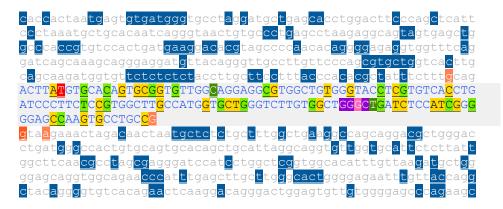
Ejemplo comparación de resultados predictores in sillico

Cambio de estudio RHCE c.939+1G>T (chr1:25388975 G/T, COSV53799157 o NM 001330430.4:c.939+1G>T)

Exón 6 e intrones adyacentes:



El cambio se encuentra en la posición justo después del exón 6 (la **g** en color rojo).

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	t strand				
	pos 5'->3'	phase strand	confidence	5' exon intron 3'	Donor splice sites, direct strand	
	83 439	1 + 0 +	0.00 0.93	CACAATCAGG^GTAACTGTGC GTGCCTGCCG^GTAAGAAACT		xon intron 3' AGG^GTAACTGTGC
Donor splice	sites, comple	ement strand			Donor splice sites, complement strand	
pos 3'->5' 357 347 265	pos 5'->3' 382 392 474	phase strand 0 - 1 - 0 -	confidence 0.55 0.71 0.47	5' exon intron 3' AAGGGATCAG^GTGACACGAG GTGACACGAG^GTACCCACAG AGGAAGCAAG^GTAGAGAGAG	357 382 0 - 0.36 AAGGGAT 347 392 1 - 0.70 GTGACAC	xon intron 3' CAG^GTGACACGAG GAG^GTACCCACAG AAG^GTAGAGAGAG
Acceptor spli	ce sites, dir				Acceptor splice sites, direct strand	
	300	phase strand 0 +	0.94	5' intron exon 3' TTCTTTGCAG^ACTTATGTGC	· · · · · · · · · · · · · · · · · · ·	ron exon 3' CAG^ACTTATGTGC
Acceptor spli	ce sites, com	mplement strand	d -		Acceptor splice sites, complement strand	
pos 3'->5' 433 411 406	pos 5'->3' 306 328 333	phase strand 0 - 1 - 0 -	confidence 0.25 0.17 0.17	5' intron exon 3' TTACCGGCAG^GCACTTGGCT CCCGATGGAG^ATCAGCCCAG TGGAGATCAG^CCCAGCCACA	411 328 1 - 0.07 CCCGATG	ron exon 3' GAG^ATCAGCCCAG CAG^CCCAGCCACA

Pérdida sitio *donor* (en rojo) de la secuencia WT a la mutante. Este sitio coincide con el sitio *donor* del exón, por lo que, además teniendo en cuenta la *confidence* del otro sitio predicho, lo más probable es que se esté perdiendo el exón completo.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 85.53.15.54.27132.0:

Start	End	Score	Exon Intron
76	90	0.91	aatcagg gt aactgt
432	446	1.00	cctgccg gt aagaaa

Donor site predictions for 85.53.15.54.26979.0:

Start	End	Score	Exon	Intron
76	90	0.91	aatcag	g gt aactgt

Acceptor site predictions for 85.53.15.54.27132.0:

Acceptor site predictions for 85.53.15.54.26979.0:

Start	End	Score	Intron	Exon	Start	End	Score	Intron	Exon
205	245	0.53	acagggttgcctt	gttccc ag cgtgctggtcacttgcagca	205	245	0.53	acagggttgccttgttccc	agcgtgctggtcacttgcagca
280	320	0.99	ccacacgctattt	ctttgc ag acttatgtgcacagtgcggt	280	320	0.99	ccacacgctatttctttgc	agacttatgtgcacagtgcggt

Pérdida sitio *donor* (en rojo) de la secuencia WT a la mutante. Este sitio coincide con el sitio *donor* del exón, por lo que o se está perdiendo el exón completo o, si se empleara el otro sitio predicho, se produciría la inclusión de un exón críptico de 218 nt.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
tgccg(g/t)taaga	gtaaga	ttaaga	27516	58%

Human Splicing Finder

Broken WT Donor Site	Alteration of the WT Donor site, most probably affecting splicing

Algorithm/Matix	position	sequences	variation
HSF Donor site (matrix GT)	chr1:25388978	- REF : CCGGTAAGA - ALT : CCGTTAAGA	91.16 > 64.02 => -29.77%
MaxEnt Donor site	chr1:25388978	- REF : CCGGTAAGA - ALT : CCGTTAAGA	10.1 > 1.59 => -84.26%

SVM-BPfinder

con i	2007	cc die	t he see he see	y cont ppt off ppt len ppt scr svm	cen			son in	2007	cc dia	st bo son bo ser	y_cont ppt_off ppt_len ppt_scr svm	cen		
seq_i			st bp_seq bp_scr		_			seq_io	agez		st bp_seq bp_scr				
wt	14	403	ggatcagca	-2.65631212097 0.527638190955 26	13	24	-1.8900583	mut	14	403	ggatcagca	-2.65631212097 0.530150753769 26	13	24	-1.8892468
wt	14	383	atgttacag	-1.99366012052 0.542328042328 6	13	24	-0.35988787	mut	14	383	atgttacag	-1.99366012052 0.544973544974 6	13	24	-0.35903345
wt	14	352	tggtcactt	-0.150943414816 0.538904899135 16	32	58	0.044272737	mut	14	352	tggtcactt	-0.150943414816 0.541786743516 16	32	58	0.045203482
wt	14	310	cctttaccc	-0.680423085566 0.518032786885 9	12	23	-0.052750974	mut	14	310	cctttaccc	-0.680423085566 0.52131147541 9	12	23	-0.051692059
wt	14	283	gacttatgt	-0.404603822609 0.503597122302 49	20	31	-2.4068231	mut	14	283	gacttatgt	-0.404603822609 0.507194244604 49	20	31	-2.4056613
wt	14	232	gtgtcacct	0.35371237824 0.515418502203 1	17	28	0.9042834	mut	14	232	gtgtcacct	0.35371237824 0.519823788546 1	17	28	0.90570617
wt	14	227	acctgatcc	1.18156024726 0.513513513514 1	12	25	1.1998622	mut	14	227	acctgatcc	1.18156024726 0.518018018018 1	12	25	1.201317
wt	14	179	ggctgatct	1.70877188798 0.488505747126 50	13	25	-1.703404	mut	14	179	ggctgatct	1.70877188798 0.494252873563 21	11	14	0.031629124
wt	14	146	cggtaagaa	-0.559711448659 0.489361702128 17	13	25	-0.50250147	mut	14	147	ccgttaaga	-3.52790715244 0.485915492958 18	13	25	-1.7291017
wt	14	130	aactaatgc	2.33048933798 0.512 1 13	25	1.6492	2334	mut	14	146	cgttaagaa	-0.987313936044 0.489361702128 17	13	25	-0.66992804
wt	14	111	ggctgaagg	0.445006669724 0.481132075472 70	12	23	-3.4852075	mut	14	130	aactaatgc	2.33048933798 0.512 1 13	25	1.649	2334
wt	14	86	acctgatgg	1.49086365031 0.518518518519 45	12	23	-1.4811723	mut	14	111	ggctgaagg	0.445006669724 0.481132075472 70	12	23	-3.4852075
wt	14	56	gcattaggc	-3.15873910529 0.529411764706 15	12	23	-1.3992437	mut	14	86	acctgatgg	1.49086365031 0.518518518519 45	12	23	-1.4811723
wt	14	32	ctcttattg	-1.05752687218 0.518518518519 27	0	0	-1.5538823	mut	14	56	gcattaggc	-3.15873910529 0.529411764706 15	12	23	-1.3992437
wt	14	23	gcttcaacg	-0.739755939636 0.444444444444 18	0	0	-0.88369878	mut	14	32	ctcttattg	-1.05752687218 0.518518518519 27	0	0	-1.5538823
			- 0					mut	14	23	gcttcaacg	-0.739755939636 0.444444444444 18	0	0	-0.88369878

Aparece un nuevo BP en la secuencia mutante (en 147), aunque tiene puntuación negativa, por lo que no lo tendremos en cuenta. El otro cambio es en 146, donde cambia el BP porque está la mutación (la tercera base pasa de g a t). Como las puntuaciones son negativas, no se tienen en cuenta.

Variant Effect Predictor tool

ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	splice_donor_variant	RHCE	ENSG00000188672	Transcript	ENST00000294413.13	protein_coding	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	splice_donor_variant	RHCE	ENSG00000188672	Transcript	ENST00000340849.8	protein_coding	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	intron_variant	RHCE	ENSG00000188672	Transcript	ENST00000346452.8	protein_coding	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	splice_donor_variant	RHCE	ENSG00000188672	Transcript	ENST00000349320.7	protein_coding	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	splice_donor_variant	RHCE	ENSG00000188672	Transcript	ENST00000349438.8	protein_coding	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	splice_donor_variant	RHCE	ENSG00000188672	Transcript	ENST00000413854.5	protein_coding	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	splice_donor_variant, non_coding_transcript_variant	RHCE	ENSG00000188672	Transcript	ENST00000527187.5	retained_intron	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	intron_variant	RHCE	ENSG00000188672	Transcript	ENST00000527747.5	protein_coding	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:939+1G>T	1:25388975- 25388975	Α	splice_donor_variant, NMD_transcript_variant	RHCE	ENSG00000188672	Transcript	ENST00000533771.1	nonsense_mediated_decay	-	-	rs751221592, COSV53799157
ENST00000413854.5:c.	:.939+1G>T	1:25388975- 25388975	Α	regulatory_region_variant	-	-	RegulatoryFeature	ENSR00000351164	TF_binding_site	-	-	rs751221592, COSV53799157

ESEfinder

Se obtiene un único resultado donde se encuentra la posición de interés (g detrás de las mayúsculas que indican el exón) con puntuaciones positivas para las matrices 5'SS:

414	414	
CATCGGGGGAGCCAAGTGCCTGCCGgtaag 4.10120	CATCGGGGGAGCCAAGTGCCTGCCGgtaag -13.27090 CATCGGGGGAGCCAAGTGCCTGCCGgtaag 4.27770 CATCGGGGGAGCCAAGTGCCTGCCGgtaag -17.	00700
(-325) CATCOOCOGACCCAACTCCCTCCCGGCAAG	-325) CATCGGGGGAGCCAAGTGCCTGCCGGCAAGTGCCTGCCGGCAAGTGCCTGCC	.00700
1/:	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	

Si buscamos la resultado equivalente en la secuencia mutante, se observa que las puntuaciones para las matrices 5'SS se reducen un poco:

			Le			
414		414		414	11	414
111 CATCGGGGAGCCAAGTGCCTGCCG++==a	3.93550	*11 CATCGGGGAGCCAAGTGCCTGCCG++884	10 70460	717 CATCGGGGAGCCAAGTGCCTGCCG++==«	3 00570	*11 CATCCCCCAACCCAACTCCCCC++ 17 40350
/ 22E/ CAICGGGGGGGGCCAAGIGCCIGCCGCGGG	3.93330	/ 225 LECATORGEOGRAGICARGIGCOIGCOGLIANG	-13./2400	/ 225 CAICGGGGGGGGCCAAGIGCCIGCCGCGGG	3.905/0	/ ane CAICGGGGGGGCCAAGIGCCIGCCGTTAAG -1/.40350
(-323)		(-323)		(-323)		(-325)

Por lo tanto, el sitio se está debilitando por lo que podría afectar al splicing.