

Ejemplo comparación de resultados predictores in silico

Cambio de estudio ZMYND8 c.3272-1970T>A (chr20:47212867 T/A, COSV53687647 o NM_001281769.2: c.3272-1970T>A)

Exón 22 e intrones adyacentes:

```
tgcat[taaaag[tggaca[atgggtgac[atgaatgaaggtgcca[actggttctaataacagta
gcctgttgatgaac[aaaaatagtaagca[cttgggggttg[gcggaat[tcag[ctctctgac
agactctgaggctacctgtg[ctatctcttcag[ttagcaaaagggtgtgacaaagca[actg[gc
ctatg[ccgcaaccaccacagaccaccagc[gcaccccaactacccgcgcccagaag[gttaa
gtatgaa[ccctgtctt[acgaaagctg[ccgggggc[ttctt[ctgggtatgcagaaggcttggtc
tgtgtgtgtc[cggtatgtgttcctttgtgggtgtgtc[ttctcctg[ttctttt[agttgt
gggtcttctctt[gcac[ctgttt[cttctt[cgctgtg[ctctggt[gcacatcacaaa[ctgcat
taggaaa[tgagagcgagcatgtggcctggcttctt[gtatggagat[cgctcct[gcgctctt[ttc
ataaaagggtcctagatgaaca[ccgattcc[ctctgtggtgattggctatggcagccagtg
tcc[cggtgtcct[ctccata[tgaggagtgagggcctgcagatccctccattagc[agcctgg
ttctgggagagagat[actgaac[tcag[cttggccctgggtccca[tttcttccag[taac[cgct
gggtgtctctg[ccctca[ccagagctggaaggac[ctgatgtgggttga[ccacatcactgta
ggat[tttctc[ccagatggggatgaccagctgeatcactccc[agagtc[ctccggagaaa
taaaagcataggcatgg[ccg[ctaagggaagaaggcataagaaaggcagattgaggtattgtt
aagaagac[tc[gc[cttgcctccagct[ctgcctaccatagtattgctccctttt[gtcac
tgatgcacatgactaaagaagagttctct[ctgagtaactacatgactttggagatctttg
tcattctcttcac[gccttatcttcagaactcatgattaccaca[acatg[ctttttacca
ggcagtaagaaatgcag[ccccctaggtagtagcaagggtccc[ctccttggc[ccactcac
tcttgccctcaacctctt[tcacacaaacattatcaccca[cccttaagaaattaaaaa[taa
ttctgtcc[agtc[ctgtctgaac[ttagta[catgatcaggatgtgatggga[ctcttagggct
tggctggaagggttctccagtcagc[catc[tagcagag[ctgcagatctgggctgggtgtt
ggctaaagtgtct[tcacagadac[ctcatt[cggtc[ctccttcag[cttctt[actt[attt
cttactcagtcacta[ctcagctccttgtcc[atgtgtccttgaagccat[cc[taggtcttat
tctgatt[ctgaattctt[agtcaccataagctt[gtcctta[ccgggagtcagtg[gtgtg
tgttccaggtggac[ctaaccattcttctcctt[batgat[ctttc[cttgggtggacaag
tgtgatttgggt[gttaaggcca[tttttcaagttgccta[acattgataaaaagaatccac
taa[gggaagttagactg[catgccaaatttcagtgcttttctcagggg[ccagggttgad[cc
cagaagtgcatagggtg[cttgggtg[cgttcagtgatgtagtaaggatacacggttccaga
cacggcacaaagaaatgcctt[cttggcagttcagac[cttttctacaaatctttattctc
tgtgagcatgacaccctacc[gttttttcaaccggtt[tcgttaaato[ctaaaaa[taagt
tgttctacctttc[gttctacag[accacagatgcaggg[agtcag[cggtggcagtggtt
gaac[catcaat[gggaggggaagggtgtt[ctgtgcaggaggggcaggtcagattgcag
gatagttgtgtcagcgccaggtgggcaggc[cgctaaa[acatt[gggttt[cc[ccctcag
ACCATTC[CGGAGTAAATAATCCAGTTGGAGCAGCAGTGATGAGAAGAGGGGATCGACAC
GTTCCGATCACAACACAGTACCAGCACGAACAGCCTCCTCCGAAAGAGTCTCGGCTGG
ACACCTTCTGGGACTAGCAAGTGAATCGGACACAAACCAACCCCACCCCATTGGGAGAAAAA
```

El cambio se encuentra en la primera línea del intrón 21 (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'
237	1	+	0.95	GCCCAGAAAGT	^	GTAAGTATGA	H
282	1	+	0.82	CTTCTTCTCG	^	GTATGCAGAA	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
2144	17	-	-	0.00	TCCCAATGGG	^	GTGGGTGGTT	
1759	402	2	-	0.34	TGAAAAACG	^	GTAGGGTGTC	
1422	739	1	-	0.34	CTGACTCCCG	^	GTAAGGACAA	
1120	1041	0	-	0.55	ATTTCTTAAG	^	GTGGGTGATA	
875	1286	2	-	0.34	CAATACTATG	^	GTAGGCAGAG	
678	1483	0	-	0.34	TCCAGCTCTG	^	GTAGGGCAG	

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
111	2	+	0.17	GGAATTCCAG	^	CTCTCTGACA	
122	1	+	0.19	TCTCTGACAG	^	ACTCTGAGGC	
130	0	+	0.20	AGACTCTGAG	^	GCTACCTGTG	
152	1	+	0.96	ATCTCTTCAG	^	TTAGCAAAAG	
156	2	+	0.41	CTTCAGTTAG	^	CAAAAGGTGT	
162	2	+	0.19	TTAGCAAAAG	^	GTGTGACAAG	
172	0	+	0.17	GTGTGACAAG	^	CAACCTGCCT	
200	1	+	0.22	ACCACCACAG	^	ACCACCAAGC	
652	1	+	0.25	TTTCTCCAG	^	TACCCGCTGG	
681	0	+	0.16	CCCTCACCAG	^	AGCTGGAAGG	
734	1	+	0.43	TTTCTCCAG	^	ATGGGGATGA	
987	2	+	0.16	TATCTTCCAG	^	AACTCATGAT	
1021	0	+	0.26	TTTTTACCAG	^	GCAGTAAGAA	
1232	1	+	0.17	AGCCATCTAG	^	CAGAGCTGCA	
1235	1	+	0.17	CATCTAGCAG	^	AGCTGCAGAT	
1237	0	+	0.17	TCTAGCAGAG	^	CTGCAGATCT	
1243	0	+	0.17	AGAGCTGCAG	^	ATCTGGGCTG	
1604	0	+	0.53	CTTTCTCCAG	^	GGGCCAAGGT	
1823	2	+	0.39	TGTTCTACAG	^	ACCCACAGAT	
1831	1	+	0.43	AGACCCACAG	^	ATGCAGGGCA	
1980	0	+	0.64	TCCCCCTCAG	^	ACCATTCCCG	
2059	2	+	0.19	ACAACACCAG	^	TACCAAGCAG	
2065	2	+	0.18	CCAGTACCAG	^	CACGAAGAGC	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
1892	269	1	-	0.27	TCCTGCACAG	^	GAAACCACT	
1788	373	1	-	0.26	TATTTTTTAG	^	GATTTAACGA	
684	1477	1	-	0.23	GTCCCTCCAG	^	CTCTGGTGAG	
603	1558	1	-	0.43	TCTCTCCAG	^	AACCAAGGCTG	
63	2098	2	-	0.00	TCATCAACAG	^	GCTACTGTTA	

Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
237	1	+	0.95	GCCCAGAAAGT	^	GTAAGTATGA	H
282	1	+	0.82	CTTCTTCTCG	^	GTATGCAGAA	

Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
2144	17	-	-	0.00	TCCCAATGGG	^	GTGGGTGGTT	
1759	402	2	-	0.34	TGAAAAACG	^	GTAGGGTGTC	
1422	739	1	-	0.34	CTGACTCCCG	^	GTAAGGACAA	
1120	1041	0	-	0.55	ATTTCTTAAG	^	GTGGGTGATA	
875	1286	2	-	0.34	CAATACTATG	^	GTAGGCAGAG	
678	1483	0	-	0.34	TCCAGCTCTG	^	GTAGGGCAG	

Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
111	2	+	0.17	GGAATTCCAG	^	CTCTCTGACA	
122	1	+	0.19	TCTCTGACAG	^	ACTCTGAGGC	
130	0	+	0.25	AGACTCTGAG	^	GCTACCTGTG	
152	1	+	0.97	ATCTCTTCAG	^	TTAGCAAAAG	
156	2	+	0.41	CTTCAGTTAG	^	CAAAAGGTGT	
162	2	+	0.19	TTAGCAAAAG	^	GTGTGACAAG	
172	0	+	0.17	GTGTGACAAG	^	CAACCTGCCT	
200	1	+	0.22	ACCACCACAG	^	ACCACCAAGC	
652	1	+	0.25	TTTCTCCAG	^	TACCCGCTGG	
681	0	+	0.16	CCCTCACCAG	^	AGCTGGAAGG	
734	1	+	0.43	TTTCTCCAG	^	ATGGGGATGA	
987	2	+	0.16	TATCTTCCAG	^	AACTCATGAT	
1021	0	+	0.26	TTTTTACCAG	^	GCAGTAAGAA	
1232	1	+	0.17	AGCCATCTAG	^	CAGAGCTGCA	
1235	1	+	0.17	CATCTAGCAG	^	AGCTGCAGAT	
1237	0	+	0.17	TCTAGCAGAG	^	CTGCAGATCT	
1243	0	+	0.17	AGAGCTGCAG	^	ATCTGGGCTG	
1604	0	+	0.53	CTTTCTCCAG	^	GGGCCAAGGT	
1823	2	+	0.39	TGTTCTACAG	^	ACCCACAGAT	
1831	1	+	0.43	AGACCCACAG	^	ATGCAGGGCA	
1980	0	+	0.64	TCCCCCTCAG	^	ACCATTCCCG	
2059	2	+	0.19	ACAACACCAG	^	TACCAAGCAG	
2065	2	+	0.18	CCAGTACCAG	^	CACGAAGAGC	

Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
1892	269	1	-	0.27	TCCTGCACAG	^	GAAACCACT	
1788	373	1	-	0.26	TATTTTTTAG	^	GATTTAACGA	
684	1477	1	-	0.23	GTCCCTCCAG	^	CTCTGGTGAG	
603	1558	1	-	0.43	TCTCTCCAG	^	AACCAAGGCTG	
63	2098	2	-	0.00	TCATCAACAG	^	GCTACTGTTA	

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
75	89	0.97	aaaaata	g t aagcac
156	170	0.74	gcaaaag	g t gtgaca
230	244	0.99	cagaagt	g t aagtat
275	289	0.85	cttcctg	g t atgcag
307	321	0.47	tgtccgt	g t atgtgt
1040	1054	0.70	cccctag	g t agtagc
1653	1667	0.51	tgatgta	g t aaggat
1935	1949	0.44	ggcgtag	g t gggcag

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
132	172	0.86	ctacctgtgctatctcttc	a gttagcaaaaggtgtgacaag
180	220	0.42	cctatgcccccaaccaccac	a gaccaccagccgacccccaac
269	309	0.55	gggcttcttctggtatgc	a gaaggcttgttctgtgtgtgt
466	506	0.82	gtcctcgctcttttcataa	a gggtcctagatgaacaccgat
632	672	0.66	ccctgggtccatttcttcc	a gtacccgctgggtgtctctgc
661	701	0.59	ggtggtctctgcccacc	a gagctggaaggacttgatgtg
714	754	0.82	cactgtaggattttctccc	a gatggggatgaccagctgcat
845	885	0.73	agactccgcccctgcccc	a gctctgcctaccatagtattg
967	1007	0.64	tcttcactgccttatcttcc	a gaactcatgattaccacaca
1001	1041	0.87	ccacacatgcctttttacc	a ggcagtaagaaatgcagcccc
1260	1300	0.71	tggctaaagtgcctttcac	a gacacctattcggctcttcc
1285	1325	0.87	ctcattcggctcttcttcc	a gcttcttacttatttcttac
1309	1349	0.71	cttcacttatttcttactc	a gtcactactcagctccttgtc
1429	1469	0.76	tcagtgggtgtgtgttccc	a ggtggaacttaaccattcttct
1584	1624	0.97	aatttcagtgcttttctcc	a ggggccaaggttgaccacaga
1803	1843	0.99	ttctacccttctgttctac	a gaccacagatgcagggcagt
1811	1851	0.44	ttctgtttacagaccac	a gatgcagggcagtcagcggtg
1960	2000	0.92	cacattgggttttccccctc	a gaccattcccggagtaataaa

Donor site predictions for mut :

Start	End	Score	Exon	Intron
75	89	0.97	aaaaata	g t aagcac
156	170	0.74	gcaaaag	g t gtgaca
230	244	0.99	cagaagt	g t aagtat
275	289	0.85	cttcctg	g t atgcag
307	321	0.47	tgtccgt	g t atgtgt
1040	1054	0.70	cccctag	g t agtagc
1653	1667	0.51	tgatgta	g t aaggat
1935	1949	0.44	ggcgtag	g t gggcag


Acceptor site predictions for mut :

Start	End	Score	Intron	Exon
132	172	0.86	ctacctgtgctatctcttc	a gttagcaaaaggtgtgacaag
180	220	0.42	cctatgcccccaaccaccac	a gaccaccagccgacccccaac
269	309	0.55	gggcttcttctggtatgc	a gaaggcttgttctgtgtgtgt
466	506	0.82	gtcctcgctcttttcataa	a gggtcctagatgaacaccgat
632	672	0.66	ccctgggtccatttcttcc	a gtacccgctgggtgtctctgc
661	701	0.59	ggtggtctctgcccacc	a gagctggaaggacttgatgtg
714	754	0.82	cactgtaggattttctccc	a gatggggatgaccagctgcat
845	885	0.73	agactccgcccctgcccc	a gctctgcctaccatagtattg
967	1007	0.64	tcttcactgccttatcttcc	a gaactcatgattaccacaca
1001	1041	0.87	ccacacatgcctttttacc	a ggcagtaagaaatgcagcccc
1260	1300	0.71	tggctaaagtgcctttcac	a gacacctattcggctcttcc
1285	1325	0.87	ctcattcggctcttcttcc	a gcttcttacttatttcttac
1309	1349	0.71	cttcacttatttcttactc	a gtcactactcagctccttgtc
1429	1469	0.76	tcagtgggtgtgtgttccc	a ggtggaacttaaccattcttct
1584	1624	0.97	aatttcagtgcttttctcc	a ggggccaaggttgaccacaga
1803	1843	0.99	ttctacccttctgttctac	a gaccacagatgcagggcagt
1811	1851	0.44	ttctgtttacagaccac	a gatgcagggcagtcagcggtg
1960	2000	0.92	cacattgggttttccccctc	a gaccattcccggagtaataaa

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
taaag(t/a)tgga	agttgg	agatgg	26609	53%

Human Splicing Finder

 Alteration of auxiliary sequences	Significant alteration of ESE / ESS motifs ratio (10)	
Algorithm/Matix	position	sequence
ESE_9G8 (New ESE Site)	chr20:47212868	GATGGA
ESS_hnRNPA1 (New ESS Site)	chr20:47212868	GATGGA
EIE (New ESE Site)	chr20:47212868	GATGGA
IIE (ESS Site Broken)	chr20:47212868	GTTGGA
EIE (New ESE Site)	chr20:47212869	AGATGG
IIE (ESS Site Broken)	chr20:47212869	AGTTGG
Fas ESS (New ESS Site)	chr20:47212869	AGATGG
ESE_Tra2 (New ESE Site)	chr20:47212870	AAGAT
RESCUE ESE (New ESE Site)	chr20:47212870	AAGATG
EIE (New ESE Site)	chr20:47212870	AAGATG
IIE (ESS Site Broken)	chr20:47212870	AAGTTG
ESE_9G8 (New ESE Site)	chr20:47212871	AAAGAT
RESCUE ESE (New ESE Site)	chr20:47212871	AAAGAT
EIE (New ESE Site)	chr20:47212871	AAAGAT
Sironi_motif1 (New ESS Site)	chr20:47212871	AAAGATGG
EIE (New ESE Site)	chr20:47212872	TAAAGA

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr		
wt	25	54	gcattaaag		-3.0639161255	0.387755102041	49	0	0	-3.7742743	
wt	25	53	cattaaagt		-0.381793832823	0.395833333333	48	0	0	-2.6581893	
wt	25	37	tggtgacat		1.4799215329	0.4375	32	0	0	-0.90300948	
wt	25	32	acatgaatg		-1.7412907212	0.444444444444	27	0	0	-1.8455319	
wt	25	28	gaatgaagg		-1.05071057807	0.478260869565	23	0	0	-1.3110221	
mut	25	54	gcattaaag		-3.0639161255	0.367346938776	49	0	0	-3.7808656	
mut	25	53	cattaaaga		-0.982504320265	0.375	48	0	0	-2.9001244	
mut	25	37	tggtgacat		1.4799215329	0.4375	32	0	0	-0.90300948	
mut	25	32	acatgaatg		-1.7412907212	0.444444444444	27	0	0	-1.8455319	
mut	25	28	gaatgaagg		-1.05071057807	0.478260869565	23	0	0	-1.3110221	

Se produce un cambio en las predicciones a causa de la presencia de la mutación pero estás solo hace que la puntuación de la región en la que se encuentra sea aún más negativa, por lo que no se va a tener en cuenta.

Variant Effect Predictor tool

ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000262975.8	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000311275.11	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000352431.6	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000355972.8	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000360911.7	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000372023.7	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000396281.8	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000446994.6	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000458360.6	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000461685.5	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000467200.6	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000471951.6	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000536340.5	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000540497.5	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000611941.4	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000617418.4	protein_coding	-	-	-	-	-	-	COSV53687647
ENST00000360911.7:c.3272-1970T>A	20:47212867-47212867	T	intron_variant	ZMYND8	ENSG00000101040	Transcript	ENST00000619049.4	protein_coding	-	-	-	-	-	-	COSV53687647

ESEfinder

Se observa un único resultado con puntuaciones positivas para las matrices 5'. Cuando se compara está puntuación con la equivalente en la secuencia mutante, se observa que ha aumentado, por lo que podría estar fortaleciéndose un sitio *donor*:

⁶ (-2155)	taaagttggacaatgggtgacatgaatgaag	2.61850	⁶ (-2155)	taaagttggacaatgggtgacatgaatgaag	-15.29150	⁶ (-2155)	taaagttggacaatgggtgacatgaatgaag	2.29110	⁶ (-2155)	taaagttggacaatgggtgacatgaatgaag	-16.48480
⁶ (-2155)	taaagatggacaatgggtgacatgaatgaag	3.00370	⁶ (-2155)	taaagatggacaatgggtgacatgaatgaag	-17.68670	⁶ (-2155)	taaagatggacaatgggtgacatgaatgaag	2.67660	⁶ (-2155)	taaagatggacaatgggtgacatgaatgaag	-19.03270