-5.45045
250 (-199)	TGTGTCT	-7.48192	250 (-199)	TGTGTCT	-4.88132	250 (-199)	TGTGTCTG	0.37570	250 (-199)	TGTGTCT	-0.20182
251 (-198)	GTGTCTG	-1.65465	251 (-198)	GTGTCTG	-1.94904	251 (-198)	GTGTCTGT	-3.04362	251 (-198)	GTGTCTG	-5.24494
252 (-197)	TGTCTGT	-2.98915	252 (-197)	TGTCTGT	-1.92007	252 (-197)	TGTCTGTG	2.15853	252 (-197)	TGTCTGT	1.82812
253 (-196)	GTCTGTG	-2.01254	253 (-196)	GTCTGTG	-1.77570	253 (-196)	GTCTGTGG	0.62644	253 (-196)	GTCTGTG	-5.84928
247 (-202)	GTATGTT	-3.94901	247 (-202)	GTATGTT	-3.33305	247 (-202)	GTATGTTT	-0.32441	247 (-202)	GTATGTT	-5.83547
248 (-201)	TATGTTT	-6.57699	248 (-201)	TATGTTT	-5.07555	248 (-201)	TATGTTTC	-3.97433	248 (-201)	TATGTTT	-3.36392
249 (-200)	ATGTTTC	-6.58612	249 (-200)	ATGTTTC	-4.98741	249 (-200)	ATGTTTCT	-3.87937	249 (-200)	ATGTTTC	-3.14161
250 (-199)	TGTTTCT	-7.02421	250 (-199)	TGTTTCT	-4.82213	250 (-199)	TGTTTCTG	2.84300	250 (-199)	TGTTTCT	-0.86213
251 (-198)	GTTTCTG	-3.71844	251 (-198)	GTTTCTG	-3.63664	251 (-198)	GTTTCTGT	-1.30498	251 (-198)	GTTTCTG	-3.68342
252 (-197)	TITCTGT	-3.66445	252 (-197)	TITCTGT	-2.64955	252 (-197)	TITCTGTG	1.51086	252 (-197)	TITCTGT	2.33761
253 (-196)	TTCTGTG	-3.38572	253 (-196)	TTCTGTG	-1.98256	253 (-196)	TTCTGTGG	-1.42320	253 (-196)	TTCTGTG	-3.34495

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	2	1	34	631.8785	23	-27.2187	4	11	40	477.1432	33	37.8100	65	88	0.74
mut	9	5	1	34	649.9512	27	-30.0507	4	11	39	461.7917	33	37.8100	76	87	0.87

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

HOT-SKIP

tcctcagaacgttgatggcagttgcaggtccatataaagggaccaaagcacattgtatcctcatctatagtcatgctgaa
agtaggagaaaagtgcacatctttattatggcagagagaattttctgaactatttatggacaacagtc aaacaacaattcttt
gtacttttttttttccttagTCTTTCCTTTGAAGCAGCAAGTATGATGAGCAAGCTTTCTCACAAGCATTGGTTTTAAAT
TATGGAGTATGTGTCTGTGGAGACGAGAgtaagtaaaactacaggcctttcctaatagcctttctcagagcatctgtttttgt
ttatatagaaaattcagtttcaggatcacagctaggtgtcagtgtaaaactataatttaacaggagttaagtatttttgaa
actgaaaacactgtaggactattcagttatatcttgtgaaaaaggaaa

285	73	G	AGTATGTGTCTGTGG	TATGTGTCTGT	0	0	0	6	123.5445	2	-3.1671	0	0	1	15.3514	0	0.0000	8	1	8.00
286	73	A	AGTATGTATCTGTGG	TATGTATCTGT	1	0	0	3	36.0398	2	-3.0994	0	0	0	0.0000	0	0.7886	6	0	6.00
287	73	C	AGTATGTCTCTGTGG	TATGTCTCTGT	2	0	0	5	62.0738	2	-2.9241	0	0	0	0.0000	0	0.4602	9	0	9.00
288	73	T	AGTATGTTTCTGTGG	TATGTTTCTGT	4	3	0	6	141.6172	6	-5.9991	0	0	0	0.0000	0	0.0000	19	0	19.00