Ejemplo comparación de resultados predictores in sillico

Cambio de estudio OPRM1 c.864+1405G>A (chr6:154092877 G/A, COSV57673208 o NM_000914.5: c.864+1405G>A)

Exón 3 e intrones adyacentes:

```
cttaagttagctctggtcaaggctaaaaatgaatgagcaaaatggcagtattaacacctt
atgacataattaaatgttgctgctaatttttcctttaaattcctttcttctag
GTTCCATAGATTGTACACTAACATTCTCTCATCCAACCTGGTACTGGGAAAACCTGCTGA
AGATCTGTGTTTCATCTTCGCCTTCATTATGCCAGTGCTCATCATTACCGTGTGCTATG
GACTGATGATCTTGCGCCTCAAGAGTGTCCGCATGCTCTCTGGCTCCAAAGAAAAGGACA
GGAATCTTCGAAGGATCACCAGGATGGTGCTGGTGGTGGTGGCTGTGTTCATCGTCTGCT
GGACTCCCATTCACATTTACGTCATCATTAAAGCCTTGGTTACAATCCCAGAAACTACGT
TCCAGACTGTTTCTTGGCACTTCTGCATTGCTCTAGGTTACACAAACAGCTGCCTCAACC
CAGTCCTTTATGCATTTCTGGATGAAAACTTCAAACGATGCTTCAGAGAGTTCTGTATCC
CAACCTCTTCCAACATTGAGCAACAAAACTCCACTCGAATTCGTCAGAACACTAGAGACC
ACCCCTCCACGGCCAATACAGTGGATAGAACTAATCATCAG
gtacgcagtctctagaattaggtatatctactggggatgacataaaaattataaggcttt
gtgctaaactaggagtttaatccattatagaggatgagaatggagggaagaggggaagca
aattgtggttctagtgttagagaagaggtttgtttataaaactgtgttctttatatttga
ctgtacatattcatttaggtataaagatacaccaatgagaaatccatgaaactattcaaa
ataactattttatggcctttacttctatgcaaaatttatgactttagcacattatagaa
ataattctgatctagaatccttttcattttccccagaattattatataattcatagatgt
tgctgcaataccctcttatttctcaaaagccagtcttgtctggttctgtgattaaaga
gagagggtgagtgccttgcccactgtggtcatggatgcaagatattcacagaaaattagc
atcatggaaaagagagaaaagattaaaaaatgaccatccctaacttttttcttaagctattgg
tttgctacctgaactctagcacaaatatcaaacatattagactgaataatatattattaa
tatgtgaattattaatctaataaatttttattagattaaacaattttttaacagacctcatgc
ttgttggagataatagggtgaatagcgactagctttttggaaggtgtatttctgtgcccag
acctacacagaaaattcaagacttatacacactattataatcatatgggaaatcaacaag
gaggaaagagaaaaggctgtctccaagttagtattacttgattaagccactggttggcta
aatttgacctctgatctataaaatcaatcaaatacatggcttaatttttaatattaatct
aagtaagtaaattttaaataatgatgcagcatccaaaaatcacagcctgtataattttt
ttaacagtgaattcaaattcaaacttcctatgaagtaaaggctggagtcagaatcttggc
taaagttcctagagtcaggcgtcaggactgtgaggacagatggctccggagaaatgaata
gcaagtcaaatgattgaggagctaggggagccctttgccaccacagggtgcagtggcttc
ctcctgctgtctcttgaccacaccctgtctctcactgtcccctctttaccccactctgcc
atccagcccagcccaggctcaacctttctaattgaccagaggactaatgtaatatggccc
agatacatccgtcatgggcgacattaatcctcatacccctggtattcagtgatgctcaag
ggaacacaggaactcacacccctagagagcatgaggtaatgatggcagacccacttcca
gacaaagggaaattaacttattatgtctcagttctaattccatttcagaaatgaatgcca
gaaagaaatggttttgactctgctcttttctctacgttttttttccacatcaagactgct
```

El cambio se encuentra en la penúltima línea del intrón 3 (la **g** en color granate).

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'
           635        2     +        0.83        TAATCATCAG^GTACGCAGTC
          1061        2     +        0.70        AAGAGAGAGG^GTGAGTGCCT
          1537        1     +        0.46        ATTAATCTAA^GTAAGTAAAT
          1991        0     +        0.44        AGAGCATGAG^GTAATGATGG

Donor splice sites, complement strand
-------------------------------------
  pos 3'->5'  pos 5'->3'  phase strand  confidence  5'      exon intron      3'
     222        1913        0     -        0.71        ATAGCACACG^GTAATGATGA
     150        1985        1     -        0.70        CCAGTACCAG^GTTGGATGAG
     135        2000        0     -        0.61        ATGAGAGAAT^GTTAGTGTAC

Acceptor splice sites, direct strand
------------------------------------
        pos 5'->3'  phase strand  confidence  5'      intron exon      3'
           113        2     +        0.94        TTTCTTCTAG^GTTCCATAGA
           122        2     +        0.18        GGTTCCATAG^ATTGTACACT
           175        0     +        0.19        CCTGCTGAAG^ATCTGTGTTT
           449        1     +        0.33        ATTGCTCTAG^GTTACACAAA
           462        2     +        0.18        ACACAAACAG^CTGCCTCAAC
           476        1     +        0.17        CTCAACCCAG^TCCTTTATGC
           970        2     +        0.33        TTTTCCCCAG^AATTATTATA
          1760        1     +        0.28        GCCACCACAG^GGTGCAGTGG
          2062        0     +        0.00        TCCATTTCAG^AAATGAATGC

Acceptor splice sites, complement strand
----------------------------------------
  pos 3'->5'  pos 5'->3'  phase strand  confidence  5'      intron exon      3'
    1977        158        1     -        0.17        TGCTCTCTAG^GGGGTGTGAG
    1357        778        0     -        0.43        TTCTGTGTAG^GTCTGGGCAC
    1042       1093        1     -        0.56        TTAATCACAG^AACCAGAGCA
     662       1473        0     -        0.33        TCCCCAGTAG^ATATACCTAA
     502       1633        1     -        0.07        TCGTTTGAAG^TTTTCATCCA
     491       1644        0     -        0.14        TTTCATCCAG^AAATGCATAA
     357       1778        1     -        0.07        TGAATGGGAG^TCCAGCAGAC
     352       1783        0     -        0.17        GGGAGTCCAG^CAGACGATGA
     349       1786        0     -        0.17        AGTCCAGCAG^ACGATGAACA
      23       2112        -     -        0.00        TCATTTTTAG^CCTTGACCAG
```

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.0.139.574098.0 :

Start	End	Score	Exon Intron
628	642	1.00	tcatcag**gt**acgcag
649	663	0.85	gaattag**gt**atatct
775	789	0.70	agaagag**gt**ttgtta
826	840	0.72	catttag**gt**ataaag
1054	1068	0.99	agagagg**gt**gagtgc
1305	1319	0.73	taatagg**gt**gaatag
1434	1448	0.94	tctccaa**gt**tagtat
1530	1544	0.98	aatctaa**gt**aagtaa
1984	1998	0.83	gcatgag**gt**aatgat

Donor site predictions for 10.42.3.123.574086.0 :

Start	End	Score	Exon Intron
628	642	1.00	tcatcag**gt**acgcag
649	663	0.85	gaattag**gt**atatct
775	789	0.70	agaagag**gt**ttgtta
826	840	0.72	catttag**gt**ataaag
1054	1068	0.99	agagagg**gt**gagtgc
1305	1319	0.73	taatagg**gt**gaatag
1434	1448	0.94	tctccaa**gt**tagtat
1530	1544	0.98	aatctaa**gt**aagtaa
1984	1998	0.83	gcatgag**gt**aatgat

Acceptor site predictions for 10.42.0.139.574098.0 :

Start	End	Score	Intron Exon
93	133	1.00	ctttaaattcctttcttct**ag**gttccatagattgtacacta
102	142	0.92	cctttcttctaggttccat**ag**attgtacactaacattctct
429	469	0.79	ggcacttctgcattgctct**ag**gttacacaaacagctgcctc
950	990	0.96	aatcctttcattttcccc**ag**aattattatataattcatag
1147	1187	0.57	ccatccctaacttttctta**ag**ctattggtttgctacctgaa
1334	1374	0.40	aaggtgtatttctgtgccc**ag**acctacacagaaaattcaag
1581	1621	0.42	cctgtataatttttttaac**ag**tgaattcaaattcaaacttc
2025	2065	0.67	aattaacttattatgtctc**ag**ttctaattccatttcagaaa
2042	2082	0.69	tcagttctaattccatttc**ag**aaatgaatgccagaaagaaa

Acceptor site predictions for 10.42.3.123.574086.0 :

Start	End	Score	Intron Exon
93	133	1.00	ctttaaattcctttcttct**ag**gttccatagattgtacacta
102	142	0.92	cctttcttctaggttccat**ag**attgtacactaacattctct
429	469	0.79	ggcacttctgcattgctct**ag**gttacacaaacagctgcctc
950	990	0.96	aatcctttcattttcccc**ag**aattattatataattcatag
1147	1187	0.57	ccatccctaacttttctta**ag**ctattggtttgctacctgaa
1334	1374	0.40	aaggtgtatttctgtgccc**ag**acctacacagaaaattcaag
1581	1621	0.42	cctgtataatttttttaac**ag**tgaattcaaattcaaacttc
2025	2065	0.62	aattaacttattatatctc**ag**ttctaattccatttcagaaa
2042	2082	0.69	tcagttctaattccatttc**ag**aaatgaatgccagaaagaaa

Cambia uno de los sitios *acceptor* de una secuencia a otra por la presencia del cambio de estudio. Como está afectando a un *acceptor* que no participa en el *splicing* normal del transcrito, podemos concluir que la mutación no afecta al *splicing*.

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
attat(g/a)tctca	tgtctc	tatctc	29360	70%

Human Splicing Finder

Alteration of auxiliary sequences	Significant alteration of ESE / ESS motifs ratio (2)	
Algorithm/Matix	**position**	**sequence**
PESS (ESS Site Broken)	chr6:154092870	TTATTATG
PESS (New ESS Site)	chr6:154092870	TTATTATA
PESS (ESS Site Broken)	chr6:154092871	TATTATGT
PESS (New ESS Site)	chr6:154092871	TATTATAT
EIE (ESE Site Broken)	chr6:154092872	ATTATG
EIE (ESE Site Broken)	chr6:154092873	TTATGT
PESS (ESS Site Broken)	chr6:154092873	TTATGTCT
PESS (New ESS Site)	chr6:154092873	TTATATCT
ESS_hnRNPA1 (ESS Site Broken)	chr6:154092874	TATGTC
PESS (ESS Site Broken)	chr6:154092874	TATGTCTC
IIE (ESS Site Broken)	chr6:154092875	ATGTCT
IIE (ESS Site Broken)	chr6:154092876	TGTCTC

SVM-BPfinder

seq_id	agez	ss_dist	bp_seq	bp_scr	y_cont	ppt_off	ppt_len	ppt_scr	svm_scr
wt	55	106	aaattaact	-2.17611658923	0.60396039604	9	23	33	-0.51747518
wt	55	105	aattaactt	0.777076092455	0.61	8	23	33	0.70408831
wt	55	101	aacttatta	-2.49649775018	0.604166666667	4	23	33	-0.32636122
wt	55	98	ttattatgt	-1.91074043991	0.602150537634	1	23	33	0.092234196
wt	55	91	gtctcagtt	0.00264363490964	0.593023255814	2	15	26	0.70995743
wt	55	84	ttctaattc	1.64504957275	0.594936708861	1	9	19	1.3517424
wt	55	74	atttcagaa	-3.1655658429	0.565217391304	25	37	71	-1.5761751
wt	55	67	aaatgaatg	-0.911494358367	0.612903225806	18	37	71	-0.23511028
wt	55	44	ttttgactc	1.22016232967	0.769230769231	1	31	63	1.651568
mut	55	106	aaattaact	-2.17611658923	0.60396039604	9	23	33	-0.51747518
mut	55	105	aattaactt	0.777076092455	0.61	8	23	33	0.70408831
mut	55	101	aacttatta	-2.49649775018	0.604166666667	4	23	33	-0.32636122
mut	55	98	ttattatat	-2.90501267985	0.602150537634	1	23	33	-0.29707033
mut	55	91	atctcagtt	-0.417293417995	0.593023255814	2	15	26	0.54553222
mut	55	84	ttctaattc	1.64504957275	0.594936708861	1	9	19	1.3517424
mut	55	74	atttcagaa	-3.1655658429	0.565217391304	25	37	71	-1.5761751
mut	55	67	aaatgaatg	-0.911494358367	0.612903225806	18	37	71	-0.23511028
mut	55	44	ttttgactc	1.22016232967	0.769230769231	1	31	63	1.651568

Cambian dos de los BP en la secuencia WT. El primero pasa de tener puntuación positiva a negativa por la presencia de la mutación, mientras que el otro reduce su puntuación cuando tiene la mutación. Podría estar perdiéndose un BP.

Variant Effect Predictor tool

ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000229768.9	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000330432.12	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000337049.8	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	downstream_gene_variant	OPRM1	ENSG00000112038	Transcript	ENST00000360422.8	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000414028.6	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000419506.6	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	downstream_gene_variant	OPRM1	ENSG00000112038	Transcript	ENST00000428397.6	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000434900.6	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000435918.6	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000452687.6	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000518759.5	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant, NMD_transcript_variant	OPRM1	ENSG00000112038	Transcript	ENST00000519083.5	nonsense_mediated_decay	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant, non_coding_transcript_variant	OPRM1	ENSG00000112038	Transcript	ENST00000519613.5	processed_transcript	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	downstream_gene_variant	OPRM1	ENSG00000112038	Transcript	ENST00000520282.5	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000520708.5	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	downstream_gene_variant	OPRM1	ENSG00000112038	Transcript	ENST00000521106.1	processed_transcript	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000522236.1	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	downstream_gene_variant	OPRM1	ENSG00000112038	Transcript	ENST00000522382.1	processed_transcript	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000522555.5	protein_coding	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant, NMD_transcript_variant	OPRM1	ENSG00000112038	Transcript	ENST00000522739.5	nonsense_mediated_decay	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant, NMD_transcript_variant	OPRM1	ENSG00000112038	Transcript	ENST00000524150.2	nonsense_mediated_decay	-	-	COSV57673208
ENST00000520708.5:c.864+1405G>A	6:154092877-154092877	A	intron_variant	OPRM1	ENSG00000112038	Transcript	ENST00000524163.5	protein_coding	-	-	COSV57673208

ESEfinder

La única predicción para la secuencia WT que tiene puntuación positiva es en las matrices 3', pero, cundo se comparan las puntuaciones con las de la secuencia mutante, no se ven prácticamente alteradas, por lo que no se tendrán en cuenta.

| 2031 (-104) | cttattatgtctcagttctaattccatttc | -17.40670 | 2031 (-104) | cttattatgtctcagttctaattccatttc | 4.54840 | 2031 (-104) | cttattatgtctcagttctaattccatttc | -18.42890 | 2031 (-104) | cttattatgtctcagttctaattccatttc | 4.28690 |
| 2031 (-104) | cttattatatctcagttctaattccatttc | -17.65010 | 2031 (-104) | cttattatatctcagttctaattccatttc | 4.54900 | 2031 (-104) | cttattatatctcagttctaattccatttc | -18.66620 | 2031 (-104) | cttattatatctcagttctaattccatttc | 4.18720 |