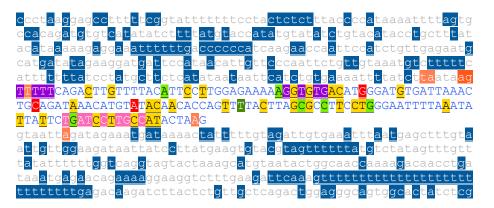**Ejemplo comparación de resultados predictores in sillico**

Cambio de estudio MFSD8 c.554-1G>T (chr4:127939998 G/T, COSV99614626 o NM_001371596.2: c.554-1G>T)

Exón 6 e intrones adyacentes:

El cambio se encuentra en la primera posición justo antes del exón 6 (la **g** en color rojo/naranja).

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              Donor splice sites, direct strand
-------------------------------              -------------------------------
        pos 5'->3'  phase strand  confidence  5'      exon intron   3'              pos 5'->3'  phase strand  confidence  5'      exon intron   3'
            386        2      +        0.37         CCATACTAAG^GTAATTAGAT                386        2      +        0.37         CCATACTAAG^GTAATTAGAT

Donor splice sites, complement strand              Donor splice sites, complement strand
-------------------------------------              -------------------------------------
  pos 3'->5'  pos 5'->3'  phase strand  confidence  5'      exon intron   3'    pos 3'->5'  pos 5'->3'  phase strand  confidence  5'      exon intron   3'
      51         635        0      -        0.00       TATAAAGCAG^GTATGTACAG          51         635        0      -        0.00       TATAAAGCAG^GTATGTACAG

Acceptor splice sites, direct strand              Acceptor splice sites, direct strand
------------------------------------              ------------------------------------
   No acceptor site predictions above threshold.             pos 5'->3'  phase strand  confidence  5'    intron exon     3'
                                                                  248        0      +        0.17      ATTTTTTCAG^ACTTGTTTTA

Acceptor splice sites, complement strand              Acceptor splice sites, complement strand
----------------------------------------              ----------------------------------------
  pos 3'->5'  pos 5'->3'  phase strand  confidence  5'    intron exon     3'    pos 3'->5'  pos 5'->3'  phase strand  confidence  5'    intron exon     3'
      562        124        2      -        0.33      CATTTATCAG^GTTGTCTTTT          562        124        2      -        0.33      CATTTATCAG^GTTGTCTTTT
```

Se detecta en la secuencia mutante un sitio *acceptor* (en azul) que no estaba en la secuencia WT. Este se encuentran desplazado por delante, es decir, hacia el interior del exón con respecto al sitio *acceptor* del exón. Por lo tanto, en el caso de que se empleara este en vez del normal, lo más probable es que se estuvieran perdiendo los ocho nucleótidos del principio del exón 6.

Splice Site Prediction by Neural Network (NNSplice)

Donor site predictions for wt :

Start	End	Score	Exon	Intron
270	284	0.83	gaaaaag**gt**gtgaca	
379	393	0.99	tactaag**gt**aattag	
469	483	0.76	atgaagt**gt**acgtag	
515	529	0.58	tggtcag**gt**agtact	

Donor site predictions for mut :

Start	End	Score	Exon	Intron
270	284	0.83	gaaaaag**gt**gtgaca	
379	393	0.99	tactaag**gt**aattag	
469	483	0.76	atgaagt**gt**acgtag	
515	529	0.58	tggtcag**gt**agtact	

Acceptor site predictions for wt :

Start	End	Score	Intron	Exon
400	440	0.54	atgataaaactatttttgt**ag**attgtgaaatttaatgagct	
479	519	0.97	cgtagtttttttatgtctat**ag**tttgtttatattttttggtc	
501	541	0.88	ttgtttatattttttggtc**ag**gtagtactaaagcatgtaat	
616	656	0.94	ttttttttttttttttttg**ag**acaagatcttactctgttgc	

Acceptor site predictions for mut :

Start	End	Score	Intron	Exon
400	440	0.54	atgataaaactatttttgt**ag**attgtgaaatttaatgagct	
479	519	0.97	cgtagtttttttatgtctat**ag**tttgtttatattttttggtc	
501	541	0.88	ttgtttatattttttggtc**ag**gtagtactaaagcatgtaat	
616	656	0.94	ttttttttttttttttttg**ag**acaagatcttactctgttgc	

Spliceman

Point mutation	Wildtype (wt)	Mutation (mt)	L1 distance	Ranking (L1)
aataa(g/t)ttttt	agtttt	attttt	32870	92%

Human Splicing Finder

🔴 Broken WT Acceptor Site	Alteration of the WT Acceptor site, most probably affecting splicing

Algorithm/Matix	position	sequences	variation
HSF Acceptor site (matrix AG)	chr4:127940009	- REF : TATCTTAATAAGTT - ALT : TATCTTAATAATTT	74.7 > 46.84 => -37.3%

SVM-BPfinder

```
seq_id  agez   ss_dist  bp_seq    bp_scr  y_cont          ppt_off  ppt_len  ppt_scr  svm_scr
wt      23     69       ttctcatat  0.404829721522  0.609375         9        9        15       0.32715088
wt      23     64       atataataa  -1.75189148877  0.627118644068   4        9        15       -0.19508512
wt      23     61       taataattc  0.120839138196  0.642857142857   1        9        15       0.73315547
wt      23     57       aattcatct  -0.783806165659 0.634615384615   11       9        21       -0.20080591
wt      23     50       ctgtgaaaa  -1.22157370831  0.644444444444   4        9        21       0.074050155
wt      23     42       attttatct  -3.97130308554  0.675675675676   11       6        17       -1.4728637
wt      23     37       atcttaata  -2.60071158764  0.65625 6        6        17       -0.62599469
wt      23     36       tcttaataa  -1.06390416857  0.677419354839   5        6        17       0.045873458
wt      23     33       taataagtt  -1.04082579095  0.714285714286   2        6        17       0.25671143
wt      23     25       ttttcagac  -1.65123590362  0.7     3        17       32       0.089532501
mut     23     69       ttctcatat  0.404829721522  0.625   9        9        15       0.3321973
mut     23     64       atataataa  -1.75189148877  0.64406779661    4        9        15       -0.18961103
mut     23     61       taataattc  0.120839138196  0.660714285714   1        9        15       0.7389228
mut     23     57       aattcatct  -0.783806165659 0.653846153846   11       9        21       -0.19459494
mut     23     50       ctgtgaaaa  -1.22157370831  0.666666666667   4        9        21       0.08122728
mut     23     42       attttatct  -3.97130308554  0.702702702703   10       7        20       -1.3728889
mut     23     37       atcttaata  -2.60071158764  0.6875  5        7        20       -0.524656
mut     23     36       tcttaataa  -1.06390416857  0.709677419355   4        7        20       0.14753771
mut     23     33       taataattt  -0.468856119354 0.75    1        7        20       0.58344512
mut     23     25       ttttcagac  -1.65123590362  0.7     3        17       32       0.089532501
```

El cambio que hay entre ambas secuencias se da por la presencia de la mutación que hace que el BP tenga mejor *score*, probablemente haciéndolo más fuerte.

Variant Effect Predictor tool

ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641503.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641508.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641509.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641538.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	downstream_gene_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641558.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641590.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641658.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641686.2	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641690.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641695.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641742.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641743.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641748.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641753.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641774.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641776.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	upstream_gene_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641830.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641843.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641869.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641870.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641882.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641928.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant	MFSD8	ENSG00000164073	Transcript	ENST00000641949.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	upstream_gene_variant	MFSD8	ENSG00000164073	Transcript	ENST00000642012.1	retained_intron	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000642034.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	splice_acceptor_variant	MFSD8	ENSG00000164073	Transcript	ENST00000642042.1	protein_coding	-	-	-	-	-	-	CS120920, COSV99614626
ENST00000641743.1:c.554-1G>T	4:127939998-127939998	A	intron_variant, NMD_transcript_variant	MFSD8	ENSG00000164073	Transcript	ENST00000642078.1	nonsense_mediated_decay	-	-	-	-	-	-	CS120920, COSV99614626

ESEfinder

Solo se encuentran una predicción con puntuación positiva para las matrices 3'SS (no se va a tener encuentra 226 porque solo en una de las dos matrices se obtiene puntuación positiva). Si se compara está puntuación con la equivalente a la secuencia mutante se observa que ha aumentado, por lo que lo más probable es que se esté fortaleciendo un sitio *acceptor,* lo que puede estar afectando al *splicing*.

234 (-452)	taataagTTTTTCAGACTTGTTTTACATTC	-14.93020	234 (-452)	taataagTTTTTCAGACTTGTTTTACATTC	1.02320	234 (-452)	taataagTTTTTCAGACTTGTTTTACATTC	-15.66840	234 (-452)	taataagTTTTTCAGACTTGTTTTACATTC	0.59250
234 (-452)	taataatTTTTTCAGACTTGTTTTACATTC	-15.58000	234 (-452)	taataatTTTTTCAGACTTGTTTTACATTC	2.68240	234 (-452)	taataatTTTTTCAGACTTGTTTTACATTC	-16.30870	234 (-452)	taataatTTTTTCAGACTTGTTTTACATTC	2.19800