

Ejemplo comparación de resultados predictores in silico

Cambio de estudio CFTR c.3718-2477C>T (chr7:117280016 G/A, rs949635279 o NM_000492G>A)

Exón 23 e intrones adyacentes:

```
..... .catctgttgcaagtattaaaatggcagagtaagacaccctga
aaggaaatgttctattcatgggtacaatgcaattacagctagcaccatttcaactgtt
taactttcaacatattattttgattatcttgatccaacattctcagggaggagggtgcat
tgaagttattagaaaaactgacttagatttagggatgtcttaaaagcttatttgcggg
aagtactctagccttattcaacagatcactggagaagcctggaaaaacaaatcccggaac
taattattatgtgccagttatataaacaagaagactttgttggtacaaacagtgattc
cttgcccttgaaaaatgtgtcagatcatgtcattaccagcagttcaatgatataaggaa
accaagagtaatagctaaaacctttaaagctaaaccaaagatttacaattgcctcttcac
ccagtccttcccaacctaataaactgagttctcttaaaaatttttagtatttttttctgaaga
aaagggaacatggacatttatctaatctcatttagaaatctgactaatgataaacaggat
ttagacctcaagcacttcttaccaaaattcttgatatgaccttatagcaaattactttca
cctgttgaacttttctttctttttattccctgtacctcacctgcactgggcataattcaag
ttgcttatacaacacttttactatgtgttagaaaaatcatgaacatgatgaatgtgtt
gtgcaacatgagctgattcataaatgaaaatgtgcatgaaattccacaattattttaaaa
ttggagtttatctagcaattgaacaaaattgattaaatccattatttggttagatcagct
aaattacataagtttcattcatctgctcataaatccatccattcttccatctggctatcc
ttagtcaattcaaataaatatttatggggcactttgggttaagccaagtgctaaagaattca
atgcaaaacaagatagactccctgtccttggttgaacttataattttggtacaaacaaa
gcaataatcaagaaaataaaaaaagtagtgattgtgattaataatataagaagaattca
acagagttattgtacttaacatttgattgatctgattttctcagttgtctgagaacaaaca
tttgtgaaaatctcattgttagagttcttacgatggatagggggtcaactgtgtcattatt
gcttatcagcttatcccaagacctaagtttattaccagatttgcaaatagtgttcaataaa
ttactcttattaagggttggtatgtactctaaaacatttatgggggtcccttcactgggtt
ctgggttacaaacttacttttctatgatgacatagtatagaatttgagagtgaatattta
gaagttcatttttattataatttttgaagtattgatatgagtgaaattagaattttaaa
aagaacaaactgtccttcactacagattgaaaagcattatactaaaagaccattttgc
tcagttatagtatataaaaggccaatgacttaaaaaaaaattatgttaaggagaaggaaac
aaccattttatcagtgccactaactgtcagccagttttttcagtggtcagttaatgactg
cagtagtggttctaccttgctcaaaacacccctcctcaagttctggcatctaagctgacatc
agaacacagagtggtggggctctctgtgggtcactctctagcattgatctectcatgcagtg
```

Se ha descrito que este cambio causa la inclusión de un exón críptico (84 pb). La mutación crea un sitio donador nuevo y resulta en la inclusión de un exón críptico que contiene un PTC¹.

El cambio se encuentra en la primera línea del intrón (la primera g en color azul).

Se va a obtener los resultados que produce analizar esta variable con los diferentes predictores y ver cuál de ellos es más preciso.

catggtgctctacgtctatgctatgttcttatggtcttttaggtaacagaataatttct
tttcttttcttactatatacattttgctttctgaaattcccttctcgccaatccagggtgaa
tgtcagaatgtgatttgacaactgtccaaagtactcattcactgaaggagggttaaggcct
tcgcscacactgccccttctctgggaatatactgctgcctgaacatatcattgtttattgcc
aggcttgaacttcaccaaatattttattaggggtcaacatcbaaatattagaactatttc
agattaatttttaagtcgtatccactttgggtactagatcaaattgcagggtctctgcttc
tggcttgagcdtatgtttagagatgatgtgcatgaagacactctttgctttcccttatg
caaatgggcattttcaatctttttgtcattagtaaaaggtcagtgataaagggaagtctgc
atcaggggtccaattccttatggccagtttctctattctgttccaagggtgtttgtctcc
atatatcaacattgggtcaggattgaaagtgtgcaacaagggttgatgaataagtgaaaa
tcttccactggtgacaggataaaatattccaatgggtttttattgaagtacaatactgaat
tatgtttatggcatggtaacctatagtcacagaagtatcccatcacttttaccttatag
GTGGGCTCTTGGGAAGAACTGGATCAGGGAAGAGTACTTGTATCAGCTTTTTGAGA
CTACTGAACACTGAAGGAGAAATCCAGATCGATGGTGTGTCTTGGGATTCAATAACTTTG
CAACACTGGAGGAAAGCCTTTGGACTGATACACAG
gtgagcaaaaggacttagccagaaaaaggcaactaaattatatttttaactgctatttg
atacttgtactcaagaaattcatattactctgcaaaatatatttgttatgcattgctgtc
ttttttctccagtgccagtttttctcataggcagaaaagatgtctctaaaagtttggaattc

NetGene2

Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
195	2	+	0.34	TAGATTTAGG	^	GTATGTCTTA	
978	1	+	0.47	GGCACTTTGG	^	GTAAGCCAGG	
1952	2	+	0.83	CTGAGGAGTG	^	GTAAGGCCCTT	
2239	1	+	0.44	ATTAGTAAAG	^	GTCAAGTGATA	
2657	1	+	0.83	GATACCACAG	^	GTGAGCAAAA	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
2493	344	0	-	0.47	CACCTATAAG	^	GTAAGAGTGA	
1794	1043	1	-	0.45	TAGCATAGAC	^	GTGAGAGCAC	
1377	1460	2	-	0.47	TCATAGAAAA	^	GTAAGTTTGT	
680	2157	1	-	0.53	CCCAGTGCAG	^	GTGAGGTACA	
602	2235	1	-	0.49	AAGAATTTTG	^	GTAAGAAGTG	

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
244	1	+	0.43	TATTCAACAG	^	ATCACTGAGA	
1895	0	+	0.49	GCCAATCCAG	^	GTGAATGTCA	
2022	2	+	0.16	TTATTGCCAG	^	GCTTGAACCT	
2500	1	+	0.83	TACCTTATAG	^	GTGGGCCTCT	
2788	0	+	0.00	TTTTCTCCAG	^	TGCAGTTTTTC	
2793	2	+	0.00	TCCAGTGCAG	^	TTTTCTCATA	
2804	1	+	0.00	TTTCTCATAG	^	GCAGAAAAA	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
2571	266	1	-	0.25	TCTCCTTCAG	^	TGTTTCAGTAG	
2151	686	2	-	0.28	CTAAACATAG	^	GCTCAAGCCA	
1168	1669	0	-	0.23	TTGTTCTCAG	^	ACAACTGAGA	
258	2579	2	-	0.43	GTTTTTCCAG	^	GCTTCTCAGT	

Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
27	-	+	0.00	AAAAATGGCAA	^	GTAAGACACC	
195	2	+	0.34	TAGATTTAGG	^	GTATGTCTTA	
978	1	+	0.47	GGCACTTTGG	^	GTAAGCCAGG	
1952	2	+	0.83	CTGAGGAGTG	^	GTAAGGCCCTT	
2239	1	+	0.44	ATTAGTAAAG	^	GTCAAGTGATA	
2657	1	+	0.83	GATACCACAG	^	GTGAGCAAAA	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
2493	344	0	-	0.47	CACCTATAAG	^	GTAAGAGTGA	
1794	1043	1	-	0.45	TAGCATAGAC	^	GTGAGAGCAC	
1377	1460	2	-	0.47	TCATAGAAAA	^	GTAAGTTTGT	
680	2157	1	-	0.53	CCCAGTGCAG	^	GTGAGGTACA	
602	2235	1	-	0.49	AAGAATTTTG	^	GTAAGAAGTG	

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
244	1	+	0.43	TATTCAACAG	^	ATCACTGAGA	
1895	0	+	0.49	GCCAATCCAG	^	GTGAATGTCA	
2022	2	+	0.16	TTATTGCCAG	^	GCTTGAACCT	
2500	1	+	0.83	TACCTTATAG	^	GTGGGCCTCT	
2788	0	+	0.00	TTTTCTCCAG	^	TGCAGTTTTTC	
2793	2	+	0.00	TCCAGTGCAG	^	TTTTCTCATA	
2804	1	+	0.00	TTTCTCATAG	^	GCAGAAAAA	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
2571	266	1	-	0.25	TCTCCTTCAG	^	TGTTTCAGTAG	
2151	686	2	-	0.28	CTAAACATAG	^	GCTCAAGCCA	
1168	1669	0	-	0.23	TTGTTCTCAG	^	ACAACTGAGA	
258	2579	2	-	0.43	GTTTTTCCAG	^	GCTTCTCAGT	

Aparece un sitio *donor* nuevo en la secuencia mutada (derecha). Si existiera un sitio *donor* anterior a esta posición que no se había utilizado, puede emplearse para incluirse un exón críptico a la secuencia del mRNA.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 85.53.15.54.9590.0 :

Start	End	Score	Exon	Intron
188	202	0.99	atttagg	gtatgtct
971	985	0.98	actttgg	gttaagcca
980	994	0.41	aagccag	gtgctaag
1578	1592	0.94	aaattat	gtaaggag
1816	1830	0.89	tctttag	gttaacaag
1889	1903	0.81	aatccag	gtgaatgt
1945	1959	0.99	aggagt	gttaaggcc
2232	2246	0.96	agtaaa	gtcagtg
2353	2367	0.71	caacaag	gtttgaat
2588	2602	0.50	atcgat	gtgtgtct
2650	2664	1.00	accacag	gtgagcaa

Donor site predictions for 85.53.15.54.9601.0 :

Start	End	Score	Exon	Intron
20	34	0.84	atggcaag	gttaagaca
188	202	0.99	atttagg	gtatgtct
971	985	0.98	actttgg	gttaagcca
980	994	0.41	aagccag	gtgctaag
1578	1592	0.94	aaattat	gtaaggag
1816	1830	0.89	tctttag	gttaacaag
1889	1903	0.81	aatccag	gtgaatgt
1945	1959	0.99	aggagt	gttaaggcc
2232	2246	0.96	agtaaa	gtcagtg
2353	2367	0.71	caacaag	gtttgaat
2588	2602	0.50	atcgat	gtgtgtct
2650	2664	1.00	accacag	gtgagcaa

Aparece un sitio *donor* nuevo en la secuencia mutada (derecha). Si existiera un sitio *donor* anterior a esta posición que no se había utilizado, puede emplearse para incluirse un exón críptico a la secuencia del mRNA.

Acceptor site predictions for 85.53.15.54.9590.0 :

Start	End	Score	Intron	Exon
224	264	0.46	tactctagccttattcaac	agatcactgagaagcctggaaa
444	484	0.58	acaaattgcctcttcatcc	agtctttcccaacctaaaaact
499	539	0.82	ttttagtatttttttctga	agaaaagggaacatggacattt
535	575	0.47	catttatctaactctcatt	agaaatctgactaatgataaca
853	893	0.40	attaaatccattatttgtt	agatcagctaaattacataagt
1143	1183	0.61	tgattgatctgattttctc	agttgtctgagaacaacattt
1229	1269	0.72	tgtgtcattattgtcttatc	agcttatccaaagacctagtt
1488	1528	0.98	caaaactgtccttactac	agattgaaaagcattatactaa
1623	1663	0.59	actgtcagccagtttttc	agtggtcagttaatgactgcag
1802	1842	0.95	ctatgttcttatggctttt	agtaacaagaataattttctt
2002	2042	0.81	catatcattgtttattgcc	aggcttgaacttcaccaaatta
2213	2253	0.94	tttcaatctttttgtcatt	agtaaaggctcagtataaagga
2287	2327	0.40	gtttctctattctgttcca	aggtgtttgtctccatataatc
2480	2520	1.00	cccatcacttttaccttat	aggtgggcctcttgggaagaac
2768	2808	0.94	attgtgtcttttttctcc	agtcagttttctcataggcag
2784	2824	0.86	tccagtgcagttttctcat	agcgagaaaagatgtctctaaa

Acceptor site predictions for 85.53.15.54.9601.0 :

Start	End	Score	Intron	Exon
224	264	0.46	tactctagccttattcaac	agatcactgagaagcctggaaa
444	484	0.58	acaaattgcctcttcatcc	agtctttcccaacctaaaaact
499	539	0.82	ttttagtatttttttctga	agaaaagggaacatggacattt
535	575	0.47	catttatctaactctcatt	agaaatctgactaatgataaca
853	893	0.40	attaaatccattatttgtt	agatcagctaaattacataagt
1143	1183	0.61	tgattgatctgattttctc	agttgtctgagaacaacattt
1229	1269	0.72	tgtgtcattattgtcttatc	agcttatccaaagacctagtt
1488	1528	0.98	caaaactgtccttactac	agattgaaaagcattatactaa
1623	1663	0.59	actgtcagccagtttttc	agtggtcagttaatgactgcag
1802	1842	0.95	ctatgttcttatggctttt	agtaacaagaataattttctt
2002	2042	0.81	catatcattgtttattgcc	aggcttgaacttcaccaaatta
2213	2253	0.94	tttcaatctttttgtcatt	agtaaaggctcagtataaagga
2287	2327	0.40	gtttctctattctgttcca	aggtgtttgtctccatataatc
2480	2520	1.00	cccatcacttttaccttat	aggtgggcctcttgggaagaac
2768	2808	0.94	attgtgtcttttttctcc	agtcagttttctcataggcag
2784	2824	0.86	tccagtgcagttttctcat	agcgagaaaagatgtctctaaa

GENSCAN → no da resultados para este cambio

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

MaxEntScan

MAXENT: -17.30 MDD: -12.10 MM: -8.96 WMM: -8.48 MAXENT: -12.27 MM: -11.90 WMM: -8.76

Spliceman


Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
atggc(g/a)agtaa	atggcg	atggca	28566	65%

En el análisis del efecto del cambio, se obtiene una puntuación elevada (65%) para el cambio, 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

 Alteration of auxiliary sequences	Significant alteration of ESE / ESS motifs ratio (4)	
Algorithm/Matix	position	sequence
EIE (ESE Site Broken)	chr7:117639962	CGCCAT
EIE (New ESE Site)	chr7:117639964	CTCACC
ESE_ASF (New ESE Site)	chr7:117639964	CTCACCA
ESE_ASFB (New ESE Site)	chr7:117639964	CTCACCA
EIE (New ESE Site)	chr7:117639965	ACTCAC
ESE_SRp40 (New ESE Site)	chr7:117639965	ACTCACC

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr				
wt	13	74	gtattaaaa		-3.8497059368	0.420289855072	31	12	19				-2.7550696
wt	13	73	tattaaaaat		-0.837525156699	0.426470588235	30	12	19				-1.5103639
wt	13	61	gagtaagac		-0.537185481398	0.464285714286	18	12	19				-0.62097369
wt	13	51	ccctgaaag		1.16447246111	0.45652173913	8	12	19				0.67578151
wt	13	33	tattcatgg		0.0547516643368	0.464285714286	28	0	0				-1.1991866
wt	13	17	caattacag		-0.324115538156	0.5	12	0	0				-0.32322319
mut	13	74	gtattaaaa		-3.8497059368	0.420289855072	31	12	19				-2.7550696
mut	13	73	tattaaaaat		-0.837525156699	0.426470588235	30	12	19				-1.5103639
mut	13	61	aagtaagac		-0.972380575823	0.464285714286	18	12	19				-0.79137314
mut	13	51	ccctgaaag		1.16447246111	0.45652173913	8	12	19				0.67578151
mut	13	33	tattcatgg		0.0547516643368	0.464285714286	28	0	0				-1.1991866
mut	13	17	caattacag		-0.324115538156	0.5	12	0	0				-0.32322319

La diferencia entre ambas es la el cabio en una de los resultados de **g** a **a**. Sin embargo, como sigue teniendo una puntuación baja, no se considera. Por lo tanto, no tendrá efecto en el *splicing*.

IntSplice

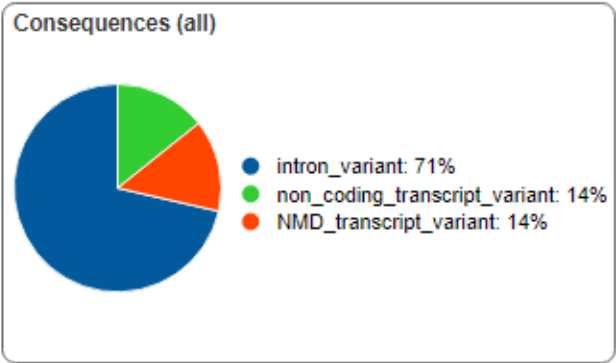
SNV at chr7:117280016 can't be predicted by IntSplice.

Prediction shows either Abnormal or Normal.

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

Variant Effect Predictor tool

Category	Count
Variants processed	1
Variants filtered out	0
Novel / existing variants	0 (0.0) / 1 (100.0)
Overlapped genes	2
Overlapped transcripts	5
Overlapped regulatory features	0



Se trata de una variante intrónica que está afectando al sitio de *splicing*, por lo que va a provocar que se altere el *splicing* normal.

Esto se observa en que en los resultados se obtiene que es una *non-coding transcript 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 (14% de los resultados) se encarga de degradar estos transcritos², dado que aparece un codón de parada

2. Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci USA. 2003;100:189–192

Uploaded variant	Location	Allele	Consequence	Symbol	Gene	Feature type	Feature	Biotype
ENST00000003084.6:c.3719-2477G>A	7:117639962-117639962	A	intron_variant	CFTR	ENSG00000001626	Transcript	ENST00000003084.11	protein_coding
ENST00000003084.6:c.3719-2477G>A	7:117639962-117639962	A	intron_variant	CFTR	ENSG00000001626	Transcript	ENST00000426809.5	protein_coding
ENST00000003084.6:c.3719-2477G>A	7:117639962-117639962	A	intron_variant, non_coding_transcript_variant	AC000061.1	ENSG000000083622	Transcript	ENST00000456270.1	lncRNA
ENST00000003084.6:c.3719-2477G>A	7:117639962-117639962	A	intron_variant, NMD_transcript_variant	CFTR	ENSG00000001626	Transcript	ENST00000647720.1	nonsense_mediated_decay
ENST00000003084.6:c.3719-2477G>A	7:117639962-117639962	A	intron_variant	CFTR	ENSG00000001626	Transcript	ENST00000649781.1	protein_coding

ESEfinder

Cuando se buscan los posibles sitios de *splicing*, solo se obtienen dos resultados con la posición de interés con puntuación positiva en las matrices de 5'SS: tgcagtattaaatggc **g**agtaagacaccc (2.23320 y 2.24520) y gtattaaatggc **g**agtaagacaccctgaa (4.72270 y 4.64940). Si buscamos los resultados equivalentes en las predicciones para la secuencia mutada, vemos que, para la primera las puntuaciones suben un poco (2.92170 y 2.92060), mientras que para la segunda suben considerablemente (7.02820 y 7.01200). Por lo tanto, se está generando un sitio 5'SS mucho más fuerte que el original, lo que hará que sea más probable la alteración del *splicing*.

EX-SKIP

Predictor para secuencias exónicas.

HOT-SKIP

Predictor para secuencias exónicas.