

Ejemplo comparación de resultados predictores in silico

Cambio de estudio ELP4 c.259+1970A>G (chr 11:31522061 A/G, rs970055506 o NM_001288725.2: c.259+1970A>G)

Exón 2 e intrones adyacentes:

```
gttgttca caact actgtttt aaagtatttgaagt gcc tttcctt cat aaagcctaa ctt
ttatgcta aaggtaat tggaa aggtaaattaat tttct tctggttttg ttttcatttccag
GTGGAGGT TAGCCGT TGGAA CAGTCTCTTAATTG
gttagt acaaaaatac atgctttt gctctcctatc attgtttt cctgtt gtgcataatcat
tttatccaat tccatttccaa aattattt taaatcacta acacgacagaagt taagataaa
ttagagagga attgactt taagttct attttaattcta atgttg agctgaactt gaaata
aattagg gcaatatatct gaacatgtag ctcttataa cttaaaaatgtt catagttg att
gtgctccatt ttttttaaaa agactaaa attaatgt gaattg atttttttgg gagaacctca
gtattttaa agtgggattatta cagggatca atttgg gaaattgaaatac caagtc catcg
tatatat catatcattttat cattatctttt ttattgaaac aaaagtaacaaagt tagttaa
tcataatttt agttcatctttta agtg gcaagagtcct taatttaagta tataaaccataatt
gcttctcttctttttt tagtatat aaggtcacagaa ctcttgaatt gaaataaacctt gt
cagttatata atcaaa gpcagccatatta atattttaatt catgtatctg taacataatctat
tgttctactttt gtctctctg atataatataa attt acgtagtt tttcacgtgacacct
taaaatattt gaaaatac ctattttttcatct caaatattct ccaactaaatggacatc
attccttaacctgg tttttatagtttag gggttttcacacacacct cactgttctgattatttct
ttttatgca ccttctagtc tggcattcattttctcct gaatcatgcccc gacctaaatac
agtatt cdaaatgtaat atgacctgaaatcaccatactgg gactataaaaat aaatatgtt
aatctagacagag ttatttagtaccaggd acdaaaaagg tttagtttctatttta ttgacccc
taaggtaac aaaaaaagt tttttggaaga attttt cccaag agtgagt aaaaata caaagt
ggtgatgctt aaaaattttggcat gta cgtagaaga attt gtctagtaaa agggaaaaa cgttatat
agaa cattttattt cctcatctat gctggataatt aaagagtaaat attattagtagtgaga
gctgccttttta ttttaactctta cgttgggtcatac attatgc aaagtgc attatgttttt
tttttttagc ttaatacaaca aacctcaga agata ctaaaatata cttcatgttatctccc
tttatad atgaggaaa tgggctcaca aggatggggt agttttt ccaactattc catagtt
aatctgataa accaggactt gaattctgtgtgattccaaa cccattttttttt gtaactat
gtctctaa agttttctc catgtttttgttt taaatag caataaaac gtaaaaatagg aaattaca
tcccttttttag atttgcagtttca agtagtaaa agtaaaa ctca cattttatggag attt
tttcttttaattgga ttttcaa attttct caa catgaaggaa tcccttgataaa attcatgcac
aaataaggat caaagttg agcactgtaaaaatc cttatttcacat gatc agcatttctcc
attcaggt caaaaaagttt tgtctttaa aatataaatcaat atacaggagtggttaaaaa
catcttcagttgtac agtattttgttagatc agttttatatt agttgtagcatattg ctatta
aaaatgtaa tatattagtttact taaattatgccattta atca ittttgggtcctgtgttct
ctcaactc agccattgat tcccgtgttggtcttggctgtgaagt adctgccaacttctt
accagtcatttactttacat caatacaaatataatcttcttgga aatttgcctagagttt agt
gcctgagta tccagatagtaga caatttttaataaaaagcta agtaaaaat tgtctttatga
```

El cambio se encuentra en la última fila del intrón 2 (la **a** en color rojo subrayada de azul).

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'
72		1	+	0.00	TATGCTAAAG	^	GTAATGGAAA	
84		1	+	0.00	AATGGAAAAG	^	GTAATTAAT	
157		1	+	0.56	CTTCTAATTG	^	GTTAGTACAA	
1094		0	+	0.46	CACCAAAAAG	^	GTTAGTTTCT	
1121		0	+	0.41	GACCCCTAAG	^	GTAACAAAAA	
1159		0	+	0.36	TTCCCCAAGA	^	GTGAGTAAAA	
1203		1	+	0.31	TGGCATGTAC	^	GTAAGAATTT	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
2033		104		2	-	0.37	TTGTATTGAT	^	GTAAGTAAAT	
2018		119		2	-	0.63	TAAATGACTG	^	GTAAGAAGTT	
1642		495		0	-	0.60	CTCCATAAAT	^	GTGAGTTTTA	
1319		818		2	-	0.41	ATGACCAACA	^	GTAAGAGTTA	
916		1221		0	-	0.46	GAACAGTGAG	^	GTGTGTGAAA	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
120		1	+	0.71	TCATTTCCAG	^	GTGGAGGTTT	
654		2	+	0.43	TCTTTTTTAG	^	TATATATAGG	
1365		0	+	0.53	TTTTTTTTTAG	^	CCTAATACAA	

Acceptor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
1798		339		0	-	0.17	TATTTTAAAG	^	ACAAAACTTT	
1116		1021		1	-	0.56	GTTACCTTAG	^	GGGTCAATAA	

Donor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
72		1	+	0.00	TATGCTAAAG	^	GTAATGGAAA	
84		1	+	0.00	AATGGAAAAG	^	GTAATTAAT	
157		1	+	0.56	CTTCTAATTG	^	GTTAGTACAA	
1094		0	+	0.46	CACCAAAAAG	^	GTTAGTTTCT	
1121		0	+	0.41	GACCCCTAAG	^	GTAACAAAAA	
1159		0	+	0.36	TTCCCCAAGA	^	GTGAGTAAAA	
1203		1	+	0.31	TGGCATGTAC	^	GTAAGAATTT	

Donor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	exon	intron	3'
2033		104		2	-	0.37	TTGTATTGAT	^	GTAAGTAAAT	
2018		119		2	-	0.63	TAAATGACTG	^	GTAAGAAGTT	
1642		495		0	-	0.60	CTCCATAAAT	^	GTGAGTTTTA	
1319		818		2	-	0.41	ATGACCAACA	^	GTAAGAGTTA	
916		1221		0	-	0.46	GAACAGTGAG	^	GTGTGTGAAA	

Acceptor splice sites, direct strand

pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
120		1	+	0.71	TCATTTCCAG	^	GTGGAGGTTT	
654		2	+	0.43	TCTTTTTTAG	^	TATATATAGG	
1365		0	+	0.53	TTTTTTTTTAG	^	CCTAATACAA	

Acceptor splice sites, complement strand

pos	3'→5'	pos	5'→3'	phase	strand	confidence	5'	intron	exon	3'
1798		339		0	-	0.17	TATTTTAAAG	^	ACAAAACTTT	
1116		1021		1	-	0.56	GTTACCTTAG	^	GGGTCAATAA	

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.3.123.574260.0 :

Start	End	Score	Exon	Intron
65	79	0.95	gctaaag	gtaatgga
77	91	0.98	ggaaaag	gtaaatta
150	164	0.98	ctaattg	gttagtac
1087	1101	0.97	caaaaag	gttagttt
1114	1128	0.96	ccctaag	gttaacaaa
1152	1166	0.95	cccaaga	gtgagtaa
1196	1210	0.79	catgtac	gtagaat

Donor site predictions for 10.42.2.148.574248.0 :

Start	End	Score	Exon	Intron
65	79	0.95	gctaaag	gtaatgga
77	91	0.98	ggaaaag	gtaaatta
150	164	0.98	ctaattg	gttagtac
1087	1101	0.97	caaaaag	gttagttt
1114	1128	0.96	ccctaag	gttaacaaa
1152	1166	0.95	cccaaga	gtgagtaa
1196	1210	0.79	catgtac	gtagaat

Acceptor site predictions for 10.42.3.123.574260.0 :

Start	End	Score	Intron	Exon
51	91	0.62	agcctaacttttatgctaa	aggtaatggaaaaggtaaatta
100	140	0.98	ctggttttgtttcatttcc	aggtggagggttagccgttgga
634	674	0.96	attgccttccttctttttt	agtatatataggtcacagaatc
643	683	0.96	cttcttttttagtatatat	aggtcacagaatcttggaattg
883	923	0.44	taacctggttttatatgtt	agggttttcacacacctcactg
933	973	0.74	ttctttttatgcaccttct	agtctggcattcattttctcct
1345	1385	1.00	cattatgtttttttttttt	agcttaatacaacaacctcaga
1454	1494	0.72	agttttccactattccat	agttaatctgataaccaggact
1550	1590	0.82	ctcacatgtttgtttaat	agcaataaacagtaaaatagga
1587	1627	0.83	aggaattacatcccttttt	agattattgcagtttcaagtag
1597	1637	0.41	tcccttttttagattattgc	agtttcaagtagtaaaagtaaaa
1763	1803	0.40	atcagcattttctccattcc	agtcaaaaagttttgtctttaa

Acceptor site predictions for 10.42.2.148.574248.0 :

Start	End	Score	Intron	Exon
51	91	0.62	agcctaacttttatgctaa	aggtaatggaaaaggtaaatta
100	140	0.98	ctggttttgtttcatttcc	aggtggagggttagccgttgga
634	674	0.96	attgccttccttctttttt	agtatatataggtcacagaatc
643	683	0.96	cttcttttttagtatatat	aggtcacagaatcttggaattg
883	923	0.44	taacctggttttatatgtt	agggttttcacacacctcactg
933	973	0.74	ttctttttatgcaccttct	agtctggcattcattttctcct
1345	1385	1.00	cattatgtttttttttttt	agcttaatacaacaacctcaga
1454	1494	0.72	agttttccactattccat	agttaatctgataaccaggact
1550	1590	0.82	ctcacatgtttgtttaat	agcaataaacagtaaaatagga
1587	1627	0.83	aggaattacatcccttttt	agattattgcagtttcaagtag
1597	1637	0.41	tcccttttttagattattgc	agtttcaagtagtaaaagtaaaa
1763	1803	0.40	atcagcattttctccattcc	agtcaaaaagttttgtctttaa

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
aaaat(a/g)tgtct	aaaata	aaaatg	28797	66%

Human Splicing Finder

<div><div></div>New Donor splice site</div> Activation of a cryptic Donor site. Potential alteration of splicing			
Algorithm/Matix	position	sequences	variation
HSF Donor site (matrix GT)	chr11:31522058	- REF : AAT A GTCT - ALT : AAT G GTCT	42.16 > 69.3 => 64.37%

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr	
wt	17	105	tacttacat		0.0121085273173	0.48	70	5	15	-3.7296
wt	17	100	acatcaata		-4.0349495049	0.473684210526	65	5	15	-4.9997626
wt	17	89	atataatct		-0.781235150664	0.488095238095	54	5	15	-3.0248438
wt	17	63	agtttagtg		-3.75061789496	0.448275862069	28	5	15	-2.5546017
wt	17	55	gcctgagta		0.202596405266	0.44	20	5	15	-0.5030177
wt	17	30	ttttaata		-4.70317476988	0.4	25	0	0	-2.893007
wt	17	29	ttttaataa		-1.13340847886	0.416666666667	24	0	0	-1.4265937
wt	17	26	taataaaag		-0.875425972321	0.428571428571	21	0	0	-1.1318415
wt	17	19	agctaaagta		-0.291948643561	0.5	14	0	0	-0.43722497
wt	17	15	aagtaaaat		-1.08163594501	0.6	10	0	0	-0.46093452
mut	17	105	tacttacat		0.0121085273173	0.48	70	5	15	-3.7296
mut	17	100	acatcaata		-4.0349495049	0.473684210526	65	5	15	-4.9997626
mut	17	89	atataatct		-0.781235150664	0.488095238095	54	5	15	-3.0248438
mut	17	63	agtttagtg		-3.75061789496	0.448275862069	28	5	15	-2.5546017
mut	17	55	gcctgagta		0.202596405266	0.44	20	5	15	-0.5030177
mut	17	30	ttttaata		-4.70317476988	0.4	18	8	13	-2.3288128
mut	17	29	ttttaataa		-1.13340847886	0.416666666667	17	8	13	-0.86239945
mut	17	26	taataaaag		-0.875425972321	0.428571428571	14	8	13	-0.56764728
mut	17	19	agctaaagta		-0.291948643561	0.5	7	8	13	0.12696926
mut	17	15	aagtaaaat		-1.08163594501	0.6	3	8	13	0.10325971

Variant Effect Predictor tool

ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000350638.10	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000379163.10	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000395934.2	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant, NMD_transcript_variant	ELP4	ENSG00000109911	Transcript	ENST00000474374.5	nonsense_mediated_decay	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000638184.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000638347.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant, NMD_transcript_variant	ELP4	ENSG00000109911	Transcript	ENST00000638376.1	nonsense_mediated_decay	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000638482.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant, non_coding_transcript_variant	ELP4	ENSG00000109911	Transcript	ENST00000638508.1	processed_transcript	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant, NMD_transcript_variant	ELP4	ENSG00000109911	Transcript	ENST00000638917.1	nonsense_mediated_decay	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant, NMD_transcript_variant	ELP4	ENSG00000109911	Transcript	ENST00000638984.1	nonsense_mediated_decay	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000639570.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000639878.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant, NMD_transcript_variant	ELP4	ENSG00000109911	Transcript	ENST00000640081.1	nonsense_mediated_decay	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000640231.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000640342.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000640533.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant, NMD_transcript_variant	ELP4	ENSG00000109911	Transcript	ENST00000640790.1	nonsense_mediated_decay	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant, non_coding_transcript_variant	ELP4	ENSG00000109911	Transcript	ENST00000640921.1	retained_intron	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000640954.1	protein_coding	rs970055506, COSV63351076
ENST00000640954.1:c.259+1970A>G	11:31522061-31522061	G	intron_variant	ELP4	ENSG00000109911	Transcript	ENST00000640961.2	protein_coding	rs970055506, COSV63351076

ESEfinder

La única predicción para la secuencia WT que tiene puntuación positiva es en las matrices 5'. Cuando se comparan las puntuaciones con las de la secuencia mutante, estas han aumentado, por lo que podría estar activándose un sitio *donor*.

2104 (-33)	tttaataaaaagctaagtaaaatgtgtotta	0.99240	2104 (-33)	tttaataaaaagctaagtaaaatgtgtotta	-17.24240	2104 (-33)	tttaataaaaagctaagtaaaatgtgtotta	0.24710	2104 (-33)	tttaataaaaagctaagtaaaatgtgtotta	-21.18720
2104 (-33)	tttaataaaaagctaagtaaaatgtgtotta	1.21900	2104 (-33)	tttaataaaaagctaagtaaaatgtgtotta	-17.25020	2104 (-33)	tttaataaaaagctaagtaaaatgtgtotta	0.61160	2104 (-33)	tttaataaaaagctaagtaaaatgtgtotta	-21.26730