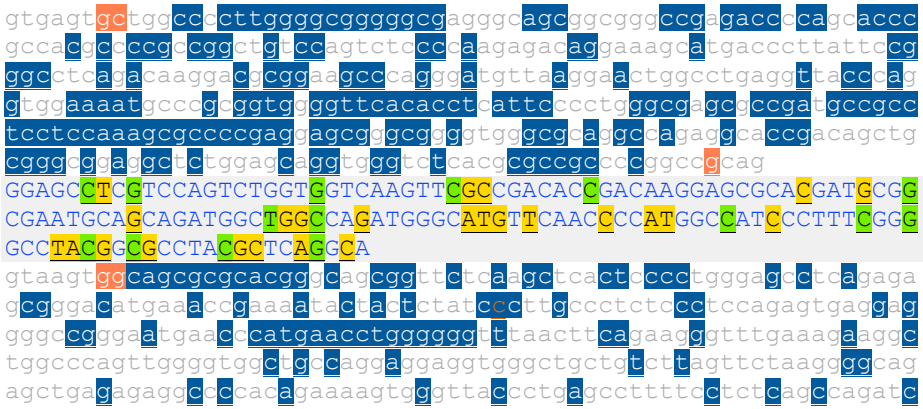
**Ejemplo comparación de resultados predictores in sillico**

Cambio de estudio CELF4 c.801+93C>A (chr18:37274218 C/A, COSV58463837 o NM_020180.4: c.801+93C>A)

Exón 6 e intrones adyacentes:



El cambio se encuentra en la segunda línea del intrón 6 (la **c** en color naranja subrayada de 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' |
|---|---|---|---|---|
| 492 | 0 | + | 0.06 | CTACGCTCAG^GCAGTAAGTG |
| 495 | 0 | + | 0.95 | CGCTCAGGCA^GTAAGTGGCA H |

Donor splice sites, complement strand

| pos 3'->5' | pos 5'->3' | phase | strand | confidence | 5' exon intron 3' |
|---|---|---|---|---|---|
| 209 | 586 | 1 | - | 0.65 | GGGGAATGAG^GTGTGAACCC |

Acceptor splice sites, direct strand

| pos 5'->3' | phase | strand | confidence | 5' intron exon 3' |
|---|---|---|---|---|
| 350 | 0 | + | 0.91 | CCGGCCGCAG^GGAGCCTCGT |
| 354 | 1 | + | 0.42 | CCGCAGGGAG^CCTCGTCCAG |
| 364 | 2 | + | 0.55 | CCTCGTCCAG^TCTGGTGGTC |
| 377 | 0 | + | 0.34 | GGTGGTCAAG^TTCGCCGACA |
| 395 | 0 | + | 0.07 | CACCGACAAG^GAGCGCACGA |
| 605 | 1 | + | 0.56 | CTCCCTCCAG^AGTGAGGAGG |

Acceptor splice sites, complement strand

| pos 3'->5' | pos 5'->3' | phase | strand | confidence | 5' intron exon 3' |
|---|---|---|---|---|---|
| 724 | 71 | 2 | - | 0.00 | TGCCCCTTAG^AACTAAGACA |
| 692 | 103 | 1 | - | 0.30 | CTCCTGGCAG^CCACCCCAAC |
| 539 | 256 | 0 | - | 0.16 | AGGCTCCCAG^GGGAGTGAGC |
| 463 | 332 | 1 | - | 0.07 | GCCCCGAAAG^GGATGGCCAT |
| 427 | 368 | 1 | - | 0.20 | ATCTGGCCAG^CCATCTGCTG |

Donor splice sites, direct strand

| pos 5'->3' | phase | strand | confidence | 5' exon intron 3' |
|---|---|---|---|---|
| 492 | 0 | + | 0.06 | CTACGCTCAG^GCAGTAAGTG |
| 495 | 0 | + | 0.95 | CGCTCAGGCA^GTAAGTGGCA H |

Donor splice sites, complement strand

| pos 3'->5' | pos 5'->3' | phase | strand | confidence | 5' exon intron 3' |
|---|---|---|---|---|---|
| 209 | 586 | 1 | - | 0.65 | GGGGAATGAG^GTGTGAACCC |

Acceptor splice sites, direct strand

| pos 5'->3' | phase | strand | confidence | 5' intron exon 3' |
|---|---|---|---|---|
| 350 | 0 | + | 0.91 | CCGGCCGCAG^GGAGCCTCGT |
| 354 | 1 | + | 0.42 | CCGCAGGGAG^CCTCGTCCAG |
| 364 | 2 | + | 0.55 | CCTCGTCCAG^TCTGGTGGTC |
| 377 | 0 | + | 0.34 | GGTGGTCAAG^TTCGCCGACA |
| 395 | 0 | + | 0.07 | CACCGACAAG^GAGCGCACGA |
| 605 | 1 | + | 0.53 | CTCCCTCCAG^AGTGAGGAGG |

Acceptor splice sites, complement strand

| pos 3'->5' | pos 5'->3' | phase | strand | confidence | 5' intron exon 3' |
|---|---|---|---|---|---|
| 724 | 71 | 2 | - | 0.00 | TGCCCCTTAG^AACTAAGACA |
| 692 | 103 | 1 | - | 0.29 | CTCCTGGCAG^CCACCCCAAC |
| 539 | 256 | 0 | - | 0.17 | AGGCTCCCAG^GGGAGTGAGC |
| 463 | 332 | 1 | - | 0.07 | GCCCCGAAAG^GGATGGCCAT |
| 427 | 368 | 1 | - | 0.20 | ATCTGGCCAG^CCATCTGCTG |

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

**Donor site predictions for 10.42.2.148.574650.0 :**

| Start | End | Score | Exon | Intron |
|---|---|---|---|---|
| 315 | 329 | 0.79 | ggagcag | gtgggtct |
| 488 | 502 | 0.97 | tcaggca | gtaagtgg |
| 600 | 614 | 0.43 | ctccaga | gtgaggag |

**Donor site predictions for 10.42.1.119.574638.0 :**

| Start | End | Score | Exon | Intron |
|---|---|---|---|---|
| 315 | 329 | 0.79 | ggagcag | gtgggtct |
| 488 | 502 | 0.97 | tcaggca | gtaagtgg |
| 600 | 614 | 0.43 | ctccaga | gtgaggag |

**Acceptor site predictions for 10.42.2.148.574650.0 :**

| Start | End | Score | Intron | Exon |
|---|---|---|---|---|
| 330 | 370 | 0.46 | cacgcgccgccccggccgc | aggagcctcgtccagtctggt |
| 585 | 625 | 0.92 | tcccttgccctctccctcc | agagtgaggaggggccgggaat |
| 708 | 748 | 0.56 | ggctgctgtcttagttcta | aggggcagagctgagagaggcc |

**Acceptor site predictions for 10.42.1.119.574638.0 :**

| Start | End | Score | Intron | Exon |
|---|---|---|---|---|
| 330 | 370 | 0.46 | cacgcgccgccccggccgc | aggagcctcgtccagtctggt |
| 585 | 625 | 0.91 | tcacttgccctctccctcc | agagtgaggaggggccgggaat |
| 708 | 748 | 0.56 | ggctgctgtcttagttcta | aggggcagagctgagagaggcc |

Cambia uno de los sitios *acceptor* (en azul) por la presencia de la mutación. Este no es el sitio *acceptor* que se utiliza de normal en el *splicing*, por lo que es muy probable que la mutación no afecte al proceso.

**Spliceman**

| Point mutation | Wildtype (wt) | Mutation (mt) | L1 distance | Ranking (L1) |
|---|---|---|---|---|
| ctatc(c/a)cttgc | ccttgc | acttgc | 26489 | 52% |

**Human Splicing Finder**

| New Donor splice site | Activation of a cryptic Donor site. Potential alteration of splicing |
|---|---|

| Algorithm/Matix | position | sequences | variation |
|---|---|---|---|
| HSF Donor site (matrix GT) | chr18:37274214 | - REF : CAAGGGATA<br>- ALT : CAAGTGATA | 47.56 > 74.7<br>=> 57.06% |

**SVM-BPfinder**

| seq_id | agez | ss_dist | bp_seq | bp_scr | y_cont | ppt_off | ppt_len | ppt_scr | svm_scr | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| wt | 45 | 50 | gagtgagga | | -1.97795127781 | 0.377777777778 | 45 | 0 | 0 | | -3.0990968 |
| wt | 45 | 33 | gaatgaacc | | -0.0241067721628 | 0.5 | 28 | 0 | 0 | | -1.2185287 |
| wt | 45 | 25 | ccatgaacc | | -0.0695934192418 | 0.5 | 20 | 0 | 0 | | -0.72995232 |
| mut | 45 | 73 | ctatcactt | | -1.34049984529 | 0.485294117647 | 1 | 16 | 33 | | 0.27776914 |
| mut | 45 | 50 | gagtgagga | | -1.97795127781 | 0.377777777778 | 45 | 0 | 0 | | -3.0990968 |
| mut | 45 | 33 | gaatgaacc | | -0.0241067721628 | 0.5 | 28 | 0 | 0 | | -1.2185287 |
| mut | 45 | 25 | ccatgaacc | | -0.0695934192418 | 0.5 | 20 | 0 | 0 | | -0.72995232 |

## Variant Effect Predictor tool

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000334919.9 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000361795.9 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000420428.7 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | upstream_gene_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000586009.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | upstream_gene_variant | AC090386.1 | ENSG00000267202 | Transcript | ENST00000586610.1 | lncRNA | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | upstream_gene_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000587074.1 | processed_transcript | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | upstream_gene_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000587657.1 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000587819.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | downstream_gene_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000587924.1 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | upstream_gene_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000588591.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000588597.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant, non_coding_transcript_variant | AC090386.1 | ENSG00000267202 | Transcript | ENST00000588766.5 | lncRNA | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000589229.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | upstream_gene_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000589386.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | 3_prime_UTR_variant, NMD_transcript_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000590112.5 | nonsense_mediated_decay | 4/10 | 894 | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000591282.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000591287.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | upstream_gene_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000591421.5 | retained_intron | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | upstream_gene_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000593271.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000601019.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000601392.5 | protein_coding | - | - | COSV58463837 |
| ENST00000591282.5:c.801+93C>A | 18:37274218-37274218 | T | intron_variant | CELF4 | ENSG00000101489 | Transcript | ENST00000603232.6 | protein_coding | - | - | COSV58463837 |

## ESEfinder

No se obtiene ningún resultado positive con la posición de interés ni en la secuencia mutante ni en la WT.