

## Ejemplo comparación de resultados predictores in silico

Cambio de estudio DGKI c.1403-841A>T (chr7:137488530 A/T, rs78033912 o NM\_001321708.2: c.1403-841A>T)

Exón 23 e intrones adyacentes:

```
caacactgtcttatcatttgccttgcacttaccagttggttaattcaaaaatttgaagggt
tttatgttagatcatggatttaatcgtgtaattgtactttttaaatagagaaattttaga
tttttttttaagttttgggaagaagactatttttcagtaatttcaggactcttgtgattct
atctgaagtaggttctatacaataaagccactgtgtgtgtgttgacactacctaagtagga
tattttataaaggtgcataaacttagtcttgccttattgttttctgaggtcctactgttacc
actcttaacattggccatccttaaaattagttggattttttaagctttgctcttctagaa
gctgttttttgttcttttctttctaacattttatgccttttagttctgatattggccacca
ggttagcaaagctgctttagtagctcaagtcataaaataggaaaactgtttagacagac
aaaaagatgaagttgtctgcattcaggaaggaggagtaagtgtccagaatcctcaatgt
ttaattcccccgggcactgcatcctcatgttgcctattgcaaatgcaagcataaatttagg
cataaagcataaaaatgcaagcataattatcaactgtcctatgagcagctccctgttctg
ccaatccatcgttactgagacaattcacttcatctcagggccagtgaaacacaatgagat
gaggggcagggtggcaaggggcaacatggatcatgctagggaagtgtgtgacaactgcaaa
caagaagagactgaaataataaaaaggtgttagggagaaagtacttctttctttttacgg
acatatctaaaaaaccataataattatgagattgttttttaaccaagagaaacacattt
atgctgatatagaaaaatcaaatatgtttattttagattttttcttcttcttcaaccttag
CGATACCTCTGGGTATTCTAGTTGTGCGTGGAGACTGTGATTTGGAGACTTGCCGTATGT
ACATAGA CCGCCTACAGGAG
gtgagtttattttagccaattttttattctcctcaactttccatttttgtaaacaattacc
agcatatatatttgcttctgaaatgatgcatttggagaggaagtggaaatgtgggtggct
ttgtccttcaagcatacgtgttatgtcaacctctccactgctctgttgttccactctc
```

El cambio se encuentra en la segunda línea del intrón 22 (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'
422	0	+	0.32	TGGCCACCAG	^	GTAGCAACAG	
517	1	+	0.54	GAAGGAAGGA	^	GTAAGTGTCC	
1041	0	+	0.83	CCTACAGGAG	^	GTGAGTTTAT	

### Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
966	255	1	-	0.32	ATACCCAGAG	^	GTATCGCTAA	
308	913	1	-	0.39	GATGGCCAAT	^	GTAAGAGTGG	
32	1189	1	-	0.00	TAACCAACTG	^	GTAAGTAGCA	

### Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
698	1	+	0.33	TTCATCTCAG	^	GGCCAGTGAA	
960	1	+	0.20	TCACCCTTAG	^	CGATACCTCT	

### Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
791	430	2	-	0.25	ATTATTTT	CAG^TCTCTTCTTG		
774	447	1	-	0.15	TTGTTTG	CAG^TTGTACACACA		
231	990	1	-	0.26	TCCTACTTAG	^	GTAAGTGCAA	

### Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
422	0	+	0.32	TGGCCACCAG	^	GTAGCAACAG	
517	1	+	0.54	GAAGGAAGGA	^	GTAAGTGTCC	
1041	0	+	0.83	CCTACAGGAG	^	GTGAGTTTAT	

### Donor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	exon	intron	3'
966	255	1	-	0.32	ATACCCAGAG	^	GTATCGCTAA	
308	913	1	-	0.39	GATGGCCAAT	^	GTAAGAGTGG	
32	1189	1	-	0.00	TAACCAACTG	^	GTAAGTAGCA	

### Acceptor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
132	2	+	0.15	TTTTTTTAAG	^	TTTTGGAAAG	
698	1	+	0.33	TTCATCTCAG	^	GGCCAGTGAA	
960	1	+	0.20	TCACCCTTAG	^	CGATACCTCT	

### Acceptor splice sites, complement strand

pos 3'→5'	pos 5'→3'	phase	strand	confidence	5'	intron	exon	3'
791	430	2	-	0.25	ATTATTTT	CAG^TCTCTTCTTG		
774	447	1	-	0.15	TTGTTTG	CAG^TTGTACACACA		
231	990	1	-	0.26	TCCTACTTAG	^	GTAAGTGCAA	

Se activa un sitio *acceptor* en la secuencia mutante. Si se empleara este en vez del normal para el exón 23, se incluiría un exón críptico de 828 pb (si se empleara el *donor* del exón).

## Splice Site Prediction by Neural Network (NNSplice)

### Donor site predictions for 10.42.2.148.574719.0 :

Start	End	Score	Exon	Intron
181	195	0.70	atctgaa	<b>gt</b> aggttc
245	259	0.66	ttataag	<b>gt</b> gcataa
415	429	0.84	ccaccag	<b>gt</b> agcaac
510	524	0.95	ggaagga	<b>gt</b> aagtgt
1034	1048	0.99	acaggag	<b>gt</b> gagttt
1174	1188	0.44	atacgct	<b>gt</b> atgtca

### Acceptor site predictions for 10.42.2.148.574719.0 :

Start	End	Score	Intron	Exon
112	152	0.63	attttagat	ttttttt <b>ag</b> ttttggaaagaagactat
145	185	0.75	gactat	ttttcagtat <b>ag</b> gactcttgatctatctg
338	378	0.56	tttaagct	ttgcctctt <b>ag</b> aagctgtttttgttcttt
382	422	0.76	tctaacatt	ttatgcctt <b>ag</b> ttctgatattggccaccagg
678	718	0.88	cgacaatt	cacttcatt <b>ag</b> ggccagtgaacacaaatgag
940	980	0.72	ttcttctt	ccttcaccct <b>ag</b> cgatacctctgggtattcta
961	1001	0.42	cgatacct	ctgggtatt <b>ag</b> ttgtgcgtggagactgtgat

### Donor site predictions for 10.42.3.123.574731.0 :

Start	End	Score	Exon	Intron
181	195	0.70	atctgaa	<b>gt</b> aggttc
245	259	0.66	ttataag	<b>gt</b> gcataa
415	429	0.84	ccaccag	<b>gt</b> agcaac
510	524	0.95	ggaagga	<b>gt</b> aagtgt
1034	1048	0.99	acaggag	<b>gt</b> gagttt
1174	1188	0.44	atacgct	<b>gt</b> atgtca

### Acceptor site predictions for 10.42.3.123.574731.0 :


Start	End	Score	Intron	Exon
112	152	0.94	attttagat	ttttttt <b>ag</b> ttttggaaagaagactat
145	185	0.75	gactat	ttttcagtat <b>ag</b> gactcttgatctatctg
338	378	0.56	tttaagct	ttgcctctt <b>ag</b> aagctgtttttgttcttt
382	422	0.76	tctaacatt	ttatgcctt <b>ag</b> ttctgatattggccaccagg
678	718	0.88	cgacaatt	cacttcatt <b>ag</b> ggccagtgaacacaaatgag
940	980	0.72	ttcttctt	ccttcaccct <b>ag</b> cgatacctctgggtattcta
961	1001	0.42	cgatacct	ctgggtatt <b>ag</b> ttgtgcgtggagactgtgat

Uno de los sitios *acceptor* (en azul) se ve alterado por la presencia de la mutación. Esto hace que el sitio tenga mayor *score*, lo que podría provocar que el *spliceosome* lo reconociera y se produjera la inclusión de un exón críptico (cuyo tamaño dependería del *donor* que se empleara, por ejemplo, de 828 nt si se empleara el *donor* del exón).

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
tgtag(a/t)ttttt	attttt	tttttt	29359	70%

Human Splicing Finder

 New Acceptor splice site	Activation of a cryptic Acceptor site. Potential alteration of splicing (cryptic exon activation)		
Algorithm/Matix	position	sequences	variation
MaxEnt Acceptor site	chr7:137488537	- REF : TTTGTAGATTTTTTTTAAAGTTT - ALT : TTTGTAGTTTTTTTTTAAAGTTT	3.96 > 5.64 => 42.42%

## SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr			
wt	23	87	cttttaaaa	-4.70827108165	0.536585365854	19	9	27	-2.2195717			
wt	23	86	ttttaaaat	-1.48716558778	0.543209876543	18	9	27	-0.89291885			
wt	23	81	aaatgagaa	-2.3894228687	0.565789473684	13	9	27	-0.92241108			
wt	23	59	tttttaagt	-4.32406925658	0.555555555556	17	9	20	-2.0016259			
wt	23	58	ttttaagtt	-1.56707208489	0.566037735849	16	9	20	-0.85544755			
wt	23	33	ttttcagta	-2.14439090512	0.642857142857	21	8	17	-1.4011239			
wt	23	25	atttcagga	-3.29626517601	0.65	13	8	17	-1.3434436			
mut	24	88	cttttaaaa	-4.70827108165	0.542168674699	18	10	30	-2.1265226			
mut	24	87	ttttaaaat	-1.48716558778	0.548780487805	17	10	30	-0.79987384			
mut	24	82	aaatgagaa	-2.3894228687	0.571428571429	12	10	30	-0.82934395			
mut	24	60	tttttaagt	-4.32406925658	0.545454545455	17	9	20	-2.0048882			
mut	24	59	ttttaagtt	-1.56707208489	0.555555555556	16	9	20	-0.85883297			
mut	24	34	ttttcagta	-2.14439090512	0.620689655172	21	8	17	-1.4082833			
mut	24	26	atttcagga	-3.29626517601	0.619047619048	13	8	17	-1.3534403			

## Variant Effect Predictor tool

ENST00000453654.6:c.1403-841A>T	<a href="#">7:137488530-137488530</a>	A	intron_variant	DGKI	<a href="#">ENSG00000157680</a>	Transcript	<a href="#">ENST00000288490.9</a>	protein_coding	-	-	<a href="#">rs78033912</a>
ENST00000453654.6:c.1403-841A>T	<a href="#">7:137488530-137488530</a>	A	intron_variant	DGKI	<a href="#">ENSG00000157680</a>	Transcript	<a href="#">ENST00000424189.6</a>	protein_coding	-	-	<a href="#">rs78033912</a>
ENST00000453654.6:c.1403-841A>T	<a href="#">7:137488530-137488530</a>	A	intron_variant	DGKI	<a href="#">ENSG00000157680</a>	Transcript	<a href="#">ENST00000446122.5</a>	protein_coding	-	-	<a href="#">rs78033912</a>
ENST00000453654.6:c.1403-841A>T	<a href="#">7:137488530-137488530</a>	A	intron_variant	DGKI	<a href="#">ENSG00000157680</a>	Transcript	<a href="#">ENST00000453654.6</a>	protein_coding	-	-	<a href="#">rs78033912</a>
ENST00000453654.6:c.1403-841A>T	<a href="#">7:137488530-137488530</a>	A	intron_variant, non_coding_transcript_variant	DGKI	<a href="#">ENSG00000157680</a>	Transcript	<a href="#">ENST00000460662.2</a>	processed_transcript	-	-	<a href="#">rs78033912</a>
ENST00000453654.6:c.1403-841A>T	<a href="#">7:137488530-137488530</a>	A	intron_variant	DGKI	<a href="#">ENSG00000157680</a>	Transcript	<a href="#">ENST00000614521.1</a>	protein_coding	-	-	<a href="#">rs78033912</a>

## ESEfinder

Solo se observa una predicción con puntuación positiva para las matrices 3'SS:

118 (-1103)	agattttttttaagttttggaaagaagac	-16.53120	118 (-1103)	agattttttttaagttttggaaagaagac	4.30310	118 (-1103)	agattttttttaagttttggaaagaagac	-17.04630	118 (-1103)	agattttttttaagttttggaaagaagac	3.92880
----------------	-------------------------------	-----------	----------------	-------------------------------	---------	----------------	-------------------------------	-----------	----------------	-------------------------------	---------

Si observamos la predicción equivalente para la secuencia mutante, se comprueba que la puntuación aumenta de manera considerable, por lo que puede que se esté fortaleciendo un sitio 3'SS (*acceptor*):

118 (-1103)	agtttttttttaagttttggaaagaagac	-16.90590	118 (-1103)	agtttttttttaagttttggaaagaagac	6.43670	118 (-1103)	agtttttttttaagttttggaaagaagac	-17.45140	118 (-1103)	agtttttttttaagttttggaaagaagac	6.12680
----------------	-------------------------------	-----------	----------------	-------------------------------	---------	----------------	-------------------------------	-----------	----------------	-------------------------------	---------