

Ejemplo comparación de resultados predictores in silico

Cambio de estudio BRCA1 c.2106C>G (chr17:43045738 C/G, COSV58786146 o NM_007294.4: c.2106C>G)

Exón 22 e intrones adyacentes:

```
taaaaatacaaaaatttagctgggtgtgatggcatgtgcctgtaattccagctactcagga
ggcagagacaggagaatttgcttgaacccaggaggcgagggttgaatgagccgagattgco
ccatcacactctagcctcggcgacagagcaagactcgtctcaaaaaaaaaaaaaaaaaa
ttagcttctacctcattaatcctaagaactcatacaaccaggacctggagtcgattgat
tagagcctagtccaggagaaatgaattgacactaatctctgcttgcttctctgtctccag
CAATGCGCAGATGTTGTGAGGCACCTGTGGTGACCCGAGAGTGGGTGTTGGACAGTGTAG
CACTCTACCACTGCCAGGAGCTGACACCTACCTGATACCCAGATCCCACACGCCACT
ACTGACTGCAGCCAGCCACAGGTA CAGAGCCACAGGACCCCAAGAAATGAGCTTACAAAGT
GGCCTTTCAGGCCCTGGGAGCTCCTCTCACTCTTCAGTCTTC TACTGTCCTGGCTACT
AAATATTTTATGTACATCAGCCTGAAAAGGACTTCTGGCTATGCAAGGTCCCTTAAAGA
TTTTCTGCTTGAAGTCTCCCTTGAAATCTGCCATGAGCACAAATTAATGTAATTTTTC
ACCTGAGAAGATTTTAAACCAATTTAAACGCCACCAATTGAGCAAGATGCTGATT CATT
TTTATCAGCCCTATTTCTTCTATTCAGGCTGTTGTTGGCTTAGGGCTGGAAGCACAGAT
GGCTTGCCCTCAAGAGAATAGCTGGTTTCCCTAAGTTTACTTCTCTAAACCCCTGTGTTT
ACAAAGGCAGAGAGTCAGAACCTTCAATGGAAGGAGAGTGCTTGGGATCGATTATGTGAC
TTAAAGTCAGAAATAGTCCTTGGGCAGTTCTCAAAATGTTGGAGTGCAACATTGGGGAGAA
ATTCTGAGGCAGGTATTAGAAATGAAAAGGAAACTTGAAAACCTGGGCATGGTGGCTCACG
CCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCAGATCACTGCAGGTCAGGAGTTTCG
AAACCAAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAGAAATTAGCCGG
TCATGTGGTGGACACCTGTAATCCAGCTACTCAGGTGGCTAAGGCAGGAGAATCACTT
CAGCCCGGGAGGTGGAGGTTGCAGTGAGCCAAGATCATACCCACGCACTCCAGCCTGGGT
GACAGTGAGACTGTGGCTCAAAAAAAAAAAAAAAAAAAGGAAATGAAACTAGAAGAGAT
TTCTAAAAGTCTGAGATATATTTGCTAGATTCTAAAGAAATGTGTTCTAAACAGCAGAA
GATTTTCAAGAACCGGTTTCCAAAGACAGTCTTCTAATTCCTATTAGTAATAAGTAAAA
TGTTTATTGTTGTAGCTCTGGTATATAATCCATTCCCTTAAATATAAGACCTCTGGCA
TGAAATATTTTCAATATCTATAAAATGACAGATCCCACCAAGGAAGGAGCTGTTGCTTTCTT
GAGGTGATTTTTTTTCTTTGCTCCCTGTGCTGAAACCATAACGCTTCAATAAATAATTTT
GCTTGCTGAAGGAAGAAAAAGTGTTTTTCATAAACCCATTATCCAGGACTGTTTATAGCT
GTTGGAAGGACTAGGTCTTCCCTAGCCCCCCCAGTGTGCAAGGGCAGTGAAGACTTGATT
GTACAAAATACGTTTTGTAAATGTTGTGCTGTTAACACTGCAATAAACTTGGTAGCAAA
```

El cambio se encuentra en la segunda línea del exón 7 (la primera **c** en color morado).

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

CACTTCCA

ccatgaatgactgttcttgagacttaggccagccgactttctcagagcccttttcactgtg
cttcagtctccactctgtaaaatgggggtaatgatagtatctacctcctaggatttatt
gaggcagcttaaataccttttgatattcctgttgctgccaaaacaattgttgcaagggtc
agaagtctgagggtggctcaactgtttctttgtttcaggttcatgaggccaaaataaagg
tgttcgcagggcgtgttccttctagaggctctgggtccttgcagttctaggactaagat

NetGene2

Desaparece uno de los sitios *acceptor* de la secuencia wt a la mutante. Como se encuentra en el interior del exón, es muy probable que no afecte al *splicing*.

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'		
442	1	+	0.92	CCAGCCACAG	^GTACAGAGCC	H		442	1	+	0.92	CCAGCCACAG	^GTACAGAGCC	H			
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'
392	269	0	-	0.65	GGGGTATCAG	^GTAGGTGTCC		392	269	0	-	0.63	GGGGTATCAG	^GTAGGTGTCC			
191	470	2	-	0.31	GATTAATGAG	^GTAGAAGCTA		191	470	2	-	0.31	GATTAATGAG	^GTAGAAGCTA			
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'		
300	1	+	0.40	CTGTCTCCAG	^CAATTGGGCA			300	1	+	0.43	CTGTCTCCAG	^CAATTGGGCA				
311	0	+	0.32	AATTGGGCAG	^ATGTGTGAGG			311	0	+	0.34	AATTGGGCAG	^ATGTGTGAGG				
320	0	+	0.31	GATGTGTGAG	^GCACCTGTGG			320	0	+	0.34	GATGTGTGAG	^GCACCTGTGG				
339	1	+	0.19	GTGACCCGAG	^AGTGGGTGTT			339	1	+	0.19	GTGACCCGAG	^AGTGGGTGTT				
341	0	+	0.19	GACCCGAGAG	^TGGGTGTTGG			341	1	+	0.19	GTGACCCGAG	^AGTGGGTGTT				
355	2	+	0.07	TGTTGGACAG	^TGAGCACTC			341	0	+	0.19	GACCCGAGAG	^TGGGTGTTGG				

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 85.53.81.31.10311.0 :

Start	End	Score	Exon	Intron
435	449	0.89	gccacag	gtacagag
644	658	0.77	aattatg	gt aatttt

Donor site predictions for 85.53.81.31.10323.0 :

Start	End	Score	Exon	Intron
435	449	0.89	gccacag	gtacagag
644	658	0.77	aattatg	gt aatttt

Acceptor site predictions for 85.53.81.31.10311.0 :

Start	End	Score	Intron	Exon
280	320	0.76	gcttggttctctgtctcc	agcaattgggcagatgtgtgag
498	538	0.75	ggagctcctctcactcttc	agtccttctactgtcctggcta

Acceptor site predictions for 85.53.81.31.10323.0 :

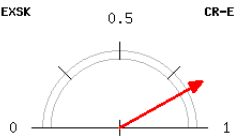
Start	End	Score	Intron	Exon
280	320	0.76	gcttggttctctgtctcc	agcaattgggcagatgtgtgag
498	538	0.75	ggagctcctctcactcttc	agtccttctactgtcctggcta

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
gcact(c/g)tacca	cactct	cactgt	31178	48%

CRYP-SKIP

Exon length (bp)	360
PESS (≤ -2.62) density	1.11
NN 5'ss score density	0.49
SF2/ASF score density	11.07
FAS-ESS (hex2) density	2.78
EIE score density	492.42
Probability of cryptic splice site activation (P_{CR-E})	0.84



Human Splicing Finder

Alteration of auxiliary sequences		Significant alteration of ESE / ESS motifs ratio (-4)
Algorithm/Matix	position	sequence
EIE (ESE Site Broken)	chr17:43045738	CTACCA
ESE_SRp40 (ESE Site Broken)	chr17:43045738	CTACCAG
EIE (ESE Site Broken)	chr17:43045740	CTCTAC
PESE (New ESE Site)	chr17:43045740	CTGTACCA
EIE (ESE Site Broken)	chr17:43045741	ACTCTA
ESE_ASF (ESE Site Broken)	chr17:43045742	CACTCTA

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	31	138	cactaatct		2.28872196415	0.518796992481	1	24	44	1.8120758	mut	30	137	cactaatct	2.28872196415	0.507575757576	1	24	44	1.8084517
wt	31	93	gtgtgaggc		-1.08118976567	0.477272727273	42	9	12	-2.4141533	mut	30	92	gtgtgaggc	-1.08118976567	0.459770114943	67	9	12	-4.0022641
wt	31	79	tggtgacc		2.74252220032	0.486486486486	28	9	12	-0.027837023	mut	30	78	tggtgacc	2.74252220032	0.465753424658	53	9	12	-1.6169912
wt	31	16	acctgatac		1.15817448354	0.636363636364	11	0	0	0.3645032	mut	30	15	acctgatac	1.15817448354	0.6	10	0	0	0.41605714

Variant Effect Predictor tool

ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000352993.7	protein_coding	22/22	2225	2106	702	L	CTC/CTG	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000357654.9	protein_coding	23/23	5645	5532	1844	L	CTC/CTG	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	3_prime_UTR_variant, NMD_transcript_variant	BRCA1	ENSG00000012048	Transcript	ENST00000461221.5	nonsense_mediated_decay	23/23	5633	-	-	-	-	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	3_prime_UTR_variant	BRCA1	ENSG00000012048	Transcript	ENST00000468300.5	protein_coding	22/22	2340	-	-	-	-	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000471181.7	protein_coding	24/24	5827	5595	1865	L	CTC/CTG	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000491747.6	protein_coding	23/23	2319	2220	740	L	CTC/CTG	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000493795.5	protein_coding	22/22	5623	5391	1797	L	CTC/CTG	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000586385.5	protein_coding	8/8	606	462	154	L	CTC/CTG	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000591534.5	protein_coding	11/11	1107	1005	335	L	CTC/CTG	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000591849.5	protein_coding	5/5	388	231	77	L	CTC/CTG	COSV58786146
ENST00000352993.7:c.2106C>G	17:43045738-43045738	C	synonymous_variant	BRCA1	ENSG00000012048	Transcript	ENST00000644379.1	protein_coding	15/15	1919	1920	640	L	CTC/CTG	COSV58786146

ESEfinder

Los resultados que tienen puntuación positiva en la secuencia wt para las matrices 3’Ss:

<div>346 (-315)</div> <div>TGTTGGACAGTGTAGCACTCTACCACTGCC</div> <div>-26.94150</div>	<div>346 (-315)</div> <div>TGTTGGACAGTGTAGCACTCTACCACTGCC</div> <div>0.48370</div>	<div>346 (-315)</div> <div>TGTTGGACAGTGTAGCACTCTACCACTGCC</div> <div>-22.92440</div>	<div>346 (-315)</div> <div>TGTTGGACAGTGTAGCACTCTACCACTGCC</div> <div>0.06060</div>
<div>357 (-304)</div> <div>GTAGCACTCTACCACTGCCAGGAGCTGGAC</div> <div>-30.76660</div>	<div>357 (-304)</div> <div>GTAGCACTCTACCACTGCCAGGAGCTGGAC</div> <div>3.76390</div>	<div>357 (-304)</div> <div>GTAGCACTCTACCACTGCCAGGAGCTGGAC</div> <div>-29.32470</div>	<div>357 (-304)</div> <div>GTAGCACTCTACCACTGCCAGGAGCTGGAC</div> <div>3.44200</div>
<div>363 (-298)</div> <div>CTCTACCACTGCCAGGAGCTGGACACCTAC</div> <div>-10.63890</div>	<div>363 (-298)</div> <div>CTCTACCACTGCCAGGAGCTGGACACCTAC</div> <div>6.26500</div>	<div>363 (-298)</div> <div>CTCTACCACTGCCAGGAGCTGGACACCTAC</div> <div>-7.03630</div>	<div>363 (-298)</div> <div>CTCTACCACTGCCAGGAGCTGGACACCTAC</div> <div>5.90170</div>

Para la secuencia mutante, todas las puntuaciones descienden:

<div>346 (-315)</div> <div>TGTTGGACAGTGTAGCACTGTACCACTGCC</div> <div>-23.13860</div>	<div>346 (-315)</div> <div>TGTTGGACAGTGTAGCACTGTACCACTGCC</div> <div>0.11700</div>	<div>346 (-315)</div> <div>TGTTGGACAGTGTAGCACTGTACCACTGCC</div> <div>-19.14460</div>	<div>346 (-315)</div> <div>TGTTGGACAGTGTAGCACTGTACCACTGCC</div> <div>-0.25530</div>
<div>357 (-304)</div> <div>GTAGCACTGTACCACTGCCAGGAGCTGGAC</div> <div>-30.79090</div>	<div>357 (-304)</div> <div>GTAGCACTGTACCACTGCCAGGAGCTGGAC</div> <div>1.89300</div>	<div>357 (-304)</div> <div>GTAGCACTGTACCACTGCCAGGAGCTGGAC</div> <div>-29.34270</div>	<div>357 (-304)</div> <div>GTAGCACTGTACCACTGCCAGGAGCTGGAC</div> <div>1.49580</div>
<div>363 (-298)</div> <div>CTGTACCACTGCCAGGAGCTGGACACCTAC</div> <div>-10.76200</div>	<div>363 (-298)</div> <div>CTGTACCACTGCCAGGAGCTGGACACCTAC</div> <div>4.97090</div>	<div>363 (-298)</div> <div>CTGTACCACTGCCAGGAGCTGGACACCTAC</div> <div>-7.17990</div>	<div>363 (-298)</div> <div>CTGTACCACTGCCAGGAGCTGGACACCTAC</div> <div>4.63240</div>

Por lo tanto, puede que se esté debilitando un sitio *acceptor*.

En cuando a los ESE, las puntuaciones descienden, por lo que puede que se estén debilitando un sitio ESE:

359 (-302)	AGCACTC	-0.68999	359 (-302)	AGCACTC	-0.40431	359 (-302)	AGCACTCT	-2.63218	359 (-302)	AGCACTC	-4.11350
360 (-301)	GCACTCT	-5.47194	360 (-301)	GCACTCT	-4.03494	360 (-301)	GCACTCTA	2.19576	360 (-301)	GCACTCT	0.88698
361 (-300)	CACTCTA	2.01819	361 (-300)	CACTCTA	2.19705	361 (-300)	CACTCTAC	-1.18084	361 (-300)	CACTCTA	-7.22045
362 (-299)	ACTCTAC	-8.25322	362 (-299)	ACTCTAC	-5.90538	362 (-299)	ACTCTACC	-2.84607	362 (-299)	ACTCTAC	2.57132
363 (-298)	CTCTACC	-4.29241	363 (-298)	CTCTACC	-1.67692	363 (-298)	CTCTACCA	1.97787	363 (-298)	CTCTACC	-0.27279
364 (-297)	TCTACCA	-2.97461	364 (-297)	TCTACCA	-2.10034	364 (-297)	TCTACCAG	-1.92425	364 (-297)	TCTACCA	-0.71623
365 (-296)	CTACCAG	-1.14152	365 (-296)	CTACCAG	0.58381	365 (-296)	CTACCAGT	-3.09218	365 (-296)	CTACCAG	3.35085

359 (-302)	AGCACTG	0.78987	359 (-302)	AGCACTG	0.58471	359 (-302)	AGCACGT	-4.30774	359 (-302)	AGCACTG	-3.27086
360 (-301)	GCACTGT	-2.89420	360 (-301)	GCACTGT	-2.38764	360 (-301)	GCACTGTA	1.28045	360 (-301)	GCACTGT	1.19260
361 (-300)	CACTGTA	1.41407	361 (-300)	CACTGTA	1.98912	361 (-300)	CACTGTAC	-2.07852	361 (-300)	CACTGTA	-8.03803
362 (-299)	ACTGTAC	-10.16825	362 (-299)	ACTGTAC	-7.21932	362 (-299)	ACTGTACC	-5.54421	362 (-299)	ACTGTAC	0.84699
363 (-298)	CTGTACC	-4.53864	363 (-298)	CTGTACC	-2.05819	363 (-298)	CTGTACCA	-0.32970	363 (-298)	CTGTACC	-0.48603
364 (-297)	TGTACCA	-1.69637	364 (-297)	TGTACCA	-1.12208	364 (-297)	TGTACCAG	0.11183	364 (-297)	TGTACCA	-1.53778
365 (-296)	GTACCAG	-2.71883	365 (-296)	GTACCAG	-2.31439	365 (-296)	GTACCAGT	-1.06367	365 (-296)	GTACCAG	1.20320

EX-SKIP

No se puede calcular porque la secuencia es demasiado larga.

HOT-SKIP

>wt

```
taaaaatacaaaaattagctgggtgtgatggcatgtgcctgtaattccagctactcaggaggcagagacaggagaattgc
ttgaaccaggaggcggaggttgatgagccgagattgcgcatcacactctagcctcggcgacagagcaagactccgtc
tcaaaaaaaaaaaaaaaaaaattagcttctacctcattaatcctaagaactcatacaaccaggaccctggagtcgattgat
tagagcctagtccaggagaatgaattgacactaatctctgcttggttctctgtctccagCAATTGGGCAGATGTGTGAG
GCACCTGTGGTGACCCGAGAGTGGGTGTTGGACAGTGTAGCACTCTACCAGTGCCAGGAGCTGGACACCTACCTGATACC
CCAGATCCCCCAGGCCACTACTGACTGCAGCCAGCCACAGGTACAGAGCCACAGGACCCCAAGAATGAGCTTACAAAGT
GGCCTTTCCAGGCCCTGGGAGCTCCTCTCACTCTTCAGTCTTCTACTGTCTGGCTACTAAATATTTTATGTACATCAG
CCTGAAAAGGACTTCTGGCTATGCAAGGGTCCCTTAAAGATTTTCTGCTTGAAGTCTCCCTGGAAATCTGCCATGAGCA
CAAAATTATGGTAATTTTTCACCTGAGAAGATTTTAAAACCATTTAAACGCCACCAATTGAGCAAGATGCTGATTCATTA
TTTATCAGCCCTATTCTTTCTATTAGGCTGTTGTTGGCTTAGGGCTGGAAGCACAGAGTGGCTTGGCCTCAAGAGAATA
GCTGGTTTCCCTAAGTTTACTTCTCTAAAACCTGTGTTTCAAAAGGCAGAGAGTCAAGCCCTTCAATGGAAGGAGAGTG
CTTGGGATCGATTATGTGACTTAAAGTCAGAAATAGTCTTGGGCAGTTCTCAAATGTTGGAGTGGAAACATTGGGGAGGAA
ATTCTGAGGCAGGTATTAGAAATGAAAAGGAACTTGAAACCTGGGCATGGTGGCTCACGCCTGTAAATCCAGCACTTTG
GGAGGCCAAGGTGGGCAGATCACTGGAGGTGAGGAGTTCGAAACAGCCTGGCCAACATGGTGAAACCCCATCTCTACTA
AAAATACAGAAATTAGCCGGTCATGGTGGTGGACACCTGTAATCCAGCTACTCAGGTGGCTAAGGCAGGAGAATCACTT
CAGCCCAGGAGGTGGAGGTTGCAGTGAGCCAAGATCATACCACGGCACTCCAGCCTGGGTGACAGTGAGACTGTGGCTCA
AAAAAAAAAAAAAAAAAAAAAGGAAAAATGAACTAGAAAGAGATTTCTAAAAGTCTGAGATATATTTGCTAGATTTCTAAAGAA
TGTGTTCTAAAACAGCAGAAGATTTTCAAGAACCAGTTTCCAAAGACAGTCTTCTAATTCCTCATTAGTAATAAGTAAAA
TGTTTATTGTTGTAGCTCTGGTATATAATCCATTCTCTTAAAATATAAGACCTCTGGCATGAATATTTTATATCTATAA
AATGACAGATCCCACAGGAAGGAGCTGTTGCTTTCTTTGAGGTGATTTTCTCTTTGCTCCCTGTTGCTGAAACCAT
ACAGCTTCATAAATAATTTGCTTGCTGAAGGAAGAAAAAGTGTTTTTCATAAACCATTATCCAGGACTGTTTATAGCT
GTTGGAAGGACTAGGTCTTCCCTAGCCCCCAGTGTGCAAGGGCAGTGAAGACTTGATTGTACAAAATACGTTTTGTAA
ATGTTGTCTGTAACTGCAAAATAAAGTGTAGCAAACTTCCAccatgaatgactgttcttgagacttaggccag
ccgactttctcagagccttttctactgtgcttcagtcctccactctgtaaaatgggggtaatgatagtatctacctcctag
gatttattgaggcagcttaaataccttttgtatttcctgttgctgcaaaaacaaattgttgcaaggtcagaagtctgagg
tggtcgaactgtttcttgggttcaggtttcatgaggcaaaataaaggtgttcgagggcggtgttccttctagaggctc
tggttccttgagttcttaggactaagat
```