

Secuencias: [NP_009231.2](#) y [NP_009225.1](#)

Mediante el programa Needle alineamos las dos secuencias y observamos que hay un hueco en el alineamiento.

```

NP_009225.1    1401 QHNLIKLQQEMAELEAVLEQHGSQPSNSYPSIISDSSALEDLRNPEQSTS    1450
                |||||||||||||||||||||||||||||||||||||||||||||||||||
NP_009231.2    1401 QHNLIKLQQEMAELEAVLEQHGSQPSNSYPSIISDSSALEDLRNPEQSTS    1450

NP_009225.1    1451 EK-----AVLTSQKSSEYPISQNPEGLSADKFEV    1479
                ||                               : |||||||||||||||
NP_009231.2    1451 EKDSIHGQRNNSMFSKRPREHISVLTSQKSSEYPISQNPEGLSADKFEV    1500

NP_009225.1    1480 SADSSTSKNKEPGVERSSPSKCPSLDDRWMHSCGSLQNRNYPSEELI    1529
                |||||||||||||||||||||||||||||||||||||||||||||||
NP_009231.2    1501 SADSSTSKNKEPGVERSSPSKCPSLDDRWMHSCGSLQNRNYPSEELI    1550

```

¿A qué corresponde el hueco en el alineamiento?

Las dos secuencias corresponden a la isoforma 1 y isoforma 2 del gen BRCA1 y por lo tanto el hueco debería reflejar cambios en el splicing.

Para saber el tipo de splicing alineamos primero las secuencias del mRNA correspondiente

NP_009225.1 -> [NM_007294.4](#)

NP_009231.2 -> [NM_007300.4](#)

mRNA:

```

NM_007300.4    4401 |||||||||||||||||||||||||||||||||||||||||||||||    4450
                CCTTCCATCATAAGTGACTCTTCTGCCCTTGAGGACCTGCGAAATCCAGA
NM_007294.4    4451 ACAAAGCACATCAGAAAAAG-----    4470
                |||||||||||||||
NM_007300.4    4451 ACAAAGCACATCAGAAAAAGATTCGCATATACATGGCCAAAGGAACAAC    4500
NM_007294.4    4471 -----CAGTATTAAC TTCACAG    4487
                |||||||||||||||
NM_007300.4    4501 CCATGTTTTCTAAAAGGCCTAGAGAACATATATCAGTATTAAC TTCACAG    4550

```

Nos fijamos en la posición del hueco, que es la 4471 (la última base es la 4470). Si buscamos en las anotaciones de las dos isoformas (https://www.ncbi.nlm.nih.gov/nuccore/NM_007294.4 https://www.ncbi.nlm.nih.gov/nuccore/NM_007300.4) vemos que en esta posición empieza un exon en ambas isoformas. Sin embargo tienen longitudes diferentes. **4471-4536 (isoforma 1) y 4471-4597 (isoforma 2).**

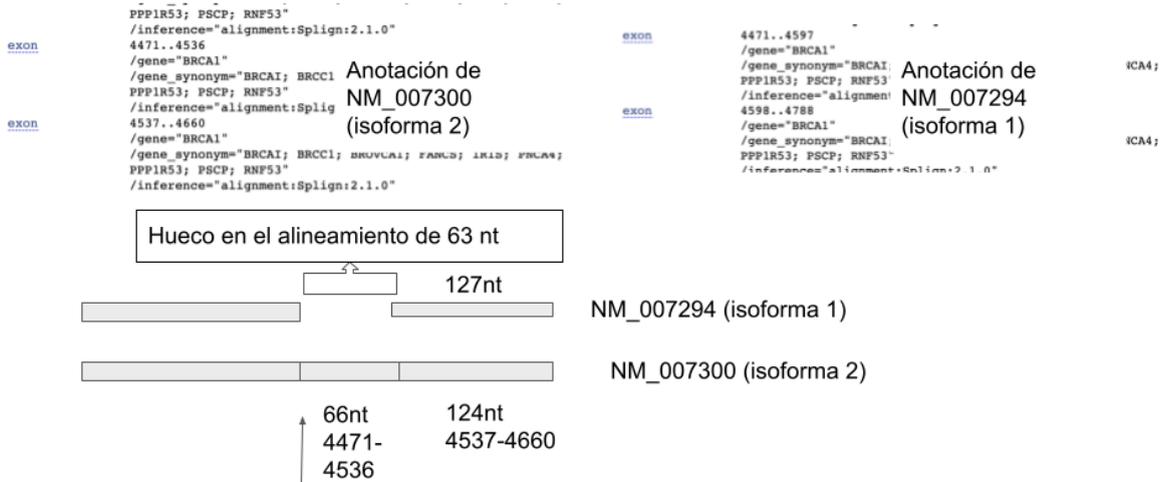
```
exon          4471..4536
              /gene="BRCA1"
              /gene_synonym="BRCAI; BRCC1; BROVCA1; FANCS; IRIS; PNCA4;
              PPP1R53; PSCP; RNF53"
              /inference="alignment:Splign:2.1.0"
exon          4537..4660
              /gene="BRCA1"
              /gene_synonym="BRCAI; BRCC1; BROVCA1; FANCS; IRIS; PNCA4;
              PPP1R53; PSCP; RNF53"
```

→ longitud exon: **4536-4471 +1 = 66 nt**

-> longitud exon: **4660 - 4537 +1 = 124nt**

```
exon          4471..4597
              /gene="BRCA1"
              /gene_synonym="BRCAI; BRCC1; BROVCA1; FANCS; IRIS; PNCA4;
              PPP1R53; PSCP; RNF53"
              /inference="alignment:Splign:2.1.0"
```

Longitud exon: **4597 - 4471 +1 = 127 nt**



Dos eventos de splicing: i) exon skipping (uso de un exon adicional de 66 nt en la isoforma 2), ii) corte-empalme en 5' del siguiente exon

Confirmación con el Navegador Genómico

Verificación mediante el Navegador genómico: introducimos BRCA1

- i) Desactivamos la pista GENCODE V36
- ii) seleccionamos 'full' en NCBI RefSeq

GENCODE V36
NCBI RefSeq

Zoom en el siguiente exon

Exon adicional en la isoforma 2

Extremos diferentes en el extremo 5' (¡el gen se ubica en la hebra menos!)

Eventos de splicing y sus frecuencias:

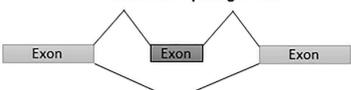
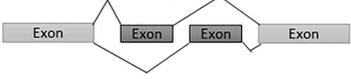
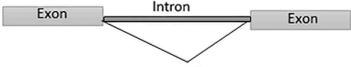
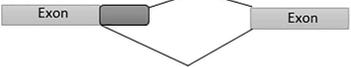
Alternative Splicing Events		Observed Frequency		Potential Outcomes
		Humans	Plants	
	(i) Exon Skipping (ES)	~40%	~8%	Diversifying proteome complexity
	(ii) Mutually Exclusive Exons (MXE)	~10.7%	NA	Diversifying proteome complexity
	(iii) Intron Retention (IR)	~5%	~60%	(i) Sequester in the nucleus and produce on demand, (ii) PTC+ degraded by NMD pathway, or (iii) may generate truncated protein
	(iv) Alternative 5' Splice Site (A5SS)	~7.9%	~7.5%	Diversifying proteome complexity
	(v) Alternative 3' Splice Site (A3SS)	~18.4%	~15.5%	Diversifying proteome complexity
	(vi) Exon Intron (EI)	~4%	~4%	(i) Generates complete protein domain, (ii) intrinsically disordered protein, or (iii) post-translational modifications to improve substrate specificity

Figura tomado de: <https://www.frontiersin.org/articles/10.3389/fpls.2019.00708/full>

a) ¿Por qué hay un desemparejamiento al final del hueco?

Si una región codificante sufre una inserción o deleción que es múltiple de 3, entonces no cambia la pauta de lectura. Sin embargo, si los nucleótidos eliminados no corresponden a una ristra de codones, entonces veremos adicionalmente un desemparejamiento. Consideramos la siguiente secuencia de la eliminamos 3 nucleótidos (abajo con ---)

ATGAAGTTG**GAT**ATC

ATG---TTGGATATC → ATGTTGGATATC → se han eliminado en fase y por eso no hay desemparejamiento en el alineamiento

ATGAAGTTG**GAT**ATC

ATGA---TGGATATC → ATG**ATG**GATATC

→ no se han eliminado en fase y aparece un nuevo codón (en azul) que alineando las proteínas se nos manifiesta como desemparejamiento