

## Ejemplo comparación de resultados predictores in silico

Cambio de estudio SORL1 c.5914C>G (chr11:121621088 C/G, COSV52750647 o NM\_003105.6: c.5914C>G)

Exón 44 e intrones adyacentes:

```
ttggatgtgcttgcttgggtccaaacgtacdtgctgatgtcacgataaatgtggccctccct
agaggggttaaaaaaattttagttgataactctatgtgcaattcctatgagtgaatagttat
ttaaagattggatcttgggtcactccaaaggccttccaaagaatggagtcagatttcgaggac
tcaaactcttggttgctgtgggtcccagctattaataatctactacattgtccaccattgaact
acaaagagtgaaaaaatgtcacactttctgattatcattcacttggtcttttttgggtcctag
TTGTATGCAGTTGCAGTCAAAGATCTCATAGAAAGACTGACAGGAGCTACAAAGTAAAA
TCCCGTAAACAGCACTGTGGAATACACCTTAACAAGTTGGAGCCTGGCGGGAATACCAC
ATCATTGTCCAACCTGGGCAACATGAGCAAGATTCAGCATAAAATTACACAG
gtaagcaggagagagaggtagaggtctctcacacagcctgccctcagtgctgtgctcgcta
gacctccagagacagacttttcattaggcggttggtttcacaggagtgcaaacgctgccc
ttcaatatttctacaggggcagagtagtgacagaggctccagggtctgagdcggtggcggtg
tggccagggcacagtggtgctgggagtggttatcgttttctttatcttaagtcttttt
aactcccctcgctgggttttgatattactaaaggaaaagaggaggatgccaagagagga
```

El cambio se encuentra en la primera línea del exón 44 (la **c** en color rojo subrayada en amarillo).

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'
476	1	+	0.83	ATTACCACAG	^	GTAAGCAGGA	
491	1	+	0.37	CAGGAGAGAG	^	GTTAGAGGTC	

### Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
28	748	-	-	0.00	ACATCAGCAG	^	GTAGGTTGGA	

### Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
300	0	+	0.90	TTGGTCCTAG	^	TTGTATGCAG	
310	1	+	0.19	TTGTATGCAG	^	TTGCAGTCAA	
316	1	+	0.20	GCAGTTGCAG	^	TCAAAGATCT	
322	1	+	0.19	GCAGTCAAAG	^	ATCTCATAAG	
332	2	+	0.18	ATCTCATAAG	^	AAAGACTGAC	
336	0	+	0.17	CATAAGAAAG	^	ACTGACAGGA	
344	2	+	0.17	AGACTGACAG	^	GAGCTACAAA	
347	2	+	0.07	CTGACAGGAG	^	CTACAAAGTA	
577	0	+	0.33	TGTTTCACAG	^	GGAGTGCAAA	
611	2	+	0.43	ATTCCTACAG	^	GGGCAGAGTA	

### Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
742	34	2	-	0.00	TTTCCTTTAG	^	TAATATCAAA	
433	343	0	-	0.17	TGTTCCCCAG	^	TTGGACAATG	
239	537	0	-	0.25	CTCTTTGTAG	^	TTCAATGGTG	
59	717	2	-	0.00	AACCCTCTAG	^	GGAGGGCCAC	

### Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
476	1	+	0.83	ATTACCACAG	^	GTAAGCAGGA	
491	1	+	0.37	CAGGAGAGAG	^	GTTAGAGGTC	

### Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
28	748	-	-	0.00	ACATCAGCAG	^	GTAGGTTGGA	

### Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
300	0	+	0.95	TTGGTCCTAG	^	TTGTATGCAG	
310	1	+	0.33	TTGTATGCAG	^	TTGCAGTCAA	
316	1	+	0.34	GCAGTTGCAG	^	TCAAAGATGT	
322	1	+	0.19	GCAGTCAAAG	^	ATGTCATAAG	
332	2	+	0.18	ATGTCATAAG	^	AAAGACTGAC	
336	0	+	0.17	CATAAGAAAG	^	ACTGACAGGA	
344	2	+	0.07	AGACTGACAG	^	GAGCTACAAA	
577	0	+	0.33	TGTTTCACAG	^	GGAGTGCAAA	
611	2	+	0.43	ATTCCTACAG	^	GGGCAGAGTA	

### Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
742	34	2	-	0.00	TTTCCTTTAG	^	TAATATCAAA	
433	343	0	-	0.17	TGTTCCCCAG	^	TTGGACAATG	
239	537	0	-	0.25	CTCTTTGTAG	^	TTCAATGGTG	
59	717	2	-	0.00	AACCCTCTAG	^	GGAGGGCCAC	

Desaparece un sitio *acceptor* que está dentro del exón y que no se utiliza en el *splicing* normal con *confidence* muy baja, por lo que no estará afectando al *splicing*.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 85.53.81.31.12789.0 :

Start	End	Score	Exon	Intron
469	483	1.00	accacag	gt aagcag
484	498	0.75	gagagag	gt tagagg

Donor site predictions for 85.53.81.31.12829.0 :

Start	End	Score	Exon	Intron
469	483	1.00	accacag	gt aagcag
484	498	0.75	gagagag	gt tagagg

Acceptor site predictions for 85.53.81.31.12789.0 :

Start	End	Score	Intron	Exon
280	320	0.97	cttgttcttttttggtcct	agttgtatgcagttgcagtcaa
557	597	0.94	cattaggcgtttgtttcac	aggagtgcaaacgctgccctt
591	631	0.93	tgcccttcaatattcctac	aggggcagagtagtgcagaggc

Acceptor site predictions for 85.53.81.31.12829.0 :

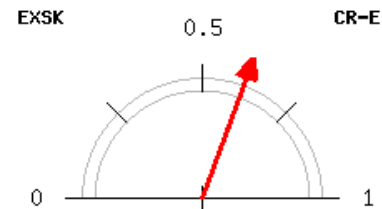
Start	End	Score	Intron	Exon
280	320	0.97	cttgttcttttttggtcct	agttgtatgcagttgcagtcaa
557	597	0.94	cattaggcgtttgtttcac	aggagtgcaaacgctgccctt
591	631	0.93	tgcccttcaatattcctac	aggggcagagtagtgcagaggc

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
aagat(c/g)tcata	aagatc	aagatg	27800	60%

## CRYP-SKIP

Exon length (bp)	175
PESS ( $\leq -2.62$ ) density	1.14
NN 5'ss score density	0.16
SF2/ASF score density	8.76
FAS-ESS (hex2) density	1.14
EIE score density	505.35
Probability of cryptic splice site activation ( $P_{CR-E}$ )	0.62




Parece que hay un sitio críptico de *splicing* dentro del propio exón, pero el cambio de interés no lo toma en consideración, por lo que no debe considerar que tenga algún efecto en el *splicing*.

>wt



## Human Splicing Finder

<div>  New Donor splice site </div> <div> Activation of a cryptic Donor site. Potential alteration of splicing </div>			
Algorithm/Matix	position	sequences	variation
HSF Donor site (matrix GT)	chr11:121621085	- REF : GATC <b>C</b> TCATA - ALT : GATG <b>T</b> CATA	42.64 > 69.78 => 63.65%

## SVM-BPfinder

seq_id	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	15	172	cgctcaaaag		-1.80124240521	0.38921556886	103	9	14	-6.5670917	mut	15	172	cgctcaaaag		-1.80124240521	0.38323352934	103	9	14	-6.5690257
wt	15	163	attctcataa		-0.555356661457	0.386075949367	94	9	14	-5.5105997	wt	15	163	attctcataa		-2.63919560734	0.386075949367	94	9	14	-6.3265211
wt	15	160	tcataagaa		-2.48718769318	0.387096774194	91	9	14	-6.0767781	mut	15	160	tcataagaa		-2.48718769318	0.387096774194	91	9	14	-6.0767781
wt	15	150	gactgacag		2.58687416211	0.4	81	9	14	-3.4528925	wt	15	150	gactgacag		2.58687416211	0.4	81	9	14	-3.4528925
wt	15	133	aagtaaaat		-1.08163594501	0.4140625	64	9	14	-3.8086744	mut	15	133	aagtaaaat		-1.08163594501	0.4140625	64	9	14	-3.8086744
wt	15	123	ccgtaaacg		1.81098911944	0.406779661017	54	9	14	-2.0454439	mut	15	123	ccgtaaacg		1.81098911944	0.406779661017	54	9	14	-2.0454439
wt	15	100	cccttaaca		-0.819723204664	0.378947368421	31	9	14	-1.6286197	wt	15	100	cccttaaca		-0.819723204664	0.378947368421	31	9	14	-1.6286197
wt	15	99	cccttaaca		0.55597581725	0.382978723404	30	9	14	-1.0253682	mut	15	99	cccttaaca		0.55597581725	0.382978723404	30	9	14	-1.0253682
wt	15	67	acatcattg		-2.94150739129	0.370967741935	62	0	0	-4.5546455	mut	15	67	acatcattg		-2.94150739129	0.370967741935	62	0	0	-4.5546455
wt	15	46	acatgagca		-1.95345905536	0.341463414634	41	0	0	-2.8480421	wt	15	46	acatgagca		-1.95345905536	0.341463414634	41	0	0	-2.8480421
wt	15	28	gcataaaaa		-1.97889410212	0.304347826087	23	0	0	-1.7306187	mut	15	28	gcataaaaa		-1.97889410212	0.304347826087	23	0	0	-1.7306187
wt	15	22	aaattacca		-0.42873957248	0.294117647059	17	0	0	-0.747174	wt	15	22	aaattacca		-0.42873957248	0.294117647059	17	0	0	-0.747174

## Variant Effect Predictor tool

ENST00000260197.11:c.5914C>G	<a href="#">11:121621088-121621088</a>	G	missense_variant	SORL1	<a href="#">ENSG00000137642</a>	Transcript	<a href="#">ENST00000260197.12</a>	protein_coding	44/48	5932	5914	1972	L/V	CTC/GTC	<a href="#">rs766895956</a> , <a href="#">COSV52750647</a>
ENST00000260197.11:c.5914C>G	<a href="#">11:121621088-121621088</a>	G	non_coding_transcript_exon_variant	SORL1	<a href="#">ENSG00000137642</a>	Transcript	<a href="#">ENST00000524633.1</a>	retained_intron	4/5	1080	-	-	-	-	<a href="#">rs766895956</a> , <a href="#">COSV52750647</a>
ENST00000260197.11:c.5914C>G	<a href="#">11:121621088-121621088</a>	G	missense_variant	SORL1	<a href="#">ENSG00000137642</a>	Transcript	<a href="#">ENST00000525532.5</a>	protein_coding	24/28	2984	2746	916	L/V	CTC/GTC	<a href="#">rs766895956</a> , <a href="#">COSV52750647</a>
ENST00000260197.11:c.5914C>G	<a href="#">11:121621088-121621088</a>	G	missense_variant	SORL1	<a href="#">ENSG00000137642</a>	Transcript	<a href="#">ENST00000527934.1</a>	protein_coding	15/19	2180	1759	587	L/V	CTC/GTC	<a href="#">rs766895956</a> , <a href="#">COSV52750647</a>
ENST00000260197.11:c.5914C>G	<a href="#">11:121621088-121621088</a>	G	non_coding_transcript_exon_variant	SORL1	<a href="#">ENSG00000137642</a>	Transcript	<a href="#">ENST00000528339.5</a>	retained_intron	6/6	3551	-	-	-	-	<a href="#">rs766895956</a> , <a href="#">COSV52750647</a>
ENST00000260197.11:c.5914C>G	<a href="#">11:121621088-121621088</a>	G	missense_variant	SORL1	<a href="#">ENSG00000137642</a>	Transcript	<a href="#">ENST00000532694.5</a>	protein_coding	21/25	2833	2452	818	L/V	CTC/GTC	<a href="#">rs766895956</a> , <a href="#">COSV52750647</a>
ENST00000260197.11:c.5914C>G	<a href="#">11:121621088-121621088</a>	G	missense_variant	SORL1	<a href="#">ENSG00000137642</a>	Transcript	<a href="#">ENST00000534286.5</a>	protein_coding	21/25	2655	2644	882	L/V	CTC/GTC	<a href="#">rs766895956</a> , <a href="#">COSV52750647</a>
ENST00000260197.11:c.5914C>G	<a href="#">11:121621088-121621088</a>	G	non_coding_transcript_exon_variant	SORL1	<a href="#">ENSG00000137642</a>	Transcript	<a href="#">ENST00000534754.5</a>	retained_intron	5/5	2076	-	-	-	-	<a href="#">rs766895956</a> , <a href="#">COSV52750647</a>

## ESEfinder

Para la búsqueda de sitios de *splicing*, solo se obtienen resultados positivos para las matrices 3'SS en tres resultados que después, comparando con la secuencia mutante, se reducen excepto en 296 que se mantienen. Por lo tanto, puede que se esté debilitando un sitio *acceptor*.

296 (-480)	cctagTTGTATGCAGTTGCAGTCAAAGATC	-13.84880	296 (-480)	cctagTTGTATGCAGTTGCAGTCAAAGATC	4.31180	296 (-480)	cctagTTGTATGCAGTTGCAGTCAAAGATC	-14.78990	296 (-480)	cctagTTGTATGCAGTTGCAGTCAAAGATC	4.25310
302 (-474)	TGTATGCAGTTGCAGTCAAAGATCTCATAA	-16.60300	302 (-474)	TGTATGCAGTTGCAGTCAAAGATCTCATAA	1.67830	302 (-474)	TGTATGCAGTTGCAGTCAAAGATCTCATAA	-17.43980	302 (-474)	TGTATGCAGTTGCAGTCAAAGATCTCATAA	1.62000
308 (-468)	CAGTTGCAGTCAAAGATCTCATAAGAAAGA	-11.74430	308 (-468)	CAGTTGCAGTCAAAGATCTCATAAGAAAGA	0.18170	308 (-468)	CAGTTGCAGTCAAAGATCTCATAAGAAAGA	-12.45650	308 (-468)	CAGTTGCAGTCAAAGATCTCATAAGAAAGA	0.01670

296 (-480)	cctagTTGTATGCAGTTGCAGTCAAAGATG	-13.87080	296 (-480)	cctagTTGTATGCAGTTGCAGTCAAAGATG	4.34560	296 (-480)	cctagTTGTATGCAGTTGCAGTCAAAGATG	-14.81150	296 (-480)	cctagTTGTATGCAGTTGCAGTCAAAGATG	4.28370
302 (-474)	TGTATGCAGTTGCAGTCAAAGATGTCATAA	-16.86000	302 (-474)	TGTATGCAGTTGCAGTCAAAGATGTCATAA	1.80740	302 (-474)	TGTATGCAGTTGCAGTCAAAGATGTCATAA	-17.66180	302 (-474)	TGTATGCAGTTGCAGTCAAAGATGTCATAA	1.74310
308 (-468)	CAGTTGCAGTCAAAGATGTCATAAGAAAGA	-8.13120	308 (-468)	CAGTTGCAGTCAAAGATGTCATAAGAAAGA	0.08100	308 (-468)	CAGTTGCAGTCAAAGATGTCATAAGAAAGA	-8.82740	308 (-468)	CAGTTGCAGTCAAAGATGTCATAAGAAAGA	-0.08220

En las predicciones para las ESE donde se encuentra la posición de interés (319 a 325), no hay ningún resultado destacable.

315 (-461)	AGTCAAA	-4.76835	319 (-457)	AAAGATC	-6.10163	315 (-461)	AGTCAAAG	1.35833	319 (-457)	AAAGATC	-1.58144
316 (-460)	GTCAAAG	-1.93899	320 (-456)	AAGATCT	-1.80290	316 (-460)	GTCAAAGA	-1.41327	320 (-456)	AAGATCT	-4.10283
317 (-459)	TCAAAGA	-2.92060	321 (-455)	AGATCTC	-3.82989	317 (-459)	TCAAAGAT	-6.01389	321 (-455)	AGATCTC	-2.53356
318 (-458)	CAAAGAT	0.74398	322 (-454)	GATCTCA	-3.43105	318 (-458)	CAAAGATC	-4.08117	322 (-454)	GATCTCA	-2.42701
319 (-457)	AAAGATC	-7.98508	323 (-453)	ATCTCAT	-1.83400	319 (-457)	AAAGATCT	-4.47613	323 (-453)	ATCTCAT	-4.50400
320 (-456)	AAGATCT	-1.61703	324 (-452)	TTCATA	-4.31647	320 (-456)	AAGATCTC	-3.68639	324 (-452)	TTCATA	0.87284
321 (-455)	AGATCTC	-5.45087	325 (-451)	CTCATAA	1.37821	321 (-455)	AGATCTCA	0.63471	325 (-451)	CTCATAA	-1.50421
322 (-454)	GATCTCA	-3.39325	326 (-450)	TCATAAG	-5.93910	322 (-454)	GATCTCAT	1.47935	326 (-450)	TCATAAG	3.37194
323 (-453)	ATCTCAT	-2.42047	327 (-449)	CATAAGA	1.41773	323 (-453)	ATCTCATA	2.09437	327 (-449)	CATAAGA	-1.29704
324 (-452)	TTCATA	-6.03417	328 (-448)	ATAAGAA	-1.91870	324 (-452)	TTCATAA	-1.87743	328 (-448)	ATAAGAA	-2.11042
325 (-451)	CTCATAA	0.36510	329 (-447)	TAAGAAA	-4.88288	325 (-451)	CTCATAAG	-1.98182	329 (-447)	TAAGAAA	-0.61620

315 (-461)	AGTCAAA	-4.76835	319 (-457)	AAAGATG	-5.11260	315 (-461)	AGTCAAAG	1.35833	319 (-457)	AAAGATG	-0.73880
316 (-460)	GTCAAAG	-1.93899	320 (-456)	AAGATGT	-0.15559	316 (-460)	GTCAAAGA	-1.41327	320 (-456)	AAGATGT	-3.79721
317 (-459)	TCAAAGA	-2.92060	321 (-455)	AGATGTC	-4.03782	317 (-459)	TCAAAGAT	-6.01389	321 (-455)	AGATGTC	-3.35114
318 (-458)	CAAAGAT	0.74398	322 (-454)	GATGTCA	-4.74500	318 (-458)	CAAAGATG	-1.82101	322 (-454)	GATGTCA	-4.15133
319 (-457)	AAAGATG	-6.50523	323 (-453)	ATGTCAT	-2.21527	319 (-457)	AAAGATGT	-6.15169	323 (-453)	ATGTCAT	-4.71724
320 (-456)	AAGATGT	0.96071	324 (-452)	TGTCATA	-3.33821	320 (-456)	AAGATGTC	-4.60170	324 (-452)	TGTCATA	0.05130
321 (-455)	AGATGTC	-6.05499	325 (-451)	GTCATAA	-1.51999	321 (-455)	AGATGTCA	-0.26297	325 (-451)	GTCATAA	-3.65186
322 (-454)	GATGTCA	-5.30828	326 (-450)	TCATAAG	-5.93910	322 (-454)	GATGTGTC	-1.21879	326 (-450)	TCATAAG	3.37194
323 (-453)	ATGTCAT	-2.66670	327 (-449)	CATAAGA	1.41773	323 (-453)	ATGTCATA	-0.21320	327 (-449)	CATAAGA	-1.29704
324 (-452)	TGTCATA	-4.75594	328 (-448)	ATAAGAA	-1.91870	324 (-452)	TGTCATAA	0.15865	328 (-448)	ATAAGAA	-2.11042
325 (-451)	GTCATAA	-1.21221	329 (-447)	TAAGAAA	-4.88288	325 (-451)	GTCATAAG	0.04669	329 (-447)	TAAGAAA	-0.61620

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	22	319.0874	12	-17.1219	3	21	72	884.3681	67	89.8206	37	163	0.23
mut	2	1	0	24	341.7336	14	-19.0755	3	22	72	890.3440	67	89.3848	41	164	0.25

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

HOT-SKIP

93	25	C	CAAAGATCTCATAAG	AAGATCTCATA	0	0	0	0	0.0000	0	-0.7533	0	1	1	10.6478	2	2.4358	0	4	0.00
94	25	A	CAAAGATATCATAAG	AAGATATCATA	0	0	0	0	0.0000	0	-0.8437	0	1	2	17.7090	2	2.4136	0	5	0.00
95	25	T	CAAAGATTTCATAAG	AAGATTTCATA	0	0	0	1	13.1403	2	-2.9666	0	0	1	12.9383	1	0.8005	3	2	1.50
96	25	G	CAAAGATGTCATAAG	AAGATGTCATA	0	0	0	2	22.6462	2	-2.7069	0	2	1	16.6237	2	2.0000	4	5	0.80