

## Ejemplo comparación de resultados predictores in silico

Cambio de estudio NLRX1 c.2627G>T (chr11:119183138 G/T, COSV52705323 o NM\_001282144.2: c.2627G>T)

Exón 10 e intrones adyacentes:

```
ctgatcgcactctgcagccaggagatgagttgtgaggccccctgactttccatccatcta  
cccccgggccctccttctcagagctctactgaatggcatcgactttctctctctctgcag  
CCTCTACTTCAATGAGCTGAGCTCAGAGGGCCGCCAGGTCTTGCGAGACTTGGGGGGTGC  
TGCTGAAGGTGGTGCCCGGGTGGTGGTGTCACTGACAGAGGGGACGGCGGGTGTCAGAATA  
CTGGTCAGTGATCCTCAGTGAAGTCCAGCGGAACCTCAATAGCTGGGATCGGGCCCGGGT  
TCAGCGGACACCTTGAGCTCCTACTGCGGGATCTGGAAGATAGCGGGGTGCCACCCTTAA  
TCCITGGCGCAAGGCCAGCTGCTCGAGTGGAGGGCGAGGTCAAGGCCCTCTGGAGCA  
GCTGGGAAGCTCTGGAAGCTGAGACACTGGCGGCAGGCACCTAGCTATGTGACCACTGGC  
CCTAAACCTTTTCCCTCTGTGGCCTCCTGGCTTGCACTGCTCCTCTAGAAAGATTCTTT  
CAGGTCTGGAGGCAGAGGAATGGGCATAGCTGAGCCAGTTGCCCTCCTAGGGCATGTTTG  
ACCAGGACTGAGTCTGGAATCTCCAAGTTAAAGATGGTGAATCAATGCTTCGGGCTTGGA  
GATGGAACATGCCTCTCTCATTCAGCTAGAAGGACCAAAGCATGTGGCATTGGATGG  
CCAGAGTGCCTGAAGCACCCTCAACCTTGCTCCCCCTCCTCTCAAAGAGCCTCTG  
ACTGTGTCACCAAGGGGCTCACATCTTATGTCTGCCATGCCAGGGGTGTGCATCCAGA  
TGTGTTGGAAGCTTCCCCTCCTGCTTATGCTCACCTGTGGACACCGAGGATGCCCTCAC  
ATTGGTCTTTTCTCTCATCCTCATGCCCTTTTGCCACAATGGTATGATGGCTTGTTAG  
CCCCTCGAGGCAGATGCACCTGACTTGCTGCTATTAAAAAGCCGTGTGCCTTCTACCA  
A  
ttgtggcctcttctttactgcctgggtgcttccaagaaagggcactgttcttggtctttct  
ggttaaactaccaaccatgatctattgcaagtggttcctggggcactaggttcagaggcag
```

El cambio se encuentra en la primera línea del exón 10 (la **g** 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'
401	0	+	0.61	GGAGGGCGAG	^	GTGAGGGCCC	
944	0	+	0.34	TGCCACAATG	^	GTATGATGGC	

### Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
875	265	2	-	0.47	GTGTCCACAG	^	GTGAGCATAA	
749	391	0	-	0.32	GGGAGGCAAG	^	GTTGGTAGTG	

### Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
81	2	+	0.00	TCCTTCTCAG	^	AGCTCTACTG	
120	2	+	0.71	CTCCTGCCAG	^	CCTCTACTTC	
236	1	+	0.17	GCGGTGTGAG	^	AATACTGGTC	
248	1	+	0.07	TACTGGTCAG	^	TGATCCTCAG	
258	2	+	0.07	TGATCCTCAG	^	TGAAGTCCAG	
590	2	+	0.33	GCCCTCCTAG	^	GGCATGTTTG	
687	2	+	0.20	CTCCATTGAG	^	CTAGAAGGAC	
691	1	+	0.16	ATTCAGTAG	^	AAGGACCAAA	
823	2	+	0.27	GCCATGCCAG	^	GGGTGTCGCC	

### Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
688	452	0	-	0.16	GTCTTCTAG	^	CTGAATGGAG	
546	594	1	-	0.33	CTGCCTCCAG	^	ACCTGAAGGA	
432	708	2	-	0.18	CAGCTTCCAG	^	AGCTTCCCAG	
430	710	1	-	0.17	GCTTCCAGAG	^	CTTCCCAGCT	
422	718	0	-	0.23	AGCTTCCCAG	^	CTGCTCCAGG	
413	727	0	-	0.20	GCTGCTCCAG	^	GAGGGCCCTG	
410	730	0	-	0.20	GCTCCAGGAG	^	GGCCCTGACC	
383	757	0	-	0.76	CCACTCGCAG	^	CAGCTGGGCC	
380	760	0	-	0.44	CTCGCAGCAG	^	CTGGGCCCTT	
363	777	2	-	0.07	TTGCGCCAAG	^	GATTAAGGGT	

### Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
401	0	+	0.61	GGAGGGCGAG	^	GTGAGGGCCC	
944	0	+	0.34	TGCCACAATG	^	GTATGATGGC	

### Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
875	265	2	-	0.47	GTGTCCACAG	^	GTGAGCATAA	
749	391	0	-	0.32	GGGAGGCAAG	^	GTTGGTAGTG	

### Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
81	2	+	0.00	TCCTTCTCAG	^	AGCTCTACTG	
120	2	+	0.71	CTCCTGCCAG	^	CCTCTACTTC	
236	1	+	0.07	GCGGTGTGAG	^	AATACTGGTC	
248	1	+	0.07	TACTGGTCAG	^	TGATCCTCAG	
590	2	+	0.33	GCCCTCCTAG	^	GGCATGTTTG	
687	2	+	0.20	CTCCATTGAG	^	CTAGAAGGAC	
691	1	+	0.16	ATTCAGTAG	^	AAGGACCAAA	
823	2	+	0.27	GCCATGCCAG	^	GGGTGTCGCC	

### Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
688	452	0	-	0.16	GTCTTCTAG	^	CTGAATGGAG	
546	594	1	-	0.33	CTGCCTCCAG	^	ACCTGAAGGA	
432	708	2	-	0.18	CAGCTTCCAG	^	AGCTTCCCAG	
430	710	1	-	0.17	GCTTCCAGAG	^	CTTCCCAGCT	
422	718	0	-	0.23	AGCTTCCCAG	^	CTGCTCCAGG	
413	727	0	-	0.20	GCTGCTCCAG	^	GAGGGCCCTG	
410	730	0	-	0.20	GCTCCAGGAG	^	GGCCCTGACC	
383	757	0	-	0.76	CCACTCGCAG	^	CAGCTGGGCC	
380	760	0	-	0.44	CTCGCAGCAG	^	CTGGGCCCTT	
363	777	2	-	0.07	TTGCGCCAAG	^	GATTAAGGGT	
183	957	1	-	0.17	CCACCTTCAG	^	CAGCACCCCC	
180	960	1	-	0.17	CCTTCAGCAG	^	CACCCCCCAA	
169	971	0	-	0.07	ACCCCCCAAG	^	TCTCGCAAGA	

Desaparece uno de los sitios *acceptor* en la secuencia mutante. Este está dentro del exón y no participa en el *splicing* normal, por lo que la mutación no tendrá efecto en este.

## Splice Site Prediction by Neural Network (NNSplice)

### Donor site predictions for 10.42.2.148.573365.0 :

Start	End	Score	Exon	Intron
237	251	0.50	aatactg	<b>gt</b> cagtga
630	644	0.42	aaagatg	<b>gt</b> gaatca
937	951	0.85	cacaatg	<b>gt</b> atgatg

### Donor site predictions for 10.42.3.123.573377.0 :

Start	End	Score	Exon	Intron
237	251	0.50	aatactg	<b>gt</b> cagtga
630	644	0.42	aaagatg	<b>gt</b> gaatca
937	951	0.85	cacaatg	<b>gt</b> atgatg

### Acceptor site predictions for 10.42.2.148.573365.0 :

Start	End	Score	Intron	Exon
61	101	0.94	ccctcgggcccctccttctc	<b>ag</b> agctctactgaatggcatcg
100	140	0.78	cgactttctctctcctgcc	<b>ag</b> cctctacttcaatgagctga
509	549	0.67	ggcttgcaactgctccctct	<b>ag</b> aaagattccttcaggtctgg
523	563	0.65	cctctagaaagattccttc	<b>ag</b> gtctggaggcagagggaatgg
570	610	0.43	ctgagccagttgccctcct	<b>ag</b> ggcatgtttgaccaggactg
667	707	0.88	acatgcctcctctccattc	<b>ag</b> ctagaaggaccaaagcatgt
752	792	0.89	tgcctccccctcctctcaa	<b>ag</b> agcctctgactgtgtcacca
803	843	0.73	atcttatgtctgccatgcc	<b>ag</b> gggtgtcgccatccagatgt

### Acceptor site predictions for 10.42.3.123.573377.0 :


Start	End	Score	Intron	Exon
61	101	0.94	ccctcgggcccctccttctc	<b>ag</b> agctctactgaatggcatcg
100	140	0.78	cgactttctctctcctgcc	<b>ag</b> cctctacttcaatgagctga
509	549	0.67	ggcttgcaactgctccctct	<b>ag</b> aaagattccttcaggtctgg
523	563	0.65	cctctagaaagattccttc	<b>ag</b> gtctggaggcagagggaatgg
570	610	0.43	ctgagccagttgccctcct	<b>ag</b> ggcatgtttgaccaggactg
667	707	0.88	acatgcctcctctccattc	<b>ag</b> ctagaaggaccaaagcatgt
752	792	0.89	tgcctccccctcctctcaa	<b>ag</b> agcctctgactgtgtcacca
803	843	0.73	atcttatgtctgccatgcc	<b>ag</b> gggtgtcgccatccagatgt

## Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
gctga(g/t)ctcag	gctgag	gctgat	29906	73%

CRYP-SKIP

Human Splicing Finder

 Alteration of auxiliary sequences	Significant alteration of ESE / ESS motifs ratio (-2)	
Algorithm/Matix	position	sequence
IIE (New ESS Site)	chr11:119183133	GCTGAT
Sironi_motif1 (ESS Site Broken)	chr11:119183133	GCTGAGCT
ESE_ASFB (ESE Site Broken)	chr11:119183134	CTGAGCT
IIE (New ESS Site)	chr11:119183135	TGATCT
ESE_SC35 (New ESE Site)	chr11:119183136	GATCTCAG
Sironi_motif3 (New ESS Site)	chr11:119183136	GATCTCAG
ESE_SRp40 (New ESE Site)	chr11:119183138	TCTCAGA
Sironi_motif3 (New ESS Site)	chr11:119183138	TCTCAGAG

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr	
wt	19	37	acttcaatg		-2.14672886083	0.4375	32	0	0	-2.3230145
wt	19	33	caatgagct		0.0544561822258	0.464285714286	28	0	0	-1.1993023
wt	19	28	agctgagct		0.801705935009	0.478260869565	23	0	0	-0.58571349
wt	19	23	agctcagag		-0.704675079044	0.444444444444	18	0	0	-0.86996297
mut	19	37	acttcaatg		-2.14672886083	0.46875	32	0	0	-2.3129217
mut	19	33	caatgagct		0.0544561822258	0.5	28	0	0	-1.1877676
mut	19	28	agctgatct		1.34433483292	0.521739130435	23	0	0	-0.35920644
mut	19	23	atctcagag		-0.696781786491	0.444444444444	18	0	0	-0.86687239

Los BP en amarillo son distintos por la presencia de la mutación pero siguen teniendo puntuaciones negativas por lo que no los tendremos en cuenta.

Variant Effect Predictor tool

ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	missense_variant	NLRX1	<a href="#">ENSG00000160703</a>	Transcript	<a href="#">ENST00000292199.6</a>	protein_coding	10/10	2868	2627	876	S/I	AGC/ATC	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	upstream_gene_variant	PDZD3	<a href="#">ENSG00000172367</a>	Transcript	<a href="#">ENST00000322712.4</a>	protein_coding	-	-	-	-	-	-	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	upstream_gene_variant	PDZD3	<a href="#">ENSG00000172367</a>	Transcript	<a href="#">ENST00000355547.10</a>	protein_coding	-	-	-	-	-	-	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	missense_variant	NLRX1	<a href="#">ENSG00000160703</a>	Transcript	<a href="#">ENST00000409109.6</a>	protein_coding	10/10	3253	2627	876	S/I	AGC/ATC	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	missense_variant	NLRX1	<a href="#">ENSG00000160703</a>	Transcript	<a href="#">ENST00000409265.8</a>	protein_coding	10/10	2789	2627	876	S/I	AGC/ATC	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	missense_variant	NLRX1	<a href="#">ENSG00000160703</a>	Transcript	<a href="#">ENST00000409991.5</a>	protein_coding	10/10	2842	2627	876	S/I	AGC/ATC	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	downstream_gene_variant	NLRX1	<a href="#">ENSG00000160703</a>	Transcript	<a href="#">ENST00000469103.3</a>	processed_transcript	-	-	-	-	-	-	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	downstream_gene_variant	NLRX1	<a href="#">ENSG00000160703</a>	Transcript	<a href="#">ENST00000481860.6</a>	retained_intron	-	-	-	-	-	-	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	upstream_gene_variant	PDZD3	<a href="#">ENSG00000172367</a>	Transcript	<a href="#">ENST00000525131.5</a>	protein_coding	-	-	-	-	-	-	<a href="#">COSV52705323</a>
ENST00000409109.5:c.2627G>T	<a href="#">11:119183138-119183138</a> T	intron_variant	NLRX1	<a href="#">ENSG00000160703</a>	Transcript	<a href="#">ENST00000525863.1</a>	protein_coding	-	-	-	-	-	-	<a href="#">COSV52705323</a>

ESEfinder

Se estaría activando un sitio *acceptor* al pasar la puntuación en la matriz 3' de casi 0 a positiva de la secuencia WT a la mutante:

<div>132 (-1008)</div> ATGAGCTGAGCTCAGAGGGCCGCCAGGTCT	-21.72180	<div>132 (-1008)</div> ATGAGCTGAGCTCAGAGGGCCGCCAGGTCT	0.13640	<div>132 (-1008)</div> ATGAGCTGAGCTCAGAGGGCCGCCAGGTCT	-19.84080	<div>132 (-1008)</div> ATGAGCTGAGCTCAGAGGGCCGCCAGGTCT	-0.19220
<div>132 (-1008)</div> ATGAGCTGATCTCAGAGGGCCGCCAGGTCT	-22.23130	<div>132 (-1008)</div> ATGAGCTGATCTCAGAGGGCCGCCAGGTCT	2.67790	<div>132 (-1008)</div> ATGAGCTGATCTCAGAGGGCCGCCAGGTCT	-20.33630	<div>132 (-1008)</div> ATGAGCTGATCTCAGAGGGCCGCCAGGTCT	2.39930

En cuanto a las ESE, se dan cambios entre ambas secuencias:

135 (-1005)	AGCTGAG	-2.72282	135 (-1005)	AGCTGAG	-1.52012	135 (-1005)	AGCTGAGC	-2.28590	135 (-1005)	AGCTGAG	-3.44934
136 (-1004)	GCTGAGC	-6.66263	136 (-1004)	GCTGAGC	-5.30465	136 (-1004)	GCTGAGCT	-3.14383	136 (-1004)	GCTGAGC	0.20583
137 (-1003)	CTGAGCT	1.68524	137 (-1003)	CTGAGCT	2.50546	137 (-1003)	CTGAGCTC	-4.38810	137 (-1003)	CTGAGCT	-3.76409
138 (-1002)	TGAGCTC	-6.35388	138 (-1002)	TGAGCTC	-3.85839	138 (-1002)	TGAGCTCA	-2.13016	138 (-1002)	TGAGCTC	-0.82363
139 (-1001)	GAGCTCA	-1.32946	139 (-1001)	GAGCTCA	-1.74346	139 (-1001)	GAGCTCAG	2.00088	139 (-1001)	GAGCTCA	-3.98853
140 (-1000)	AGCTCAG	-2.11871	140 (-1000)	AGCTCAG	-1.31219	140 (-1000)	AGCTCAGA	0.42205	140 (-1000)	AGCTCAG	-2.63175
141 (-999)	GCTCAGA	-2.54096	141 (-999)	GCTCAGA	-2.56631	141 (-999)	GCTCAGAG	1.43581	141 (-999)	GCTCAGA	0.39106

135 (-1005)	AGCTGAT	-2.34928	135 (-1005)	AGCTGAT	-1.31245	135 (-1005)	AGCTGATC	0.03410	135 (-1005)	AGCTGAT	-5.83107
136 (-1004)	GCTGATC	-8.78266	136 (-1004)	GCTGATC	-6.84795	136 (-1004)	GCTGATCT	-3.95753	136 (-1004)	GCTGATC	-1.81672
137 (-1003)	CTGATCT	-0.23437	137 (-1003)	CTGATCT	0.90160	137 (-1003)	CTGATCTC	-4.24966	137 (-1003)	CTGATCT	-1.45525
138 (-1002)	TGATCTC	-5.89617	138 (-1002)	TGATCTC	-3.79920	138 (-1002)	TGATCTCA	0.33714	138 (-1002)	TGATCTC	-1.48394
139 (-1001)	GATCTCA	-3.39325	139 (-1001)	GATCTCA	-3.43105	139 (-1001)	GATCTCAG	3.73952	139 (-1001)	GATCTCA	-2.42701
140 (-1000)	ATCTCAG	-2.79401	140 (-1000)	ATCTCAG	-2.04167	140 (-1000)	ATCTCAGA	-0.22563	140 (-1000)	ATCTCAG	-2.12226
141 (-999)	TCTCAGA	-3.91414	141 (-999)	TCTCAGA	-2.77316	141 (-999)	TCTCAGAG	-0.61383	141 (-999)	TCTCAGA	2.89539

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	7	27	14	179	2484.1319	59	-95.8501	54	64	288	3087.8652	305	419.7110	286	711	0.40
mut	7	27	14	181	2504.9791	59	-95.8501	54	64	288	3087.8652	303	418.2827	288	709	0.41

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

## HOT-SKIP

ctgatcgactctgcagccaggagatgagttgtgaggccccctgactttccatccatctaccctcgggcccctccttctca  
gagctctactgaatggcatcgactttctctctcctgccagCCTCTACTTCAATGAGCTGAGCTCAGAGGGCCGCCAGGTC  
TTGCGAGACTTGGGGGGTGCTGCTGAAGGTGGTGCCCGGGTGGTGGTGTCAGTACAGAGGGGACGGCGGTGTGAGAATA  
CTGGTCAGTGATCCTCAGTGAAGTCCAGCGGAACCTCAATAGCTGGGATCGGGCCGGGTTGAGCGACACCTTGAGCTCC  
TACTGCGGGATCTGGAAGATAGCCGGGGTGCCACCCTTAATCCTTGGCGCAAGGCCAGCTGCTGCGAGTGGAAGGCGAG  
GTGAGGGCCCTCCTGGAGCAGCTGGGAAGCTCTGGAAGCTGAGACACTGGCGGCAGGCACCTAGCTATGTGACCACTGGC  
CCTAAACCTTTTCCCTCTGTGGCCTCCTGGCTTGCACTGCTCCCTCTAGAAAGATTCTTCAGGTCTGGAGGCAGAGGAA  
TGGGCATAGCTGAGCCAGTTGCCCTCTAGGGCATGTTTGACCAGGACTGAGTCTGGAATCTCCAAGTTAAAGATGGTGA  
ATCAATGCTTCGGGCTTGGAGATGGAACATGCCTCCTCTCCATTGAGCTAGAAGGACCAAGCATGTGGCATTGGATGG  
CCAGAGTGCCCTGAAGCACCCTACCAACCTTGCCCTCCCCCTCCTCTCAAAGAGCCTCTGACTGTGTACCAAGGGGCTC  
ACATCTTATGTCTGCCATGCCAGGGGTGTGCGCATCCAGATGTGTTGGAAGCTTCCCCTCCTGCCTTATGCTCACCTGTG  
GACACCGAGGATGCCCTCACATTGGTGCTTTCTCCTCATCCTCATGCCCCCTTTGCCACAATGGTATGATGGCTTGGTAG  
CCCCCTGAGGCAGATGCACCTGACTTGCTGCTATTAAAAAGCCGTGTGCCTTCTACCAAttgtggcctcttctttactgc  
ctggtgcttccaagaaagggcactgttcttggtcttcttggttaaaactaccaacatgatctattgcaagtgggtcttg  
gcactaggttcagaggcag