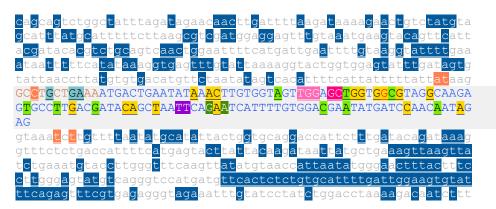
#### Ejemplo comparación de resultados predictores in sillico

Cambio de estudio KRAS c.38G>A (chr12:25245347 G/A, rs112445441 o NM\_033360.4: c.38G>A)

#### Exón 2 e intrones adyacentes:



El cambio se encuentra en la primera línea del exón 2 (la **g** en color rojo subrayada de amarillo).

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 s	sites, direct	strand	d -			Donor splice s	sites, direct	strand		
	pos 5'->3' 199 217 423	phase 1 1 0	strand + + +	confidence 0.61 0.37 0.65	5' exon intron 3' TCATATAAAG^GTGAGTTTGT GTATTAAAAG^GTACTGGTGG AACAATAGAG^GTAAATCTTG		pos 5'->3' 199 217 423	phase strand 1 + 1 + 0 +	confidence 0.58 0.37 0.70	5' exon intron 3' TCATATAAAG^GTGAGTTTGT GTATTAAAAG^GTACTGGTGG AACAATAGAG^GTAAATCTTG
Donor splice s	sites, comple	ment st	trand			Donor splice s	sites, comple	ment strand		
pos 3'->5' 554	pos 5'->3' 169	phase 0	strand -	confidence 0.41	5' exon intron 3' GAAACCCAAG^GTACATTTCA	pos 3'->5' 554	pos 5'->3' 169	phase strand 0 -	confidence 0.41	5' exon intron 3' GAAACCCAAG^GTACATTTCA
Acceptor splic	ce sites, dir	ect str	rand			Acceptor splic	ce sites, dir	ect strand		
	pos 5'->3' 300 458	phase 1 1	strand + +	confidence 0.18 0.19	5' intron exon 3' TTATTATAAG^GCCTGCTGAA ACTGGTGCAG^GACCATTCTT		pos 5'->3' 300 336 458	phase strand 1 + 1 + 1 +	confidence 0.23 0.07 0.18	5' intron exon 3' TTATTATAAG^GCCTGCTGAA CTTGTGGTAG^TTGGAGCTGG ACTGGTGCAG^GACCATTCTT

Aparece un sitio *acceptor* en la secuencia mutante. Este está en medio del exón por lo que, en el caso de que se utilizara este y no el normal, se perderían 36 pb del exón.

## **Splice Site Prediction by Neural Network (NNSplice)**

# Donor site predictions for 85.53.81.31.9334.0:

# **Donor site predictions for 85.53.81.31.9371.0:**

Start	End	Score	Exon Intron	Start	End	Score	Exon Intron
192	206	0.99	tataaag <b>gt</b> gagttt	192	206	0.99	tataaag <b>gt</b> gagttt
210	224	0.79	ttaaaag <b>gt</b> actggt	210	224	0.79	ttaaaag <b>gt</b> actggt
416	430	0.99	aatagag <b>gt</b> aaatct	416	430	0.99	aatagag <b>gt</b> aaatct

# Acceptor site predictions for 85.53.81.31.9334.0:

# Acceptor site predictions for 85.53.81.31.9371.0:

Start	End	Score	Intron	Exon	Start	End	Score	Intron	Exon
280	320	0.69	tttcattatttttattata	<b>ag</b> gcctgctgaaaatgactgaa	280	320	0.69	tttcattatttttattata	<b>3g</b> gcctgctgaaaatgactgaa
456	496	0.67	caggaccattctttgatac	<b>ag</b> ataaaggtttctctgaccat	456	496	0.67	caggaccattcttgatac	<b>ag</b> ataaaggtttctctgaccat

### **Spliceman**

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
tggtg(g/a)cgtag	ggtggc	ggtgac	27567	59%

#### **CRYP-SKIP**

#### Results for sequence wt

Exon length (bp)	122	EVOK
PESS (<=-2.62) density	0.00	<b>EXSK</b> 0.5
NN 5'ss score density	0.04	
SF2/ASF score density	7.94	
FAS-ESS (hex2) density	3.28	// \
EIE score density	410.12	// \
Probability of cryptic splice site activation (PcR-E)	0.37	0
>wt		
	0,29	0.07

ttcattatttttattataagGCCTGCTGAAAATGACTGAATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGGCAAGA GTGCCTTGACGATACAGCTAATTCAGAATCATTTTGTGGACGAATATGATCCAACAATAGAGgtaaatc

#### **Human Splicing Finder**

New Donor splice site Activa	tion of a cryptic Donor si	te. Potential alteration of splic	ting
Algorithm/Matix	position	sequences	variation
MaxEnt Donor site	chr12:25245348	- REF : GGCGTAGGC - ALT : GACGTAGGC	-0.29 > 3.52 => 1313.79%
HSF Donor site (matrix GT)	chr12:25245353	- REF : CTGGTGGCG - ALT : CTGGTGACG	71.59 > 81.89 => 14.39%

Parece que hay un sitio críptico de *splicing* dentro del propio exón, pero el cambio de interés (justo en el siguiente nt) no lo toma en consideración, por lo que no debe considerar que tenga algún efecto en el *splicing*.

CR-E

0.5

## SVM-BPfinder

seq_id	agez	ss_dis	t bp_seq bp_scr	y_cont ppt_off ppt_len ppt_sc	r svm_s	cr			seq_io	agez	ss_dis	st bp_seq bp_scr	y_cont ppt_off	ppt_len ppt_scr	svm_sc	r		
wt	15	154	ttttcatta	-2.14414808302 0.44966442953	1	11	23	-0.1415629	mut	15	154	ttttcatta	-2.14414808302	0.44966442953	1	11	23	-0.1415629
wt	15	151	tcattattt	-3.16680438209 0.445205479452	1	8	19	-0.58068463	mut	15	151	tcattattt	-3.16680438209	0.445205479452	1	8	19	-0.58068463
wt	15	145	tttttatta	-4.64419963387 0.428571428571	. 96	9	16	-7.2058152	mut	15	145	tttttatta	-4.64419963387	0.428571428571	96	9	16	-7.2058152
wt	15	142	ttattataa	-3.40175220342 0.423357664234	93	9	16	-6.5311271	mut	15	142	ttattataa	-3.40175220342	0.423357664234	93	9	16	-6.5311271
wt	15	139	ttataaggc	-1.77318081181 0.425373134328	90	9	16	-5.7029186	mut	15	139	ttataaggc	-1.77318081181	0.425373134328	90	9	16	-5.7029186
wt	15	129	tgctgaaaa	0.481644311634 0.41935483871	80	9	16	-4.1890085	mut	15	129	tgctgaaaa	0.481644311634	0.41935483871	80	9	16	-4.1890085
wt	15	123	aaatgactg	1.10723640726 0.432203389831	. 74	9	16	-3.56012	mut	15	123	aaatgactg	1.10723640726	0.432203389831	74	9	16	-3.56012
wt	15	119	gactgaata	0.297230706651 0.429824561404	70	9	16	-3.6248505	mut	15	119	gactgaata	0.297230706651	0.429824561404	70	9	16	-3.6248505
wt	15	113	atataaact	-1.22794910316 0.435185185185	64	9	16	-3.8405092	mut	15	113	atataaact	-1.22794910316	0.435185185185	64	9	16	-3.8405092
wt	15	69	ccttgacga	1.41053075025 0.453125	20	9	16	-0.016499459	mut	15	89	tggtgacgt	2.3610579405	0.440476190476	40	9	16	-0.91437476
wt	15	57	agctaattc	1.17370942078 0.461538461538	8	9	16	0.65307092	mut	15	69	ccttgacga	1.41053075025	0.453125	20	9	16	-0.016499459
wt	15	53	aattcagaa	-2.47812047387 0.4375 4	9	16	-0.531	36354	mut	15	57	agctaattc	1.17370942078	0.461538461538	8	9	16	0.65307092
wt	15	47	gaatcattt	-2.2186001094 0.452380952381	. 1	6	13	-0.26299544	mut	15	53	aattcagaa	-2.47812047387	0.4375 4	9	16	-0.531	36354
wt	15	29	atatgatcc	-0.837615331155 0.458333333333	24	0	0	-1.2973196	mut	15	47	gaatcattt	-2.2186001094	0.452380952381	1	6	13	-0.26299544
									mut	15	29	atatgatcc	-0.837615331155	0.458333333333	24	0	0	-1.2973196

Hay un nuevo BP en la secuencia mutante, pero tiene puntuación negativa, por lo que no lo tendremos en cuenta.

## **Variant Effect Predictor tool**

ENST00000557334.5:c.38G>A	12:25245347- T 25245347	missense_variant	KRAS	ENSG00000133703 Transcript	ENST00000256078.10 pro	otein_coding 2	2/6	228	38	13	G/D	GGC/GAC rs112445441, CM125166, COSV55497357, COSV55497388, COSV55522580
ENST00000557334.5:c.38G>A	12:25245347- T 25245347	missense_variant	KRAS	ENSG00000133703 Transcript	ENST00000311936.8 pro	otein_coding 2	2/5	228	38	13	G/D	GGC/GAC rs112445441, CM125166, COSV55497357, COSV55497388, COSV55522580
ENST00000557334.5:c.38G>A	12:25245347- T 25245347	missense_variant	KRAS	ENSG00000133703 Transcript	ENST00000556131.1 pro	otein_coding 2	2/3	215	38	13	G/D	GGC/GAC rs112445441, CM125166, COSV55497357, COSV55497388, COSV55522580
ENST00000557334.5:c.38G>A	<u>12:25245347-</u> T <u>25245347</u>	missense_variant	KRAS	ENSG00000133703 Transcript	<u>ENST00000557334.5</u> pro	otein_coding 2	2/3	235	38	13	G/D	GGC/GAC <u>rs112445441</u> , <u>CM125166</u> , <u>COSV55497357</u> , <u>COSV55497388</u> , <u>COSV55522580</u>

#### **ESEfinder**

Hay tres resultados con puntuaciones positivas en las matrices 5'SS:

331 (-392)	TGGTAGTTGGAGCTGGTGGCGTAGGCAAGA	1.32130	331 (-392) TGGTAGTTGGAGCTGGTGGCGTAGGCAAGA	-10.04130	331 (-392) TGGTAGTTGGAGCTGGTGGCGTAGGCAAGA	1.20920	331 (-392) TGGTAGTTGGAGCTGGTGGCGTAGGCAAGA -11.43830
336 (-387)	GTTGGAGCTGGTGGCGTAGGCAAGAGTGCC	1.20390	336 (-387) GTTGGAGCTGGTGGCGTAGGCAAGAGTGCC	-32.95400	336 (-387) GTTGGAGCTGGTGGCGTAGGCAAGAGTGCC	1.06100	336 (-387) GTTGGAGCTGGTGGCGTAGGCAAGAGTGCC -31.65020
340 (-383)	GAGCTGGTGGCGTAGGCAAGAGTGCCTTGA	4.48380	340 (-383) GAGCTGGTGGCGTAGGCAAGAGTGCCTTGA	-0.89750	340 (-383) GAGCTGGTGGCGTAGGCAAGAGTGCCTTGA	4.30110	340 GAGCTGGTGGCGTAGGCAAGAGTGCCTTGA -0.94060

Si buscamos las predicciones equivalentes para la secuencia mutante, vemos que para 331 suben pero para las otras dos se mantienen prácticamente igual.

331 (-392)	TGGTAGTTGGAGCTGGTGACGTAGGCAAGA	3.62390	331 (-392) TGGTAGTTGGAGCTGGTGACGTAGGCAAGA -10.258	331 (-392) TGGTAGTTGGAGCTGGTGACGTAGGCAAGA	3.60440	331 (-392) TGGTAGTTGGAGCTGGTGACGTAGGCAAGA -11.59820
336 (-387)	GTTGGAGCTGGTGACGTAGGCAAGAGTGCC	3.50940	336 (-387) GTTGGAGCTGGTGACGTAGGCAAGAGTGCC -16.374	336 GTTGGAGCTGGTGACGTAGGCAAGAGTGCC	3.42360	336 (-387) GTTGGAGCTGGCGTAGGCAAGAGTGCC -17.11290
340 (-383)	GAGCTGGTGACGTAGGCAAGAGTGCCTTGA	4.39050	340 (-383) GAGCTGGTGACGTAGGCAAGAGTGCCTTGA -0.810	0 (-383) GAGCTGGTGACGTAGGCAAGAGTGCCTTG	A 4.22980	340 (-383) GAGCTGGTGACGTAGGCAAGAGTGCCTTGA -0.78430

Por lo tanto, es probable que se esté fortaleciendo un sitio donor críptico.

En cuenta a los sitios ESE, se observa que las puntuaciones para 346 se hacen mayores, por lo que puede que se esté activando un ESE.

342 (-381)	GCTGGTG	-5.38319	342 (-381) GCTGGTG -4.15254	342 (-381) GCTGGTGG -3.79819	342 (-381) GCTGGTG -3.52864
343 (-380)	CTGGTGG	-0.93875	343 (-380) CTGGTGG 0.62148	343 (-380) CTGGTGGC -7.48497	343 CTGGTGG 1.07681
344 (-379)	TGGTGGC	-2.31648	344 (-379) TGGTGGC -1.08007	344 (-379) TGGTGGCG -0.59383	344 (-379) TGGTGGC -2.88777
345 (-378)	GGTGGCG	-4.56267	345 (-378) GGTGGCG -3.27828	345 (-378) GGTGGCGT -2.29326	345 (-378) GGTGGCG -2.63326
346 (-377)	GTGGCGT	0.38121	346 GTGGCGT -0.25726	346 (-377) GTGGCGTA -0.56694	346 (-377) GTGGCGT -4.94382
347 (-376)	TGGCGTA	-0.77254	347 TGGCGTA 0.05577	347 (-376) TGGCGTAG -1.55536	347 TGGCGTA -4.06479
348 (-375)	GGCGTAG	-4.17226	348 GGCGTAG -2.94563	348 (-375) GGCGTAGG -0.60252	348 (-375) GGCGTAG -1.93491

342 (-381)	GCTGGTG	-5.38319	342 (-381) GCTGGTG -4.152	4 (-38	GCTGGTGA	-4.24809	342 (-381)	GCTGGTG	-3.52864
343 (-380)	CTGGTGA	-0.21197	343 (-380) CTGGTGA 1.056	4 (-38)	CTGGTGAC	-6.18808	343 (-380)	CTGGTGA	-1.30493
344 (-379)	TGGTGAC	-4.89421	344 (-379) TGGTGAC -2.859	2 (-37	TGGTGACG	-0.95531	344 (-379)	TGGTGAC	-3.45560
345 (-378)	GGTGACG	-6.48228	345 (-378) GGTGACG -4.9840	6 (-37)	GGTGACGI	-1.86800	345 (-378)	GGTGACG	-0.07870
346 (-377)	GTGACGT	3.28979	346 (-377) GTGACGT 1.462	9 (-37	GTGACGTA	-0.56694	346 (-377)	GTGACGT	-4.78852
347 (-376)	TGACGTA	-2.83633	347 (-376) TGACGTA -1.327	8 (-37	ITGACGTAG	-0.25847	347 (-376)	TGACGTA	-1.45599
348 (-375)	GACGTAG	-3.72504	348 (-375) GACGTAG -3.2438	5 (-37	GACGTAGG	-0.96338	348 (-375)	GACGTAG	-3.38008

## **EX-SKIP**

Seq	PESS	FAS-ESS hex2	FAS-ESS hex3	IIE	IIE	NI-ESS trusted	NI-ESS all	PESE	RESCUE -ESE	EIE	EIE	NI-ESE trusted	NI-ESE all	ESS	ESE	ESS/ESE
	(count)	(count)	(count)	(count)	(sum)	(count)	(sum)	(count)	(count)	(count)	(sum)	(count)	(sum)	(total)	(total)	(ratio)
wt	0	4	3	33	510.6662	16	-23.7770	8	15	44	500.3453	49	58.2806	56	116	0.48
mut	0	3	2	30	465.7483	15	-21.9393	8	15	43	492.3333	53	62.0733	50	119	0.42

Allele wt has a higher chance of exon skipping than allele mut.

## **HOT-SKIP**

189	49	G	GCTGGTGGCGTAGGC	TGGTGGCGTAG	0	1	1	4	54.2346	1	-1.8376	0	0	1	8.0120	0	0.8397	7	1	7.00
190	49	Α	GCTGGTGACGTAGGC	TGGTGACGTAG	0	0	0	1	9.3168	0	0.0000	0	0	0	0.0000	4	4.6324	1	4	0.25
191	49	С	GCTGGTGCCGTAGGC	TGGTGCCGTAG	0	0	0	2	14.6412	0	0.0000	0	0	0	0.0000	0	1.2747	2	0	2.00
192	49	T	GCTGGTGTCGTAGGC	TGGTGTCGTAG	0	0	0	3	43.5254	0	-1.1609	0	1	0	0.0000	1	1.0000	3	2	1.50