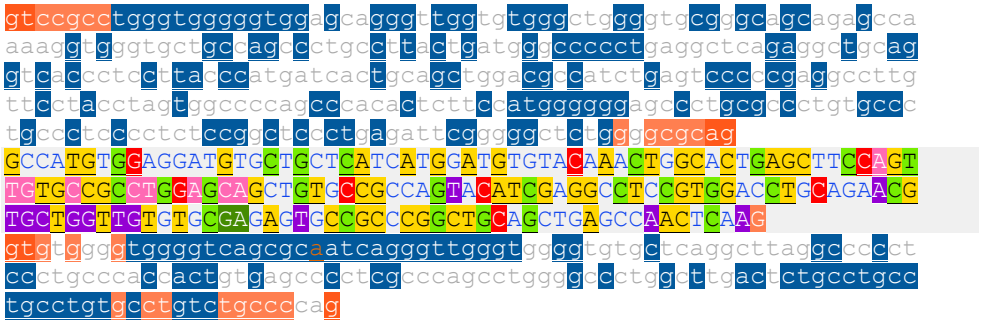
Ejemplo comparación de resultados predictores in sillico

Cambio de estudio LZTR1 c.1785+21A>G (chr22:20994748 A/G, rs178292 o NM_006767.4: c.1785+21A>G)

Exón 15 e intrones adyacentes:



El cambio se encuentra en la primera línea del intrón 15 (la **a** en color naranja subrayada en 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'
65	1	+	0.00	GAGCCAAAAG^GTGGGTGCTG
448	1	+	0.00	CTGCAGCTGA^GCCAACTCAA
459	0	+	0.94	CCAACTCAAG^GTGTGGGGTG H
461	2	+	0.24	AACTCAAGGT^GTGGGGTGGG
466	1	+	0.19	AAGGTGTGGG^GTGGGGTCAG
471	0	+	0.24	GTGGGGTGGG^GTCAGCGCAA

Donor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' exon intron 3'
187	414	1	-	0.41	GGGCCACTAG^GTAGGAACAA
134	467	0	-	0.47	GTGATCATGG^GTAAGGAGGG

Acceptor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5' intron exon 3'
266	0	+	0.55	GCTCCCTGAG^ATTCGGGGGC
288	1	+	0.34	TGGGGCGCAG^GCCATGTGGA
299	0	+	0.19	CCATGTGGAG^GATGTGCTGC

Acceptor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'
453	148	0	-	0.19	ACACCTTGAG^TTGGCTCAGC
444	157	0	-	0.32	GTTGGCTCAG^CTGCAGCCGG
438	163	0	-	0.34	TCAGCTGCAG^CCGGGCGGCA
411	190	0	-	0.20	ACACAACCAG^CACGTTCTGC
399	202	0	-	0.49	CGTTCTGCAG^GTCCACGGAG

Donor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5' exon intron 3'
65	1	+	0.00	GAGCCAAAAG^GTGGGTGCTG
448	1	+	0.00	CTGCAGCTGA^GCCAACTCAA
459	0	+	0.94	CCAACTCAAG^GTGTGGGGTG H
461	2	+	0.24	AACTCAAGGT^GTGGGGTGGG
466	1	+	0.19	AAGGTGTGGG^GTGGGGTCAG
471	0	+	0.24	GTGGGGTGGG^GTCAGCGCGA

Donor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' exon intron 3'
187	414	1	-	0.41	GGGCCACTAG^GTAGGAACAA
134	467	0	-	0.47	GTGATCATGG^GTAAGGAGGG

Acceptor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5' intron exon 3'
266	0	+	0.55	GCTCCCTGAG^ATTCGGGGGC
288	1	+	0.34	TGGGGCGCAG^GCCATGTGGA
299	0	+	0.19	CCATGTGGAG^GATGTGCTGC

Acceptor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'
453	148	0	-	0.20	ACACCTTGAG^TTGGCTCAGC
444	157	0	-	0.34	GTTGGCTCAG^CTGCAGCCGG
438	163	0	-	0.34	TCAGCTGCAG^CCGGGCGGCA
411	190	0	-	0.19	ACACAACCAG^CACGTTCTGC
399	202	0	-	0.39	CGTTCTGCAG^GTCCACGGAG

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.1.119.574590.0 :

Start	End	Score	Exon	Intron
58	72	0.96	ccaaaag	**gt**gggtgc
452	466	0.83	actcaag	**gt**gtgggg
526	540	0.60	caccact	**gt**gagccc

Acceptor site predictions for 10.42.1.119.574590.0 :

Start	End	Score	Intron	Exon
171	211	0.56	cgaggccttgttcctacct	**ag**tggccccagcccacactctt

Donor site predictions for 10.42.2.148.574602.0 :

Start	End	Score	Exon	Intron
58	72	0.96	ccaaaag	**gt**gggtgc
452	466	0.83	actcaag	**gt**gtgggg
526	540	0.60	caccact	**gt**gagccc

Acceptor site predictions for 10.42.2.148.574602.0 :

Start	End	Score	Intron	Exon
171	211	0.56	cgaggccttgttcctacct	**ag**tggccccagcccacactctt

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
agcgc(a/g)atcag	caatca	cgatca	26673	53%

Human Splicing Finder

Type	Interpretation
No significant impact on splicing signals.	No significant impact on splicing signals.

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr
wt	26	129	agctgagcc		0.905318440489	0.532258064516	66	20	31	-2.9607414
wt	26	120	aactcaagg		-0.651503836258	0.530434782609	57	20	31	-3.0012149
wt	26	102	gggtcagcg		-0.946399830243	0.577319587629	39	20	31	-1.9621684
wt	26	93	caatcaggg		-2.33396206765	0.590909090909	30	20	31	-1.9313907
wt	26	72	tgctcaggc		0.376263127657	0.65671641791	9	20	31	0.48030912
wt	26	66	ggcttaggc		-1.93653968107	0.672131147541	3	20	31	-0.040494057
wt	26	40	ctgtgagcc		0.0546228351802	0.657142857143	2	10	17	0.6671759

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr
mut	26	129	agctgagcc		0.905318440489	0.532258064516	66	20	31	-2.9607414
mut	26	120	aactcaagg		-0.651503836258	0.530434782609	57	20	31	-3.0012149
mut	26	102	gggtcagcg		-0.946399830243	0.577319587629	39	20	31	-1.9621684
mut	26	93	cgatcaggg		-2.53423033138	0.590909090909	30	20	31	-2.0098052
mut	26	72	tgctcaggc		0.376263127657	0.65671641791	9	20	31	0.48030912
mut	26	66	ggcttaggc		-1.93653968107	0.672131147541	3	20	31	-0.040494057
mut	26	40	ctgtgagcc		0.0546228351802	0.657142857143	2	10	17	0.6671759

El único cambio entre ambos es el BP que detecta sobre la posición de interés, por lo que la mutación marca la diferencia en el resultado, pero el *score* baja, por lo que es menos probable que en la secuencia WT de que sea un BP.

Variant Effect Predictor tool

ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	downstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000414985.5	nonsense_mediated_decay	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, NMD_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000415354.6	nonsense_mediated_decay	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, NMD_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000415817.2	nonsense_mediated_decay	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, NMD_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000439171.5	nonsense_mediated_decay	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	upstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000452988.5	nonsense_mediated_decay	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	downstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000461510.1	retained_intron	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	upstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000463909.1	retained_intron	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, non_coding_transcript_variant	AC002470.2	ENSG00000285314	Transcript	ENST00000479606.5	lncRNA	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	downstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000480895.1	processed_transcript	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, non_coding_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000491432.5	retained_intron	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	downstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000492480.1	retained_intron	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, non_coding_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000495142.6	retained_intron	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	downstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000497716.5	nonsense_mediated_decay	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	downstream_gene_variant	THAP7	ENSG00000184436	Transcript	ENST00000498406.1	retained_intron	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	upstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000498649.1	retained_intron	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, NMD_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000642151.1	nonsense_mediated_decay	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, non_coding_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000643578.1	retained_intron	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, non_coding_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000643710.1	processed_transcript	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	downstream_gene_variant	LZTR1	ENSG00000099949	Transcript	ENST00000644435.1	protein_coding	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant	LZTR1	ENSG00000099949	Transcript	ENST00000646124.2	protein_coding	-	-	rs178292
ENST00000646124.1:c.1785+21A>G	22:20994748-20994748	G	intron_variant, non_coding_transcript_variant	LZTR1	ENSG00000099949	Transcript	ENST00000646506.1	retained_intron	-	-	rs178292

ESEfinder

Para la secuencia WT se encuentran 4 resultados con la posición de interés y puntuación positiva en las matrices 5'SS (451, 456, 471 y 476) y un resultado para las matrices 3'SS(470):

451 (-150)	AACTCAAGgtgtggggtggggtcagcgcaa	5.07410	451 (-150)	AACTCAAGgtgtggggtggggtcagcgcaa	-23.69270	451 (-150)	AACTCAAGgtgtggggtggggtcagcgcaa	4.60130	451 (-150)	AACTCAAGgtgtggggtggggtcagcgcaa	-22.94080
456 (-145)	AAGgtgtggggtggggtcagcgcaatcagg	3.93040	456 (-145)	AAGgtgtggggtggggtcagcgcaatcagg	-26.06920	456 (-145)	AAGgtgtggggtggggtcagcgcaatcagg	3.71470	456 (-145)	AAGgtgtggggtggggtcagcgcaatcagg	-24.89110
470 (-131)	ggtcagcgcaatcagggttgggtggggtgt	-10.22620	470 (-131)	ggtcagcgcaatcagggttgggtggggtgt	0.83990	470 (-131)	ggtcagcgcaatcagggttgggtggggtgt	-8.03540	470 (-131)	ggtcagcgcaatcagggttgggtggggtgt	0.57680
471 (-130)	gtcagcgcaatcagggttgggtggggtgtg	3.58250	471 (-130)	gtcagcgcaatcagggttgggtggggtgtg	-16.48630	471 (-130)	gtcagcgcaatcagggttgggtggggtgtg	3.75010	471 (-130)	gtcagcgcaatcagggttgggtggggtgtg	-14.85800
476 (-125)	cgcaatcagggttgggtggggtgtgctcag	5.02830	476 (-125)	cgcaatcagggttgggtggggtgtgctcag	-18.71000	476 (-125)	cgcaatcagggttgggtggggtgtgctcag	4.93240	476 (-125)	cgcaatcagggttgggtggggtgtgctcag	-17.27440

Si se buscan los resultados equivalentes para la secuencia mutante, se observa que las puntuaciones positivas de las matrices 5'SS aumentan mientras que las de las matrices 3'SS disminuyen:

451 (-150)	AACTCAAGgtgtggggtggggtcagcgcga	5.41460	451 (-150)	AACTCAAGgtgtggggtggggtcagcgcga	-23.61260	451 (-150)	AACTCAAGgtgtggggtggggtcagcgcga	5.07100	451 (-150)	AACTCAAGgtgtggggtggggtcagcgcga	-22.91890
456 (-145)	AAGgtgtggggtggggtcagcgcgatcagg	4.17150	456 (-145)	AAGgtgtggggtggggtcagcgcgatcagg	-25.68220	456 (-145)	AAGgtgtggggtggggtcagcgcgatcagg	4.18490	456 (-145)	AAGgtgtggggtggggtcagcgcgatcagg	-24.55390
470 (-131)	ggtcagcgcgatcagggttgggtggggtgt	-10.13290	470 (-131)	ggtcagcgcgatcagggttgggtggggtgt	0.75270	470 (-131)	ggtcagcgcgatcagggttgggtggggtgt	-7.96410	470 (-131)	ggtcagcgcgatcagggttgggtggggtgt	0.42050
471 (-130)	gtcagcgcgatcagggttgggtggggtgtg	3.82590	471 (-130)	gtcagcgcgatcagggttgggtggggtgtg	-16.48690	471 (-130)	gtcagcgcgatcagggttgggtggggtgtg	3.98740	471 (-130)	gtcagcgcgatcagggttgggtggggtgtg	-14.75830
476 (-125)	cgcgatcagggttgggtggggtgtgctcag	5.26170	476 (-125)	cgcgatcagggttgggtggggtgtgctcag	-18.24620	476 (-125)	cgcgatcagggttgggtggggtgtgctcag	5.12030	476 (-125)	cgcgatcagggttgggtggggtgtgctcag	-16.66680

Aun así, son variaciones muy pequeñas, por lo que es muy probable que no estén afectando al *splicing.*