Ejemplo comparación de resultados predictores in sillico

Cambio de estudio CALD1 c.1171-428G>C (chr 7:134957641 G/C, COSV62652451o NM_033138.4: c.1171-428G>C)

Exón 9 e intrones adyacentes:

```
gagtgcagtggcacaatcttggctcactgcaatctctccctcccaggttcaagcaattct
cccacctcagccttccaagtagctggggctacaggcgcgtgtcaccacaccagctactt
tttgtatttttagtagagatagggtttcaccatgttggccaggctggtcttgaactcct
gacctcaagtgatttgcctgcctcggcctctcaaagtgctaggattataggtgtgagcct
cagcgctcagccaattatgttagtcttaaactttaattattttgccactgatttatattg
ttttagaggcattttattccatttctgtctggtctttattttgccttgaaacttgatt
tttttttttttacacgttttctctatcctgtatcaatcctaccttcattcccccaaatct
cactatcaaataccttatatgtagttttgctttctagcagagctacgcacagtaactcaa
atacccttatatgctcttcggtgtgtttaacaggttcaggatgatattggcctgggctg
agaaacactgtttataagcttgaaatattaattgggtttatttctttgtaattcctag
ATAGAAGAGCGAGCAGAATTTTTGAATAAGTCTGTGCAGAAAAG
gtaaatatgcttgatggttcaattgagctaatcagctagcatatgtatgagaagattctc
tctcatattttatatgtatgtgtttacttttttag
```

El cambio se encuentra en la tercera línea del intrón 8 (la **g** 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'
          645           2     +        0.63          TGCAGAAAAG^GTAAATATGC

Donor splice sites, complement strand
-------------------------------------
  pos 3'->5'  pos 5'->3'  phase strand  confidence  5'      exon intron    3'
    402         339         1     -        0.47          GGGAATGAAG^GTAGGATTGA

Acceptor splice sites, direct strand
------------------------------------
          pos 5'->3'  phase strand  confidence  5'      intron exon    3'
          460           0     +        0.33          TTTCTAGCAG^AGCTACGCAC
          600           0     +        0.33          TAATTCCTAG^ATAGAAGAGC

Acceptor splice sites, complement strand
----------------------------------------
  pos 3'->5'  pos 5'->3'  phase strand  confidence  5'      intron exon    3'
    632         109         1     -        0.33          TTCTGCACAG^ACTTATTCAA
    597         144         1     -        0.43          TTCTATCTAG^GAATTACAAA
    533         208         0     -        0.67          CTCAGCCCAG^GCCAATATCA
```

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.1.119.574818.0 :

Start	End	Score	Exon	Intron
40	54	0.40	ctcccag**gt**tcaagc	
252	266	0.43	caattat**gt**tagtct	
494	508	0.56	ctcttcg**gt**gtgttt	
638	652	1.00	agaaaag**gt**aaatat	

Donor site predictions for 10.42.2.148.574830.0 :

Start	End	Score	Exon	Intron
40	54	0.40	ctcccag**gt**tcaagc	
252	266	0.43	caattat**gt**tagtct	
494	508	0.56	ctcttcg**gt**gtgttt	
638	652	1.00	agaaaag**gt**aaatat	

Acceptor site predictions for 10.42.1.119.574818.0 :

Start	End	Score	Intron	Exon
26	66	0.95	actgcaatctctccctccc**ag**gttcaagcaattctcccacc	
50	90	0.54	caagcaattctcccacctc**ag**ccttccaagtagctggggct	
117	157	0.97	actttttgtattttttagt**ag**agatagggtttcaccatgtt	
286	326	0.67	cactgatttatattgtttt**ag**aggcatttttattccatttc	
440	480	0.93	tgtagttttgctttctagc**ag**agctacgcacagtaactcaa	
493	533	0.91	gctcttcggtgtgtttaac**ag**gttcagggatgatattggcc	
580	620	0.95	attttcttttgtaattcct**ag**atagaagagcgagcagaatt	

Acceptor site predictions for 10.42.2.148.574830.0 :

Start	End	Score	Intron	Exon
26	66	0.95	actgcaatctctccctccc**ag**gttcaagcaattctcccacc	
50	90	0.54	caagcaattctcccacctc**ag**ccttccaagtagctggggct	
117	157	0.97	actttttgtattttttagt**ag**agatagggtttcaccatgtt	
286	326	0.67	cactgatttatattgtttt**ag**aggcatttttattccatttc	
440	480	0.93	tgtagttttgctttctagc**ag**agctacgcacagtaactcaa	
493	533	0.91	gctcttcggtgtgtttaac**ag**gttcagggatgatattggcc	
580	620	0.95	attttcttttgtaattcct**ag**atagaagagcgagcagaatt	

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
gtctt(g/c)aactc	tgaact	tcaact	27144	56%

Human Splicing Finder

Alteration of auxiliary sequences	Significant alteration of ESE / ESS motifs ratio (6)	
Algorithm/Matix	**position**	**sequence**
ESE_SC35 (New ESE Site)	chr7:134957636	GTCTTCAA
RESCUE ESE (New ESE Site)	chr7:134957637	TCTTCA
IIE (ESS Site Broken)	chr7:134957637	TCTTGA
Sironi_motif3 (New ESS Site)	chr7:134957637	TCTTCAAC
IIE (ESS Site Broken)	chr7:134957638	CTTGAA
ESE_SRp40 (New ESE Site)	chr7:134957638	CTTCAAC
IIE (ESS Site Broken)	chr7:134957639	TTGAAC
EIE (New ESE Site)	chr7:134957640	TCAACT

SVM-BPfinder

```
seq_id  agez    ss_dist bp_seq  bp_scr  y_cont  ppt_off ppt_len ppt_scr svm_scr
wt      12      172     gtgtcacca       -0.00518096213453       0.562874251497  11      18      37      0.22994589
wt      12      143     tttttagta       -4.6484671846   0.528985507246  43      11      17      -3.8109283
wt      12      126     gtttcacca       0.19210677241   0.553719008264  26      11      17      -0.83155524
wt      12      102     tcttgaact       -0.428095221544 0.556701030928  2       11      17      0.44572919
wt      12      94      tcctgacct       3.57645241859   0.550561797753  11      20      31      1.572453
wt      12      89      acctcaagt       -0.863583404532 0.535714285714  6       20      31      0.14566543
wt      12      84      aagtgattt       -0.947514897218 0.556962025316  1       20      31      0.43615628
wt      12      63      ctctcaaag       0.0721748585815 0.48275862069   58      0       0       -3.085348
wt      12      49      ggattatag       -2.97019189301  0.522727272727  44      0       0       -3.3774932
wt      12      40      gtgtgagcc       -0.163572380789 0.571428571429  35      0       0       -1.693155
wt      12      34      gcctcagcg       1.12975143618   0.551724137931  29      0       0       -0.81333163
wt      12      27      cgctcagcc       1.83106548125   0.545454545455  22      0       0       -0.097670656
wt      12      19      caattatgt       -0.835874382556 0.571428571429  14      0       0       -0.62712826
mut     12      172     gtgtcacca       -0.00518096213453       0.568862275449  11      18      37      0.23187984
mut     12      143     tttttagta       -4.6484671846   0.536231884058  36      18      26      -3.2816571
mut     12      126     gtttcacca       0.19210677241   0.561983471074  19      18      26      -0.3019552
mut     12      102     tcttcaact       -1.1273512789   0.556701030928  2       11      17      0.17193739
mut     12      94      tcctgacct       3.57645241859   0.550561797753  11      20      31      1.572453
mut     12      89      acctcaagt       -0.863583404532 0.535714285714  6       20      31      0.14566543
mut     12      84      aagtgattt       -0.947514897218 0.556962025316  1       20      31      0.43615628
mut     12      63      ctctcaaag       0.0721748585815 0.48275862069   58      0       0       -3.085348
mut     12      49      ggattatag       -2.97019189301  0.522727272727  44      0       0       -3.3774932
mut     12      40      gtgtgagcc       -0.163572380789 0.571428571429  35      0       0       -1.693155
mut     12      34      gcctcagcg       1.12975143618   0.551724137931  29      0       0       -0.81333163
mut     12      27      cgctcagcc       1.83106548125   0.545454545455  22      0       0       -0.097670656
mut     12      19      caattatgt       -0.835874382556 0.571428571429  14      0       0       -0.62712826
```

El BP pasa a ser más débil en la secuencia mutante, por lo que podría estar afectando al *splicing*.

Variant Effect Predictor tool

ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000361675.7	protein_coding	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000361901.6	protein_coding	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000393118.6	protein_coding	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000417172.5	protein_coding	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000422748.5	protein_coding	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000424922.5	protein_coding	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant, NMD_transcript_variant	CALD1	ENSG00000122786	Transcript	ENST00000430085.5	nonsense_mediated_decay	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000436461.6	protein_coding	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant, NMD_transcript_variant	CALD1	ENSG00000122786	Transcript	ENST00000443197.5	nonsense_mediated_decay	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000466704.1	protein_coding	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	upstream_gene_variant	CALD1	ENSG00000122786	Transcript	ENST00000473714.1	retained_intron	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant, non_coding_transcript_variant	CALD1	ENSG00000122786	Transcript	ENST00000482470.5	retained_intron	-	-	rs931855774, COSV62652451
ENST00000361901.6:c.1171-428G>C	7:134957641-134957641	C	intron_variant	CALD1	ENSG00000122786	Transcript	ENST00000495522.1	protein_coding	-	-	rs931855774, COSV62652451

ESEfinder

Solo se obtiene una predicción con puntuaciones positivas para las matrices 3'SS:

149 (-592) caccatgttggccaggctggtcttgaactc	-0.29140	149 (-592) caccatgttggccaggctggtcttgaactc	2.91850	149 (-592) caccatgttggccaggctggtcttgaactc	-0.67190	149 (-592) caccatgttggccaggctggtcttgaactc	2.37610

Estas se ven muy ligeramente reducidas en las predicciones equivalentes para la secuencia mutante, por lo que no se tendrán en cuenta.

149 (-592) caccatgttggccaggctggtcttcaactc	-0.24410	149 (-592) caccatgttggccaggctggtcttcaactc	2.73790	149 (-592) caccatgttggccaggctggtcttcaactc	-0.59870	149 (-592) caccatgttggccaggctggtcttcaactc	2.26100