

## Ejemplo comparación de resultados predictores in silico

Cambio de estudio TP53 c.266G>A (chr17:7674220 G/A, rs11540652 o NM\_000546.6: c.266G>A)

Exón 3 e intrones adyacentes:

```
aggagaatggcggtgaacctggcggtggagcttgcagtgagctgagatcacgccactgca
ctccagcctgggcgacagagcgagattccatctcaaaaaaaaaaaaaaggcctccct
gcttgccacaggtctcccgaaggcgcactggcctcatcttgggcctgtgttatctcctag
GTTGGCTCTGACTGTACCACCATCCACTACAACCTACATGTGTACAGTTCTGCATGGGC
GGCATGAACCGGAGGCCATCCTCACCATCATCACACTGGAAGACTCCAG
gtcaggagccacttgccacctgcacactggcctgctgtgcccagcctctgcttgctc
tgaccctgggcccacctcttaacgatttcttccatactactaccatccacctctcatc
acatcccggcggggaatctccttactgctccactcagttctcttttctctggctttg
```

El cambio se encuentra en la segunda línea del exón 3 (la **g** en color rojo y subrayada en amarillo a la izquierda de la primera g en 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 strand										Donor splice sites, direct strand													
-----										-----													
	pos	5'->3'	phase	strand	confidence	5'	exon	intron	3'		pos	5'->3'	phase	strand	confidence	5'	exon	intron	3'				
	291		2	+	0.91	AAGACTCCAG		^GTCAGGAGCC	H		291		2	+	0.91	AAGACTCCAG		^GTCAGGAGCC	H				
-----										-----													
Donor splice sites, complement strand										Donor splice sites, complement strand													
-----										-----													
	pos	3'->5'	pos	5'->3'	phase	strand	confidence	5'	exon	intron	3'		pos	3'->5'	pos	5'->3'	phase	strand	confidence	5'	exon	intron	3'
	373		98		1	-	0.63	GAAGAAATCG		^GTAAGAGGTG		373		98		1	-	0.63	GAAGAAATCG		^GTAAGAGGTG		
	366		105		2	-	0.37	TCGGTAAGAG		^GTGGGCCAG		366		105		2	-	0.37	TCGGTAAGAG		^GTGGGCCAG		
-----										-----													
Acceptor splice sites, direct strand										Acceptor splice sites, direct strand													
-----										-----													
	pos	5'->3'	phase	strand	confidence	5'	intron	exon	3'		pos	5'->3'	phase	strand	confidence	5'	intron	exon	3'				
	131		0	+	0.56	CTTGCCACAG		^GTCTCCCAA			131		0	+	0.54	CTTGCCACAG		^GTCTCCCAA					
	180		0	+	1.00	TATCTCCTAG		^GTTGGCTCTG	H		180		0	+	0.98	TATCTCCTAG		^GTTGGCTCTG	H				
	227		2	+	0.17	TGTGTAACAG		^TTCCTGCATG			227		2	+	0.17	TGTGTAACAG		^TTCCTGCATG					
-----										-----													
Acceptor splice sites, complement strand										Acceptor splice sites, complement strand													
-----										-----													
	pos	3'->5'	pos	5'->3'	phase	strand	confidence	5'	intron	exon	3'		pos	3'->5'	pos	5'->3'	phase	strand	confidence	5'	intron	exon	3'
	285		186		0	-	0.18	TGACCTGGAG		^TCTTCCAGTG		285		186		0	-	0.17	TGACCTGGAG		^TCTTCCAGTG		
	277		194		2	-	0.19	AGTCTTCCAG		^TGTGATGATG		277		194		2	-	0.18	AGTCTTCCAG		^TGTGATGATG		
	262		209		0	-	0.19	TGATGGTGAG		^GATGGGCCTC		262		209		0	-	0.19	TGATGGTGAG		^GATGGGCCTC		

Aunque dos de los *acceptor* bajan su confianza un poco, los sitios siguen siendo los mismos, por lo que no causa mayores consecuencias.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
284	298	0.90	actccag	gtcaggag

Donor site predictions for mut :

Start	End	Score	Exon	Intron
284	298	0.90	actccag	gtcaggag

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
111	151	0.97	ggcctccctgcttgccac	aggtctccccaggcgactgg
160	200	0.96	tgggcctgtgttatctcct	aggttggtctgtactgtaccac
429	469	0.58	ctccttactgctcccactc	agttttcttttctctggctttg

Acceptor site predictions for mut :

Start	End	Score	Intron	Exon
111	151	0.97	ggcctccctgcttgccac	aggtctccccaggcgactgg
160	200	0.96	tgggcctgtgttatctcct	aggttggtctgtactgtaccac
429	469	0.58	ctccttactgctcccactc	agttttcttttctctggctttg

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
gaacc(g/a)gaggc	cggagg	cagagg	30592	78%

Human Splicing Finder

 New Acceptor splice site		Activation of a cryptic Acceptor site. Potential alteration of splicing	
Algorithm/Matix	position	sequences	variation
HSF Acceptor site (matrix AG)	chr17:7674230	- REF : GGCATGAACCGGAG - ALT : GGCATGAACCAAGAG	42.91 > 70.78 => 64.95%

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr			
wt	13	62	gcatgaacc		-0.554895229849	0.526315789474	9	18	24	0.0083901318		
wt	13	44	tcctcacca		2.05523814278	0.512820512821	39	0	0	-1.0965077		
wt	13	38	ccatcatca		-1.79062982431	0.484848484848	33	0	0	-2.2315911		
wt	13	35	tcatcacac		-0.982973709553	0.466666666667	30	0	0	-1.7313328		
wt	13	15	aggtcagga		-3.39037967028	0.5	10	0	0	-1.3972139		
mut	13	62	gcatgaacc		-0.554895229849	0.526315789474	9	18	24	0.0083901318		
mut	13	44	tcctcacca		2.05523814278	0.512820512821	39	0	0	-1.0965077		
mut	13	38	ccatcatca		-1.79062982431	0.484848484848	33	0	0	-2.2315911		
mut	13	35	tcatcacac		-0.982973709553	0.466666666667	30	0	0	-1.7313328		
mut	13	15	aggtcagga		-3.39037967028	0.5	10	0	0	-1.3972139		

## Variant Effect Predictor tool

ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	downstream_gene_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000604348.5</a>	protein_coding	-	-	-	-	-	-	-	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000610292.4</a>	protein_coding	6/10	993	626	209	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000610538.4</a>	protein_coding	7/12	876	626	209	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000610623.4</a>	protein_coding	3/8	625	266	89	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000618944.4</a>	protein_coding	3/8	625	266	89	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000619186.4</a>	protein_coding	3/7	625	266	89	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000619485.4</a>	protein_coding	7/11	879	626	209	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000620739.4</a>	protein_coding	7/11	933	626	209	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000622645.4</a>	protein_coding	7/12	876	626	209	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	
ENST00000619186.4:c.266G>A	<a href="#">17-7674220-7674220</a>	T	missense_variant, NMID_transcript_variant	TP53	<a href="#">ENSG000000141510</a>	Transcript	<a href="#">ENST00000635293.1</a>	nonsense_mediated_decay	7/12	886	626	209	R/Q	CGG/CAG	rs11540652, CM920675, COSV52661091, COSV52661580, COSV52675468, COSV52802300	

## ESEfinder

No hay puntuaciones positivas para ninguna de las secuencias.

En cuanto a los ESE, se producen algunas alteraciones que pueden estar afectando al *splicing*:

245 (-226)	TGAACCG	-2.42315	245 (-226)	TGAACCG	-1.25361	245 (-226)	TGAACCGG	-1.62681	245 (-226)	TGAACCG	1.89123
246 (-225)	GAACCGG	0.98144	246 (-225)	GAACCGG	-0.10348	246 (-225)	GAACCGGA	1.39492	246 (-225)	GAACCGG	-0.18362
247 (-224)	AACCGGA	2.48624	247 (-224)	AACCGGA	1.65143	247 (-224)	AACCGGAG	1.50263	247 (-224)	AACCGGA	-4.32378
248 (-223)	ACCGGAG	-4.45876	248 (-223)	ACCGGAG	-2.55758	248 (-223)	ACCGGAGG	-4.52911	248 (-223)	ACCGGAG	-1.96748
249 (-222)	CCGGAGG	-1.54168	249 (-222)	CCGGAGG	0.27017	249 (-222)	CCGGAGGC	-8.58655	249 (-222)	CCGGAGG	1.63459
250 (-221)	CGGAGGC	3.08488	250 (-221)	CGGAGGC	3.68554	250 (-221)	CGGAGGCC	-5.30017	250 (-221)	CGGAGGC	-2.42886
251 (-220)	GGAGGCC	-6.04253	251 (-220)	GGAGGCC	-3.96346	251 (-220)	GGAGGCCC	-1.05945	251 (-220)	GGAGGCC	-2.42862

  

245 (-226)	TGAACCA	-1.69637	245 (-226)	TGAACCA	-0.81824	245 (-226)	TGAACCAG	-0.32992	245 (-226)	TGAACCA	-0.49050
246 (-225)	GAACCAG	-1.59630	246 (-225)	GAACCAG	-1.88314	246 (-225)	GAACCAGA	1.03343	246 (-225)	GAACCAG	-0.75145
247 (-224)	AACCAGA	0.56663	247 (-224)	AACCAGA	-0.05496	247 (-224)	AACCAGAG	1.92790	247 (-224)	AACCAGA	-1.76922
248 (-223)	ACCAGAG	-1.55019	248 (-223)	ACCAGAG	-0.83782	248 (-223)	ACCAGAGG	-4.52911	248 (-223)	ACCAGAG	-1.81218
249 (-222)	CCAGAGG	-3.60547	249 (-222)	CCAGAGG	-1.11358	249 (-222)	CCAGAGGC	-7.28965	249 (-222)	CCAGAGG	4.24338
250 (-221)	CAGAGGC	3.53210	250 (-221)	CAGAGGC	3.38732	250 (-221)	CAGAGGCC	-5.66102	250 (-221)	CAGAGGC	-3.87403
251 (-220)	AGAGGCC	-6.97040	251 (-220)	AGAGGCC	-4.20101	251 (-220)	AGAGGCCC	-2.81152	251 (-220)	AGAGGCC	-0.97390

EX-SKIP

Seq	PESS (count)	FAS-ESS hex2 (count)	FAS-ESS hex3 (count)	IIE (count)	IIE (sum)	NI-ESS trusted (count)	NI-ESS all (sum)	PESE (count)	RESCUE -ESE (count)	EIE (count)	EIE (sum)	NI-ESE trusted (count)	NI-ESE all (sum)	ESS (total)	ESE (total)	ESS/ESE (ratio)
wt	1	3	2	12	154.5907	6	-10.0546	9	12	33	407.2447	41	53.6785	24	95	0.25
mut	1	3	2	12	154.5907	6	-10.0546	10	15	34	422.8006	41	54.0224	24	100	0.24

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

HOT-SKIP

aggagaatggcgtgaacctgggcgggtggagcttgacgtgagctgagatcacgccactgcactccagcctgggcgacagag  
cgagattccatctcaaaaaaaaaaaaaaaaaaggcctccccctgcttgccacagggtctccccaaggcgccactggcctcatctt  
gggcctgtgttatctctctagGTTGGCTCTGACTGTACCACCATCCACTACAACATCATGTGTAAACAGTTCTTGCATGGGCT  
GGCATGAACCGGAGGCCCATCCTCACCATCATCACACTGGAAGACTCCAGgtcaggagccacttgccaccctgcacactg  
gcctgctgtgccccagcctctgcttgccctctgacccctggggcccacctcttaccgatttcttcatactactacctatccatcc  
acctctcatcacatccccgggcggggaatctccttactgctccactcagttttctttctctgtgctttgg