

## Ejemplo comparación de resultados predictores in silico

Cambio de estudio NCKAP1 c.253C>T (chr2:183003310 C/T, COSV62941240 o NM\_013436.5:c.253C>T)

Exón 4 e intrones adyacentes:

```
tttttaatttttaagagttcaattgatactttttttccaaacactttcagattttcctt
aatagttagcgcaatttggtactaatatgcttttgatttgtttatagcctcacactgtcag
ttgctgatcactatttagcagttgctgtctcgagtactttggttttttaaaaagcgcgtaaat
catctagaaatatctctaaaagccttggtctccatatagcagtcataaaaagaagatagttt
gtagtataaattaaatacttttgcttctctatgagcaatgcattctaattttgtttttacag
CAACAGCTTGCACTACAGAAAGAAAAATCAAGAGATTCTGAAAAATCTGGCATTATAT
TACTTCACATTGTAGATGTTATGGAATTTAAG
gtatgtatttttaatatataatatttaacacttcagaagtagattctaaatgttttattaatt
gtgtgaattgctgtgtcagttgtaaataataaaattgaaattcttcagaactgaagttt
aacacataccttatttggtttggttttctgaatttaaacagattttatacagataaaagat
taaaataagtattttcttttcag
```

El cambio se encuentra en el exón 4 (la primera **c** en color verde).

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'		
394	0	+	0.83	GGAATTTAAG^GTATGTATTT	394	0	+	0.83	GGAATTTAAG^GTATGTATTT		
Donor splice sites, complement strand					Donor splice sites, complement strand						
-----					-----						
pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' exon intron 3'	pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' exon intron 3'
522	74	1	-	0.00	AACAAATAAG^GTATGTGTTA	522	74	1	-	0.00	AACAAATAAG^GTATGTGTTA
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'		
65	0	+	0.00	TCCTTAATAG^TAGCGCAATT	65	0	+	0.00	TCCTTAATAG^TAGCGCAATT		
300	0	+	0.85	GTTTTTACAG^CAACAGCTTG	300	0	+	0.82	GTTTTTACAG^CAACAGCTTG		
306	0	+	0.17	ACAGCAACAG^CTTGCACAAC	306	0	+	0.16	ACAGCAACAG^CTTGCACAAT		
Acceptor splice sites, complement strand					Acceptor splice sites, complement strand						
-----					-----						
pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'	No acceptor site predictions above threshold.					
316	280	1	-	0.93	CTTCTGTAG^TTGTGCAAGC						

Desaparece el sitio *acceptor* de la hebra complementaria. Este se encuentra dentro del exón, coincidiendo con la posición de interés. Como no participa en el proceso de *splicing* normal, es probable que la mutación no esté afectando al *splicing*.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 85.53.15.54.26305.0 :

Start	End	Score	Exon	Intron
387	401	1.00	atttaag	gtatgtat

Donor site predictions for 85.53.15.54.26355.0 :

Start	End	Score	Exon	Intron
387	401	1.00	atttaag	gtatgtat

Acceptor site predictions for 85.53.15.54.26305.0 :

Start	End	Score	Intron	Exon
31	71	0.77	tttttttcctaacacttttc	agattttccttaatagtagcgc
45	85	0.66	ctttcagattttccttaat	agtagcgcaatttggtactaat
86	126	0.90	atgctttgatttgttttat	agcctcacactgtcagttgctg
280	320	0.92	attctaatttggtttttac	agcaacagcttgcaacaactaca
356	396	0.85	tatattacttcacatttgt	agatgttatggaatttaaggta

Acceptor site predictions for 85.53.15.54.26355.0 :

Start	End	Score	Intron	Exon
31	71	0.77	tttttttcctaacacttttc	agattttccttaatagtagcgc
45	85	0.66	ctttcagattttccttaat	agtagcgcaatttggtactaat
86	126	0.90	atgctttgatttgttttat	agcctcacactgtcagttgctg
280	320	0.89	attctaatttggtttttac	agcaacagcttgcaacaattaca
356	396	0.85	tatattacttcacatttgt	agatgttatggaatttaaggta

No hay cambios.

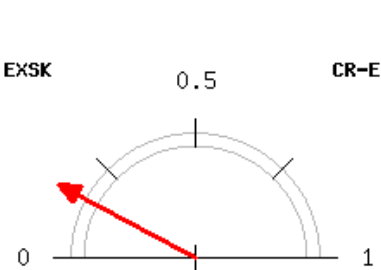
Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
cacaa(c/t)tacag	actaca	attaca	28703	66%

CRYP-SKIP

Results for sequence wt

Exon length (bp)	93
PESS ( $\leq -2.62$ ) density	3.23
NN 5'ss score density	0.00
SF2/ASF score density	3.02
FAS-ESS (hex2) density	1.08
EIE score density	723.59
Probability of cryptic splice site activation ( $P_{CR-E}$ )	0.16



No hay efecto de la mutación en el *splicing*.

>wt  
ttctaattttgtttttacagCAACAGCTTGACAACTACAGAAAGAAAAATCAGAGATTCTGAAAAATCTGGCATTATAT  
TACTTCACATTTGTAGATGTTATGGAATTTAAGgtatgta

Human Splicing Finder



Alteration of auxiliary sequences		Significant alteration of ESE / ESS motifs ratio (-7)
Algorithm/Matix	position	sequence
Sironi_motif1 (ESS Site Broken)	chr2:183003310	CTACAGAA
RESCUE ESE (ESE Site Broken)	chr2:183003311	ACTACA
RESCUE ESE (ESE Site Broken)	chr2:183003312	AACTAC
ESE_SC35 (ESE Site Broken)	chr2:183003312	AACTACAG
EIE (ESE Site Broken)	chr2:183003313	CAACTA
RESCUE ESE (ESE Site Broken)	chr2:183003314	ACAACT
EIE (ESE Site Broken)	chr2:183003314	ACAACT
ESE_9G8 (ESE Site Broken)	chr2:183003315	CACAAC
EIE (ESE Site Broken)	chr2:183003315	CACAAC

## SVM-BPfinder

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_48	236	tagtaaaat	-0.655994023546	0.502164502165	8	16	30	0.080198403	mut_48	236	tagtaaaat	-0.655994023546	0.502164502165	8	16	30	0.080198403	
wt_48	231	aaatataat	-3.05381608375	0.504424778761	3	16	30	-0.54144067	mut_48	231	aaatataat	-3.05381608375	0.504424778761	3	16	30	-0.54144067	
wt_48	230	aattaataa	-1.5801248805	0.506666666667	2	16	30	0.099601493	mut_48	230	aattaataa	-1.5801248805	0.506666666667	2	16	30	0.099601493	
wt_48	211	ctatgagca	-1.20461780604	0.490291262136	7	20	32	-0.05618457	wt_48	211	ctatgagca	-1.20461780604	0.490291262136	7	20	32	-0.05618457	
wt_48	197	ttctaattt	1.0553543152	0.489583333333	1	12	25	1.1427179	mut_48	197	ttctaattt	1.0553543152	0.489583333333	1	12	25	1.1427179	
wt_48	186	ttttcacag	-2.47574195632	0.469613259669	63	15	22	-4.1987664	wt_48	186	ttttcacag	-2.47574195632	0.469613259669	63	15	22	-4.1987664	
wt_48	150	aaatcagag	-2.93352639442	0.496551724138	27	15	22	-2.0905707	wt_48	165	caattacag	-0.324115538156	0.46875742	15	22	22	-2.0273171	
wt_48	140	ttctgaaaa	0.10046204897	0.503703703704	17	15	22	-0.26732782	mut_48	150	aaatcagag	-2.93352639442	0.496551724138	27	15	22	-2.0905707	
wt_48	126	gcattatat	-2.74496552013	0.512396694215	3	15	22	-0.49246314	wt_48	140	ttctgaaaa	0.10046204897	0.503703703704	17	15	22	-0.26732782	
wt_48	121	atttacttt	-1.82934800956	0.508620689655	1	12	18	-0.045841859	wt_48	126	gcattatat	-2.74496552013	0.512396694215	3	15	22	-0.49246314	
wt_48	116	acttcacat	-1.16577152797	0.495495495495	32	6	13	-1.7990865	wt_48	121	atttacttt	-1.82934800956	0.508620689655	1	12	18	-0.045841859	
wt_48	101	atgttatgg	-2.61285848179	0.489583333333	17	6	13	-1.4181241	wt_48	116	acttcacat	-1.16577152797	0.495495495495	32	6	13	-1.7990865	
wt_48	92	aatttaagg	-4.20751026383	0.494252873563	8	6	13	-1.4713127	wt_48	101	atgttatgg	-2.61285848179	0.489583333333	17	6	13	-1.4181241	
wt_48	91	atttaaggt	-2.42222900354	0.5	7	6	13	-0.7071362	wt_48	92	aatttaagg	-4.20751026383	0.494252873563	8	6	13	-1.4713127	
wt_48	78	attttaata	-5.1666216293	0.493150684932	37	8	14	-3.6735416	mut_48	91	atttaaggt	-2.42222900354	0.5	7	6	13	-0.7071362	
wt_48	77	ttttaatat	-0.63666895298	0.5	36	8	14	-1.8343405	wt_48	78	attttaata	-5.1666216293	0.493150684932	37	8	14	-3.6735416	
wt_48	70	atataatat	-1.2551519652	0.507692307692	29	8	14	-1.6309332	wt_48	77	ttttaatat	-0.63666895298	0.5	36	8	14	-1.8343405	
wt_48	64	tatttaaca	-3.21777038232	0.491525423729	23	8	14	-2.0248226	wt_48	70	atataatat	-1.2551519652	0.507692307692	29	8	14	-1.6309332	
wt_48	63	atttaacac	0.525200512851	0.5	22	8	14	-0.49323721	wt_48	64	tatttaaca	-3.21777038232	0.491525423729	23	8	14	-2.0248226	
wt_48	56	acttcagaa	-3.30341226882	0.470588235294	15	8	14	-1.5587309	mut_48	63	atttaacac	0.525200512851	0.5	22	8	14	-0.49323721	
wt_48	44	gatttaaat	-3.7270233635	0.512820512821	3	8	14	-0.95137506	wt_48	56	acttcagaa	-3.30341226882	0.470588235294	15	8	14	-1.5587309	
wt_48	43	attttaagt	-1.42149293897	0.526315789474	2	8	14	0.019005971	wt_48	44	gatttaaat	-3.7270233635	0.512820512821	3	8	14	-0.95137506	
wt_48	36	tgtttaata	-3.73990524016	0.516129032258	14	10	13	-1.6609478	wt_48	43	attttaagt	-1.42149293897	0.526315789474	2	8	14	0.019005971	
wt_48	33	ttttaattat	-3.32242127387	0.5	11	10	13	-1.3127973	mut_48	36	tgtttatta	-3.73990524016	0.516129032258	14	10	13	-1.6609478	
wt_48	32	tatttaattg	0.196622193144	0.518518518519	10	10	13	0.13435381	wt_48	33	ttttaattat	-3.32242127387	0.5	11	10	13	-1.3127973	
wt_48	24	gtgtgaatt	-0.955994067854	0.526315789474	2	10	13	0.19195493	wt_48	32	tatttaattg	0.196622193144	0.518518518519	10	10	13	0.13435381	
									mut_48	24	gtgtgaatt	-0.955994067854	0.526315789474	2	10	13	0.19195493	

El único cambio que hay es la aparición de un BP en la secuencia mutante, pero tiene una puntuación negativa, por lo que no se tendrá en cuenta.

## Variant Effect Predictor tool

ENST00000360982.2:c.253C>T	<a href="#">2:183003310-183003310</a>	A	 <a href="#">synonymous_variant</a>	NCKAP1	<a href="#">ENSG000000061676</a>	Transcript	<a href="#">ENST00000360982.2</a>	protein_coding	4/32	1012	253	85	L	CTA/TTA	<a href="#">rs1327765186, COSV62941240</a>
ENST00000360982.2:c.253C>T	<a href="#">2:183003310-183003310</a>	A	<a href="#">synonymous_variant</a>	NCKAP1	<a href="#">ENSG000000061676</a>	Transcript	<a href="#">ENST00000361354.9</a>	protein_coding	3/31	593	235	79	L	CTA/TTA	<a href="#">rs1327765186, COSV62941240</a>
ENST00000360982.2:c.253C>T	<a href="#">2:183003310-183003310</a>	A	 <a href="#">regulatory_region_variant</a>	-	-	RegulatoryFeature	<a href="#">ENSR00001041848</a>	promoter_flanking_region	-	-	-	-	-	-	<a href="#">rs1327765186, COSV62941240</a>

## ESEfinder

Aparecen dos predicciones con puntuaciones positivas para las matrices de 3'SS (292, 307):

292 (-304)	tttttacagCAACAGCTTGCACAACTACAG	-11.51300	292 (-304)	tttttacagCAACAGCTTGCACAACTACAG	5.17230	292 (-304)	tttttacagCAACAGCTTGCACAACTACAG	-11.89950	292 (-304)	tttttacagCAACAGCTTGCACAACTACAG	4.80680
307 (-289)	CTTGCACTACAGAAAGAAAAATCAGAG	-21.38000	307 (-289)	CTTGCACTACAGAAAGAAAAATCAGAG	4.29130	307 (-289)	CTTGCACTACAGAAAGAAAAATCAGAG	-18.40890	307 (-289)	CTTGCACTACAGAAAGAAAAATCAGAG	4.25190

Cuando buscamos el resultado equivalente para la secuencia mutante, las puntuaciones no se ven a penas alteradas:

292 (-304)	tttttacagCAACAGCTTGCACAATTACAG	-11.71270	292 (-304)	tttttacagCAACAGCTTGCACAATTACAG	4.57560	292 (-304)	tttttacagCAACAGCTTGCACAATTACAG	-12.22930	292 (-304)	tttttacagCAACAGCTTGCACAATTACAG	4.22060
307 (-289)	CTTGCAATTACAGAAAGAAAAATCAGAG	-21.89020	307 (-289)	CTTGCAATTACAGAAAGAAAAATCAGAG	4.40020	307 (-289)	CTTGCAATTACAGAAAGAAAAATCAGAG	-18.86120	307 (-289)	CTTGCAATTACAGAAAGAAAAATCAGAG	4.28240

Por lo tanto, no hay cambios en los sitios de *splicing*.

En cuanto a las ESE, para los resultados donde se encuentra la posición de interés, solo se obtienen dos resultados con puntuaciones positivas en más de una matriz (311 y 316):

310 (-286)	GCACAAC	-5.46651	310 (-286)	GCACAAC	-7.32534	310 (-286)	GCACAAC	-1.24892	310 (-286)	GCACAAC	2.40960
311 (-285)	CACAACT	1.61160	311 (-285)	CACAACT	1.13438	311 (-285)	CACAACTA	0.44184	311 (-285)	CACAACT	-2.95094
312 (-284)	ACAACTA	-1.72320	312 (-284)	ACAACTA	-2.07160	312 (-284)	ACAACTAC	-6.05761	312 (-284)	ACAACTA	-2.43550
313 (-283)	CAACTAC	-1.78575	313 (-283)	CAACTAC	-4.02258	313 (-283)	CAACTACA	-0.07876	313 (-283)	CAACTAC	2.04480
314 (-282)	AACTACA	-2.95702	314 (-282)	AACTACA	-3.46843	314 (-282)	AACTACAG	2.61237	314 (-282)	AACTACA	-4.45947
315 (-281)	ACTACAG	-2.69876	315 (-281)	ACTACAG	-3.25609	315 (-281)	ACTACAGA	-4.65026	315 (-281)	ACTACAG	0.35368
316 (-280)	CTACAGA	0.88451	316 (-280)	CTACAGA	-0.36073	316 (-280)	CTACAGAA	-0.09595	316 (-280)	CTACAGA	3.27392

Si buscamos las predicciones equivalentes para la secuencia mutante, se observa que las puntuaciones se mantienen parecidas.

310 (-286)	GCACAAT	-5.47194	310 (-286)	GCACAAT	-4.26981	310 (-286)	GCACAATT	-0.60448	310 (-286)	GCACAAT	0.87050
311 (-285)	CACAATT	1.59209	311 (-285)	CACAATT	1.71559	311 (-285)	CACAATTA	-1.28718	311 (-285)	CACAATT	-4.66787
312 (-284)	ACAATTA	-4.59532	312 (-284)	ACAATTA	-3.53499	312 (-284)	ACAATTAC	-6.81685	312 (-284)	ACAATTA	-0.94425
313 (-283)	CAATTAC	-5.47990	313 (-283)	CAATTAC	-3.04051	313 (-283)	CAATTACA	-0.30960	313 (-283)	CAATTAC	-0.33983
314 (-282)	AATTACA	-5.77845	314 (-282)	AATTACA	-5.02588	314 (-282)	AATTACAG	2.04343	314 (-282)	AATTACA	-3.11119
315 (-281)	ATTACAG	-2.65316	315 (-281)	ATTACAG	-2.44998	315 (-281)	ATTACAGA	-3.26186	315 (-281)	ATTACAG	0.04162
316 (-280)	TTACAGA	-3.31121	316 (-280)	TTACAGA	-2.22054	316 (-280)	TTACAGAA	-0.11708	316 (-280)	TTACAGA	3.63061

Por lo tanto, no hay cambios en el *splicing*.

## 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	2	1	0	21	296.6577	16	-19.5984	3	22	37	647.2721	36	43.5254	40	98	0.41
mut	2	1	0	21	296.6577	16	-19.7104	2	19	34	571.3015	32	39.2206	40	87	0.46

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

## HOT-SKIP

>wt

```
tttttaattttttaagagttcatttgatacttttttccctaacactttcagattttccttaatagtagcgcaatttggtgta  
ctaataatgctttgatttgttttatagcctcacactgtcagttgctgatcactattagcagttgctgtctcgagtactttg  
gttttttaaaaagccgtaatcatctagaaatattctaaaagcttggtgcttccatatagcagtcataaaaagaagatagttt  
gtagtaaaattaaatacttttgttctctatgagcaatgcattcctaattttgtttttacagCAACAGCTTGCACAACTACA  
GAAAGAAAAATCAGAGATTCTGAAAAATCTGGCATTATATTACTTCACATTTGTAGATGTTATGGAATTTAAGgtatgta  
ttttaatatataatatttaacacttcagaagtagatttaaagttttattaattgtgtgaattgctgtgtcagttgtaaat  
aataaaattgaaattcttccagaactgaagtttaacacataccttatttggttggttttctgaatttaaacagatttat  
acagataaaaagattaaaataagtatttcttttcag
```

>wt

```
tttttaattttttaagagttcatttgatacttttttccctaacactttcagattttccttaatagtagcgcaatttggtgta  
ctaataatgctttgatttgttttatagcctcacactgtcagttgctgatcactattagcagttgctgtctcgagtactttg  
gttttttaaaaagccgtaatcatctagaaatattctaaaagcttggtgcttccatatagcagtcataaaaagaagatagttt  
gtagtaaaattaaatacttttgttctctatgagcaatgcattcctaattttgtttttacagCAACAGCTTGCACAATTACA  
GAAAGAAAAATCAGAGATTCTGAAAAATCTGGCATTATATTACTTCACATTTGTAGATGTTATGGAATTTAAGgtatgta  
ttttaatatataatatttaacacttcagaagtagatttaaagttttattaattgtgtgaattgctgtgtcagttgtaaat  
aataaaattgaaattcttccagaactgaagtttaacacataccttatttggttggttttctgaatttaaacagatttat  
acagataaaaagattaaaataagtatttcttttcag
```