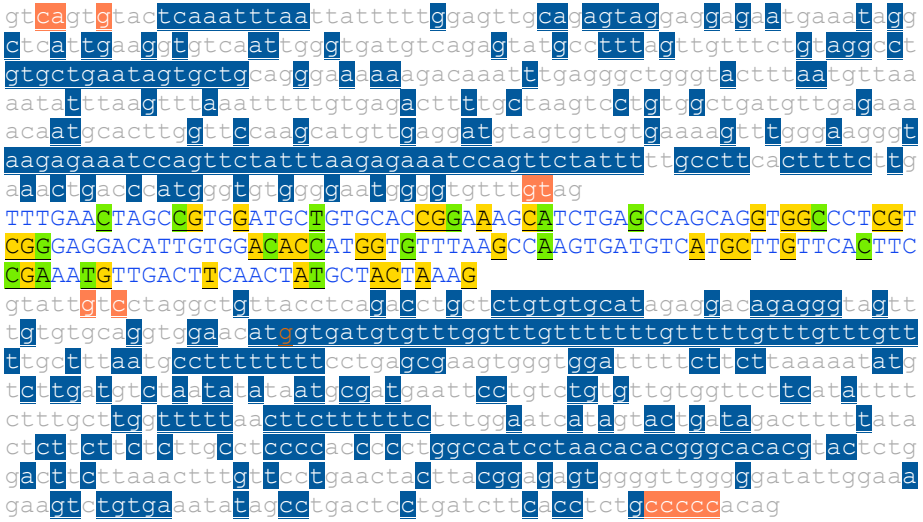
Ejemplo comparación de resultados predictores in sillico

Cambio de estudio ATXN2L c.616+79G>T (chr16:28826469 G/T, COSV57380552 o NM_001308230.2: c.616+79G>T)

Exón 5 e intrones adyacentes:



El cambio se encuentra en la segunda línea de intrón después del exón 5 (la **g** en color naranja 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'
299	1	+	0.67	TTTGGGAAGG^GTAAGAGAAA
550	1	+	0.82	GCTACTAAAG^GTATTGTCCT

Donor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' exon intron 3'
941	80	2	-	0.00	CCCACTCTCC^GTAAGTAGTT
368	653	2	-	0.31	ACACCCATGG^GTCAGTTTCA

Acceptor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5' intron exon 3'
116	2	+	0.25	GTTTCTGTAG^GCCTGTGCTG
398	0	+	0.25	GTGTTTGTAG^TTTGAACTAG
408	1	+	0.34	TTTGAACTAG^CCGTGGATGC
432	1	+	0.19	CACCGGAAAG^CATCTGAGCC
440	0	+	0.18	AGCATCTGAG^CCAGCAGGTG
444	1	+	0.17	TCTGAGCCAG^CAGGTGGCCC
447	1	+	0.07	GAGCCAGCAG^GTGGCCCTCG

Acceptor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'
582	439	0	-	0.26	ATGCACACAG^AGCAGGTCTG
544	477	2	-	0.07	ATACCTTTAG^TAGCATAGTT
541	480	2	-	0.07	CCTTTAGTAG^CATAGTTGAA
536	485	1	-	0.07	AGTAGCATAG^TTGAAGTCAA
530	491	1	-	0.17	ATAGTTGAAG^TCAACATTTC
134	887	1	-	0.78	TTCCCTGCAG^CACTATTCAG

Donor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5' exon intron 3'
299	1	+	0.67	TTTGGGAAGG^GTAAGAGAAA
550	1	+	0.82	GCTACTAAAG^GTATTGTCCT

Donor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' exon intron 3'
941	80	2	-	0.00	CCCACTCTCC^GTAAGTAGTT
368	653	2	-	0.31	ACACCCATGG^GTCAGTTTCA

Acceptor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5' intron exon 3'
116	2	+	0.25	GTTTCTGTAG^GCCTGTGCTG
398	0	+	0.25	GTGTTTGTAG^TTTGAACTAG
408	1	+	0.34	TTTGAACTAG^CCGTGGATGC
432	1	+	0.19	CACCGGAAAG^CATCTGAGCC
440	0	+	0.18	AGCATCTGAG^CCAGCAGGTG
444	1	+	0.17	TCTGAGCCAG^CAGGTGGCCC
447	1	+	0.07	GAGCCAGCAG^GTGGCCCTCG

Acceptor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'
582	439	0	-	0.26	ATGCACACAG^AGCAGGTCTG
544	477	2	-	0.07	ATACCTTTAG^TAGCATAGTT
541	480	2	-	0.07	CCTTTAGTAG^CATAGTTGAA
536	485	1	-	0.07	AGTAGCATAG^TTGAAGTCAA
530	491	1	-	0.17	ATAGTTGAAG^TCAACATTTC
134	887	1	-	0.78	TTCCCTGCAG^CACTATTCAG

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.3.123.574551.0 :

Start	End	Score	Exon	Intron
292	306	0.97	gggaagg	gtaagaga
543	557	0.78	actaaag	gtattgtc

Acceptor site predictions for 10.42.3.123.574551.0 :

Start	End	Score	Intron	Exon
96	136	0.95	tgcctttagttgtttctgt	agcctgtgctgaatagtgctg

Donor site predictions for 10.42.2.148.574539.0 :

Start	End	Score	Exon	Intron
292	306	0.97	gggaagg	gtaagaga
543	557	0.78	actaaag	gtattgtc

Acceptor site predictions for 10.42.2.148.574539.0 :

Start	End	Score	Intron	Exon
96	136	0.95	tgcctttagttgtttctgt	agcctgtgctgaatagtgctg

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
aacat(g/t)gtgat	aacatg	aacatt	27796	60%

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	67	54	tggtgatgt		1.33155803905	0.734693877551	1	7	11	1.1996063

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr
mut	67	54	ttgtgatgt		0.83930924355	0.734693877551	1	7	11	1.0068677

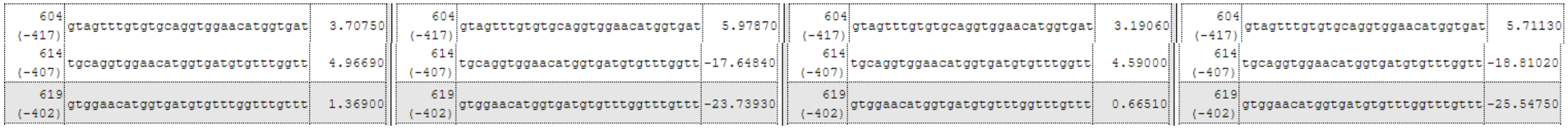
El BP que detecta con el cambio tiene peor puntuación con la secuencia mutante, por lo que puede que se esté debilitando.

Variant Effect Predictor tool

ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000325215.10	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000336783.8	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000340394.12	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000382686.8	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000395547.6	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	downstream_gene_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000561539.1	retained_intron	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	non_coding_transcript_exon_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000562867.5	retained_intron	5/5	473	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant, non_coding_transcript_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000563314.5	retained_intron	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	downstream_gene_variant	AC133550.2	ENSG00000260570	Transcript	ENST00000563565.1	lncRNA	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	upstream_gene_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000564284.1	retained_intron	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000564304.5	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000564656.5	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant, NMD_transcript_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000565971.5	nonsense_mediated_decay	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	downstream_gene_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000566080.1	retained_intron	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	upstream_gene_variant	AC145285.3	ENSG00000260796	Transcript	ENST00000568183.1	lncRNA	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000568266.5	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000570200.5	protein_coding	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	upstream_gene_variant	ATXN2L	ENSG00000168488	Transcript	ENST00000570284.1	processed_transcript	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	upstream_gene_variant	AC145285.6	ENSG00000275807	Transcript	ENST00000614819.1	lncRNA	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	intron_variant, non_coding_transcript_variant	AC009093.11	ENSG00000288630	Transcript	ENST00000675058.1	lncRNA	-	-	COSV57380552
ENST00000570200.5:c.616+79G>T	16:28826469-28826469	T	regulatory_region_variant	-	-	RegulatoryFeature	ENSR00001000737	CTCF_binding_site	-	-	COSV57380552

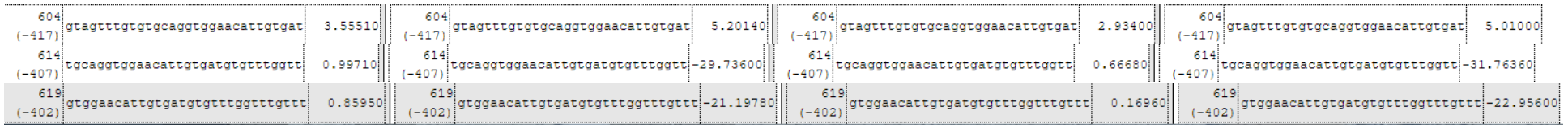
ESEfinder

Se obtienen puntuaciones positivas para la secuencia WT en tres de los resultados para las matrices 5'SS (en 604 también son positivas las matrices 3'SS):



604 (-417)	gtagtttgtgtgcaggtggaacatggtgat	3.70750	604 (-417)	gtagtttgtgtgcaggtggaacatggtgat	5.97870	604 (-417)	gtagtttgtgtgcaggtggaacatggtgat	3.19060	604 (-417)	gtagtttgtgtgcaggtggaacatggtgat	5.71130
614 (-407)	tgcaggtggaacatggtgatgtgtttggtt	4.96690	614 (-407)	tgcaggtggaacatggtgatgtgtttggtt	-17.64840	614 (-407)	tgcaggtggaacatggtgatgtgtttggtt	4.59000	614 (-407)	tgcaggtggaacatggtgatgtgtttggtt	-18.81020
619 (-402)	gtggaacatggtgatgtgtttggtttgttt	1.36900	619 (-402)	gtggaacatggtgatgtgtttggtttgttt	-23.73930	619 (-402)	gtggaacatggtgatgtgtttggtttgttt	0.66510	619 (-402)	gtggaacatggtgatgtgtttggtttgttt	-25.54750

Si buscamos las predicciones equivalentes para la secuencia mutante, para 604 las puntuaciones descienden en todas las matrices un poco, en 614 las puntuaciones disminuyen considerablemente, mientras que en 619 descienden menos que la anterior pero más que en la primera.



604 (-417)	gtagtttgtgtgcaggtggaacattgtgat	3.55510	604 (-417)	gtagtttgtgtgcaggtggaacattgtgat	5.20140	604 (-417)	gtagtttgtgtgcaggtggaacattgtgat	2.93400	604 (-417)	gtagtttgtgtgcaggtggaacattgtgat	5.01000
614 (-407)	tgcaggtggaacattgtgatgtgtttggtt	0.99710	614 (-407)	tgcaggtggaacattgtgatgtgtttggtt	-29.73600	614 (-407)	tgcaggtggaacattgtgatgtgtttggtt	0.66680	614 (-407)	tgcaggtggaacattgtgatgtgtttggtt	-31.76360
619 (-402)	gtggaacattgtgatgtgtttggtttgttt	0.85950	619 (-402)	gtggaacattgtgatgtgtttggtttgttt	-21.19780	619 (-402)	gtggaacattgtgatgtgtttggtttgttt	0.16960	619 (-402)	gtggaacattgtgatgtgtttggtttgttt	-22.95600

Por lo tanto, se está debilitando o casi perdiendo un sitio 5'SS (*donor*) en la secuencia mutante (614), que puede estar afectando al *splicing* (aunque dado que está en el interior del intrón es poco probable).