

Ejemplo comparación de resultados predictores in silico

Cambio de estudio PGF c.150C>T (chr14:74949519 C/T o NM_001207012.1: c.150C>T)

Exón 3 e intrones adyacentes:

```
cagatgagctaaggggcccgcaggcatttcggaggggattcacattaggagttcagagttga
gggagctcggatggggctgggggctgaacctggaaggccagcaggggccttgggagccagg
tttagggcttcaaggggcagggtcctgtgtggcctgacccagcccccttgcttccttcag
TGGTACCGTTCCAGGAAGTGTGGGC CGCAGCTACTGCCGGGCGCTGGAGAGGCTGGTGG
ACGT CGTGTCCGAGTACCCAGCGAGGTGGAGCACATGTTCAGCCATCCTGTGTCTCCC
TGCTGCGCTGACCGGCTGCTGCGCGATGAGAATCTGCAC TGTGTGCGGTGGAGACGG
CCAAATGTCACCA TGCAG
gtaggtccataccctgcccagggggccacccaatctgcccactagaagggtacctggagag
ggaagaagataggcccagggtcccagggtggagcctgactccaggccagtgctcctacc
acccagcccagggtcctcttcttctaagccctccgatggctctgagccactcatctcc
```

El cambio se encuentra en la primera fila del exón 7 (la c 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' |
|-----|-------|-------|--------|------------|------------|-------------|--------|----|
| 378 | | 0 | + | 0.99 | CACCATGCAG | ^GTAGGTCCAT | H | |
| 426 | | 0 | + | 0.93 | CCACTAGAAG | ^GTACCTGGAG | H | |

Donor splice sites, complement strand

No donor site predictions above threshold.

Acceptor splice sites, direct strand

| pos | 5'→3' | phase | strand | confidence | 5' | intron | exon | 3' |
|-----|-------|-------|--------|------------|------------|-------------|------|----|
| 125 | | 0 | + | 0.07 | CCAGGTTTAG | ^GGCTTCAAGG | | |
| 134 | | 0 | + | 0.18 | GGGCTTCAAG | ^GGGCAGGTCT | | |
| 140 | | 0 | + | 0.19 | CAAGGGGCAG | ^GTCTCTGCTG | | |
| 161 | | 0 | + | 0.44 | CCTGACCCAG | ^CCCCCTTGCT | | |
| 180 | | 1 | + | 1.00 | TTCCTCGCAG | ^TGGTACCCTT | H | |
| 194 | | 0 | + | 0.26 | ACCCTTCCAG | ^GAAGTGTGGG | | |
| 198 | | 1 | + | 0.07 | TTCCAGGAAG | ^TGTGGGGCCG | | |

Acceptor splice sites, complement strand

| pos 3'→5' | pos 5'→3' | phase | strand | confidence | 5' | intron | exon | 3' |
|-----------|-----------|-------|--------|------------|------------|-------------|------|----|
| 430 | 128 | 2 | - | 0.97 | CCCTCTCCAG | ^GTACCTTCTA | | |
| 419 | 139 | 1 | - | 0.42 | TACCTTCTAG | ^TGGGCAGATT | | |
| 412 | 146 | 2 | - | 0.19 | TAGTGGGCAG | ^ATTCGGTGGC | | |
| 391 | 167 | 2 | - | 0.30 | CCCTGGGCAG | ^GGTATGGACC | | |
| 320 | 238 | 0 | - | 0.15 | ATCGCCGCAG | ^CAGCCGGTGC | | |

Donor splice sites, direct strand

| pos | 5'→3' | phase | strand | confidence | 5' | exon | intron | 3' |
|-----|-------|-------|--------|------------|------------|-------------|--------|----|
| 378 | | 0 | + | 0.99 | CACCATGCAG | ^GTAGGTCCAT | H | |
| 426 | | 0 | + | 0.93 | CCACTAGAAG | ^GTACCTGGAG | H | |

Donor splice sites, complement strand

No donor site predictions above threshold.

Acceptor splice sites, direct strand

| pos | 5'→3' | phase | strand | confidence | 5' | intron | exon | 3' |
|-----|-------|-------|--------|------------|------------|-------------|------|----|
| 134 | | 0 | + | 0.17 | GGGCTTCAAG | ^GGGCAGGTCT | | |
| 140 | | 0 | + | 0.19 | CAAGGGGCAG | ^GTCTCTGCTG | | |
| 161 | | 0 | + | 0.43 | CCTGACCCAG | ^CCCCCTTGCT | | |
| 180 | | 1 | + | 1.00 | TTCCTCGCAG | ^TGGTACCCTT | H | |
| 194 | | 0 | + | 0.32 | ACCCTTCCAG | ^GAAGTGTGGG | | |
| 198 | | 1 | + | 0.18 | TTCCAGGAAG | ^TGTGGGGCCG | | |

Acceptor splice sites, complement strand

| pos 3'→5' | pos 5'→3' | phase | strand | confidence | 5' | intron | exon | 3' |
|-----------|-----------|-------|--------|------------|------------|-------------|------|----|
| 430 | 128 | 2 | - | 0.97 | CCCTCTCCAG | ^GTACCTTCTA | | |
| 419 | 139 | 1 | - | 0.42 | TACCTTCTAG | ^TGGGCAGATT | | |
| 412 | 146 | 2 | - | 0.19 | TAGTGGGCAG | ^ATTCGGTGGC | | |
| 391 | 167 | 2 | - | 0.30 | CCCTGGGCAG | ^GGTATGGACC | | |

Se pierde uno de los sitios *acceptor* (en rojo) en la secuencia mutante. Este se encuentra por delante del exón, por lo que no participa en el *splicing* normal y perderlo no tendrá importancia. El resto de los sitios que se detectan son los mismos, aunque cambian las confianzas de cada uno de los sitios. La predicción que coincide con el *acceptor* del exón es la que tiene confianza 1 y es la única que se mantiene intacta en la secuencia mutante, por lo que es probable que no haya cambios en el *splicing*.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for 10.42.3.123.573095.0 :

| Start | End | Score | Exon | Intron |
|-------|-----|-------|---------|------------------|
| 371 | 385 | 0.98 | catgcag | g taggtcc |
| 419 | 433 | 0.87 | ctagaag | g tacctgg |

Donor site predictions for 10.42.0.139.573107.0 :

| Start | End | Score | Exon | Intron |
|-------|-----|-------|---------|------------------|
| 371 | 385 | 0.98 | catgcag | g taggtcc |
| 419 | 433 | 0.87 | ctagaag | g tacctgg |

Acceptor site predictions for 10.42.3.123.573095.0 :

| Start | End | Score | Intron | Exon |
|-------|-----|-------|---------------------|--------------------------------|
| 160 | 200 | 0.96 | agcccccttgcttcctcgc | a gtgggtacccttcaggaagtg |
| 174 | 214 | 0.67 | ctcgcagtggtacccttcc | a ggaagtgtggggccgcagcta |
| 506 | 546 | 0.56 | ccagggcctcttcttccta | a gccctccgatggctctgagcc |

Acceptor site predictions for 10.42.0.139.573107.0 :

| Start | End | Score | Intron | Exon |
|-------|-----|-------|---------------------|--------------------------------|
| 160 | 200 | 0.96 | agcccccttgcttcctcgc | a gtgggtacccttcaggaagtg |
| 174 | 214 | 0.67 | ctcgcagtggtacccttcc | a ggaagtgtggggccgcagcta |
| 506 | 546 | 0.56 | ccagggcctcttcttccta | a gccctccgatggctctgagcc |


Spliceman

| Point mutation | Wildtype (wt) | Mutation (mt) | L1 distance | Ranking (L1) |
|-----------------|---------------|---------------|-------------|--------------|
| agcta(c/t)tgccg | ctgccg | ttgccg | 26235 | 50% |

CRYP-SKIP

Parece que hay un sitio críptico de *splicing* dentro del propio exón, pero el cambio de interés (la primera **a** en minúsculas detrás de las mayúsculas, que indican el exón) no lo toma en consideración, por lo que no debe considerar que tenga algún efecto en el *splicing*.

Human Splicing Finder

|  Alteration of auxiliary sequences | | Significant alteration of ESE / ESS motifs ratio (-4) |
|---|----------------|---|
| Algorithm/Matix | position | sequence |
| ESE_ASF (ESE Site Broken) | chr14:74949519 | CTGCCGG |
| ESE_ASFB (ESE Site Broken) | chr14:74949519 | CTGCCGG |
| EIE (New ESE Site) | chr14:74949520 | ATTGCC |
| Sironi_motif3 (ESS Site Broken) | chr14:74949520 | ACTGCCGG |
| IIE (New ESS Site) | chr14:74949521 | TATTGC |
| IIE (New ESS Site) | chr14:74949522 | CTATTG |
| ESE_SRp40 (ESE Site Broken) | chr14:74949522 | CTACTGC |
| ESE_SC35 (ESE Site Broken) | chr14:74949524 | AGCTACTG |

SVM-BPfinder

seq_id agez ss_dist bp_seq bp_scr y_cont ppt_off ppt_len ppt_scr svm_scr

Variant Effect Predictor tool

| | | | | | | | | | | | | | | | |
|-----------------------------|--------------------------------------|---|------------------------------------|-----|---------------------------------|------------|------------------------------------|-----------------|-----|-----|-----|----|---|---------|---|
| ENST00000238607.10:c.150C>T | 14:74949519-74949519 | A | synonymous_variant | PGF | ENSG00000119630 | Transcript | ENST00000238607.10 | protein_coding | 3/7 | 485 | 150 | 50 | Y | TAC/TAT | - |
| ENST00000238607.10:c.150C>T | 14:74949519-74949519 | A | synonymous_variant | PGF | ENSG00000119630 | Transcript | ENST00000405431.2 | protein_coding | 3/7 | 153 | 153 | 51 | Y | TAC/TAT | - |
| ENST00000238607.10:c.150C>T | 14:74949519-74949519 | A | synonymous_variant | PGF | ENSG00000119630 | Transcript | ENST00000553716.5 | protein_coding | 3/6 | 537 | 153 | 51 | Y | TAC/TAT | - |
| ENST00000238607.10:c.150C>T | 14:74949519-74949519 | A | non_coding_transcript_exon_variant | PGF | ENSG00000119630 | Transcript | ENST00000555234.1 | retained_intron | 2/3 | 465 | - | - | - | - | - |
| ENST00000238607.10:c.150C>T | 14:74949519-74949519 | A | upstream_gene_variant | PGF | ENSG00000119630 | Transcript | ENST00000555253.1 | retained_intron | - | - | - | - | - | - | - |
| ENST00000238607.10:c.150C>T | 14:74949519-74949519 | A | synonymous_variant | PGF | ENSG00000119630 | Transcript | ENST00000555567.6 | protein_coding | 3/7 | 508 | 153 | 51 | Y | TAC/TAT | - |
| ENST00000238607.10:c.150C>T | 14:74949519-74949519 | A | non_coding_transcript_exon_variant | PGF | ENSG00000119630 | Transcript | ENST00000556939.5 | retained_intron | 3/4 | 508 | - | - | - | - | - |
| ENST00000238607.10:c.150C>T | 14:74949519-74949519 | A | non_coding_transcript_exon_variant | PGF | ENSG00000119630 | Transcript | ENST00000557748.5 | retained_intron | 2/5 | 366 | - | - | - | - | - |

ESEfinder

Solo se obtiene una predicción con puntuación positiva para 3'. Si se comparan las puntuaciones con las de la secuencia mutante, disminuyen muy poco, por lo que podría estar debilitando un sitio *acceptor*:

| | | | | | | | | | | | |
|---------------|---------------------------------|----------|---------------|---------------------------------|---------|---------------|---------------------------------|----------|---------------|---------------------------------|---------|
| 197 (-361) | AGTGTGGGGCCGCGAGCTACTGCCGGGCGCT | -4.05130 | 197 (-361) | AGTGTGGGGCCGCGAGCTACTGCCGGGCGCT | 2.89210 | 197 (-361) | AGTGTGGGGCCGCGAGCTACTGCCGGGCGCT | -4.05610 | 197 (-361) | AGTGTGGGGCCGCGAGCTACTGCCGGGCGCT | 2.77820 |
| 197 (-361) | AGTGTGGGGCCGCGAGCTATTGCCGGGCGCT | -3.84770 | 197 (-361) | AGTGTGGGGCCGCGAGCTATTGCCGGGCGCT | 2.42800 | 197 (-361) | AGTGTGGGGCCGCGAGCTATTGCCGGGCGCT | -3.93700 | 197 (-361) | AGTGTGGGGCCGCGAGCTATTGCCGGGCGCT | 2.32980 |

En cuando a las ESE, la posición está entre 210 y 216 y no hay cambios destacables entre la secuencia WT (arriba) y la mutante (abajo):

| | | | | | | | | | | | |
|---------------|---------|----------|---------------|---------|----------|---------------|----------|----------|---------------|---------|----------|
| 206 (-352) | CCGCAGC | -1.10651 | 210 (-348) | AGCTACT | -2.88649 | 206 (-352) | CCGCAGCT | -4.21284 | 210 (-348) | AGCTACT | -3.01430 |
| 207 (-351) | CGCAGCT | 2.60677 | 211 (-347) | GCTACTG | -2.22486 | 207 (-351) | CGCAGCTA | 0.37743 | 211 (-347) | GCTACTG | -2.55576 |
| 208 (-350) | GCAGCTA | -4.05230 | 212 (-346) | CTACTGC | -0.43736 | 208 (-350) | GCAGCTAC | -4.30554 | 212 (-346) | CTACTGC | 4.56729 |
| 209 (-349) | CAGCTAC | -1.95879 | 213 (-345) | TACTGCC | -2.64433 | 209 (-349) | CAGCTACT | -3.18593 | 213 (-345) | TACTGCC | -4.42531 |
| 210 (-348) | AGCTACT | -4.26889 | 214 (-344) | ACTGCCG | -4.28616 | 210 (-348) | AGCTACTG | 3.99633 | 214 (-344) | ACTGCCG | 0.46059 |
| 211 (-347) | GCTACTG | -1.87050 | 215 (-343) | CTGCCGG | 3.74721 | 211 (-347) | GCTACTGC | -5.16068 | 215 (-343) | CTGCCGG | 1.30988 |
| 212 (-346) | CTACTGC | -2.56737 | 216 (-342) | TGCCGGG | 1.54498 | 212 (-346) | CTACTGCC | -1.81437 | 216 (-342) | TGCCGGG | 0.55274 |
| 213 (-345) | TACTGCC | -4.20076 | 217 (-341) | GCCGGGC | -1.52940 | 213 (-345) | TACTGCCG | 2.26820 | 217 (-341) | GCCGGGC | -3.69701 |
| 214 (-344) | ACTGCCG | -6.16466 | 218 (-340) | CCGGGCG | 0.32925 | 214 (-344) | ACTGCCGG | -2.92357 | 218 (-340) | CCGGGCG | -1.22559 |
| 215 (-343) | CTGCCGG | 3.50000 | 219 (-339) | CGGGCGC | 2.17372 | 215 (-343) | CTGCCGGG | -1.76742 | 219 (-339) | CGGGCGC | -1.76657 |
| 216 (-342) | TGCCGGG | 0.86693 | 220 (-338) | GGGCGCT | -0.06906 | 216 (-342) | TGCCGGGC | -1.99114 | 220 (-338) | GGGCGCT | -4.85220 |

| | | | | | | | | | | | |
|---------------|---------|----------|---------------|---------|----------|---------------|----------|----------|---------------|---------|----------|
| 206 (-352) | CCGCAGC | -1.10651 | 210 (-348) | AGCTATT | -2.78249 | 206 (-352) | CCGCAGCT | -4.21284 | 210 (-348) | AGCTATT | -4.73123 |
| 207 (-351) | CGCAGCT | 2.60677 | 211 (-347) | GCTATTG | -4.03665 | 207 (-351) | CGCAGCTA | 0.37743 | 211 (-347) | GCTATTG | -1.06451 |
| 208 (-350) | GCAGCTA | -4.05230 | 212 (-346) | CTATTGC | -1.69211 | 208 (-350) | GCAGCTAT | -4.30554 | 212 (-346) | CTATTGC | 2.18266 |
| 209 (-349) | CAGCTAT | -0.10539 | 213 (-345) | TATTGCC | -4.71319 | 209 (-349) | CAGCTATT | -2.54149 | 213 (-345) | TATTGCC | -3.07704 |
| 210 (-348) | AGCTATT | -3.81118 | 214 (-344) | ATTGCCG | -4.03738 | 210 (-348) | AGCTATTG | 2.26731 | 214 (-344) | ATTGCCG | 0.14853 |
| 211 (-347) | GCTATTG | -4.39423 | 215 (-343) | TTGCCGG | 0.64216 | 211 (-347) | GCTATTGC | -5.91992 | 215 (-343) | TTGCCGG | 1.66657 |
| 212 (-346) | CTATTGC | -4.02469 | 216 (-342) | TGCCGGG | 1.54498 | 212 (-346) | CTATTGCC | -2.04521 | 216 (-342) | TGCCGGG | 0.55274 |
| 213 (-345) | TATTGCC | -6.51077 | 217 (-341) | GCCGGGC | -1.52940 | 213 (-345) | TATTGCCG | 1.69927 | 217 (-341) | GCCGGGC | -3.69701 |
| 214 (-344) | ATTGCCG | -5.56173 | 218 (-340) | CCGGGCG | 0.32925 | 214 (-344) | ATTGCCGG | -1.53517 | 218 (-340) | CCGGGCG | -1.22559 |
| 215 (-343) | TTGCCGG | 0.54952 | 219 (-339) | CGGGCGC | 2.17372 | 215 (-343) | TTGCCGGG | -1.78855 | 219 (-339) | CGGGCGC | -1.76657 |
| 216 (-342) | TGCCGGG | 0.86693 | 220 (-338) | GGGCGCT | -0.06906 | 216 (-342) | TGCCGGGC | -1.99114 | 220 (-338) | GGGCGCT | -4.85220 |

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 | 0 | 4 | 2 | 36 | 539.1709 | 7 | -12.9646 | 19 | 5 | 63 | 739.1964 | 67 | 95.7670 | 49 | 154 | 0.32 |
| mut | 0 | 4 | 2 | 38 | 563.5265 | 7 | -13.9268 | 19 | 5 | 64 | 729.9196 | 66 | 94.0661 | 51 | 154 | 0.33 |

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

HOT-SKIP

cagatgagctaaggggcccgcaggcatttcggagggatcacattaggagttcagagttgaggagctcggatggggctgg
gggctgaacctggaaggccagcagggccttggagccaggtttagggttcaaggggcaggtctctgctggcctgaccca
gcccccttgcttctcgcagTGGTACCTTCCAGGAAGTGTGGGGCCGCAGCTACTGCCGGGCGCTGGAGAGGCTGGTGG
ACGTGCTGTCCGAGTACCCAGCGAGGTGGAGCACATGTTAGCCCATCCTGTGTCTCCCTGCTGCGCTGCACCGGCTGC
TGCGGCGATGAGAATCTGCACCTGTGTGCCGGTGGAGACGGCCAATGTACCATGCAGgtaggtccataccctgcccaggg
ggccaccgaatctgcccactagaaggtacctggagaggggaagaagataggcccaggtcccagggctggagcctgactccc
aggccagtgtctcctaccacccagcccagggcctcttcttcctaagccctccgatggctctgagcccactcatctcc