

## Ejemplo comparación de resultados predictores in silico

Cambio de estudio: NF1 c.288+1137C>T (chr17:29487248)

### Exón 3 e intrones adyacentes

```
cagtttagttatttgcacatatgaatatgtggctgtgttccagttatattgaaatattt
aataatttgggacccctcttttcaaaacaaaacagcattgttatattggcgttcatttagtg
cttattaaaataagttttcaaaacaaatagttctcttgttaataaaatcacatatgtagtattca
ttttgaggatagtgagtagtctaaattcactgtgtgtgtatagcatgtttttgctataaaa
aattaagactttgatcatttcgtggcataaaagtatagttaaatccccaattcaagattc
tggtacaggtctatattgtgtgtctaaatcctacatatatttttaaagcatgaagcaaaaacag
catcttttactgttaacaggttaattggcagactcttaataaatgccaattctgtttgcctt
agacttttagtttttatgaggattaggtataaaatcagaaataatatatttaagtatagttat
aatctgggaggtaaaatggaagactattgttgatcttttagttcttccacttttcagatgt
gtgttgattggttagcagaaagtgaactaaccttttatgtctgaatatcttttctgttag
AGAATATTGGAGAAGCTGCTGAAAAAATTTATATCTCTCTCAGTTGATTATTTGGAT
ACACTCGAAAAATGTCCTTGGG
gtaagtaaatgtagcttaagttaggcaggttttgtgaatttgaatcttgagaatgatcttat
gtccgaagtacagatgtgtggaccaagaggacagtcctatggacttttgtctgagacatat
aaatatgagttttgttaataatagctgacctggtgacagacaatttttcatgagaaaattt
tctcatcttttgttaaatataaacagatatttaagtttactataaaaatgttatgtttcagcc
acgtatctgtctctcaggttttttaggaaaatatttttatgaagagatcactatttgattta
acggatctctcagagagttgtgggagagtgcgattcagaagtcgaagagtaattctgtat
ttctctctatatgaatatgattttaaatattctacttttgaaattatgtaatgagaact
taaaaattaaacctatataaggaatagatgttaccttgagaacacctgttttatgtagacac
acaaataaggataggtatgtggaggaggagagtaccaaacaattatttctctgtgtttcc
tgtaagttgttaccacagtgattttcaacatgacttaatttgaactttttcaagagtat
ttcctattgtgtaaagaaagtatttctatacatcattatcttaattcttctggaacatg
gtaagagtttatataaaatatataaattaaataacctctgttttaagtgtgcctaattcca
tgtgtattgaaattaaactcatttcttatgttctctcttcccttgggtcagaattcagaat
tttgttttgattcagaattgtctccctcttccatatgttgcagttctctatagtgaac
tagttgaaagagaccttaatatatgaggactgatagttgtgttgagaagtggtgtgga
gcatggggtagggcattgaaaaaaaaatttatattgacacaactgtgtatccctagggaa
ggctgtaagcgaaaaaatgttaagcataatatgtatatatatatttcttttgagacagagt
ctcactctgcggcacaggctgaagtgacagtgccaccctcctcggtcattgcaacctcaac
ttccgggttcaagctattctcatgcctcagcctcccaagtagctgggattacaggcatg
taccaccatgctcagctatttttttttgtattttcataagccaatgggggttcactatgttg
```

Se ha descrito que este cambio produce la inclusión de un exón críptico de 118 pb. La mutación crea un *donor splice site* fuerte el cual interacciona con un pre-existente *acceptor cryptic splice site* fuerte. Esto lleva a la inclusión de un exón crítico que contiene un codón de STOP prematuro (o por sus siglas PTC, *premature termination codon*)<sup>1</sup>.

El cambio se encuentra en la penúltima línea del intrón (la **c** en rojo).

1. Svaasand EK, Engebretsen LF, Ludvigsen T, Brechan W, Sjurgen W (2015) A novel deep intronic mutation introducing a cryptic exon causing neurofibromatosis type 1 in a family with highly variable phenotypes: a case study. Hereditary Genet 4:152

## NetGene2

The sequence: wt has the following composition:

Length: 1884 nucleotides.

30.5% A, 14.7% C, 18.6% G, 36.2% T, 0.0% X, 33.3% G+C

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
192	1	+		0.35	TTTGAGGATA	^	GTGAGTACTT	
491	0	+		0.45	AATCTGGGAG	^	GTAAGTAAAT	
685	1	+		0.79	TCTTGCTGGG	^	GTAAGTAAAT	
1180	1	+		0.53	TATGCGATAG	^	GTATGTGGAG	
1226	0	+		0.45	CTGGTTTCCT	^	GTAAGTTGTT	
1345	0	+		0.53	CTGGAACATG	^	GTAAGAGTTT	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
1136	749	0	-	0.44	TGTTCTCAAG	^	GTAACATCTA	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
536	1	+		0.33	CACATTTTCAG	^	ATGTGTGTTG	
949	2	+		0.14	AGGTTTTTAG	^	GAAAAATATT	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
1647	238	1	-	0.33	TCGCTTACAG	^	CCTTCCTAG	
1335	550	1	-	0.26	CATGTTCCAG	^	GAAGAATTAA	
794	1091	0	-	0.14	TATGTCTCAG	^	ACAAAAGTCC	
787	1098	1	-	0.07	CAGACAAAAG	^	TCCATAGGAC	
780	1105	2	-	0.17	AAGTCCATAG	^	GACTGTCCTC	

The sequence: mut has the following composition:

Length: 1884 nucleotides.

30.5% A, 14.6% C, 18.6% G, 36.3% T, 0.0% X, 33.2% G+C

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
192	1	+		0.35	TTTGAGGATA	^	GTGAGTACTT	
491	0	+		0.45	AATCTGGGAG	^	GTAAGTAAAT	
685	1	+		0.79	TCTTGCTGGG	^	GTAAGTAAAT	
1180	1	+		0.53	TATGCGATAG	^	GTATGTGGAG	
1226	0	+		0.45	CTGGTTTCCT	^	GTAAGTTGTT	
1345	0	+		0.53	CTGGAACATG	^	GTAAGAGTTT	
1820	0	+		0.00	GGGATTACAG	^	GTATGTACCA	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
1136	749	0	-	0.44	TGTTCTCAAG	^	GTAACATCTA	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
536	1	+		0.33	CACATTTTCAG	^	ATGTGTGTTG	
949	2	+		0.14	AGGTTTTTAG	^	GAAAAATATT	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
1647	238	1	-	0.33	TCGCTTACAG	^	CCTTCCTAG	
1335	550	1	-	0.26	CATGTTCCAG	^	GAAGAATTAA	
794	1091	0	-	0.14	TATGTCTCAG	^	ACAAAAGTCC	
787	1098	1	-	0.07	CAGACAAAAG	^	TCCATAGGAC	
780	1105	2	-	0.17	AAGTCCATAG	^	GACTGTCCTC	

Al comparar los resultados para ambas secuencias, vemos que aparece un nuevo sitio *donor* (en rojo) en la hebra directa en la secuencia mutante. Existe un sitio *aceptor* anterior (en verde) que no tiene ningún efecto en el *splicing* en la secuencia *wild type* pero, con este nuevo sitio *donor*, se puede producir la introducción de un exón críptico en la secuencia de 870 pb.

## Splice Site Prediction by Neural Network (NNSplice)

### Donor site predictions for 80.102.134.127.20150.0 :

Start	End	Score	Exon	Intron
185	199	0.80	gaggata	<b>gt</b> gagtac
484	498	0.69	ctgggag	<b>gt</b> aaaatg
678	692	1.00	tgctggg	<b>gt</b> aagtaa
1173	1187	0.99	gcgatag	<b>gt</b> atgtgg
1219	1233	0.94	gtttcct	<b>gt</b> aagttg
1338	1352	1.00	gaacatg	<b>gt</b> aagagt
1656	1670	0.98	aaaaaat	<b>gt</b> aagcat

### Donor site predictions for 80.102.134.127.20162.0 :

Start	End	Score	Exon	Intron
185	199	0.80	gaggata	<b>gt</b> gagtac
484	498	0.69	ctgggag	<b>gt</b> aaaatg
678	692	1.00	tgctggg	<b>gt</b> aagtaa
1173	1187	0.99	gcgatag	<b>gt</b> atgtgg
1219	1233	0.94	gtttcct	<b>gt</b> aagttg
1338	1352	1.00	gaacatg	<b>gt</b> aagagt
1656	1670	0.98	aaaaaat	<b>gt</b> aagcat
1813	1827	1.00	attacag	<b>gt</b> atgtac

### Acceptor site predictions for 80.102.134.127.20150.0 :

Start	End	Score	Intron	Exon
288	328	0.50	caattcaagattctggtac	<b>ag</b> gtctatatgtgtgtctaaat
402	442	0.95	tgccatttctgtttgcctt	<b>ag</b> actttagttttatgaggat
516	556	0.97	ttttagtctttcacttttc	<b>ag</b> atgtgtgttgattggtagca
580	620	0.79	tctgaatatctttctgtt	<b>ag</b> agaatatttgagaagctgc
625	665	0.53	aaaaatttatatctctctc	<b>ag</b> ttgattatattggatacact
921	961	0.68	agccacgtatctgtctctc	<b>ag</b> gtttttaggaaaatatttta
1434	1474	0.46	ttctctcttcctttggctc	<b>ag</b> aattcagaattttgtttga
1459	1499	0.79	cagaattttgttttgattc	<b>ag</b> aattgcttcctcttcacat
1498	1538	0.85	atatgttgagttctctat	<b>ag</b> tgaactagtgtgaagagac
1681	1721	0.70	tatatctttcttttgagac	<b>ag</b> agtcctcactctgcggcacag

### Acceptor site predictions for 80.102.134.127.20162.0 :

Start	End	Score	Intron	Exon
288	328	0.50	caattcaagattctggtac	<b>ag</b> gtctatatgtgtgtctaaat
402	442	0.95	tgccatttctgtttgcctt	<b>ag</b> actttagttttatgaggat
516	556	0.97	ttttagtctttcacttttc	<b>ag</b> atgtgtgttgattggtagca
580	620	0.79	tctgaatatctttctgtt	<b>ag</b> agaatatttgagaagctgc
625	665	0.53	aaaaatttatatctctctc	<b>ag</b> ttgattatattggatacact
921	961	0.68	agccacgtatctgtctctc	<b>ag</b> gtttttaggaaaatatttta
1434	1474	0.46	ttctctcttcctttggctc	<b>ag</b> aattcagaattttgtttga
1459	1499	0.79	cagaattttgttttgattc	<b>ag</b> aattgcttcctcttcacat
1498	1538	0.85	atatgttgagttctctat	<b>ag</b> tgaactagtgtgaagagac
1681	1721	0.70	tatatctttcttttgagac	<b>ag</b> agtcctcactctgcggcacag

Al comparar los resultados para ambas secuencias (*wild type* a la izquierda, mutante a la derecha), vemos que aparece un nuevo sitio *donor* (en rojo) en la secuencia mutante. Existe un sitio *aceptor* anterior (en verde) que no tiene ningún efecto en el *splicing* en la secuencia *wild type* pero, con este nuevo sitio, se puede producir la introducción de un exón críptico en la secuencia de 120 pb.

GENSCAN → no da resultados para este cambio

Predicted genes/exons:

Gn.Ex Type S .Begin ...End .Len Fr Ph I/Ac Do/T CodRg P.... Tscr..  
-----  
NO EXONS/GENES PREDICTED IN SEQUENCE

Predicted genes/exons:

Gn.Ex Type S .Begin ...End .Len Fr Ph I/Ac Do/T CodRg P.... Tscr..  
-----  
NO EXONS/GENES PREDICTED IN SEQUENCE

MaxEntScan

Necesitamos una región 5’ o 3’ del intrón/exón y, como se trata de un cambio intrónico profundo, no se puede emplear esta herramienta.

Spliceman

>sec  
cagtttagtattgcacatatgaatatgtggcgtgtccagttatattgaaatattttaataattgggaccccttttcaaaacaaaacagcattgttatattggcggtcatttagtgcctttaaataaagttttcaaaccatagtcctgttaataaatacatatgtagtattcatttgaggatagtgagtacacttaaaatcactgtgggtctatagcatgttttgcataaaaataagacatttgatcatttcgtggcataaaagtatagttaaatccccaattcaagattctgtaca

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
acagg(c>t)atgta	caggca	caggta	28749	66%

En el análisis del efecto del cambio, se obtiene una puntuación elevada (66%) para el cambio C>T, por lo que puede estar afectando al *splicing*.

CRYP-SKIP

Se emplea para el exón y las regiones flanqueantes a este, por lo que este predictor no va a ser útil para una variante intrónica profunda.

Human Splicing Finder

Impact Prediction

Format: text Download Copy

Type	Interpretation
New Donor splice site	Activation of a cryptic Donor site. Potential alteration of splicing

Algorithm/Matix	position	sequences	variation
HSF Donor site (matrix GT)	chr17:31160226	- REF : CAGG <b>C</b> ATGT - ALT : CAGG <b>T</b> ATGT	65.66 > 92.8 => 41.33%
MaxEnt Donor site	chr17:31160226	- REF : CAGG <b>C</b> ATGT - ALT : CAGG <b>T</b> ATGT	2.05 > 9.8 => 378.05%

Con este cambio se produce la activación de un sitio *donor* críptico que puede alterar el *splicing* si existe un sitio *acceptor* anterior que no se haya utilizado, produciendo la inclusión de un exón críptico al mRNA.

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr
mut	24	88	ggttcaagc	-1.25556085582	0.614457831325	3	23	33	0.22614623
mut	24	77	ttctcatgc	1.73973944427	0.611111111111	1	14	19	1.3940418
mut	24	70	gcctcagcc	0.909174006341	0.6	46	19	40	-1.5875441
mut	24	48	ggattacag	-1.54965413017	0.651162790698	24	19	40	-1.1412044
mut	24	26	tgctcagct	1.19026800706	0.761904761905	2	19	40	1.3599341

Los BP predichos para esa región no incluyen el fragmento con el cambio de interés, por lo que no debe estar afectando al *splicing*.

## IntSplice

### SNV at chr7:29487248 can't be predicted by IntSplice.

Prediction shows either Abnormal or Normal.

Prediction	Genomic Mutation	Ensembl 64 Transcript ID and Exon No.
------------	------------------	---------------------------------------

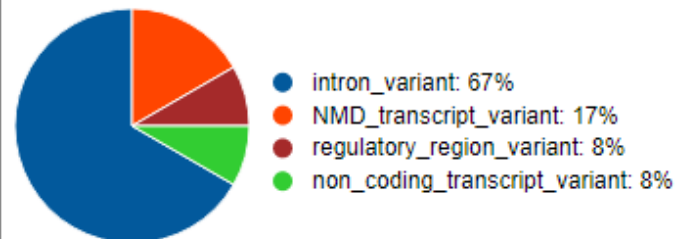
Input queries: hg19, chr7, 29487248

## Variant Predictor Effect tool

### Summary statistics

Category	Count
Variants processed	1
Variants filtered out	0
Novel / existing variants	-
Overlapped genes	1
Overlapped transcripts	8
Overlapped regulatory features	1

### Consequences (all)



Se identifica que tiene efecto en el *splicing*. Si nos fijamos en las consecuencias que produce, el 8% de los resultados indican que es una *non coding transcript exon variant*, es decir, una variante que se encuentra en un transcrito que no se suele transcribir porque no es el mayoritario. El método de NMD (17% de los resultados) se encarga de degradar estos transcritos<sup>5</sup>, dado que aparece un codón de parada (50% de las consecuencias predichas).

Show/hide columns (25 hidden)									
Uploaded variant	Location	Allele	Consequence	Symbol	Gene	Feature type	Feature	Biotype	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	intron_variant	NF1	<a href="#">ENSG00000196712</a>	Transcript	<a href="#">ENST00000356175.7</a>	protein_coding	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	intron_variant	NF1	<a href="#">ENSG00000196712</a>	Transcript	<a href="#">ENST00000358273.9</a>	protein_coding	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	intron_variant	NF1	<a href="#">ENSG00000196712</a>	Transcript	<a href="#">ENST00000431387.8</a>	protein_coding	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	intron_variant	NF1	<a href="#">ENSG00000196712</a>	Transcript	<a href="#">ENST00000487476.5</a>	protein_coding	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	intron_variant, non_coding_transcript_variant	NF1	<a href="#">ENSG00000196712</a>	Transcript	<a href="#">ENST00000489712.6</a>	processed_transcript	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	intron_variant	NF1	<a href="#">ENSG00000196712</a>	Transcript	<a href="#">ENST00000490416.2</a>	protein_coding	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	intron_variant, NMD_transcript_variant	NF1	<a href="#">ENSG00000196712</a>	Transcript	<a href="#">ENST00000495910.6</a>	nonsense_mediated_decay	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	intron_variant, NMD_transcript_variant	NF1	<a href="#">ENSG00000196712</a>	Transcript	<a href="#">ENST00000579081.5</a>	nonsense_mediated_decay	
ENST00000358273.4:c.288+1137C>T	<a href="#">17:31160230-31160230</a>	T	regulatory_region_variant	-	-	RegulatoryFeature	<a href="#">ENSR00001008386</a>	enhancer	

## ESEfinder

Tabular results									
5SS_U2_human threshold: 6.67			3SS_U2_human threshold: 6.632			5SS_U2_mouse threshold: 6.679			BranchSite threshold: 0
Position*/Site/Score			Position*/Site/Score			Position*/Site/Score			Position*/Site/Score
10 (-30)	tagctgggattacaggtatgtaccaccatg	10.70550	10 (-30)	tagctgggattacaggtatgtaccaccatg	4.27490	10 (-30)	tagctgggattacaggtatgtaccaccatg	10.68560	2 (-38) cccccaa 4.12690

Excepto para la última matriz, el cambio aparece en los resultados dados para el mejor sitio de *splicing*. Da mayor puntuación en las matrices de 5'SS, por lo que podría estar generando, con mayor probabilidad, un sitio *acceptor* de *splicing*, aunque no se puede descartar el sitio *donor*.

## EX-SKIP

### EX-SKIP - Results for submitted sequences

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)
mut	8	0	0	30	464.1346	22	-28.8932	9	12	34	508.8407	24	28.2796	60	79	0.76
wt	8	0	0	30	464.1346	22	-28.8932	9	12	34	508.8407	24	28.2796	60	79	0.76

Both alleles have a comparable chance of exon skipping.

## HOT-SKIP

Es un predictor para mutaciones exónicas, por lo que no nos sirve en este caso.



