

Ejemplo comparación de resultados predictores in silico

Cambio de estudio KDR c.802C>T (chr4:55113478 C/T, COSV55781128 o NM_002253.4: c.802C>T)

Exón 7 e intrones adyacentes:

```
atttatttctcctgaaaacaaaatatatttggtgtctgcatgaggittaaagttctggtgt
ccctgttttttagcattaaataatgtttaccaaaagccagatcttaattctgtgtgttacta
gaagttattgggtaatgttataatgctgtgctttgggaagttcagtcacatctttttttcag
CATCAGCATAAGAAACTTGTAAACCGAGACCTAAAAACCCAGTCTGGGAGTGAGATGAAG
AAATTTTGTAGCACCCTTAACCTATAGATGTTGTAAACCCGAGTGACCAAGGATTGTACACC
TGTGCAGCATCCAGTGGGCTGATGACCAAGAAGAACAGCACATTGTTCAGGGTCCATG
gtaagctatgggtcttggaattattctgtgccttgacaagtgaataattttaaataaatt
taggtcacttagtgatttcctatctttgttcattcagaagatagtttctagtttttcttggt
aggagagccacatgacctagaggtcaagagcatagctttgtagtcagggaacttgggttca
```

El cambio se encuentra en la primera línea del exón 7 (la c 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	+	0.83	AGGGTCCATG		GTAAGCTATG	

Donor splice sites, complement strand

No donor site predictions above threshold.

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
180		0	+	0.48	TTTTTTTCAG		CATCAGCATA	
186		0	+	0.18	TCAGCATCAG		CATAAGAAAC	
192		0	+	0.07	TCAGCATAAG		AAACTTGTAA	
480		1	+	0.00	TTCTTGTTAG		GGAGGCCACA	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
495	44	1	-	0.00	TGACCTCTAG		GTCATGTGGC	
390	149	0	-	0.18	ACTTGTCAAG		GCACAGAATA	
384	155	0	-	0.19	CAAGGCACAG		AATAATTTCC	
371	168	1	-	0.20	AATTTCCAAG		ACCATAGCTT	
364	175	2	-	0.20	AAGACCATAG		CTTACCATGG	

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
359		1	+	0.83	AGGGTCCATG		GTAAGCTATG	

Donor splice sites, complement strand

No donor site predictions above threshold.

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
180		0	+	0.28	TTTTTTTCAG		CATTAGCATA	
480		1	+	0.00	TTCTTGTTAG		GGAGGCCACA	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
495	44	1	-	0.00	TGACCTCTAG		GTCATGTGGC	
390	149	0	-	0.18	ACTTGTCAAG		GCACAGAATA	
384	155	0	-	0.19	CAAGGCACAG		AATAATTTCC	
371	168	1	-	0.20	AATTTCCAAG		ACCATAGCTT	
364	175	2	-	0.20	AAGACCATAG		CTTACCATGG	

Por un lado, el sitio *acceptor* del exón (en verde) cambia a causa de la presencia de la mutación en la secuencia mutante, lo que hace que su nivel de confianza se reduzca. Además, desaparecen dos de los sitios *acceptor* (en rojo), que no son los que se utilizan para el exón, por lo que no se tendrán en cuenta. Por lo tanto, un sitio *acceptor* más débil podría estar afectando al *splicing* siempre y cuando existiera otro sitio *acceptor* más fuerte que el complejo que realiza el *splicing* fuera capaz de detectar antes que él, o bien que esa debilidad haga imposible su reconocimiento por el complejo y se salte el exón completo.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.2.148.572545.0 :

Start	End	Score	Exon	Intron
352	366	1.00	gtccatg	gt ^a aagcta

Donor site predictions for 10.42.3.123.572557.0 :

Start	End	Score	Exon	Intron
352	366	1.00	gtccatg	gt ^a aagcta

Acceptor site predictions for 10.42.2.148.572545.0 :

Start	End	Score	Intron	Exon
52	92	0.68	ttctgggtgtccctgtttttt	agcattaaataatgtttaccaa
160	200	0.82	tcagtcaactctttttttc	agcatcagcataaagaaacttgt
433	473	0.41	attcctattttgttcatttc	agaagatagtttctagtttttc
460	500	0.99	gtttctagtttttcttggtt	aggaggccacatgacctagag

Acceptor site predictions for 10.42.3.123.572557.0 :

Start	End	Score	Intron	Exon
52	92	0.68	ttctgggtgtccctgtttttt	agcattaaataatgtttaccaa
160	200	0.88	tcagtcaactctttttttc	agcattagcataaagaaacttgt
433	473	0.41	attcctattttgttcatttc	agaagatagtttctagtttttc
460	500	0.99	gtttctagtttttcttggtt	aggaggccacatgacctagag

Se produce un cambio en el *acceptor* (en rojo) a causa de la presencia de la mutación. Esta hace que el *score* para este sitio sea mayor, por lo que podría estar reforzándolo como el sitio *acceptor* del exón, lo cual no tendría efecto en el *splicing*.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
agcat(c/t)agcat	agcatc	agcatt	26632	53%

CRYP-SKIP

Human Splicing Finder

Alteration of auxiliary sequences		Significant alteration of ESE / ESS motifs ratio (-2)
Algorithm/Matix	position	sequence
ESS_hnRNPA1 (New ESS Site)	chr4:55113478	TAGCAT
RESCUE ESE (ESE Site Broken)	chr4:55113481	CATCAG
ESE_SRp55 (ESE Site Broken)	chr4:55113483	AGCATC
PESE (New ESE Site)	chr4:55113485	TCAGCATT

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr
wt	20	51	agttcagtc	-1.1989305333	0.521739130435	2	15	31	0.26304091
wt	20	47	cagtcaact	-1.00903623219	0.52380952381	2	11	30	0.32874635
wt	20	34	ttttcagca	-1.43635880732	0.379310344828	29	0	0	-1.8737697
wt	20	28	gcatcagca	-2.34559242658	0.347826086957	23	0	0	-1.8601562
wt	20	22	gcataagaa	-2.42378228924	0.352941176471	17	0	0	-1.5093293
mut	20	51	agttcagtc	-1.1989305333	0.521739130435	2	15	31	0.26304091
mut	20	47	cagtcaact	-1.00903623219	0.52380952381	2	11	30	0.32874635
mut	20	34	ttttcagca	-1.43635880732	0.379310344828	29	0	0	-1.8737697
mut	20	28	gcattagca	-2.70362756626	0.347826086957	23	0	0	-2.0003439
mut	20	22	gcataagaa	-2.42378228924	0.352941176471	17	0	0	-1.5093293

El único cambio que hay entre ambos es en el BP que se detecta dónde se encuentra la mutación, pero la presencia de este cambio hace que el *score* sea mucho más bajo por lo que no tendrá efecto en el *splicing*.

Variant Effect Predictor tool

ENST00000263923.5:c.802C>T	4.55113478-55113478	A	stop_gained	KDR	ENSG00000128052	Transcript	ENST00000263923.5	protein_coding	7/30	1104	802	268	Q/*	CAG/TAG	COSV55781128
ENST00000263923.5:c.802C>T	4.55113478-55113478	A	non_coding_transcript_exon_variant	KDR	ENSG00000128052	Transcript	ENST00000512566.1	retained_intron	7/13	802	-	-	-	-	COSV55781128
ENST00000263923.5:c.802C>T	4.55113478-55113478	A	non_coding_transcript_exon_variant	KDR	ENSG00000128052	Transcript	ENST00000647068.1	processed_transcript	7/30	815	-	-	-	-	COSV55781128

ESEfinder

Se obtienen dos resultados con puntuaciones positivas para las matrices 3' (166 y 172) y uno con puntuaciones positivas para las matrices 5' (184):

166 (-373)	aactctttttttcagCATCAGCATAAGAAA	-29.70240	166 (-373)	aactctttttttcagCATCAGCATAAGAAA	9.16630	166 (-373)	aactctttttttcagCATCAGCATAAGAAA	-26.53580	166 (-373)	aactctttttttcagCATCAGCATAAGAAA	8.78660
172 (-367)	ttttttcagCATCAGCATAAGAACTTGTA	-25.55170	172 (-367)	ttttttcagCATCAGCATAAGAACTTGTA	5.74660	172 (-367)	ttttttcagCATCAGCATAAGAACTTGTA	-22.43830	172 (-367)	ttttttcagCATCAGCATAAGAACTTGTA	5.52960
184 (-355)	CAGCATAAGAACTTGTAACCGAGACCTA	2.98630	184 (-355)	CAGCATAAGAACTTGTAACCGAGACCTA	-25.97310	184 (-355)	CAGCATAAGAACTTGTAACCGAGACCTA	2.76230	184 (-355)	CAGCATAAGAACTTGTAACCGAGACCTA	-28.31380

Si se comparan estos valores con los de la secuencia mutante observamos que para las todas las puntuaciones se reducen:

166 (-373)	aactctttttttcagCATTAGCATAAGAAA	-29.49880	166 (-373)	aactctttttttcagCATTAGCATAAGAAA	8.70220	166 (-373)	aactctttttttcagCATTAGCATAAGAAA	-26.41670	166 (-373)	aactctttttttcagCATTAGCATAAGAAA	8.33820
172 (-367)	ttttttcagCATTAGCATAAGAACTTGTA	-27.70250	172 (-367)	ttttttcagCATTAGCATAAGAACTTGTA	4.15830	172 (-367)	ttttttcagCATTAGCATAAGAACTTGTA	-24.54020	172 (-367)	ttttttcagCATTAGCATAAGAACTTGTA	3.91590
184 (-355)	TAGCATAAGAACTTGTAACCGAGACCTA	2.24830	184 (-355)	TAGCATAAGAACTTGTAACCGAGACCTA	-25.74850	184 (-355)	TAGCATAAGAACTTGTAACCGAGACCTA	2.03000	184 (-355)	TAGCATAAGAACTTGTAACCGAGACCTA	-28.09560

Por lo tanto, podría estar debilitándose un sitio *acceptor*, pero los resultados son contradictorios.

Además, se ven alteraciones en los sitios ESE, que pueden alterar al *splicing*:

178 (-361)	cagCATC	-1.50109	178 (-361)	cagCATC	-0.26818	178 (-361)	cagCATCA	-1.54105	178 (-361)	cagCATC	-1.77299
179 (-360)	agCATCA	-1.46478	179 (-360)	agCATCA	-0.89571	179 (-360)	agCATCAG	0.21909	179 (-360)	agCATCA	-2.44442
180 (-359)	gCATCAG	-4.77908	180 (-359)	gCATCAG	-3.81792	180 (-359)	gCATCAGC	-2.68291	180 (-359)	gCATCAG	-0.86937
181 (-358)	CATCAGC	-1.44484	181 (-358)	CATCAGC	-0.41247	181 (-358)	CATCAGCA	1.01130	181 (-358)	CATCAGC	1.81108
182 (-357)	ATCAGCA	-0.22047	182 (-357)	ATCAGCA	-0.02133	182 (-357)	ATCAGCAT	-2.82719	182 (-357)	ATCAGCA	-4.24377
183 (-356)	TCAGCAT	-6.23642	183 (-356)	TCAGCAT	-3.87630	183 (-356)	TCAGCATA	-3.06958	183 (-356)	TCAGCAT	-0.08646
184 (-355)	CAGCATA	0.70555	184 (-355)	CAGCATA	1.15621	184 (-355)	CAGCATAA	-1.91972	184 (-355)	CAGCATA	-3.31209

178 (-361)	cagCATT	0.35231	178 (-361)	cagCATT	0.92852	178 (-361)	cagCATTAA	-0.89661	178 (-361)	cagCATT	-3.31209
179 (-360)	agCATTAA	-1.00708	179 (-360)	agCATTAA	-0.79171	179 (-360)	agCATTAG	-1.50992	179 (-360)	agCATTAA	-4.16135
180 (-359)	gCATTAG	-7.30280	180 (-359)	gCATTAG	-5.62972	180 (-359)	gCATTAGC	-3.44215	180 (-359)	gCATTAG	0.62188
181 (-358)	CATTAGC	-2.90216	181 (-358)	CATTAGC	-1.66722	181 (-358)	CATTAGCA	0.78046	181 (-358)	CATTAGC	-0.57355
182 (-357)	ATTAGCA	-2.53049	182 (-357)	ATTAGCA	-2.09019	182 (-357)	ATTAGCAT	-3.39612	182 (-357)	ATTAGCA	-2.89549
183 (-356)	TTAGCAT	-5.63349	183 (-356)	TTAGCAT	-3.62752	183 (-356)	TTAGCATA	-1.68118	183 (-356)	TTAGCAT	-0.39852
184 (-355)	TAGCATA	-2.24493	184 (-355)	TAGCATA	-1.94884	184 (-355)	TAGCATAA	-1.94085	184 (-355)	TAGCATA	-2.95539

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	3	3	0	42	622.0635	15	-23.2563	19	21	69	1044.8947	82	98.3847	63	191	0.33
mut	3	3	0	42	622.0635	15	-23.5657	19	20	69	1044.8947	78	94.5902	63	186	0.34

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

HOT-SKIP

at ttattttctcctgaaacaaaatatatttgttgtctgcatgagggttaaaagtcttggtgtccctgttttagcattaaat
aatgtttaccaaagcccagatttaattctgtgtgttactagaagttattgggtaatgttatatgctgtgctttggaagtt
cagtcaactcttttttccagCATCAGCATAAGAACTTGTAACCGAGACCTAAAAACCCAGTCTGGGAGTGAGATGAAG
AAATTTTGAGCACCTTAACCTATAGATGGTGTAAACCGGAGTGACCAAGGATTGTACACCTGTGCAGCATCCAGTGGGCT
GATGACCAAGAAGAACAGCACATTTGTGAGGGTCCATGgtaagctatggctcttggaattattctgtgccttgacaagt
agataatttaaataaatttaggtcacttagtgattccattttgttcattcagaagatagtttctagttttctttag
ggaggccacatgacctagagggtcaagagcatagcctttagtcaggaaacttggttca