

Ejemplo comparación de resultados predictores in sillico

Cambio de estudio ATP5MF-PTCD1 c.121+1825T>A (chr7:99458261 T/A, COSV52854368)

Exón 2 e intrones adyacentes:

```
tgccatatgtgttaactgggagattattcagtggggtatTTTTCTGactcttccgtttcag
TACCACTGAAGGACAAGAACTTCTGGAAGTCAAAGTGGGAGCTGCCAAGCTGGATCT
TGATGCGGGACTTCAGTCCATGCGCATTTTCGGAGCGTTCAAGAG
gtcagagccctgttgatgtgtgtaaatgaaagcaaatctctgggtcttcagatctctttt
ttgcatgaattaatTTTgggcaatgaaggccttggtgtctgaaagcagtaagttatgtag
aggatgacagggaggatgctagccctggacttctgtggttgatggccctttcaagg
ctggctctgagatattacagccaaagaaatggtttcctgttgcttggaaatcagtatgtctg
cctctgtcttagagaaatcctaagacttctttaagaggaaatgaattggaaacgtagtg
ggcattagtctataatatgatgctctctccctgcagaaacttcaggacaatatTTTgaa
atggcctatcttggcctgggtgtggtggtcttatggcctgtaacccggcacttcgggaggcc
aaggctggaggattgcttgagccagctttagaccagcctaggcaatatggagagaCCCCA
tctctataaaaaaaaaccacacaagaaaaattagctggctgtgtgtgtgcatgctgtagt
cccagctactcaggaggctaagctcacagccaaaggggagaggtggaatccaggtagg
cctagttcaggagctggtatgtgcttatagttcaggctaagggatgagagaccttactg
gcatTTTgtgcttgtcatctttatcagccagtgaatgcaggccgaggggcttaacagcca
gagcaggatagggtccctgaatgtcagccagactcaactgtgtgtctcaacttcactcaaa
tgtgaagcccagcagggcagtgagcgcctcttgcggttgagggttactaccggtactaca
acaagtacatcaatgtgaagaaggggagcatctcggggattaccatgggtgctggcatgct
acgtgctcttttagctactctcttttctacaagcatctcagtgagtgctctgcggcgtct
tgcttttagttcccatgagaggggtggtgactgatttcattagatacagcagcccac
ttcttctgagggtgaggggacctttagttaaagttccttatgtttccacctaaaagaattg
gagggaacctatcagagtacagtatgtgggatattgtttgaaatgagaaaattgtgaCAAA
gagaacacaggaaaaatcaagatgaagccagggttaagagagttacttagaagcattcttta
aaatacagcacactgtgtaaaaatttggtcagattttcctagtagccaatgcagaaagagaa
atgagttgagtgggataatctgtgtaacccagaaaagcatggctgttcgggtctgaggct
tgagaggagcttgtctgtgtgggagattggcgggagggtgtcggtggcagcctctgactagg
ctgtttctgacagtgtagtgtaacccccctacccacactcactccattctgtggcacc
agggcggtcctgcctctgaaggacgggttctctggggctgtttgtctgagcgtgtgttc
ctctgttcctcttgaaaagaaaagtggttgccagggtggcaggttggtctctgaggtgtct
ttgtgcccccggtctgtctgattctgcagagacacgggcaggcgggtcatgggtcatct
ctgaagggaattctcaggaaggctttgtgtgatctcagcctgtctcctgccatgctgtgc
cttcactgtaaccttttaagatacttaacatcttgccctcctgacttcagagcacgagcg
gctccgcacaaatcacactgaagagacacactctgcacccccccacccacagaccttggtc
cgagcccttcgtgaggaaacaCAATctcaatcgttgctgaatcctttcatatcctaatag
gaattaaacctcaaataaaacatgactggtacgtgtgctgattgcctgcctgtctgcgt
agccattatcccattgtccagggtgtcctgtggcacagggtggggggagggtcgtcttcc
```

El cambio se encuentra en la antepenúltima línea del intrón 2 (la **t** en color granate).

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'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
169	1	+	0.61	TTTCAAAGAG	^	GTCAGAGCCT		169	1	+	0.61	TTTCAAAGAG	^	GTCAGAGCCT	
276	1	+	0.58	TCTGAAAGCA	^	GTAAGTTATG		276	1	+	0.58	TCTGAAAGCA	^	GTAAGTTATG	
764	1	+	0.47	TGGAATCCAG	^	GTAGGCCTAG		764	1	+	0.47	TGGAATCCAG	^	GTAGGCCTAG	
785	0	+	0.44	TCAGGAGCTG	^	GTATGTGCTT		785	0	+	0.44	TCAGGAGCTG	^	GTATGTGCTT	
1108	1	+	0.97	AAGCATCTCA	^	GTGAGTGCCT	+	1108	1	+	0.97	AAGCATCTCA	^	GTGAGTGCCT	+
1340	0	+	0.41	TGAAGCCAGG	^	GTAAGAGAG		1340	0	+	0.41	TGAAGCCAGG	^	GTAAGAGAG	
2057	2	+	0.00	AACATGACTG	^	GTACGTGTGC		2057	2	+	0.00	AACATGACTG	^	GTACGTGTGC	
2128	-	+	0.00	GTGGCACAAG	^	GTGGGGGAG		2128	-	+	0.00	GTGGCACAAG	^	GTGGGGGAG	

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'
1876	273	1	-	0.83	AGGCAAGATG	^	GTAAGTATCT		1876	273	1	-	0.83	AGGCAAGATG	^	GTAAGTATCT	
1586	563	0	-	0.45	TGGGAGTGAG	^	GTGGGGGTAG		1586	563	0	-	0.45	TGGGAGTGAG	^	GTGGGGGTAG	
1187	962	0	-	0.36	TCAGAAGAAG	^	GTGGGCTGCT		1187	962	0	-	0.36	TCAGAAGAAG	^	GTGGGCTGCT	
826	1323	2	-	0.41	ACAAATGCCA	^	GTAAGGTCTC		826	1323	2	-	0.41	ACAAATGCCA	^	GTAAGGTCTC	

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'
60	1	+	0.00	TCCGTTTCAG	^	TACCAAGTGAA		60	1	+	0.00	TCCGTTTCAG	^	TACCAAGTGAA	
506	1	+	0.27	TCCCTGCCAG	^	AACCTCAGGA		506	1	+	0.27	TCCCTGCCAG	^	AACCTCAGGA	
955	1	+	0.19	AAATGTGAAG	^	CCCAGCAGGG		955	1	+	0.19	AAATGTGAAG	^	CCCAGCAGGG	
960	1	+	0.20	TGAAGCCAG	^	CAGGGCAGTG		960	1	+	0.20	TGAAGCCAG	^	CAGGGCAGTG	
963	1	+	0.32	AGCCCAGCAG	^	GGCAGTGAGC		963	1	+	0.32	AGCCCAGCAG	^	GGCAGTGAGC	
968	0	+	0.34	AGCAGGGCAG	^	TGAGCGCCTC		968	0	+	0.34	AGCAGGGCAG	^	TGAGCGCCTC	
990	1	+	0.95	GCGTTTGCAG	^	GTTACTACCG		990	1	+	0.95	GCGTTTGCAG	^	GTTACTACCG	
1610	2	+	0.33	GTGGCACCAG	^	GGCGGTCCTG		1610	2	+	0.33	GTGGCACCAG	^	GGCGGTCCTG	
1898	0	+	0.23	CTGACTTCAG	^	AGCACGAGCG		1898	0	+	0.17	CTGACTTCAG	^	AGCACGAGCG	
2028	1	+	0.33	ATCCTAATAG	^	GAATTAACCT		2028	1	+	0.33	ATCCTAATAG	^	GAATTAACCT	
2090	2	+	0.00	TCTGCCGTAG	^	CCATTATCCC		2090	2	+	0.00	TCTGCCGTAG	^	CCATTATCCC	
2109	0	+	0.00	CATGTTCCAG	^	GGGCTCCTGT		2109	0	+	0.00	CATGTTCCAG	^	GGGCTCCTGT	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'	pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
2116	33	1	-	0.00	TGTGCCACAG	^	GAGCCCCCTGG		2116	33	1	-	0.00	TGTGCCACAG	^	GAGCCCCCTGG	
1789	360	0	-	0.26	TTCCCTTCAG	^	AGATGACCCA		1789	360	0	-	0.26	TTCCCTTCAG	^	AGATGACCCA	
1753	396	2	-	0.23	GTCTCTGCAG	^	AATCAGCAGA		1753	396	2	-	0.23	GTCTCTGCAG	^	AATCAGCAGA	
1680	469	0	-	0.29	TTCTTTCCAG	^	AAGGAACAGA		1680	469	0	-	0.29	TTCTTTCCAG	^	AAGGAACAGA	
1677	472	0	-	0.39	TTTCCAGAAG	^	GAACAGAGGG		1677	472	0	-	0.39	TTTCCAGAAG	^	GAACAGAGGG	
1671	478	0	-	0.18	GAAGGAACAG	^	AGGGAACACA		1671	478	0	-	0.18	GAAGGAACAG	^	AGGGAACACA	
1669	480	2	-	0.18	AGGAACAGAG	^	GGAACACACG		1669	480	2	-	0.18	AGGAACAGAG	^	GGAACACACG	
1654	495	2	-	0.07	ACACGCTCAG	^	ACAACACAGCC		1654	495	2	-	0.07	ACACGCTCAG	^	ACAACACAGCC	
1503	646	2	-	0.25	CTCCCACCAG	^	ACAAGCTCCT		1503	646	2	-	0.25	CTCCCACCAG	^	ACAAGCTCCT	
188	1961	1	-	0.77	TCATTTACAG	^	ACATCCACAA		188	1961	1	-	0.77	TCATTTACAG	^	ACATCCACAA	

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.0.139.574182.0 :

Start	End	Score	Exon	Intron
162	176	0.75	caaagag	gt cagagc
269	283	0.97	gaaagca	gt aagtta
757	771	0.98	aatccag	gt aggcct
778	792	0.93	ggagctg	gt atgtgc
1006	1020	0.53	acaacaa	gt acatca
1101	1115	0.93	catctca	gt gagtg
1333	1347	0.81	agccagg	gt aaagag
2050	2064	0.97	atgactg	gt acgtgt
2121	2135	0.60	gcacaag	gt gggggg

Donor site predictions for 10.42.1.119.574194.0 :

Start	End	Score	Exon	Intron
162	176	0.75	caaagag	gt cagagc
269	283	0.97	gaaagca	gt aagtta
757	771	0.98	aatccag	gt aggcct
778	792	0.93	ggagctg	gt atgtgc
1006	1020	0.53	acaacaa	gt acatca
1101	1115	0.93	catctca	gt gagtg
1333	1347	0.81	agccagg	gt aaagag
2050	2064	0.97	atgactg	gt acgtgt
2121	2135	0.60	gcacaag	gt gggggg

Acceptor site predictions for 10.42.0.139.574182.0 :

Start	End	Score	Intron	Exon
40	80	0.94	tttctcactcttccgtttc	ag taccagtgaaggacaagaaa
400	440	0.74	gtatgtctgcctctgtctt	ag agaaatcctaagacttcttt
486	526	0.78	tgatgtctctcctccctgcc	ag aacttcaggacaaatatttg
778	818	0.47	ggagctggatgtgtcttat	ag ttcaggctatgggatgaga
834	874	0.44	tgtgcttgtcatctttatc	ag ccagtgaatgcaggcgagg
970	1010	0.94	gagcgctcttgcgtttgc	ag gttactaccgtactacaac
1061	1101	0.47	ggcatgctacgtgctcttt	ag ctactcctttcctacaagc
1117	1157	0.80	tctgcggcgtcttgccttt	ag ttcccatgagagggtggtgg
1388	1428	0.58	aatttggctcgattttcct	ag tagccaatgcagaaagagaa
1738	1778	0.51	ccgggtctgctgattctgc	ag agacaccggcaggcggtca
1878	1918	0.82	atcttgcttctgacttc	ag agcacgagcggctccgaaa
2008	2048	0.69	aatcctttcatatccta	ag gaattaacctccaataaaa
2089	2129	0.82	agccattatcccatgttcc	ag gggctcctgtggcacaaggt

Acceptor site predictions for 10.42.1.119.574194.0 :

Start	End	Score	Intron	Exon
40	80	0.94	tttctcactcttccgtttc	ag taccagtgaaggacaagaaa
400	440	0.74	gtatgtctgcctctgtctt	ag agaaatcctaagacttcttt
486	526	0.78	tgatgtctctcctccctgcc	ag aacttcaggacaaatatttg
778	818	0.47	ggagctggatgtgtcttat	ag ttcaggctatgggatgaga
834	874	0.44	tgtgcttgtcatctttatc	ag ccagtgaatgcaggcgagg
970	1010	0.94	gagcgctcttgcgtttgc	ag gttactaccgtactacaac
1061	1101	0.47	ggcatgctacgtgctcttt	ag ctactcctttcctacaagc
1117	1157	0.80	tctgcggcgtcttgccttt	ag ttcccatgagagggtggtgg
1388	1428	0.58	aatttggctcgattttcct	ag tagccaatgcagaaagagaa
1738	1778	0.51	ccgggtctgctgattctgc	ag agacaccggcaggcggtca
1878	1918	0.82	atcttgcttctgacttc	ag agcacgagcggctccgaaa
2008	2048	0.69	aatcctttcatatccta	ag gaattaacctccaataaaa
2089	2129	0.82	agccattatcccatgttcc	ag gggctcctgtggcacaaggt

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
cacaa(t/a)ctcaa	tctcaa	actcaa	28657	65%

Human Splicing Finder

Alteration of auxiliary sequences		Significant alteration of ESE / ESS motifs ratio (4)	
Algorithm/Matix	position	sequence	
EIE (New ESE Site)	chr7:99458261	ACTCAA	
EIE (New ESE Site)	chr7:99458262	AACTCA	
RESCUE ESE (New ESE Site)	chr7:99458263	AAACTC	
Sironi_motif3 (ESS Site Broken)	chr7:99458263	AATCTCAA	

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr				
wt	16	62	ccgtgagga		-1.52768403304	0.561403508772	27	14	27	-1.4725926			
wt	16	48	atctcaatc		-0.0747201413931	0.604651162791	13	14	27	-0.0035443643			
wt	16	37	tgctgaatc		1.55511442561	0.625	2	14	27	1.3374665			
wt	16	28	ctttcatat		-1.30703608036	0.565217391304	23	0	0	-1.3833013			
wt	16	20	tcctaatag		1.43235144722	0.533333333333	15	0	0	0.1853874			
mut	15	61	ccgtgagga		-1.52768403304	0.535714285714	27	14	27	-1.4808895			
mut	15	47	aactcaatc		-0.0294814399462	0.595238095238	13	14	27	0.011128583			
mut	15	36	tgctgaatc		1.55511442561	0.612903225806	2	14	27	1.3335596			
mut	15	27	ctttcatat		-1.30703608036	0.545454545455	22	0	0	-1.3263857			
mut	15	19	tcctaatag		1.43235144722	0.5	14	0	0	0.23792002			

La diferencia entre las secuencias es que el BP que en la WT tiene puntuación negativa, en la secuencia mutante tiene puntuación positiva, lo que podría significar que se está activando un BP.

Variant Effect Predictor tool

ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	3_prime_UTR_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000292475.8	protein_coding	4/4	377	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000292476.10	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	3_prime_UTR_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000359832.8	protein_coding	3/3	257	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000394186.3	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	intron_variant	ATP5MF-PTCD1	ENSG00000248919	Transcript	ENST00000413834.5	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	3_prime_UTR_variant, NMD_transcript_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000414062.5	nonsense_mediated_decay	3/4	240	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	intron_variant	PTCD1	ENSG00000106246	Transcript	ENST00000430982.1	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000436336.6	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	intron_variant, non_coding_transcript_variant	ATP5MF-PTCD1	ENSG00000248919	Transcript	ENST00000437572.1	processed_transcript	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000440514.1	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000441580.5	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	3_prime_UTR_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000449683.5	protein_coding	4/4	389	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	intron_variant, NMD_transcript_variant	ATP5MF-PTCD1	ENSG00000248919	Transcript	ENST00000451138.1	nonsense_mediated_decay	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000451876.5	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000452047.1	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000465132.1	retained_intron	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	intron_variant, non_coding_transcript_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000466753.1	processed_transcript	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000469897.1	retained_intron	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	CPSF4	ENSG00000160917	Transcript	ENST00000471455.1	processed_transcript	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000481899.1	retained_intron	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000485011.1	retained_intron	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	3_prime_UTR_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000488775.5	protein_coding	3/3	242	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	non_coding_transcript_exon_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000491560.1	retained_intron	3/3	592	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000523680.1	protein_coding	-	-	COSV52854368
ENST00000413834.5:c.121+1825T>A	7-99458261-99458261	T	downstream_gene_variant	ATP5MF	ENSG00000241468	Transcript	ENST00000524321.1	retained_intron	-	-	COSV52854368

ESEfinder

La única predicción para la secuencia WT que tiene puntuación positiva es en las matrices 5', pero, cuando se comparan las puntuaciones con las de la secuencia mutante, no se ven prácticamente alteradas, por lo que no se tendrán en cuenta.

1965 (-184)	ggcccgagcccctccgtgaggaacacaatc	4.39210	1965 (-184)	ggcccgagcccctccgtgaggaacacaatc	-22.62460	1965 (-184)	ggcccgagcccctccgtgaggaacacaatc	4.15680	1965 (-184)	ggcccgagcccctccgtgaggaacacaatc	-20.35410
1965 (-184)	ggcccgagcccctccgtgaggaacacaaac	4.10360	1965 (-184)	ggcccgagcccctccgtgaggaacacaaac	-22.31340	1965 (-184)	ggcccgagcccctccgtgaggaacacaaac	3.89710	1965 (-184)	ggcccgagcccctccgtgaggaacacaaac	-20.05140