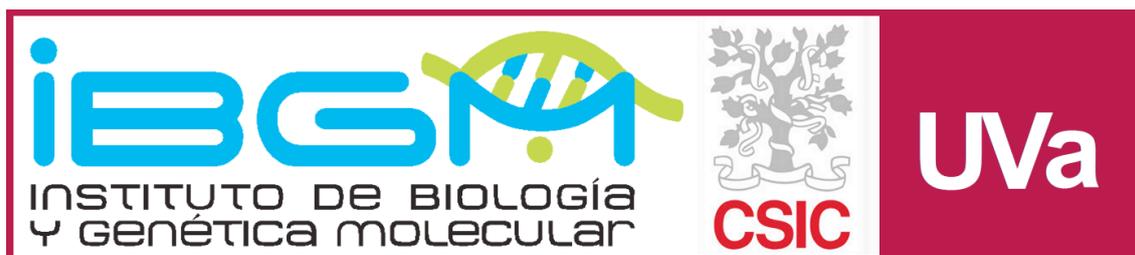


**UNIVERSIDAD DE VALLADOLID**

**FACULTAD DE MEDICINA**

DEPARTAMENTO DE BIOQUÍMICA, BIOLOGÍA MOLECULAR Y FISIOLOGÍA  
INSTITUTO DE BILOGÍA Y GENÉTICA MOLECULAR (UVa-CSIC)

---



Alteraciones de *splicing* del gen *BRCA2* en cáncer de  
mama y ovario. Diseño del nuevo vector de *splicing* pSAD  
para la construcción de minigenes híbridos.

Memoria para optar al Grado de Doctor presentada por

Alberto Acedo Bécares

2013

Director:

Dr. Eladio Andrés Velasco Sampedro



Don Eladio Andrés Velasco Sampedro, Doctor en Biología y Científico Titular del CSIC en el Instituto de Biología y Genética Molecular de Valladolid

**CERTIFICA:**

Que el trabajo titulado: “Alteraciones de *splicing* del gen *BRCA2* en cáncer de mama y ovario. Diseño del nuevo vector de *splicing* pSAD para la construcción de minigenes híbridos”.

que presenta el Licenciado en Ciencias Ambientales D. Alberto Acedo Bécares, ha sido realizado bajo su dirección y reúne, a su juicio, la originalidad y contenidos suficientes para que sea presentado ante el tribunal correspondiente y pueda optar al Grado de Doctor por la Universidad de Valladolid.

Para que así conste, y a los efectos oportunos, firma la presente en Valladolid, a 23 de Septiembre de 2013.

Dr. Eladio Andrés Velasco Sampedro



El presente trabajo ha sido financiado por el Fondo Social Europeo y la Consejería de Educación de la Junta de Castilla y León bajo el Programa Operativo Castilla y León 2007-2013, el Instituto de Salud Carlos III (Referencia del proyecto: PI10/02910), la Consejería de Educación de Castilla y León (Referencia del proyecto: CSI00410-2) y la Consejería de Sanidad (Referencia del proyecto BIO39/VA27/10).



**A MI ESPOSA,  
VICTORIA**



Varios son los años transcurridos desde mi llegada a Valladolid y todavía recuerdo con entusiasmo esa primera visita al laboratorio que sería mi hábitat. Recuerdo estos años con una sensación de felicidad y, a la vez, una melancolía inexplicable. Una etapa de crecimiento personal y profesional rodeado de excelentes personas, y es que la ilusión por la ciencia se palpa en cada persona, en cada rincón.

Muchos sois los que me habéis acompañado en esta etapa, que me habéis ayudado y apoyado, que habéis estado conmigo en este período de mi vida y que espero sigáis estando presentes en el futuro. A todos tengo que agradecer este trabajo.

En primer lugar quiero dar las gracias a mi director de tesis. Gracias, Eladio, porque me has permitido cumplir uno de mis sueños: dedicarme a la investigación. Un día confiaste en mí y trabajando contigo he aprendido mucho más que de ciencia. Tú has sabido transmitirme la ilusión y el entusiasmo por el trabajo bien hecho y ha sido un verdadero placer compartir todo este tiempo.

También quiero agradecer esta tesis a la Dra. Cristina Miner, por la confianza que ha depositado en mí y por su ayuda a lo largo de este trabajo. Y por extensión, a todo el laboratorio de diagnóstico genético del Programa de Prevención de Cáncer Hereditario del IBGM: la Dra. Mercedes Durán y la Dra. Mar Infante por darme la oportunidad de colaborar estrechamente que tantos buenos momentos nos ha dado.

Y qué puedo decir de mis compañeros de laboratorio, esa pequeña familia con los que he compartido el día a día. Los que están y los que estuvieron. Sin vosotros nada habría sido igual. De todos he aprendido mucho; de todos me llevo muchas experiencias. De entre vosotros quiero mencionar especialmente a: David, por haberme enseñado los entresijos de cada técnica; Bea, por tu paciencia conmigo; Álvaro y Cristina por los buenos ratos entre pipetas y por vuestra labor de continuidad; Noemí, Lara y Lucía por la ayuda prestada con los equipos, por compartir anécdotas, risas y reactivos; Cecilia, José y tantos otros que hemos compartido excelentes ratos fuera y dentro del laboratorio; al grupo de metabolopatías por su apoyo y cafés disfrutados a lo largo de este tiempo.

A mis amigos de toda la vida, a esa pandilla que me aguanta, me apoya y se ilusiona conmigo en cada proyecto. Siempre me habéis animado a seguir adelante desde el cariño.

A mi compañera en la vida, quien me ha sabido levantar en los baches y abrazar en los éxitos. Tú, que me haces sonreír cada mañana, que despiertas lo mejor de mí: gracias especialmente a ti.

Quisiera terminar con un gracias infinito a los más cercanos, a mi familia, a mis padres y hermano, que me consuelan cuando me frustro y se alegran más que yo cuando las cosas salen bien.

A todos los que me habéis ayudado a conseguir este reto.

Gracias por apoyarme y creer en mí.

# ÍNDICE



<b>ÍNDICE</b>	11
<b>RESUMEN</b>	17
<b>LISTADO DE ABREVIATURAS</b>	21
<b>INTRODUCCIÓN</b>	27
<b>I. Splicing: Conceptos Básicos.</b>	29
I.1. Definición de <i>splicing</i> .	29
I.2. Estructura del pre-ARNm.	29
I.2.a. Tipos de intrones.	29
I.2.b. Secuencias básicas de <i>splicing</i> .	30
I.2.c. Secuencias reguladoras de <i>splicing</i> .	30
I.3. Reacción bioquímica de <i>splicing</i> .	31
I.4. El espliceosoma.	32
I.4.a. Espliceosoma mayor.	32
I.4.b. Espliceosoma menor.	33
I.4.c. Conformación dinámica del espliceosoma: La reacción de <i>splicing</i> .	33
I.5. Exón o intrón definición.	34
I.5.a. Modelo de definición de exón.	34
I.5.b. Modelo de definición de intrón.	35
I.5.c. Reconocimiento del primer y último exón.	36
I.5.c.1. Adición de la caperuza y reconocimiento del primer exón el ARNm naciente.	37
I.5.c.2. Factores de poliadenilación del ARN y reconocimiento del último exón.	37
<b>II. Splicing Alterativo.</b>	37
II.1. Tipos de <i>splicing</i> alternativo.	38
II.2. Origen evolutivo del <i>splicing</i> alternativo.	39
II.2.a. <i>Exon shuffling</i> .	39
II.2.b. Exonización.	40
II.2.c. Transición.	40
II.3. Factores principales de la regulación del <i>splicing</i> alternativo.	42
II.3.a. Proteínas SR.	43
II.3.a.1. Mecanismos de actuación de las proteínas SR.	43
II.3.a.2. Factores implicados en la actividad de las proteínas SR.	44
II.3.b. Proteínas hnRNP.	44
II.3.b.1. Mecanismos de actuación de las hnRNPs.	45
II.4. Otros factores implicados en la regulación del <i>splicing</i> .	45
II.4.a. Factores de <i>splicing</i> específicos de tejido.	45
II.4.b. Estructura secundaria.	46
II.4.c. Parámetros relacionados con la transcripción y la condensación de la cromatina. Teoría cinética.	47
<b>III. Alteraciones en el proceso de <i>splicing</i>: Enfermedad.</b>	48
III.1. Enfermedades provocadas por la alteración del <i>splicing</i> en <i>cis</i> .	49
III.1.a. Alteración de los sitios de <i>splicing</i> . Síndrome QT-largo.	49
III.1.b. Alteración de secuencias ESE. Síndrome de Zellweger.	50
III.1.c. Alteración de secuencias ESS. Distrofia muscular de Duchenne.	50
III.2. Enfermedades provocadas por la alteración del <i>splicing</i> en <i>trans</i> .	51
III.2.a. Alteración de la secuencia de genes implicados en el proceso de <i>splicing</i> .	51
III.2.a.1. Retinitis pigmentaria.	51
III.2.a.2. Atrofia muscular espinal.	51
III.2.a.3. El síndrome de Taybi-Linder.	52
III.2.b. Alteraciones en la expresión de los genes relacionados con el <i>splicing</i> .	52
III.2.b.1. Enfermedades infecciosas que alteran la expresión de factores de <i>splicing</i> : SIDA.	53
III.3. <i>Nonsense Mediated Decay</i> .	53

III.4. <i>Splicing</i> y cáncer.....	54
III.4.a. Cáncer y mutaciones en <i>cis</i> que alteran el <i>splicing</i> .....	54
III.4.b. Expresión diferencial de factores de <i>splicing</i> y su implicación en cáncer.....	55
III.4.b.1. Cáncer y expresión de los proteínas SR.....	55
III.4.b.2. Cáncer y expresión de proteínas hnRNP.....	56
III.4.b.3. Otros factores de <i>splicing</i> relacionados con el cáncer.....	57
<b>IV. El síndrome de cáncer de mama y ovario hereditario (CMOH).....</b>	<b>57</b>
IV.1. Principales genes de predisposición a CMOH: <i>BRCA1</i> y <i>BRCA2</i> .....	58
IV.2. Espectro de predisposición en CMOH.....	58
IV.3. Modelo poligénico de predisposición a CMOH.....	59
IV.4. Estructura de los genes <i>BRCA1</i> y <i>BRCA2</i> .....	60
IV.5. <i>BRCA1</i> y <i>BRCA2</i> en la protección de la integridad del genoma.....	61
IV.6. El espectro mutacional en los genes BRCA.....	63
<b>OBJETIVOS</b>	<b>67</b>
<b>MATERIAL Y MÉTODOS</b>	<b>71</b>
<b>I. Muestras biológicas de pacientes.....</b>	<b>73</b>
I.1. Detección de mutaciones en pacientes con CMOH. Análisis de heterodúplex.....	74
I.2. Mutaciones seleccionadas del Programa de prevención de CMOH.....	75
<b>II. Análisis <i>in silico</i> del impacto de mutaciones en el <i>splicing</i>.....</b>	<b>76</b>
II.1. Bases de datos.....	81
II.1.a. Bases de datos de secuencias. Ensembl.....	81
II.1.b. Bases de datos de mutaciones. BIC, LOVD y UMD.....	82
II.2. Análisis <i>in silico</i> del impacto de mutaciones en el <i>splicing</i> .....	82
II.2.a. NNsplice.....	82
II.2.b. ESEfinder.....	82
II.2.c. ESRsearch.....	83
II.2.d. Human Splicing Finder.....	83
II.3. Estudio evolutivo de la conservación de secuencias implicadas en el <i>splicing</i> en <i>BRCA2</i> .....	84
<b>III. Ensayos funcionales de <i>splicing</i>. RT-PCR de linfocitos. Detección de isoformas de <i>splicing</i> de muestras de pacientes.....</b>	<b>84</b>
III.1. Extracción de ARN.....	84
III.2. Retrotranscripción ARN-ADNc.....	85
III.3. RT-PCR.....	85
III.4. Electroforesis en gel de agarosa.....	86
III.5. Secuenciación de ADN mediante electroforesis capilar.....	86
<b>IV. Ensayos funcionales de <i>splicing</i> mediante minigenes híbridos.....</b>	<b>87</b>
IV.1. Construcción de minigenes.....	88
IV.1.a. Tipos de vectores.....	88
IV.1.a.1. Vector pSPL3.....	88
IV.1.a.2. Nuevo vector: pSAD.....	88
IV.1.b. Minigenes en pSPL3. MGBR2 19-20 y 23-24.....	90
IV.1.c. Minigenes en pSAD.....	92
IV.1.c.1. MGBR2 2-9.....	93
IV.1.c.2. MGBR2 16-20.....	95
IV.1.c.3. MGBR2 19-27.....	97
IV.1.d. Transformación.....	101
IV.1.e. Selección de colonias.....	101
IV.1.e.1. Selección de colonias recombinantes.....	101
IV.1.e.2. Rastreo de colonias por PCR.....	101
IV.1.f. Extracción y comprobación de las construcciones.....	102
IV.2. Mutagénesis dirigida.....	103

IV.2.a. Mutaciones puntuales .....	103
IV.2.b. Microdeleciones.....	111
IV.3. Transfección .....	114
IV.3.a. Células humanas en cultivo.....	114
IV.3.b. Protocolo de transfección.....	115
IV.3.c. Extracción de ARN.....	115
IV.4. Comprobación de isoformas .....	116
IV.4.a. RT-PCR.....	116
IV.4.a.1. Electroforesis en gel .....	117
IV.4.a.2. Cuantificación de isoformas. Electroforesis capilar .....	117
IV.4.a.3. Caracterización de isoformas. Secuenciación capilar .....	118
<b>RESULTADOS</b> .....	<b>119</b>
<b>I. Resultados RT-PCR linfocitos de pacientes.....</b>	<b>121</b>
I.1. Resultado de la mutación c.8488-1G>A.....	121
I.2. Resultado mutación c.9026_9030del .....	123
I.3. Resultado mutación c.9501+9A>C .....	123
<b>II. Análisis Bioinformático.....</b>	<b>124</b>
<b>III. Ensayos funcionales de <i>splicing</i> con minigenes híbridos.....</b>	<b>128</b>
<b>IV. Ensayos funcionales de <i>splicing</i> con minigenes construidos en pSPL3 .....</b>	<b>128</b>
IV.1. Resultados del minigen MGBR2 19-20 .....	128
IV.1.a. Exón 19.....	129
IV.1.b. Exón 20.....	134
IV.1.c. Intrón 19 .....	137
IV.1.d. Resultados globales de los exones 19 y 20 .....	140
IV.2. Resultados del minigen MGBR2 23-24 .....	141
IV.2.a. Exón 23.....	141
IV.2.b. Exón 24.....	147
IV.2.c. Resultados globales de los exones 23 y 24 .....	151
<b>V. Nuevo vector pSAD .....</b>	<b>152</b>
<b>VI. Ensayos funcionales de <i>splicing</i>. Minigenes construidos en pSAD. ....</b>	<b>153</b>
VI.1. Minigen MGBR2 2-9 .....	153
VI.2. Minigen MGBR2 16-20.....	154
VI.2.a. Mapeo de elementos reguladores de <i>splicing</i> mediante microdeleciones solapantes exónicas en el MGBR2 16-20.....	155
VI.2.b. Resultado mutagénesis MGBR2 16-20.....	157
VI.3. Minigen MGBR2 19-27 .....	159
VI.3.a. Resultados del MGBR2 19-27.....	160
VI.3.a.1. Resultados del MGBR2 19-27. Mutaciones ensayadas previamente en pSPL3 .....	161
VI.3.a.2. Exón 21.....	164
VI.3.a.3. Exón 22.....	169
VI.3.a.4. Exón 25.....	174
VI.3.a.5. Exón 26.....	177
VI.3.a.6. Exón 27.....	180
VI.3.b. Mapeo de elementos reguladores de <i>splicing</i> mediante microdeleciones solapantes exónicas en el MGBR2 19-27.....	183
VI.3.c. Comparación entre células HeLa y MCF-7 .....	187
VI.3.d. Inhibición del mecanismo de <i>nonsense mediated mRNA decay</i> .....	190
<b>DISCUSIÓN</b> .....	<b>193</b>
<b>I. Correlación entre <i>splicing</i> y enfermedad en las variantes detectadas en pacientes del programa de prevención del CMOH.....</b>	<b>195</b>
<b>II. Cambio en el efecto de una mutación por alteración del <i>splicing</i> .....</b>	<b>200</b>
II.1. Mutaciones frameshift y nonsense .....	200

II.2. Mutaciones <i>missense</i> .....	201
II.3. Mutaciones sinónimas .....	202
II.4. Mutaciones de sitios de <i>splicing</i> .....	203
II.5. Mutaciones intrónicas.....	203
<b>III. Reproducibilidad de los resultados obtenidos en minigenes .....</b>	<b>206</b>
<b>IV. Técnicas de identificación de alteraciones de <i>splicing</i> provocadas por mutaciones puntuales en <i>cis</i> .....</b>	<b>211</b>
<b>V. Tipos de alteración en <i>cis</i> detectadas y efecto en el proceso de <i>splicing</i>.....</b>	<b>216</b>
<b>VI. Evolución y perspectivas del modelo de cáncer de mama y ovario hereditario. ....</b>	<b>221</b>
<b>VII. Propuesta de procedimiento clínico de evaluación de variantes de ADN ...</b>	<b>223</b>
<b>VIII. Aplicaciones y perspectivas de los estudios de <i>splicing</i>.....</b>	<b>225</b>
VIII.1. Terapias de <i>splicing</i> .....	225
VIII.2. Nuevas tecnologías de análisis global de patrones de <i>splicing</i> .....	227
 <b>CONCLUSIONES</b>	 <b>231</b>
 <b>BIBLIOGRAFÍA</b>	 <b>.235</b>

# RESUMEN



**Introducción:** El *splicing* es el proceso de maduración de ARNm previo a la traducción, en el que los intrones son eliminados del precursor de ARNm y los exones unidos de forma secuencial. La regulación de este proceso depende de múltiples y diferentes señales e implica diferentes macrocomplejos proteicos como el espliceosoma. Las alteraciones en el proceso de *splicing* han demostrado ser un importante mecanismo de patogenicidad en la gran mayoría de enfermedades de etiología genética, por lo que su conocimiento resulta de gran importancia en la práctica clínica. El objetivo central de este trabajo fue el estudio de la correlación entre alteraciones de *splicing* del gen supresor de tumores *BRCA2* y la susceptibilidad hereditaria a cáncer de mama y ovario.

**Material y Métodos:** 333 variantes de ADN detectadas en pacientes, entre las que se incluyen tres detectadas en el programa de prevención del cáncer de Castilla y León y 300 registradas en la base de datos BIC fueron analizadas mediante una estrategia combinada, que consiste en la utilización de diversas herramientas bioinformáticas junto con ensayos funcionales de ARN de linfocitos y/o minigenes híbridos. Asimismo, se han aplicado novedosas soluciones tecnológicas como la creación de un nuevo vector de *splicing* pSAD o la cuantificación relativa de isoformas de *splicing* a través de la detección de transcritos mediante electroforesis capilar. Setenta y nueve variantes seleccionadas fueron reproducidas mediante mutagénesis dirigida en los minigenes *wild type* y ensayados funcionalmente en células eucariotas.

**Resultados:** Se construyeron un total de cinco minigenes: dos en pSPL3: el MGBR2 19-20 que incluía los exones 19, 20 y sus regiones intrónicas adyacentes y el MGBR2 23-24 con los exones 23 y 24; tres en pSAD, el MGBR2 2-9 que incluye los exones 2, 3, 4, 5, 6, 7, 8 y 9 y regiones intrónicas colindantes, el MGB2 16-20 con los exones 16, 17, 18, 19 y 20 y por último el MGB2 19-27 que contiene los exones 19, 20, 21, 22, 23, 24, 25, 26, y 27 que además sustituyó a parte de exón V2 del vector.

Un total de 33 variantes de 79 seleccionadas más 6 generadas de forma artificial con el fin de identificar elementos reguladores de *splicing*, alteraron el proceso de *splicing*. Ocho variantes alteraron el *splicing* en el exón 19 y sus regiones intrónicas flanqueantes (c.8378C>A, c.8486A>T, c.8486A>C, c.8487G>A, c.8487+1G>A, c.8487+3A>G, c.8487+3A>C y c.8487+3A>T), tres en el exón 20 (c.8488-2A>G, c.8488-1G>A y c.[8609A>G;8611G>T]), nueve en el exón 23 (c.8954-3C>G, c.8954-

1\_8955delinsAA, c.8969G>A, c.9006A>T, c.[8972G>A;9006A>T], c.9076C>T, c.9117G>A, c.9117+1G>T, c.9117+1G>A), cuatro en el exón 24 (c.9118-2A>G, c.9248\_9256+7del, c.9256G>T, c.9256+1G>A), seis en el exón 21 (c.8633-24\_8634del, c.8754G>A, c.8754+4A>G, c.8754+5G>A, c.8754+5G>T, c.8754+5G>C), tres en el exón 22 (c.8755-1G>A, c.8948\_8953+5del, c.8953+1G>T), dos en el exón 25 (c.9257-1G>C y c.9501+3A>T), dos en el exón 26 (c.9502-12T>G y c.9502-2A>C) y dos en el exón 27 (c.9649-2A>G, c.9649-5A>G y c.9698G>T). Así mismo, se realizaron 54 microdeleciones en los exones 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, y 27 en busca de elementos reguladores de *splicing* exónicos. Cinco microdeleciones resultaron positivas, localizando elementos reguladores en secuencias específicas de los exones 17, 18, 19, 20 y 23.

Los minigenes de gran tamaño permiten mantener el contexto genómico natural donde se producen las reacciones de *splicing*. De hecho, en la mutación c.8488-1G>A se consiguen reproducir los patrones de *splicing* alterados obtenidos en muestras de paciente.

**Discusión:** Los resultados obtenidos en este estudio aportan luz sobre el mecanismo de patogenicidad relacionado con la alteración del *splicing* de las variantes de ADN estudiadas. En este sentido, podemos afirmar que una importante fracción de variantes de ADN está asociada con *splicing* anómalo del gen *BRCA2* y éste podría incluso ser considerado como el principal mecanismo de inactivación génica en variantes de ADN no clasificadas. Las nuevas herramientas desarrolladas en este trabajo para la realización de ensayos funcionales de *splicing*, el nuevo vector pSAD y los grandes minigenes construidos deberían ser incorporadas de rutina como instrumentos de apoyo en los estudios genéticos, ya que permiten evaluar cualquier variante de ADN en relación con el proceso de *splicing*. Por último y no menos importante, los estudios funcionales de *splicing* mediante minigenes híbridos permiten esclarecer los complejos mecanismos de regulación de la expresión génica en los eucariotas.

## *LISTADO DE ABREVIATURAS*



<b>9G8</b>	Splicing factor 9G8 or splicing factor arginine/serine-rich 7
<b>aa</b>	Aminoácido
<b>ADAR</b>	Adenosina deaminasa específica de ARN
<b>ADN</b>	Ácido desoxirribonucleico
<b>ADNc</b>	Ácido desoxirribonucleico complementario
<b>APC</b>	Adenomatous polyposis coli
<b>ARNm</b>	Ácido ribonucleico mensajero
<b>ARNt</b>	Ácido ribonucleico transferente
<b>AS</b>	<i>Splicing</i> Alternativo
<b>ATM</b>	Ataxia telangiectasia mutated
<b>ATP</b>	Adenosín trifosfato
<b>BARD1</b>	BRCA1-associated RING domain protein 1
<b>Bcl-X</b>	Anti-apoptotic Bcl-2 gene
<b>BIC</b>	Breast Cancer Information Core
<b>BIN1</b>	Myc box-dependent-interacting protein 1
<b>BP</b>	Branch point, punto de ramificación
<b>BRCA1</b>	Breast cancer 1
<b>BRCA2</b>	Breast cancer 2
<b>BRCT</b>	Dominio C-terminal de BRCA1
<b>BRIP1</b>	BRCA1 interacting protein C-terminal helicase 1
<b>BSCL2</b>	Berardinelli-Seip congenital lipodystrophy 2 (seipin)
<b>CAP</b>	Caperuza
<b>CAPER</b>	Coactivator of activatingprotein-1 and estrogen receptors
<b>CASP8</b>	aspase 8, apoptosis-related cysteine peptidase
<b>CDH1</b>	Cadherin 1, type 1, E-cadherin
<b>CDK2</b>	Cyclin-dependent kinase 2
<b>CELF</b>	CUG-BP and ETR3-like factors
<b>CF</b>	Factor de corte
<b>CHEK2</b>	Checkpoint kinase 2
<b>CLASP1</b>	Cytoplasmic linker associated protein 1
<b>CMOH</b>	Cáncer de Mama y Ovario Hereditario
<b>CPSF</b>	Factor específico de poliadenilación y corte
<b>Csft</b>	Factor de estimulación de corte
<b>CtIP</b>	Chymotrypsin-like protease
<b>D-MEM</b>	Dulbecco's Modified Eagle's Medium
<b>ddntps</b>	2'3'-dideoxirribonucleótidos trifosfato
<b>DMD</b>	Distrofia muscular de Duchenne
<b>dntps</b>	Desoxirribonucleótidos trifosfato
<b>BRCT</b>	Dominio C-terminal de BRCA1
<b>Downstream</b>	Posterior en la secuencia
<b>DSB</b>	Roturas de sobre cadena
<b>Dscam</b>	Down syndrome cell adhesion molecule
<b>EDTA</b>	Thylenediaminetetraacetic acid
<b>ESE</b>	Exonic splicing enhancer, Secuencia enhancer del splicing exónica
<b>ESS</b>	Exonic splicing silencer, Secuencia silenciadora del splicing exónica

<b>FAM20A</b>	Family with sequence similarity 20, member A
<b>FANCD2</b>	Fanconi anemia, complementation group D2
<b>FANCI</b>	Fanconi anemia, complementation group I
<b>FANCM</b>	Fanconi anemia, complementation group M
<b>FBS</b>	Suero bovino fetal
<b>FN</b>	Fibronectina
<b>Fox</b>	Neuro-oncological ventral antigen
<b>FW</b>	Forward
<b>GWAS</b>	Genome-wide association study
<b>H</b>	Domino Helicoidal
<b>H2AX</b>	Chromatin-associated histone
<b>HA-CAE</b>	Análisis de heredoduplex mediante electroforesis capilar
<b>HGMD</b>	The Human Gene Mutation Database
<b>HGVS</b>	Human Genome Variation Society
<b>HIV-1</b>	Human immunodeficiency virus 1, Virus de inmunodeficiencia humano
<b>hnRNP</b>	Heterogeneous nuclear ribonucleoproteins. Ribonucleoproteínas heterogéneas nucleares
<b>HPRP3</b>	Homolog pre-mRNA splicing factor 3
<b>hSNF5</b>	Chromatin-remodeling factor
<b>IARC</b>	Unclassified Genetic Variants Working Group.
<b>IBGM</b>	Instituto de Biología y Genética Molecular
<b>IPTG</b>	Isopropil- $\beta$ -D-tiogalactósido
<b>ISE</b>	Intronic splicing enhancer. Secuencia enhancer del splicing intrónica
<b>ISS</b>	Intronic splicing silencer. Secuencia silenciadora del splicing intrónica
<b>IVS</b>	Intervening sequence variant or Intronic variant. Variante intrónica
<b>Kb</b>	Kilo base
<b>KCNH2</b>	Potassium voltage-gated channel, subfamily H member 2
<b>Kda</b>	Kilo Dalton
<b>KH</b>	K homology domain
<b>LB</b>	Medio Luria-Bertani
<b>LOVD</b>	Leiden Open (source) Variation Database
<b>MCF7</b>	Human breast adenocarcinoma cell line
<b>MCS</b>	Multiple cloning site. Sitio de clonación múltiple o poliensamblador
<b>MD</b>	Mutagénesis dirigida
<b>MDC1</b>	Mediator of DNA-damage checkpoint 1
<b>MGBR2</b>	Minigen BRCA2
<b>MOPD I</b>	Enanismo primordial osteodisplásico tipo I
<b>MSP</b>	Proteína estimuladora de macrófagos
<b>MYC</b>	v-myc myelocytomatosis viral oncogene homolog
<b>NARF</b>	Factor de reconocimiento de prolamina A nuclear
<b>NBN</b>	Nibrin
<b>NBS1</b>	Nijmegen breakage syndrome protein 1
<b>NF1</b>	Neurofibromatosis tipo 1
<b>NF2</b>	Neurofibromatosis tipo II
<b>NGS</b>	Next Generation Sequencing
<b>NLS</b>	Secuencias de localización nuclear

<b>NMD</b>	Nonsense-mediated mRNA decay. Degradación del ARNm mediada por mutaciones terminadoras
<b>Nova</b>	Neuro-oncological ventral antigen
<b>Nt</b>	Nucleótido
<b>OB</b>	Domino de union a oligonucleótidos
<b>ORF</b>	Open reading frame
<b>P53</b>	Tumor protein 53.
<b>PALB2</b>	Partner and localizer of BRCA2
<b>PAP</b>	Polimerasa poliA
<b>PAP1</b>	Polimerasa poliA1
<b>Pb</b>	Pares de bases
<b>PCR</b>	Polymerase Chain Reaction
<b>PEX</b>	Peroxisoma
<b>PP1</b>	Proteína fosfatasa 1
<b>PP1 y 2</b>	Protein Phosphatase 1 and 2
<b>PP2</b>	Proteína fosfatasa 2
<b>PPT</b>	Tiamina pirofosfato
<b>pre-ARNm</b>	Ácido ribonucleico pre-mensajero
<b>PRPF31</b>	Pre-mRNA processing factor 31 homolog
<b>PRPF8</b>	Pre-mRNA processing factor 8 homolog
<b>PT</b>	Tracto de polipirimidina
<b>PTB</b>	Polypyrimidine tract-binding protein. Proteína de unión al tracto de polipirimidina
<b>PTCs</b>	Codones prematuros de terminación
<b>PTEN</b>	Phosphatase and tensin homolog
<b>RAD50</b>	DNA repair RAD51C
<b>RAD51C</b>	DNA repair RAD51C
<b>RAD51D</b>	DNA repair RAD51D
<b>RAP80</b>	Receptor Associated Protein 80
<b>RING</b>	Ring-type zinc finger
<b>RNF8</b>	Ring finger protein 8, E3 ubiquitin protein ligase
<b>RON</b>	Macrophage-stimulating protein receptor
<b>RRM</b>	Motivo de reconocimiento de ARN
<b>RS</b>	Arginine-serine-rich domain
<b>RT-PCR</b>	Reverse transcription polymerase chain reaction
<b>RV</b>	Reverse
<b>sam68</b>	Nuclear RNA-binding protein
<b>SC35.</b>	Splicing factor arginine/serine-rich 2 or Splicing component 35 kDa.
<b>SCD</b>	SQ-cluster domain
<b>SF1mBBP</b>	Human Splicing Factors
<b>SF2/ASF</b>	Splicing factor arginine/serine-rich 1 or pre-mRNA-splicing factor SF2
<b>SF3b1</b>	Splicing factor 3b1
<b>SIDA</b>	Síndrome de inmunodeficiencia adquirida
<b>SINE</b>	Elementos cortos intercalables
<b>siRNA</b>	Small interfering RNA. Ácido ribonucleico pequeño de interferencia.

<b>SLX4</b>	SLX4 structure-specific endonuclease subunit
<b>SMA</b>	Atrofia muscular espinal
<b>SMC</b>	Sitio de clonación múltiple
<b>SMN1</b>	Survival of motor neuron 1
<b>SMN2</b>	Survival of motor neuron 2
<b>SNP</b>	Single-nucleotide polymorphism
<b>snRNA</b>	Small nuclear ribonucleic acid. Ácido ribonucleico pequeño nuclear.
<b>snRNP</b>	Small nuclear ribonucleoprotein. Ribonucleoproteína pequeña nuclear.
<b>SNRNP200</b>	Small nuclear ribonucleoprotein 200kDa (U5)
<b>Splicing</b>	Ajuste o empalme del ARN
<b>SRE</b>	Secuencia Reguladora de Splicing
<b>SRp40</b>	Splicing factor arginine/serine-rich 5 or Pre-mRNA-splicing factor SRP40
<b>SRp55</b>	Splicing factor arginine/serine-rich 6 or Pre-mRNA-splicing factor SRP55.
<b>SRPK1</b>	Serina/treonina proteína quinasa 2
<b>SRPK2</b>	Serina/treonina proteína quinasa 2
<b>SRSF1</b>	Serine/arginine-rich splicing factor 1
<b>SS 3'</b>	Splice site 3'
<b>SS 5'</b>	Splice Site 5'
<b>STAR</b>	Signal transduction and activation of RNA proteins
<b>STK11</b>	Serine/threonine kinase 11
<b>SWI/SNF</b>	SWItch/Sucrose Non Fermentable
<b>T</b>	domino Torre
<b>Tat</b>	HIV transactivating regulatory gene
<b>TDP-43</b>	TAR-DNA binding protein 43 kDa
<b>TP53</b>	Tumor protein p53
<b>UCG</b>	Unidad de Consejo Genético
<b>UMD</b>	Universal Mutation Database
<b>Upstream</b>	Anterior en la secuencia
<b>UTR</b>	Untranslated region. Región no traducida.
<b>UVS</b>	Variante de efecto biológico desconocido
<b>VEGF</b>	Vascular endothelial growth factor
<b>WT</b>	Wild Type
<b>XGal</b>	5-bromo-4-cloro-3-indolil- $\beta$ -D-galactopiranosido
<b>XRCC2</b>	X-ray repair complementing defective repair in Chinese hamster cells 2

# INTRODUCCIÓN



## I. *Splicing*: Conceptos Básicos

En el año 1977, Phillip A. Sharp y Richard J. Roberts observaron por primera vez, a través del microscopio electrónico, como determinadas secuencias de ADN de un gen no estaban presentes en sus correspondientes ARNm transcritos. Este sorprendente hallazgo se realizó de forma casual mientras se efectuaban estudios independientes acerca de la transcripción de genes tardíos en adenovirus tipo 2 (Berget *et al.*, 1977; Chow *et al.*, 1977). Ambos investigadores recibieron el premio Nobel de medicina en 1993 por el descubrimiento del *splicing*.

### I.1. Definición de *splicing*

El *Splicing* es el proceso por el cual las secuencias intrónicas son identificadas y eliminadas de los transcritos del precursor de las moléculas de ARN mensajero (pre-ARNm) y las secuencias exónicas flanqueantes a los intrones son unidas de forma secuencial, para dar lugar a un ARNm maduro, que posteriormente será traducido a proteína (Gilbert, 1978).

El *splicing* puede considerarse constitutivo cuando un exón en particular siempre va a formar parte de el ARNm o alternativo cuando un exón puede ser eliminado en una proporción de los transcritos. El *splicing* alternativo es un mecanismo prevalente en los eucariotas superiores, en el que a partir de un único gen pueden ser expresadas múltiples proteínas diferentes (Nilsen y Graveley, 2010). La flexibilidad del proceso de *splicing* ha jugado un papel fundamental en la evolución de los organismos superiores, promoviendo la expansión del proteoma e interviniendo en la regulación de la expresión génica (Lareau *et al.*, 2004).

Las alteraciones en el proceso de *splicing* han demostrado ser un importante mecanismo de patogenicidad en la gran mayoría de enfermedades de etiología genética. El conocimiento y estudio de los procesos que alteran el correcto procesado del ARNm resulta de gran importancia en la práctica clínica (Baralle *et al.*, 2009).

### I.2. Estructura del pre-ARNm

#### I.2.a. Tipos de intrones

La eliminación de los intrones en el proceso de *splicing* puede ocurrir de forma autocatalítica, como por ejemplo en los intrones del grupo I y II de algunos genomas

bacterianos, plantas, algas y hongos. Sin embargo, en la mayoría de los genomas eucariotas se encuentran intrones que son reconocidos y eliminados del pre-ARNm mediante reacciones catalizadas por el “espliceosoma”, uno de los mayores complejos macromoleculares de la célula (Roy *et al.*, 2006). Existe además un cuarto tipo de intrones presente en los ARN de transferencia (ARNt) de los eucariotas, éstos no requieren de reacciones de transesterificación, puesto que son eliminados a través de endonucleasas y los exones unidos por medio de una enzima ligasa.

### I.2.b. Secuencias básicas de *splicing*

La integridad del marco abierto de lectura (ORF) requiere la identificación precisa de los límites entre las regiones intrónicas y exónicas. Los pre-ARNm de los vertebrados facilitan la información necesaria para el reconocimiento de un intrón por parte de la maquinaria de *splicing* a través de pequeñas secuencias intrónicas altamente conservadas en la evolución, como se muestra en la Figura 1. En el comienzo del intrón podemos encontrar el sitio de *splicing* 5' o donador (SS 5'), mientras que el extremo opuesto son necesarias tres secuencias asociadas para facilitar el reconocimiento del final del intrón: el sitio de *splicing* 3' o aceptor (SS 3'); el punto de ramificación (BP), localizado a 18-40 nucleótidos *upstream* del sitio aceptor y el tracto de polipirimidina (PT), situado entre el punto de ramificación y el sitio aceptor (Ast, 2004).

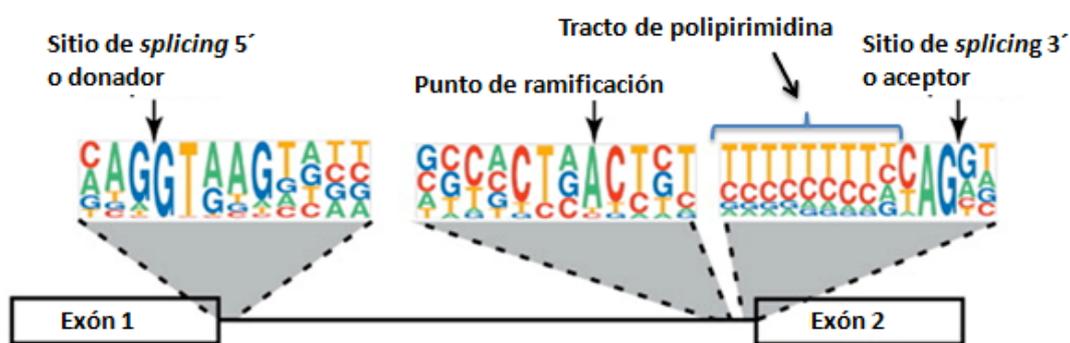


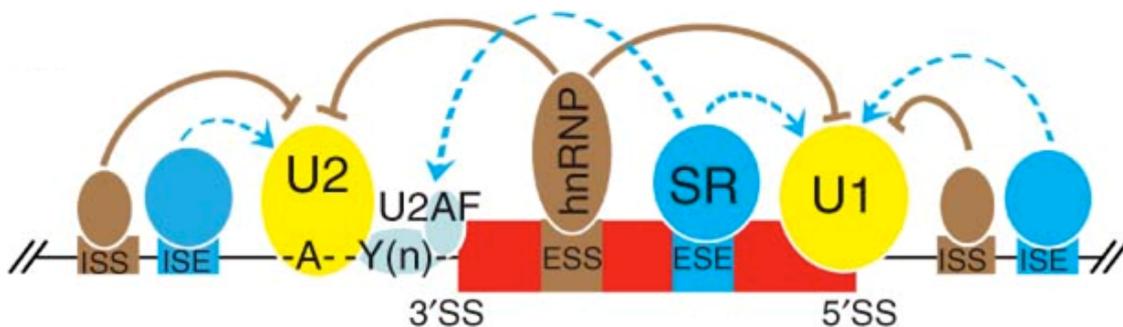
Figura 1. Secuencias básicas de reconocimiento de intrón por parte del espliceosoma (Padgett, 2012).

### I.2.c. Secuencias reguladoras de *splicing*

Existen además otras secuencias (SRE) a lo largo del pre-ARNm que son sitios de unión degenerados a los que pueden unirse diversas proteínas que actúan como reguladoras del proceso de *splicing* a través de su interacción con el espliceosoma

(Figura 2), que por ello son claves en este proceso (Smith y Valcarcel, 2000; Wang y Burge, 2008). En función de la localización de las SRE en las regiones intrónicas o exónicas y su efecto estimulador o represor del reconocimiento de los sitios de *splicing* por parte del espliceosoma, los podemos clasificar en:

- Secuencias exónicas que estimulan el *splicing* (ESE-Exonic splicing enhancer).
- Secuencias intrónicas que estimulan el *splicing* (ISE-Intronic splicing enhancer).
- Secuencias exónicas que reprimen el *splicing* (ESS-Exonic splicing silencer).
- Secuencias intrónicas que reprimen el *splicing* (ISS-Intronic splicing silencer).



**Figura 2. Secuencias reguladoras de *splicing*.** En azul se muestran secuencias que estimulan el *splicing* (ESE y ISE) a través de su unión con proteínas del tipo SR y que facilitan el reconocimiento de los sitios de *splicing* por parte de las subunidades del espliceosoma. En marrón, secuencias (ISS y ESS) que impiden el reconocimiento de los sitios de *splicing* (Lu *et al.*, 2011).

### I.3. Reacción bioquímica de *splicing*

Desde el punto de vista bioquímico el proceso de *splicing* ocurre a través de dos reacciones consecutivas de transesterificación. En la primera reacción tiene lugar el ataque nucleofílico del grupo hidroxilo 2' de la adenina del punto de ramificación al enlace fosfodiéster que une el sitio 5' del intrón con el último nucleótido del primer exón (verde-Figura 3). En esta reacción los exones continúan unidos al intrón pero se forma una estructura en lazo o *ariat*. El proceso continúa con la segunda reacción de transesterificación. El grupo hidroxilo del último nucleótido del exón 1 produce un ataque nucleofílico al enlace fosfodiéster que une el sitio 3' del intrón con el primer nucleótido del segundo exón (rojo-Figura 3). El resultado de esta reacción es la unión de los dos exones y la escisión del intrón (Moore *et al.*, 1993).

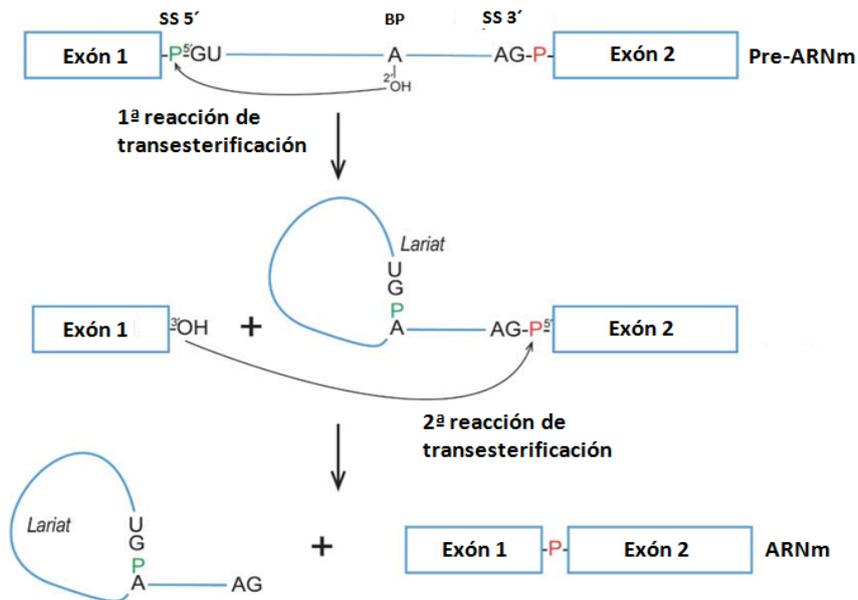


Figura 3. Reacción bioquímica del *splicing*. Doble reacción de transesterificación (Pérez-Valle y Vilardell, 2012).

#### I.4. El espliceosoma

Hace más de 25 años Osheim y Bayer observaron el espliceosoma por primera vez a través de micrografías. Ellos pudieron comprobar cómo determinadas partículas proteicas se unían al ARNm naciente durante el proceso de la transcripción. (Osheim *et al.*, 1985). Estas partículas parecían fusionarse y al hacerlo provocaban un lazo que dejaba fuera las regiones intrónicas del pre-ARNm (Bayer y Osheim, 1988).

##### I.4.a. Espliceosoma mayor

En la mayoría de los eucariotas coexisten dos tipos diferentes de espliceosomas, el más frecuente o espliceosoma mayor está formado por cinco pequeñas ribonucleoproteínas nucleares ricas en uridina, las snRNPs-U1, U2, U4, U5 y U6, en combinación con un número variable de proteínas asociadas (Jurica y Moore, 2003; Hegele *et al.*, 2012). El espliceosoma no es una estructura que exista de *novo*, su composición y conformación varía de forma dinámica a través de la reacción de *splicing*. Las distintas conformaciones y complejos que aparecen en la formación del espliceosoma han sido caracterizadas tanto *in vitro* como *in vivo* y serán descritos en el apartado I.4.c. (Tardiff y Rosbash, 2006; Hicks *et al.*, 2005).

### I.4.b. Espliceosoma menor

El espliceosoma menor procesa un determinado tipo de intrones minoritario llamados intrones U12. Los intrones U12 representan menos del 0,5% de todos los intrones de cualquier genoma y se diferencian del resto de intrones por las secuencias consenso de los sitios de *splicing* (Figura 4), AT en el extremo 5' y AC en el 3' (Wu y Krainer, 1997).

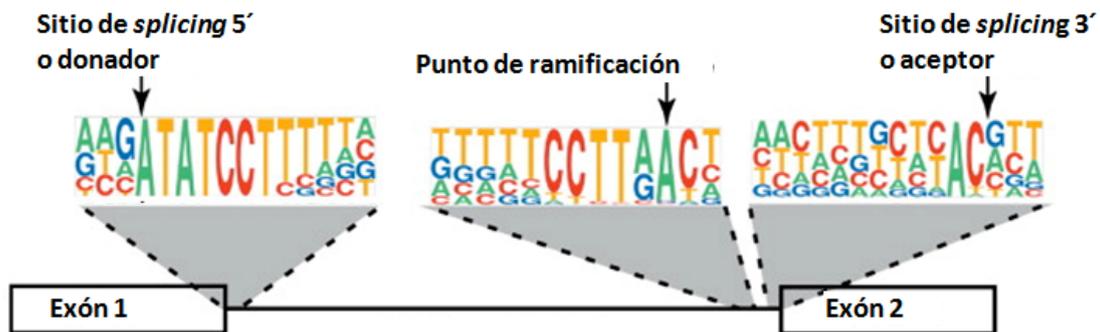


Figura 4. Secuencias básicas de reconocimiento de intrones U12 por parte del espliceosoma menor (Padgett, 2012).

Además, es importante destacar la alta conservación evolutiva de las secuencias del sitio donador y del punto de ramificación. Salvo por la snRNP U5, el resto de ribonucleoproteínas que conforman el espliceosoma menor, U11, U12, U4atac y U6tac, son específicas. Sin embargo, como se puede apreciar en la Figura 5, sus estructuras secundarias son muy similares a las de sus equivalentes del espliceosoma mayor (Janne *et al.*, 2012). Cabe la posibilidad de una controvertida diferencia entre los dos espliceosomas, algunos autores han propuesto que el espliceosoma menor actúa también fuera del núcleo (König *et al.*, 2007), pero actualmente no hay evidencias de que los eventos de *splicing* puedan ocasionalmente ocurrir o ser completados en el citoplasma (Steitz *et al.*, 2008).

### I.4.c. Conformación dinámica del espliceosoma: La reacción de *splicing*

El proceso de *splicing* sucede de manera muy similar tanto a través del espliceosoma mayor como menor. En ambos casos tiene lugar la aparición de cuatro complejos catalíticos diferentes que marcan las distintas fases del proceso de *splicing*:

- El **complejo E** establece el inicio de la reacción de *splicing*, para su conformación la

ribonucleoproteína snRNP-U1 reconoce y se une a la secuencia del sitio de *splicing* 5'. Por otro lado actúan factores no ribonucleoproteicos. La proteína SF1/mBBP se une al punto de ramificación y el heterodímero U2AF (U2AF65/35) se une tanto al tracto de polipirimidina como al sitio de *splicing* 3'.

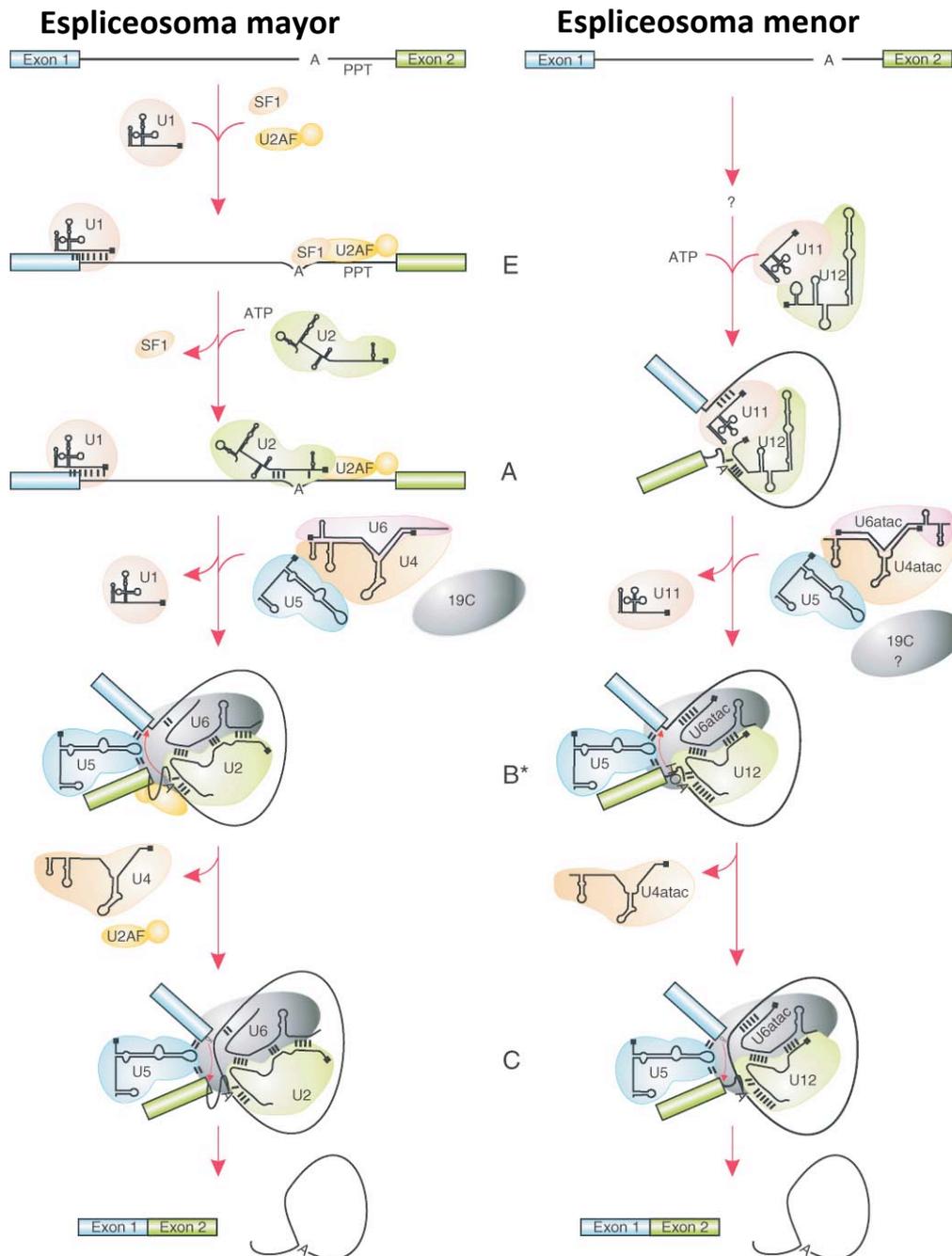
- A continuación, la snRNP-U2 desplaza la proteína SF1/mBBP y se une al punto de ramificación a través de una reacción dependiente de ATP formando el **complejo A**.
- El **complejo B** se forma por la unión del complejo tri-snRNP U4-U6/U5 al sitio donador.
- El **complejo C** se produce por la remodelación de las snRNAs y los factores proteicos constituyentes tras el desplazamiento de las snRNP U1 y U4 que depende de la asociación con el complejo Prp19/CDC5L. Este hecho crea un sitio activo que cataliza las reacciones de transesterificación que unirán los exones y liberará el intrón en su estructura de lazo o *ariat* (Will y Lührmann, 2011).

### I.5. Exón o intrón definición

Las principales preguntas que pueden surgir inmediatamente después de conocer el proceso de *splicing* son si el espliceosoma reconoce primero los exones y posteriormente los intrones son procesados o si este hecho ocurre al contrario. La arquitectura del genoma humano ha ayudado en parte a contestar esta cuestión.

#### I.5.a. Modelo de definición de exón

En 1990, Susan Berget comprobó que al aumentar de forma artificial el tamaño de un exón de un pre-ARNm de vertebrados por encima de los 300 nt se inhibía la formación del espliceosoma y por lo tanto enunció la teoría de “definición de exón” en la cual, cuando los exones de un pre-ARNm son pequeños y los intrones largos es muy probable que la maquinaria de *splicing* se forme a través del exón (Robberson *et al.*, 1990; Berget, 1995; Sakharkar *et al.*, 2005). Otro hecho que apoya esta teoría se fundamenta en que un exón constitutivo no es reconocido si su tamaño se reduce a menos de 50 nt de longitud (Dominski y Kole, 1991). Este efecto constata que sea necesario un espacio mínimo para que no existan interferencias estéricas entre los factores que reconocen los sitios de *splicing* y que delimitan el exón (Sternner *et al.*, 1996).



**Figura 5. Ensamblado y conformación dinámica del espliceosoma mayor y menor durante el proceso de *splicing*.** La figura muestra las diferentes snRNP que participan del proceso así como la estructura secundaria que adoptan en la formación de los diferentes complejos (E-A-B-C) durante el procesado del ARNm a lo largo de la reacción de *splicing* (Turunen *et al.*, 2012).

### I.5.b. Modelo de definición de intrón

Por otro lado, en los eucariotas inferiores la arquitectura genómica se caracteriza por pequeños intrones y largos exones, por lo que en este caso se ha propuesto el modelo de “definición de intrón” donde los sitios de *splicing* son reconocidos a través del intrón. Esta teoría ha sido refrendada de forma experimental tanto en levaduras

como en *Drosophila*, donde al aumentar el tamaño de los intrones se ha constatado un aumento en los defectos de *splicing* (Dominski y Kole, 1991). También se ha observado que los exones de *Drosophila* flanqueados por intrones muy largos sufren *splicing* alternativo de forma más habitual si se los compara con los exones flanqueados por intrones cortos. El efecto contrario se ha observado al estudiar los perfiles de *splicing* en humanos. En este caso, la activación de sitios de *splicing* alternativos es menos frecuente si los intrones flanqueantes son largos (Fox-Walsh *et al.*, 2005).

A nivel práctico, el mecanismo de ensamblado del espliceosoma es el mismo ya sea a través del modelo de definición de exón o de intrón y ambos modelos pueden aparecer en el proceso de *splicing* de una misma molécula de pre-ARNm. En estos casos el entorno genómico en el cual el exón está localizado podría tener un profundo efecto en su reconocimiento.

Un estudio reciente ha determinado que las diferencias en el contenido de GC entre los exones y sus intrones flanqueantes ayudan a marcar los exones para que sean reconocidos por la maquinaria de *splicing* en el modelo de definición de exón. Esta marca es menos importante en los exones rodeados de intrones cortos, ya que el proceso de *splicing* en este caso ocurre a través del modelo de definición de intrón (Amit *et al.*, 2012).

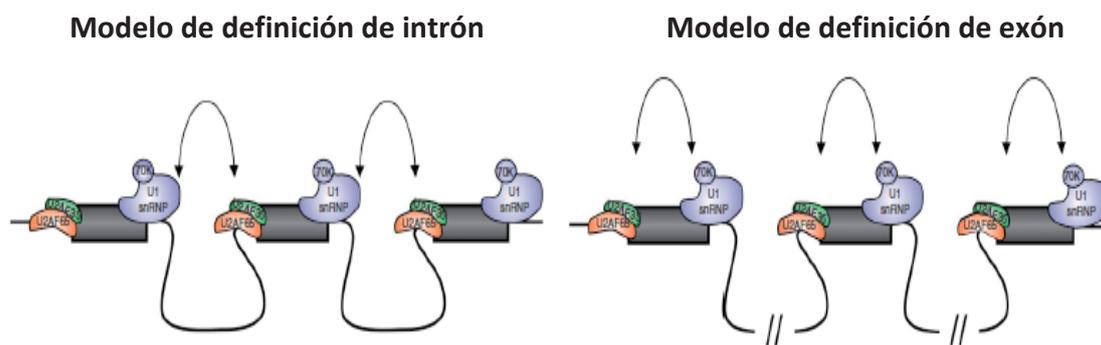


Figura 6. Representación esquemática del modelo de definición de intrón y del modelo de definición de exón (De Conti *et al.*, 2012).

### 1.5.c. Reconocimiento del primer y último exón

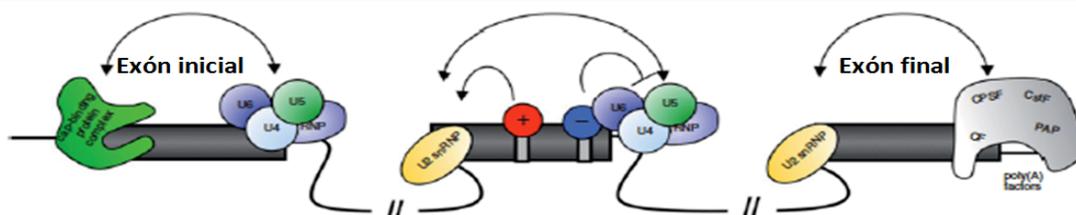
El primer y el último exón de una molécula de pre-ARNm constan únicamente de un sitio de *splicing*, por lo que su reconocimiento es diferente que el de los exones intermedios.

### I.5.c.1. Adición de la caperuza y reconocimiento del primer exón el ARNm naciente

En el caso de la definición del primer exón, el problema tiene una elegante solución. Transcurrido el inicio de la transcripción, el primer exón sufre la adición de una estructura llamada caperuza (CAP) en el extremo 5' (McCracken *et al.*, 1997). La CAP es un nucleótido modificado de guanina al que se une el complejo nuclear 7-metilguanosa fosfato para facilitar el reconocimiento del sitio 5' del exón por parte de la subunidad del espliceosoma snRNP U1 (Lewis *et al.*, 1996).

### I.5.c.2. Factores de poliadenilación del ARN y reconocimiento del último exón

Los mecanismos que regulan la definición del sitio 3' del exón terminal son más complejos y no están del todo definidos. Los datos actuales sugieren que en el reconocimiento del sitio 3' participan los factores de poliadenilación del ARNm, entre los cuales se encuentra el factor específico de poliadenilación y corte (CPSF), el factor de estimulación de corte (CstF), el factor de corte (CF) y la polimerasa poliA (PAP) (Pandya-Jones 2011).



**Figura 7. Reconocimiento del primer y último exón.** Representación esquemática del reconocimiento del primer exón a través del complejo de la caperuza y del último exón a través de diferentes factores implicados en la poliadenilación (De Conti *et al.*, 2012).

## II. *Splicing* Alternativo

El potencial de codificación de proteínas de nuestro genoma es cuantitativamente mínimo; se calcula que menos del 2% del genoma humano se traduce a proteínas (Elgar y Vavouri, 2008). Sin embargo, se ha estimado que unas cien mil proteínas diferentes son sintetizadas a partir de veinticuatro mil genes. (Modrek y Lee, 2002). Este fenómeno se produce gracias a la inclusión alternativa de diferentes exones procedentes de un mismo pre-ARNm durante el proceso de *splicing* o lo que se conoce como ***splicing* alternativo** (AS).

El AS permite un aumento en la variabilidad de los transcritos expresados por un genoma y por lo tanto en una mayor adaptabilidad. La importancia del *splicing* como factor multiplicador de la expresión génica se refleja en la complejidad de los organismos a lo largo del árbol filogenético. Diferentes estudios basados en secuenciación de ARN han confirmado esta asertación incluso entre los genomas de organismos pluricelulares. Por ejemplo, el porcentaje de genes con AS en humanos se estima que es del 95% frente al 60% y 25% en los genes de *Drosophila melanogaster* y *Caenorhabditis elegans*, respectivamente (Pan *et al.*, 2008; Graveley *et al.*, 2011; Ramani *et al.*, 2011; Barash, 2010). Los distintos ARNm producidos a partir del AS de un gen van a codificar proteínas con funciones distintas, incluso en algunos casos antagónicas (Singh y Valcarcel, 2005).

### II.1. Tipos de *splicing* alternativo

Como muestra la Figura 8, existen diferentes tipos de AS que podemos clasificar en cuatro grupos:

- El primer tipo es el conocido como exón *skipping* y se muestra en la Figura 8A. En este evento un exón conocido no es incluido en el ARNm final durante el proceso de *splicing*. El 40% del los procesos de AS ocurre de este modo en los eucariotas superiores, pero es muy raro en el resto. El incremento y prevalencia del exón *skipping* en la filogenia de los eucariotas sugiere que este tipo de AS es el que más contribuye a la complejidad fenotípica (Alekseyenko, *et al.*, 2007).
- El segundo y tercer tipo lo conforman la selección alternativa de los sitios de *splicing* 3' y 5' (Figuras 8B y 8C) representando entre el 18,4% y el 7,9% de todos los eventos de AS en los eucariotas superiores.
- El cuarto tipo corresponde con la retención de intrón. Aquí un intrón es incluido durante la maduración del ARNm (Figura 8D). Este tipo de AS apenas representa el 5% de los eventos de AS conocidos. Sin embargo, la retención de intrón es el evento de AS más frecuente en plantas, lo que sugiere que el AS no tiene un papel relevante en su evolución y en este caso los eventos de duplicación del genoma son los que contribuyen al incremento de la riqueza de su transcriptoma y proteoma (Kim *et al.*, 2008).

- Otra clase de eventos de AS ha sido descrita siendo su frecuencia muy baja. En ellos el proceso de AS se produce a través del uso de un promotor alternativo, una señal alternativa de poliadenilación, exones mutuamente excluyentes (Figuras 8E, F y G) o la reacción entre dos transcritos primarios en *trans* (Labrador y Corces, 2003).

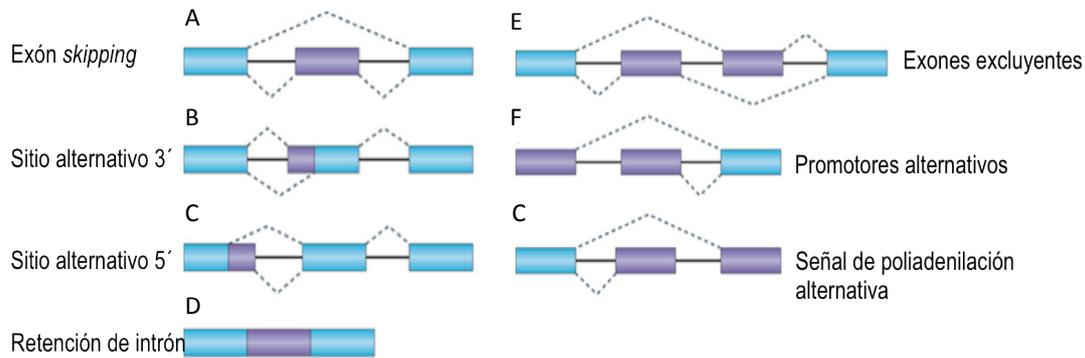


Figura 8. Tipos de patrones de *splicing* alternativo (Keren *et al.*, 2010).

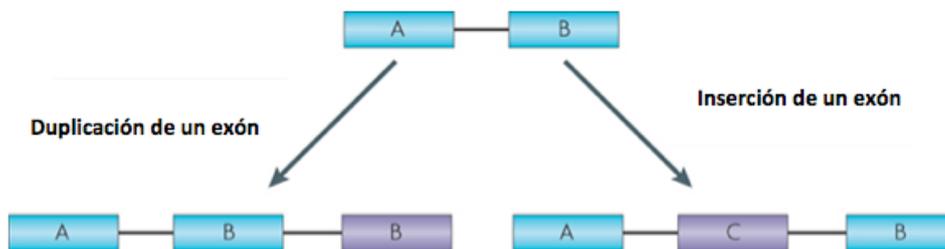
## II.2. Origen evolutivo del *splicing* alternativo

Análisis comparativos entre el genoma humano y el de ratón han mostrado que el AS se puede asociar con eventos recientes de ganancia o pérdida de exones (Thanaraj *et al.*, 2003). Actualmente se conocen tres mecanismos por los cuales un exón alternativo puede ser incluido en el proceso de *splicing* y serán descritos a continuación.

### II.2.a. *Exon shuffling*

El *Exon shuffling* o arrastre de exón es un proceso por el cual un exón es insertado en un gen o un exón es duplicado dentro del mismo gen. En 1978, Walter Gilbert fue el primero en proponer que el *exon shuffling* podría crear una nueva proteína quimérica otorgando una ventaja evolutiva a un organismo (Gilbert, 1978). Un estudio donde se han analizado nueve especies, incluidos vertebrados e invertebrados, concluye que el *exon shuffling* ha sido un mecanismo muy común en la evolución de los eucariotas y ha contribuido de forma sustancial a la complejidad del proteoma (Liu *et al.*, 2004). Además, se ha estimado que el 19% de los intrones de los genes eucariotas contienen exones que provienen del mecanismo de *exon shuffling*, donde el evento más frecuente es la duplicación en tándem de exones (Figura 9).

Éste se ha encontrado en el 10% de los genes humanos (Letunic *et al.*, 2002). La diversidad que puede generar la duplicación de exones se refleja en el gen *dscam* de *Drosophila melanogaster* donde múltiples exones mutuamente excluyentes proporcionan más de 38.016 isoformas diferentes mediante AS (Schmucker *et al.*, 2000).



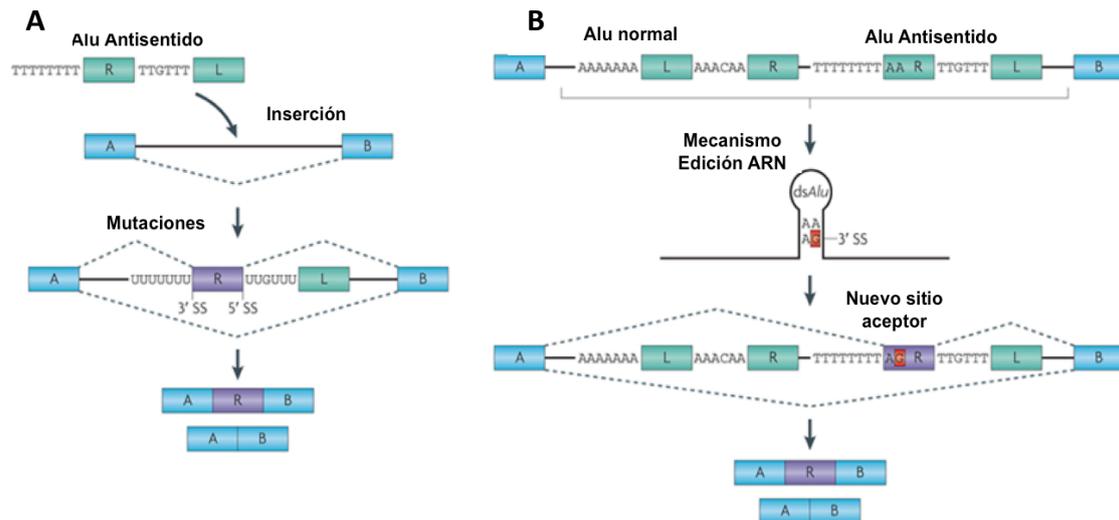
**Figura 9. Distintos procesos de inclusión de un exón a través del mecanismo de *exon skipping*.** Duplicación de exón en tándem e inserción de exón. El nuevo exón se muestra en color morado.

### II.2.b. Exonización.

La exonización es la forma de ganar un exón “de la nada” a través de secuencias genómicas que logran ser identificadas como exones. Se conocen dos mecanismos por los cuales se crean sitios de *splicing* 3' y 5' que pueden ser reconocidos por el espliceosoma y así exonizar una determinada secuencia: mutaciones al azar en secuencias intrónicas y mecanismos de corrección del ARN. Sendos procesos han actuado sobre los elementos transponibles, especialmente la familia de elementos cortos intercalables o SINE a la que pertenecen las secuencias Alu y se presentan en la Figura 10 (Sorek *et al.*, 2002; Nekrutenko *et al.*, 2001).

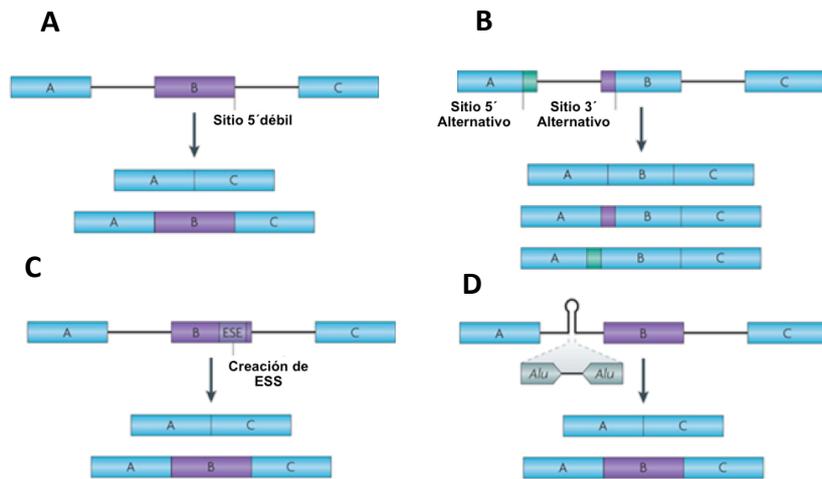
### II.2.c. Transición

En los dos procesos anteriores los exones que llegan a formar parte del *splicing* son generados *de novo*, mientras que en la transición los exones alternativos son derivados de los mismos exones constitutivos. La transición puede ocurrir por dos mecanismos:



**Figura 10. Ejemplos de exonización de secuencias intrónicas (Karen *et al.*, 2010).** **A) Exonización por mutaciones intrónicas en secuencias Alu.** Las secuencias consenso de los elementos Alu contienen múltiples secuencias similares a los sitios de *splicing*. El 85% de las exonizaciones ocurren en el brazo derecho (R) en orientación antisentido de un elemento Alu, puesto que en esta orientación la región rica en poliA se convertiría en un tracto de polipirimidina (PT). En el caso de que se produzcan determinadas mutaciones *downstream* del potencial PT, en regiones similares al sitio 3' de *splicing* y/o a unos 120 nt de éste, en el hipotético sitio 5', los nuevos sitios de *splicing* podrían ser reconocidos y por lo tanto, parte de la secuencia Alu exonizada (Gal-Mark *et al.*, 2009). **B) Exonización por mecanismos de corrección de ARN en secuencias Alu.** El segundo mecanismo para la creación de sitios de *splicing* requiere una estructura secundaria de ARN de doble cadena que se puede conformar cuando dos secuencias Alu aparecen insertadas de forma consecutiva y en orientaciones opuestas en un mismo intrón. Esta estructura es reconocida por la enzima ADAR una adenosina deaminasa específica de ARN y se produce la edición de adenosina a inosina. La inosina es reconocida como una guanina por la maquinaria biológica (Bass, 2002), por lo que una determinada secuencia AA puede pasar a ser AG, creando un nuevo sitio de *splicing* (Athanasiadis *et al.*, 2004). La exonización del exón 8 del gen *NARF* es un buen ejemplo de este proceso (Moller-Krull *et al.*, 2008).

- La acumulación de mutaciones en los sitios de *splicing*. Las mutaciones en los sitios de *splicing* pueden: debilitar los sitios de *splicing* existentes (Figura 11A) y aparecer isoformas alternativas; crear nuevos sitios de *splicing* alternativos (Figura 11B) y por último mutaciones en exones o intrones pueden alterar las secuencias reguladoras de *splicing* (Figura 11C) como las ESE o ISE o crear ESS o ISS (Ke *et al.*, 2008).
- La creación de estructuras secundarias en el ARN que impiden el reconocimiento de un exón (Figura 11D). Éste es el caso de las estructuras secundarias formadas en el ARNm precursor, por ejemplo por la cercanía de dos secuencias Alu en orientación opuesta (Lev-Maor *et al.*, 2008).



**Figura 11. Diferentes procesos de transición por los que un exón puede sufrir AS (Keren *et al.*, 2010).** A) Mutaciones que debilitan sitios de *splicing*. B) Mutaciones que crean sitios de *splicing* alternativos. C) Mutaciones que crean o eliminan ESE, ESS, IES, ISS. D) Estructuras secundarias del ARNm.

### II.3. Factores principales de la regulación del *splicing* alternativo

La selección de exones alternativos y por lo tanto la elección de los sitios de *splicing* está controlada por una complicada red de regulación en la que participan elementos reguladores de *splicing* en *cis* y en *trans*.

Las principales señales que definen los sitios de *splicing* han sido ampliamente comentadas en el apartado I.b. de esta introducción y como ya adelantamos en el mismo, se ha demostrado, mediante diferentes aproximaciones bioquímicas y bioinformáticas, la existencia de secuencias tanto intrónicas como exónicas que contienen información adicional para el reconocimiento de los sitios de *splicing* (Zhang *et al.*, 2003; Reed y Maniatis, 1986; Wang y Burge, 2008). Estas secuencias reguladoras de *splicing* (SRE) son reconocidas por diferentes proteínas de unión a ARN llamadas factores de transactivación de *splicing* que determinan si un exón debe ser seleccionado o no por el espliceosoma.

Aunque se conocen otras proteínas de unión a ARN que contribuyen a la regulación del *splicing* las proteínas de las familias SR y hnRNP son las más destacadas en relación al reconocimiento de secuencias SRE. Además su expansión evolutiva se relaciona con el incremento del *splicing* alternativo en organismos superiores (Busch y Hartel, 2012).

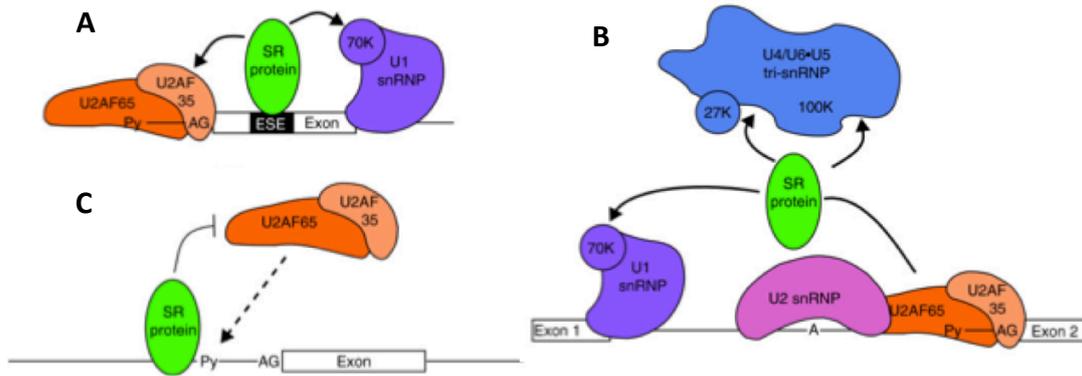
### II.3.a. Proteínas SR

La mayor parte de las secuencias, ya sean exónicas o intrónicas (ESE y ISE), que favorecen el *splicing* son reconocidas por los miembros de la familia de proteínas SR a través de su dominio RRM (RNA Recognition Motif), el cual les permite identificar secuencias específicas del ARN. Además se caracterizan por la presencia de dominios RS ricos en repeticiones de Arginina-Serina. En los seres humanos se han definido doce proteínas SR en función de la presencia de estos dominios. Otros factores que contienen dominios RS también han sido identificados, con o sin dominios RRM que carecen de la capacidad de influir en el *splicing*.

#### II.3.a.1. Mecanismos de actuación de las proteínas SR

Las proteínas SR contribuyen tanto al *splicing* constitutivo como alternativo favoreciendo la inclusión de un exón mediante la interacción de su dominio RS con el dominio RS de otro factor de *splicing*. Se han descrito tres mecanismos por el cual las proteínas SR actúan sobre el *splicing*:

- Las proteínas SR pueden facilitar el reconocimiento de los sitios de *splicing* por parte del espliceosoma una vez que se han unido a secuencias ESE o ISE a través de su interacción directa con los componentes del espliceosoma snRNP U1 o U2AF como se muestra en la Figura 12A (Lam y Hertel 2002).
- También se ha descrito que algunas proteínas SR pueden, sin llegar a unirse al ARN, facilitar la incorporación del tri-snRNP U4/U6-U5 en la formación del espliceosoma o interactuar simultáneamente con snRNP U1 y U2AF para facilitar el apareamiento de éstos con los sitios de *splicing* según se aprecia en la Figura 12B (Roscigno *et al.*, 1995).
- Por último, se ha observado que la asociación de las proteínas SR con determinadas secuencias intrónicas inhibe la formación del espliceosoma mediante impedimento estérico según se observa en la Figura 12C (Ibrahim *et al.*, 2005).



**Figura 12. Mecanismos de acción de las proteínas SR.** Los procesos A y B facilitan la unión de diferentes snRNPs que conforman el espliceosoma, mientras que el mecanismo C inhibe la unión de snRNPs por impedimento estérico.

### II.3.a.2. Factores implicados en la actividad de las proteínas SR

La actividad de las proteínas SR depende de dos factores:

- La localización concreta de las secuencias ESE o ISE en el transcrito. Se ha determinado que el aumento de la distancia a la que se encuentra el sitio de unión para las proteínas SR respecto a los sitios de *splicing* disminuye la probabilidad de que el exón sea incluido.
- El estado de fosforilación de los dominios SR. Los dominios RS de las proteínas SR pueden ser fosforilados por quinasas como SRPK1 y 2 (Aubol *et al.*, 2003). El nivel de fosforilación condiciona la especificidad en la unión a ARN y es determinante en la interacción con otras proteínas y esencial en la translocación al núcleo de proteínas SR (Gui *et al.*, 1994). También existen proteínas capaces de desfosforilar las proteínas SR hiperfosforiladas, como las fosfatasa PP1 y PP2. Esta actividad es fundamental en la regulación de variantes de *splicing* implicadas en el control de la apoptosis (Chalfant *et al.*, 2002).

### II.3.b. Proteínas hnRNP

Al igual que las proteínas SR, un grupo de ribonucleoproteínas heterogéneas nucleares o hnRNP reconocen, a través de sus dominios RRM, secuencias tanto intrónicas o exónicas (ISS, ESS) y actúan por lo general, reprimiendo el reconocimiento de los sitios de *splicing* por parte del espliceosoma (Bush y Hertel, 2012). Una excepción son las hnRNP E/K que interactúan con el ARN a través de dominios KH. También es frecuente encontrar en las hnRNPs otros dominios como la “caja” RGG que

contiene repeticiones Arg-Gly-Gly y otros dominios adicionales ricos en prolina o glicina de interacción con el ARN.

Debido a su gran diversidad, la clasificación de las hnRNP no está tan definida como la de las proteínas SR; de forma habitual su tipificación se limita a las hnRNP identificadas por Dreyfus y sus colaboradores (Choi y Dreyfuss, 1984), omitiendo otras hnRNPs de unión a ARN como las proteínas CELF, Fox, Nova o TDP-43 (Barreau *et al.*, 2006; Kuroyanagi, 2009; Hallegger *et al.*, 2010; Buratti y Baralle, 2010).

### **II.3.b.1. Mecanismos de actuación de las hnRNPs**

Después de su unión al ARN, las hnRNPs interactúan con otras proteínas modulando el *splicing*, pero el mecanismo concreto por el cual actúan sólo se conoce en un número reducido de casos. Podemos citar dos mecanismos básicos: mediante la represión estérica de las subunidades del espliceosoma en su intento de unión a los sitios de *splicing* (House y Lynch, 2006) y la creación de un bucle en las regiones exónicas (Martinez-Contreras *et al.*, 2006). Se sabe que algunas hnRNPs pueden activar o reprimir el reconocimiento de un exón dependiendo de su localización relativa, pero se desconoce el mecanismo por el que esto ocurre.

## **II.4. Otros factores implicados en la regulación del *splicing***

### **II.4.a. Factores de *splicing* específicos de tejido**

La expresión de reguladores de *splicing* concretos en diferentes tejidos, incluso en distintos tipos celulares o etapas concretas del desarrollo, da lugar a eventos de *splicing* alternativo específicos.

Los factores de *splicing* Nova1 y Nova2 han sido los primeros factores de *splicing* específicos de tejido descritos en mamíferos. Se expresan de forma concreta en las neuronas, en concreto Nova2 es específico de neurocórteX cerebral. Los factores Nova regulan el *splicing* de las proteínas implicadas en la inhibición neuronal y la plasticidad sináptica, como represores de neurotransmisores, proteínas implicadas en la liberación de neurotransmisores y otras proteínas relacionadas con la función y estructura de la sinapsis (Ule *et al.*, 2005).

Otro ejemplo son los reguladores de *splicing* epiteliales ESRP1 y ESRP2 los cuales son reprimidos transcripcionalmente durante la transición epitelio-mesénquima.

Estos factores promueven de forma mayoritaria la inclusión de exones o el reconocimiento de sitios 3' alternativos. De esta forma, favorecen eventos de *splicing* en genes que codifican proteínas relacionadas con la organización del citoesqueleto, la adhesión, polaridad y migración celular (Warzecha *et al.*, 2009).

#### II.4.b. Estructura secundaria

Los diagramas que han ido acompañando la lectura a través de las diferentes figuras y que representan diferentes reacciones en el proceso de *splicing* son una simple representación de la realidad. *In vivo* el pre-ARNm adopta complejas estructuras secundarias y terciarias pero su estudio es aún más complejo (Buratti y Baralle, 2004). Los algoritmos de predicción del plegado de las moléculas de ARN, como *mfold* y *evofold* (Zuker, 1989; Pedersen *et al.*, 2006), ayudan a identificar la estructura secundaria termodinámicamente más estable que puede conformarse. Combinando las predicciones estructurales de ARN con estudios de conservación de elementos relacionados con el *splicing*, Shepard y Hertel hallaron que los sitios de *splicing* alternativo estaban enriquecidos en estructuras secundarias conservadas (Shepard y Hertel, 2008). Otro estudio reciente ha determinado que los exones alternativos muestran un alto contenido en GC proporcionando estructuras de ARN estables que impiden la conformación del espliceosoma en sitios de *splicing* alternativos (Zhang *et al.*, 2011). Además, se ha descrito que las estructuras secundarias puntuales de ARNm pueden afectar al reconocimiento de elementos ESE y ESS por parte de proteínas SR o hnRNP, respectivamente (Hiller *et al.*, 2007). La estructura del ARNm también puede ser alterada por la unión de pequeños ARN sensibles a la presencia de metabolitos llamados *riboswitches*. Los *riboswitches* fueron descubiertos en el año 2002 (Winkler *et al.*, 2002); están localizados dentro del transcrito que ellos mismos regulan, y actúan *in cis* para controlar la expresión de la secuencia codificante dentro de un transcrito. Dicha regulación está mediada por la unión directa de tiamina pirofosfato (PPT) en los eucariotas. De esta forma participa en la regulación del *splicing* (Wachter, 2010).

#### II.4.c. Parámetros relacionados con la transcripción y la condensación de la cromatina.

##### Teoría cinética

En la Figura 13 podemos observar diferentes aspectos de la estructura de la cromatina y la dinámica de la transcripción que afectan al proceso de *splicing* y que serán descritos a continuación.

- **Secuencias promotoras:** Se han observado diferentes niveles de inclusión del exón 33 del gen de la fibronectina (*FN*) cuando es transcrito a partir de diferentes secuencias promotoras. Una explicación para este hecho es que diferentes promotores podrían reclutar complejos transcripcionales de composición variada que pueden interferir con el proceso de *splicing* (Cramer *et al.*, 1997).
- **Activadores transcripcionales:** Los activadores transcripcionales que se unen a las secuencias promotoras, como las proteínas CAPER pueden alterar los patrones de *splicing* de los ARNm producidos por la unidad transcripcional *downstream*, pero no se sabe si este efecto es directo o debido a la modificación de la cinética de la transcripción (Dowhan *et al.*, 2005).
- **Factores de remodelación de nucleosomas:** Los factores de remodelación de los nucleosomas que requieren enzimas como SWI/SNF para la activación de las secuencias promotoras pueden interferir en el *splicing* al unirse a subunidades del espliceosoma como snRNP U1 y U5 (Batsche *et al.*, 2006).
- **Densidad y estructura de la cromatina:** Las estructuras densas de cromatina dificultan la unión de la ARN polimerasa II al ADN y de esta forma se disminuye el ratio de transcripción y afecta al proceso de *splicing* según la **teoría cinética de la transcripción** (Luco *et al.*, 2010). Dicha teoría se basa en que, cuando la ARN polimerasa II trabaja a altos ratios de transcripción, allí donde se encuentre un exón alternativo entre dos constitutivos la competición entre el sitio 3' constitutivo más fuerte y el sitio 3' alternativo es a favor del constitutivo, siendo el exón alternativo no incluido. Mientras que cuando en la misma situación el proceso ocurre con un ratio de transcripción bajo, no existe competición entre los sitios 3' y el exón alternativo es incorporado con mayor facilidad (De la Mata *et al.*, 2003).
- **Señales extracelulares:** En el ejemplo anterior del gen *FN* se ha observado que siRNA dirigidos a regiones intrónicas *upstream* del exón 33, provocan la condensación de la cromatina y la inclusión del exón 33 (Allo *et al.*, 2009).

También estímulos despolarizantes que conducen a un estado abierto de la cromatina favorecen la inclusión del exón 8 en el gen *NCAM* (Schor *et al.*, 2009).

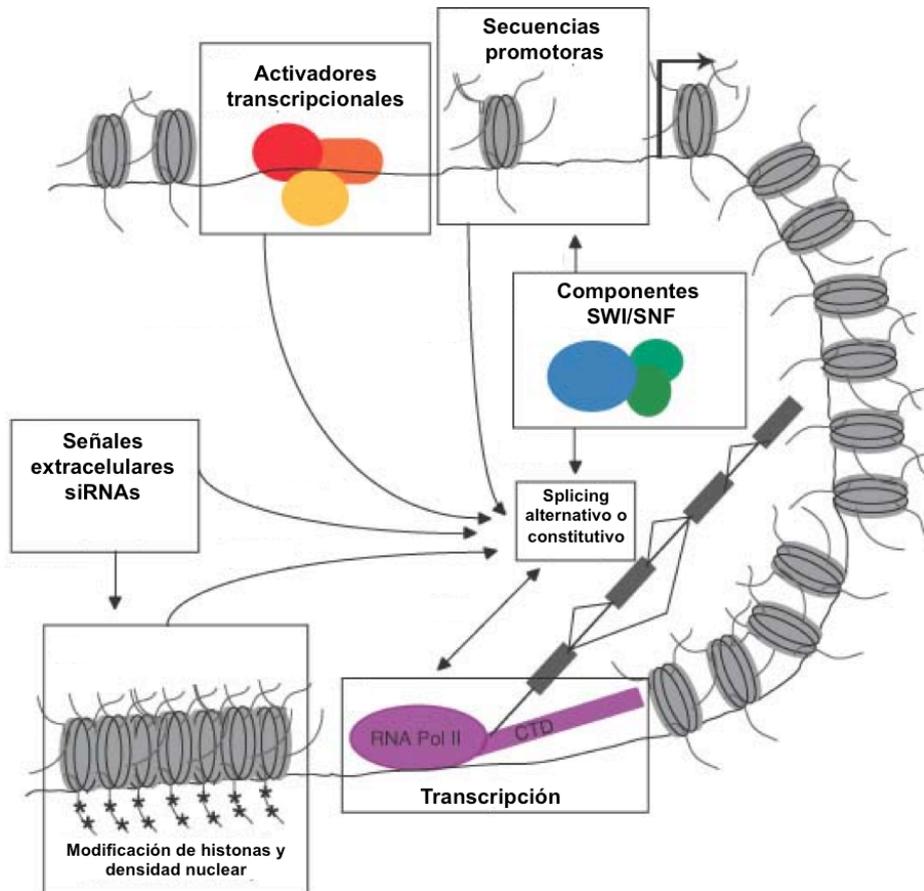


Figura 13. Parámetros relacionados con la transcripción y la condensación de la cromatina implicados en la regulación del *splicing* (Pandya-Jones, 2011).

### III. Alteraciones en el proceso de *splicing*

El proceso de *splicing* puede diferir entre las personas debido a la variación genética interindividual o a la aparición de mutaciones que pueden alterar tanto los elementos en *cis* como los reguladores de *splicing* en *trans*. Mientras que los efectos de una variante de ADN en *cis* se espera que sean locales (restringidos a un exón específico), en *trans* suelen ser globales, esto se debe a que estos reguladores de *splicing* controlan miles de eventos de *splicing* alternativo en el transcriptoma (Chen y Manley, 2009).

La generación de eventos de *splicing* aberrantes o la modificación del ratio de diferentes isoformas en un tejido concreto o estado del desarrollo son causa de

enfermedad y pueden ser determinantes tanto en la severidad como en la susceptibilidad hereditaria (Faustino y Cooper, 2003).

Según la base de datos *SpliceDisease* (<http://cmbi.bjmu.edu.cn/sdisease>) más de 300 genes contienen alguna mutación que afecta a su proceso normal de *splicing*. Estas mutaciones son las causantes de 370 enfermedades (Wang *et al.*, 2012). Se estima que al menos el 10% de las mutaciones que causan enfermedades hereditarias afecta a los sitios principales de *splicing* en las fronteras intrón-exón (Krawczak *et al.*, 2007). Parece que las alteraciones en las secuencias reguladores de *splicing* tienen un papel aún más relevante representando más del 60% según otros autores (López-Bigas *et al.*, 2005). No obstante, este porcentaje puede ser aún mayor, ya que las mutaciones que afectan al proceso de *splicing* han sido habitualmente ignoradas, porque tradicionalmente sólo se consideraban relevantes aquellas mutaciones con un efecto predecible sobre la proteína. Un estudio reciente ha revelado que al menos el 25% de las mutaciones *missense* y *nonsense* conocidas y relacionadas directamente con diferentes patologías alteran secuencias específicas de regulación del *splicing* dentro de los exones, lo que sugiere un papel mucho más relevante del *splicing* en relación con la etiopatogenicidad de las enfermedades (Sterne-Weiler *et al.*, 2011).

### III.1. Enfermedades provocadas por la alteración del *splicing* en *cis*

Las mutaciones en *cis* producen transcritos de ARNm aberrantes debido a la pérdida total o parcial de exones, o también a la retención parcial o total de intrones. Estas mutaciones afectan a las secuencias consenso que reconoce la maquinaria de *splicing* o a secuencias reguladoras de *splicing*. A continuación comentaremos algunos ejemplos de enfermedades provocadas por alteraciones de *splicing* en *cis*.

#### III.1.a. Alteración de los sitios de *splicing*. Síndrome QT-largo

El síndrome de QT-largo es una enfermedad que provoca taquicardias y está causada por anomalías estructurales en los canales de sodio/potasio del corazón. La mutación +6T>C *downstream* del sitio donador del exón 7 del gen *KCNH2* impide la unión de la subunidad del espliceosoma snRNP U1. Aunque cabría pensar que en este caso el exón 7 no fuera incluido en el ARNm maduro, ocurre que una secuencia ESE en el extremo 5' del exón 7 es reconocida por una proteína SR y junto con la ausencia de

snRNP U1 se produce la retención del intrón 7 y en consecuencia, la aparición de una proteína truncada y la enfermedad (Zhang *et al.*, 2004).

### III.1.b. Alteración de secuencias ESE. Síndrome de Zellweger

El síndrome de Zellweger es una enfermedad genética que afecta al desarrollo cerebral y además produce alteraciones en hígado y riñones. A nivel bioquímico se trata de una patología que afecta a la biogénesis de la peroxidasa debido a que uno de los trece genes relacionados con la formación del peroxisoma (*PEX*) resulta estar alterado. Un estudio de secuenciación masiva de todos los genes *PEX* en enfermos del síndrome de Zellweger detectó en cuatro pacientes la misma mutación en homocigosis en el exón 1 del gen *PEX6*. Se trata de la delección c.802\_815 que elimina una secuencia ESE de unión de la proteína SF2/ASF. En este caso la desaparición del ESE a la que se une SF2/ASF impide el reconocimiento del sitio natural de *splicing* y favorece el uso de sitios aceptores críptico dentro del exón 1 por parte de la maquinaria de *splicing*. Los resultados de RT-PCR confirmaron la presencia de dos transcritos que eliminan 264 y 657 nt del exón 1, produciendo proteínas inestables y como consecuencia, el clásico fenotipo del síndrome de Zellweger (Levesque *et al.*, 2012).

### III.1.c. Alteración de secuencias ESS. Distrofia muscular de Duchenne

La distrofia muscular de Duchenne es una enfermedad hereditaria recesiva ligada al cromosoma X, provocada por mutaciones en el gen de la distrofina que codifica una proteína encargada de conectar los filamentos de actina con la matriz extracelular en el músculo estriado. En algunos pacientes se ha detectado que la sustitución c.4250A>T provoca un codón de parada prematuro en la traducción del ARNm. Sin embargo, esta mutación en realidad crea un sitio ESS reconocido por la proteína hnRNPA1 y el resultado es que el exón 31 del gen no es incluido en el ARNm maduro (Disset *et al.*, 2006). Curiosamente, los ARNm que carecen del exón 31 conservan la pauta de lectura y codifican una proteína distrofina parcialmente funcional que esta asociada a una forma más leve de la enfermedad. Ejemplos como éste demuestran la necesidad de estudiar los efectos de cualquier mutación sobre el *splicing*.

### III.2. Enfermedades provocadas por la alteración del *splicing* en *trans*

La pérdida de funcionalidad de las proteínas implicadas en el *splicing* a través de mutaciones en sus genes codificantes o los cambios en la expresión de genes relacionados con la regulación del *splicing* puede causar la muerte prematura en neonatos o la aparición de enfermedades severas.

#### III.2.a. Alteración de la secuencia de genes implicados en el proceso de *splicing*

Las mutaciones que alteran la funcionalidad de las proteínas que forman parte del espliceosoma o de sus factores auxiliares, por lo general suelen ser letales. Sin embargo, mutaciones en proteínas, relacionadas con el ensamblado o la función del espliceosoma, que intervienen en la regulación específica del *splicing* en tipos celulares concretos provocan enfermedades asociadas a los mismos (Faustino y Cooper, 2003). Seguidamente se detallarán algunos ejemplos. A continuación mencionaremos tres ejemplos distintos de enfermedades causadas por este tipo de alteración del *splicing* en *trans*.

##### III.2.a.1. Retinitis pigmentaria

La retinitis pigmentaria es una enfermedad que en su variante de heredabilidad autosómica dominante afecta a multitud de genes implicados en funciones específicas de la retina. El resultado es la degeneración de las células fotorreceptoras del tipo bastón. En un 5% de los casos la causa de la enfermedad se debe a mutaciones en 5 genes (*PRPF31*, *PRPF8*, *HPRP3*, *PAP1* y *SNRNP200*), que codifican para proteínas esenciales en la formación de los complejos tri-snRNP tanto del espliceosoma mayor como menor impidiendo su actividad catalítica (Mordes *et al.*, 2007).

##### III.2.a.2. Atrofia muscular espinal

La Atrofia muscular espinal (SMA) es una grave enfermedad neurodegenerativa hereditaria que afecta principalmente a las motoneuronas  $\alpha$  del asta anterior de la médula espinal. Aunque se ha descrito la existencia de heterogeneidad genética, está causada generalmente por mutaciones en el gen de la supervivencia de motoneuronas o *SMN1* (>90% casos se deben a deleciones completas del gen) perdiendo su actividad

funcional. Este gen está implicado en el ensamblado de las snRNP del espliceosoma. Curiosamente *SMN2*, un gen parálogo de *SMN1*, difiere de éste en varios cambios de nucleótido, uno de ellos consiste en una transición de citosina a timina en el sexto nucleótido del exón 7 que provoca la pérdida de este exón en la mayor parte de los transcritos maduros. Este evento ha sido atribuido tanto a la pérdida de una secuencia ESE reconocida por SF2/ASF (Cartegni y Krainer, 2002) como a la creación de un ESS a la que se une la proteína hnRNPA/B (Kashima y Manley, 2003). Numerosas investigaciones se están centrando en alterar el *splicing* de *SMN2* para inducir la inclusión del exón 7 y restaurar los niveles de la proteína SMN en los pacientes afectador por esta enfermedad (Wirth *et al.*, 2006).

### III.2.a.3. El síndrome de Taybi-Linder

El enanismo primordial osteodisplásico tipo I (MOPD I) o síndrome de Taybi-Linder es una rara displasia ósea autosómica recesiva y es la primera enfermedad descrita que está provocada por mutaciones en una snRNP del espliceosoma y en concreto por defectos funcionales del espliceosoma menor. El descubrimiento fue realizado por dos grupos de investigadores, que analizaron mediante secuenciación masiva el genoma de diferentes pacientes y localizaron ocho mutaciones diferentes en una región muy pequeña en un intrón del gen *CLASP1*, en el cromosoma 2. Curiosamente, este intrón contiene el gen que codifica el snRNP U4atac, un componente esencial del espliceosoma menor, que es transcrito en dirección opuesta al gen *CLASP1*. Todas las mutaciones menos una se localizaron en una región del ARNm de U4atac que forma una estructura secundaria de ARN de doble cadena muy conservada en la evolución, sugiriendo que esta región tiene una función esencial. Estudios funcionales *in vivo* comprobaron que estas mutaciones afectaban únicamente al *splicing* de los intrones dependientes del espliceosoma menor, viendo disminuida su actividad en más del 90% (He *et al.*, 2001; Ederly *et al.*, 2011).

### III.2.b. Alteraciones en la expresión de los genes relacionados con el *splicing*

Otro mecanismo de actuación en *trans* es la variación de la expresión de genes que codifican proteínas relacionadas con el *splicing*. La sobreexpresión de elementos reguladores de *splicing* provoca alteraciones en los ratios de transcritos en genes que

sufren *splicing* alternativo específico de tejido. Estos cambios pueden estar inducidos de forma interna, por alteración de los propios mecanismos celulares, por ejemplo en cáncer, el cual trataremos de forma amplia en el siguiente apartado o, como veremos a continuación, a través de agentes externos infecciosos.

### **III.2.b.1. Enfermedades infecciosas que alteran la expresión de factores de *splicing*: SIDA**

Una enfermedad distinta a las comentadas con anterioridad, puesto que no es causada por la alteración del ADN de la persona que la padece, es el síndrome de inmunodeficiencia adquirida. El SIDA es una enfermedad provocada por la infección del retrovirus HIV-1. El genoma del HIV-1 está formado por un pre-ARNm que contiene diferentes sitios alternativos de *splicing* 3' y 5' con el que consigue generar más de cuarenta ARNm maduros (Stoltzfus y Madsen, 2006). Durante la infección, el VIH-1 provoca cambios en los niveles de expresión de proteínas reguladoras de *splicing* como las SR, proteínas de la célula huésped que son utilizadas por el virus para obtener gran cantidad de transcritos a partir de un solo pre-ARNm (Dowling *et al.*, 2008). Por ejemplo, la proteína SF2/ASF reconoce una secuencia ESE en el exón 5 del virus promoviendo su inclusión en el ARNm maduro. Debido a la importancia del proceso de *splicing* en la infección del VIH-1, actualmente se están desarrollando terapias dirigidas a la inhibición selectiva de proteínas SR (Bakkour *et al.*, 2007).

### **III.3. Nonsense Mediated Decay**

Una fracción de las alteraciones que tienen como resultado el *splicing* aberrante provoca la aparición de codones de parada prematuros en los transcritos (PTCs: Premature Termination Codons). En la célula existe un mecanismo regulador que se encarga de degradarlos. El *Nonsense Mediated Decay* (NMD) está altamente conservado en la evolución e impide la aparición de proteínas truncadas generadas a partir de ARNm que contienen PTCs que proceden tanto por mutaciones *nonsense* o *frameshift* como por mecanismos de *splicing* (Huang y Wilkinson., 2012). Su actividad varía entre los diferentes tipos celulares. Se ha podido comprobar que en células HeLa la eficiencia es mucho más alta que en células MCF7, por lo que el NMD podría tener un papel relevante en la regulación de la expresión génica (Linde *et al.*, 2007).

### III.4. *Splicing* y cáncer

Las alteraciones en el proceso de *splicing* en genes relacionados con el ciclo celular (migración, división y crecimiento celular, reparación de ADN, apoptosis,) están directamente relacionadas con el cáncer (Pajares *et al.*, 2007). Cuando estos genes contienen mutaciones en secuencias reguladoras de *splicing* o existen cambios en la expresión de proteínas que intervienen en la regulación del *splicing* de los mismos, pueden aparecer proteínas truncadas, deficientes, con una nueva funcionalidad o disminuir la abundancia relativa de la proteína funcional. En consecuencia, puede surgir un proceso de desarrollo y progresión tumoral. La detección de variantes específicas de *splicing* puede contribuir tanto al diagnóstico temprano de la enfermedad como a la identificación de nuevas dianas terapéuticas (Kim y Kim., 2012).

#### III.4.a. Cáncer y mutaciones en *cis* que alteran el *splicing*

A pesar de que un gran número de mutaciones ha sido detectado en genes implicados en cáncer, como los genes supresores de tumores, los programas de rastreo de mutaciones en estos genes se ha centrado en la detección de mutaciones exónicas con alteración directa en la proteína, quedando ignoradas mutaciones intrónicas y exónicas que pueden afectar a elementos reguladores de *splicing* a los sitios naturales de *splicing*. Actualmente se han acumulado muchos ejemplos de mutaciones asociadas al proceso de *splicing* en múltiples genes implicados en tumorigénesis.

- El cambio de nucleótidos G>A en el nucleótido +1 del sitio donador del exón 7 provoca que éste no sea incluido en el ARNm maduro procedente del gen *hSNF5* que codifica una subunidad del complejo de remodelación de la cromatina dependiente de ATP y al que se ha caracterizado como un gen supresor de tumores. La pérdida del exón 7 es responsable de la aparición de una proteína truncada que causa tumores cerebrales infantiles (Taylor *et al.*, 2000).
- Una mutación parecida en el sitio aceptor del exón 4 del gen *APC* en el que ocurre un cambio AG>AT provoca que el exón 4 no sea reconocido. Esta mutación está relacionada con el desarrollo de cáncer de hígado y colorrectal (Kurahashi *et al.*, 1995).
- Una mutación relacionada con cáncer de mama se ha encontrado en el gen que codifica

el receptor de estrógeno. La mutación intrónica A>G crea un sitio críptico de *splicing* provocando la inserción de 69 nucleótidos en el ARNm maduro (Wang *et al.*, 1997).

- En el gen de la Neurofibromatosis tipo II (*NF2*), relacionado con tumores del sistema nervioso central se ha detectado una mutación que crea un punto de ramificación alternativo y provoca el reconocimiento de un exón críptico que es incluido en ARNm maduro (De Klein *et al.*, 1998).
- Además también se han detectado mutaciones en secuencias ESS y ESE relacionadas con el cáncer. Por ejemplo, secuencias consenso para proteínas SR (SC35 Y SF2/ASF) han sido alteradas por una doble mutación en medio de un exón del gen de la Neurofibromatosis tipo I (*NF1*) provocando un alto nivel de exclusión de ese exón en los transcritos (Colapietro *et al.*, 2003).

### III.4.b. Expresión diferencial de factores de *splicing* y su implicación en cáncer

#### III.4.b.1. Cáncer y expresión de los proteínas SR

El factor de *splicing* 1 (SRSF1 o SF2/ASF) tiene actividad oncogénica, puesto que se encuentra regulado positivamente en muchos tipos de tumores. La sobreexpresión de *SRSF1* está ligada a la alteración de los patrones de *splicing* de importantes oncogenes y genes supresores de tumores. El resultado es una inhibición de la apoptosis y el incremento de la tasa de crecimiento celular. A continuación se citan varios ejemplos seleccionados de la bibliografía (Kaida *et al.*, 2012):

- *BIN1* es un gen supresor de tumores que interactúa con el protooncogén *MYC* suprimiendo su actividad. Se ha observado que en el melanoma, las células sobreexpresan SRSF1 promoviendo la inclusión de un exón alternativo 12A en el ARNm de *BIN1* (Figura 14A). La inclusión de este exón alternativo hace que BIN1 tenga menos afinidad por MYC y disminuya su actividad antitumoral (Karni *et al.*, 2007).
- Además distintas proteínas SR pueden intervenir sobre un mismo pre-ARNm. El resultado son diferentes isoformas de *splicing* con funciones opuestas (Figura 14B). Un ejemplo relacionado con el cáncer implica al factor de crecimiento endotelial vascular (VEGF). El término VEGF engloba a dos familias de proteínas (VEGFxxx y VEGFxxx) procedentes del *splicing* alternativo de un solo gen *VEGF* de 8 exones. Esta clasificación se basa en que el gen *VEGF* contiene dos posibles sitios 3' de *splicing* en el exón 8. La selección de estos sitios está regulada por proteínas SR y va

a determinar la función de VEGF. La sobreexpresión de SRF51 y SRp40 implica el reconocimiento del sitio proximal y que la familia de proteínas resultante sea pro-angiogénica (VEGFxxx). En cambio, SRp55 promueve la selección del sitio de *splicing* distal produciendo isoformas con propiedades anti-angiogénicas (Nowak *et al.*, 2008).

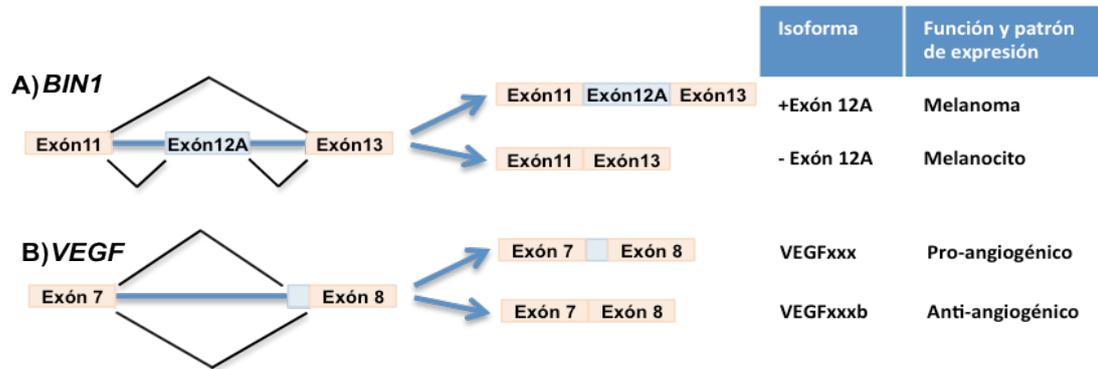


Figura 14. Actividad oncogénica de los genes *BIN1* y *VEGF* regulada por sobreexpresión de proteínas SR.

### III.4.b.2. Cáncer y expresión de proteínas hnRNP

La falta de regulación en las proteínas hnRNP también contribuye a la carcinogénesis a través de múltiples vías. Muchas hnRNPs están involucradas en la regulación de múltiples genes relacionados con el ciclo celular. Por ejemplo la hnRNP H se encuentra sobreexpresada en células de glioma y actuaría como un oncogén modulando el *splicing* alternativo del gen *RON* que codifica un receptor con actividad tirosina quinasa cuyo ligando es la proteína estimuladora de macrófagos (MSP) y relacionado con la movilidad, disociación e invasión celular. La hnRNP H se une a una secuencia ESS del exón 11 de *RON* impidiendo que éste forme parte del ARNm maduro. Esta isoforma de *RON* induce la invasión y migración celular (Ghigna *et al.*, 2005).

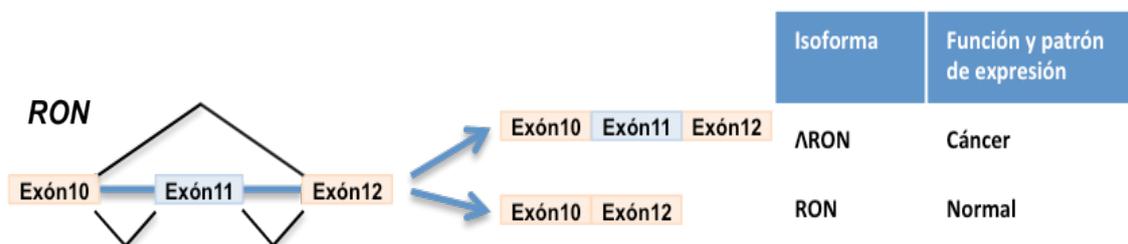


Figura 15. Actividad oncogénica de *RON* regulada por la sobreexpresión de hnRNP H.

### III.4.b.3. Otros factores de *splicing* relacionados con el cáncer

Existen ejemplos de otros factores relacionados con el cáncer y que actúan a través de la regulación del *splicing*. Por ejemplo *Sam68* es un miembro de la familia de proteínas STAR implicadas en la transducción de señales y la activación del ARN. *Sam68* es considerado un potencial supresor de tumores, puesto que interfiere en el *splicing* del regulador de la apoptosis *Bcl-X*. *Bcl-X* tiene dos sitios donadores alternativos en el exón 2. La isoforma de ARNm de mayor tamaño de *Bcl-X* es antiapoptótica (*Bcl-XL*) mientras que la de menor tamaño es pro-apoptótica (*Bcl-XS*). *Sam68* se une al ARNm del *Bcl-X* favoreciendo el reconocimiento del sitio 5' que tiene como resultado la isoforma *Bcl-XS* (Figura 16). Se ha observado que la inactivación de *Sam68* conduce a una transformación neoplásica y su sobreexpresión conlleva la detención del ciclo celular y la apoptosis (Paronetto *et al.*, 2007).

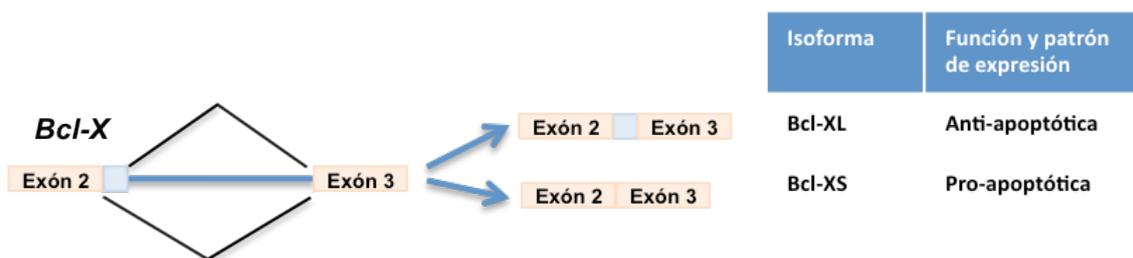


Figura 16. Actividad apoptótica de *Bcl-X* regulada por la sobreexpresión de *Sam68*.

## IV. El síndrome de cáncer de mama y ovario hereditario (CMOH)

El cáncer de mama es el tumor maligno más frecuente en occidente y la principal causa de muerte por cáncer en las mujeres. Se ha estimado que una de cada diez mujeres padece cáncer de mama a lo largo de su vida. Diferentes estudios epidemiológicos han demostrado que el riesgo de una mujer, que tiene familiares en primer grado que han padecido cáncer de mama, es el doble en comparación con la población normal. Aunque este incremento del riesgo puede atribuirse a factores ambientales, estudios realizados en gemelos han indicado que este aumento en la predisposición se debe a factores hereditarios (Stratton y Rahman, 2008). Hoy sabemos que entre el 5% y el 10% de los casos de cáncer de mama son de origen hereditario (Apostolou y Fostira, 2013).

#### IV.1. Principales genes de predisposición a CMOH: *BRCA1* y *BRCA2*

En los años noventa diferentes grupos identificaron, mediante estudios de ligamiento genético, los dos principales genes de predisposición a cáncer de mama y ovario hereditario *BRCA1* y *BRCA2* (Miki *et al.*, 1994; Wooster *et al.*, 1995). Ambos están implicados en la reparación de roturas de doble hebra en el ADN (Gudmundsdottir y Ashworth, 2006). Las mutaciones deletéreas en los genes *BRCA* multiplican el riesgo de padecer cáncer de mama en un rango que oscila entre 10 y 30 veces. Estos datos implican que el riesgo vital de padecer la enfermedad es del 85%, comparado con un 3% en la población general. Además, las mutaciones en estos genes también están ligadas a un incremento en el riesgo de padecer cáncer de ovario; éste, aunque menor, es bastante elevado y se ha estimado que puede alcanzar el 50%. Otros tipos de cáncer como el de próstata y el pancreático están asociados a mutaciones en los *BRCA* multiplicando por 4 el riesgo relativo de padecer estas patologías (Foulkes, 2008).

La predisposición a CMOH se transmite como un rasgo autosómico dominante a través de las familias portadoras de mutaciones en estos genes. A nivel celular, *BRCA1* y *BRCA2* actúan como genes recesivos (Gudmundsdottir y Ashworth, 2006). Lo más frecuente es la aparición de mutaciones en heterocigosis, siendo necesaria la inactivación del alelo normal a nivel somático para el desarrollo del cáncer. La aparición de mutaciones en homocigosis por ejemplo en *BRCA2*, da lugar a síndromes concretos como la anemia de Fanconi (Howlett *et al.*, 2002).

#### IV.2. Espectro de predisposición en CMOH

Las mutaciones en los genes *BRCA1* y *BRCA2* son únicamente responsables del 16% de los casos de CMOH. Mutaciones en otros cuatro genes implicados en síndromes concretos (*TP53*-síndrome de Li-Fraumeni, *PTEN*-síndrome Cowden, *STK11*-síndrome de Peutz-Jeghers y *CDH1*-cáncer gástrico difuso hereditario) están asociadas a un elevado riesgo de padecer cáncer de mama. Estas mutaciones son raras y en conjunto representan un pequeño incremento en el espectro de predisposición (Thompson y Easton, 2004).

La pregunta que de forma inmediata surge es cómo explicar el 80% de casos restantes con un fuerte historial familiar que implica un riesgo familiar de CMOH. Dos aproximaciones diferentes han ayudado a dilucidar parte del resto del espectro de predisposición.

La primera se ha centrado en buscar genes relacionados con la vía funcional en la que participan los genes *BRCA*. Esta estrategia ha identificado diferentes genes mutados que confieren un riesgo moderado de padecer la enfermedad y su frecuencia de aparición es muy baja. Entre ellos podemos citar *ATM* (Renwick *et al.*, 2006), *BRIP1* (Seal *et al.*, 2006), *CHEK2* (Meijers-Heijboer *et al.*, 2002), *PALB2* (Erkko *et al.*, 2007), *RAD51C* (Somyajit *et al.*, 2010) o de forma más reciente *Abraxas* (Solyom *et al.*, 2012), *XCCR2* (Park *et al.*, 2012) o *RAD51D* (Osher *et al.*, 2012), entre otros. Al igual que se ha mencionado en el caso de *BRCA2*, mutaciones bialélicas en *BRIP1* (Levitus *et al.*, 2005) y *PALB2* (Reid *et al.*, 2007) causan raros subgrupos de anemia de Fanconi donde el riesgo de padecer cáncer de mama es elevado. También, mutaciones en ambas copias del gen *ATM* provocan ataxia talangiectasia (Renwick *et al.*, 2006). Sin embargo no hay evidencias de ello en *CHEK2* o *BRCA1*. De hecho, las mutaciones en los dos alelos del gen *BRCA1* parecen ser letales durante el proceso de desarrollo embrionario (Haken *et al.*, 1996).

En la segunda aproximación se han realizado estudios genómicos de asociación alélica para identificar SNP asociados a cáncer de mama y ovario hereditario. En estos estudios se han identificado más de veinte variantes comunes, cuya frecuencia de aparición en la población oscila entre el 5 y el 50% (Rosa-Rosa *et al.*, 2009). Dichos SNPs multiplican por 1,5 el riesgo de padecer CMOH y en conjunto se estima que contribuyen a un 4% del espectro de predisposición (Fanale *et al.*, 2011).

### **IV.3. Modelo poligénico de predisposición a CMOH**

Aunque el desarrollo de nuevas tecnologías como la ultrasecuenciación de ADN está contribuyendo al descubrimiento de genes relacionados con CMOH, el espectro de predisposición de esta enfermedad no ha sido todavía completado. El carácter poligénico de la enfermedad y los datos actualmente disponibles sugieren que la coincidencia de diferentes alelos de baja predisposición en una misma paciente actuarían de forma multiplicativa sobre el riesgo de padecer CMOH (Ponder *et al.*, 2005).

#### IV.4. Estructura de los genes *BRCA1* y *BRCA2*

En 1990, se realizó la localización del gen *BRCA1* en el brazo largo del cromosoma 17 mediante análisis de ligamiento en familias con un fuerte historial en cáncer de mama y ovario (Hall *et al.*, 1990). Cuatro años más tarde fue clonado para su caracterización por científicos de la compañía Myriad Genetics (Miki *et al.*, 1994). *BRCA1* es un gen que consta de una secuencia genómica de 81,19 kb, contiene 23 exones, 22 de ellos codificantes. La transcripción de *BRCA1* genera un ARNm de 7094nt que se traduce en una proteína de 1863 aminoácidos y un peso molecular de 220 KDa. A nivel de proteína podemos distinguir varios dominios funcionales (Figura 17):

- En el extremo amino terminal aparece un dominio RING que interactúa con el dominio RING de la proteínas BARD1 y una secuencia de localización nuclear (NLS).
- La región central de *BRCA1* contiene un sitio de fosforilación S988 mediado por la proteína CHEK2.
- Por su parte, en el extremo carboxilo terminal podemos encontrar un dominio helicoidal de interacción con el complejo formado por las proteínas PALB2 y *BRCA2*, un dominio SCD (SQ/TQ *domain*) que presenta hasta diez sitios de fosforilación por ATM y por último un dominio BRCT que proporciona diferentes funciones a *BRCA1* en función de su interacción con una de estas tres proteínas: Abraxas, CtIP o BRIP1.

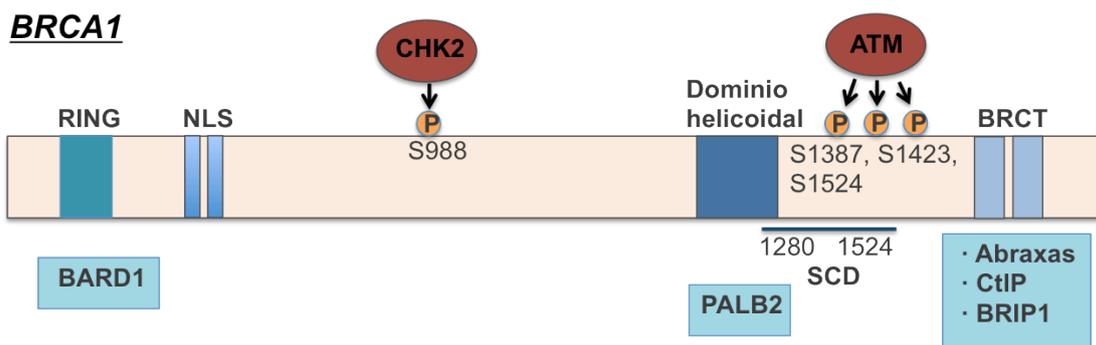


Figura 17. Dominios funcionales de la proteína *BRCA1*.

El gen *BRCA2* se localiza en el brazo largo del cromosoma 13, su tamaño es de 84,19 kb y consta de 27 exones, 26 de ellos codificantes. La transcripción de *BRCA2* genera un ARNm de 10930 nt que se traduce en una proteína de 3418 aminoácidos y

384 KDa de peso molecular. Al igual que la proteína BRCA1, BRCA2 tiene varios dominios funcionales de interacción con otras proteínas, esenciales para sus funciones.

- En el extremo amino terminal existe una secuencia de aminoácidos de interacción con la proteína PALB2.
- En la región central existen ocho motivos de repetición BRC esenciales para la unión de la proteína recombinasa RAD51. Además contiene un dominio de unión a ADN formado por un dominio helicoidal (H), tres dominios de unión a oligonucleótidos (OB) y un dominio en torre (T) que facilita la unión de BRCA2 a ADN mono y bicatenario.
- Por último, en el extremo carboxilo terminal contiene señales de localización nuclear y un sitio de fosforilación dependiente de quinasas (CDK2) al que también se une RAD51.

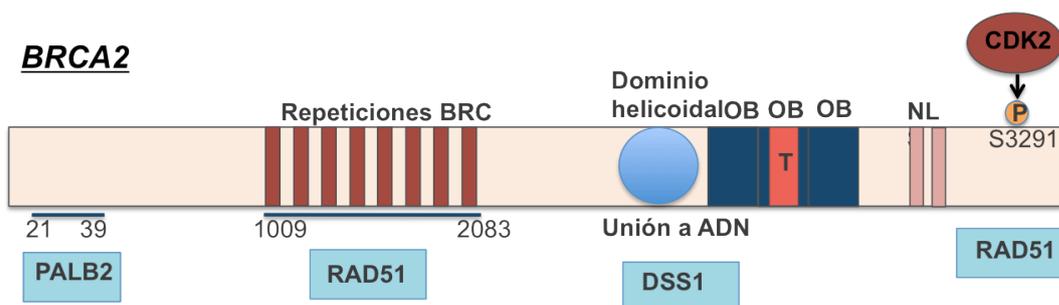


Figura 18. Dominios funcionales de la proteína BRCA2.

#### IV.5. BRCA1 y BRCA2 en la protección de la integridad del genoma

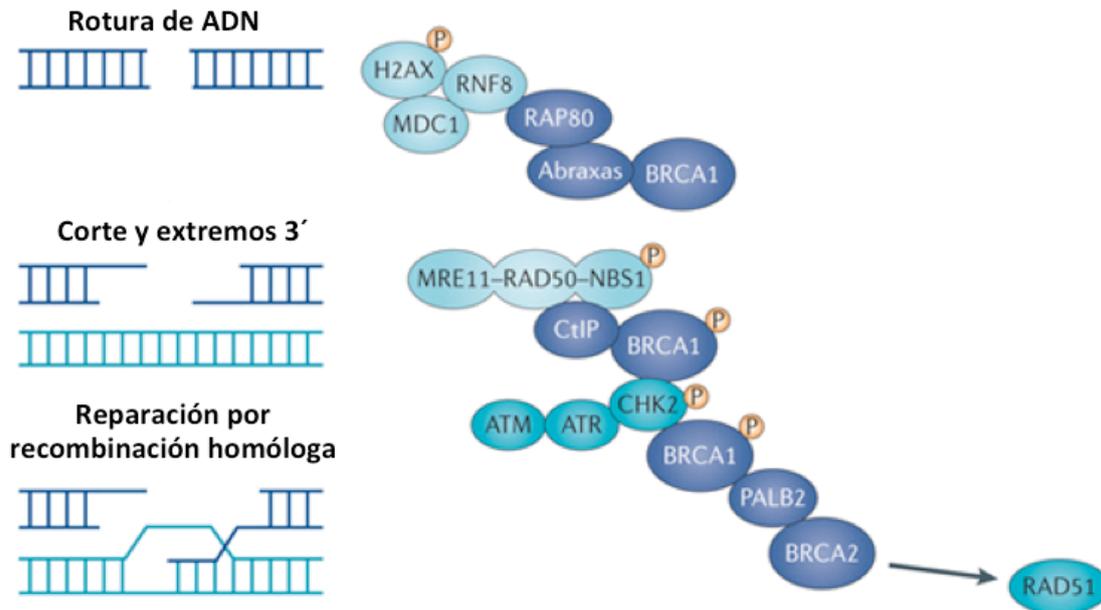
Los dos principales genes de predisposición a CMOH están implicados en el mecanismo más importante de protección de la integridad del genoma durante la proliferación celular: la recombinación homóloga.

La recombinación homóloga es un proceso fundamental en la reparación de roturas de doble hebra en el ADN (DSBs: Double-Strand Breaks). Las DSBs son el peor daño que puede ocurrir en el ADN; pueden aparecer durante el proceso de replicación o por explosión a radiación ionizante u otros agentes genotóxicos. Aunque las DSBs pueden ser reparadas por otros mecanismos, el proceso de recombinación homóloga es el único que no introduce errores en la secuencia de nucleótidos (Roy *et al.*, 2012).

El proceso de reparación de DBSs consta de tres pasos bien diferenciados. En primer lugar diferentes sensores detectan el daño, de forma consecutiva; los

mediadores de la señalización reclutan efectores que, por último, repararán el daño y activan puntos de control del ciclo celular. Como podemos apreciar en la Figura 19, BRCA2 es esencial en la fase de reparación del daño, mientras que BRCA1 está involucrado en varios macrocomplejos que son mediadores clave en todas las fases de la reparación por recombinación homóloga.

- El complejo formado por BRCA1-Abraxas-RAP80 se asocia con histonas ubiquitinadas cerca de los sitios de ADN dañados. Esta acción depende de la fosforilación de la histona H2AX y el complejo que forma con MDC1 y RNF8 (Wang *et al.*, 2007).
- Una vez señalizada la DBSs, otro complejo en el que participa BRCA1 junto con CtIP se une al complejo MRN formado por las proteínas MER11-RAD50-NBS1 cuya función es la de escindir parte de los extremos del punto de rotura y producir una pequeña elongación de los extremos 3' de ambas cadenas (Yun y Hiom, 2009).
- La última etapa se completa cuando RAD51 se une a los extremos de ADN de cadena sencilla para catalizar el proceso de recombinación. Este proceso está regulado por BRCA2 (Thorslund *et al.*, 2010) a través de su interacción con PALB2 y otros complejos adicionales en los que participa CHEK2 y BRCA1 (Oliver *et al.*, 2009). Diferentes parálogos de RAD51 (RAD51B, RAD51C, RAD51D, y RAD52) son capaces de encontrar, a lo largo de todo el genoma, secuencias homólogas a las del punto de rotura donde se realizará la recombinación.



**Figura 19. Complejos proteicos implicados en la reparación de DBSs mediante recombinação homóloga.** En azul claro se muestran los complejos de señalización del daño en el ADN, en azul oscuro, los mediadores de señalización que activan los efectores.

Las proteínas mencionadas en esta sección en relación a la reparación de DBSs mediante recombinação homóloga son aquellas codificadas por los genes cuyas mutaciones son responsables, en mayor o menor grado, del espectro de predisposición a CMOH (ver sección 4.3.). No existe aún una explicación que responda a por qué los daños en estos genes producen de forma específica cáncer de mama y ovario. La teoría más aceptada sugiere que en las células epiteliales de mama y ovario se produce un mayor estrés oxidativo debido a fuertes señales de crecimiento mediadas por estímulo hormonal durante el ciclo menstrual. El estrés oxidativo provoca daños en el ADN e incrementa la presencia de mecanismos de reparación como la recombinação homóloga que cuando fallan aumentan el riesgo de padecer CMOH (Roy *et al.*, 2012).

#### IV.6. El espectro mutacional en los genes BRCA

En la actualidad existen diferentes bases de datos donde se registran las diferentes variaciones de ADN detectadas en los genes *BRCA*s: BIC (<http://research.nhgri.nih.gov/bic/>), LOVD (<http://brca.iarc.fr/LOVD>) o HGMD (<http://www.hgmd.cf.ac.uk/>) son algunos ejemplos. La base de datos BIC es específica de los genes *BRCA*s, la más antigua y la que mayor número de entradas tiene. En el mes de diciembre de 2012 hay anotadas 3.765 variantes distintas, 1771 *BRCA1* y 1994

en *BRCA2*. La mayor parte de las mutaciones registradas son cambios de nucleótido pero también aparecen algunas pequeñas deleciones e inserciones. Las mutaciones han sido clasificadas en función del efecto que ejercen durante la traducción a proteína. Se pueden encontrar mutaciones que producen un cambio en el pauta de lectura (*Frameshift*), un codón de parada prematuro (*Nonsense*), cambio de aminoácido (*Missense*), sinónimas si no producen cambio de aminoácido, inserciones o deleciones que conservan la pauta de lectura (*Ins/Del in frame*), mutaciones intrónicas, en la región UTR y por último aquellas que alteran el proceso de *splicing*. En número concreto de mutaciones registradas para cada uno de estos efectos y su importancia clínica se muestra en la Tabla 1.

**Tabla 1. Importancia clínica y clasificación de las mutaciones incluidas en la base de datos BIC.**

Tipo de mutación	BRCA1				BRCA2			
	Importancia clínica			Total	Importancia clínica			Total
	si	Desconocida	No		Si	Desconocida	No	
<i>Frameshift</i>	577	5	0	582	584	6	0	590
<i>Nonsense</i>	195	2	1	198	193	2	0	195
<i>Missense</i>	19	566	18	603	9	845	32	886
<i>Splice</i>	10	33	4	47	3	1	0	4
IVS	91	160	8	259	43	146	3	192
Sinónimas	2	39	9	50	2	55	27	84
<i>Ins/Del In frame</i>	1	24	0	25	0	32	0	32
UTR	0	6	1	7	0	8	3	11
<b>Total</b>	<b>895</b>	<b>835</b>	<b>41</b>	<b>1771</b>	<b>834</b>	<b>1095</b>	<b>65</b>	<b>1994</b>

Según los datos globales mostrados en la tabla 1, el 46% de las mutaciones registradas tienen importancia clínica. En su gran mayoría son mutaciones *frameshift* y *nonsense*. De *splicing* un 0,3%, que puede elevarse al 4% si tenemos en cuenta que todas las mutaciones clasificadas como intrónicas y de importancia clínica son mutaciones que alteran los nucleótidos con mayor conservación en sitios donadores y aceptores de *splicing* (AG-GT). Además, existe un 51% de las variantes de ADN anotadas de las que se desconoce su significancia clínica, la gran mayoría *missense* (Tabla 1). En un estudio *in silico* se detectó que en el 60% de las secuencias ESE altamente conservadas en el gen *BRCA1* se han descrito mutaciones *missense*, por lo que podrían tener un efecto en el desarrollo de CMOH a través de la alteración del *splicing* (Pettigrew *et al.*, 2005). También, existen ejemplos en el bibliografía que confirman que mutaciones

previamente catalogadas en relación a su efecto directo en la proteína afectan al *splicing*. A continuación se muestran algunos ejemplos relacionados con el gen *BRCA2*, puesto que es el objeto de estudio de este trabajo:

- La mutación c.145C>T en el exón 3 del gen *BRCA2* se encuentra registrada en la base de datos BIC y está clasificada como *nonsense*. Se ha demostrado que realmente afecta al *splicing* y provoca que en un alto porcentaje de transcritos no sea incluido el exón 3. Según las estimaciones bioinformáticas, la mutación elimina una secuencia ESE a la que se uniría la proteína SF2/ASF. El resultado no es una proteína truncada por un codón de stop prematuro, sino una delección en fase de lectura que afecta al dominio de unión a PALB2 del extremo N- terminal de *BRCA2* (Sanz *et al.*, 2010).
- Otro ejemplo es la mutación c.470\_474del5 en el exón 5 del gen *BRCA2* registrada en la base de datos del BIC y clasificada como *frameshift*. Esta mutación afecta al *splicing* produciendo la exclusión mayoritaria del exón 5 debido a la eliminación de una secuencia ESE (Sanz *et al.*, 2010).
- La mutación c.8331G>A es una mutación del exón 18 de *BRCA2* y está clasificada como sinónima de efecto clínico desconocido. Ensayos funcionales de *splicing* han demostrado que altera el sitio donador del exón 18 provocando que éste no sea incluido en el transcrito maduro. La alteración en el proceso de *splicing* provoca una proteína truncada (Sanz *et al.*, 2010).
- Se han demostrado que mutaciones clasificadas como *missense* y de efecto clínico desconocido, por ejemplo la c.517G>T en el exón 7 de *BRCA2*, alteran el *splicing*. En este ejemplo concreto se ha descrito que un porcentaje significativo de transcritos no incluye el exón 7 debido a la alteración que esta mutación provoca en el sitio aceptor de este exón (Gaildrat *et al.*, 2012).
- Las mutaciones intrónicas que afectan a los sitios donador y aceptor de *splicing* (GT-AG) son prácticamente catalogadas como de importancia clínica porque afectan al proceso de *splicing*, pero mutaciones fuera de esos sitios no. No obstante, mutaciones intrónicas profundas como la c.6937+594T>G en el intrón 12 del gen *BRCA2* afectan al *splicing* activando sitios crípticos de *splicing* dentro del intrón. En este caso particular la mutación provoca que se incluya un pseudoexón en el transcrito maduro (Anczukow *et al.*, 2012). Otras secuencias intrónicas más cercanas a los exones y claves en el proceso de *splicing*, como el tracto de polipirimidina,

también pueden ser afectadas. Por ejemplo, es el caso de la mutación c.68-7T>A que afecta al tracto de polipirimidina del intrón 2 del gen BRCA2 y promueve que el exón 3 no sea incluido en parte de los transcritos (Sanz *et al.*, 2010).

En el intento por determinar la patogenicidad de una variante que afecta al proceso de *splicing* se debe tener cuenta no sólo el efecto que producen las isoformas aberrantes en la traducción a proteína, sino también el porcentaje de transcrito alterado respecto al natural y otros factores como la presencia de la variante en la población control y la cosegregación de la mutación con la enfermedad dentro de una misma familia.

En el presente trabajo se analizó el impacto de variantes de ADN en el proceso de *splicing*. Para ello se diseñó una estrategia combinada consistente en la utilización de diversas herramientas bioinformáticas junto con ensayos funcionales de ARN de linfocitos y/o minigenes híbridos. Asimismo, se han aplicado novedosas soluciones tecnológicas como la creación de un nuevo vector de *splicing* pSAD o la cuantificación relativa de isoformas de *splicing* a través de la detección de transcritos mediante electroforesis capilar. Con todo, se intentó determinar la patogenicidad de numerosas mutaciones detectadas en el gen *BRCA2* en pacientes con CMOH. Por último, el conocimiento generado ha contribuido a elucidar diferentes mecanismos de regulación esenciales para el correcto *splicing* de este gen.

# OBJETIVOS



- I. Análisis del efecto sobre el *splicing* de variantes de ADN del gen *BRCA2* detectadas en pacientes CMOH procedentes del Programa de Prevención de Cáncer Hereditario de Castilla y León y reportadas en las bases de datos internacionales.
- II. Selección de variantes candidatas con impacto en el *splicing* mediante análisis bioinformático con los programas disponibles.
- III. Diseño y validación de un nuevo vector de *splicing* para la realización de ensayos funcionales mediante minigenes híbridos.
- IV. Construcción de minigenes híbridos del gen *BRCA2* de predisposición a CMOH.
- V. Análisis del efecto de las mutaciones de los sitios de *splicing* y otros elementos reguladores (*enhancers* o silenciadores) sobre los patrones de *splicing*.
- VI. Contribución del *splicing* aberrante de los genes *BRCA1* y *BRCA2* a la susceptibilidad genética a cáncer de mama/ovario.



# *MATERIAL Y MÉTODOS*



## I. Muestras biológicas de pacientes

Las muestras biológicas utilizadas en este estudio y punto de partida del mismo proceden del programa de prevención de Cáncer Hereditario de la Junta de Castilla y León.

Actualmente, este programa consta de tres unidades de consejo genético (UCG) y dos laboratorios de genética molecular, uno en nuestro centro, el Instituto de Biología y Genética y Molecular (IBGM) de Valladolid, y otro en el Centro del Cáncer de Salamanca. En concreto, las UCG situadas en el hospital Yagüe de Burgos y el hospital Río Hortega de Valladolid se encuentran asociadas con el laboratorio de genética molecular del IBGM de Valladolid.

Las UCG ofrecen asesoramiento genético y apoyo psicológico a pacientes y familiares que cumplen determinados criterios que establecen un riesgo moderado-alto de padecer, entre otros, cáncer de mama y ovario hereditario (Tabla 1). En función de la historia familiar y datos clínicos las UCG determinan la conveniencia de realizar el estudio genético de los genes *BRCA1* y *BRCA2* en las pacientes con riesgo de padecer CMOH. El resultado del estudio establece las recomendaciones de seguimiento y profilaxis disponibles que impliquen un beneficio clínico para el paciente.

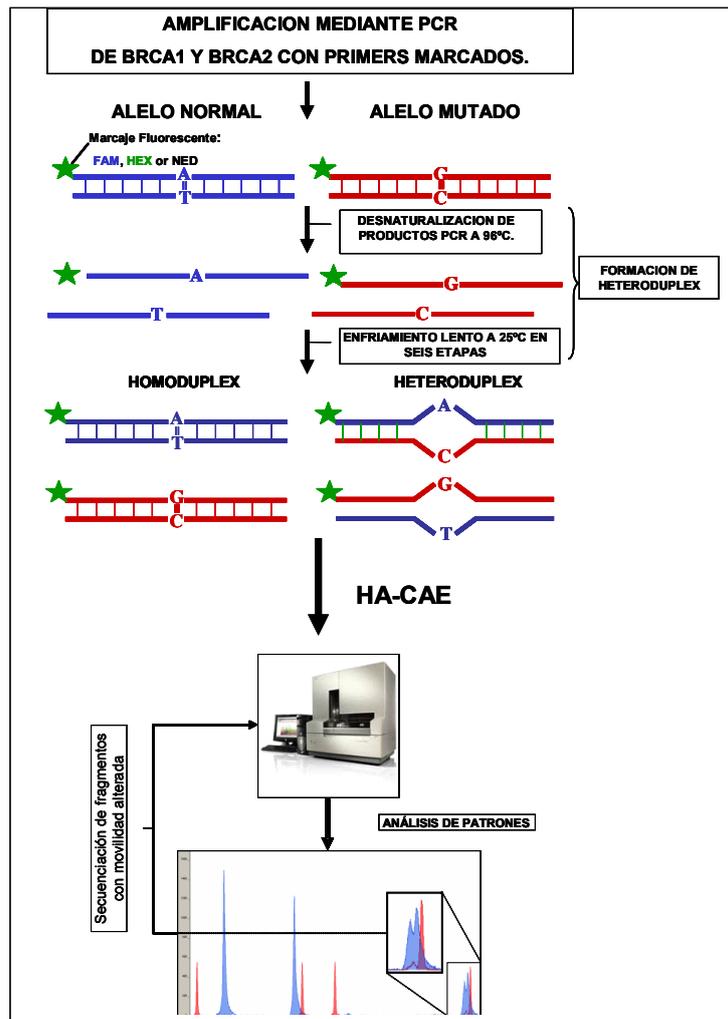
**Tabla 1. Criterios de selección de pacientes de alto riesgo.**

Criterios de selección de pacientes de alto riesgo
<b>ÚNICO CASO</b>
A. Cáncer de mama bilateral, dos primarios ipsilaterales o mama y ovario en el mismo individuo.
B. Cáncer de mama u ovario diagnosticado antes de los 40 años.
C. Cáncer de mama en el varón.
<b>DOS CASOS</b>
D. Dos casos de cáncer de mama, uno de ellos diagnosticado antes de los 50 años, (familiares de 1 <sup>er</sup> grado).
E. Cáncer de mama antes de los 50 y ovario en la misma familia (familiares de 1 <sup>er</sup> y 2 <sup>o</sup> grado).
F. Dos casos de cáncer de ovario a cualquier edad (familiares de 1 <sup>er</sup> y 2 <sup>o</sup> grado).
<b>TRES CASOS</b>
G. Tres o más casos de cáncer de mama (dos en familiares de 1 <sup>er</sup> grado) y/o ovario (familiares de 1 <sup>er</sup> y 2 <sup>o</sup> grado), diagnosticados a cualquier edad.
<b>OTROS CASOS</b>
H. Otros casos de agregación familiar de cáncer de mama y/u ovario o de otros tipos de cáncer.

### I.1. Detección de mutaciones en pacientes con CMOH. Análisis de heterodúplex

El laboratorio de genética del cáncer del IBGM ha desarrollado una técnica propia para la detección de mutaciones en los genes *BRCA1* y *BRCA2*: el análisis de heterodúplex mediante electroforesis capilar o HA-CAE (Velasco *et al.*, 2007). En general el protocolo consta de cuatro procesos:

- **Extracción de ADN** a partir de 1 ml de sangre anticoagulada con EDTA utilizando un robot de extracción MagNA Pure Compact (Roche, Basel, Suiza) junto con el kit MagNA Pure Compact Nucleic Acid Isolation 1-Large Volume (Roche).
- **PCR múltiplex** o amplificación de varios fragmentos de ADN en un mismo tubo en una sola reacción. En el gen *BRCA1* se realiza la amplificación de 31 fragmentos divididos en 10 grupos, mientras que 42 fragmentos son agrupados en 13 grupos para la amplificación de *BRCA2*. La PCR múltiplex es un paso previo a la HA-CAE y se utilizan un cebador fluorescente por cada producto PCR.
- **Análisis de heterodúplex.** Una vez finalizado el proceso de amplificación, en caso de existir una mutación en heretocigosis, se forman cuatro tipos de moléculas de ADN bicatenario: dos homodúplex y dos heterodúplex. El desapareamiento entre las bases nitrogenadas en las moléculas heterodúplex (*wild type*-mutación) provoca un retraso en la en la movilidad electroforética respecto a las homodúplex. La detección de patrones electroforéticos alterados advierte de la presencia de una mutación (Figura 1).



**Figura 1. Esquema del Análisis de Heterodúplex mediante electroforesis capilar (HA-CAE).** Tras la amplificación utilizando cebadores marcados, las dos dobles hebras de ADN amplificado (normal o wt, en azul y mutante en rojo) se desnaturalizan y rehibridan lentamente. Se producen entonces dos moléculas de ADN homodúplex (wt-wt, mut-mut) y dos heterodúplex o moléculas híbridas (mut-wt, wt-mut). Mediante electroforesis capilar pueden ser detectadas diferencias de movilidad entre las distintas especies formadas indicando la presencia de una mutación.

## I.2. Mutaciones seleccionadas del Programa de prevención de CMOH

En los últimos diez años, 2.200 individuos pertenecientes a 1.400 familias diferentes han sido analizados por el procedimiento HA-CAE en Valladolid. En total se han detectado 253 variantes diferentes, 104 en *BRCA1* y 149 en *BRCA2*. El 68% son mutaciones de efecto clínico desconocido. Tres mutaciones detectadas en el Programa de Prevención de Cáncer Hereditario de la Junta de Castilla y León se han seleccionado para realizar este estudio:

- La mutación c.8488-1G>A, detectada en una paciente que desarrolló cáncer de mama bilateral a los 66 y 74 años respectivamente y de ovario a los 77. Esta variante no ha sido detectada con anterioridad en pacientes con CMOH por lo que su

significancia clínica era incierta. La mutación se sitúa en el último nucleótido del intrón 19 del gen *BRCA2*, por lo que presuntamente podría afectar al sitio aceptor de *splicing* del exón 20.

- La mutación intrónica c.9501+9A>C fue detectada 9 nt *upstream* del inicio del intrón 25 del gen *BRCA2*. La cercanía de esta mutación al sitio donador del exón 25 podría influir en el correcto reconocimiento del exón 25 durante el proceso de *splicing*.

- La mutación c.9026\_9030del en el exón 23 del gen *BRCA2*, aunque catalogada como *frameshift* se ha incluido en el estudio porque en estudios previos realizados en nuestro laboratorio se ha descrito que pequeñas deleciones exónicas pueden afectar al proceso de *splicing*, como por ejemplo la c.470\_474del (Sanz *et al.*, 2010).

## II. Análisis *in silico* del impacto de mutaciones en el *splicing*

Diversas herramientas informáticas construidas en base a datos experimentales permiten evaluar la probabilidad que tiene un determinado cambio de ADN de afectar a secuencias básicas de los sitios canónicos de *splicing* y otros elementos reguladores implicados en el proceso de *splicing*. En el proceso de análisis se han comparado las predicciones de los diferentes elementos en *cis* detectados a partir de la secuencia consenso de referencia con las secuencias que incluyen las mutaciones que queremos analizar. Siguiendo esta metodología se han analizado 343 variantes de ADN con los programas detallados en el apartado II.2 de esta sección, de las cuales 333 se encuentran registradas en diferentes bases de datos (apartado II.1.b). Además, en colaboración con el Dr. Miguel de la Hoya (Hospital Clínico San Carlos, Madrid) se incluyó el estudio de la variante c.7806-14T>C localizada en el sitio aceptor del exón 17. Las siguientes tablas reflejan las mutaciones concretas analizadas (Tablas 2 a 10).

**Tabla 2. Mutaciones estudiadas *in silico* del exón 19.** A excepción de las mutaciones artificiales, el resto se encuentran anotadas en la base de datos BIC. Las mutaciones destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing* con minigenes híbridos.

Mutaciones analizadas exón 19			
<b>Frameshift</b>			
c.8340_8343delTAAC	<b>c.8394_8396delTAGinsAA</b>	c.8436_8437insA	c.8474_8474delC
c.8343_8343delC	c.8395_8395delA		
<b>Nonsense</b>			
c.8485C>T			
<b>Missense</b>			
c.8350C>T	c.8363G>C	<b>c.8398C&gt;T</b>	<b>c.8438G&gt;A</b>
c.8351G>A	<b>c.8368A&gt;T</b>	<b>c.8399C&gt;G</b>	c.8443G>A
c.8356G>C	c.8375T>C	c.8420C>T	c.8452G>A
c.8356G>A	c.8377G>A	c.8428A>C	c.8458G>T
c.8359C>T	<b>c.8378G&gt;A</b>	c.8428A>G	c.8462T>C
c.8360G>A	c.8382C>G	c.8432A>G	c.8482A>G
<b>c.8362T&gt;C</b>	c.8393C>T	<b>c.8435G&gt;A</b>	<b>c.8486A&gt;T</b>
<b>Intrónicas</b>			
c.8332-66T>C	<b>c.8487+4T&gt;C</b>	c.8487+19A>G	<b>c.8487+1G&gt;A</b>
c.8332-64A>G	<b>c.8487+3A&gt;G</b>	c.8487+47C>T	<b>c.8487+8G&gt;A</b>
<b>Sinónimas</b>			
c.8460A>C	c.8487A>G		
<b>Artificiales</b>			
<b>c.8484A&gt;T</b>	<b>c.8486A&gt;C</b>	<b>c.8487+3A&gt;T</b>	<b>c.8487+3A&gt;C</b>

**Tabla 3. Mutaciones estudiadas *in silico* del exón 20.** A excepción de las mutaciones artificiales, el resto se encuentran anotadas en la base de datos BIC. Las mutaciones destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing*. La mutación destacada con asterisco se ha detectado en pacientes del Programa de Prevención de Cáncer Hereditario de Castilla y León.

Mutaciones Exón 20			
<b>Frameshift</b>			
c.8501_8501delC	c.8535_8538delAGAG	c.8575_8575delC	c.8593_8594insT
c.8505_8505delA	c.8537_8538delAG	c.8578_8579delAA	
c.8533_8534delAG	<b>c.8560_8560delT</b>	c.8585_8586insT	
<b>Nonsense</b>			
c.8489G>A	c.8572C>T	c.8594T>A	
c.8504C>A	c.8575C>T	c.8629G>T	
<b>Missense</b>			
c.8503T>C	c.8525G>A	c.8548G>A	c.8593T>G
c.8518A>G	c.8528A>G	c.8567A>C	c.8599A>C
c.8524C>T	<b>c.8539G&gt;A</b>	<b>c.8572C&gt;A</b>	c.8618T>G
c.8525G>T	c.8545A>G	c.8591C>T	
<b>Intrónicas</b>			
c.8632+6A>G	c.8632+24A>G	<b>c.8488-1G&gt;A*</b>	
<b>Artificiales</b>			
<b>c.8512T&gt;A</b>	<b>c.8518delA</b>	<b>c.8609A&gt;G;8611G&gt;T</b>	<b>c.8488-2A&gt;G</b>

**Tabla 4. Mutaciones estudiadas *in silico* del exón 21.** A excepción de la mutación artificial y la mutación destacada en cursiva que pertenece a la base de datos LOVD, el resto se encuentran anotadas en la base de datos BIC. Las mutaciones destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing*.

Mutaciones exón 21			
<b>Frameshift</b>			
c.8636_8637insA <i>c.8645del4</i>	c.8676_8676delA	c.8713_8716delTATG	<b>c.8748_8749insTTAC</b>
<b>Missense</b>			
c.8647C>T	c.8663G>T	c.8687G>A	<b>c.8732C&gt;T</b>
<b>c.8662C&gt;T</b>	c.8665G>T	c.8702G>A	
<b>c.8663G&gt;C</b>	c.8668C>A	<b>c.8723T&gt;G</b>	
<b>Intrónicas</b>			
c.8633-26A>G	c.8633-16C>G	<b>c.8754+4A&gt;G</b>	<b>c.8754+5G&gt;A</b>
<b>c.8633-24_8634del</b>	c.8633-2A>G	<b>c.8754+5G&gt;T</b>	
<b>Sinónimas</b>			
<b>c.8754G&gt;A</b>			
<b>Artificial</b>			
<b>c.8754+5G&gt;C</b>			

**Tabla 5. Mutaciones estudiadas *in silico* del exón 22.** Todas las mutaciones seleccionadas se encuentran anotadas en la base de datos BIC. Las mutaciones destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing*.

Mutaciones exón 22			
<b>Frameshift</b>			
<b>c.8756_8756delG</b>	c.8800ins346	c.8904_8904delC	c.8940_8941insA
c.8789_8789delA	c.8848_8848delAinsCT	c.8912_8912delA	c.8946_8947insA
<b>Nonsense</b>			
<b>c.8770G&gt;T</b>	c.8878C>T	c.8951C>G	
c.8773C>T	c.8933C>A		
<b>Missense</b>			
<b>c.8756G&gt;T</b>	c.8806T>A	c.8850G>T	<b>c.8915T&gt;G</b>
<b>c.8764A&gt;G</b>	c.8808G>C	c.8851G>A	c.8917C>T
<b>c.8774A&gt;G</b>	c.8824G>A	c.8893G>C	c.8918G>A
<b>c.8775G&gt;C</b>	c.8830A>T	c.8905G>A	c.8944A>C
<b>Splicing</b>			
<u>c.8948_8953+5del</u>			
<b>Intrónicas</b>			
<u>c.8755-1G&gt;A</u>	c.8755-66T>C	c.8953+98T>C	
<b>c.8755-79G&gt;A</b>	c.8953+16C>T	<b>c.8953+1G&gt;T</b>	

**Tabla 6. Mutaciones estudiadas *in silico* del exón 23.** A excepción de las mutación doble artificial, el resto se encuentran anotadas en la base de datos BIC. Las mutaciones destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing*. La mutación destacada con asterisco se ha detectado en pacientes del Programa de Prevención de Cáncer Hereditario de Castilla y León.

Mutaciones exón 23			
<b>Frameshift</b>			
<b>c.8961_8964delGAGT</b>	c.9016_9017delTA	c.9054_9055delTA	c.9069_9076delTAACATAC
c.8980_8983delTCAG	<b>c.9026_9030delATCAT*</b>	c.9057_9057delA	c.9098_9099insA
c.9015_9016delAT	c.9027_9027delT	c.9066_9067insAA	c.9099_9100delTC
<b>Nonsense</b>			
<b>c.8969G&gt;A</b>	c.8999T>A	c.9041C>A	<b>c.9100C&gt;T</b>
c.8970G>A	c.9018C>A	<b>c.9076C&gt;T</b>	
<b>Missense</b>			
<b>c.8962A&gt;G</b>	c.9032T>C	c.9082G>C	c.9104A>G
<b>c.8972G&gt;A</b>	c.9038C>T	c.9085G>A	c.9104A>C
<b>c.9004G&gt;A</b>	c.9059C>G	c.9086C>T	<b>c.9116C&gt;T</b>
<b>c.9006A&gt;T</b>	c.9076C>G	<b>c.9101A&gt;G</b>	
<b>Intrónicas</b>			
<b>c.8954-3C&gt;G</b>	c.9117+9C>T	<b>c.9117+1G&gt;A</b>	
<b>c.8954-1_8955delinsAA</b>	<b>c.9117+1G&gt;T</b>		
<b>Sinónimas</b>			
c.8997G>A	<b>c.9117G&gt;A</b>		
<b>Ins/del En marco de lectura</b>			
c.8975_9100del	c.9115_9116insTAC		
<b>Artificiales</b>			
<b>c.9872G&gt;A;9006A&gt;T</b>			

**Tabla 7. Mutaciones estudiadas *in silico* del exón 24.** Todas las mutaciones seleccionadas se encuentran anotadas en la base de datos BIC. Las mutaciones destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing* con minigenes híbridos.

Mutaciones exón 24			
<b>Frameshift</b>			
c.9157_9157delG	<b>c.9177_9177delA</b>	<b>c.9252_9255delAACAAinsTT</b>	c.9253_9254insA
<b>Nonsense</b>			
c.9148C>T	c.9196C>T	<b>c.9256G&gt;T</b>	
<b>c.9182T&gt;A</b>	c.9207T>A		
<b>Missense</b>			
c.9118G>A	c.9175A>G	<b>c.9196C&gt;G</b>	c.9227G>T
c.9154C>T	c.9187C>T	<b>c.9213G&gt;T</b>	c.9235G>A
c.9155G>A	c.9190G>T	c.9215T>A	c.9242T>C
c.9161C>A	c.9190G>A	<b>c.9218A&gt;G</b>	c.9247A>G
<b>c.9172A&gt;G</b>	c.9195T>A	<b>c.9227G&gt;A</b>	c.9249A>T
<b>Splicing</b>			
<b>c.9248_9256+7del</b>			
<b>Intrónicas</b>			
<b>c.9118-6C&gt;T</b>	c.9118-14T>C	c.9256+16T>C	
<b>c.9118-2A&gt;G</b>	<b>c.9256+1G&gt;A</b>		
<b>Sinónimas</b>			
c.9237T>C			

**Tabla 8. Mutaciones estudiadas *in silico* del exón 25.** Todas las mutaciones seleccionadas se encuentran anotadas en la base de datos BIC. Las mutaciones destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing* con minigenes híbridos. La mutación destacada con asterisco se ha detectado en pacientes del Programa de Prevención de Cáncer Hereditario de Castilla y León.

Mutaciones exón 25			
<b>Frameshift</b>			
c.9269_9270insT	c.9384_9385insG	c.9426_9427delTT	<b>c.9466_9466delC</b>
<b>c.9275_9278delATTT</b>	c.9401_9401delG	c.9435_9436delGT	
c.9310_9311delAA	c.9403_9403delC	c.9455_9456delAG	
c.9360_9360delT	c.9408_9408delT	c.9456_9457insAG	
<b>Nonsense</b>			
c.9276T>G	c.9294C>A	c.9376C>T	c.9481A>T
c.9286G>T	c.9317G>A	c.9380G>A	
c.9294C>G	c.9356T>G	c.9382C>T	
<b>Missense</b>			
c.9263C>T	<b>c.9301C&gt;G</b>	c.9367A>G	c.9449C>T
c.9271G>A	c.9302T>G	c.9371A>T	<b>c.9454G&gt;A</b>
c.9275A>C	c.9307A>G	c.9374T>A	c.9455A>G
c.9275A>G	c.9309A>G	c.9385C>G	c.9458G>C
c.9285C>G	c.9350A>C	c.9401G>T	c.9477C>A
c.9292T>C	c.9353T>C	c.9425A>G	c.9500A>C
<b>Intrónicas</b>			
<b>c.9257-18C&gt;A</b>	c.9257-10_9257-9insT	<b>c.9501+9A&gt;C*</b>	
c.9257-19_9257-17delinsCC	<b>c.9257-1G&gt;C</b>	<b>c.9501+3A&gt;T</b>	
c.9257-16T>C	c.9257-10_9257-10delT	<b>c.9501+4A&gt;G</b>	
<b>Sinónimas</b>			
c.9414A>G	c.9501G>A	<b>c.9301C&gt;G</b>	

**Tabla 9. Mutaciones estudiadas *in silico* del exón 26.** A excepción la mutación destacada en cursiva que pertenece a la base de datos UMD, el resto se encuentran anotadas en la base de datos BIC. Las mutaciones destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing*.

Mutaciones exón 26			
<b>Framsehift</b>			
c.9501G>A	c.9513_9516delACTT	c.9580_9581delCC	
<b>Nonsense</b>			
c.9599C>G			
<b>Missense</b>			
c.9509A>G	c.9583A>G	<b>c.9611C&gt;G</b>	c.9637A>G
c.9542T>G	c.9586A>G	<b>c.9613_9614delGCinsCT</b>	<b>c.9550C&gt;G</b>
c.9562G>A	c.9592T>C	c.9616C>G	
c.9581C>A	<b>c.9610A&gt;G</b>	c.9634G>C	
<b>Intrónicas</b>			
c.9648+10T>G	<b>c.9502-12T&gt;G</b>	c.9648+84G>A	c.9648+106_9648+106delT
<b>c.9502-2A&gt;C</b>	c.9502-71G>T	c.9648+54G>A	
<b>Sinónimas</b>			
<b>c.9646C&gt;T</b>			
<b>Ins/del En marco de lectura</b>			
c.9513_9515delACT			

**Tabla 10. Mutaciones estudiadas *in silico* del exón 27.** Las mutaciones destacadas en cursiva pertenecen a la base de datos UMD, el resto se encuentran anotadas en la base de datos BIC. Las mutaciones

destacadas en negrita se han seleccionado para realizar ensayos funcionales de *splicing* con minigenes híbridos.

Mutaciones exón 27			
<b>Framseshift</b>			
c.9666_9666delT	c.9676_9676delT	c.9682delA	
c.9672_9673insA	c.9699_9702delTATG		
<b>Nonsense</b>			
c.9682delA			
<b>Missense</b>			
c.9658C>T	c.9718G>A	c.9680A>G	c.9727C>T
c.9677A>G	c.9728C>T	<b>c.9698G&gt;T</b>	c.9730G>A
c.9690A>T	c.9730G>A	c.9700A>G	
c.9706A>G	c.9656C>T	c.9718G>T	
<b>Intrónicas</b>			
<b>c.9649-2A&gt;G</b>	<b>c.9649-5A&gt;G</b>		
<b>Sinónimas</b>			
c.9720T>C	c.9738C>T	c.9738C>A	

## II.1. Bases de datos

La información necesaria para la realización del estudio bioinformático se ha recabado de diferentes bases de datos de ámbito internacional. Por un lado, ha sido necesario consultar la secuencia consenso para el gen *BRCA2* y por otro, listar mutaciones detectadas en análisis genéticos de pacientes de CMOH de todo el mundo.

### II.1.a. Bases de datos de secuencias. Ensembl

Ensembl ([www.ensembl.org](http://www.ensembl.org)) es la base de datos utilizada para la consulta de la secuencia consenso del gen *BRCA2* en el genoma humano. Este repositorio lo estableció el Wellcome Trust Sanger Institute a raíz de la ingente cantidad de secuencias aportadas por el proyecto Genoma Humano (Hubbard *et al.*, 2002). Ensembl se actualiza con regularidad e incluye numerosas anotaciones. En ella también se encuentran depositadas las secuencias de otros genomas eucariotas que han sido utilizados para el estudio de secuencias implicadas en el *splicing* y conservadas evolutivamente. La Tabla 11 muestra los números de acceso a las secuencias del gen *BRCA2* de los genomas utilizados.

Tabla 11. Números de acceso de las secuencias del gen *BRCA2* de diferentes especies.

Especie	Nº Acceso BRCA2
<i>Homo sapiens</i>	ENSG00000139618

<i>Bos Taurus</i>	ENSBTAG00000000988
<i>Canis familiaris</i>	ENSCAFG00000006368
<i>Felix Catus</i>	ENSFCAG00000003304
Gallus gallus	ENSGALG00000017073
Mus musculus	ENSMUSG00000041147
Rattus norvegicus	ENSRNG000000001111

### II.1.b. Bases de datos de mutaciones. BIC, LOVD y UMD

La base de datos que contiene un mayor número de registros de mutaciones diferentes en *BRCA2* es **B**reast **C**ancer **I**nformation **C**ore o **BIC** (Szabo *et al.*, 2000). La última actualización del año 2012 contenía casi 2000 variaciones de ADN diferentes en este gen. Sin embargo, existen otras bases de datos como **L**eiden **O**pen **V**ariation **D**atabase o **LOVD** (Fokkema *et al.*, 2005) y **U**niversal **M**utation **D**atabase o **UMD** (Beroud *et al.*, 2000) donde se pueden encontrar anotadas algunas mutaciones que no aparecen en BIC. Existen diferentes nomenclaturas entre las diferentes bases de datos, pero se han utilizado las normas recomendadas por la Human Genome Variation Society (<http://www.hgvs.org/mutnomen/>).

## II.2. Análisis *in silico* del impacto de mutaciones en el *splicing*

### II.2.a. NNsplice

NNsplice ([http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)) es una herramienta *on-line* desarrollada por la Universidad de Berkeley dentro del proyecto Genoma de *Drosophila* en el año 1997 (Reese *et al.*, 1997). NNsplice aplica el modelo de Markov para predecir posibles sitios de *splicing* aceptor y donador de una secuencia determinada en formato FASTA. En su origen se implementó en el programa Genie (Reese *et al.*, 1997) para la búsqueda de genes multiexónicos. En este trabajo se ha usado para determinar el efecto que una mutación puede tener sobre el reconocimiento de los sitios canónicos de *splicing*.

### II.2.b. ESEfinder

ESEfinder (<http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>) es un recurso web desarrollado por Adrian Krainer y sus colaboradores (Cartegni *et al.*, 2003) que facilita el reconocimiento rápido de putativas secuencias exónicas de unión a proteínas SR. En concreto, detecta ESE relacionadas con las proteínas SF2/ASF, SC35,

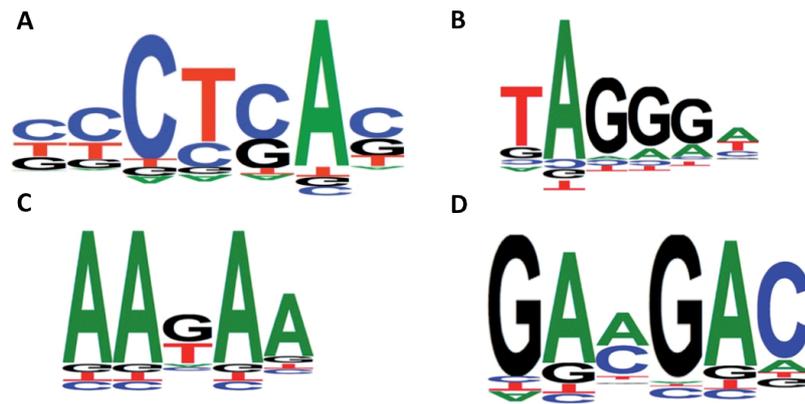
SRp40 y SRp55. En el diseño de ESEfinder se han utilizado matrices basadas en valores de frecuencia de aparición de nucleótidos en secuencias ESE, obtenidos a partir de experimentos funcionales SELEX (*Systematic Evolution of Ligands by Exponential enrichment*; Tuerk y Gold, 1990). El resultado es un valor relativo que estima la probabilidad de que una de las cuatro proteínas mencionadas reconozca un sitio ESE dentro de la secuencia a estudio.

### II.2.c. ESRsearch

ESRsearch (<http://esrsearch.tau.ac.il/>) es un programa informático de aplicación *web* que permite la detección de secuencias reguladoras de *splicing* en una determinada secuencia. Implementa cuatro algoritmos diferentes de búsqueda de ESRs: PESX (Zhang y Chasin, 2004), ExonScan (Wang *et al.*, 2004), Rescue-ESE (Fairbrother *et al.*, 2004), PESR (Goren *et al.*, 2006). ESRsearch es capaz de detectar 285 secuencias ESR diferentes, tanto silenciadoras (ESS) como potenciadoras de *splicing* (ESE). Dichas secuencias fueron seleccionadas de acuerdo a su abundancia dentro de los genes humanos y a su conservación evolutiva.

### II.2.d. Human Splicing Finder

Human Splicing Finder (HSF, [www.http://www.umd.be/HSF/](http://www.umd.be/HSF/)) es una herramienta bioinformática en red que permite identificar los motivos de *splicing* en el pre-ARNm y predecir el efecto que tienen las mutaciones que aparecen en estas secuencias. HSF agrupa todas las matrices disponibles para la predicción de secuencias auxiliares relacionadas con el *splicing* e incorpora algunas nuevas como la predicción de secuencias de unión a las proteínas 9G8 y Tra2- $\beta$  y la ribonucleoproteína hnRPA1. Además, se incluyen nuevas matrices que expresan la probabilidad de que sean reconocidos determinados sitios aceptor, donador y puntos de ramificación por parte de los diferentes componentes del espliceosoma (Renwick *et al.*, 2009).



**Figura 2.** Nuevas matrices incluidas en HSF. El tamaño de cada nucleótido es proporcional a la probabilidad de que el mismo constituya A) el punto de ramificación, secuencias de reconocimiento de B) hnRNPA1, C) Tra2-β y D) 9G8.

### II.3. Estudio evolutivo de la conservación de secuencias implicadas en el *splicing* en *BRCA2*

El estudio evolutivo de conservación de secuencias en el gen *BRCA2* se realizó con la herramienta de alineación de secuencias ClustalW (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>; Thompson *et al.*, 2002) usando las secuencias anotadas en Ensembl y cuyos números de acceso se han mostrado en la Tabla 11. Estudios previos han demostrado que la conservación evolutiva de determinadas secuencias está relacionada con la importancia de las mismas en el proceso de *splicing* (Ast *et al.*, 2004). La detección de secuencias conservadas que coinciden con aquellas determinadas por las previsiones bioinformáticas puede aumentar la probabilidad de que dichas secuencias tengan importancia funcional en el proceso de *splicing*.

### III. Ensayos funcionales de *splicing*. RT-PCR de linfocitos. Detección de isoformas de *splicing* de muestras de pacientes

Los ensayos funcionales de *splicing* mediante RT-PCR implican un conjunto de técnicas que permite la detección de isoformas de *splicing* producidas por una mutación en el ADN, a partir de muestras biológicas de ARN.

#### III.1. Extracción de ARN

Salvo con la excepción de la paciente que porta la variante c.9501+9A>C de la cual fue imposible obtener una muestra biológica, la extracción de ARN se efectuó a

partir de 2-3 ml de sangre anticoagulada con EDTA mediante el kit QIAamp RNA Blood Mini (Qiagen, Hilden, Alemania) siguiendo las instrucciones del fabricante. Como control se utilizó ARN extraído de células HeLa y MCF10A en cultivo con el kit Nucleo Spin RNA II (Macherey-Nagel, Düren, Alemania).

### III.2. Retrotranscripción ARN-ADNc

Una vez medida la concentración y calidad del ARN mediante un espectrofotómetro Nanodrop (ThermoScientific, Waltham, MA, EE.UU) se obtuvo ADN complementario (ADNc) utilizando el kit comercial High Capacity (Applied Biosystems, Austin, TX, EE.UU) que incluye una enzima retrotranscriptasa, siguiendo las instrucciones del fabricante.

### III.3. RT-PCR

La RT-PCR es una variante de la PCR en la que se amplifica el ADNc resultante de la retrotranscripción. Para la amplificación del ADNc obtenido del paciente que porta la mutación c.8488-1G>A se diseñó un pareja de cebadores solapantes al exón 19 y al exón 23 respectivamente, abarcando una secuencia de 653 pb. Además, con el propósito de identificar posibles variables de *splicing* que impliquen una retención de intrón se diseñó un cebador en sentido *reverse* (RV) solapante con el final del intrón 19 y el comienzo del exón 20. Este nuevo cebador se utilizó en combinación con el previamente diseñado para el exón 19. Por su parte, la amplificación del ADNc obtenido de la muestra que porta la mutación c.9026\_9030del se realizó con una pareja de cebadores diseñada sobre los exones 22 y el 26 que abarca una secuencia de 773 pb. En ambos casos, las reacciones de amplificación se realizaron con la enzima GoTaq (Promega, Madison, WI, EE.UU) en un termociclador Veriti (Applied Biosystems) con el programa: 95°C-2'; 35 ciclos x [95°C 30"-58°C 30"-72°C 1'30"];72°C 10'.

**Tabla 12. Secuencia cebadores Forward (FW) y Reverse (RV) utilizados en ensayos funcionales RT-PCR.**

Cebadores ensayos funcionales RT-PCR BRCA2	
Cebador	Secuencias 5'- 3'
RT_BR2_EX19_FW	AAACTTGGATTCTTTTCCTGAC
RT_BR2_EX23_RV	ATTCTGTATCTCTTTTCCTTCTGTT
RT_BR2_IVS_19_EX20RV	TGTCTTCTCCATCCACTGTAAT
RT_BR2_EX22_FW	GTGAAGAGCAGTTAAGAGCCT
RT_BR2_EX26_RV	CTTGTTTTCTGCTTCATTGC

Tabla 13. Mezcla de reacción de RT-PCR.

RT-PCR	
Buffer	1x
Cl <sub>2</sub> Mg	2-2,5 mM
ADNc molde	1-2 µl
Cebador FW	0,2 µM
Cebador RV	0,2 µM
dNTPs	0,2 mM
Enzima GoTaq	1,5 Unidades
Agua	
Volumen final	25 µl

#### III.4. Electroforesis en gel de agarosa

El resultado de la amplificación RT-PCR se comprobó mediante electroforesis en gel de agarosa. Los geles utilizados contienen una concentración 1,2% (p/v) de agarosa en tampón TBE x 1 (Tris-borato 45 mM pH 8,0; ácido bórico 45mM, EDTA 1 mM). En la elaboración de los geles, se añade 5 µl de Redsafe (INtRON Biotechnology, Seoul, Korea) por cada 100 µl de solución. El Redsafe es un agente químico de unión a ADN y que emite fluorescencia verde. Los productos de RT-PCR se mezclaron con el *buffer* de carga Green-GoTaq que contiene los colorantes Xylene Cyanol y Orange G. Las electroforesis se realizaron a 100V y una vez finalizadas los fragmentos se visualizaron en un transiluminador ultravioleta con ayuda del sistema GelDoC (BioRad, Hercules, CA, EE.UU) para digitalizar las imágenes. El tamaño de los fragmentos visualizados se estimó comparando los mismos con un marcador de tamaño molecular conocido (*ladder* de 1kb, Invitrogen, Carlsbad, CA, EE.UU).

#### III.5. Secuenciación de ADN mediante electroforesis capilar

La secuenciación de ADN en los ensayos funcionales de RT-PCR se ha utilizado para la caracterización de isoformas que presentan un *splicing* anormal.

Previa purificación de la reacción de RT-PCR con el kit High Pure PCR Product Purification (Roche, Basel, Suiza), se realizaron las reacciones de secuencia con el kit BigDye Terminator Cycle Sequencing v3.1 (Applied Biosystems) siguiendo las instrucciones del fabricante y utilizando como cebadores los mismos usados para la reacción RT-PCR de cada mutación.

La reacción de secuencia se basa en una amplificación en la cual se incorporan didesoxinucleótidos (ddntps) o terminadores de cada una de las cuatro bases

nitrogenadas marcadas con fluorocromos diferentes que impiden la elongación de la polimerasa. Los fragmentos a los cuales se han incorporado pueden ser detectados por el secuenciador automático tras su estimulación por láser. Los ddntps no incorporados en la reacción de secuencia deben ser eliminados ya que pueden interferir en la carrera electroforética de la secuenciación automática debido a su fluorescencia. Para tal fin, se utilizaron perlas magnéticas CleanSEQ (Agencourt Bioscience, Beverly, MA, EEUU) según las instrucciones del fabricante.

La carrera electroforética capilar se llevó a cabo en un secuenciador automático ABI Prism® 3130 (Applied Biosystems) con cuatro capilares de 36 cm, utilizando POP-7 (Applied Biosystems) como polímero de electroforesis y tampón de carrera con EDTA (Applied Biosystems).

**Tabla 14. Condiciones de carrera de la electroforesis capilar para secuenciación.**

Parámetros	Valor
Temperatura del horno	50 °C
Voltaje de inyección	1 Kvolts
Tiempo de inyección	22 segundos
Voltaje de carrera	12,2 Kvolts
Tiempo de carrera	6.500 segundos

Los datos brutos de la carrera electroforética fueron procesados mediante el programa informático Data Collection 3.0 y Sequencing Analysis v5.2 (Applied Biosystems). El análisis de los cromatogramas se ha realizado con la aplicación Chromas Pro (<http://technelysium.com.au/>; Technelysium, South Brisbane QLD, Australia).

#### **IV. Ensayos funcionales de *splicing* mediante minigenes híbridos**

Los ensayos funcionales de *splicing* mediante minigenes híbridos agrupan a un conjunto de técnicas que permiten estudiar el efecto que una mutación produce en el proceso de *splicing* sin la necesidad de utilizar una muestra biológica del paciente que contiene la variante a estudio. Este es el caso de la paciente del programa de prevención de cáncer hereditario de Castilla y León en la que se detectó la mutación c.9501+9A>C . Además, todas las mutaciones seleccionadas de las bases de datos

internacionales en relación a su potencial impacto en el *splicing in silico* fueron estudiadas con esta metodología.

#### IV.1. Construcción de minigenes

Los minigenes son vectores plasmídicos de expresión en eucariotas que contienen un fragmento del gen que se desea estudiar.

##### IV.1.a. Tipos de vectores

##### IV.1.a.1. Vector pSPL3

El vector pSPL3 fue cedido a nuestro laboratorio por las Dras. Belén Pérez y Lourdes R. Desviat (Universidad Autónoma de Madrid) y su derivado pSPL3b (Burn *et al.*, 1995) fue enviado por la Dra. Rita Brandao (Maastricht University). Consta de: un origen de replicación para *E. Coli*; secuencias del virus SV40 que promueven la replicación y la transcripción en células humanas; un marcador de selección (gen de resistencia a ampicilina); 2 exones, V1 procedente de la  $\beta$ -Globina de conejo y V2 del gen HIV1-tat y un sitio de clonación múltiple (MCS) que se encuentra inmerso dentro de una secuencia intrónica del gen HIV1-tat (Figura 3). Las posibles señales de *splicing* del fragmento insertado en el plásmido provocan reacciones heterólogas entre los sitios del vector y los del fragmento introducido.

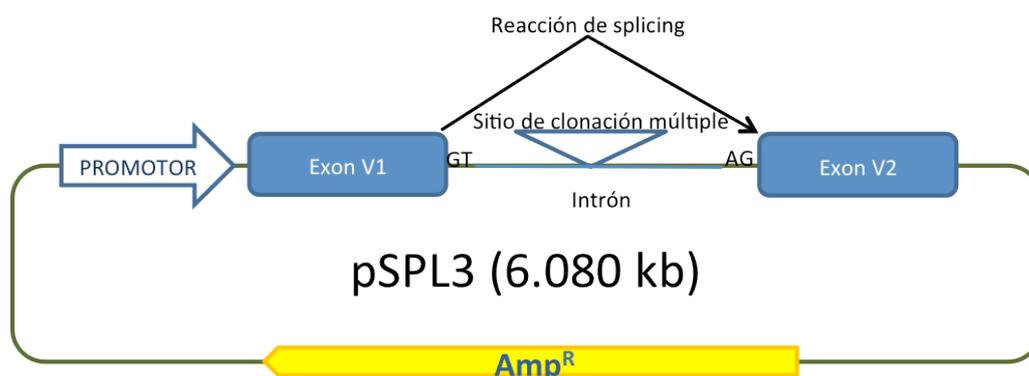


Figura 3. Esquema del vector pSPL3.

##### IV.1.a.2. Nuevo vector: pSAD

El manejo continuado del vector pSPL3 reveló la existencia de determinados problemas. Algunos se han descrito con anterioridad, como sitios crípticos dentro del intrón del vector que son reconocidos en el proceso de *splicing* (Burn *et al.*, 1995).

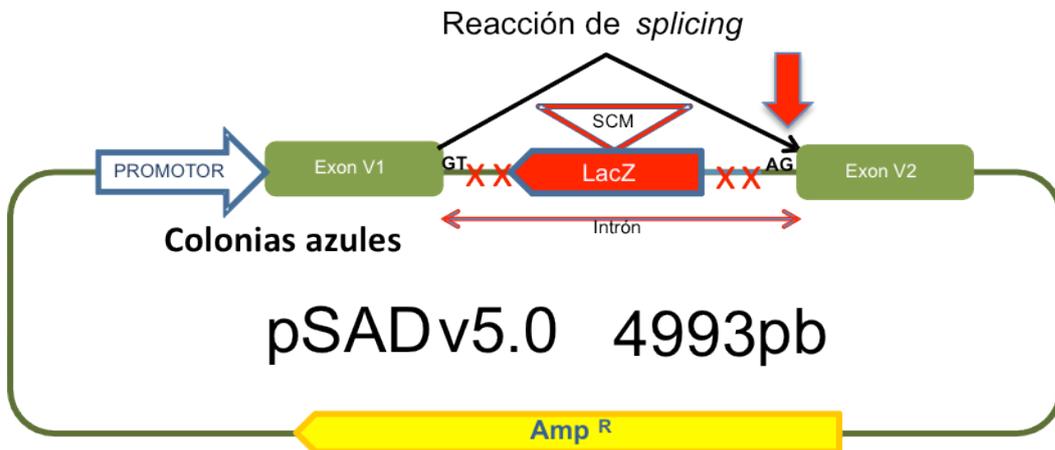
Otros, como la aparición de transcritos en los que los exones clonados no son incluidos, la selección de colonias que contienen el minigen o la dificultad de clonar grandes fragmentos se detectaron en el trabajo diario con este vector.

Con el objetivo de mejorar el vector se realizaron varias modificaciones del mismo (Figura 4):

- Reducción de tamaño de intrón (delección 1.2 kb) mediante mutagénesis dirigida. Este proceso permite manejar mejor el vector y clonar fragmentos más grandes.
- Doble selección: Amp<sup>R</sup> más un segundo marcador de selección (color de colonias azules/blancas). Al marcador de selección de colonias transformadas mediante ampicilina se le añadió parte del gen de la  $\beta$ -galactosidasa, de esta forma en presencia de IPTG 0.1 mM (Fermentas) y X-Gal 40  $\mu$ g/ml (Fermentas) en la placa de agar, podemos diferenciar aquellas colonias recombinantes que han incluido el inserto, ya que presentan color blanco, de las que no lo han incorporado que son de color azul.
- Diferentes sitios crípticos de *splicing* se han eliminado mediante mutagénesis dirigida para evitar interferencias con aquellos propios del inserto clonado.
- La sustitución del sitio de clonación múltiple ha aumentado la versatilidad de enzimas que pueden usarse en el proceso de clonación. En concreto, se ha logrado incrementar de 10 a 27 las secuencias únicas y específicas de reconocimiento de enzimas restricción. Además, se han eliminado 4 sitios de restricción duplicados.
- Por último, se llevó a cabo una mejora del sitio aceptor débil del exón V2 mediante triple mutagénesis. Según la aplicación NNsplice la probabilidad relativa de detección de este sitio por parte de la maquinaria de *splicing* es de 0,64. La modificación de este sitio aumentó la probabilidad de que sea reconocido a 0.99.

Las modificaciones realizadas han supuesto una gran diferenciación respecto al vector original. Este hecho ligado a la posibilidad de aplicación comercial de esta herramienta han propiciado la protección mediante patente del nuevo vector con el nombre pSAD (*Splicing And Disease*) en la Oficina Española de Patentes y Marcas, número de solicitud P201231427. En la patente se plasma de forma precisa todos los pasos necesarios que han posibilitado la creación de pSAD y se incluyeron las secuencias del plásmido y de todos sus derivados intermedios. Todas las modificaciones del vector se realizaron con la técnica de mutagénesis dirigida por PCR

y dos variantes asociadas a la misma: inserción y delección mediante PCR. Estas técnicas serán desarrolladas con amplitud en los capítulos siguientes. Con el objeto de protección de la patente en este documento no se reflejará de forma detallada las modificaciones que han originado el nuevo vector pSAD.



**Figura 4. Esquema del vector pSAD.** En color rojo se representan las principales modificaciones realizadas respecto al vector pSPL3.

#### IV.1.b. Minigenes en pSPL3. MGBR2 19-20 y 23-24

El proceso de construcción de minigenes en pSPL3 se realizó mediante PCR de regiones genómicas utilizando la enzima de alta fidelidad Pfx50™ (Invitrogen). Se amplificaron dos regiones de ADN genómico a partir de 250 ng de ADN: una de 1188 pb que comprende los exones 19 y 20 y otra de 717 pb que incluye los exones 23 y 24. Para la amplificación de estos fragmentos se han usado cebadores (Tablas 15 y 16) específicos a los que se les ha añadido secuencias en su extremo 5' con sitios de corte para las enzimas de restricción XhoI y BamHI (Fermentas, Burlington, ON, Canada). Las reacciones de amplificación se realizaron en un termociclador Veriti (Applied Biosystems) con el programa: 95°C 2'; 35 ciclos x [95°C 30"- 59°C 30"- 68°C 3']; 68°C 20' con la mezcla de reacción presentada en la Tabla 17.

**Tabla 15. Secuencia y tamaño de los cebadores de amplificación de la región genómica que incluye los exones 19 y 20.** Subrayada se muestran las secuencias de los sitios de restricción.

Cebadores BRCA2-pSPL3 Clonación Minigen 19-20		
Cebador	secuencia 5' - 3'	Tamaño
MGBR2_ex19-20_XhoI-FW	CACACACTCGAGATAGCATT <u>A</u> AGA <u>A</u> CTTGTAGCA	1188 pb
MGBR2_ex19-20-BamHI-RV	CACACAGGATCCATTACA <u>A</u> ATGGCTTAGACCTGA	

**Tabla 16. Secuencia y tamaño de los cebadores de amplificación de la región genómica que incluye los exones 23 y 24.** Subrayada se muestran las secuencias de los sitios de restricción.

Cebadores BRCA2-pSPL3 Clonación Minigen 23-24		
Cebador	secuencia 5' - 3'	Tamaño
MGBR2_ex23-24_XhoI-FW	CACACACTCGAGAATGC <u>T</u> TTGTTTTATCAGTTTT	717 pb
MGBR2_ex23-24_BamHI-RV	CACACAGGATCCAA <u>T</u> TTGCCAACTGGTAGCTC	

**Tabla 17. Mezcla de amplificación de secuencias genómicas clonadas en pSPL3.**

PCR Pfx50™	
Buffer	1x
Cebador FW	0,2 µM
Cebador RV	0,2 µM
dNTPs	1 mM
Enzima Pfx50™	1,5 Unidades
Agua	
Volumen final	25 µl

El tamaño de la regiones genómicas amplificadas fue comprobado mediante electroforesis en tampón TBE x1 de gel de agarosa al 1% teñido con Bromuro de Etidio. Las electroforesis se realizaron a 100V y una vez finalizadas los fragmentos se visualizaron en un transiluminador ultravioleta con ayuda del sistema GelDoC (BioRad) para digitalizar las imágenes. El tamaño de los fragmentos visualizados se estimó comparando los mismos con el marcador de tamaño molecular 1 Kb plus DNA ladder (Invitrogen).

A continuación, se procedió a la inserción de regiones genómicas amplificadas en el vector pSPL3. El proceso de inserción comenzó con la digestión tanto de los fragmentos amplificados como el vector con las enzimas de restricción XhoI y BamHI (Fermentas) de acuerdo con las instrucciones del fabricante. La aplicación Webcutter se utilizó para detectar sitios de restricción tanto del vector como en el fragmento amplificado. De esta forma comprobó la existencia de dianas únicas para las enzimas de restricción seleccionadas.

Los productos de las digestiones fueron purificados de forma individual con las columnas High Pure PCR Product Purification (Roche) siguiendo el protocolo

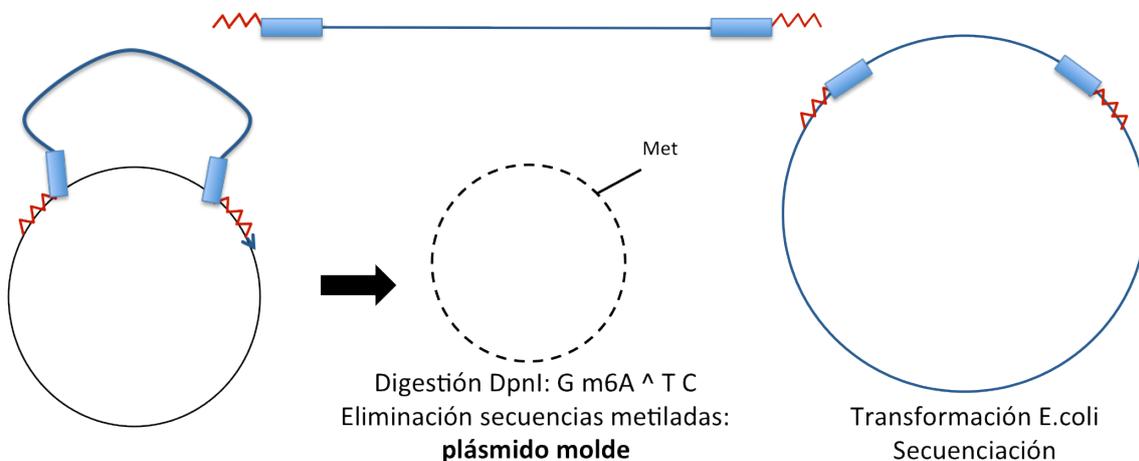
recomendado. El proceso de ligación del vector con los productos amplificados se realizó con la enzima ligasa T4 (Fermentas) partiendo de 200 ng de plásmido en una relación molar 1:3 respecto al inserto.

#### IV.1.c. Minigenes en pSAD

Los tres minigenes construidos a partir del vector pSAD, contienen en conjunto 21 de los 27 exones del gen BRCA2. La incorporación de un gran número de exones en cada minigen fue compleja y requirió la construcción de minigenes intermedios. Por ello, se describe de forma individual la elaboración de cada uno de ellos.

El clonaje de las diferentes regiones amplificadas se realizó en varias etapas y requirió la creación de minigenes intermedios. Los fragmentos amplificados se han insertado mediante dos procedimientos diferentes:

- PCR de inserción descrito en la Figura 5.
- Inserción con enzimas de restricción más ligación con el kit FastLink (Epicentre). Este método es el mismo utilizado en la construcción de los minigenes anteriores.



**Figura 5. PCR de inserción.** La PCR de inserción es un procedimiento similar al descrito por Agilent en el kit GeneMorph II EZClone Domain Mutagenesis Kit. Esta metodología utiliza el fragmento amplificado como “megacebador”. Las “colas” o secuencias no solapantes durante la amplificación de ADN genómico, se unen de forma específica a la región donde queremos introducir el amplificado. A continuación se realiza una PCR de alta fidelidad con la enzima Pfu Turbo en un termociclador Veriti (Applied Biosystems) con el programa: 95°C 2’; 20 ciclos de 95°C 30”- 59°C 30”- 68°C 2’ por Kb ;68°C 20 y con la mezcla presentada en la tabla 24. El plásmido molde es eliminado mediante digestión con la enzima de restricción sensible a metilación DpnI.

#### IV.1.c.1. MGBR2 2-9

El MGBR2 2-9 contiene los exones 2, 3, 4, 5, 6, 7, 8 y 9 y parte de sus intrones flanqueantes. Su construcción requirió la amplificación de cinco regiones genómicas diferentes, mediante PCR de alta fidelidad con la enzima Phusion Hot Start II (Thermo Scientific) a partir de 250ng de ADN.

- La primera amplificación de 2365 pb contiene los exones 4, 5, 6 y 7 y se realizó con cebadores específicos que incluyen secuencias reconocidas por las enzimas de restricción EcoRI (FW) y XhoI (RV).
- Un fragmento de 808 pb en el que se incluye el exón 3 se amplificó de forma individual con cebadores específicos que incluyen sitios de reconocimiento de las enzimas de restricción EcoRI y BspEI (FW) y EcoRI (RV).
- La amplificación de una región de 572 pb con cebadores específicos incluyó al exón 8. Los cebadores utilizados contienen sitios de restricción para las enzimas XhoI (FW) y KpnI y HindIII (RV).
- El exón 9 fue amplificado en un fragmento de 649 pb con cebadores específicos que contienen secuencias de restricción para las enzimas HindIII (FW) y KpnI (RV).
- Por último, se amplificó una región de 715 pb que incluye al exón 2 con cebadores específicos con los sitios de restricción de las enzimas EagI (FW) y BspEI (RV).

Las reacciones de amplificación se realizaron en un termociclador Veriti (Applied Biosystems) con el programa: 98°C 30"; 35 ciclos x [98°C 10"- 58-60°C 15" - 72°C 15" por kb]; 72°C 1' con la Mezcla presentada en la Tabla 19.

**Tabla 18. Secuencia de cebadores y tamaño de amplificado de los amplificadores usados para la construcción del MGBR2 2-9.**

Cebadores BRCA2-pSAD Clonación Minigen 2-9		
Cebador	secuencia 5' - 3'	Tamaño
MGBR2_ex4-7_EcoRI_AatII-FW	CACACAGAATTCCACACAGACGCTCTCTGGTACATTGGTAGGTTTTCA	2364 pb
MGBR2_ex4-7_XhoI-RV	CACACACTCGAGCCAAATACTGCCTCTTAAACGCGCGTTTAAGAGGC AGTATTTGGCTCGAGTGTGTG	
MGBR2_ex3_EcoRI_BspI_BstBI-FW	CACACAGAATTCACACATCCGGACACACATTCGAAGCCATCTTGTA ACTTTTGTGAA	808 pb
MGBR2_ex3_EcoRI-RV	CACACAGAATTTCTGGCATAAGAAAGACGTAATT	
MGBR2_ex8_XhoI-FW	CACACACTCGAGGTCATAACCTCACAGCATCATCT	572 pb
MGBR2_ex8_KpnI_HindIII-RV	CACACAGGTACCCACACAAAGCTTCAACTTTAACAGAGAGACAGCA	
MGBR2_ex9_HindIII-FW	CACACAAAGCTTTCCTAGTGGTGCAAGATTTTCAT	649 pb
MGBR2_ex9_KpnI-RV	CACACAGGTACCTGGTGCCAATTA AAAAGAGTAG	
MGBR2_ex2_EagI-FW	CACACACGGCCGAAACTAAATCGTATGAAAATCCT	715 pb
MGBR2_ex2_BspEI-RV	CACACATCCGGATAAGCTACATAACAAAAGCAC	

Tabla 19. Mezcla de amplificación de secuencias genómicas en la construcción del MBBR2 2-9.

PCR Phusion	
Buffer HF	1x
Cebador FW	0,2 $\mu$ M
Cebador RV	0,2 $\mu$ M
dNTPs	1 mM
Enzima Phusion	1 Unidad
Agua	
Volumen final	50 $\mu$ l

El tamaño de la regiones genómicas amplificadas fue comprobado mediante electroforesis en tampón TBE de gel de agarosa al 1% teñido con Redsafe (INtRON Biotechnology). Las electroforesis se realizaron a 100V y una vez finalizadas los fragmentos se visualizaron en un transiluminador ultravioleta con el sistema GelDoC (BioRad). El tamaño de los fragmentos visualizados se estimó comparando los mismos con el marcador de tamaño molecular 1 Kb plus DNA ladder (Invitrogen).

La inserción de las diferentes regiones no se ha realizado de una sola vez, sino que se ha requerido la creación de minigenes intermedios en varias etapas. La clonación de todos los fragmentos amplificados para la construcción del MGBR2 2-9 se ha realizado mediante corte con enzimas de restricción (New England Biolabs, Ipswich, MA, EE.UU.) y ligación con el kit FastLink (Epicentre, Madison, WI, EE.UU.). Además, tanto el vector como los productos de la amplificación y los minigenes intermedios, una vez cortados, fueron purificados mediante la extracción de bandas de gel de agarosa de bajo punto de fusión al 0,8% (Agarose LM; Conda/Pronadisa, Torrejón de Ardoz, España) y el uso del kit Wizard SV Gel and PCR clean-up system (Promega) según las instrucciones del fabricante.

El procedimiento de inserción de los diferentes fragmentos que conforman el MGBR2 2-9 se ha realizado en diferentes pasos:

- En primer lugar se procedió a la digestión tanto el vector pSAD como del fragmento amplificado que contiene los exones 4, 5, 6 y 7 con las enzimas *EcoRI* y *XhoI*. La ligación de ambos constituyó el minigen intermedio MGIBR2 4-7.

- Tanto el MGiBR2 4-7 como el fragmento amplificado de ADN genómico que incluye el exón 3 se digirieron con la enzima *EcoRI*. La ligación de ambos resultó en el MGiBR2 3-7.
- El MGiBR2 3-7 fue digerido junto con el producto de la amplificación que contiene el exón 8 con las enzimas *XhoI* y *KpnI*. La ligación de ambos constituyó el MGiBR2 3-8.
- El fragmento amplificado con el exón 9 fue introducido en el MGiBR2 3-8 tras ser cortados previamente con las enzimas de restricción *HindIII* y *KpnI*.
- El minigen final MGBR2 2-9 es el resultado de insertar el fragmento que contiene el exón 2 en el MGiBR2 3-8. Ambos, fueron digeridos con las enzimas *EagI* y *BspEI* y seguidamente ligados.

#### IV.1.c.2. MGBR2 16-20

La construcción del MGBR2 16-20 que contiene los exones 16 ,17 ,18 ,19 y 20 ha requerido la amplificación de tres regiones genómicas mediante PCR de alta fidelidad con la encima Phusion (ThermoScientific) a partir de 250ng de ADN. La amplificación de los fragmentos que incluyen los exones 17 y 18 por un lado y 16 por otro se ha realizado con cebadores especiales para el clonaje inserción de fragmentos a través de PCR. Dichos cebadores constan de una secuencia que hibrida de forma específica con los extremos de la región genómica que queremos amplificar y de una secuencia específica o “cola” de unión a la región del vector donde queremos realizar la inserción. Con esta clase de cebadores se amplificó un fragmento de 1556 pb que incluye a los exones 17 y 18 y otra de 805 pb que contiene el exón 16. Por su parte, el fragmento de 1188 pb que incluye los exones 19 y 20 se amplificó con cebadores específicos que contienen dianas de restricción de las enzimas *XhoI* y *KpnI*. Las reacciones de amplificación se realizaron en un termociclador Veriti (Applied Biosystems) con el programa: 98°C 30”;35 ciclos de 98°C 10”- 58-60°C 15” - 72°C 15” por kb; 72°C 1’ con la mezcla presentada en la tabla 19.

**Tabla 20. Secuencia de cebadores y tamaño de amplificado de los amplificadores usados para la construcción del MGBR2 16-20.**

Cebadores BRCA2-pSAD Clonación Minigen 16-20		
Cebador	secuencia 5' - 3'	Tamaño
MGBR2_ex17-18_Ins-FW	GCTCTAGAACTAGTGGATCCCCCGGTCAGTATGATACTTTGATACATGT	1556 pb
MGBR2_ex17-18_Ins-RW	ATAAGCTTGATATCGAATTCCCTGCAGTGGAAATAGGGATCTGATCAA	
MGBR2_ex16_Ins-FW	GCTCTAGAACTAGTGGATCCCCCGGCCCTTTTGGTTTCCCATCTAAGT	805 pb
MGBR2_ex16_Ins-RW	GACATGTATCAAAGTATCATACTGAATAAATGCCTAAGAAAAATGT	
MGBR2_ex19-20_XhoI-FW	CACACACTCGAGATAGCATTAAAGAACTTGTAGCA	1188 pb
MGBR2_ex19-20-KpnI-RV	CACACAGGTACCATTACAAATGGCTTAGACCTGA	

El tamaño de productos PCR fue comprobado mediante electroforesis en tampón TBE de gel de agarosa al 1% teñido con Redsafe (INtRON Biotechnology). Las electroforesis se realizaron a 100V y una vez finalizadas los fragmentos se visualizaron en un transiluminador ultravioleta con ayuda del sistema GelDoC (BioRad) para digitalizar las imágenes. El tamaño de los fragmentos se estimó comparando los mismos con el marcador de tamaño molecular conocido como *ladder* de 1kb (Invitrogen).

- En la construcción del MGBR2 16-20 el primer fragmento introducido en el vector pSAD fue el que incluye los exones 17 y 18. Este proceso se realizó mediante PCR de inserción (Figura 5) originando el minigen intermedio MG<sub>i</sub>BR2 17-18.
- A continuación, se construyó el MG<sub>i</sub>BR2 16-18 al añadir el fragmento de ADN genómico que contiene el exón 16 al MG<sub>i</sub>BR2 17-18 mediante PCR de inserción.
- Por último, tanto el amplicón que contienen el los exones 19 y 20 como el MG<sub>i</sub>BR2 16-18 fueron digeridos con las enzimas de restricción XhoI y BamHI. La ligación de ambos fragmentos constituyó el minigen final MGBR2 16-20.

Todos los fragmentos amplificados fueron purificados mediante el aislamiento de bandas de ADN en gel de agarosa de bajo punto de fusión al 0,8%, Agarose LM (Conda/Pronadisa) y el uso del kit Wizard SV Gel and PCR clean-up system (Promega) según las instrucciones del fabricante como paso previo antes de realizar las inserciones.

#### IV.1.c.3. MGBR2 19-27

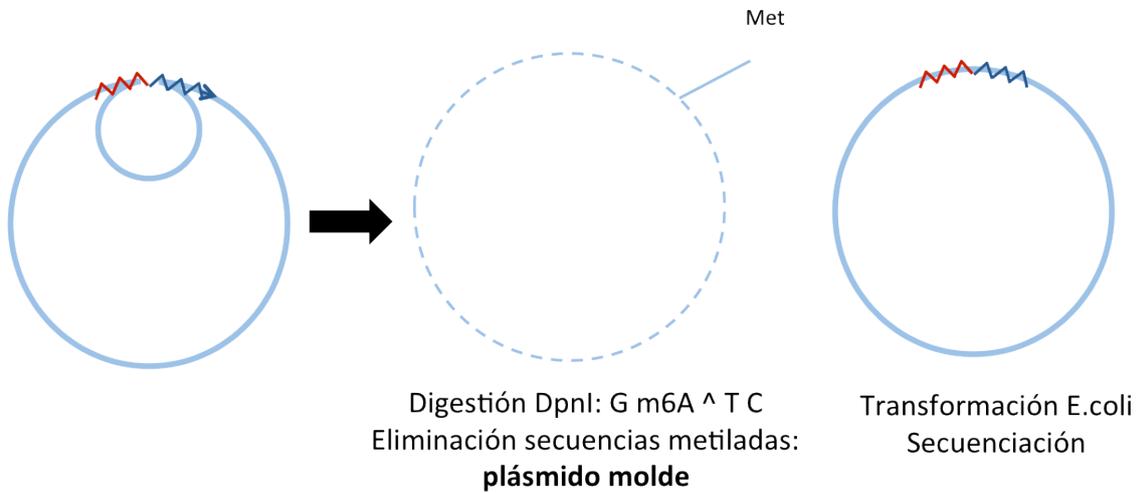
El minigen MGBR2 19-27 contiene los últimos 9 exones del gen *BRCA2*. Tanto su diseño como su construcción ha supuesto un verdadero reto técnico. En el proceso de construcción podemos diferenciar dos fases.

En la primera fase se produce la inserción de los exones 19, 20, 21, 22, 23 y 24 en el vector pSPL3., para lo cual se amplifican dos regiones mediante PCR de alta fidelidad con la polimerasa Phusion (ThermoScientific) a partir de 250ng de ADN. La primera región amplificada de ADN genómico consta de 3784 pb y contiene los exones 21, 22, 23 y 24. Dicha amplificación se realizó con cebadores específicos que incluyen en sus extremos secuencias de reconocimiento de las enzimas de restricción XhoI y BamHI. Por otro lado, se amplificó una región genómica que contiene los exones 19 y 20 mediante cebadores específicos propios del proceso de inserción mediante PCR, donde en los extremos están formados por secuencias que únicamente hibridan con las secuencias del vector donde se desea realizar la inserción. Las reacciones de amplificación se realizaron en un termociclador Veriti (Applied Biosystems) con el programa: 98°C 30";35 ciclos de 98°C 10" - 58-60°C 15" - 72°C 15" por kb]; 72°C 1' con la mezcla presentada en la Tabla 19.

El proceso de construcción de los minigenes intermedios en la primera etapa de creación del MGBR2 19-27 consta de las siguientes etapas:

- Tanto el vector pSPL3 como el fragmento 21-24 fueron digeridos con las enzimas de restricción XhoI y BamHI. La ligación de ambos con el kit FastLink (Epicentre) constituyó el minigen intermedio MGIBR2 21-24.
- El tamaño del inserto 21-24 es el más grande que se ha clonado directamente en este trabajo llegando a las 3784 pb. Con el fin de disminuir el tamaño del MGIBR2 21-24 antes de proseguir introduciendo más insertos, se procedió a eliminar 2000 pb del intrón 21. La delección de este fragmento se realizó mediante una técnica que hemos denominado "delección por PCR" y que se muestra en la Figura 6. Utilizando los cebadores MGBR2\_IVS21\_del2000-FW y RV mostrados en la tabla 21. El resultado es un MGIBR2 21-24\_del\_2000.

- En la última inserción de la primera fase se incluyó el fragmento que contiene los exones 19 y 20 en el MGiBR2 21-24\_del\_2000 mediante PCR de inserción, procedimiento descrito en la Figura 5. Por lo que finalmente se obtuvo el MGi 19-24.



**Figura 6. Delección por PCR.** La delección por PCR es un procedimiento en el cual se diseñan cebadores de unos 50 pb de longitud. La mitad del cebador hibrida con una región del plásmido inmediatamente anterior al inicio de la región que se quiere deleccionar y la otra mitad se une de forma específica a la secuencia posterior al final de la región que se quiere eliminar. La hibridación del cebador con el vector genera un bucle en la secuencia del vector que se quiere eliminar. A continuación, se realiza un PCR con la enzima de alta fidelidad Pfu Turbo en un termociclador Veriti (Applied Biosystems) con el programa: 95°C 2'; 35 ciclos de 95°C 30"- 59°C 30"- 68°C 2' por Kb ; 68°C 20 y con la mezcla presentada en la Tabla 24. El plásmido molde es eliminado mediante digestión con la enzima de restricción sensible a metilación DpnI.

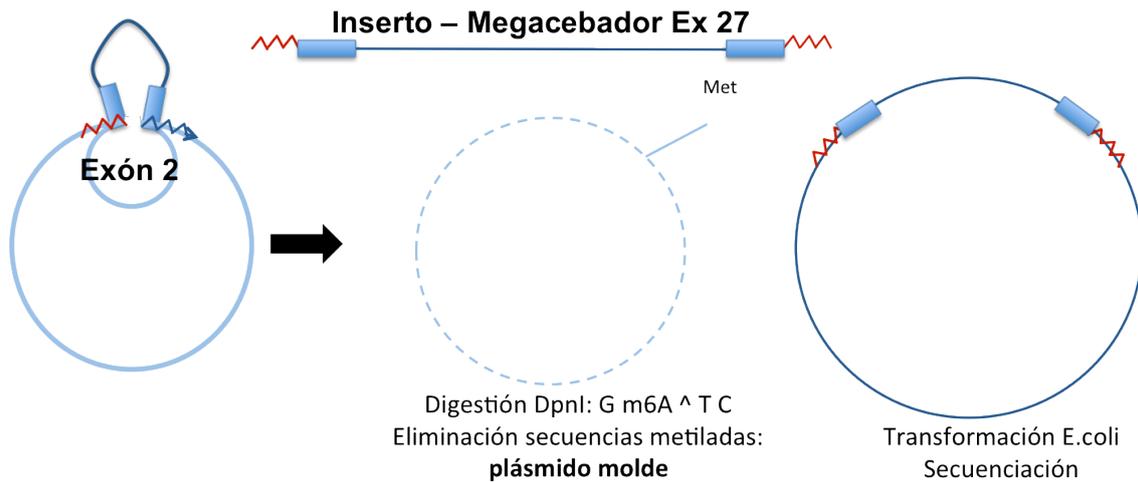
**Tabla 21. Secuencia de cebadores y tamaño de amplificado de los amplificadores usados para la construcción del MGBR2 19-27.**

Cebadores BRCA2-pSAD Clonación Minigen 19-27		
Cebador	secuencia 5' - 3'	Tamaño
MGBR2_ex21-24_XhoI-FW	CACACACTCGAGCCTTCTTTGGGTGTTTTATGCT	3784 pb
MGBR2_ex21-24-BamHI-RV	CACACAGGATCCAAATCAGAGGTTCAAAGAGGCT	
MGBR2_IVS21_del2000-FW	TTAAAATGCAGGCCAAAACCTTGTTTTAGCCATCTGTAATGTAGTTGGTGA	del 2000 pb
MGBR2_IVS21_del2000-RV	TCACCAACTACATTACAGATGGCTAAAAACAAGTTTTGCCTGCATTTTAA	
MGBR2_ex19-20_Ins-FW	TCTGGAGCTCGAGCACACATCCGGAATAGCATTAAAGAACTTGTAGCA	1214 pb
MGBR2_ex19-20_Ins-RV	CCAAGCATAAAAACACCCAAAGAAGGATTACAAATGGCTTAGACCTGA	
MGBR2_EX19-24_XhoI-FW	CACACACTCGAGCACACATCCGGAATAGCATTAAAGAACTTGTAGCA	2913 pb
MGBR2_EX19-24_KpnI-RV	CACACAGGATCCCACACAGGTACCAAATCAGAGGTTCAAAGAGGCT	
MGBR2_ex25_Ins-FW	GTAAGCCTCTTTGAACCTCTGATTTAATCTGTACTCTGTAGCAAT	997 pb
MGBR2_ex25_Ins-RV	CAAAAGGGAACAAAAGCTGGGTACCTCAGTGTCTCTATCTGGGATT	
MGBR2_ex26-KpnI-FW	CACACAGGTACCAGTAAATAGAGCTAGGACTTGA	850 pb
MGBR2_ex26-KpnI-RV	CACACAGGTACCTTGATTTGCATCTACTGTGATT	
MGBR2_ex27_Ins&del-FW	CCAGCTTTTGTTCCTTTTGTGAGGGTCTTCCTTCCTTTTCATGTCATTT	994 pb
MGBR2_ex27_Ins&del-RV	ATCTGTCTCTGTCTCTCTCCACCTAATTTCTTTTCTCATTTGTGCAAC	

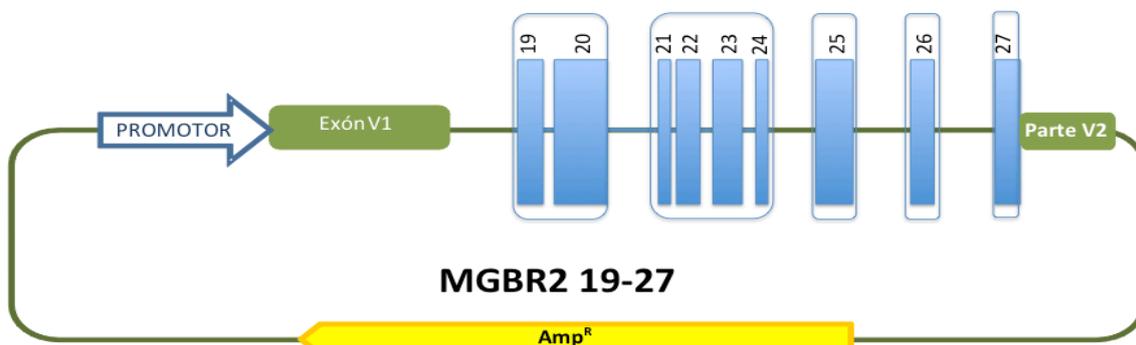
La segunda fase de la construcción del MGBR2 19-27 se inicia con la amplificación mediante cebadores específicos del fragmento total insertado en el MG<sub>i</sub> 19-24 y que contiene los exones 19, 20, 21, 22, 23 y 24. Dichos cebadores contienen sitios de restricción para las enzimas XhoI y KpnI. Además, se amplificaron otras tres regiones genómicas. En la región amplificada que contiene el exón 26 se han utilizado cebadores específicos que contienen en sus extremos unas secuencias de reconocimiento para la enzimas de restricción KpnI. Mientras que en la amplificación de los fragmentos que contienen los exones 25 y 27 se han utilizado cebadores específicos para la técnica PCR de inserción. Las reacciones de amplificación se realizaron con la enzima con la enzima Phusion (Thermo Scientific) a partir de 250ng de ADN, en un termociclador Veriti (Applied Biosystems) con el programa: 98°C 30";35 ciclos de 98°C 10"- 58-60°C 15" - 72°C 15" por kb; 72°C 1' con la mezcla presentada en la Tabla 19.

Las etapas que llevaron a la construcción final del MGBR2 19-27 fueron las siguientes:

- El fragmento de ADN amplificado del MG<sub>i</sub> 19-24 junto con el vector pSAD fueron digeridos con las enzimas XhoI y KpnI. Seguidamente, se procedió a la ligación de ambos con el kit FastLink (Epicentre). El resultado es un MG<sub>i</sub> 19-24 construido en pSAD y nombrado como MG<sub>i</sub> 19-24\_pS para distinguirlo del construido en pSPL3.
- El fragmento amplificado de ADN genómico fue insertado en el MG<sub>i</sub> 19-24\_pS a través de el protocolo de PCR de inserción según el procedimiento descrito en la Figura 5 y por lo que finalmente se obtuvo el MG<sub>i</sub> 19-25.
- Al MG<sub>i</sub> 19-25 le fue agregado al amplicón que contenía el exón 26 mediante el kit FastLink (Epicentre), previa digestión de ambos con la enzima de restricción KpnI. El resultado fue el MG<sub>i</sub> 19-26.
- En el último proceso de ligación se combinaron las técnicas de inserción y delección por PCR. El fragmento amplificado se utilizó como megacebador según el proceso de inserción por PCR, de modo que el exón 27 con su sitio aceptor natural sustituye a parte del exón 2 de del vector pSAD, eliminando su sitio aceptor y manteniendo las secuencias de hibridación de los cebadores específicos del vector (Figura 7).



**Figura 7. Proceso de inserción y delección mediante PCR en un solo paso.** El procedimiento es muy similar al de la PCR de inserción, se diseñan cebadores de unos 50 pb de longitud. Parte del cebadores hibridan en los extremos de la región genómica que se quiere insertar. La amplificación de esta región genera un megacebador cuyos extremos hibridan con una región del plásmido inmediatamente anterior y posterior de la región que se quiere deleccionar e insertar la región genómica amplificada. La hibridación de los extremos del megacebador con el vector genera un bucle en éste y en la secuencia del vector que se quiere eliminar. A continuación, se realiza un PCR con la enzima de alta fidelidad Pfu Turbo en un termociclador Veriti (Applied Biosystems) con el programa: 95°C 2'; 35 ciclos de 95°C 30"- 59°C 30"- 68°C 2' por Kb ; 68°C 20 y con la mezcla presentada en la tabla 24. El plásmido molde es eliminado mediante digestión con la enzima de restricción sensible a metilación DpnI.



**Figura 8. Esquema del MGBR2 19-27 con el exón 2 del pSAD fusionado al último exón del gen BRCA2.**

El tamaño de la regiones genómicas amplificadas fue comprobado mediante electroforesis a 100V en tampón TBEx1 de gel de agarosa al 1% usando como estándar de tamaño 1 kb Plus DNA ladder (Invitrogen).

Tanto los fragmentos amplificados y usados como megacebadores como aquellos amplificados con cebadores que incluyen dianas de restricción y el vector pSPL3, pSAD y el MG<sub>i</sub> 19-25 fueron digeridos con las enzimas de restricción correspondientes y purificados mediante el aislamiento de bandas de ADN en gel de

agarosa de bajo punto de fusión al 0,8%, Agarose LM (Conda/Pronadisa) con el kit Wizard SV Gel and PCR clean-up system (Promega) antes del clonaje.

#### IV.1.d. Transformación

Las células competentes de *Escherichia coli* DH5 $\alpha$  (Invitrogen), cuyo genotipo es: F-  $\phi$ 80*lacZ* $\Delta$ M15  $\Delta$ (*lacZYA-argF*)U169 *recA1 endA1 hsdR17*( $r_k^-$ ,  $m_k^+$ ) *phoA supE44 thi-1 gyrA96 relA1*  $\lambda^-$ , fueron transformadas siguiendo el protocolo facilitado por la casa comercial y crecidas medio S.O.C. (Hanahan, 1983), con todos los minigenes finales e intermedios contruidos según el orden de creación descrito en los apartados anteriores.

#### IV.1.e. Selección de colonias

##### IV.1.e.1. Selección de colonias recombinantes

Las células transformadas fueron sembradas en placas de medio LB (Luria-Bertani) cuya composición es: 1% Triptona, 0,5% extracto de levadura, 1% NaCl premezclado con agar-LB-Miller (Merck/VWR, West Chester, PA, EE.UU.) y a las que se les añade ampicilina a la concentración final de 100  $\mu$ g/ml como método de selección de colonias transformadas. Las bacterias fueron incubadas a 37°C durante 18 horas. Únicamente en las placas en las que se sembraron las primeras construcciones realizadas en pSAD como el MGiBR2 4-7; MGiBR2 18-19 y MGiBR2 19-24\_pS se añadió además de ampicilina, IPTG (Fermentas) y XGal (Fermentas). En estos casos las colonias blancas fueron seleccionadas frente a las de color azul que no incluían el inserto.

##### IV.1.e.2. Rastreo de colonias por PCR

El rastreo de colonias por PCR es una forma sencilla y rápida de constatar que el proceso de inserción se ha realizado con éxito. El método consiste en una amplificación directa por PCR con la enzima Phire (ThermoScientific) del extracto resultante del hervido de cada colonia aislada. Se han utilizado dos tipos de combinaciones para la PCR:

- La primera consiste en una amplificación con dos cebadores del vector flanqueantes al sitio de clonación múltiple del vector cuya secuencia se muestra en la Tabla 22.

Con esta amplificación se pueden distinguir los vectores con el inserto deseado del resto por el tamaño del amplificado.

**Tabla 22. Cebadores flanqueantes al sitio de clonación múltiple (SCM) de los vectores pSPL3 y pSAD.**

Combinaciones de Cebadores flanqueantes al sitio de clonación múltiple en el pSPL3 y pSAD	
Cebador	Secuencias 5' - 3'
SCM-FW	CCTTGGGATGTTGATGAT
SCM-RV	TTGCTTCCTTCCACACAG

- La segunda consiste en una amplificación que combina un cebador flanqueante al sitio de clonación múltiple con otro cebador localizado en el inserto.

Las reacciones de amplificación se realizaron en un termociclador Veriti (Applied Biosystems) con el programa: 98°C 30"; 35 ciclos de 98°C 10"- 58-60°C 15" - 72°C 15" por kb; 72°C 1' con la mezcla presentada en la tabla 23.

**Tabla 23. Mezcla de amplificación en el rastreo de colonias por PCR.**

PCR de colonias	
Buffer	5x
Cebador FW	0,2 µM
Cebador RV	0,2 µM
dNTPs	1 mM
Enzima Phire	1 Unidad
Agua	
Volumen final	25 µl

En todos los casos el tamaño de la regiones genómicas amplificadas fue comprobado mediante electroforesis en tampón TBE<sub>x</sub>1 de gel de agarosa al 1,5 % con el marcador de tamaño molecular conocido o de 1kb Plus DNA ladder (Invitrogen).

#### IV.1.f. Extracción y comprobación de las construcciones

Para la comprobación de las colonias bacterianas seleccionadas se procedió, en primer lugar, a la extracción de ADN plasmídico. Los minigenes construidos en pSPL3 fueron extraídos con el kit de kit NucleoSpin Plasmid (Macherey-Nagel), mientras que para extraer las construcciones en pSAD se ha utilizado el kit PureYield™ Plasmid Miniprep System (Promega). En ambos casos, la extracción se realizó a partir de cultivo bacteriano en saturación tras dieciocho horas a 37°C en medio líquido LB (Caldo-LB-Miller, Merk) suplementado con ampicilina a una concentración final de 100 µg/ml.

Las reacciones de secuencia se efectuaron con el kit BigDye Terminator Cycle Sequencing v3.1 (Applied Biosystems) con los cebadores flanqueantes al sitio de clonación múltiple de la Tabla 22 siguiendo las indicaciones del fabricante. Los ddntps no incorporados en la reacción de secuencia fueron eliminados mediante el método de perlas magnéticas del kit CleanSEQ (Agencourt Bioscience) según las instrucciones del fabricante. La carrera electroforética capilar se llevó a cabo en un secuenciador automático ABI Prism® 3130 (Applied Biosystems) con capilares de 36 cm, utilizando POP-7 (Applied Biosystems) como polímero de electroforesis y tampón de carrera con EDTA (Applied Biosystems) utilizando las condiciones de la Tabla 14.

Los datos brutos de la carrera se procesaron con el programa informático Data Collection 3.0 (Applied Biosystems). El análisis y la elaboración de los cromatogramas se han realizado con la aplicación Chromas Pro (Technelysium).

## IV.2. Mutagénesis dirigida

La técnica de mutagénesis dirigida es una adaptación del protocolo del kit QuickChange (Stratagene) que permite generar cualquier alteración detectada en el ADN de un paciente o mutaciones artificiales en los minigenes construidos. De esta forma, no es necesario obtener muestras biológicas de los pacientes.

### IV.2.a. Mutaciones puntuales

El protocolo de mutagénesis dirigida consta de una amplificación mediante PCR de alta fidelidad con la enzima Pfu Turbo, de un minigen *wild type* (wt) con cebadores que introducen la mutación que se desea reproducir. Las reacciones de amplificación se realizaron en un termociclador Veriti (Applied Biosystems) con el programa: 95°C 2'; 20 ciclos de 95°C 30"- 55°C 30"- 68°C 2' por kb; 68°C 20 y con las mismas condiciones de la mezcla presentada en la Tabla 24. Excepcionalmente, las mutaciones reproducidas en el MGBR2 19-27 se generaron siguiendo las recomendaciones del kit de mutagénesis QuikChange Lightning Site-Directed Mutagenesis (Agilent). Una vez finalizada la amplificación, en ambos procedimientos se elimina el plásmido molde con la enzima de restricción sensible a metilación *DpnI* (Invitrogen). Las células DH5 $\alpha$  (Invitrogen) se transformaron con los minigenes que portaban las mutaciones

siguiendo el proceso descrito en el apartado IV.1.d. Por último, los minigenes con las mutaciones introducidas fueron extraídos y comprobados mediante secuenciación de la región de interés donde se ha realizado la mutagénesis de forma similar a la referida en el apartado IV.1.f.

**Tabla 24. Mezcla de amplificación PCR de mutagénesis, PCR inserción y delección.**

PCR Pfu Turbo	
Buffer	1x
Plásmido molde	100-200 ng
Cebador FW	0,2 $\mu$ M
Cebador RV	0,2 $\mu$ M
dNTPs	1 mM
Enzima Pfu Turbo	1,5 Unidades
Agua	
Volumen final	25 $\mu$ l

Un total de 186 cebadores específicos de mutagénesis dirigida han sido diseñados para estudiar las 93 mutaciones en los MGBR2 19-20, MGBR2 23-24, MGBR2 16-20 y MGBR2 19-27 seleccionadas en los estudios bioinformáticos previos. En ciertos casos se han realizado varias mutagénesis sobre el mismo minigen para estudiar el efecto combinado de varias mutaciones en el *splicing*, como es el caso de las delecciones realizadas en el intrón 19.

**Tabla 25. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 19 del MGBR2 19-20.**

Cebadores Mutagénesis Dirigida Minigen 19-20 <i>BRCA2</i> . Exón 19		
Mutación	Cebador	secuencia 5' - 3'
c.8362T>C	MD_BR2_c.8362T>C_FW	AGTACTCGGCCTGCTCGCCGGTATACCAAACCTGGATT
	MD_BR2_c.8362T>C_RV	AATCCAAGTTTGGTATACCGGCGAGCAGGCCGAGTACT
c.8368A>T	MD_BR2_c.8368A>T_FW	TGTCGCTGGTATTCCAAACCTGGATTCTTTCCTGAC
	MD_BR2_c.8368A>T_RV	GTCAGGAAAGAATCCAAGTTTGGAAATACCAGCGAGCA
c.8378G>A	MD_BR2_c.8378G>A_FW	CGTGGTATACCAAACCTGAATTCTTTCCTGACCCCTAGAC
	MD_BR2_c.8378G>A_RV	GTCTAGGGTCAGGAAAGAATTCAAGTTTGGTATACCAGCG
c.8394_8396delTAGinsAA	MD_BR2_c.8394_8396delTAGinsAA_FW	AAACTTGGATTCTTTCCTGACCCAAACCTTTCCTCTGTC
	MD_BR2_c.8394_8396delTAGinsAA_RV	GCAGAGGAAAAGGTTTGGGTGAGAAAGAATCCAAGTTT
c.8398C>T	MD_BR2_c.8398C>T_FW	TTCTTTCCTGACCCCTAGATCTTTCCTCTGCCCTTATCA
	MD_BR2_c.8398C>T_RV	TGATAAGGGCAGAGGAAAAGATCTAGGGTCAGGAAAGAA
c.8399C>G	MD_BR2_c.8399C>G_FW	ATTCTTTCCTGACCCCTAGACGTTTTTCTCTGCCCTTAT
	MD_BR2_c.8399C>G_RV	ATAAGGGCAGAGGAAAACGCTCTAGGGTCAGGAAAGAAT
c.8435G>A	MD_BR2_c.8435G>A_FW	CTTTTCAGTGATGAAGGAAATGTTGGTTGTGTTGATGTAA
	MD_BR2_c.8435G>A_RV	TTACATCAACACAACCAACATTTCTTTCATCACTGAAAAG
c.8438G>A	MD_BR2_c.8438G>A_FW	CTTTTCAGTGATGGAGAAAATGTTGGTTGTGTTGATGT
	MD_BR2_c.8438G>A_RV	ACATCAACACAACCAACATTTTCTCCATCACTGAAAAG
c.8486A>T	MD_BR2_c.8486A>T_FW	CAAAGAGCATACCCTATACTGGTATGATGTATTCTTGAAAC
	MD_BR2_c.8486A>T_RV	GTTTCAAGAATACATCATACCAGTATAGGGTATGCTCTTTG
c.8487+1G>A	MD_BR2_c.8487+1G>A_FW	TCAAAGAGCATACCCTATACAGATATGATGTATTCTTGAAA
	MD_BR2_c.8487+1G>A_RV	TTTCAAGAATACATCATATCTGTATAGGGTATGCTCTTTGA
c.8487+3A>G	MD_BR2_c.8487+3A>G_FW	AAGAGCATACCCTATACAGGTGTGATGTATTCTTGAAACTT
	MD_BR2_c.8487+3A>G_RV	AAGTTTCAAGAATACATCACACCTGTATAGGGTATGCTCTT
c.8487+4T>C	MD_BR2_c.8487+4T>C_FW	AAGAGCATACCCTATACAGGTACGATGTATTCTTGAAACTT
	MD_BR2_c.8487+4T>C_RV	AAGTTTCAAGAATACATCGTACCTGTATAGGGTATGCTCTT
c.8487+8G>A	MD_BR2_c.8487+8G>A_FW	GCATACCCTATACAGGTATGATATATTCTTGAAACTTACC
	MD_BR2_c.8487+8G>A_RV	GGTAAGTTTCAAGAATATATCATACCTGTATAGGGTATGTC
c.8484A>T	MD_BR2_c.8484A>T_FW	TTCAAAGAGCATACCCTATTCAGGTATGATGTATTCTTGA
	MD_BR2_c.8484A>T_RV	TCAAGAATACATCATACCTGAATAGGGTATGCTCTTTGAA
c.8486A>C	MD_BR2_c.8486A>C_FW	CAAAGAGCATACCCTATACCGGTATGATGTATTCTTGAAA
	MD_BR2_c.8486A>C_RV	TTTCAAGAATACATCATACCGGTATAGGGTATGCTCTTTG
c.8487+3A>C	MD_BR2_c.8487+3A>C_FW	AAGAGCATACCCTATACAGGTCTGATGTATTCTTGAAACT
	MD_BR2_c.8487+3A>C_RV	AGTTTCAAGAATACATCAGACCTGTATAGGGTATGCTCTT
c.8487+3A>T	MD_BR2_c.8487+3A>T_FW	AAGAGCATACCCTATACAGGTTTGTGATGTATTCTTGAAACT
	MD_BR2_c.8487+3A>T_RV	AGTTTCAAGAATACATCAAACCTGTATAGGGTATGCTCTT

**Tabla 26. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el intrón 19 del MGBR2 19-20.**

Cebadores Mutagénesis Dirigida Minigen 19-20 <i>BRCA2</i> . Intrón 19.		
Mutación	Cebador	secuencia 5' - 3'
c.8487+54_8487+355del	MD_BR2_c.8487+54_8487+355del_FW	TATTTCTTTCTTTTGATACAATTAATGATACAATTAAGTGAATGTTATA
	MD_BR2_c.8487+54_8487+355del_RV	TATACCATTCAAGTTAATGTATCATTAATGTATCAAAAAGAAAGAAATA
c.8487+31_8487+42del	MD_BR2_c.8487+31_8487+42del_FW	ATTCTTGAACTTACCATATAGATACAATTAATTTGTTTGTGTTG
	MD_BR2_c.8487+31_8487+42del_RV	CAAACAAACAAATTAATTTGATCTATATGGTAAGTTTCAAGAAT
c.8488-49_8488-44del	MD_BR2_c.8488-49_8488-44del_FW	GCCTGGCCTGATACAATTAAGTTTATATATGTGACTTTTTTTGG
	MD_BR2_c.8488-49_8488-44del_RV	CCAAAAAAGTCACATATATAAAGTTAATTTGATCAGGCCAGGC

**Tabla 27. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 20 del MGBR2 19-20.**

Cebadores Mutagénesis Dirigida Minigen 19-20 <i>BRCA2</i> Exón 20		
Mutación	Cebador	secuencia 5' - 3'
c.8488-1G>A	MD_BR2_c.8488-1G>A_FW	GTGTAACACATTATTACAATGGATGGAGAAGACATCATC
	MD_BR2_c.8488-1G>A_RV	GATGATGTCTTCTCCATCCATTGTAATAATGTGTTACAC
c.8488-2A>G	MD_BR2_c.8488-2A>G_FW	GTGTAACACATTATTACGATGGATGGAGAAGACATCATC
	MD_BR2_c.8488-2A>G_RV	GATGATGTCTTCTCCATCCATCGTAATAATGTGTTACAC
c.8512T>A	MD_BR2_c.8512T>A_FW	GAGAAGACATCATCTGGAATATACATATTTTCGCAATGA
	MD_BR2_c.8512T>A_RV	TCATTGCGAAATATGTATATTTCCAGATGATGTCTTCTC
c.8518delA	MD_BR2_c.8518delA_FW	AAGACATCATCTGGATTATACTATTTTCGCAATGAAAGA
	MD_BR2_c.8518delA_RV	TCTTTCATTGCGAAATAGTATAATCCAGATGATGTCTT
c.8539G>A	MD_BR2_c.8539G>A_FW	TTATACATATTTTCGCAATGAAAGAGAGAAAGAAAAGGAAG
	MD_BR2_c.8539G>A_RV	CTTCCTTTTCTTTCTCTCTTTTCATTGCGAAATATGTATAA
c.8560_8560delT	MD_BR2_c.8560_8560delT_FW	AAGAAAAGGAAGCAGCAAAAATGTGGAGGCCCAAC
	MD_BR2_8c.8560_8560delT_RV	GTTGGCCCTCCACATTTTTGCTGCTTCTTTTCTT
c.8572C>A	MD_BR2_c.8572C>A_FW	AAGCAGCAAAATATGTGGAGGCCAAACAAAAGAGACTA
	MD_BR2_c.8572C>A_RV	TAGTCTTTTGTGTTGGCCCTCCACATATTTGCTGCTT
c.8609A>G;8611G>T	MD_BR2_c.8609A>G;8611G>T_FW	GCCTTATTCCTACTAAAATTCGGTAGGAATTTGAAGACATG
	MD_BR2_c.8609A>G;8611G>T_RV	CATGTCTTCAAAATTCCTACCGAATTTTAGTGAATAAGGC

**Tabla 28. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 23 del MGBR2 23-24.**

Cebadores Mutagénesis Dirigida Minigen 23-24 <i>BRCA2</i> Exón 23		
Mutación	Cebador	secuencia 5' - 3'
c.8954-3C>G	MD_BR2_c.8954-3C>G_FW	TTTCATCTTTCTCCAAGAGTTATACTGAGTATTTGGC
	MD_BR2_c.8954-3C>G_RV	GCCAAATACTCAGTATAACTCTTTGGAGAAAGATGAGAA
c.8954-1_8955delinsAA	MD_BR2_c.8954-1_8955delinsAA_FW	TCTCATCTTTCTCCAACAAAATACTGAGTATTTGGCGT
	MD_BR2_c.8954-1_8955delinsAA_RV	ACGCCAAATACTCAGTATTTTGTGGAGAAAGATGAGA
c.8961_8964delGAGT	MD_BR2_c.8961_8964delGAGT_FW	TTTCTCCAACAGTTATACTATTTGGCGTCCATCATCAG
	MD_BR2_c.8961_8964delGAGT_RV	CTGATGATGGACGCCAAATAGTATAACTGTTGGAGAAA
c.8969G>A	MD_BR2_c.8969G>A_FW	AACAGTTATACTGAGTATTTAGCGTCCATCATCAGATTT
	MD_BR2_c.8969G>A_RV	AAATCTGATGATGGACGCTAAATACTCAGTATAACTGTT
c.8972G>A	MD_BR2_c.8972G>A_FW	TACTGAGTATTTGGCATCCATCATCAGATTT
	MD_BR2_c.8972G>A_RV	AAATCTGATGATGGATGCCAAATACTCAGTA
c.9004G>A	MD_BR2_c.9004G>A_FW	ATATTCTCTGTTAACAAAAGGAAAGAGATACAG
	MD_BR2_c.9004G>A_RV	CTGTATCTCTTTCCCTTTTGTAAACAGAGAATAT
c.9006A>T	MD_BR2_c.9006A>T_FW	CTCTGTTAACAGATGGAAAGAGATACAG
	MD_BR2_c.9006A>T_RV	CTGTATCTCTTTCCATCTGTTAACAGAG
c.9872G>A;9006A>T	MD_BR2_c.9872G>A;9006A>T_FW	Combinación
	MD_BR2_c.9872G>A;9006A>T_RV	Combinación
c.9076C>T	MD_BR2_c.9076C>T_FW	ATCTGAAAGAGCTAACATATAGTTAGCAGCGACAAAAAA
	MD_BR2_c.9076C>T_RV	TTTTTTGTCGCTGCTAACTATATGTTAGCTCTTTTCAGAT
c.9100C>T	MD_BR2_c.9100C>T_FW	CAGCGACAAAAAAACTTAGTATCAACAACCTACCGGTAC
	MD_BR2_c.9100C>T_RV	GTACCGGTAGTTGTTGATACTAAGTTTTTTTTGTCGCTG
c.9101A>G	MD_BR2_c.9101A>G_FW	GCAGCGACAAAAAAACTCGGTATCAACAACCTACCGGTAC
	MD_BR2_c.9101A>G_RV	GTACCGGTAGTTGTTGATACTGAGTTTTTTTTGTCGCTGC
c.9116C>T	MD_BR2_c.9116C>T_FW	CTCAGTATCAACAACCTACTGGTACAAACCTTTCATTG
	MD_BR2_c.9116C>T_RV	CAATGAAAGGTTTGTACCAGTAGTTGTTGATACTGAG
c.9117G>A	MD_BR2_c.9117G>A_FW	ACTCAGTATCAACAACCTACCAGTACAACCTTTCATTGT
	MD_BR2_c.9117G>A_RV	ACAATGAAAGGTTTGTACTGGTAGTTGTTGATACTGAGT
c.9117+1G>T	MD_BR2_c.9117+1G>T_FW	TCAGTATCAACAACCTACCGTTACAACCTTTCATTGTAA
	MD_BR2_c.9117+1G>T_RV	TTACAATGAAAGGTTTGTAAACGGTAGTTGTTGATACTGA
c.9117+1G>A	MD_BR2_c.9117+1G>A_FW	TCAGTATCAACAACCTACCGATACAACCTTTCATTGTAA
	MD_BR2_c.9117+1G>A_RV	TTACAATGAAAGGTTTGTATCGGTAGTTGTTGATACTGA

**Tabla 29. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 24 del MGBR2 23-24.**

Cebadores Mutagénesis Dirigida Minigen 23-24 <i>BRCA2</i> Exón 24		
Mutación	Cebador	secuencia 5´- 3´
c.9118-2A>G	MD_BR2_c.9118-2A>G_FW	TTTTGTTTTGTTTTCTGTGGGTTTCAGATGAAATTTTAT
	MD_BR2_c.9118-2A>G_RV	ATAAAATTTTCATCTGAAACCCACAGAAAACAAAACAAA
c.9118-6C>T	MD_BR2_c.9118-6C>T_FW	TGAATTTTTGTTTTGTTTTTGTAGGTTTCAGATGAAAT
	MD_BR2_c.9118-6C>T_RV	ATTTTCATCTGAAACCTACAAAAACAAAACAAAATTC
c.9148C>T	MD_BR2_9148C>T_FW	AATTTTATTTTCAGATTTACTAGCCACGGGAGCCCCTTC
	MD_BR2_9148C>T_RV	GAAGGGGCTCCCGTGGCTAGTAAATCTGAAATAAAATT
c.9177_9177delA	MD_BR2_c.9177_9177delA_FW	CCCCTTCACTTCAGCAATTTTGTAGATCCAGACTTTTCAG
	MD_BR2_c.9177_9177delA_RV	CTGAAAGTCTGGATCTAAAAATTGCTGAAGTGAAGGGG
c.9227G>A	MD_BR2_c.9227G>A_FW	TTCTGAGGTGGACCTAATAGTATTTGTCGTTTCTGTGT
	MD_BR2_c.9227G>A_RV	ACAACAGAAACGACAAATACTATTAGGTCCACCTCAGAA
c.9248_9256+7del	MD_BR2_c.9248_9256+7del_FW	AGATGAAATTTTATTTTCAGATTTACTTCACTTCAGCAAAT TTTTAGATCC
	MD_BR2_c.9248_9256+7del_RV	GGATCTAAAAATTTGCTGAAGTGAAGTAAATCTGAAATAA AATTCATCT
c.9252_9255delAACA insTT	MD_BR2_c.9252_9255delAACAinsTT_FW	TCGTTTCTGTTGTGAAAAATTTGGTAATGCACAATATAGT
	MD_BR2_c.9252_9255delAACAinsTT_RV	ACTATATTGTGCATTACCAATTTTTCACAACAGAAACGA
c.9256G>T	MD_BR2_c.9256G>T_FW	TTTCTGTTGTGAAAAAACATGTAATGCACAATATAGTT
	MD_BR2_c.9256G>T_RV	AACTATATTGTGCATTACATGTTTTTTTCACAACAGAAA
c.9256+1G>A	MD_BR2_c.9256+1G>A_FW	TCTGTTGTGAAAAAACAGATAATGCACAATATAGTTAA
	MD_BR2_c.9256+1G>A_RV	TAACTATATTGTGCATTATCTGTTTTTTTCACAACAGA
c.9196C>G	MD_BR2_c.9196C>G_FW	GATCCAGACTTTGAGCCATCTTGTTTC
	MD_BR2_c.9196C>G_RV	GAACAAGATGGCTCAAAGTCTGGATC
c.9213G>T	MD_BR2_c.9213G>T_FW	GCCATCTTGTCTGATGTGGACCTAATAGG
	MD_BR2_c.9213G>T_RV	CCTATTAGGTCCACATCAGAACAAGATGGC
c.9218A>G	MD_BR2_c.9218A>G_FW	CATCTTGTCTGAGGTGGCCTAATAGGATTTGTCCG
	MD_BR2_c.9218A>G_RV	CGACAAATCCTATTAGGCCACCTCAGAACAAGATG
c.9172A>G	MD_BR2_c.9172A>G_FW	CGGGAGCCCCTTCACTTCGGCAAATTTTGTAGATCCAG
	MD_BR2_c.9172A>G_RV	CAATGAAAGGTTTGTACCAGTAGTTGTTGATACTGAG

**Tabla 30. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 21 del MGBR2 19-27.**

Cebadores mutaciones Exón 21 Minigen 19-27 BRCA2		
Mutación	Cebador	secuencia 5' - 3'
c.8633-24_8634del	BR2_c.8633-24_8634del-FW	TTGAATTTACAGTTTAGTGAA AACACAACAAAACCATA
	BR2_c.8633-24_8634de -RV	TATGGTTTTGTGTGTTTTCACTAAACTGTAAATTCAA
c.8723T>G	BR2_c.8723T>G-FW	AGAGCTTTATGAAGCAGGGAAGAATGCAGCAGAC
	BR2_c.8723T>G-RV	GTCTGCTGCATTCTTCCCTGCCTTCATAAAGCTCT
c.8732C>T	BR2_c.8732C>T-FW	AAGCAGTGAAGAATGTAGCAGACCCAGCTTACCT
	BR2_c.8732C>T-RV	AGGTAAGCTGGGTCTGCTACATCTCTTCACTGCTT
c.8748ins4	BR2_c.8748ins4-FW	AGACCCAGCTTACTTACCTTGAGGTGAGAGTAAGAG
	BR2_c.8748ins4-RV	CTCTTACTCTCTCACCTCAAGGTAAGTAAGCTGGGTCT
c.8754+4A>G	BR2_c.8754+4A>G-FW	GCTTACCTTGAGGTGGGAGAGTAAGAGGACATATAA
	BR2_c.8754+4A>G-RV	TTATATGTCCTCTTACTCTCCCACCTCAAGGTAAGC
c.8754+5G>A	BR2_c.8754+5G>A-FW	CAGCTTACCTTGAGGTGAAAGAGTAAGAGGACATAT
	BR2_c.8754+5G>A-RV	ATATGTCCTCTTACTCTTTCACCTCAAGGTAAGCTG
c.8754+5G>T	BR2_c.8754+5G>T-FW	CCAGCTTACCTTGAGGTGATAGAGTAAGAGGACATA
	BR2_c.8754+5G>T-RV	TATGTCCTCTTACTCTATCACCTCAAGGTAAGCTGG
c.8647C>T	BR2_c.8647C>T-FW	AGAAAACACAACAAAATCATATTTACCATCACGTGC
	BR2_c.8647C>T-RV	GCACGTGATGGTAAATATGATTTTGTGTGTTTTCT
c.8754G>A	BR2_c.8754G>A-FW	GACCCAGCTTACCTTGAAGTGAGAGAGTAAGAGGAC
	BR2_c.8754G>A-RV	GTCCTCTTACTCTCTCACTTCAAGGTAAGCTGGGTC

**Tabla 31. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 22 del MGBR2 19-27.**

Cebadores mutaciones exón 22 Minigen 19-27 BRCA2		
Mutación	Cebador	secuencia 5' - 3'
c.8755-1G>A	BR2_c.8755-1G>A-FW	ATCTTAAATGGTCACAAGGTTATTTTCAGTGAAGAGC
	BR2_c.8755-1G>A-RV	GCTCTTCACTGAAATAACCTTGTGACCATTTAAGAT
c.8773C>T	BR2_c.8773C>T-FW	GAAGAGCAGTTAAGAGTCTTGAATAATCACAGGCAA
	BR2_c.8773C>T-RV	TTGCCTGTGATTATTCAAGACTCTTAACTGCTCTTC
c.8789delA	BR2_8789delA-FW	CAGTTAAGAGCCTTGATAATCACAGGCAAATGTTG
	BR2_8789delA-RV	CAACATTTGCCTGTGATTATCAAGGCTCTTAACTG
c.8915T>G	BR2_c.8915T>G-FW	ACAACCGTGTGGAAGTGGCGTATTGTAAGCTATTC
	BR2_c.8915T>G-RV	GAATAGCTTACAATACGCCACTTCCACACGGTTGT
c.8948_8953+5del	BR2_c.8948_8953+5del-FW	GCTATTTCAAAAAAGAAAAAGTATGTAATGCTTTGTTTT
	BR2_c.8948_8953+5del-RV	AAAACAAAGCATTTACATACTTTTCTTTTTTTGAATAGC
c.8953+1G>T	BR2_c.8953+1G>T-FW	CAAAAAAGAAAAAGATTCAGTTAAGTATGTAATGCT
	BR2_c.8953+1G>T-RV	AGCATTTACATACTTAACTGAATCTTTTTCTTTTTTTG

**Tabla 32. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 25 del MGBR2 19-27.**

Cebadores mutaciones exón 25 Minigen 19-27 BRCA2		
Mutación	Cebador	secuencia 5'-3'
c. 9276T>G	MDBR2_c.9276T-G-FW	CTTGCCCTTTTCGTCTAGTTGTCAGACGAATGTTAC
	MDBR2_c.9276T-G-RV	GTAACATTCGTCTGACAACCTAGACGAAAGGGCAAG
c. 9301C>G	MDBR2_c.9301C-G-FW	ACGAATGTTACAATTTAGTGGCAATAAAGTTTGGAT
	MDBR2_c.9301C-G-RV	ATCCAAAACCTTATTGCCACTAAATGTAACATTCGT
c.9501G>A	MDBR2_c.9501G-A-FW	AATGAAAAATACTGTGAAGTAAGTTACTTTTCAGC
	MDBR2_c.9501G-A-RV	GCTGAAAAGTAACCTTACTTCAACAGTATTTTTCATT
c. 9257-1G>C	MDBR2_c.9257-1G-C-F	CTTTTTTTTCCATTCTACGACTTGCCCTTTTCGTC
	MDBR2_c.9257-1G-C-R	GACGAAAGGGCAAGTCGTAGAATGGAAAAAAG
c. 9257-18C>A	MDBR2_c.9257-18CA-FW	CTATAATAACATTCCTTTTATTTTTTTTCCATCTAGG
	MDBR2_c.9257-18CA-RV	CCTAGAATGGAAAAAATAAAGAATGTTATTATAG
c. 9501+3A>T	MDBR2_c.9500+3A-T -FW	GAAAAATACTGTTGAGGTTAGGTTACTTTTCAGCATC
	MDBR2_c.9501+3A-T-R	GATGCTGAAAAGTAACCTAACCTCAACAGTATTTTTC
c.9501+4A>G	MDBR2_c.9500+4A-G-F	AAAAATACTGTTGAGGTAGGTTACTTTTCAGCATCA
	MDBR2_c.9501+4A-G-R	TGATGCTGAAAAGTAACCTACCTCAACAGTATTTTTT
c. 9454G>A	MDBR2_c.9454G-A-FW	TTTTCTGCTAGTCCAAAAAGGGCCACTTTCAAGAGA
	MDBR2_c.9454G-A-RV	TCTCTTGAAAGTGGCCCTTTTTTGGACTAGCAGAAAA
c. 9466delC	MDBR2_c.9466delC-FW	CCAAAAGAGGGCCACTTTAAGAGACATTCACAAAAT
	MDBR2_c.9466delC-RV	ATTTTGTGAATGTCTCTTAAAGTGGCCCTCTTTTGG
c. 9275_9278del	MDBR2c. 9275_9278del-FW	GACTTGCCCTTTTCGTCTGTGACAGCAATGTTACAA
	MDBR2_c. 9275_9278del -RV	TTGTAACATTCGTCTGACAGACGAAAGGGCAAGTC
c.9501+9A>C	MDBR2_c.9501+9A>C-FW	TACTGTTGAGGTAAGGTTCCCTTTTCAGCATCACCACA
	MDBR2_c.9501+9A>C-RV	TGTGGTGATGCTGAAAAGGAACCTTACCTCAACAGTA

**Tabla 33. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 26 del MGBR2 19-27.**

Cebadores mutaciones exón 26 Minigen 19-27 BRCA2		
Mutación	Cebador	secuencia 5'-3'
c. 9502-2A>C	MDBR2_c.9502-2A-C-FW	TTTCCACTTATTTTCTTGAATATTGACATACTTTGC
	MDBR2_c.9502-2A-C-RV	GCAAAGTATGTCAATATTGGAAGAAAATAAGTGGAAA
c. 9502-12T>G	MDBR2_c.9502-12TG-FW	TAAAGCAGCTTTTCCACGTATTTTCTTAGAATATTGA
	MDBR2_c.9502-12TG-RV	TCAATATTCTAAGAAAATACGTGAAAAGCTGCTTTA
c. 9550C>G	MDBR2_c.9550C-G-FW	ACAAGCTTATGCATATAGTCATGCAAATGATCCC
	MDBR2_c.9550C-G-RV	GGGATCATTTGCATGCACATATGCATAAGCTTGT
c. 9610A>G	MDBR2_c.9610A-G-FW	GTACTTCAGGGCCGTACGCTGCTCAAATCATTCCT
	MDBR2_c.9610A-G-RV	AGGAATGATTTGAGCAGCGTACGGCCCTGAAGTAC
c. 9611C>G	MDBR2_c.9611C-G-FW	TACTTCAGGGCCGTACAGTGCTCAAATCATTCCTG
	MDBR2_c.9611C-G-RV	CAGGAATGATTTGAGCACTGTACGGCCCTGAAGTA
c. 9613G>CT	MDBR2_c.9613GC-CT-FW	ACTTCAGGGCCGTACACTCTTCAAATCATTCCTGGTAC
	MDBR2_c.9613GC-CT-RV	GTACCAGGAATGATTTGAAGAGTGTACGGCCCTGAAGT
c. 9646C>T	MDBR2_c.9646C-T-FW	GTACAGGAAACAAGCTTTTGGTAAGTTAATGTAAA
	MDBR2_c.9646C-T-RV	TTTACATTAACCTACCAAAGCTTGTTCCTGTAC

**Tabla 34. Mutaciones y secuencia de cebadores utilizados para la mutagénesis dirigida en el exón 27 del MGBR2 19-27.**

Cebadores mutaciones exón 27 Minigen 19-27 <i>BRCA2</i>		
Mutación	Cebador	secuencia 5' - 3'
c.9649-2A>G	MDBR2_ex27-c9649-2A>G-FW	CGTTTTTCATTTTTTTTATCGGATGTCTTCTCCTAATTG
	MDBR2_ex27-c9649-2A>G-RV	CAATTAGGAGAAGACATCCGATAAAAAAATGAAAACG
c.9649-5A>G	MDBR2_ex27-c9649-5A>G-FW	CTACGTTTTTCATTTTTTTTGTGATGTCTTCTCCTAA
	MDBR2_ex27-c9649-5A>G-RV	TTAGGAGAAGACATCTGACAAAAAATGAAAACGTAG
c. 9698G>T	MDBR2_ex27-c9698G>T-FW	AAAGTCCTTTATCACTTTTTATGGCCAAAAGGAAGTC
	MDBR2_ex27-c9698G>T-RV	GACTTCCTTTTGGCCATAAAAAGTGATAAAGGACTTT

**Tabla 35. Mutación c.7806-14T>C, secuencia de cebadores utilizados para la mutagénesis dirigida en el MGBR2 16-20.**

Cebadores mutación c.7806-14T>C Minigen 16-20 <i>BRCA2</i>		
Mutación	Cebador	secuencia 5' - 3'
c.7806-14T>C	MDBR2_ex17-c.7806-14T>C-FW	TTTATGATAATATTCTACCTTTATTTGTTTCAGGGCTC
	MDBR2_ex17-c.7806-14T>C-RV	GAGCCCTGAACAAATAAAGGTAGAATATTATCATAAA

#### IV.2.b. Microdeleciones

Una aplicación particular basada en los principios de la mutagénesis dirigida se ha utilizado para realizar un total de 20 microdeleciones exónicas en el MGBR2 16-20 y otras 34 en el MGBR2 19-27. Esta técnica permite detectar mediante ensayos funcionales de *splicing* secuencias exónicas esenciales (*enhancers* o silenciadores) en la regulación del *splicing*. El protocolo utilizado para las microdeleciones del MGBR2 16-20 es el mismo descrito en la Figura 6 para los procesos de delección por PCR, mientras que en el MGBR2 19-27 se realizaron con el kit de mutagénesis QuikChange Lightning Site-Directed Mutagenesis (Agilent). Los minigenes producto de las delecciones se transforman posteriormente en células DH5 $\alpha$  (Invitrogen) siguiendo el protocolo descrito en el apartado IV.1.d. El rastreo de colonias que contenían los plásmidos con las delecciones fueron seleccionados de forma similar a la desarrollada en el apartado IV.1.e. y comprobados mediante secuenciación de la región de interés como se ha descrito en el apartado IV.1.f.

**Tabla 36. Microdeleciones y secuencia de cebadores utilizados para la PCR de deleción en el MGBR2 16-20.**

Cebadores Microdeleciones Minigen 16-20 BRCA2		
Microdeleción	Cebador	secuencia 5' - 3'
c.7620_7649del	MDBR2_ex16-c.7620_7649del-FW	TTGTGTGTGTTTATTTTGTGTAGCTAAAAATTAACAGCAAAAATGCAGAG
	MDBR2_ex16-c.7848_7877del-RV	CTCTGCATTTTTTGCTGTTAATTTTTAGCTACACAAAATAAACACACACAA
c.7645_7674del	MDBR2_ex16-c.7645_7674del-FW	GTATACGTATGGCGTTTCTAACATTCTTTTCAGTTTCACACTGAAGATT
	MDBR2_ex16-c.7645_7674del-RV	AATCTTCAGTGTGAAACTGAAAAGAATGTTTAGAAACGCCATACGTATAC
c.7675_7704del	MDBR2_ex16-c.7675_7704del-FW	AAAAATTAACAGCAAAAATGCAGAGATGGTGGATGGCTCATACCCCTCAA
	MDBR2_ex16-c.7675_7704del-RV	TTGGAGGGTATGAGCCATCCACCATCTCTGCATTTTTGCTGTTAATTTTT
c.7748_7772del	MDBR2_ex16-c.7748_7772del-FW	ACTGAAAAGGAATACAGTTGGCTGAAAAGGCTGAAAAGAAGATTTTA
	MDBR2_ex16-c.7748_7772del-RV	TAAAATCTCTTTTCCAGCCTTTCCAGCCAAGTATTCTTTCCAGT
c.7773_7802del	MDBR2_ex16-c.7773_7802del-FW	ATGGTGGATGGCTCATACCCCTCCAATAGGTACTCTATGCAAAAGATTGT
	MDBR2_ex16-c.7773_7802del-RV	ACAATCTTTTGCATAGAGTACCTATTGGAGGGTATGAGCCATCCACCAT
c.7808_7837del	MDBR2_ex17-c.7808_7837del-FW	ATATTCTACTTTTATTTGTTTCAGGAGCTTATTTCTAGAATTTGGGTTTA
	MDBR2_ex17-c.7808_7837del-RV	TAAACCCAAATCTAGAAATAAGCTCCCTGAACAAATAAAGTAGAATAT
c.7833_7862del	MDBR2_ex17-c.7833_7862del-FW	CTCTGTGTGACACTCCAGGTGTGGATAATCACTATAGATGGATCATATGG
	MDBR2_ex17-c.7833_7862del-RV	CCATATGATCCATCTATAGTATTATCCACACCTGGAGTGTACACAGAG
c.7863_7918del	MDBR2_ex17-c.7863_7918del-FW	AGCTTATTTCTAGAATTTGGGTTTAAAGGAATTTGCTAATAGATGCCTAAG
	MDBR2_ex17-c.7863_7918del-RV	CTTAGGCATCTATTAGCAAAATCCTTAAACCCAAATCTAGAAATAAGCT
c.7919_7948del	MDBR2_ex17-c.7919_7948del-FW	GCAGCTATGGAATGTGCCTTTCCFAAAAGGGTGCCTTCACTAAAATA
	MDBR2_ex17-c.7919_7948del-RV	TATTTTAGTTGAAGAAGCACCTTTTAGGAAAGGCACATCCATAGCTGC
c.7944_7973del	MDBR2_ex17-c.7944_7973del-FW	AGGAATTTGCTAATAGATGCCTAAGCAGGCAAGTTTAAAGCATTACATTA
	MDBR2_ex17-c.7944_7973del-RV	TAATGTAATGCTTTAAACTTGCTGCTTAGGCATCTATTAGCAAATTCCT
c.7979_8008del	MDBR2_ex18-c.7979_8008del-FW	TGCATTTTTGTTTTTCACTTTTAGATCGGCTATAAAAAAGATAATGGAAAG
	MDBR2_ex18-c.7979_8008del-RV	CTTTCCATTATCTTTTTTATAGCCGATCTAAAAGTGAAAACAAAATGCA
c.8004_8033del	MDBR2_ex18-c.8004_8033del-FW	ATGATACGGAAATGATAGAAGCAGGGATGACACAGCTGCAAAAACACTT
	MDBR2_ex18-c.8004_8033del-RV	AAGTGTTTTTGCAGCTGTGCATCCCTGCTTCTATCAATTTCCGTATCAT
c.8034_8158del	MDBR2_ex18-c.8034_8158del-FW	CGGCTATAAAAAAGATAATGGAAAGAACTTACAGATGGTGGTATGCTGT
	MDBR2_ex18-c.8034_8158del-RV	ACAGCATACCACCATCTGTAAGTTCTTTCCATTATCTTTTTTATAGCCG
c.8150_8274del	MDBR2_ex18-c.8150_8274del-FW	AGTAGTGCAGATACCCAAAAGTGGTGGGCTCTCCTGATGCCTGTACAC
	MDBR2_ex18-c.8150_8274del-RV	GTGTACAGGCATCAGGAGAGCCCACTTTTGGGTATCTGCACTACT
c.8274_8303del	MDBR2_ex18-c.8274_8303del-FW	AGATTATTCTTCATGGAGCAGAAGTGAAGCCCCAGAATCTCTTATGTTA
	MDBR2_ex18-c.8274_8303del-RV	TAACATAAGAGATTCTGGGGCTCAAGTTCTGCTCCATGAAGAATAATCT
c.8299_8328del	MDBR2_ex18-c.8299_8328del-FW	GGTGGGCTCTCCTGATGCTGTACAAAGGTAATTAATTGCACCTTTGG
	MDBR2_ex18-c.8299_8328del-RV	CCAAGAGTGCAAATTAATTTACCTTTGTACAGGCATCAGGAGAGCCACC
c.8334_8363del	MDBR2_ex19-c.8334_8363del-FW	AATATATTTATTAATTTGTCAGATGTATACCAAAGTGGATTCTTTCT
	MDBR2_ex19-c.8334_8363del-RV	AGGAAAGAATCCAAGTTGGTATACATCTGGACAAATTAATAAATATATT
c.8455_8484del	MDBR2_ex19-c.8455_8484del-FW	TGATGGAGGAAATGTGGTGTGTTTCAGGTATGATGTATTCTTGAAACTT
	MDBR2_ex19-c.8455_8484del-RV	AAGTTTCAAGAATACATCATACCTGAACACAACCAACATTTCTCCATCA
c.8490_8519del	MDBR2_ex20-c.8490_8519del-FW	GTGTGTGTAACACATTATTACAGTGATTTTCGCAATGAAAGAGAGGAAGAA
	MDBR2_ex20-c.8490_8519del-RV	TTCTTCTCTTTTTCATTGCGAAATCACTGTAATAATGTGTACACACAC
c.8600_8629del	MDBR2_ex20-c.8600_8629del-FW	CAAAAGAGACTAGAAGCCTTATTCAAAGGTAAAATAGTTATATGGTACA
	MDBR2_ex20-c.7773_7802del-RV	TGTACCATATAACTAATTTTACCTTTGAATAAGGCTTCTAGTCTCTTTTG

Tabla 37. Microdeleciones y secuencia de cebadores utilizados para la PCR de delección en el MGBR2 19-27.

Cebadores Microdeleciones Minigen 19-27 BRCA2		
Microdelección	cebador	secuencia 5' - 3'
c. 8334_8363del	MDBR2_ex19-c.8334_8363del-FW	AATATATTATTAATTTGTCCAGATGTATACCAAACCTGGATTCTTTCT
	MDBR2_ex19-c.8334_8363del-RV	AGGAAAGAATCCAAGTTTGGTATACATCTGGACAAATTAATAAATATATT
c.8359_8388del	MDBR2_ex19-c.8359_8388del-FW	TTCTGCTAACAGTACTCGGCCTGCTGACCCTAGACCTTTTCTCTGCCCT
	MDBR2_ex19-c.8359_8388del-RV	AGGCAGAGGAAAAGGTCTAGGGTCAGCAGCCGAGTACTGTAGCAGAA
c.8430_8459del	MDBR2_ex19-c.8430_8459del-FW	CTCTGCCCTTATCATCGCTTTTTCAGAATTATTCAAAGAGCATACCCTATA
	MDBR2_ex19-c.8430_8459del-RV	TATAGGGTATGCTCTTTGAATAATTCTGAAAAGCGATATAAGGCAGAG
c.8455_8484del	MDBR2_ex19-c.8455_8484del-FW	TGATGGAGGAAATGTTGGTTGTTCAGGTATGATGTATTCTTGAACCTT
	MDBR2_ex19-c.8455_8484del-RV	AAGTTTCAAGAATACATCATACCTGAACACAAACACATTCTCCATCA
c.8490_8519del	MDBR2_ex20-c.8490_8519del-FW	GTGTGTGTAAACACATTATTACAGTATTTCGCAATGAAAGAGAGGAAGAA
	MDBR2_ex20-c.8490_8519del-RV	TTCTCTCTCTTTCATGCGAAATCACTGTAATAATGTGTACACACAC
c.8515_8544del	MDBR2_ex20-c.8515_8544del-FW	GATGGAGAAGACATCATCTGGATTAAAGGAAGCAGCAAAATATGTGGAGG
	MDBR2_ex20-c.8515_8544del-RV	CCTCCACATATTTTGTCTCTTCTTAAATCCAGATGATGTCTTCTCCATC
c.8575_8604del	MDBR2_ex20-c.8575_8604del-FW	AGCAGCAAAATATGTGGAGGCCAAATTCAGGAGGAATTTGAAGAACATG
	MDBR2_ex20-c.8575_8604del-RV	CATGTTCTTCAAATCTCTCTGAATTTGGGCTCCACATATTTTGTCTGCT
c.8600_8629del	MDBR2_ex20-c.8600_8629del-FW	CAAAAGAGACTAGAAGCCTTATTCAAAGGTAATAGTTATATGGTACA
	MDBR2_ex20-c.8600_8629del-RV	TGTACCATATAACTAATTTTACCTTTGAATAAGGCTTCTAGTCTCTTTTG
c.8635_8664del	MDBR2_ex21-c.8635_8664del-FW	TAATAATCCTTTTGTCTTTCTTAGAAGCACTAACAGACAGCAAGTCTGTG
	MDBR2_ex21-c.8635_8664del-RV	CACGAACCTGTCTTGTGTAGTCTTCTAAGAAAACAAAAGGATTATTA
c.8660_8689del	MDBR2_ex21-c.8660_8689del-FW	AACACAACAAAACCATATTTTACCATCTTTGCAAGATGGTGCAGAGCTTTA
	MDBR2_ex21-c.8660_8689del-RV	TAAAGCTCTGCACCATCTTGAAGATGGTAAATATGGTTTTGTGTGTT
c.8697_8726del	MDBR2_ex21-c.8697_8726del-FW	CAAGACAGCAAGTTCGTGCTTTGCAGAATGCAGCAGACCCAGCTTACCTT
	MDBR2_ex21-c.8697_8726del-RV	AAGGTAAGCTGGGTCTGCTGCATTCTGCAAAGCAGCAACTTGTCTGCTTG
c.8722_8751del	MDBR2_ex21-c.8722_8751del-FW	AGATGGTGCAGAGCTTTATGAAGCAGAGGTGAGAGAGTAAGAGGCATAT
	MDBR2_ex21-c.8722_8751del-RV	ATATGTCTCTTACTCTCTCACCTCTGCTTATAAAGCTCTGCACCATCT
c.8757_8786del	MDBR2_ex22-c.8757_8786del-FW	TCCAATATCTTAAATGGTACAGGGGAATAATCACAGGCAATGTTGAAT
	MDBR2_ex22-c.8757_8786del-RV	ATTCAACATTTGCCTGTGATTATCCCTGTGACCATTAAAGATATTGGA
c.8782_8811del	MDBR2_ex22-c.8782_8811del-FW	TTATTTTCAAGTGAAGAGCAGTTAAGAGATAAGAAAACAGCTCAGATCCAGT
	MDBR2_ex22-c.8782_8811del-RV	ACTGGATCTGAGCTTGTCTTCTTATCTCTTAACTGCTCTTCACTGAAATAA
c.8896_8926del	MDBR2_ex22-c.8896_8926del-FW	AAAGGAACAAGGTTATCAAGGATAGCTATTCAAAAAAAGAAAAAGATT
	MDBR2_ex22-c.8896_8926del-RV	AATCTTTTCTTTTTTTGAATAGCTATCCCTTGATAAACCTGTTCCTTTT
c.8921_8950del	MDBR2_ex22-8921_8950del-FW	GTCACAACCGTGTGGAAGTTGCGTACAGGTAAGTATGTAATGCTTTGTT
	MDBR2_ex22-8921_8950del-RV	AACAAAGCATTTACATACTTACCTGTACGCAACTCCACACGGTTGTGAC
c.8956_8985del	MDBR2_ex23-c.8956_8985del-FW	CTTCTCATCTTTCTCCAACAGTTTATATTCTCTGTAAACAGAAGGAA
	MDBR2_ex23-c.8956_8985del-RV	TTCTCTGTGTTAACAGAGAAATATAAAACTGTTGGAGAAAGATGAGAAAG
c.8981_9010del	MDBR2_ex23-c.8981_9010del-FW	ATACTGAGTATTTGGCGTCCATCATAGAGATACAGAATTTATCATCTTGC
	MDBR2_ex23-c.8981_9010del-RV	GCAAGATGATAAATCTGTATCTCTATGATGGACGCCAAACTACTCAGTAT
c.9060_9089del	MDBR2_ex23-c.9060_9089del-FW	CAACTTCAAAATCTAAAAGTAAATCAAAAAAATCAGTATCAACAACATA
	MDBR2_ex23-c.9060_9089del-RV	TAGTTGTGATACTGAGTTTTTTTTGATTTACTTTTAGATTTGAAGTTG
c.9085_9114del	MDBR2_ex23-c.9085_9114del-FW	TGAAAGAGCTAACATACAGTTAGCACCGGTACAAACCTTTCATTGTAATT
	MDBR2_ex23-c.9085_9114del-RV	AATTACAATGAAAGGTTTGTACCGGTGCTAACTGTATGTTAGCTCTTTCA
c.9120_9149del	MDBR2_ex24-c.9120_9149del-FW	AATTTTTGTTTTGTTTTCTGTAGGTGCCACGGGAGCCCTTCACTTCAGC

	MDBR2_ex24-c.9120_9149del-RV	GCTGAAGTGAAGGGGCTCCCGTGGCACCTACAGAAAACAAAACAAAATT
c.9145_9174del	MDBR2_ex24-c.9145_9174del-FW	TTCAGATGAAATTTTATTTTCAGATTAATTTTTAGATCCAGACTTTCAGC
	MDBR2_ex24-c.9145_9174del-RV	GCTGAAAGTCTGGATCTAAAAATTTAATCTGAAATAAAATTTTCATCTGAA
c.9199_9228del	MDBR2_ex24-c.9199_9228del-FW	CAAATTTTGTAGATCCAGACTTTCAGTTTGTCTGTTCTGTTGTGAAAAAA
	MDBR2_ex24-c.9199_9228del-RV	TTTTTTTTCACAACAGAAACGACAACTGAAAGTCTGGATCTAAAAATTTG
c.9224_9253del	MDBR2_ex24-c.9224_9253del-FW	CCATCTTGTCTGAGGTGGACCTAACAGGTAATGCACAATATAGTTAATT
	MDBR2_ex24-c.9224_9253del-RV	AATTAACATATATTGTGCATTACCTGTTAGGTCCACCTCAGAACAGATGG
c.9259_9288del	MDBR2_ex25-c.9259_9288del-FW	CTTTTCTTTTTTTTCCATTCTAGGATGTTACAATTTACTGGCAATAAAGT
	MDBR2_ex25-c.9259_9288del-RV	ACTTTATTGCCAGTAAATTGTAACATCTAGAATGGAAAAAAGAAAAAG
c.9284_9313del	MDBR2_ex25-c.9284_9313del-FW	CTTGCCCTTTTCGTCTATTTGTCTAGTTGGATAGACCTTAATGAGGACAT
	MDBR2_ex25-c.9284_9313del-RV	ATGTCCTCATTAAAGTCTATCCAACCTGACAAATAGACGAAAGGGCAAG
c.9444_9473del	MDBR2_ex25-c.9444_9473del-FW	CTGAGATTTTTTCTGTGTTTTCTGCATCAACAAAATGAAAAACTGT
	MDBR2_ex25-c.9444_9473del-RV	AACAGTATTTTTTCATTTTGTGAATGCAGAAAACACAGAAAAATCTCCAG
c.9469_9498del	MDBR2_ex25-c.9469_9498del-FW	TAGTCCAAAAGAGGGCCACTTTCAGAGGTAAGGTTACTTTTCAGCATCA
	MDBR2_ex25-c.9469_9498del-RV	TGATGCTGAAAAGTAACTTACCTCTTGAAAGTGGCCCTCTTTTGACTA
c.9504_9533del	MDBR2_ex26-c.9504_9533del-FW	AGCTTTTCCACTTATTTTCTTAGAACAAGCTTATGCATATACTGCATGCA
	MDBR2_ex26-c.9504_9533del-RV	TGCATGCAGTATATGCATAAGCTTGTCTAAGAAAATAAGTGGAAAAGCT
c.9529_9558del	MDBR2_ex26-c.9529_9558del-FW	TATTGACATACTTTGCAATGAAGCAAAATGATCCCAAGTGGTCCACCCCAA
	MDBR2_ex26-c.9529_9558del-RV	TTGGGGTGGACCCTTGGGATCATTTGCTTCATTGCAAAGTATGTCAATA
c.99591_9620del	MDBR2_ex26-c.9591_9620del-FW	CCAAGTGGTCCACCCCACTAAAGACATTCCTGGTACAGGAAACAAGCTT
	MDBR2_ex26-c.9591_9620del-RV	AAGCTTGTTCCTGTACCAGGAATGCTTTTAGTTGGGGTGGACCCTTGG
c.9616_9645del	MDBR2_ex26-c.9616_9645del-FW	CTGTACTTCAGGGCCGTACACTGCTCTGGTAAGTTAATGTAACTCAAGG
	MDBR2_ex26-c.9616_9645del-RV	CCTTGAGTTTACATTAACCTACCAGAGCAGTGTACGGCCCTGAAGTACAG
c.9651_9680del	MDBR2_ex27-c.9651_9680del-FW	CTACGTTTTCATTTTTTTATCAGATAAGTCCTTTATCACTTTGTATGGCC
	MDBR2_ex27-c.9651_9680del-RV	GGCCATACAAAAGTGATAAAGGACTTATCTGATAAAAAATGAAAACGTAG
c.9676_9705del	MDBR2_ex27-c.9676_9705del-FW	GTCTTCTCCTAATTGTGAGATATATAAAGGAAGTCTGTTTCCACACCTG
	MDBR2_ex27-c.9676_9705del-RV	CAGGTGTGAAAACAGACTTCCTTTTATATATCTCACAATTAGGAGAAGAC

### IV.3. Transfección

La transfección es la técnica con la que todos los minigenes construidos *wild type*, intermedios y generados por mutagénesis fueron introducidos en las células humanas en cultivo.

#### IV.3.a. Células humanas en cultivo

Las líneas celulares humanas utilizadas en los ensayos funcionales con minigenes híbridos fueron:

- HeLa: células epiteliales humanas de adenocarcinoma de cérvix (Jones *et al.*, 1971; Macville *et al.*, 1999).

- MCF-7: células tumorales obtenidas de efusión pleural de una paciente con carcinoma mamario metastásico (Brooks *et al.*, 1973).

Ambas líneas celulares se cultivaron en medio D-MEN (Invitrogen, Carlsbad, CA, EE.UU) suplementado con 10% de FBS (Invitrogen), 1% de L-glutamina (Invitrogen) y 1% de Penicilina/Estreptomicina (Invitrogen). Su crecimiento se realizó en un incubador a 37°C con 5% de CO<sub>2</sub>. Se realizaron dos subcultivos semanales mediante tripsinización (Gibco/Invitrogen) a intervalos regulares de tiempo.

#### IV.3.b. Protocolo de transfección

Las transfecciones se realizaron en células HeLa, a excepción de los plásmidos con las mutaciones c.8486A>C, c.8488-1G>A, c.9117G>A, c.9256G>T, c.8948\_8953+5del, c.8953+1G>T, c.9649-2A>G que fueron transfectados también en células MCF-7, para comparar los resultados obtenidos en ambas líneas celulares. Las transfecciones se realizaron con Lipofectamina 2000 (Invitrogen) siguiendo el protocolo indicado por el fabricante. En cada transfección se utilizaron dos tipos de medio: OPTIMEM (Gibco/Invitrogen) y D-MEN con un 10% de FBS (Invitrogen) pero sin antibióticos. Tras 24 h de cultivo, se obtuvieron las células mediante tripsinización (Gibco/Invitrogen) y centrifugación. Habría que destacar que se realizó una transfección especial con el MGBR2 19-27 y el portador de la mutación c.8948\_8953+5del con cicloheximida (Sigma, St. Louis, MO, EE.UU.) a una concentración final de 10 µg/ml al medio e incubando durante 4h antes de la extracción de ARN. La cicloheximida es una molécula que inhibe el MND en las células y se utilizó para realizar una evaluación del efecto del NMD en la degradación de las isoformas de *splicing* que portan un codón de parada prematuro.

#### IV.3.c. Extracción de ARN

El ARN total se extrajo de las células transfectadas mediante el kit Nucleospin RNA II (Macherey-Nagel) siguiendo las instrucciones del fabricante. A continuación, se obtuvo ADNc mediante el kit High Capacity (Applied Biosystems) o el kit Transcriptor First Strand cDNA Synthesis kit (Roche) a partir de ARN procedente de células transfectadas con los minigenes *wild type* y mutantes construidos en pSPL3 y pSAD .

#### IV.4. Comprobación de isoformas

Todos los patrones de *splicing* anómalos producidos tras una reacción de RT-PCR fueron caracterizados mediante tres metodologías: electroforesis en gel de agarosa, electroforesis capilar y secuenciación capilar.

##### IV.4.a. RT-PCR

Como se mencionó en el apartado III. la RT-PCR es una variante de la PCR en la que se amplifica el ADN complementario resultante de la transcripción. La diferencia principal respecto al protocolo descrito con anterioridad es que en este caso el ADNc no procede de ARN de linfocitos de un paciente, sino de las células transfectadas. Además, para todas las amplificaciones realizadas se ha usado al menos un cebador específico de la secuencia exónica del vector. Con ello, únicamente se amplifican los transcritos procedentes del vector y no los procedentes del genoma de las células transfectadas.

**Tabla 38. Secuencia de cebadores utilizados para la RT-PCR de ensayos funcionales de *splicing* con minigenes híbridos.**

Cebadores RT-PCR PSPL3	
Cebador	Secuencias 5' - 3'
RT-pSPL3-FW	TCACCTGGACAACCTCAAAG
RT-pSPL3-RV	TGAGGAGTGAATTGGTCGAA
RT-pSAD-FW	TGCTAAGCTGAGTGAAGT
RT-pSAD-RV	GGAGTGAATTGGTCGAATG

Las amplificaciones se realizaron con los cebadores mostrados en la Tabla 38 y solapantes ambos a regiones exónicas de los vectores pSPL3 o pSAD.

La excepción la marca el MGBR2 19-27 y su derivados mutantes, ya que debido al tamaño de sus transcritos, superior a las 2 kb fue necesario realizar las amplificaciones mediante diferentes parejas de cebadores en función de los exones donde se encuentra la mutación reproducida manteniendo siempre para la amplificación un cebador específico de los exones del vector. La combinación y secuencia de los cebadores se muestra en la Tabla 39.

**Tabla 39. Secuencia de cebadores utilizados para la RT-PCR de ensayos funcionales de *splicing* con el MGBR2 19-27 en combinación con RT-pSAD-FW y RT-pSAD-RV.**

Combinaciones de Cebadores RT-PCR MGBR2 19-27		
Cebadores utilizados en combinación con RT-pSAD-FW		
Cebador	Secuencias 5' - 3'	Exones analizados
RTBR2_EX22-RV	TGGATCTGAGCTTGTTCCTT	19/20/21
RTBR2-23-RV	ATTCTGTATCTCTTTCCTTCTGTT	21
RTBR2_EX25-RV	CACAGAAAAATCTCCAGCAA	22/23/24
Cebadores utilizados en combinación con RT-pSAD-RV		
Cebador	Secuencias 5' - 3'	Exones analizados
RTBR2-24-FW	TTTTTAGATCCAGACTTTCAGC	25
RTBR2_EX25-FW	TTTGCTGGAGATTTTCTGTG	26/27

Las reacciones de amplificación se realizaron con la enzima Go-Taq (Promega) en un termociclador Veriti (Applied Biosystems) con el programa: 95°C-2'; 35 ciclos 95°C 30"-58°C 30"-72°C 1'30";72°C 10'.

#### IV.4.a.1. Electroforesis en gel

El producto de las amplificaciones fue comprobado mediante electroforesis en tampón TBEx1 en geles de agarosa al 1,5% para fragmentos inferiores a 1 kb y al 1% para fragmentos superiores, con el marcador de tamaño molecular 1 kb Plus DNA ladder de 1kb (Invitrogen). Cuando alguno de los minigenes ensayados mostraba en la amplificación de RT-PCR un patrón de bandas alterado respecto al *wild type*, se procedió a su cuantificación y caracterización mediante dos técnicas diferentes: densitometría de bandas de geles de agarosa y área de picos visualizados en electroforesis capilar.

#### IV.4.a.2. Cuantificación de isoformas. Electroforesis capilar

La detección de isoformas mediante electroforesis capilar es una técnica altamente sensible que permite visualizar diferencias mínimas en el tamaño de los fragmentos analizados. Los amplicones analizados por esta metodología fueron obtenidos mediante reacciones RT-PCR en las que se ha utilizado al menos un cebador marcado con FAM (RT-pSPL3-FW o RT-pSPL3-RV de la Tabla 38). Las reacciones de amplificación se realizaron por triplicado a partir de cDNA obtenido de las células transfectadas y utilizando la polimerasa Go-Taq en un termociclador Veriti (Applied Biosystems) con el programa: 95°C-2'; 35 ciclos 95°C 30"-58°C 30"-72°C 1'30";72°C 10'.

La electroforesis capilar se realizó en una secuenciador automático ABI Prism® 3130 (Applied Biosystems) con capilares de 36 cm, utilizando POP-7 (Applied Biosystems) como polímero de electroforesis y tampón de carrera con EDTA (Applied Biosystems). Los amplicones marcados con FAM obtenidos en la RT-PCR se mezclaron con Hi-Di formamida (Applied Biosystems) y con el marcador de tamaño ROX-500 (Applied Biosystems) en el caso de fragmentos que no superaron en exceso las 500 pb o LIZ-1200 (Applied Biosystems) en el caso de fragmentos superiores. Antes de realizar la electroforesis capilar, la mezcla fue desnaturalizada 5 minutos a 95 °C y posteriormente incubada en hielo.

Los datos brutos de la carrera electroforética son procesados mediante el programa informático Data Collection 3.0 (Applied Biosystems), mientras que con el programa informático Peak Scanner™ Software v1.0 (Applied Biosystems) se han elaborado y analizado los cromatogramas. La cuantificación relativa de cada isoforma se obtuvo calculando del área de cada pico detectado y realizando la media aritmética de los tres valores obtenidos.

#### **IV.4.a.3. Caracterización de isoformas. Secuenciación capilar**

Las variaciones detectadas en el tamaño de las isoformas procedentes de los ensayos de mutaciones mediante minigenes híbridos respecto a los minigenes wt, fueron caracterizadas mediante secuenciación capilar. Las reacciones de secuencia se efectuaron con el kit BigDye Terminator Cycle Sequencing v3.1 (Applied Biosystems) con los cebadores sin marcaje utilizados para la reacción RT-PCR de cada mutación de acuerdo al apartado IV.4.a. siguiendo las indicaciones del fabricante. Los terminadores fluorescentes no incorporados en la reacción de secuencia fueron eliminados mediante perlas magnéticas CleanSEQ (Agencourt Bioscience) según las instrucciones del fabricante. La carrera electroforética capilar se llevó a cabo en un secuenciador automático ABI Prism® 3130 (Applied Biosystems) con capilares de 36 cm, utilizando POP-7 (Applied Biosystems) como polímero de electroforesis y tampón de carrera con EDTA (Applied Biosystems) utilizando las condiciones de la Tabla 14. Los datos brutos de la carrera electroforética fueron procesados mediante el programa informático Data Collection 3.0 (Applied Biosystems). El análisis y la elaboración de los cromatogramas se ha realizado con la aplicación Chromas Pro (Technelysium).

# RESULTADOS



## I. Resultados RT-PCR linfocitos de pacientes

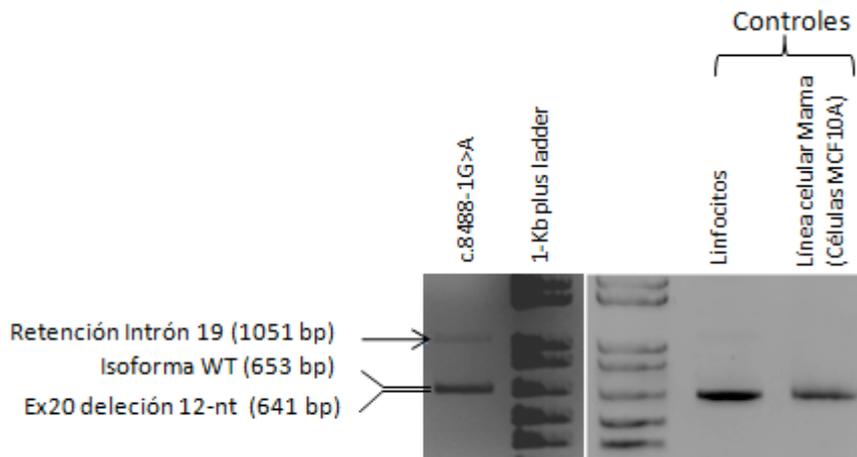
El Programa de Prevención del Cáncer de Castilla y León a través de las Unidades Consejo Genético tiene como objetivo informar de las posibles medidas preventivas y profilácticas a aquellas personas cuyo análisis genético sea positivo y por lo tanto porten una mutación en los genes *BRCA1* y *BRCA2* de predisposición a CMOH. Sin embargo, en determinados casos la falta de una relación directa entre el cambio genético detectado y su potencial efecto funcional sobre la proteína, provocan la aparición de resultados inciertos que impiden, por un lado, evaluar el riesgo asociado a CMOH y por otro, que la persona en concreto se beneficie de las medidas preventivas y profilácticas recomendadas.

En relación con los citados pretextos de incertidumbre, sobre el efecto funcional y el riesgo asociado de determinadas variantes de ADN en relación con el CMOH, fueron detectadas en el laboratorio de diagnóstico genético del Instituto de Biología y Genética Molecular, dos mutaciones situadas en regiones intrónicas, es decir, no codificantes del gen *BRCA2*. La primera, la mutación c.8488-1G>A se encuentra en el intrón 19 y no había sido descrita previamente. Por su parte, la mutación c.9501+9A>C fue detectada en el inicio del intrón 25 del gen *BRCA2*. Ambas variantes se sitúan próximas a las fronteras exón-intrón por lo que se decidió estudiar el efecto de las mismas en relación al proceso de *splicing* a través de ensayos funcionales RT-PCR del ARN extraído de cada paciente. También, se decidió analizar por esta metodología la mutación c.9026\_9030del en el exón 23 del gen *BRCA2* debido a que, aunque está catalogada como *frameshift*, en estudios previos realizados en nuestro laboratorio se habían descrito deleciones exónicas que afectaban al proceso de *splicing*, como por ejemplo la c.470\_474del (Sanz *et al.*, 2010).

### I.1. Resultado de la mutación c.8488-1G>A

Tras realizar la amplificación mediante RT-PCR entre los exones 19 y 23 y posterior electroforesis capilar, se pudo observar que el patrón de bandas resultante de la paciente portadora de la mutación c.8488-1G>A difería del obtenido en los controles (ARN de linfocitos y de células en cultivo no tumorales MCF10A). En la Figura 1 se puede distinguir claramente que existe una banda superior de mayor tamaño. El tamaño de dicha banda en relación con el marcador de peso molecular de 1kb

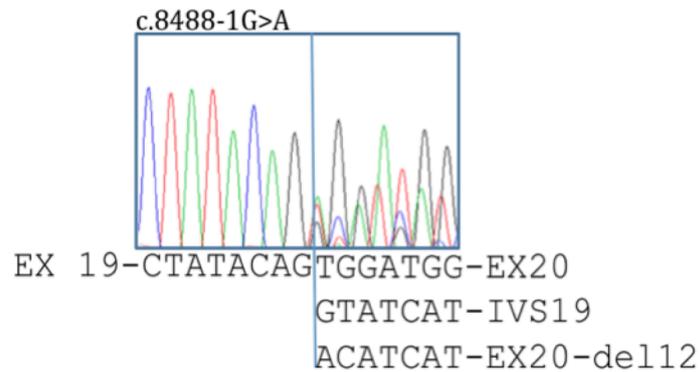
permitió deducir que el amplificado superaba un tamaño de 1000 pb, lo que podría implicar la retención del intrón 19 (398 pb).



**Figura 1. RT-PCR de linfocitos, mutación c.8488-1C>A.** Resultado del ensayo RT-PCR de linfocitos de la paciente que porta la mutación c.8488-1C>A y de dos controles. En la electroforesis en gel de agarosa se puede apreciar el patrón de bandas anómalo respecto al que presentan los controles.

Una nueva amplificación con un cebador localizado en el exón 19 y otro de hibridación específica en la frontera intrón 19 – exón 20 constató la presencia de la isoforma que contiene la retención del intrón 19 en el ADNc obtenido a partir de ARN del paciente y no así de los controles, donde no hubo amplificación (datos no mostrados).

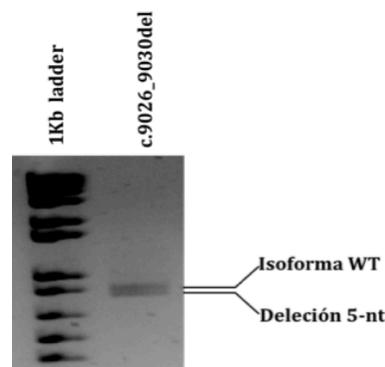
La verificación de la presencia de productos amplificados de la RT-PCR que incluyen el intrón 19 se obtuvo también mediante la secuenciación con un cebador solapante en el exón 19 en dirección *forward*. El cromatograma reveló que además de la presencia de la secuencia normal y la que incluye el intrón 19, existía un tercer transcrito en el cual no son incluidos los 12 primeros nucleótidos del exón 20 debido al uso de un aceptor críptico (Figura 2).



**Figura 2. Secuencia productos RT-PCR, mutación c.8488-1G>A.** Cromatograma de secuenciación del producto RT-PCR del paciente que porta la mutación c.8488-1G>A donde se puede apreciar la presencia de tres isoformas de *splicing* diferentes.

### I.2. Resultado mutación c.9026\_9030del

La amplificación del ADNc obtenido del ARN de la paciente portadora de la deleción c.9026\_9030del en el exón 23 del gen *BRCA2*, fue visualizada en gel de agarosa (Figura 3). El resultado propició la discriminación entre el alelo mutado (deleción de 5 nt) y el WT, no observándose ningún patrón de *splicing* anómalo.



**Figura 3. Resultado RT-PCR linfocitos, mutación c.9026\_9030del.** Resultado negativo del ensayo RT-PCR de linfocitos de la paciente que porta la mutación c.9026\_9030del.

### I.3. Resultado mutación c.9501+9A>C

No existió posibilidad alguna de obtener una muestra biológica de ARN de la paciente que portaba la mutación c.9501+9A>C. Existen factores concretos que contribuyen a que este hecho sea habitual, por ejemplo, la degradación del material altamente inestable como el ARN o el fallecimiento de la persona portadora de la variante genética. En tal caso, se decidió analizar dicha mutación a través de modelos informáticos y ensayos funcionales de *splicing* mediante minigenes híbridos, técnicas que carecen de la necesidad de muestras biológicas de pacientes.

## II. Análisis Bioinformático

Trescientas cuarenta y dos variantes de ADN fueron analizadas, entre ellas se incluyeron las tres mutaciones detectadas en el Programa de Prevención del Cáncer Hereditario: c.8488-1G>A, c.9026\_9030del, c.9501+9A>C. El resto de mutaciones analizadas proceden de las bases de datos internacionales BIC, LOVD y UMD, mientras que once fueron generadas de forma artificial.

Además, se realizó un estudio acerca de la conservación evolutiva de las secuencias detectadas de importancia para el *splicing*, mediante la alineación de la secuencia genómica de referencia para *BRCA2* en humano con las secuencias de otras especies como *Bos taurus*, *Canis familiaris*, *Felix catus*, *Gallus gallus*, *Mus musculus* y *Rattus norvegicus*. Con el propósito de analizar mediante procedimientos informáticos las consecuencias de la alteración de las secuencias conservadas, se decidió generar *in silico* nueve variantes no presentes en pacientes. En conjunto, la región analizada abarca mutaciones detectadas y generadas *de novo* entre el final del intrón 16 y el extremo 3' del gen *BRCA2* en el exón 27.

Las Tablas 1, 2 y 3 muestran ordenadas por exones las 96 variantes de ADN analizadas que fueron seleccionadas según cuatro criterios:

- Variantes que eliminan el sitio natural de *splicing*, o disminuyen drásticamente la probabilidad de que éste sea reconocido.
- Variantes que crean nuevos sitios críticos de *splicing* con una probabilidad alta de ser reconocidos por el espliceosoma.
- Variantes que eliminan secuencias potenciadores de *splicing* altamente conservadas en la evolución.
- Variantes que crean nuevos sitios silenciadores que podrían ser utilizados por factores represores de *splicing*.

**Tabla 1. Análisis bioinformático de variantes de ADN de los exones 19 y 20, junto con regiones intrónicas adyacentes que potencialmente afectan al *splicing*.** El símbolo [+] implica ganancia, mientras que [-] pérdida.

Mutación	Origen	Tipo de mutación	Resultado Bioinformático
<b>EXÓN 19</b>			
c.8362T>C	BIC	<i>Missense</i>	[-] SC35, SRp40
c.8368A>T	BIC	<i>Missense</i>	[-] SC35
<b>c.8378G&gt;A</b>	BIC	<i>Missense</i>	[-] ESE, [+] ESS (hnRNP-B)
c.8394_8396delTAGinsAA	BIC	<i>Frameshift</i>	[+] ESS
c.8398C>T	BIC	<i>Missense</i>	[-] SRp40, [+] ESS (hnRNP-I)
c.8435G>A	BIC	<i>Missense</i>	[-] SF2/ASF, [+] ESEs, [-] ESS
c.8438G>A	BIC	<i>Missense</i>	[+] 1 ESE, [+] 1 ESS
c.8484A>T	Artificial	<i>Sinónima</i>	Sitio Donador (exón 19): 0.95 → 0.95. / [-] SF2/ASF (ESEfinder)
<b>c.8486A&gt;C</b>	Artificial	<i>Missense</i>	Sitio Donador (exón 19): 0.95 → 0.92 / [-] motivo conservado de unión a SF2/ASF
<b>c.8486A&gt;T</b>	BIC	<i>Missense</i>	Sitio Donador (exón 19): 0.95 → 0.87 / [-] motivo conservado de unión a SF2/ASF
<b>c.8487G&gt;A</b>	BIC	<i>Sinónima</i>	Sitio Donador (exón 19): 0.95 → 0.40 / [-] motivo conservado de unión a SF2/ASF
<b>c.8487+1G&gt;A</b>	BIC	Intrónica	[-] Sitio Donador (exón 19)
<b>c.8487+3A&gt;G</b>	BIC	Intrónica	Sitio Donador (exón 19): 0.95 → 0.78. Disminución del 44% en la puntuación de MaxEnt
<b>c.8487+3A&gt;C</b>	Artificial	Sitio de <i>splicing</i>	Sitio Donador (exón 19): 0.95 → 0.61.
<b>c.8487+3A&gt;T</b>	Artificial	Sitio de <i>splicing</i>	Sitio Donador (exón 19): 0.95 → 0.77.
<b>DELECCIONES INTRÓN 19</b>			
c.8487+31_8487+42del	Artificial	Intrónica	Del 12-nt motivo conservado. Posible sitio de unión para hnRNP-I
c.8487+54_8487+355del o IVS19delAlu	Artificial	Intrónica	
c.8488-49_8488-44del	Artificial	Intrónica	Del. Motivo TGAATG: Elemento intrónico (IIE): silenciador.
<b>EXÓN 20</b>			
<b>c.8488-2A&gt;G</b>	BIC	Intrónica	[-] Sitio Aceptor (exón 20)
<b>c.8488-1G&gt;A</b>	Paciente		[-] Sitio Aceptor (exón 20)
c.8512T>A	Artificial	<i>Missense</i>	[-] motivo conservado de unión a SC35
c.8518del	Artificial	<i>Frameshift</i>	[-] motivo conservado de unión a SRp55
c.8539G>A	BIC	<i>Missense</i>	[-] motivos conservados de unión a SF/ASF y SRp40
c.8560del	BIC	<i>Frameshift</i>	Creación de dos silenciadores
c.8572C>A	BIC	<i>Missense</i>	[+] ESE, [+] ESS
<b>c.[8609A&gt;G;8611G&gt;T] Combinadas</b>	Artificial	-	[-] motivo conservado de unión a SRp40

**Tabla 2. Análisis bioinformático de variantes de ADN de los exones 23 y 24, junto con regiones intrónicas adyacentes que potencialmente afectan al *splicing*. El símbolo [+] implica ganancia, mientras que [-] pérdida.**

Mutación	Origen	Tipo de mutación	Resultado Bioinformático
<b>EXÓN 23</b>			
<b>c.8954-3C&gt;G</b>	BIC	Intrónica	[-] Sitio Aceptor (exón 23), [+] Sitio Aceptor críptico 2-nt <i>upstream</i>
<b>c.8954-1_8955delinsAA</b>	BIC	Intrónica	[-] Sitio Aceptor(exón 23)
c.8961_8964del	BIC	<i>Frameshift</i>	Sitio Aceptor (exón 23): 0.53→0.26 [-] SF2/ASF +SRp55, [+] ESS (PESX)
<b>c.8969G&gt;A</b>	BIC	<i>Nonsense</i>	[+] hnRNPA1 (HSF), [+] hnRNP-B (ESR)
c.8972G>A	BIC	<i>Missense</i>	[-] SC35, [-] SRp55
c.9004G>A	BIC	<i>Missense</i>	[-] SF2/ASF Parcialmente conservado y múltiples SREs
<b>c.9006A&gt;T</b>	BIC	<i>Missense</i>	[-] Motivo conservado de unión a SRp40
<b>c.[8972G&gt;A; 9006A&gt;T] combinadas</b>	Artificial	-	[-] SC35; [-] SRp40Conservado. ↑ Sitio Aceptor críptico
c.9026_9030del	BIC/ Paciente	<i>Frameshift</i>	[-] 2 Elemento exónico/ [+] 2 Elementos intrónicos
<b>c.9076C&gt;T</b>	BIC	<i>Nonsense</i>	Creación de silenciador (Rescue-ESE & PESX)
c.9100C>T	BIC	<i>Nonsense</i>	[-]SF2/ASF, [-] 1ESE / [+] 3 ESS, [+] hnRNPA1
c.9101A>G	BIC	<i>Missense</i>	[-] SF2/ASF (conservado en mamíferos)
c.9116C>T	BIC	<i>Missense</i>	Sitio Donador: 0.57→0.37
<b>c.9117G&gt;A</b>	BIC	Sinónima	[-] Sitio Donador (exón 23)
<b>c.9117+1G&gt;T</b>	BIC	Intrónica	[-] Sitio Donador (exón 23)
<b>c.9117+1G&gt;A</b>	BIC	Intrónica	[-] Sitio Donador (exón 23)
<b>EXÓN 24</b>			
<b>c.9118-2A&gt;G</b>	BIC	Intrónica	[-] Sitio Aceptor (exón 24). [+] Sitio Aceptor críptico 7-nt <i>downstream</i>
c.9148C>T	BIC	<i>Nonsense</i>	[+] 2 ESS
c.9172A>G	BIC	<i>Missense</i>	[-] SRp40
c.9177delA	BIC	<i>Frameshift</i>	[+] Sitio Aceptor interno 0,78
c.9196C>G	BIC	<i>Missense</i>	[+] ESS (hnRNP-B)
c.9213G>T	BIC	<i>Missense</i>	[+] ESEs parcialmente conservados, [+] ESS (hnRNP-H)
c.9218A>G	BIC	<i>Missense</i>	[-] ESE, [-] ESS; [+] ESS (hnRNP-B)
c.9227G>T	BIC	<i>Missense</i>	[+] 2 ESS
<b>c.9248_9256+7del</b>	BIC	Sitio de <i>splicing</i>	[-] Sitio Donador(exón 24)
c.9252_9255delAACainsTT	BIC	<i>Frameshift</i>	Sitio Donador (exón 24): 0,95→0,46
<b>c.9256G&gt;T</b>	BIC	<i>Nonsense</i>	[-] Sitio Donador (exón 24)
<b>c.9256+1G&gt;A</b>	BIC	Intrónica	[-] Sitio Donador (exón 24)

**Tabla 3. Análisis bioinformático de variantes de ADN de los exones 21, 22, 25, 26 y 27, junto con regiones intrónicas adyacentes que potencialmente afectan al *splicing*. El símbolo [+] implica ganancia, mientras que [-] pérdida.**

Mutación	Origen	Tipo de Mutación	Resultado Bioinformático
<b>EXÓN 21</b>			
<b>c.8633-24_8634del</b>	BIC	Intrónica	[-] Sitio Aceptor
c.8647C>T	BIC	<i>Missense</i>	[-] 3EES, [+] ESS
c.8723T>G	BIC	<i>Missense</i>	[+] SF2/ASF/ [+] ESS (hnRNP-F)
c.8732C>T	BIC	<i>Missense</i>	[-] SF2/ASF parcialmente conservado / [+] ESS (hnRNP-A1)
c.8748ins4	BIC	<i>Frameshift</i>	[+] SRp55/ [+] ESS
<b>c.8754G&gt;A</b>	BIC	Sinónima	[-] Sitio Donador
<b>c.8754+4A&gt;G</b>	BIC	Intrónica	[-] Sitio Donador (exón 21)
<b>c.8754+5G&gt;A</b>	BIC	Intrónica	Sitio Donador: 0,98 → 0,43 / [+]Sitio Donador críptico7-nt <i>downstream</i> 0,73
<b>c.8754+5G&gt;T</b>	BIC	Intrónica	Sitio Donador: 0,98 → 0,42 / [+]Sitio Donador críptico7-nt <i>downstream</i> 0,82
<b>EXÓN 22</b>			
<b>c.8755-1G&gt;A</b>	BIC	Intrónica	[-] Sitio Aceptor
c.8773C>T	BIC	<i>Nonsense</i>	[+] SC5 / [+] ESS ( +hnRNP-I(PTB))
c.8789delA	BIC	<i>Frameshift</i>	[-] SRp55 parcialmente conservado
c.8915T>G	BIC	<i>Missense</i>	[-] SRp55 / [+] ESS
<b>c.8948_8953+5del</b>	BIC	Sitio de <i>splicing</i>	[-] Sitio Donador
<b>c.8953+1G&gt;T</b>	BIC	Intrónica	[-] Sitio Donador
<b>EXÓN 25</b>			
c.9257-18C>A	BIC	Intrónica	[+] SRp40; [-] SRp40; [-] SRp55; [-] 2SF2/ASF, [-] 11ESE, [+] 27ESS
<b>c.9257-1G&gt;C</b>	BIC	Intrónica	[-] Sitio Aceptor / [-] SRp40, [+] SRp40, [+] SRp55, [+] 4 ESE, [-]2ESS
c.9275_9278del	BIC	<i>Frameshift</i>	Sitio Aceptor críptico 23-nt <i>downstream</i> 0,85 / [-]1ESE, [-] 4ESS, [+] 2ESS
c.9276T>G	BIC	<i>Nonsense</i>	Sitio Aceptor críptico 20-nt <i>downstream</i> 0,72 / [+] SRp40, [+] ESS
c.9301C>G	BIC	<i>Missense</i>	Sitio Aceptor críptico27-nt <i>downstream</i> 0,60 / [+] SC35, [+] 4ESS (hnRNP-A1)
c.9454G>A	BIC	<i>Missense</i>	[-] SRp40; [-] SF2/ASF, [-] 3ESE, [-] 1 ESS
c.9466delC	BIC	<i>Frameshift</i>	[-] SRp40, [-] 3 ESE, [-] y [+] 1 ESS
c.9501G>A	BIC	Sinónima	Sitio Donador: 0,99 → 0,60 / [-]SRp40;[+], SRp40;[+], SRp55, [+] 4 ESE, [-] 2 ESS
<b>c.9501+3A&gt;T</b>	BIC	Intrónica	Sitio Donador: 0,99 → 0,41
c.9501+4A>G	BIC	Intrónica	Sitio Donador: 0,99 → 0,78
c.9501+9A>C	<i>Patient</i>	Intrónica	[-] ESE, [+] 2ESS, [-] 1ESS
<b>EXÓN 26</b>			
<b>c.9502-12T&gt;G</b>	BIC	Intrónica	Sitio Aceptor: 0,93 → 0,74
<b>c.9502-2A&gt;C</b>	BIC	Intrónica	[-] Sitio Aceptor
c.9550C>G	LOVD	<i>Missense</i>	[+] 2ESS (hnRNP-A1)
c.9610A>G	BIC	<i>Missense</i>	[-] SRp40; [+] SF2/ASF
c.9611C>G	BIC	<i>Missense</i>	[-] SRp40
c.9613GC>CT	BIC	<i>Missense</i>	[-] SC35, [-] SRp40, [-] SRp55
c.9646C>T	BIC	Sinónima	[-] SF2/ASF, [-] SC35, [+] HnRNP-B
<b>EXÓN 27</b>			
c.9649-5A>G	LOVD	Intrónica	c. 9649 -5A>G
<b>c.9649-2A&gt;G</b>	BIC	Intrónica	[-] Sitio Aceptor
<b>c.9698 G&gt;T</b>	LOVD	<i>Missense</i>	[+] 2ESS

### III. Ensayos funcionales de *splicing* con minigenes híbridos

Los ensayos funcionales de *splicing* con minigenes híbridos nos permiten comprobar, sin necesidad de poseer una muestra biológica, el efecto que cualquier variante de ADN puede tener sobre el proceso de *splicing*.

En concreto, se ha analizado por este procedimiento tanto la mutación intrónica c.9501+9A>C de *BRCA2* detectada en el Programa de Prevención del Cáncer Hereditario de Castilla y León y de la que no existía muestra de ARN, como las 96 variantes de los exones 19, 20, 21, 22, 23, 24, 25, 26, y 27 e intrones flanqueantes de *BRCA2* que han resultado positivas según el análisis informático. Además, las variantes c.8488-1G>A y c.9026\_9030del, previamente ensayadas mediante RT-PCR de linfocitos, también fueron testadas en minigenes híbridos para comparar ambos procedimientos. Por otra parte, se procedió a realizar ensayos funcionales de *splicing* de 54 microdeleciones de doce exones de *BRCA2* (exones 16 a 27 ambos incluidos) con el objetivo de localizar elementos reguladores (*enhancers*) esenciales para el reconocimiento de sus respectivos exones. Con el mismo objetivo de mapear secuencias reguladoras, 11 de las 96 variantes seleccionadas son de origen artificial.

### IV. Ensayos funcionales de *splicing* con minigenes construidos en pSPL3

Dos minigenes se han construido sobre pSPL3. El primero, MGB2 19-20 incluye una región genómica que contiene los exones 19 y 20, el intrón 19 y parte de los intrones adyacentes 18 y 20 [IVS18 (247 pb) - **EX19 (156 pb)** - IVS19 (398 pb) - **EX20 (145 pb)** - IVS20 (207 pb)]. El segundo, MGBR2 23-24, incluye los exones 23 y 24 así como el intrón 23 y parte de los intrones adyacentes 22 y 24 [(IVS22 (223 pb) – **EX23 (164 pb)** – IVS23 (93 pb) - **EX24 (139 pb)**– IVS24 (98 pb)]. Ambas construcciones fueron transfectadas y los transcritos procedentes de la construcción fueron comprobados mediante RT-PCR con cebadores específicos de los exones del vector, que sólo amplifican los transcritos procedentes de los minigenes.

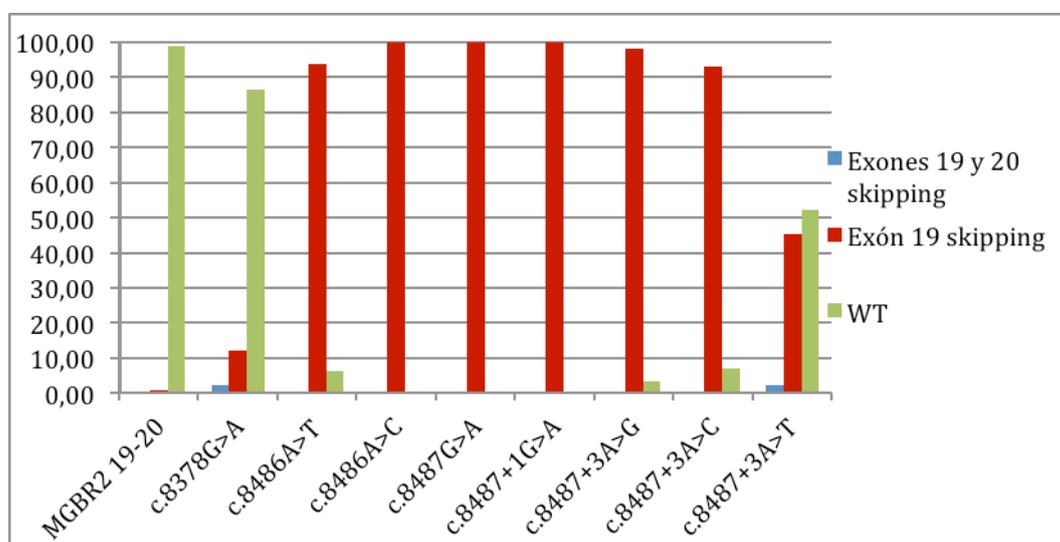
#### IV.1. Resultados del minigen MGBR2 19-20

Mediante mutagénesis dirigida se han generado 29 minigenes mutantes a partir del MGBR2 19-20 (Tabla 1). De ellos, 11 contienen mutaciones en el exón 19 (c.8362T>C, c.8368A>T, c.8378G>A, c.8394\_8396delTAGinsAA, c.8398C>T, c.8435G>A,

c.8438G>A, c.8484A>T, c.8486A>C, c.8486A>T, c.8487G>A) , 6 en el exón 20 (c.8512T>A, c.8518del, c.8539G>A, c.8560del, c.8572C>A, c.[8609A>G;8611G>T]). Además, se han generado 6 mutaciones en el intrón 19 (c.8487+1G>A, c.8487+3A>G, c.8487+3A>C, c.8487+3A>T, c.8488-2A>G, c.8488-1G>A) y 3 deleciones (c.8487+31\_8487+42del, c.8487+54\_8487+355, c.8488-49\_8488-44del) que posteriormente se combinaron (Tabla 1). Los resultados de la amplificación mediante RT-PCR del ARN procedente del proceso de transfección en células HeLa han sido comprobados mediante electroforesis en gel de agarosa, electroforesis capilar y secuenciación capilar. A continuación se muestran los resultados más relevantes.

#### IV.1.a. Exón 19

Un total de 8 (c.8378C>A, c.8486A>T, c.8486A>C, c.8487G>A, c.8487+1G>A, c.8487+3A>G, c.8487+3A>C y c.8487+3A>T) de las 15 mutaciones ensayadas en el exón 19 y sus regiones intrónicas flanqueantes más inmediatas alteraron el proceso de *splicing*. Siendo el evento más frecuente la pérdida del exón 19 en el ARN maduro o lo que es lo mismo el exón 19 *skipping*.



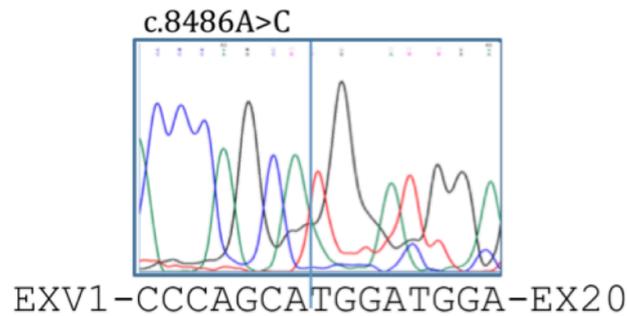
**Figura 4. Cuantificación de isoformas detectadas por mutaciones en el exón 19, MGBR2 19-20.** Diagrama de representación de la media aritmética de la cuantificación relativa de cada isoforma detectado en el cromatograma resultante de la electroforesis capilar para las mutaciones ensayadas en el exón 19 del MGBR2 19-20.

La mutación **c.8378G>A** es una variante registrada en la base de datos BIC y catalogada como *missense*. Según el estudio informático, la mutación elimina un *enhancer* o ESE que ayudaría a reconocer el exón 19 o crea un silenciador de *splicing* o

ESS de la ribonucleoproteína hnRNP-B que inhibiría la inclusión del exón 19 en el transcrito maduro (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 que porta la mutación c.8378G>A se ha podido comprobar la alteración en el proceso de *splicing* mediante electroforesis en gel de agarosa (Figura 6), electroforesis capilar (Figura 7) y secuenciación (datos no mostrados). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 4, revela que aunque la mayoría de los productos amplificados son similares al correspondiente WT, un 12% corresponden a la isoforma que carece del exón 19 y un 2% no han incluido ni el exón 19 ni el 20.

La mutación **c.8486A>T** es una variante registrada en la base de datos BIC y catalogada como *missense*. Según el estudio informático con NNSplice, la mutación disminuye la probabilidad relativa del sitio donador del exón 19 de 0,95 a 0,87. Además, eliminaba un putativo sitio conservado de unión a la proteína SF2/ASF que ayudaría a reconocer el exón 19 y, por lo tanto, que éste formara parte del transcrito maduro (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 que porta la mutación c.8486A>T se comprobó la alteración en el proceso de *splicing* tanto en gel de agarosa (Figura 6), como en electroforesis capilar (Figura 7) y secuenciación (datos no mostrados). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 4, mostró que el 94% transcritos no ha incluido el exón 19, mientras que el resto son iguales al transcrito canónico.

La mutación **c.8486A>C** es una variante artificial, no descrita en pacientes y generada *de novo*. Según el estudio informático (NNSplice) la mutación apenas disminuía la probabilidad relativa del sitio donador del exón 19 de 0,95 a 0,92. Además, eliminaba un putativo sitio conservado de unión a la proteína SF2/ASF que podría ayudar a reconocer el exón 19 (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 que porta la mutación c.8486A>C se comprobó la alteración en el proceso de *splicing* tanto en gel de agarosa (datos no mostrados), como en electroforesis capilar (Figura 7) y secuenciación (Figura 5). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 4, reveló que la práctica totalidad de los transcritos no contenían el exón 19.



**Figura 5. Secuencia del producto RT-PCR generado por la mutación c.8486A>C.** Cromatograma de secuenciación donde se observa que el exón 19 no es incluido en los transcritos maduros procedentes del MGBR2 19-20 que porta la mutación c.8486A>C. El mismo resultado se ha podido detectar en mayor o menor grado en el resto de las mutaciones positivas del exón 19.

La mutación **c.8487G>A** es una variante sinónima registrada en la base de datos BIC. Según NNsplice la mutación disminuía la probabilidad relativa del sitio donador del exón 19 de 0,95 a 0,40. Asimismo, podría eliminar un sitio conservado de unión a SF2/ASF que podría intervenir en el reconocimiento del exón 19 (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 que porta la mutación c.8487G>A se comprobó la alteración en el proceso de *splicing* en gel de agarosa (datos no mostrados), electroforesis capilar (Figura 7) y secuenciación (datos no mostrados). La cuantificación relativa de las isoformas detectadas (Figura 4), mostró que la práctica totalidad de los transcritos carecen del exón 19.

La mutación **c.8487+1G>A** es una variante registrada en la base de datos BIC y catalogada como intrónica. Según el estudio informático la variante elimina el sitio donador del exón 19 (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 que porta la mutación c.8487+1G>A se ha podido comprobar la alteración del *splicing* (Figuras 6 y 7). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 4, revela que, en la práctica totalidad de los transcritos, el exón 19 no ha sido incorporado.

La mutación **c.8487+3A>G** es una variante registrada en la base de datos BIC y catalogada como intrónica. Según NNsplice la mutación disminuye la probabilidad relativa de que el sitio donador del exón 19 sea reconocido de 0,95 a 0,78 (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 que porta la mutación c.8487+3A>G se ha podido comprobar la alteración en el proceso de *splicing* (Figuras 6 y 7).

La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 4 revela que aproximadamente el 98% de los transcritos no contienen el exón 19.

Las mutaciones **c.8487+3A>C** y **c.8487+3A>T** son variantes artificiales, no descritas en pacientes y creadas *de novo*. El objetivo de la generación de estas variantes de ADN es el estudio del efecto que pueden tener los tres tipos de sustitución nucleotídica en la posición adyacente al sitio consenso y conservado AG en el donador del exón 19.

Según el estudio informático, la mutación **c.8497+3A>C** disminuye la probabilidad relativa del sitio donador del exón 19 de 0,95 a 0,61 (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 que porta la mutación **c.8487+3A>C** se ha podido comprobar la alteración de los patrones de *splicing* (Figura 6). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 4, muestra la pérdida del exón 19 en aproximadamente el 93% de los transcritos.

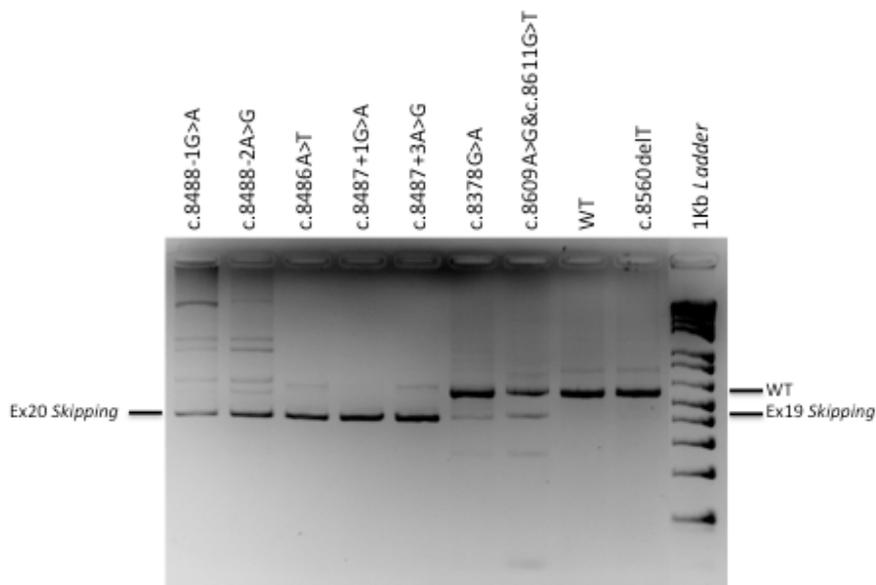
Por su parte, la variante **c.8497+3A>T**, de acuerdo a las predicciones de NNsplice disminuye el valor de reconocimiento del sitio donador del exón 19 de 0,95 a 0,77 (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 que porta la mutación **c.8487+3A>T** se ha podido comprobar la alteración en el proceso de *splicing* (Figura 7). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 4, revela que aproximadamente el 45% de los transcritos carecen del exón 19, un 52% son del mismo tamaño que el canónico y un 3% residual han perdido los exones 19 y 20 simultáneamente.

En la comparación de las tres posibles cambios de nucleótido en la posición **c.8497+3** se puede resumir que:

- El cambio **c.8497+3A>G** inhibe por completo el reconocimiento del exón 19; no así en el resto de los cambios.

- La mutación c.8497+3A>C es la que presenta un mayor porcentaje de transcritos canónicos.

El efecto sobre la predisposición a CMOH en variantes de *splicing* depende tanto de la proporción de mensajeros afectados como del cambio concreto que se produzca en el ARN y la proteína. En este caso, el riesgo relativo asociado a padecer CMOH en función del efecto en el *splicing* y relacionado con la traducción de las isoformas de ARNm maduro resultantes a proteína se espera que sea mayor para la mutación c.8497+3A>G y menor para la variante c.8497+3A>C.



**Figura 6. Resultados de la RT-PCR del minigen MGBR2 19-20 *wild type* y mutantes.** Electroforesis en gel de agarosa (1,5%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el minigen MGBR2 19-20 y variantes del mismo portadores de mutaciones generadas en los exones 19 y 20 del gen *BRCA2*.

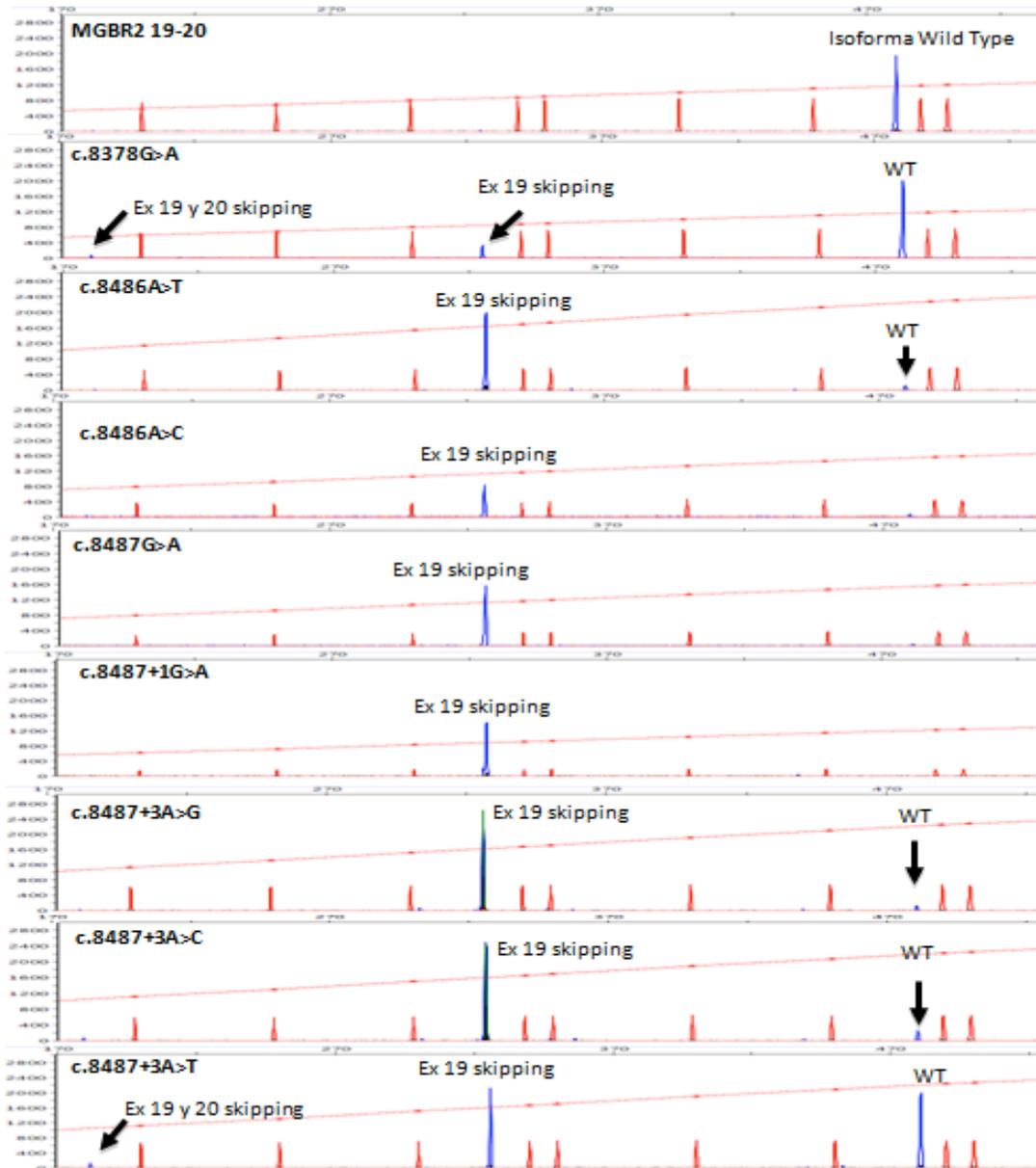
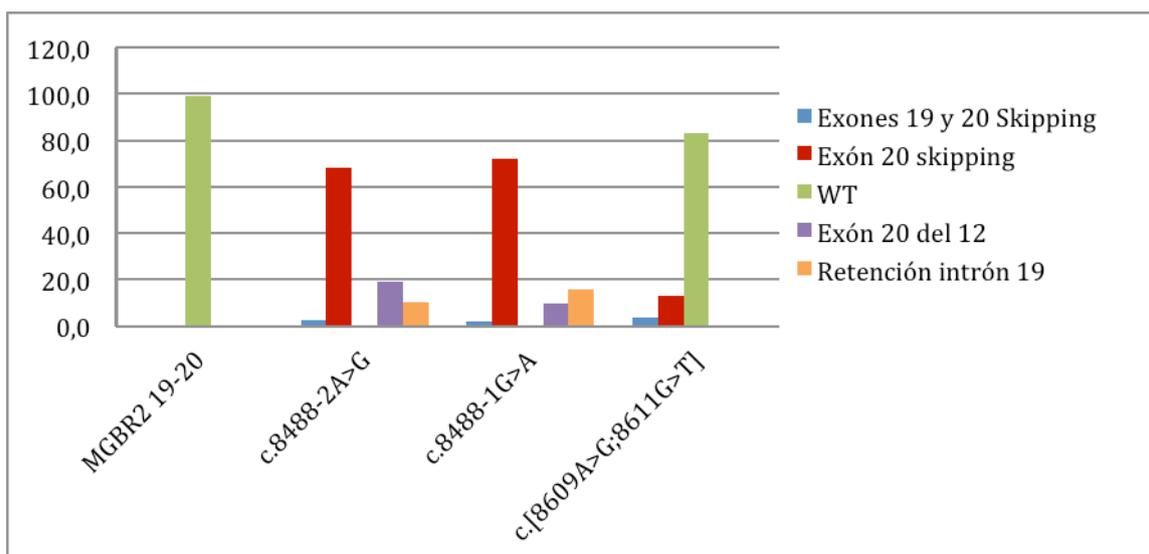


Figura 7. Electroforesis capilar de los ensayos funcionales de *splicing* de mutaciones del exón 19 de *BRCA2* (MGBR2 19-20) en células HeLa.

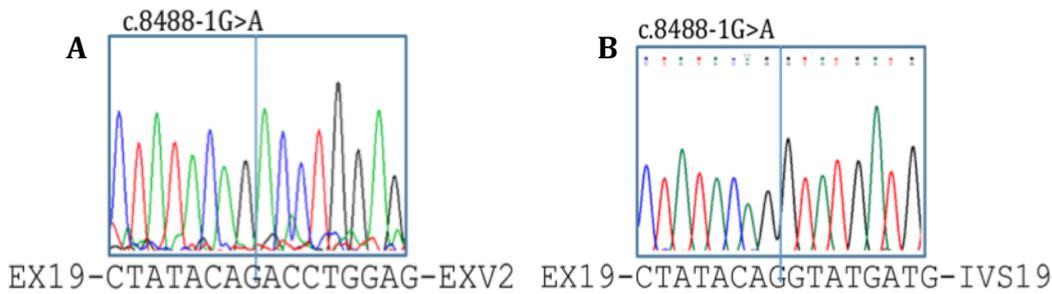
#### IV.1.b. Exón 20

Tres mutaciones (c.8488-2A>G, c-8488-1G>A y la doble mutagénesis [c.8609A>G;c.8611G>T]) de las 8 ensayadas en el exón 20 y regiones intrónicas flanqueantes inmediatas, alteraron el proceso de *splicing*, siendo el evento más frecuente la delección del exón 20 en el ARN maduro (*exón 20 skipping*).



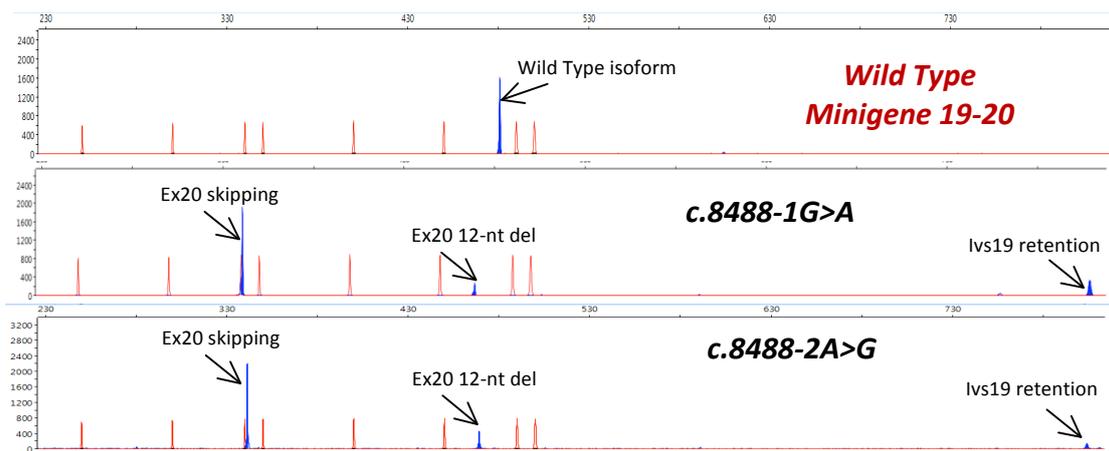
**Figura 8. Cuantificación de las isoformas producidas por mutaciones del exón 20 en el minigen MGBR2 19-20.** Representación de la media aritmética de 3 experimentos independientes de la cuantificación relativa de las isoformas detectadas en la electroforesis capilar de variantes del exón 20 del MGBR2 19-20.

La mutación **c.8488-1G>A** es una de las variantes detectadas en el laboratorio de diagnóstico de cáncer hereditario del IBGM y que previamente fue analizada mediante ensayo funcional RT-PCR de linfocitos. La variante fue detectada en una paciente que desarrolló cáncer de mama bilateral a los 66 y 74 años respectivamente y de ovario a los 77. La mutación se sitúa en la secuencia consenso para el sitio aceptor del exón 20, el cual está muy poco conservado, puesto que los programas bioinformáticos apenas lo reconocen. Además, la mutación debilitaría aún más este sitio aceptor (Tabla 1). En los resultados del ensayo funcional en el MGBR2 19-20 con la mutación c.8488-1G>A se comprobó la alteración en el proceso de *splicing* en electroforesis en gel de agarosa (Figura 6), electroforesis capilar (Figura 10) y secuenciación (Figura 9). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar, representada en la Figura 8, reveló que el 72% de los transcritos no incluyeron el exón 20, el 16% presentaban la retención del intrón 19 y un 10% carecían de los 12 primeros nucleótidos del exón 20 debido al uso de un sitio crítico exónico. Por último apareció un porcentaje residual de isoformas que no han incluido ni el exón 19 ni el 20.



**Figura 9. Secuencias de los productos RT-PCR procedentes de la mutación c.8488-1G>A.** Cromatogramas de secuenciación donde se observa la carencia del exón 20 (A) y la retención del intrón 19 (B) en los transcritos maduros procedentes del MGBR2 19-20 con la mutación c.8488-1G>A.

La mutación **c.8488-2A>G** es una variante registrada en la base de datos BIC y catalogada como intrónica. Al igual que la mutación c.8488-1 G>A, esta variante se sitúa en la secuencia consenso para el sitio aceptor débil del exón 20. La mutación debilitaría aún más este sitio aceptor (Tabla 1). En los resultados del ensayo funcional con el MGBR2 19-20 portador de c.8488-2A>G se pudo comprobar la alteración en el proceso de *splicing* en electroforesis en gel de agarosa (Figura 6), electroforesis capilar (Figura 10) y secuenciación (datos no mostrados). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar, representada en la Figura 8, reveló que el 68% de los transcritos no incluían el exón 20, el 10% presentan la retención del intrón 19 y un 19% carecían de los 12 primeros nucleótidos del exón 20 puesto que un sitio crítico es reconocido. Por último, aparecía un porcentaje residual de isoformas que no incluían los exones 19 y 20.



**Figura 10. Electroforesis capilar de los ensayos funcionales de splicing de las mutaciones c.8488-1G>A y c.8488-2A>G del exón 20 de BRCA2 (MGBR2 19-20) en células HeLa.**

Las variantes artificiales **c.8609A>G** y **c.8611G>T** se combinaron en un único minigen, ya que, según el estudio bioinformático, la conjunción de ambas eliminaría un

sitio altamente conservado de unión a la proteína SRp40 (Tabla 1) que podría facilitar el reconocimiento del exón 20 y su inclusión en el ARN maduro durante el proceso de *splicing*. En los resultados del ensayo funcional con el MGBR2 19-20 portador de las dos mutaciones c.8609A>G y c.8611G>T se pudo comprobar la alteración del *splicing* en electroforesis en gel de agarosa (Figura 8), electroforesis capilar (Figura 11) y secuenciación (datos no mostrados). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar, representada en la Figura 8, reveló que en un 13 % de los transcritos no está incluido el exón 20, mientras que el resto (83%), sería transcrito WT, a excepción de una isoforma residual sin exones 19 y 20.

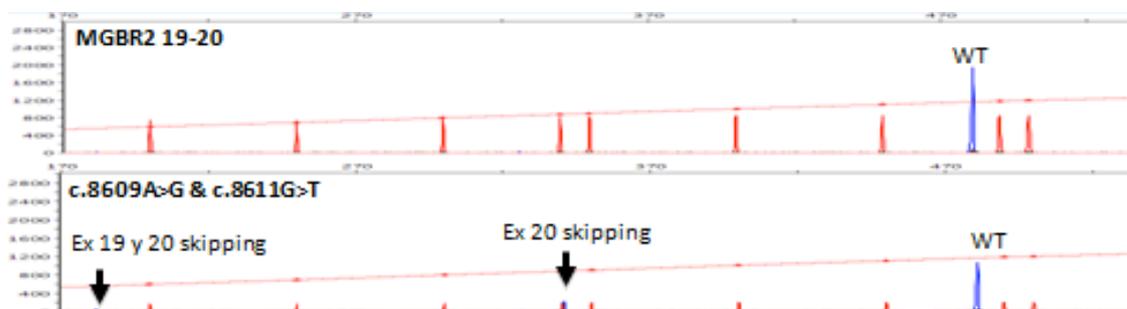


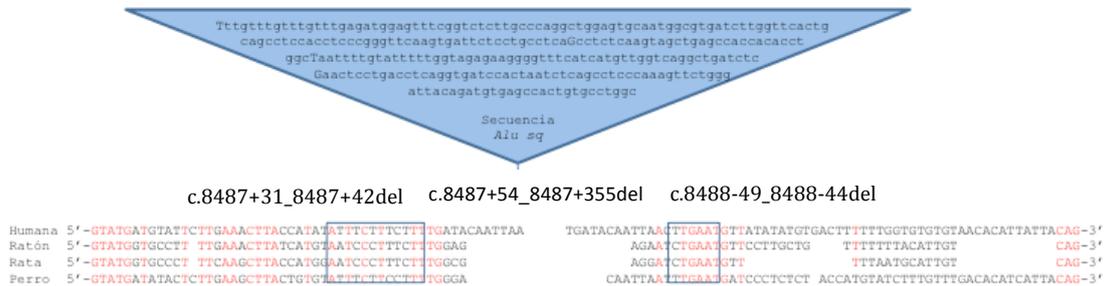
Figura 11. Electroforesis capilar de los ensayos funcionales de *splicing* de la doble mutación c.8609A>G + c.8611G>T del exón 20 de *BRCA2* (MGBR2 19-20) en células HeLa.

#### IV.1.c. Intrón 19

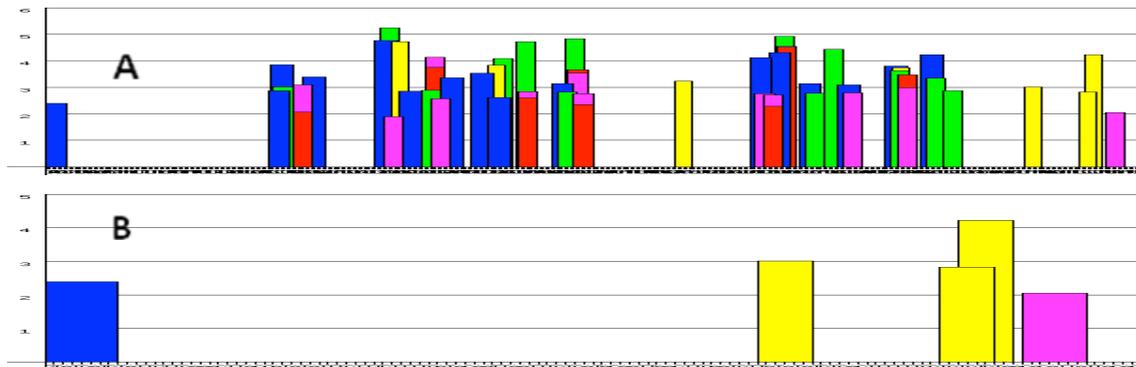
Como se ha comentado previamente, el análisis bioinformático en el estudio de las variantes c.8488-1G>A y c.8488-2A>G desveló la existencia de un sitio aceptor natural muy débil en el exón 20. Con el objetivo de detectar posibles elementos reguladores de *splicing* implicados en el reconocimiento del sitio aceptor del exón 20 se decidió estudiar en profundidad, mediante los medios informáticos disponibles, la secuencia del intrón 19.

El alineamiento de la secuencia consenso del intrón 19 humano respecto a otras especies reveló que el intrón humano era mucho más grande que en el resto de especies, en parte debido a la existencia de una secuencia Alu *Sq* insertada en orientación contraria, en mitad del intrón 19. Esta secuencia Alu contiene dos regiones ricas en putativos *enhancers*, que podrían estar facilitando el reconocimiento del sitio aceptor débil del exón 20. En el análisis de comparación evolutiva entre las secuencias

del intrón 19 y de diferentes especies, se puede constatar la existencia de dos secuencias muy conservadas, flanqueantes a la inserción Alu.



**Figura 12. Alineamiento de secuencias del intrón 19 de BRCA2 en diferentes especies de mamíferos.** Alineamiento y comparación del intrón 19 humano con las secuencias intrónicas de ratón, rata y perro. En rojo se han resaltado aquellos nucleótidos conservados, además aparecen recuadradas las secuencias que se han deletado en el MGBR2 19-20 mediante mutagénesis dirigida.

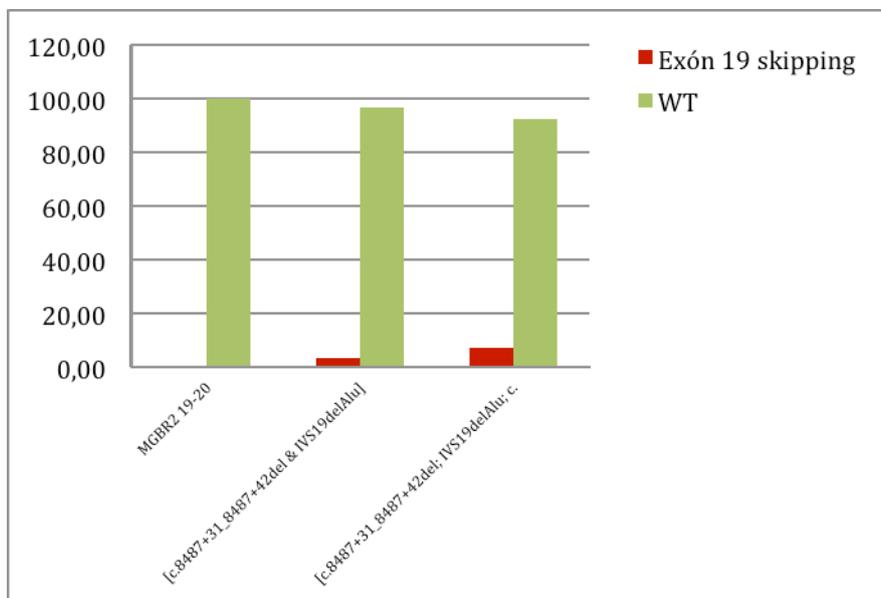


**Figura 13. Resultados del análisis bioinformático con ESEfinder del intrón 19 de BRCA2.** A- Posibles ISE de la secuencia completa del intrón 19 con una gran abundancia dentro de la secuencia Alu. B- Posibles ISE de la secuencia del intrón 19 sin la repetición Alu.

Para estudiar el efecto funcional tanto de las regiones conservadas como de los elementos reguladores detectados en la secuencia Alu (Figura 13) se procedió a realizar ensayos funcionales de *splicing* mediante tres minigenes híbridos derivados del MGBR2 19-20 en el que se deletaron de forma individual 12 pb (c.8487+31\_8487+42del) conservados en el extremo 5' del intrón 19 próximos al sitio donador del exón 19, la secuencia Alu (c.8487+54\_8487+355del o IVS19delAlu) y 6 pb (c.8488-49\_8488-44del) altamente conservados en el extremo 3' del intrón próximos al sitio aceptor del exón 20 (Figura 12). Los resultados de todos los ensayos funcionales fueron negativos.

Sin embargo, como los ESEs rara vez actúan de forma individual, se realizaron diferentes combinaciones de las tres deletaciones. El ensayo funcional con dos de los

minigenes derivados del MGBR2 19-20 sobre los que se generó la combinación de deleciones c.8487+31\_8487+42del-c.8487+54\_8487+355del y c.8487+31\_8487+42del-c.8487+54\_8487+355del-c.8488-49\_8488-44del afectaron al *splicing*. Dicha alteración fue comprobada en electroforesis en gel de agarosa (datos no mostrados), electroforesis capilar (Figura 15) y secuenciación (datos no mostrados). La cuantificación relativa de las isoformas detectadas mediante PCR semicuantitativa y electroforesis capilar representada en la Figura 14, reveló que, sorprendentemente en el 3% de los transcritos el exón 19 no ha sido incluido para el minigen con la doble deleción c.8487+31\_8487+42del-c.8487+54\_8487+355del, mientras que este porcentaje aumenta hasta el 7% en el minigen que contiene la triple deleción c.8487+31\_8487+42del-c.8487+54\_8487+355del-c.8488-49\_8488-44del. Estos datos sugieren que diferentes elementos *enhancer* del intrón 19 facilitarían el reconocimiento del exón 19. Probablemente, elementos *enhancer* del exón 20 colaborarían en el reconocimiento del sitio aceptor débil del exón 20, aunque no se pudieron identificar.



**Figura 14.** Cuantificación de las isoformas de *splicing* producidas por deleciones del intrón 19 en el minigen MGBR2 19-20. Representación de la media aritmética de 3 experimentos independientes de la cuantificación relativa de cada isoforma detectada en la electroforesis capilar para las diferentes combinaciones de deleciones ensayadas en el intrón 19 del MGBR2 19-20.

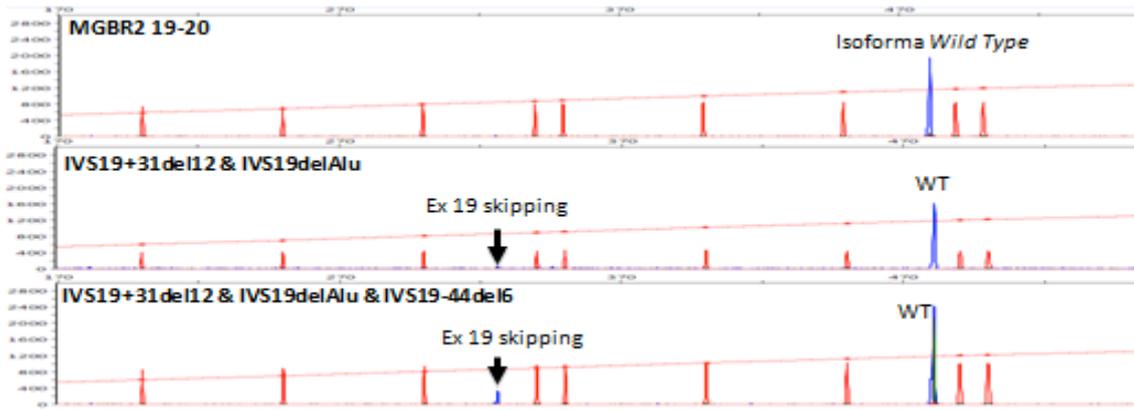
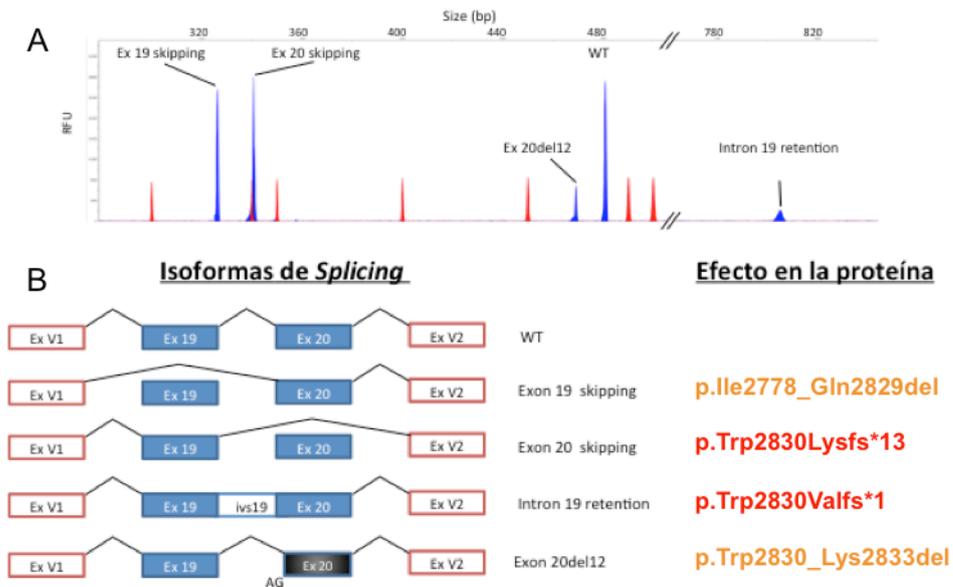


Figura 15. Electroforesis capilar de los ensayos funcionales de *splicing* de dos deleciones del intrón 19 en células HeLa.

#### IV.1.d. Resultados globales de los exones 19 y 20

En conjunto, cinco tipos de transcritos diferentes, producto de un *splicing* anómalo, fueron identificados en el análisis funcional de mutaciones mediante minigenes híbridos derivados del MGBR2 19-20. Salvo por los transcritos que no contenían los exones 19 y 20, que podrían ser atribuidos al reducido contexto genómico del minigen, el resto de isoformas detectadas es muy probable que se manifiesten también *in vivo*. Todas ellas, tienen un efecto claro sobre la traducción del ARN maduro a proteína. Las isoformas que no contienen el exón 19 y aquellas en las que no son incluidos los 12 primeros nucleótidos del exón 20 mantendrían la fase de lectura original y por lo tanto las proteínas resultantes de la traducción carecen de 52 y 4 aminoácidos, respectivamente. Por otro lado, los transcritos que no incluyen el exón 20 y las que han retenido el intrón 19 interrumpen la fase de lectura apareciendo un codón prematuro de parada de la traducción, por lo que se producirían proteínas truncadas que sean o bien inactivas o bien parcialmente funcionales.



**Figura 16. Isoformas de *splicing* detectadas en el minigen MGBR2 19-20 y efecto predecible en la proteína de acuerdo a la nomenclatura HGVS. A) cromatogramas superpuestos de varios ensayos donde se observan las diferentes isoformas detectadas en. B) Diagramas de las reacciones de *splicing* que dan lugar a las diferentes isoformas y el efecto de las mismas en la traducción a proteína.**

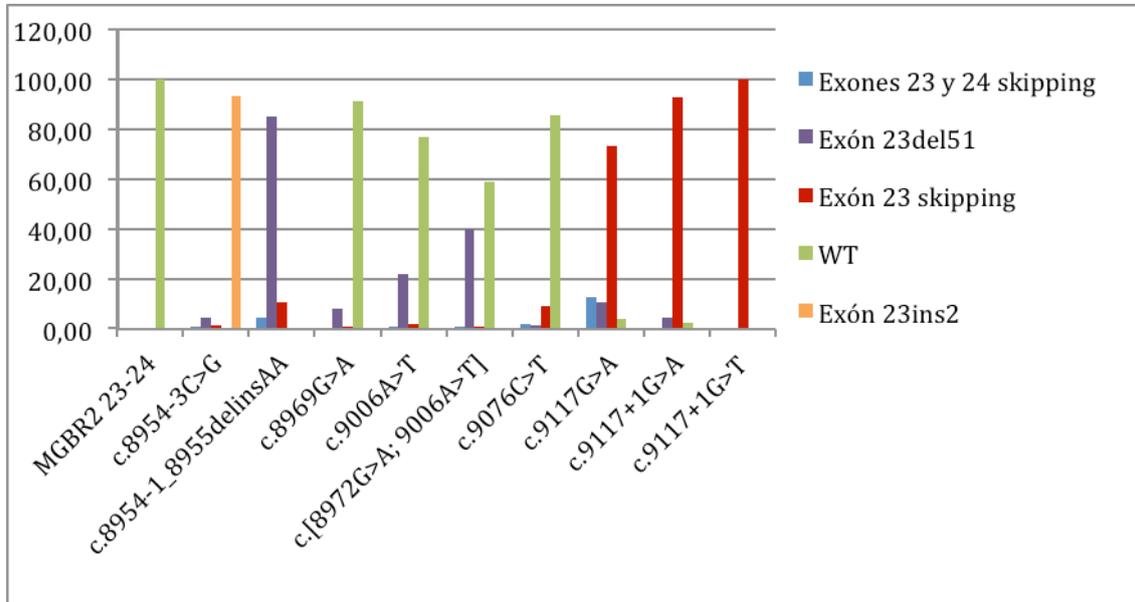
## IV.2. Resultados del minigen MGBR2 23-24

Mediante mutagénesis dirigida se han generado 28 minigenes a partir del minigen WT MGBR2 23-24 (Tabla 2). De ellos, dos contienen mutaciones en el intrón 22 (c.8954-3C>G, c.8954-1\_8955delinsAA), doce en el exón 23 (c.8961\_8964del, c.8969G>A, c.8972G>A, c.9004G>A, c.9006A>T, c.[8972G>A; 9006A>T], c.9026\_9030del, c.9076C>T, c.9100C>T, c.9101A>G, c.9116C>T, c.9117G>A), tres en el intrón 23 (c.9117+1G>T, c.9117+1G>A, c.9118-2A>G), once en el exón 24 (c.9148C>T, c.9172A>G, c.9177delA, c.9196C>G, c.9213G>T, c.9218A>G, c.9227G>T, c.9248\_9256+7del, c.9252\_9255delAACainsTT, c.9256G>T) y uno en el intrón 24 (c.9256+1G>A). Los resultados de la amplificación mediante RT-PCR del ARN procedente del proceso de transfección en células HeLa han sido comprobados mediante electroforesis en gel de agarosa, electroforesis capilar y secuenciación capilar. A continuación se muestran los resultados de mutaciones con efecto en el *splicing*.

### IV.2.a. Exón 23

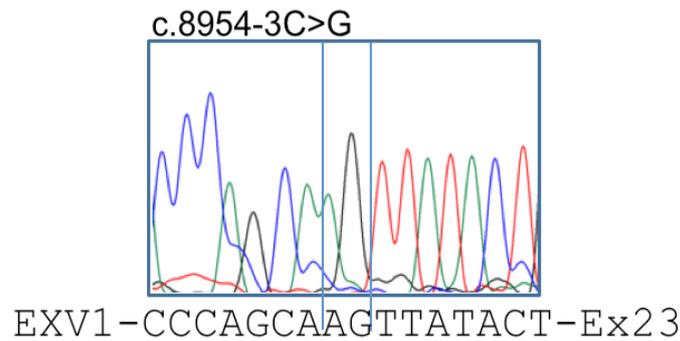
Un total de nueve mutaciones (c.8954-3C>G, c.8954-1\_8955delinsAA, c.8969G>A, c.9006A>T, [c.8972G>A;c.9006A>T], c.9076C>T, c.9117G>A, c.9117+1G>T,

c.9117+1G>A) de las dieciséis ensayadas en el exón 23 y en sus secuencias intrónicas flanqueantes alteraron el proceso de *splicing*. Los eventos más frecuentes son la delección del exón 23 en el transcrito maduro y la pérdida de los primeros 51 nt del exón 23 por el uso de un sitio críptico aceptor.



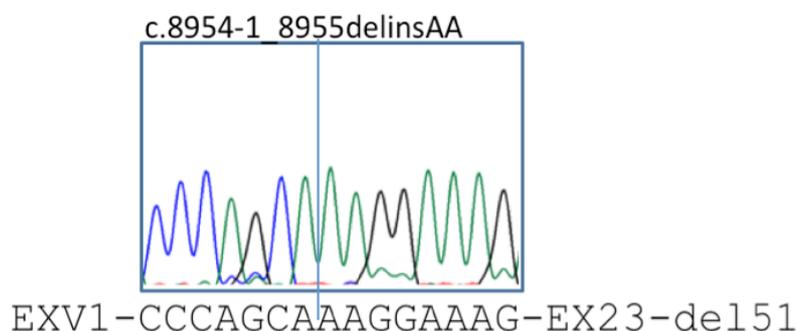
**Figura 17. Cuantificación de las isoformas detectadas producidas por mutaciones del exón 23 en el minigen MGBR2 23-24.** Representación de la media aritmética de 3 experimentos independientes de la cuantificación relativa de cada isoforma detectada en la electroforesis capilar de las mutaciones ensayadas del exón 23 de *BRCA2*.

La mutación **c.8954-3C>G** es una variante registrada en la base de datos BIC y catalogada como intrónica. Según el estudio informático, la mutación eliminaría el sitio aceptor del exón 23, pero crea un nuevo sitio críptico 2 nt *upstream* del natural (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 que porta c.8954-3C>G se pudo comprobar la existencia de una alteración del patrón (Figuras 18 y 22). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar, representada en la Figura 17, reveló que el 93% de los transcritos habían incorporado los dos últimos nucleótidos del intrón 22 debido al uso por parte de la maquinaria de *splicing* del nuevo sitio aceptor generado por la mutación. El resto de isoformas detectadas aunque minoritarias también afectaron al *splicing*. Un 4% de los transcritos no incorporaron los primeros 51 nt del exón 23, ya que el espliceosoma ha reconocido un sitio críptico 51 nt *downstream* del natural, un 2% perdió el exón 23 y un minoritario 1% no incluía los exones 23 y 24.



**Figura 18. Secuencia de los productos RT-PCR generados por la mutación c.8954-3C>G.** Cromatograma de secuenciación donde se observa la inserción de 2 nt del final del intrón 22 en transcritos maduros procedentes del MGBR2 23-24 con la mutación c.8954-3C>G.

La mutación **c.8954-1\_8955delinsAA** es una variante registrada en la base de datos BIC y catalogada como intrónica. Según el estudio informático la mutación elimina el sitio aceptor del exón 23 (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 que porta esta mutación se pudo comprobar la alteración del proceso de *splicing* (Figuras 19 y 22). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar, representada en la Figura 17, mostró que el 85% de los transcritos no habían incorporado los primeros 51 nt del exón 23 por el uso de un sitio críptico 51 nt *downstream* del natural, un 10% no incluían el exón 23, mientras que el resto no contenía los exones 23 y 24.



**Figura 19. Secuencia de los productos RT-PCR generados por la mutación c.8954-1\_8955delinsAA.** Cromatograma de secuenciación donde se observa la deleción de los primeros 51 nt del exón 23 en transcritos maduros procedentes del MGBR2 23-24 que porta la mutación c.8954-1\_8955delinsAA.

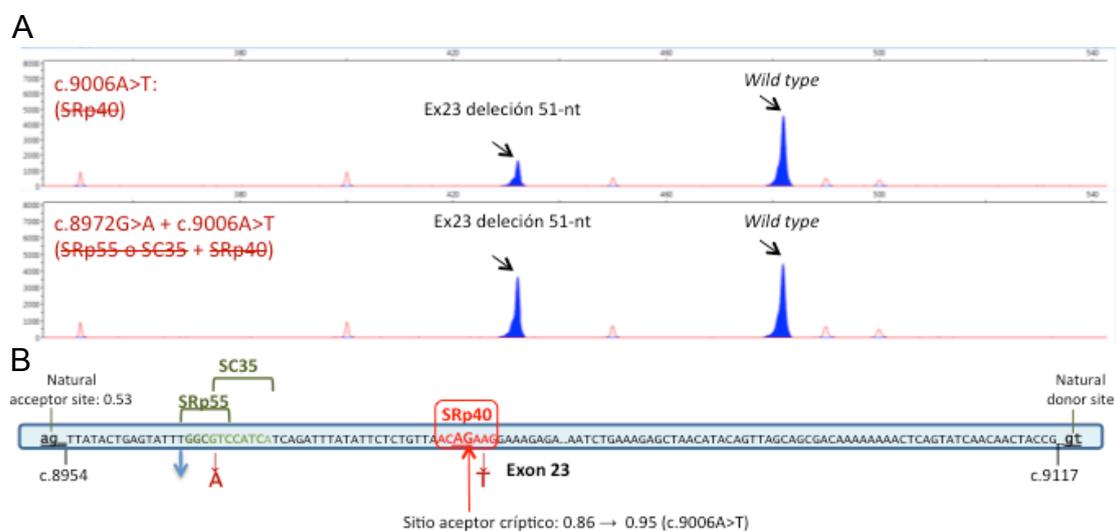
La mutación **c.8996G>A** es una variante registrada en la base de datos BIC y catalogada como *nonsense*. Según el estudio informático, la mutación crea un sitio de unión que reconocería una ribonucleoproteína del tipo hnRNP A1 o B que por su cercanía al sitio aceptor del exón 23 podría impedir su reconocimiento (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 que porta la mutación c.8996G>A

se ha podido comprobar la alteración en el proceso de *splicing* (Figura 22). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 17, reveló que una amplia mayoría (91%) de los transcritos eran canónicos, sin embargo un 8% de ellos no habían incorporado los primeros 51 nt del exón 23, ya que el espliceosoma había reconocido un sitio críptico 51 nt *downstream* del natural, y una minoría de transcritos (<1%) no contenía los exones 23 y 24 (skipping total de los exones clonados).

La mutación **c.9006A>T** es una variante registrada en la base de datos BIC y catalogada como *missense*. Según el estudio informático, la mutación elimina un *enhancer* altamente conservado de la proteína SRp40, pero también fortalece el sitio críptico aceptor situado (NNsplice: 0,86 a 0,95) a 51 nt *downstream* del aceptor natural del exón 23 (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 que porta la mutación c.9006A>T se pudo comprobar la alteración de *splicing* (Figura 22). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 17, reveló que un 76% de los transcritos era del mismo tamaño que el WT. Sin embargo, un 22% no había incorporado los primeros 51 nt del exón 23, puesto que el espliceosoma había reconocido el sitio críptico aceptor 51 nt *downstream* del natural y el resto no contenía los exones 23 y 24.

La mutación **c.8972G>A** (situada a tan sólo 3 nt de la mutación c.8969G>A que ha resultado positiva) según el estudio informático, eliminaba un motivo de reconocimiento de la proteína SC35 y otro de la proteína SRp55, pero el resultado del ensayo funcional de *splicing* con el minigen mutante fue negativo (Tabla 2). No obstante, una doble mutagénesis con la conjunción de las variantes **c.8972G>A** y **c.9006A>T** fue realizada sobre el mismo MGBR2 23-24. En este caso, se estimó un fortalecimiento del sitio aceptor críptico a 51 nt del aceptor natural. También que este sitio críptico podría estar bloqueado gracias a la unión de una proteína SRp40 que reconocería una secuencia de unión altamente conservada. Este bloqueo, que impediría de forma esteérica el reconocimiento del sitio críptico, podría ser eliminado debido a que la variante c.9006A>T eliminaría la secuencia de reconocimiento de la proteína SRp40. Además, la variante c.8972G>A estaría eliminando una secuencia ESE

que facilitaría el reconocimiento del sitio aceptor natural del exón 23 frente al sitio crítico (Figura 20). En el ensayo con el minigen que contenía ambas variantes se pudo observar una alteración en el proceso de *splicing* que duplicaba el porcentaje de las isoformas que no han incorporado los primeros 51 nt del exón 23 (40% del total; Figura 17) respecto a la mutación c.9006A>T sola. Por lo tanto, podríamos deducir que el sitio aceptor crítico no es reconocido de forma natural ya que probablemente esté bloqueado por una proteína SR y que otras proteínas *enhancer* cercanas al sitio aceptor natural sean necesarias para facilitar su reconocimiento respecto al crítico.

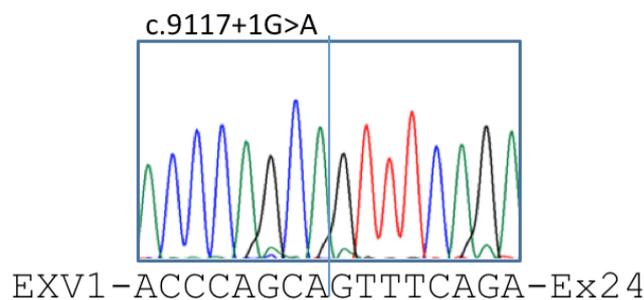


**Figura 20. Efecto conjunto de las mutaciones c.8972G>A y c.9006A>T en el exón 23.** A) Cromatogramas procedentes de transcritos amplificados de los ensayos funcionales con los MGBR2 19-20 portadores de la mutación c.9006A>T y la combinación de las mutaciones [c.8972G>A;c.9006A>T]. En la imagen se puede apreciar un incremento de la isoforma con la deleción de 51 nt en los transcritos procedentes del minigen con la doble mutagénesis. B) Elementos potenciadores de *splicing* o *enhancers* eliminados por las mutaciones c.8972G>A y c.9006A>T. Posición de los sitios naturales y críticos del exón 23 junto con la probabilidad relativa de que sean detectados por el espliceosoma.

La mutación **c.9076C>T** es una variante registrada en la base de datos BIC y catalogada como *nonsense*. Según el estudio informático la mutación elimina un silenciador (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 portador de c.9076C>T se pudo comprobar la alteración del patrón de *splicing* (Figura 22). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 17, mostró que el transcrito canónico era el más representado (88% de las isoformas), un 9% no contenía el exón 23, un 2% había perdido los exones 23 y el 24 y un 1% no había incorporado los primeros 51 nt del exón 23.

La mutación **c.9117G>A** es una variante registrada en la base de datos BIC y catalogada como sinónima. Está situada en el último nucleótido del exón 23. Según el estudio informático la mutación eliminaba el sitio donador del exón 23 (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 portador de c.9117G>A se pudo comprobar la alteración en el proceso de *splicing* (Figura 22). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 17, reveló que el 73% de las isoformas no contiene el exón 23, un 13% no había incorporado ni el exón 23 y ni el 24, un 10% no había incluido los primeros 51 nt del exón 23. Por último, el transcrito canónico constituye un 4% de los totales detectados.

Las mutaciones **c.9117+1G>T** y **c.9117+1G>A** son variantes registradas en la base de datos BIC y catalogadas intrónicas. Según el estudio informático ambas mutaciones eliminaban el sitio donador del exón 23 (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 portador de c.9117G>A se pudo comprobar que las dos variantes alteraban el proceso de *splicing* (Figuras 21 y 22). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 17, mostró que la practica totalidad de las isoformas generadas por el MGBR2 23-24 con la variante c.9117+1G>T no había incluido el exón 23. Sin embargo, la variante c.9117+1G>A generaba un porcentaje pequeño de isoformas WT (2%) y un 4% con la deleción de los primeros 51 nt del exón 2. El resto presentaban *skipping* del exón 23.



**Figura 21. Secuencia de los productos RT-PCR originados por la mutación c.9117+1G>A.** Cromatograma de secuenciación donde se observa la pérdida del exón 23 en transcritos maduros procedentes del MGBR2 23-24 que portaba la mutación c.9117+1G>A.

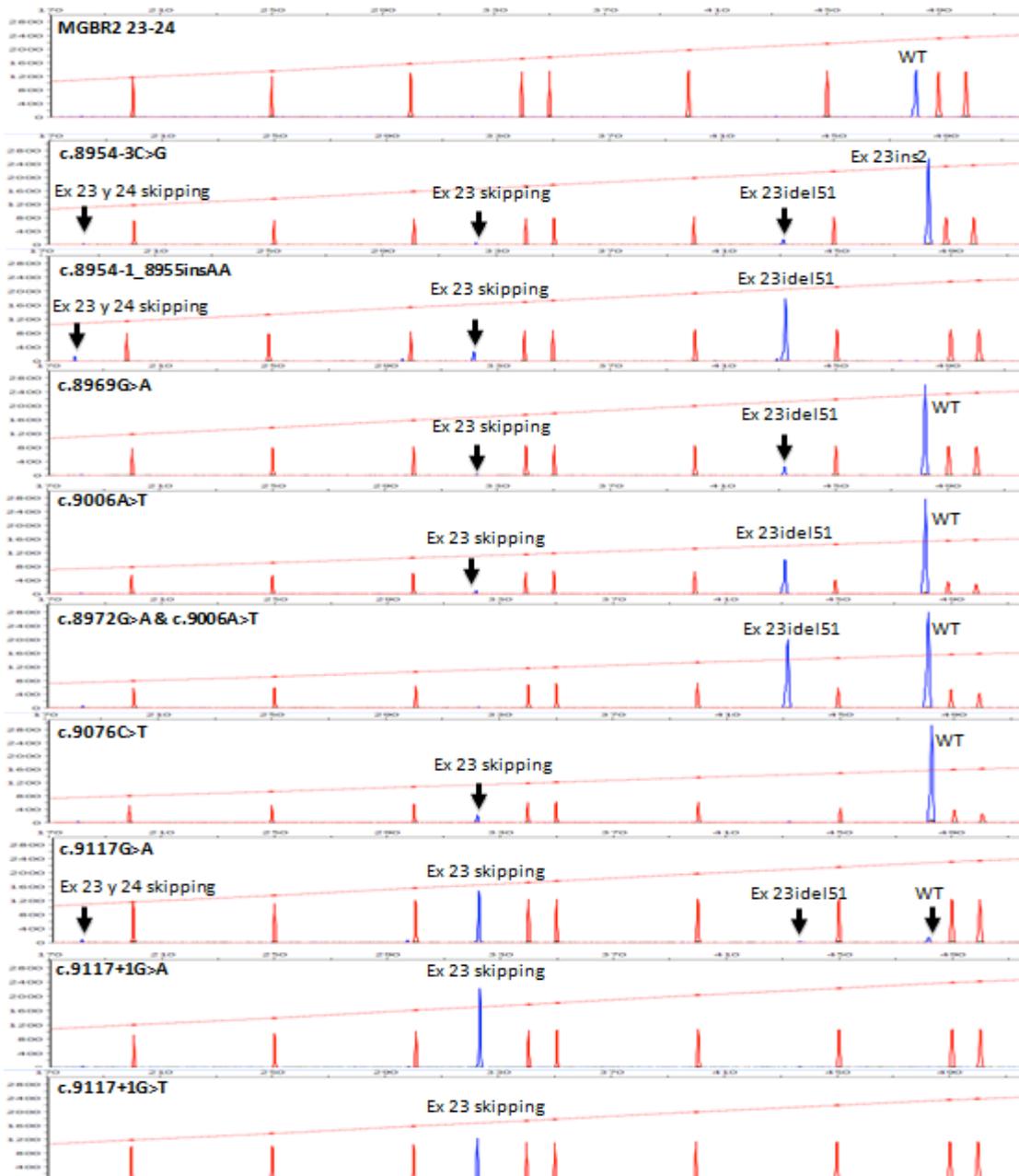
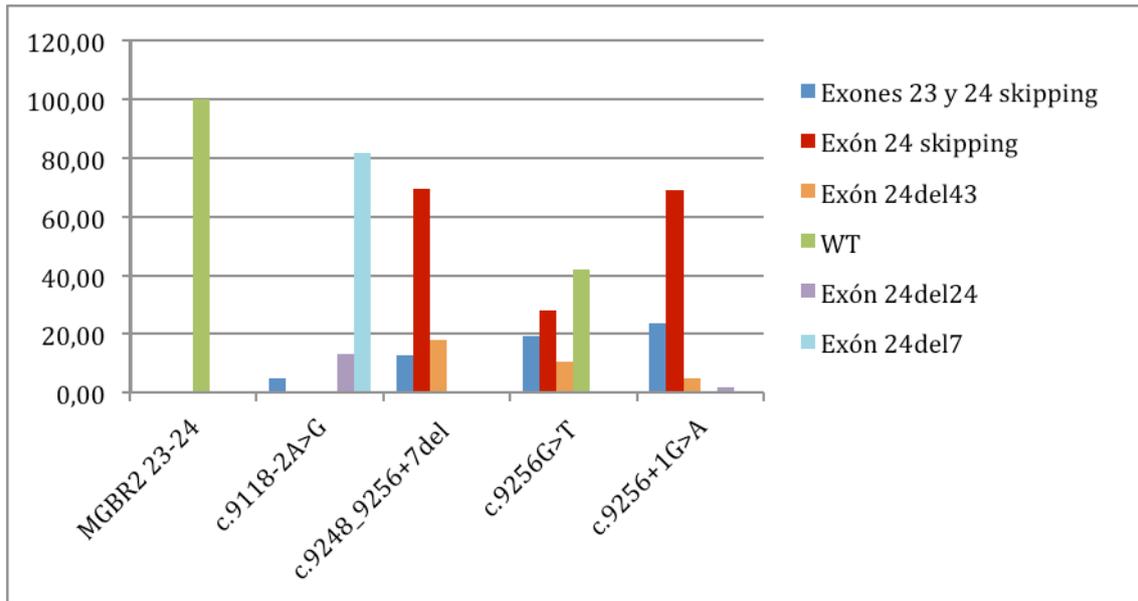


Figura 22. Electroforesis capilar de los ensayos funcionales de *splicing* de las mutaciones del exón 23 de *BRCA2* (MGBR2 23-24) en células HeLa.

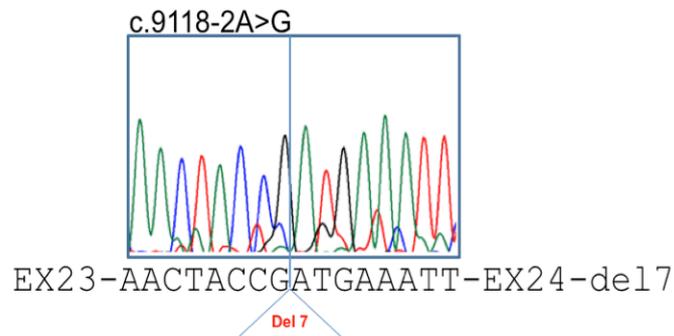
#### IV.2.b. Exón 24

De las doce variantes de ADN ensayadas en el exón 24 y en sus secuencias intrónicas flanqueantes, un total de cuatro alteraron el proceso de *splicing* (c.9118-2A>G, c.9248\_9256+7del, c.9256G>T, c.9256+1G>A). Los principales eventos de *splicing*, por su abundancia en las mutaciones ensayadas, han sido aquellos que no incluyen el exón 24 (*skipping*) en el transcrito maduro y los que no contienen los 7 primeros nucleótidos del exón 24.



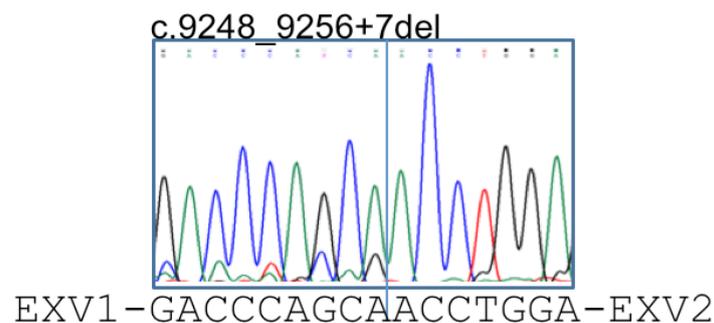
**Figura 23. Cuantificación isoformas mutaciones del exón 24, MGBR2 23-24.** Diagrama de representación de la media aritmética de la cuantificación relativa de cada isoforma o pico detectado en el cromatograma resultante de la electroforesis capilar de las mutaciones ensayadas en el exón 24 del MGBR2 23-24 y que han resultado alterar el *splicing*.

La mutación **c.9118-2A>G** es una variante registrada en la base de datos BIC y catalogada como intrónica. Según el estudio informático, la mutación eliminaba el sitio aceptor del exón 24 y crea un sitio críptico *de novo* 7 nt *downstream* del natural con un valor de NNsplice de 0,70 de ser reconocido (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 portador de c.9118-2A>G se pudo comprobar la alteración de los patrones de *splicing* (Figuras 24 y 27). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 23, reveló que el 82% de las isoformas había perdido los 7 primeros nucleótidos del exón 24, al ser reconocido el nuevo sitio críptico generado por la mutación, un 13% presentaba una delección de los primeros 24 nt del exón 24 debido al uso de un sitio aceptor críptico (NNsplice=0,59). Por último, se detectó que el 5% de las isoformas restantes había perdido los exones 23 y 24 (doble *skipping*).



**Figura 24. Secuencia de los productos RT-PCR originados por la mutación c.9118-2A>G.** Cromatograma de secuenciación donde se observa la delección de los primeros 7 nt del exón 24 en los transcritos maduros procedentes del MGBR2 23-24 que porta la mutación c.9118-2A>G.

La mutación **c.9248\_9256+7del** es una variante registrada en la base de datos BIC y catalogada como de sitio de *splicing*. Según el estudio informático, la mutación elimina el sitio donador del exón 24 (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 portador de c.9248\_9256+7del se pudo comprobar la alteración en el proceso de *splicing* (Figuras 25 y 27). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 23 mostró que el 69% de las isoformas había perdido el exón 24, un 18% no contenía los últimos 43 nt del exón 24 puesto que la maquinaria de *splicing* reconocía un sitio donador críptico situado a 43 nt *upstream* no detectado por NNSplice. Además se había detectado que el 13% de las isoformas presentaba *skipping* doble de los exones 23 y 24.

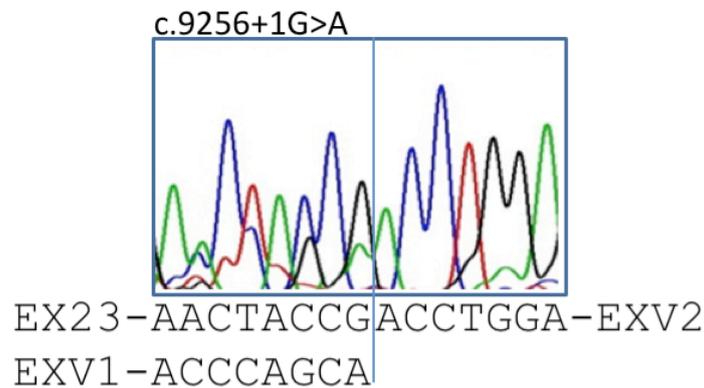


**Figura 25. Secuencia de los productos RT-PCR originados por la mutación c.9248\_9256+7del.** Cromatograma de secuenciación donde se observa el doble *skipping* de los exones 23 y 24 en transcritos maduros procedentes del MGBR2 23-24 que porta la mutación c.9248\_9256+7del.

La mutación **c.9256G>T** es una variante registrada en la base de datos BIC y catalogada como *nonsense*. Según el estudio informático, la mutación eliminaba el sitio donador del exón 24 (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 que porta la mutación c.9256G>T se pudo comprobar la alteración en el

proceso de *splicing* (Figura 27). La cuantificación relativa de las isoformas detectadas (Figura 23), reveló que el 28% de las isoformas perdió el exón 24, un 11% no contenía los últimos 43 nt del exón 24 debido al uso de un sitio donador críptico situado a 43 nt *upstream*. Además se detectó que el 19% de las isoformas restantes había perdido los exones 23 y 24, el resto (un 42%) correspondía al transcrito canónico.

La mutación **c.9256+1G>A** es una variante BIC y catalogada como intrónica. Según el estudio informático, la mutación elimina el sitio donador del exón 24 (Tabla 2). En los resultados del ensayo funcional con el MGBR2 23-24 que porta la mutación c.9256+1G>A se pudo comprobar la alteración de *splicing* (Figuras 23 y 26). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 23, reveló que el 69% de las isoformas *skipping* del exón 24, un 5% no contenía los últimos 43 nt del exón 24. Además se detectó que el 24% de las isoformas restantes habían perdido ambos exones, y el resto (2%) eran similares al transcrito WT.



**Figura 26. Secuencia de los productos RT-PCR originados por la mutación 9118-2A>G.** Cromatograma de secuenciación donde se observa la pérdida tanto del exón 23 como del 23 y 24 procedentes del MGBR2 23-24 que porta la mutación c.9256+1G>A.

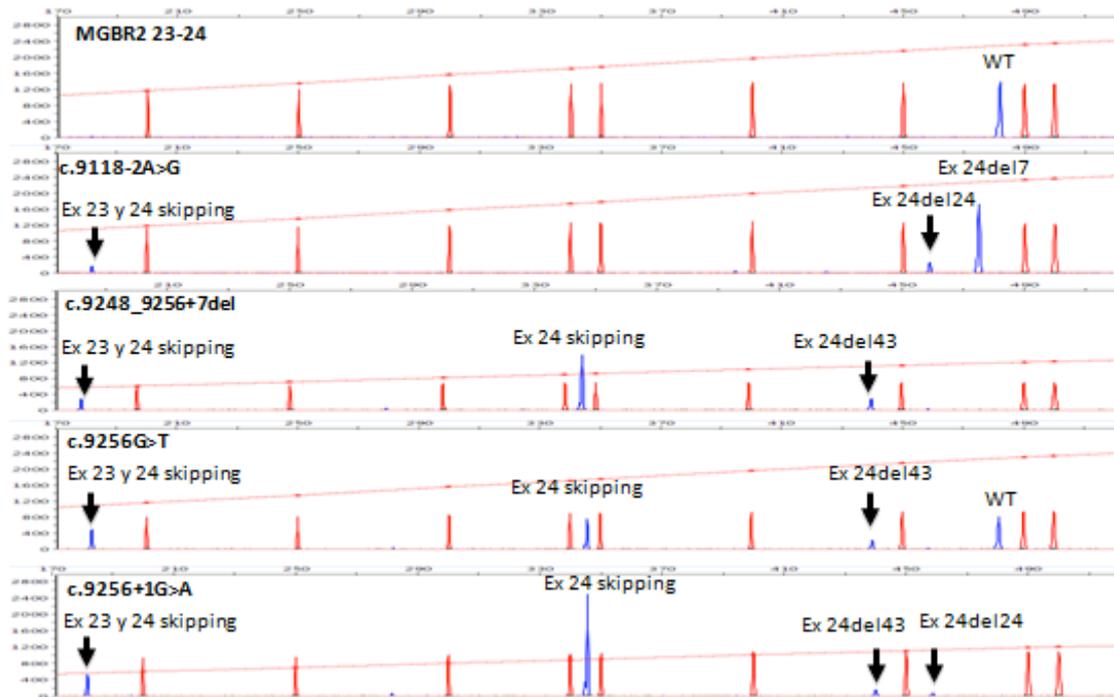
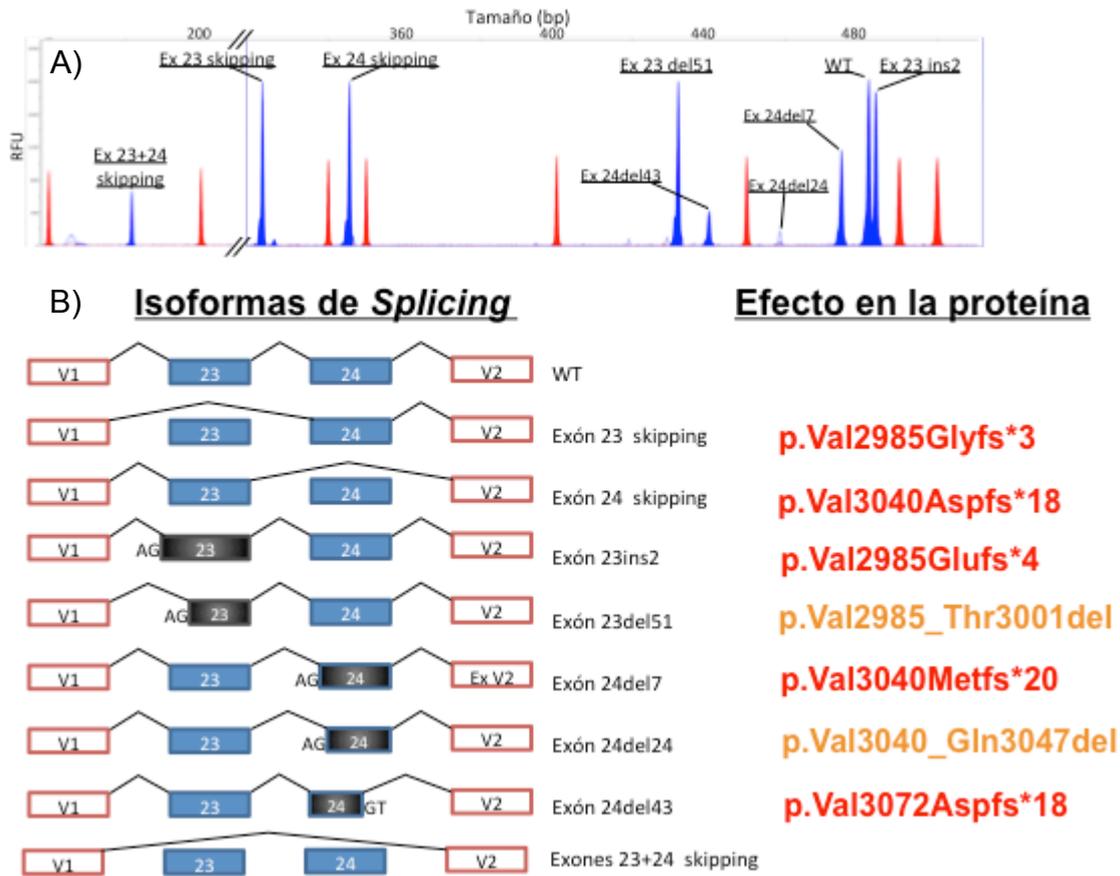


Figura 27. Electroforesis capilar de los ensayos funcionales de *splicing* de mutaciones del exón 24 de *BRCA2* (MGBR2 23-24) en células HeLa.

#### IV.2.c. Resultados globales de los exones 23 y 24

En conjunto, ocho isoformas de *splicing* además de la WT han sido identificadas en el análisis funcional de mutaciones mediante minigenes híbridos derivados del MGBR2 23-24 (Figura 28). Salvo por los transcritos que carecen de los exones 23 y 24, que pueden ser atribuidos al reducido contexto genómico en el ámbito del minigen, el resto de isoformas detectadas es muy probable que se manifieste también *in vivo*. Todas ellas, tienen un efecto claro sobre la traducción del ARN maduro a proteína. Las isoformas que carecen de los primeros 51 nt del exón 23 y las que no contienen los primeros 24 nt del exón 24 mantienen la fase de lectura y por lo tanto las proteínas resultantes de la traducción carecen de 17 aminoácidos y 8 respectivamente. Por otro lado, el resto de transcritos cambia la fase de lectura y en todos los casos aparece un codón prematuro de parada de la traducción, por lo que se producirían proteínas truncadas inactivas o parcialmente funcionales.



**Figura 28. Isoformas de *splicing* producidas por el minigen MGBR2 19-20 y efecto predecible en la proteína.** A) cromatogramas superpuestos donde se observan las diferentes isoformas detectadas en los ensayos funcionales con el MGBR2 23-24. B) Diagramas de representación de las reacciones de *splicing* que dan lugar a las diferentes isoformas y el efecto deducido en la proteína.

### V. Nuevo vector pSAD

La construcción de un nuevo vector obedeció al intento de mejorar la técnica de los ensayos funcionales de *splicing* con minigenes híbridos. Las técnicas de mutagénesis dirigida, inserciones y deleciones mediante PCR fueron determinantes para la construcción de pSAD. Todas las modificaciones introducidas fueron comprobadas mediante secuenciación. Además, se realizaron ensayos de transfección del células HeLa y posterior RT-PCR de los transcritos procedentes del vector, para confirmar la estabilidad del mismo. La secuencia completa de ADN del vector pSAD ha sido registrada mediante el procedimiento de patente. Las nuevas características del vector propiciaron el aumento del tamaño de los insertos clonados mejorando el contexto genómico de cada exón y los resultados de los minigenes construidos. El nuevo marcador de selección de coloración de colonias recombinantes con X-Gal

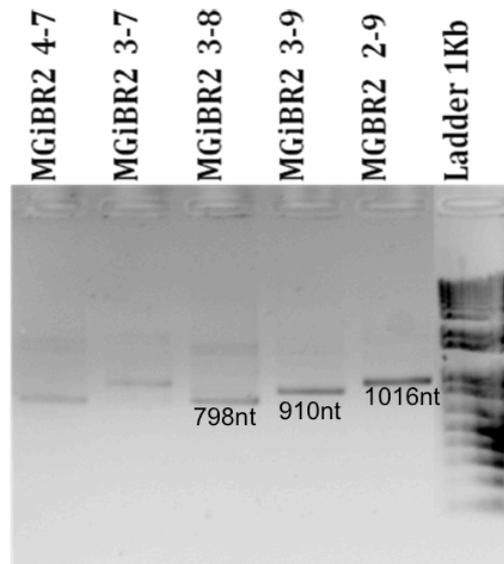
(gen de la  $\beta$ -galactosidasa o LacZ) disminuyó el tiempo y el coste necesario para la comprobación de las construcciones y sus derivados mutacionales.

## VI. Ensayos funcionales de *splicing*. Minigenes construidos en pSAD

Tres minigenes fueron construidos a partir del vector pSAD, en conjunto 21 de los 27 exones del gen *BRCA2* fueron clonados en pSAD. La construcción de minigenes multiexónicos requirió la combinación de diferentes técnicas de inserción y delección de secuencias de ADN. La clonación de grandes minigenes precisó de varias etapas de clonaje intermedias, por lo que se generaron minigenes intermedios hasta llegar a las construcciones finales. Tanto los minigenes finales como los intermedios fueron comprobados mediante secuenciación. También fueron transfectados en células HeLa y los diferentes transcritos producidos fueron visualizados mediante electroforesis en gel de agarosa de los productos RT-PCR.

### VI.1. Minigen MGBR2 2-9

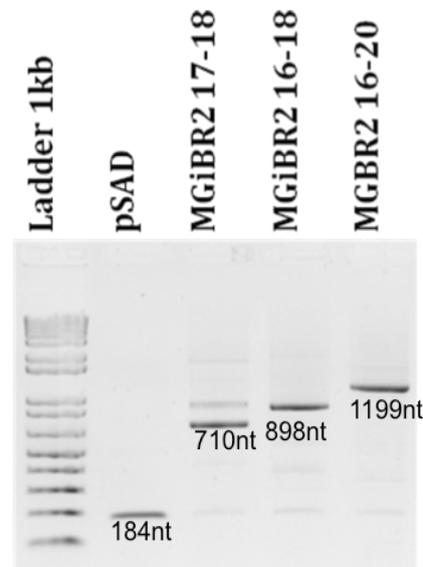
El MGBR2 2-9 contiene diferentes secuencias de ADN comprendidas entre el intrón 2 y el intrón 10 del gen *BRCA2*, incluyendo los 8 exones intermedios [IVS2 (252pb) - **EX2 (106pb)** - IVS3 (357pb) // IVS3 (210pb) - **EX3 (249pb)** - IVS4 (2077pb) - IVS4(494pb) - **EX4 (109pb)** - IVS5 (916pb) - **EX5 (50pb)** - IVS6 (91pb)- **EX6 (41pb)** - IVS7 (216pb) - **EX7 (115pb)** - IVS8 (296pb) //IVS8 (349pb) - **EX8 (50pb)** - IVS9 (190pb) // IVS9 (218pb) - **EX9 (102pb)** - IVS10 (295pb)]. El inserto total del MGBR2 2-9 consta de 4.977pb, mientras que para los mismos extremos la región genómica de *BRCA2* abarca 15.156pb. La construcción del MGB2 2-9 fue compleja y requirió generar cuatro minigenes intermedios (Tabla 4). Tanto el MGBR2 2-9 como los minigenes intermedios fueron comprobados mediante secuenciación. Tras la transfección de los mismos en células HeLa, se procedió a visualizar, por medio de electroforesis en gel de agarosa, el resultado de la amplificación de los transcritos obtenidos mediante cebadores que se solapan de forma específica en los exones del vector (Figura 29). El resultado es un minigen que produce un único transcrito por encima de 1kb y preparado para ser usado en ensayos funcionales de cualquier mutación que afecte a las regiones genómicas que contiene el MGBR2 2-9.



**Figura 29. Resultados de la RT-PCR del minigen MGBR2 2-9 *wild type* e intermedios.** Electroforesis en gel de agarosa (2%) del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 2-9 y las construcciones intermedias. Se puede observar la ganancia en nt en el tamaño del transcrito amplificado con cebadores específicos de los exones del vector, hasta superar 1 kb en el MGBR2 2-9. El MGiBR2 3-7 muestra una retención de intrón, el proceso de *splicing* se estabilizó con la inserción de nuevos exones.

## VI.2. Minigen MGBR2 16-20

El MGBR2 16-20 contiene los exones 16,17,18,19 y 20 del gen *BRCA2* y parte de las regiones intrónicas flanqueantes a los mismos [IVS16 (333pb) - **EX16 (188pb)** - IVS17 (234pb) // IVS17 (181pb) - **EX17 (171pb)** - IVS18 (485pb) - **EX18 (355pb)** - IVS19 (314pb) // IVS19 (235pb) - **EX19 (156pb)** - IVS20 (398pb) - **EX20 (145pb)** - IVS21 (207pb)]. El inserto total del MGBR2 16-20 consta de 3.402pb, mientras que para los mismos extremos, la región genómica de *BRCA2* abarca 13.899pb. Por su complejidad, la construcción final del MGBR2 16-20 requirió generar dos minigenes intermedios. Tras la transfección de los mismos en células HeLa, se visualizó, mediante electroforesis en gel de agarosa, el resultado de la amplificación de los transcritos obtenidos mediante cebadores específicos de los exones del vector (Figura 30). El resultado es un minigen que produce un único transcrito de tamaño esperado y preparado para ser usado en ensayos funcionales de cualquier mutación que afecte a las regiones genómicas que contiene el MGBR2 16-20.



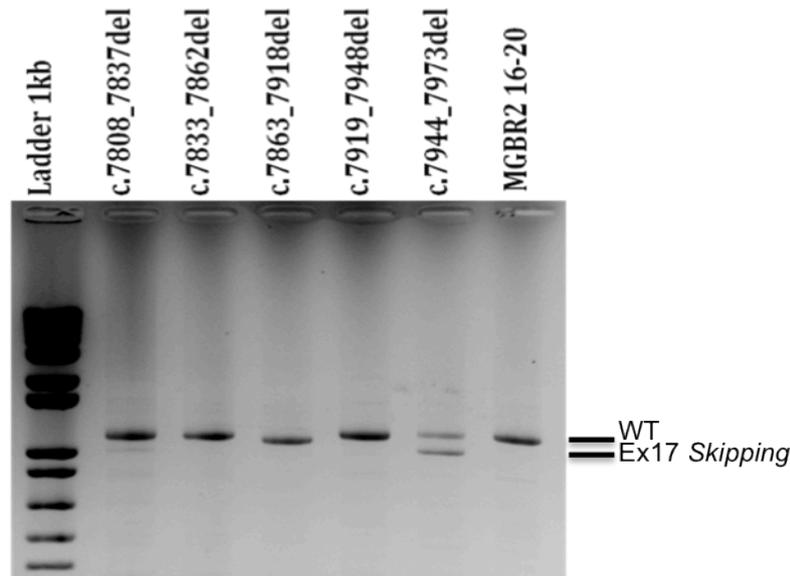
**Figura 30. Resultados de la RT-PCR del minigen MGBR2 16-20 *wild type* y construcciones intermedias.** Electroforesis en gel de agarosa (2%) del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 16-20 y los MG intermedios que han dado lugar a su construcción. Se puede observar la ganancia en nt en el tamaño del transcrito hasta superar 1,2 kb en el MGBR2 16-20. El minigen inicial MGiBR2 17-18 muestra una retención de intrón, pero las reacciones de *splicing* se estabilizaron con la inserción del exón anterior.

### VI.2.a. Mapeo de elementos reguladores de *splicing* mediante microdeleciones solapantes exónicas en el MGBR2 16-20

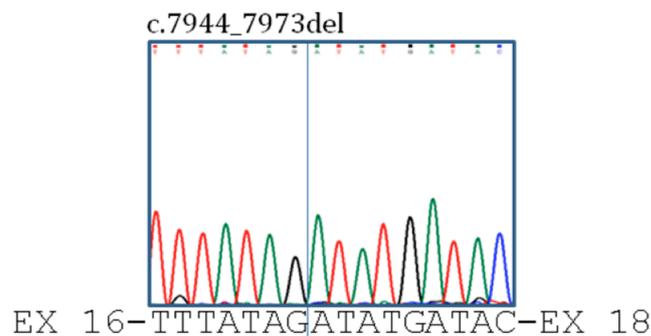
El desarrollo de nuevos protocolos aplicados en la construcción de pSAD, en concreto la técnica de generación de deleciones mediante PCR- mutagénesis dirigida, ha propiciado una nueva aproximación para la búsqueda de secuencias reguladoras de *splicing*. Específicamente, se realizaron veinte microdeleciones solapantes de 30 pb en los extremos de los exones y de hasta 125 pb en las regiones centrales. Tres microdeleciones afectaron el *splicing*, una en el exón 17 y dos en el exón 18.

La microdelección de 30 nt **c.7944\_7973del** finaliza 3 nt antes del final del exón 17. El sitio donador de este exón es especial: en lugar de un motivo GT se encuentra un GC. Según las predicciones bioinformáticas con HSF la microdelección podría haber reducido la probabilidad de que sitio donador fuera reconocido por la maquinaria de *splicing*. Las predicciones también indicaban que múltiples secuencias de reconocimiento de proteínas reguladoras tanto potenciadores como silenciadoras de *splicing* habrían sido eliminadas. En los resultados del ensayo funcional con el MGBR2 16-20 que porta la delección c.7944\_7973del se pudo comprobar (Figuras 31 y 32) que

parte de los transcritos portaban una delección del exón 17 sugiriendo que diferentes proteínas potenciadoras del tipo SR eran críticas para facilitar el reconocimiento del sitio donador del exón 17.



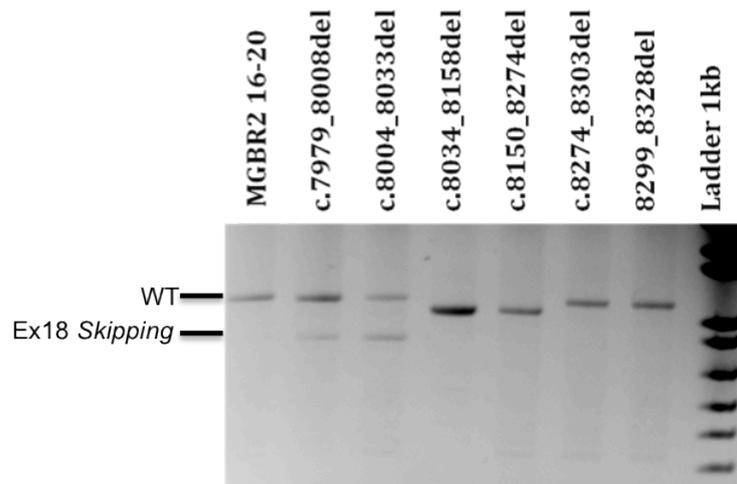
**Figura 31. Resultados de la RT-PCR del minigen MGBR2 16-20 *wild type* y microdelecciones realizadas en el exón 17.** Electroforesis en gel de agarosa (2%) del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 16-20 y los derivados del mismo a los que se le han realizado diferentes microdelecciones en el exón 17. La microdelección c.7944\_7973del alteró el *splicing* causando que parte de sus transcritos perdieran el exón 17.



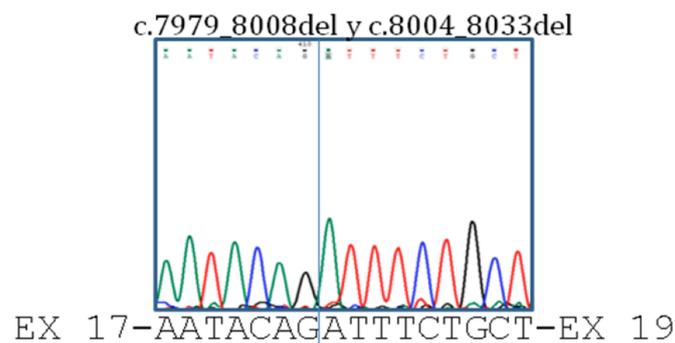
**Figura 32. Secuencia de los productos RT-PCR generados por la delección c.7944\_7973del.** Resultado de la secuenciación de los productos RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 16-20 con la microdelección c.7944\_7973del. La secuencia muestra la pérdida del exón 17 en los transcritos maduros.

Las microdelecciones de 30 nt **c.7979\_8008del** y **c.8004\_8033del** se sitúan en el extremo 5' del exón 18. La microdelección c.7979\_8008del respeta los primeros 2 nt del exón 18 y, por lo tanto, no afecta directamente al sitio aceptor. Ambas microdelecciones producirían la eliminación de múltiples secuencias reguladoras tanto de potenciadores como silenciadores de *splicing*. En los resultados del ensayo

funcional con el MGBR2 16-20 que portaba las deleciones c.7979\_8008del y c.8004\_8033del se pudo comprobar (Figuras 33 y 34) que parte de los transcritos carecían de exón 18, por lo que la teoría más apropiada para explicar este hecho, sugiere que diferentes ESEs son necesarios para facilitar el reconocimiento del sitio aceptor del exón 18.



**Figura 33. Resultados de la RT-PCR del minigen MGBR2 16-20 *wild type* y microdeleciones realizadas en el exón 18.** Resultado de electroforesis en gel de agarosa (2%) del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 16-20 y los derivados del mismo a los que se le han realizado diferentes microdeleciones en el exón 18. Las microdeleciones c.7979\_8008del y c.8004\_8033del han alterado el *splicing* causando que parte de sus transcritos maduros perdieran el exón 18.

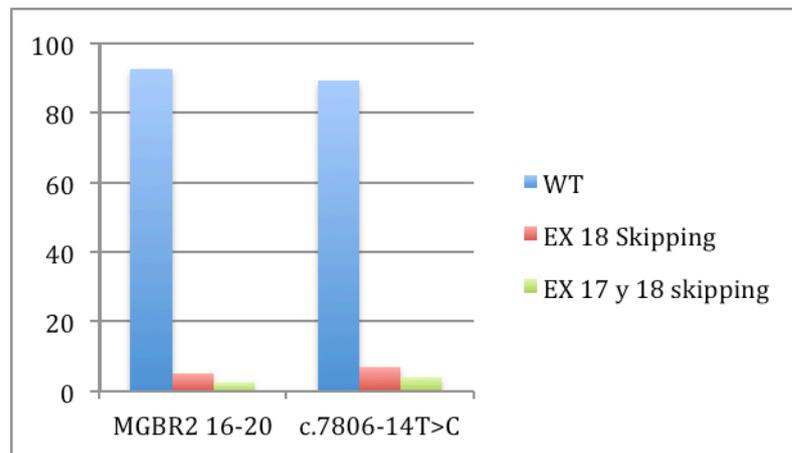


**Figura 34. Secuencia del producto RT-PCR generado por la microdelección c.7979\_8008del.** Secuenciación de los productos RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 16-20 con la microdelección c.7979\_8008del. El mismo resultado se ha obtenido con la c.8004\_8033del. La secuencia muestra la pérdida del exón 18 en los transcritos maduros.

### VI.2.b. Resultado mutagénesis MGBR2 16-20

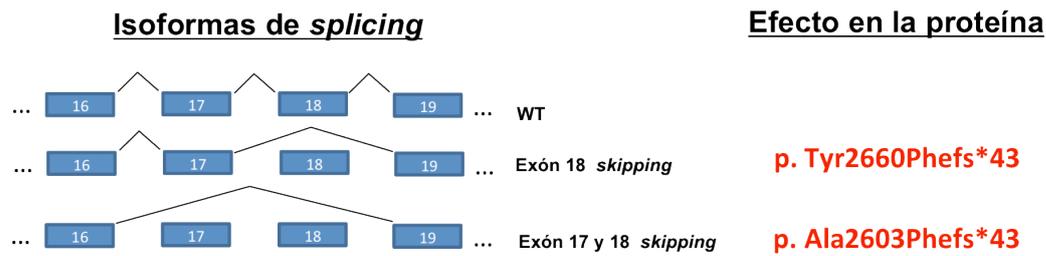
Únicamente la variante de ADN c.7806-14T>C situada en el intrón 16 de *BRCA2* fue reproducida en el MGBR2 16-20. El ensayo de la misma surgió de una colaboración con el Dr. Miguel de la Hoya. La variante c.7806-14T>C es común en la población y se

sitúa muy cercana al sitio aceptor del exón 18. A través del análisis informático se pudo detectar una escasa reducción del 4,68% (HSF) en la fortaleza del sitio aceptor. Los resultados obtenidos mediante el ensayo del MGBR2 16-20 WT y el que conteína la variante c.7806-14T>C mostraron en ambos casos la presencia del transcrito normal y dos isoformas aberrantes minoritarias: una, que no contiene los exones 17 y 18 y otra, en la que no es incluido el exón 18. Curiosamente, se ha descrito que estas dos isoformas son eventos de *splicing* alternativo que ocurren de forma natural (Walker *et al.*, 2010). La cuantificación relativa mostró un ligero incremento de las isoformas alternativas a causa de la mutación c.7806-14T>C. En concreto del 4,95% al 6,78% en la isoforma con el *skipping* del exón 18 y del 2,55% al 3,91% en la del *skipping* de los exones 17 y 18 (Figura 35).



**Figura 35. Cuantificación de las isoformas producidas por el MGBR2 16-20 y la mutación c.7806-14T>C.** Diagrama de representación de la media aritmética de la cuantificación relativa de cada isoforma detectada en la electroforesis capilar de los ensayos funcionales del MGBR2 16-20 y la mutación c.7806-14T>C ensayada sobre el mismo.

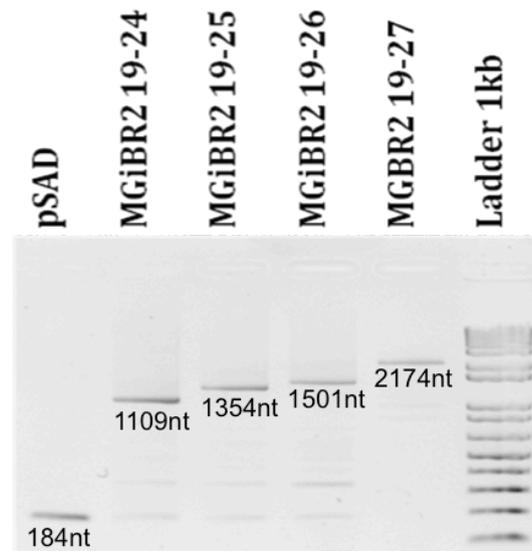
El efecto sobre la proteína de ambas isoformas es distinto. La traducción del transcrito que no contiene el exón 18 es interrumpida por la presencia de un codón de parada prematuro. El marco de lectura se mantiene intacto en los transcritos que no han incluido los exones 17 y 18. No obstante, los 175 aa codificados por estos exones forman parte del dominio BRCT que resulta esencial para la función de *BRCA2*.



**Figura 36. Isoformas de *splicing* detectadas tanto en el MGBR2 16-20 WT como aquel que contiene la mutación c.7806-14T>C y efecto predecible en la proteína.** Diagramas de representación de las reacciones de *splicing* que dan lugar a las diferentes isoformas detectadas al ensayar el MGBR2 16-20 WT y la mutación c. 7806-14T>C y el efecto de las mismas en la traducción a proteína.

### VI.3. Minigen MGBR2 19-27

El MGBR2 19-27 contiene los últimos nueve exones y parte de las regiones intrónicas flanqueantes a los mismos del gen *BRCA2* [IVS18 (247 pb) - **EX19 (156 pb)** - IVS19 (398 pb) - **EX20 (145 pb)** - IVS20 (207 pb) // IVS20 (90 pb) - **EX21 (122 pb)** - IVS21 (262 pb) // IVS21 (262 pb) - **EX22 (199 pb)** - IVS22 (234 pb) - **EX23 (164 pb)** - IVS23 (93 pb) - **EX24 (139 pb)** - IVS24 (147 pb) // IVS24 (271 pb) - **EX25 (245 pb)** - IVS25 (431 pb) // IVS25 (341 pb) - **EX26 (147 pb)** - IVS26 (344 pb) // IVS26 (221 pb) - **EX27 (723 pb)**]. Además, el exón final 27 sustituye parcialmente el exón final V2 del vector pSAD. El inserto total del MGBR2 19-27 consta de 5.620 pb, mientras que para los mismos extremos, la región genómica de *BRCA2* abarca 28.700 pb. Su diseño fue complejo y requirió la construcción de tres minigenes intermedios en pSPL3 y tres más en pSAD para conseguir el minigen final. Tras la transfección de aquellos construidos en pSAD en células HeLa, se visualizó, mediante electroforesis en gel de agarosa, el resultado de la amplificación de los transcritos obtenidos mediante cebadores específicos de los exones del vector (Figura 37). El resultado es un minigen que produce un único transcrito de 2.174 nt preparado para ser usado en ensayos funcionales de cualquier mutación que afecte a las regiones genómicas clonadas en MGBR2 19-27.



**Figura 37. Resultados de la RT-PCR del minigen MGBR2 19-27 *wild type* e intermedios.** Electroforesis en gel de agarosa (2%) del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y los MG intermedios construidos en pSAD. Se puede observar la ganancia en el tamaño del transcrito hasta los 2.174 nt del MGBR2 19-27

### VI.3.a. Resultados del MGBR2 19-27

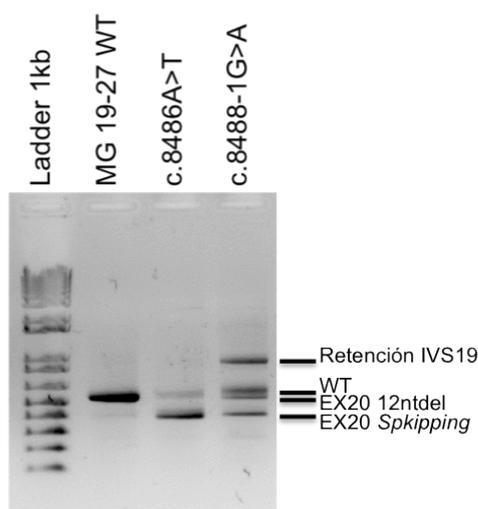
Con el propósito de validar el nuevo MGBR2 19-27 se realizaron cuatro mutagénesis distintas (c.8486A>T, c.8A88-1G>A, c.9117G>A, c.9256+1G>A) en representación de los cuatro exones ya analizados en el MGBR2 19-20 y MGBR2 23-24. Por otro lado, de las 166 mutaciones analizadas mediante análisis bioinformático y descritas en las bases de datos internacionales dentro de la región clonada en el MBGR2 19-27, 41 (Tabla 3) cumplieron los criterios establecidos para su posterior estudio mediante ensayos funcionales con minigenes híbridos: diez mutaciones en el exón 21 y sus secuencias intrónicas más próximas (c.8633-24\_8634del, c.8633-24\_8634del, c.8647C>T, c.8723T>G, c.8732C>T, c.8748ins4, c.8754G>A, c.8754+4A>G, c.8754+5G>A y c.8754+5G>T), seis en el exón 22 y sus regiones intrónicas más cercanas (c.8755-1G>A, c.8773C>T, c.8789delA, c.8915T>G, c.8948\_8953+5del y c.8953+1G>T), once para el exón 25 (c. 9257-18C>A, c. 9257-1G>C, c.9275\_9278del, c.9276T>G, c.9301C>G, c.9454G>A, c.9466delC, c.9501G>A, c.9501+3A>T, c.9501+4A>G y c.9501+9A>C), siete para el exón 26 (c.9502-12T>G, c.9502-2A>C, c.9550C>G, c.9610A>G, c.9611C>G, c.9613GC>CT y c.9646C>T) y tres en el exón 27 (c.9649 -5A>G, c 9649 -2A>G, c.9698 G>T).

Debido a que el tamaño del transcrito generado por el MGBR2 19-27 supera los 2.000 nt, se decidió realizar amplificaciones de menor tamaño. Para amplificar específicamente los transcritos procedentes del vector, se utilizó siempre un cebador específico del vector y otro localizado en uno de los exones del inserto. Los resultados de las amplificaciones mediante RT-PCR del ARN de células HeLa transfectadas se comprobaron mediante electroforesis en gel de agarosa, electroforesis capilar y secuenciación capilar. A continuación se muestran los resultados.

### VI.3.a.1. Resultados del MGBR2 19-27. Mutaciones ensayadas previamente en pSPL3

Los resultados de alteración del *splicing* del ensayo funcional con el MGBR2 19-27 que porta la mutación **c.8486A>T** fueron obtenidos mediante electroforesis en gel de agarosa (Figura 38), análisis de fragmentos (Figura 40) y secuenciación (datos no mostrados). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 39, reveló que el 80% de los transcritos había perdido el exón 19, mientras que el resto son similares al correspondiente WT.

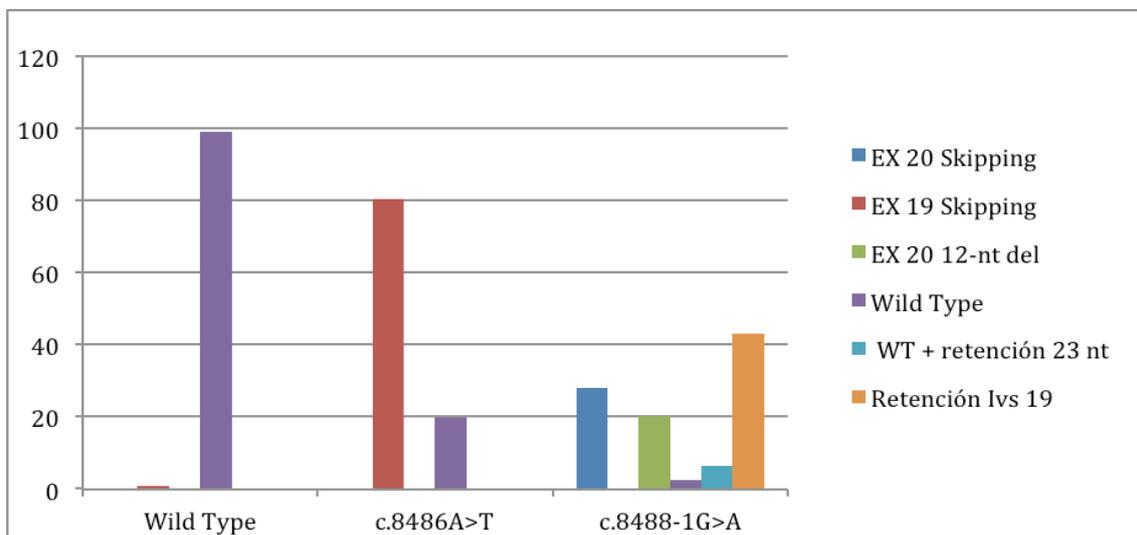
Resultados RT-PCR Mutaciones EX 19 y 20  
Cebadores: pSAD RT FW – 22 RT RV



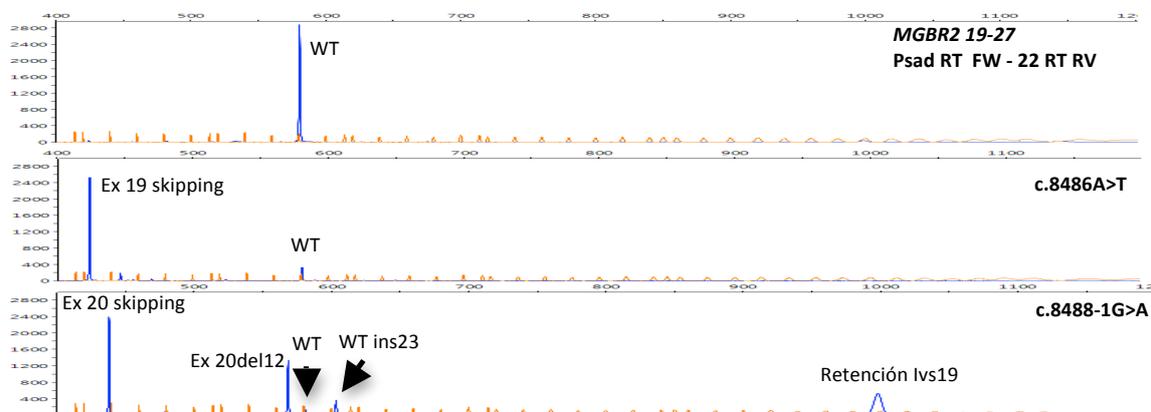
**Figura 38. Resultados de la RT-PCR del minigen MGBR2 19-27 *wild type* y de las mutaciones c.8486A>T y c.8488-1G>A ensayadas en los exones 19 y 20 respectivamente.** Electroforesis en gel de agarosa (1,5%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y dos derivados del mismo con las mutaciones c.8486A>T en el exón 19 y c.8488-1G>A en el exón 20.

Los resultados de alteración del *splicing* con el MGBR2 19-27 que porta la mutación **c.8488-1G>A** fueron obtenidos mediante electroforesis en gel de agarosa

(Figura 38), electroforesis capilar (Figura 40) y secuenciación (datos no mostrados). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar representada en la Figura 39, reveló que el 43% de los transcritos presenta retención del intrón 19, un 28% no ha incluido el exón 20, y otro 20% carece de los 12 primeros nucleótidos del exón 20, puesto que un sitio aceptor crítico es reconocido *downstream* del natural. Por último, un porcentaje minoritario del 2% fue similar al WT y un 7% de las isoformas contienen 23 nt más respecto al WT, pero no se pudo identificar la anomalía.



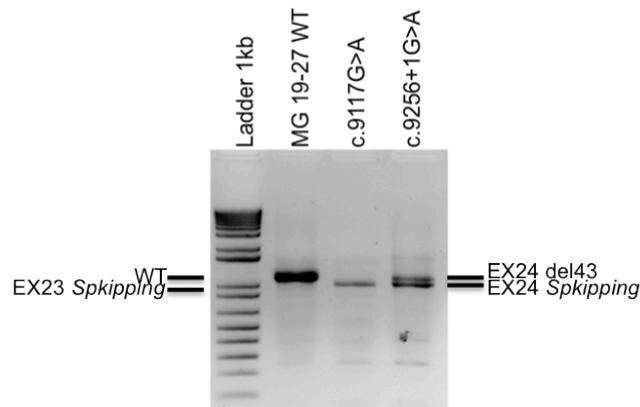
**Figura 39. Cuantificación de las isoformas producidas por las mutaciones c.8486A>T y c.8488-1G>A ensayadas en el MGBR2 19-27.** Diagrama de representación de la media aritmética de la cuantificación relativa de cada isoforma detectada en la electroforesis capilar de los ensayos funcionales de las mutaciones c.8486A>T en el exón 19 y c.8488-1G>A en el sitio aceptor del exón 20 del MGBR2 19-27.



**Figura 40. Electroforesis capilar de los ensayos funcionales de *splicing* de las mutaciones c.8486A>T y c.8488-1A>G del exón 19 e intrón 19, respectivamente, de BRCA2 (MGBR2 19-27) en células HeLa.**

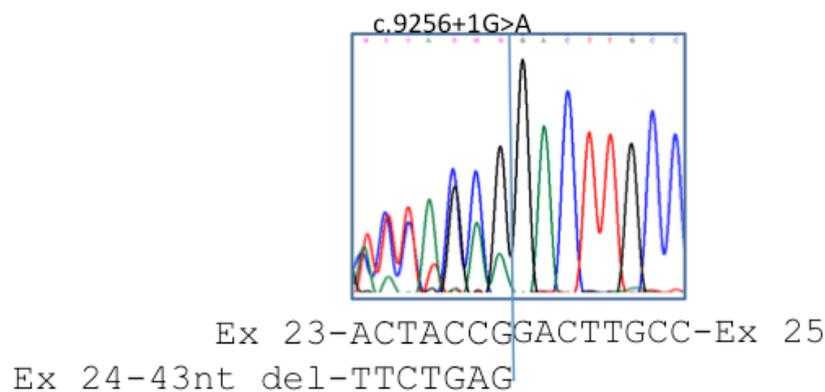
El ensayo funcional de la mutación **c.9117G>A** del exón 23 en el MGBR2 19-27 permitió detectar un único transcrito aberrante (100% de las isoformas) con delección del exón 23 (Figuras 41 a 44).

Resultados RT-PCR Mutaciones EX 23 y 24  
Cebadores: pSAD RT FW – 25 RT RV

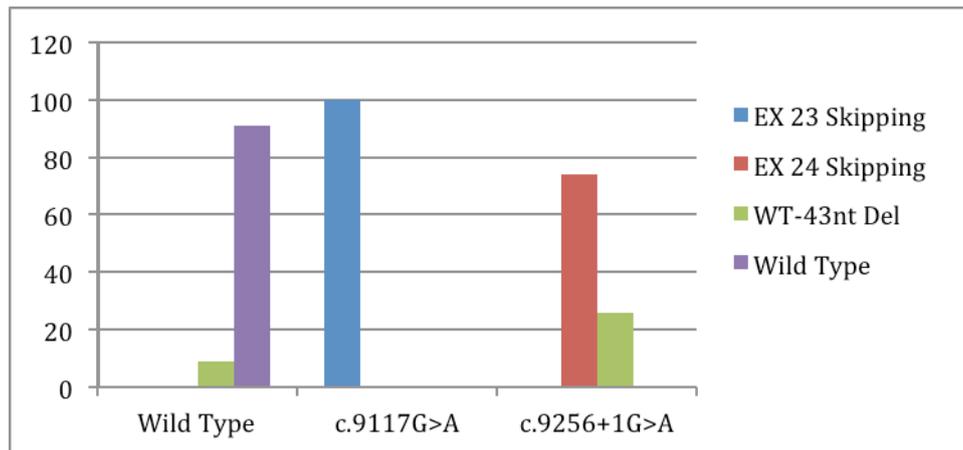


**Figura 41. Resultados de la RT-PCR del minigen MGBR2 19-27 *wild type* y de las mutaciones c.9117G>A y c.9256+1G>A ensayadas en el exón 23 e intrón 24, respectivamente.** Electroforesis en gel de agarosa (2%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y dos derivados del mismo con las mutaciones c.9117G>A y c.9256+1G>A en los sitios donadores de los exones 23 y 24, respectivamente.

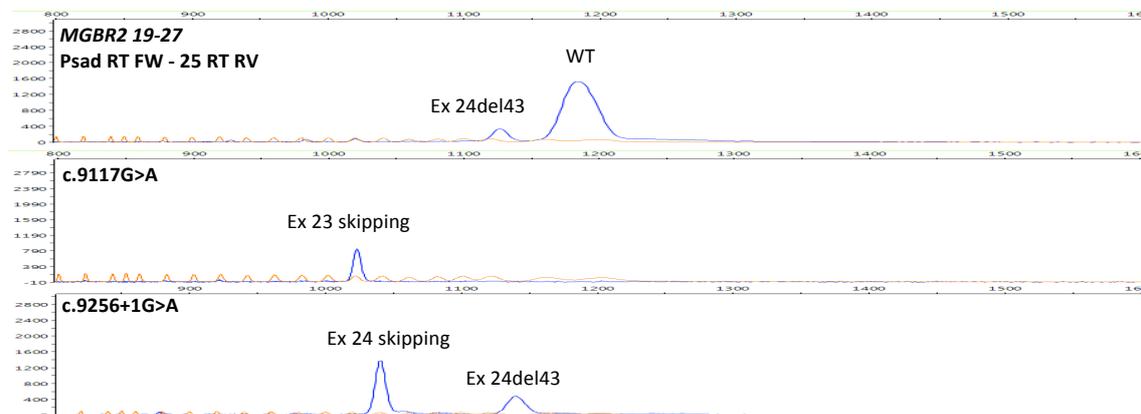
La mutación **c.9256+1G>A** en el MGBR2 19-27 generó dos isoformas aberrantes: *skipping* del exón 24 (74% de los transcritos) y pérdida de los últimos 43 nt del exón 24 (26%) debido al uso de un sitio donador críptico situado 43 nt *upstream* (Figuras 41-44).



**Figura 42. Secuencia de los productos RT-PCR generados por la mutación c.9256+1G>A.** En el cromatograma se puede apreciar la presencia de dos transcritos diferentes, uno que carece del exón 24 y otro que ha perdido los últimos 43 nt del exón 24.



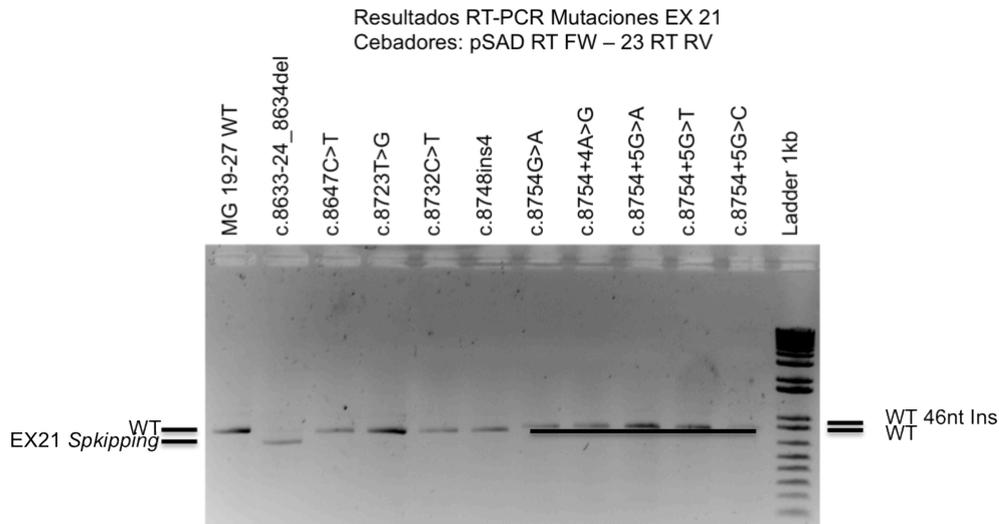
**Figura 43.** Cuantificación de isoformas producidas por las mutaciones c.9117G>A y c.9256+1G>A de los exón 23 y 24 respectivamente, ensayadas en el MGBR2 19-27. Diagrama de representación de la media aritmética de la cuantificación relativa de cada isoforma generada por las mutaciones c.9117G>A en el exón 23 y c.9256+1G>A en el sitio donador del exón 24 ensayadas en el MGBR2 19-27.



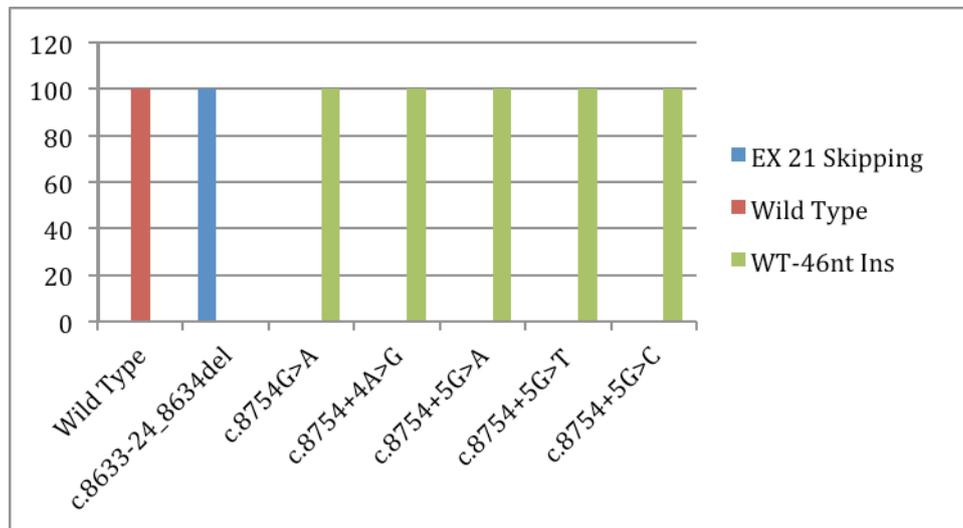
**Figura 44.** Electroforesis capilar de los ensayos funcionales de *splicing* de las mutaciones c.9117G>A y c.9256+1G>A en los exón 23 y 24 de *BRCA2* (MGBR2 19-27) en células HeLa.

### VI.3.a.2. Exón 21

Se generaron diez minigenes híbridos derivados del MGBR2 19-27 portadores de mutaciones en el exón 21 (Tabla 3). En los ensayos funcionales de *splicing* un total de seis alteraron este proceso (c.8633-24\_8634del, c.8754G>A, c.8754+4A>G, c.8754+5G>A, c.8754+5G>T, c.8754+5G>C). Los eventos de *splicing* mayoritarios en las mutaciones ensayadas fueron el *skipping* del exón 21 en el transcrito maduro y la inserción de los primeros 46 nt del intrón 21.

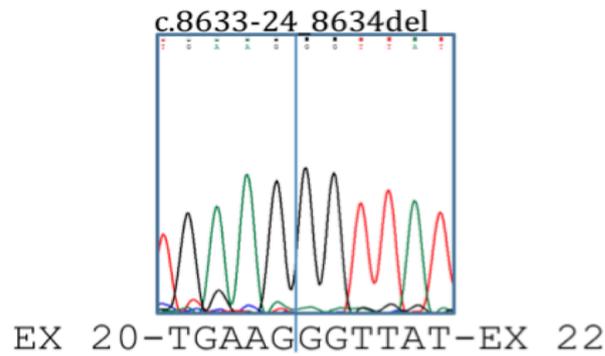


**Figura 45. Resultados RT-PCR MGBR2, mutaciones ensayadas exón 21, MGBR2 19-27.** Electroforesis en gel de agarosa (2%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y derivados del mismo con mutaciones en el exón 21.



**Figura 46. Cuantificación de las isoformas detectadas producidas por mutaciones del exón 21 en el minigen MGBR2 29-27.** Diagrama de representación de la media aritmética de la cuantificación relativa de cada isoforma o pico detectado en el cromatograma resultante de la electroforesis capilar para las mutaciones del exón 21 ensayadas en el MGBR2 19-27 y que han resultado alterar el *splicing*

La mutación **c.8633-24\_8634del** es una delección de 26 pb registrada en la base de datos BIC que elimina el sitio aceptor del exón 21 (Tabla 3). En los resultados del ensayo funcional del MGBR2 19-27 con la mutación c.8633-24\_8634del se comprobó la alteración del proceso de *splicing* (Figuras 45 a 49). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 46) reveló que el total de los transcritos generados habían perdido el exón 21.



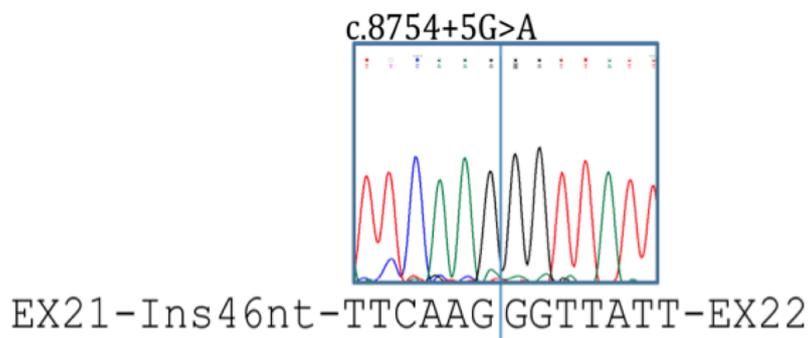
**Figura 47.** Secuencias de los productos RT-PCR procedentes de la mutación **c.8633-24\_8634del**. En el cromatograma se puede comprobar la ausencia del exón 21.

La variante **c.8754G>A** está registrada en la base de datos BIC catalogada como sinónima. Según el estudio informático, la mutación eliminaba el sitio donador del exón 21 (Tabla 3). En los resultados del ensayo funcional del MGBR2 19-27 con la mutación **c.8754G>A** se detectaron alteraciones en el proceso de *splicing* (Figuras 45 y 49). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 46), reveló que el total de los transcritos generados incluían los primeros 46 nt del intrón 21, debido al uso de un sitio donador críptico *downstream* del natural.

La mutación **c.8754+4A>G** es una variante intrónica registrada en la base de datos BIC. Según el estudio informático, **c.8754+4A>G** eliminaba el sitio donador del exón 21 (Tabla 3). El ensayo funcional del MGBR2 19-27 con la mutación **c.8754+4A>G** se identificó una alteración en los patrones de *splicing* (Figuras 45 y 49). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 46), reveló que los transcritos generados han incluido los primeros 46 nt del intrón 21 debido al uso de un sitio donador críptico *downstream* del natural.

Dos mutaciones intrónicas distintas han sido registradas en el BIC en la misma posición: **c.8754+5G>A** y de **c.8754+5G>T**. Una vez realizada la mutagénesis dirigida, detectamos mediante secuenciación una colonia que contenía el tercer cambio en la misma posición **c.8754+5G>C** y decidimos que también sería interesante ensayarla. Según el estudio informático, las tres mutaciones disminuían la probabilidad relativa de que el sitio donador del exón 21 sea reconocido de 0,98 a 0,42. Además, creaban un

sitio donador críptico 7 nt *downstream* del natural (Tabla 3). En los ensayos funcionales con tres minigenes diferentes contruidos a partir del MGBR2 19-27 y que portaban las variantes **c.8754+5G>A**, **c.8754+5G >T** y **c.8754+5G>C** se comprobó la presencia de una isoforma aberrante de *splicing* (Figuras 45, 48 y 49). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 46) indicó que el total de los transcritos generados en los tres casos presentaban una inserción de los primeros 46 nt del intrón 21 por el uso de un sitio donador críptico *downstream* del natural.



**Figura 48.** Secuencia de los productos RT-PCR generados por la mutación **c.8754+5G>A** en el MGBR2 19-27. En el cromatograma se puede comprobar que el transcrito contiene los primeros 46 nt del intrón 21.

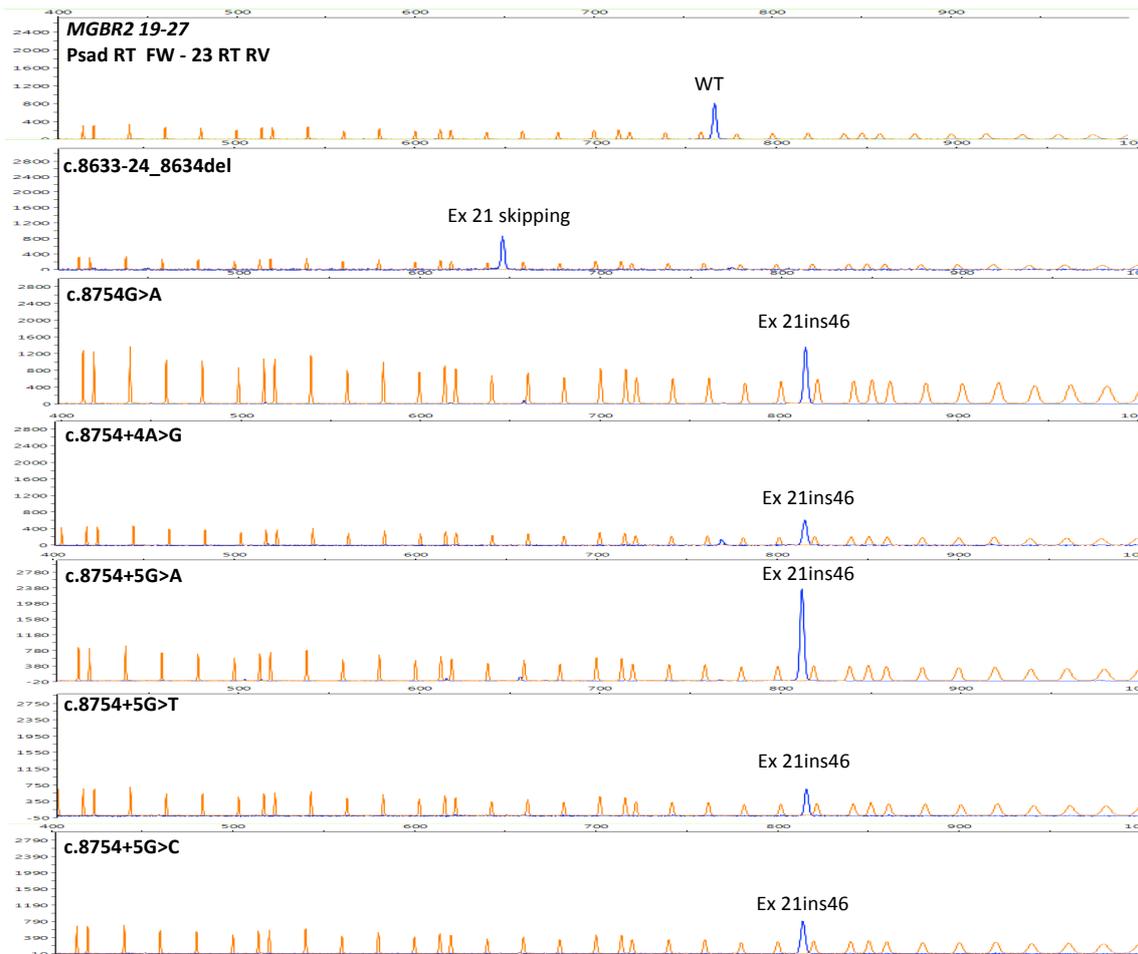


Figura 49. Electroforesis capilar de los ensayos funcionales de *splicing* de las mutaciones del exón 21 de *BRCA2* (MGBR2 19-27) en células HeLa.

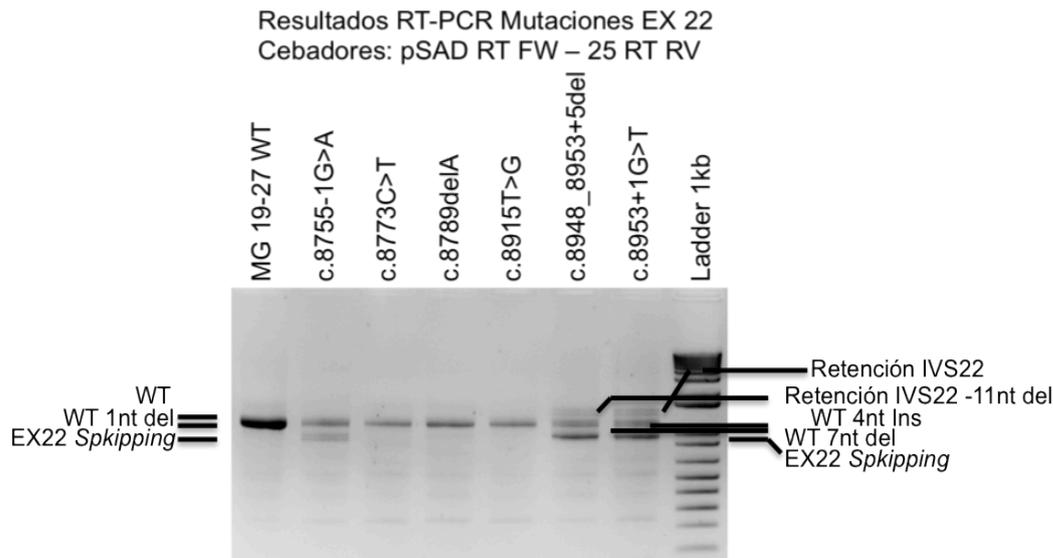
Los resultados obtenidos en el ensayo funcional de *splicing* con minigenes híbridos que han resultado positivos en el exón 21, muestran la presencia de dos tipos de transcritos aberrantes (Figura 50). La traducción de ambos transcritos es interrumpida por la presencia de un codón de parada prematuro. En estos casos, aquellos transcritos que no llegan a ser degradados por el mecanismo de *nonsense mediated mRNA decay* (NMD) producirán proteínas truncadas.



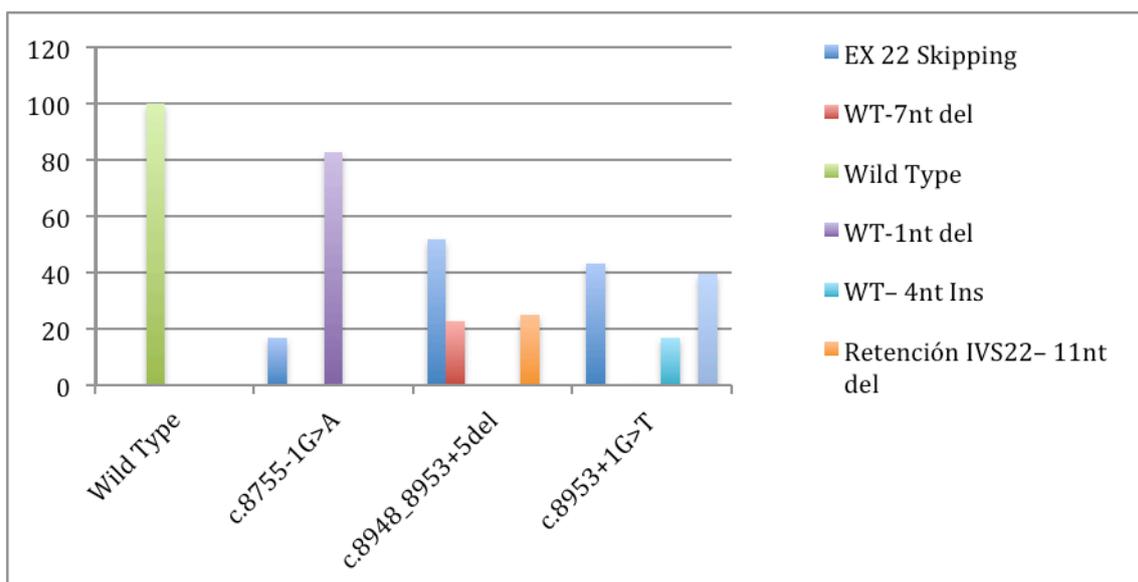
Figura 50. Isoformas de *splicing* detectadas en exón 21 del minigen MGBR2 19-20 y efecto predecible en la proteína de acuerdo a la nomenclatura HGVS. Diagramas del proceso de *splicing* que da lugar a las diferentes isoformas debido a mutaciones en el exón 21 y el efecto predecible en la proteína.

### VI.3.a.3. Exón 22

En los ensayos funcionales de *splicing* mediante minigenes híbridos derivados del MGB2 19-27 en los que se han reproducido seis variantes de ADN en el exón 22 y sus secuencias intrónicas inmediatas (Tabla 3), tres (c.8755-1G>A, c.8948\_8953+5del, c.8953+1G>T) alteraron el proceso de *splicing*. El principal evento de *splicing* común a las tres mutaciones positivas fue la delección del exón 22 en los transcritos maduros.

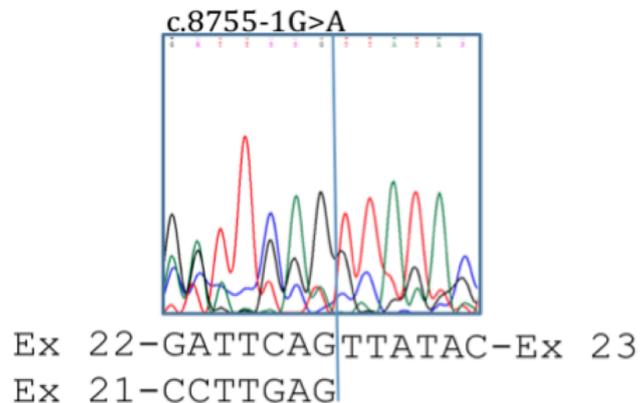


**Figura 51. Resultados de la RT-PCR del minigen MGBR2 19-20 *wild type* y mutantes del exón 22.** Electroforesis en gel de agarosa (2%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y derivados del mismo con mutaciones en el exón 22.



**Figura 52. Cuantificación de las isoformas detectadas producidas por mutaciones del exón 22 en el minigen MGBR2 19-27.** Diagrama de la media aritmética de la cuantificación relativa de cada isoforma detectada en la electroforesis capilar de las mutaciones del exón 22 ensayadas en el MGBR2 19-27.

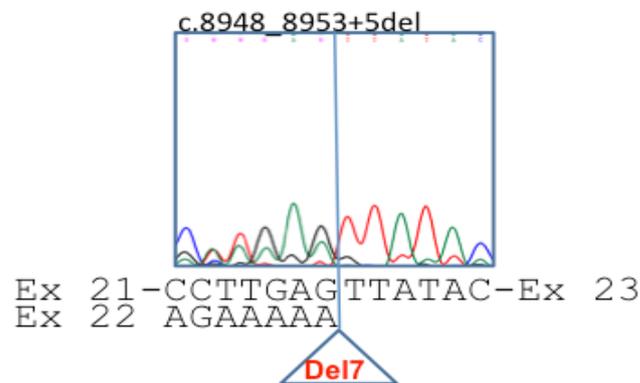
**BRCA2 c.8755-1G>A** es una variante intrónica registrada en la base de datos BIC. Según el estudio informático, la mutación eliminaba el sitio aceptor del exón 22 (Tabla 3). Los resultados del ensayo funcional con el MGBR2 19-27 con la mutación c.8755-1G>A indicaban una alteración del proceso de *splicing* (Figuras 51, 53 y 57). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 52), reveló la existencia de dos tipos de transcritos, pero sólo mediante secuenciación se ha podido afirmar que el 83% de los transcritos habían perdido el primer nucleótido del exón 22, puesto que la mutación crea un nuevo sitio aceptor 1 nt downstream del natural, mientras que el resto de transcritos había perdido el exón 22.



**Figura 53. Secuencia de los productos RT-PCR generados por la mutación.** Resultado de la secuenciación de los productos RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 con la mutación c.8755-1G>A. En el cromatograma se puede comprobar la presencia de transcritos WT y aquellos con delección del exón 22.

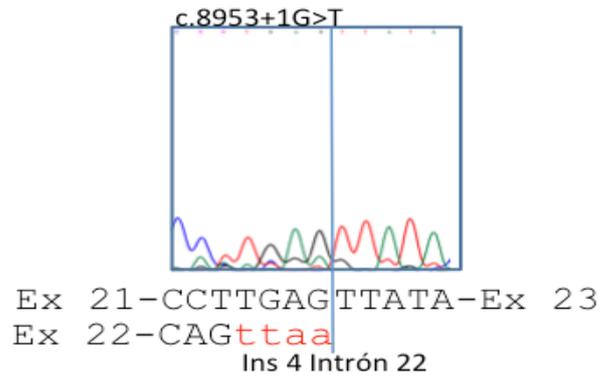
La delección **c.8948\_8953+5del** está registrada en la base de datos BIC y catalogada como de sitio de *splicing*. Según el estudio informático, la mutación eliminaba el sitio donador del exón 22 (Tabla 3). En el ensayo funcional con el MGBR2 19-27 que porta la mutación c.8948\_8953+5del se pudo comprobar la alteración de *splicing* (Figuras 51, 54 y 57). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 52) reveló que el 52% de los transcritos carecía del exón 22, en un 23% se había incluido el exón 22 sin los últimos 7 nt del exón. La causa es que la mutación en sí elimina los últimos 6 nt del exón 22 y genera un nuevo sitio donador alternativo (Figura 56A). La Guanina que debería ser el último nucleótido de este exón 22 mutado junto con la Timina de la posición c.8953+6 (puesto que los 5

primeros nucleótidos del intrón 22 también han sido eliminados a causa de la deleción) forman un nuevo sitio donador, por lo que la Guanina de la posición c.8947 tampoco es incluida (Figura 56A). Respecto al resto de los transcritos detectados, estos contienen el intrón 22 con la consiguiente deleción de 11 nt de la mutación estudiada. La retención de intrón se produjo por el uso del sitio donador del exón 23. De esta forma el exón 22, intrón 22 y exón 23 serían reconocidos por la maquinaria de *splicing* como una sola entidad.



**Figura 54. Secuencias de los productos RT-PCR procedentes de la mutación c.8949\_8953+5del.** En el cromatograma se puede comprobar la presencia de transcritos que carecen del exón 22 y aquellos que no incluyen los últimos 7 nt del exón 22.

La mutación **c.8953+1G>T** es una variante intrónica registrada en la base de datos BIC. Según el estudio informático, la mutación eliminaba el sitio donador del exón 22 (Tabla 3). En los resultados del ensayo funcional con el MGBR2 19-27 con la mutación c.8953+1G>T se comprobó la modificación del patrón de *splicing* canónico (Figuras 51, 55 y 57). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 52), reveló que el 44% de los transcritos perdió el exón 22, un 17% había incluido los primeros 4 nt del intrón 22, puesto que el espliceosoma reconoce un sitio donador a 4 nt del natural (Figura 56B), y el resto de los transcritos detectados contenía una retención del intrón 22, ya que el único sitio donador detectado a partir del aceptor del exón 22 era el correspondiente al exón 23. De esta forma, el exón 22, el intrón 22 y el exón 23 son reconocidos por la maquinaria de *splicing* como una sola entidad.



**Figura 55. Secuencia de los productos RT-PCR generados por la mutación c.8953+1G>T.** En el cromatograma se puede comprobar la presencia de transcritos que carecen del exón 22 y otros que incluyen los primeros 4 nt del intrón 22.



**Figura 56. Efecto de las mutaciones c.8948\_8953+5del y c.8953+1G>T a través de la generación de nuevos sitios donadores crípticos reconocidos por el espliceosoma.** La figura muestra el sitio donador natural GT del exón 22. A) La mutación c.8948\_8953+5del elimina 6 nt del final del exón 22 (en mayúsculas y color rojo) y 5 del inicio del intrón 22 (minúsculas y rojo). La deleción conforma un nuevo sitio donador Gt que es reconocido por el espliceosoma. B) La mutación c.8953+1G>T elimina el sitio GT del donador natural y el sitio donador críptico situado a 4 nt del original es reconocido por la maquinaria de *splicing*.

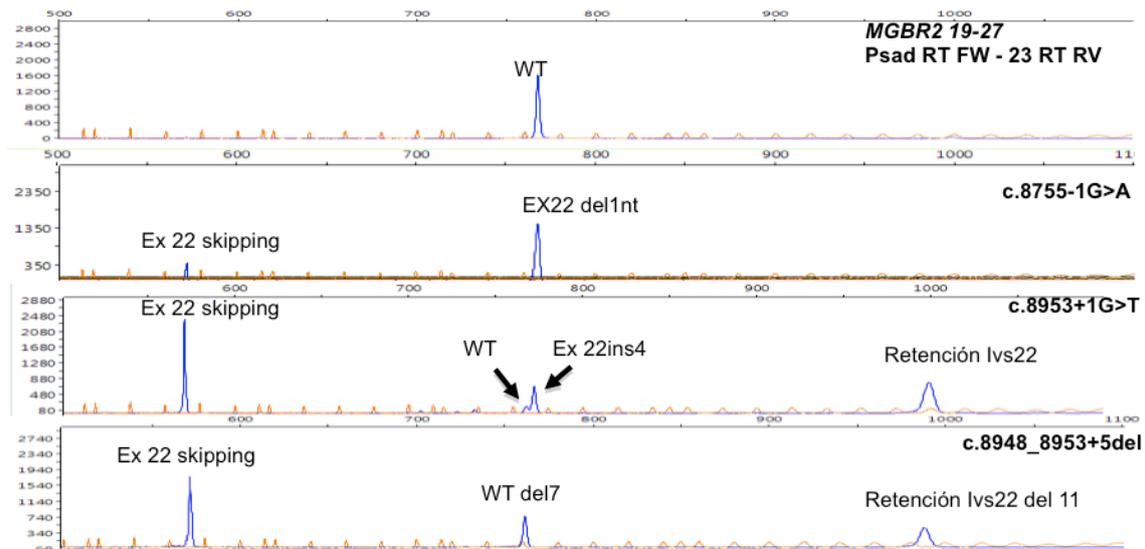


Figura 57. Electroforesis capilar de los ensayos funcionales de *splicing* de las mutaciones del exón 22 de *BRCA2* (MGBR2 19-27) en células HeLa.

Los resultados obtenidos en el ensayo funcional de *splicing* con minigenes híbridos que han resultado positivos en el exón 22 muestran la presencia de seis tipos diferentes de transcritos aberrantes (Figura 58). La traducción de todos ellos sería interrumpida por la presencia de un codón de parada prematuro. En estos casos, aquellos transcritos que no llegan a ser degradados por el mecanismo de NMD producirían proteínas truncadas.

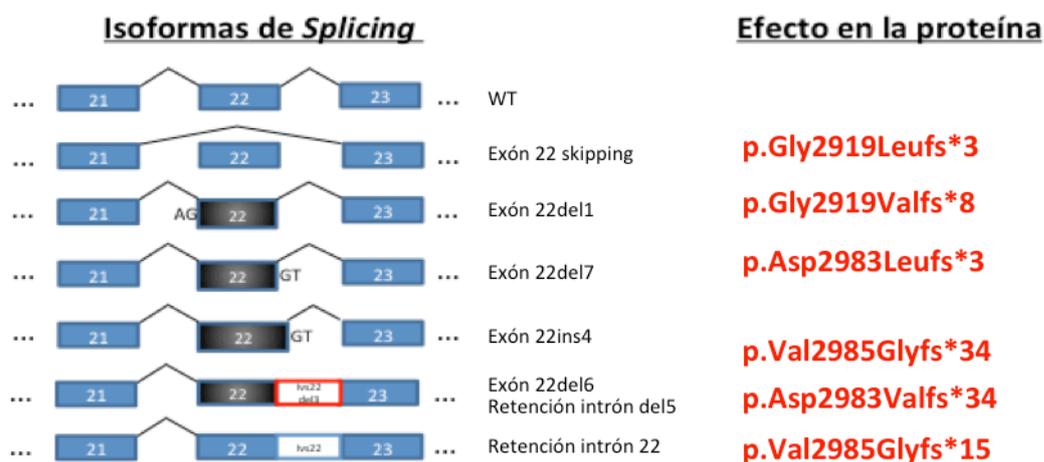
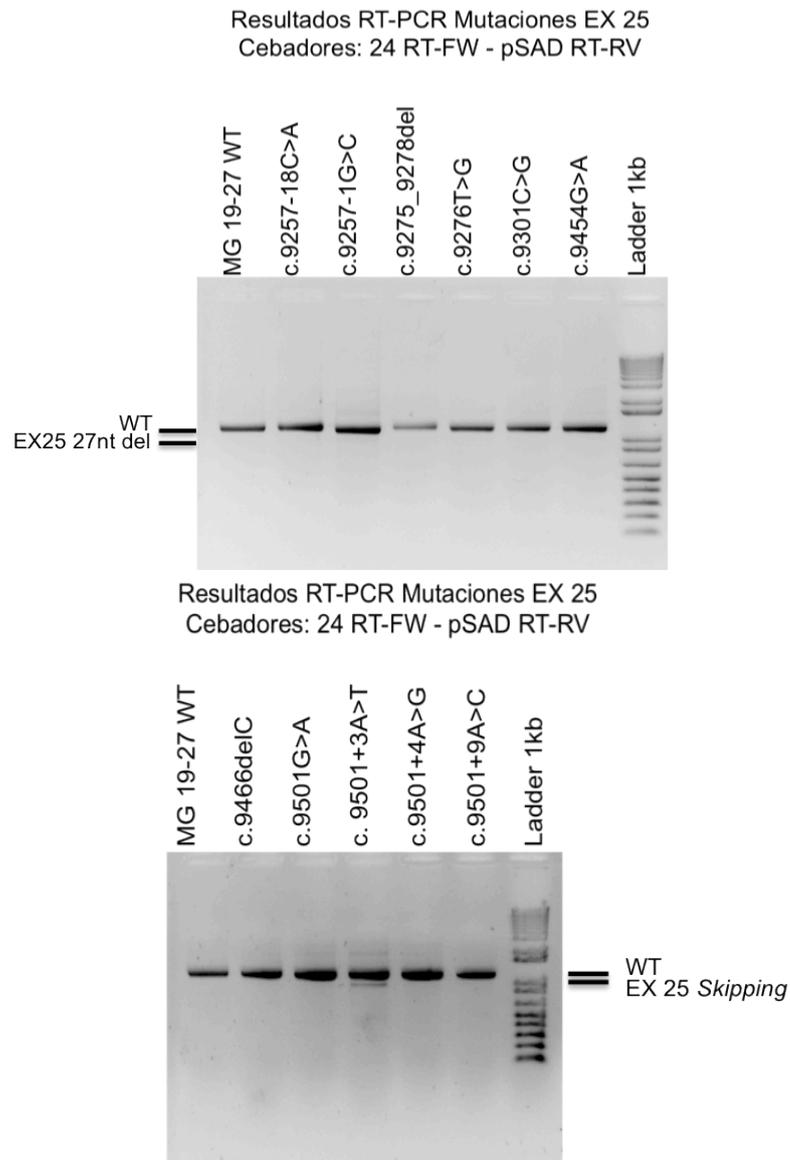


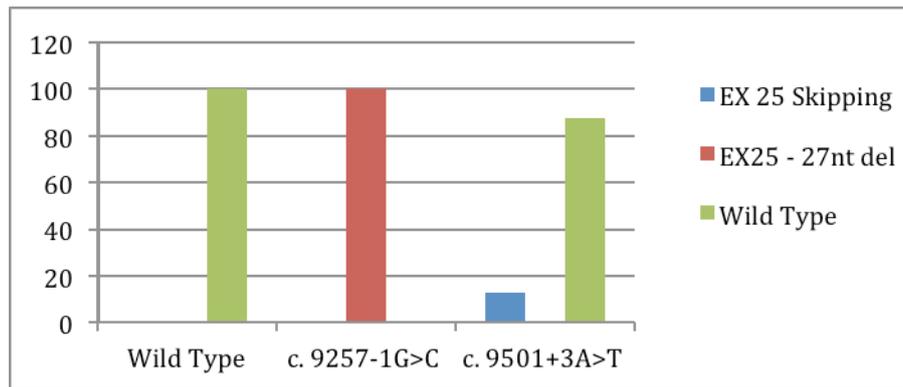
Figura 58. Isoformas de *splicing* producidas por mutaciones en el exón 22 (MGBR2 19-27) y efecto predecible en la proteína. Diagramas de representación del proceso de *splicing* que da lugar a las diferentes isoformas debido a mutaciones en el exón 22 y el efecto de las mismas en la traducción a proteína.

### VI.3.a.4. Exón 25

En los ensayos funcionales de *splicing* con once minigenes híbridos derivados del MGBR2 19-27 con mutaciones en el exón 25 (Tabla 3), dos que contenían las variantes de ADN (c.9257-1G>C y c. 9501+3A>T) alteraron el proceso de *splicing*. Los resultados se muestran a continuación.

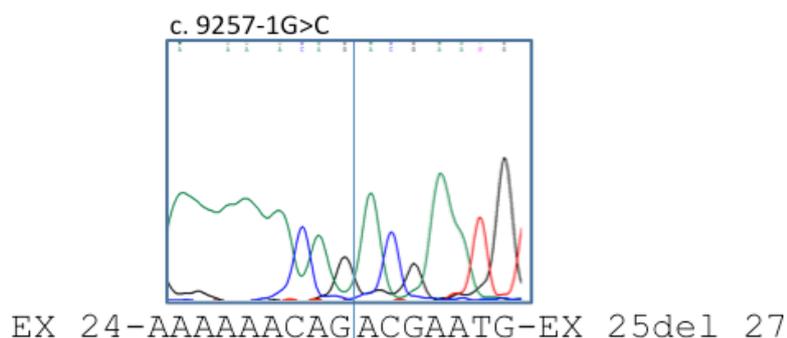


**Figura 59. Resultados de la RT-PCR del MGBR2 19-27 y mutaciones ensayadas en el exón 25.** Electroforesis en gel de agarosa (2%) de productos RT-PCR generados por el minigen MGBR2 19-27 y derivados del mismo con mutaciones en el exón 25.



**Figura 60. Cuantificación de las isoformas detectadas producidas por mutaciones del exón 25 en el MGBR2 19-27.** Diagrama de la media aritmética de la cuantificación relativa de cada isoforma detectada mediante electroforesis capilar para las mutaciones del exón 25 ensayadas en el MGBR2 19-27.

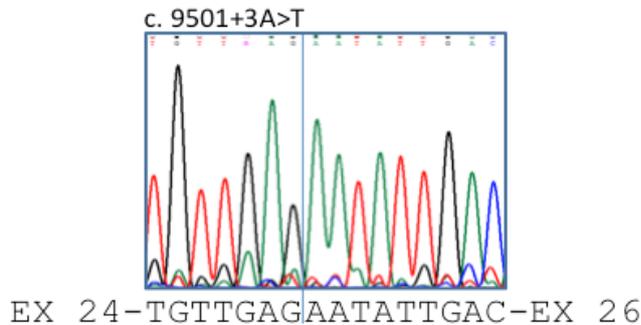
La mutación **c.9257-1G>C** es una variante intrónica registrada en la base de datos BIC. Según el estudio informático, la mutación alteraba el sitio aceptor del exón 25, además de alterar múltiples elementos reguladores de *splicing* (Tabla 3). En los resultados del ensayo funcional con el MGBR2 19-27 con la mutación c.9257-1G>C se comprobó que alteraba el proceso de *splicing* (Figura 59, 61 y 63). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 60) reveló que la totalidad de los transcritos carecía de los primeros 27 nt del exón 25, debido a que la maquinaria de *splicing* reconocía un sitio aceptor 27 nt *downstream* del natural. Las herramientas bioinformáticas predecían una probabilidad relativa muy baja (0,36) de reconocimiento de este sitio aceptor crítico por parte del espliceosoma.



**Figura 61. Secuencia de los productos RT-PCR procedentes de la mutación c.9257-1G>A.** En el cromatograma se puede comprobar la presencia de transcritos que carecen de los primeros 27 nt del exón 25.

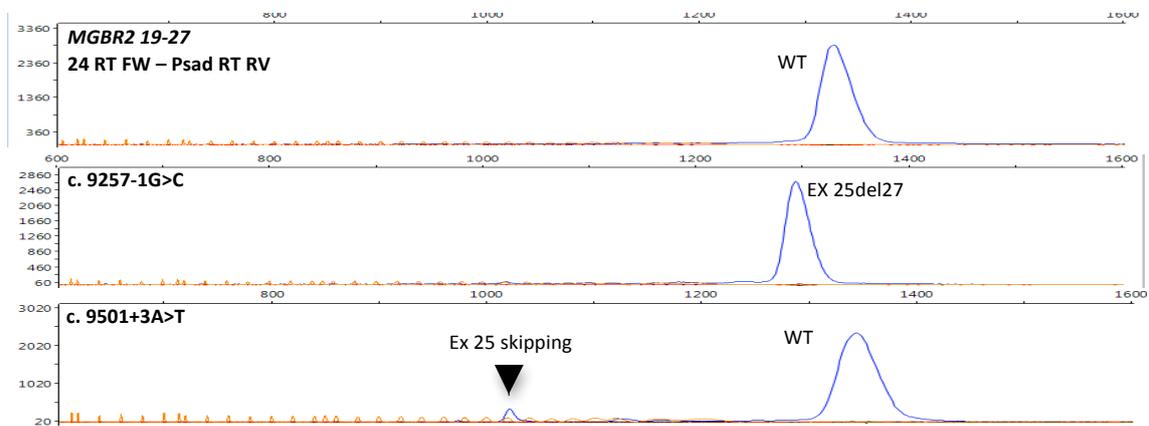
La mutación **c.9501+3A>T** es una variante intrónica registrada en la base de datos BIC. Según el estudio informático, la mutación disminuía la probabilidad relativa de reconocimiento del sitio donador del exón 25 por parte de la maquinaria de *splicing*

de 0,99 a 0,41 (Tabla 3). Los resultados del ensayo funcional con el MGBR2 19-27 que porta la mutación c.9501+3A>T indicaron que dicha variante alteraba el proceso de *splicing* (Figuras 59, 62 y 63). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 60) indicó que sólo el 13% de los transcritos había perdido el exón 25, mientras que el resto era del mismo tamaño que el WT.



**Figura 62. Secuencia de los productos RT-PCR generados por la mutación c.9501+3A>T.** En el cromatograma se puede comprobar la presencia de transcritos que no contienen el exón 25.

Las previsiones bioinformáticas para las mutaciones c.9501G>A y c.9501+4A>G (Tabla 3) eran muy parecidas, ya que sendas mutaciones disminuían el valor del sitio donador de 0,99 a 0,60 y 0,78 respectivamente. Sin embargo, no se apreció efecto en el *splicing*.



**Figura 63. Electroforesis capilar de los ensayos funcionales de *splicing* de las mutaciones positivas del exón 25, MGBR2 19-27.**

Los resultados obtenidos en el ensayo funcional de *splicing* de minigenes híbridos con variantes del exón 25 con impacto en el *splicing*, mostraban la presencia de dos tipos de transcritos aberrantes (Figura 64). Los transcritos que carecían de los primeros 27 nt del exón 25 provocaron el cambio de aminoácido en la posición 3085 y que la proteína perdiera 9 aminoácidos. Sin embargo, en los transcritos con delección

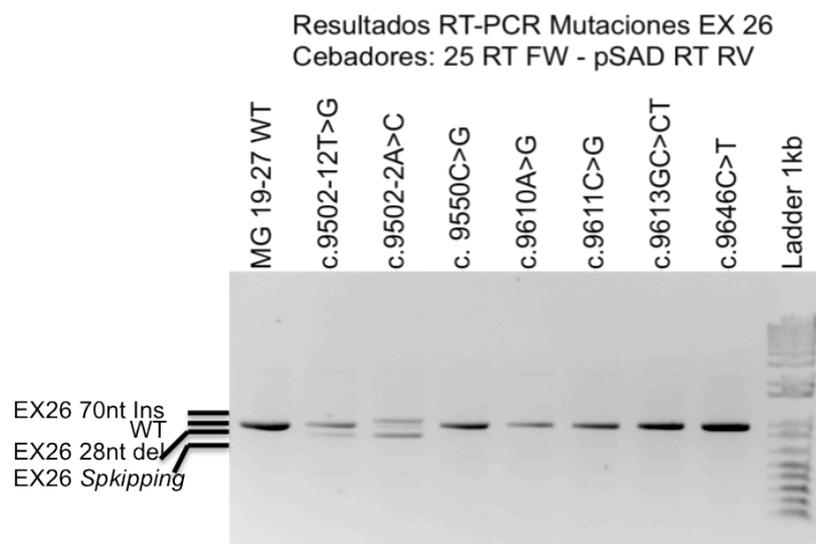
del exón 25, la traducción sería interrumpida por la presencia de un codón de parada prematuro. En este caso, aquellos transcritos que no llegan a ser degradados por el mecanismo de NMD producirían proteínas truncadas.



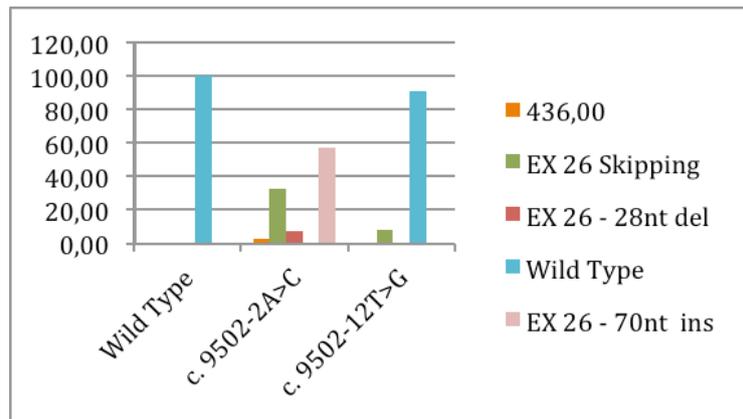
**Figura 64. Isoformas de *splicing* identificadas por mutaciones en el exón 25 (MGBR2 19-27) y su efecto predecible en la proteína.** Diagramas de representación del proceso de *splicing* que da lugar a las diferentes isoformas por mutaciones en el exón 25 y el efecto de las mismas en la traducción a proteína.

### VI.3.a.5. Exón 26

En los ensayos funcionales de *splicing* con siete minigenes híbridos derivados del MGBR2 19-27 con mutaciones en el exón 26 (Tabla 3), dos variantes (c.9502-12T>G y c.9502-2A>C) alteraron el proceso de *splicing*. Los resultados se muestran a continuación.

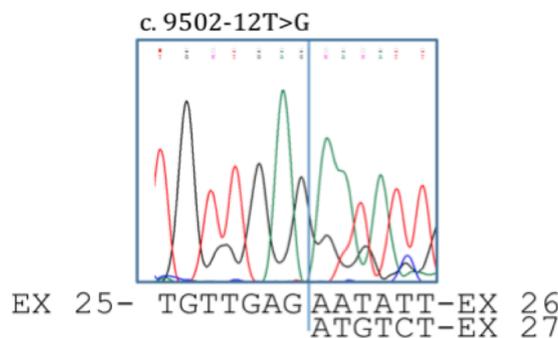


**Figura 65. Resultados de la RT-PCR del minigen MGBR2 19-27 *wild type* y mutantes del exón 26.** Electroforesis en gel de agarosa (2%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y derivados del mismo con mutaciones en el exón 26.



**Figura 66. Cuantificación de isoformas detectadas por mutaciones en el exón 26 del MGBR2 19-27.** Diagrama de la media aritmética de la cuantificación relativa de cada isoforma detectada en la electroforesis capilar para las mutaciones del exón 26 ensayadas en el MGBR2 19-27.

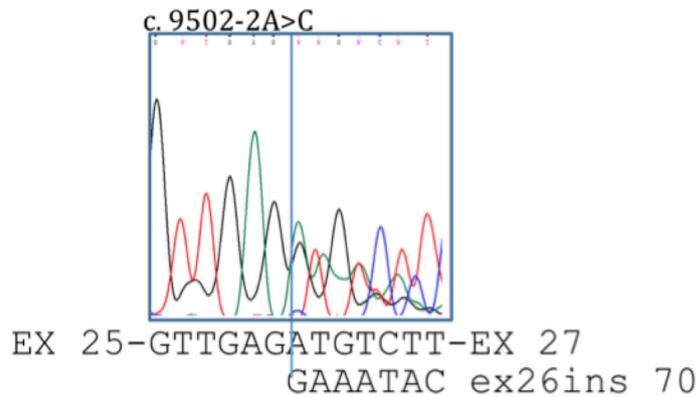
La mutación **c.9502-12T>C** es una variante intrónica registrada en la base de datos BIC. Según el estudio informático, la mutación reducía la probabilidad relativa del sitio aceptor del exón 26 de 0,93 a 0,74 (Tabla 3). Su posición, además, indica que podría estar afectando al tracto de polipirimidina. En los ensayos funcionales con el MGBR2 19-27 que portaba la mutación c.9502-12T>C se detectó la alteración en el proceso de *splicing* (Figuras 65, 67 y 69). La cuantificación relativa de las isoformas detectadas mediante electroforesis capilar (Figura 64), revela que el 8% de los transcritos carece del exón 26, mientras que el resto es similar al WT.



**Figura 67. Secuencia de los productos RT-PCR generados por la mutación c.9502-12T>G.** En el cromatograma se puede comprobar la presencia de transcritos WT y aquéllos con delección del exón 26.

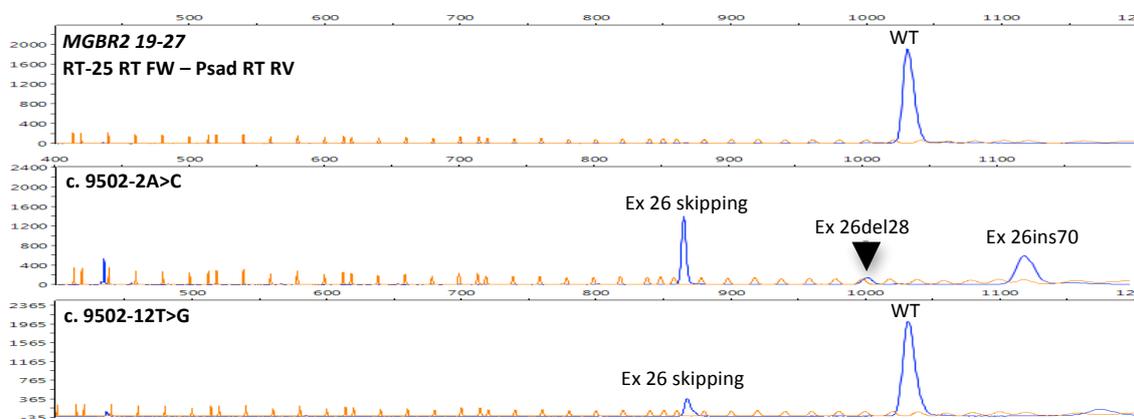
La mutación **c.9502-2A>C** es una variante intrónica registrada en la base de datos BIC. Según el estudio informático, la mutación eliminaba el sitio aceptor del exón 26 (Tabla 3). El ensayo funcional del MGBR2 19-27 con la mutación c.9502-12A>G permitió detectar un patrón de *splicing* aberrante (Figuras 65, 68 y 69). La cuantificación relativa de las isoformas (Figura 66) mostró que el 33% de los transcritos no habían incluido el exón 26, mientras que en un 57% se habían insertado los últimos

70 nt del intrón 25, debido al reconocimiento de un sitio aceptor críptico 70 nt *upstream* del natural, que según las predicciones bioinformáticas tenía un probabilidad alta de ser reconocido por el espliceosoma (NNsplice 0,81). El 8% de los transcritos incluían el exón 26 con una deleción de los primeros 28 nt. En este caso, la maquinaria de *splicing* reconocía un sitio aceptor situado 28 nt *downstream* del natural sin embargo, este sitio críptico no fue detectado por las previsiones bioinformáticas.



**Figura 68.** Secuencia de los productos RT-PCR generados por la mutación c.9502-2A>C. En el cromatograma se puede comprobar la presencia tanto de transcritos que incluyen los último 70 nt como aquéllos con *skipping* del exón 26.

Existe un porcentaje muy minoritario de transcritos de un tamaño de 436 nt de origen desconocido que ha sido detectado tanto en el MGBR2 19-27 como los derivados con las mutaciones con efecto en el *splicing*. Sin embargo, no se ha podido deducir ni comprobar su secuencia. Probablemente se trate de un producto inespecífico procedente de las condiciones de amplificación para este exón.



**Figura 69.** Electroforesis capilar de los ensayos funcionales de *splicing* de las mutaciones del exón 26 de *BRCA2* (MGBR2 19-27) con impacto en el *splicing* en células HeLa.

Los resultados obtenidos en el ensayo funcional de *splicing* con minigenes híbridos del exón 26 mostraban la presencia de tres diferentes tipos de transcritos aberrantes (Figura 70). Los transcritos sin exón 26 provocarían que la proteína tuviera una delección de 49 aminoácidos manteniéndose la pauta de lectura. Los aminoácidos codificados por el exón 26 se sitúan de forma consecutiva al dominio DBD, cuya función es la unión de la proteína BRCA2 con hebras sencillas de ADN. Sin embargo, en los transcritos que habían incluido los últimos 70 nt del intrón 25 y perdido los primeros 28 nt del exón 26, la traducción es interrumpida por la presencia de un codón de parada prematuro. En este caso, aquellos transcritos que no llegan a ser degradados por el mecanismo de NMD producirían proteínas truncadas.

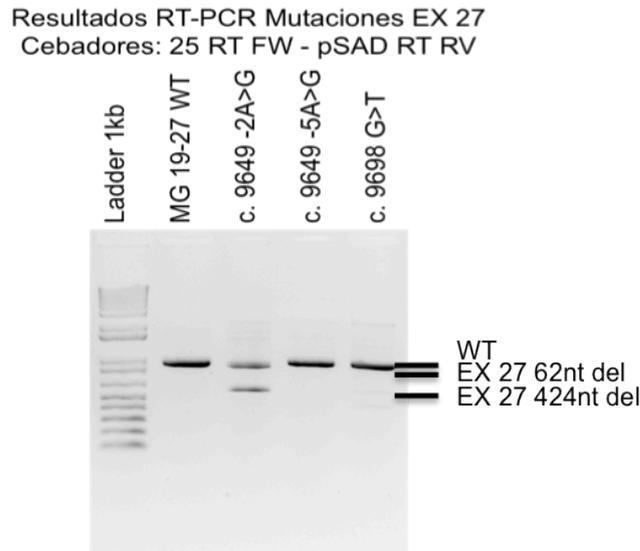


**Figura 70. Isoformas de *splicing* detectadas en el exón 26 del minigen MGBR2 19-27 y efecto predecible en la proteína.** Diagramas de representación de las reacciones de *splicing* que dan lugar a las diferentes isoformas del exón 26 y el efecto de las mismas en la traducción a proteína.

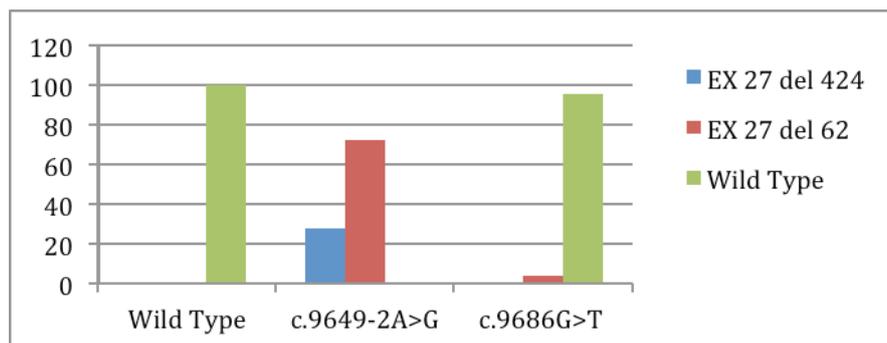
### VI.3.a.6. Exón 27

El exón 27 es el último del gen *BRCA2*. Los exones finales sólo tienen sitio aceptor. Los mecanismos por los que el extremo 3' es reconocido son complejos y necesitan de diferentes factores relacionados con la poliadenilación. Aunque en el exón 27 se han registrado múltiples variantes de ADN, sólo tres de las veinticinco variantes analizadas en los primeros 100 nt del exón (c.9649-2A>G, c.9649-5A>G y c.9698G>T; Tabla 3) fueron seleccionadas para la realización de ensayos funcionales, por su posible alteración del sitio aceptor.

Los resultados obtenidos mostraron que las mutaciones c.9649-2A>G y c.9698G>T alteraron el *splicing*. Los resultados se muestran a continuación.



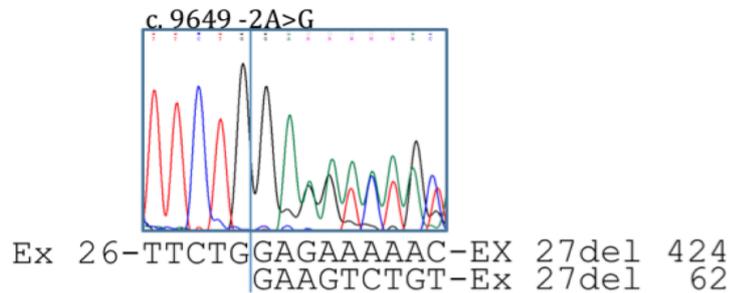
**Figura 71.** Resultados de la RT-PCR del minigen MGBR2 19-27 *wild type* y mutaciones ensayadas en el exón 27. Electroforesis en gel de agarosa (2%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y derivados del mismo con mutaciones en el exón 27.



**Figura 72.** Cuantificación de las isoformas producidas por mutaciones del exón 27 en el minigen MGBR2 19-27. Diagrama de la media aritmética de la cuantificación relativa de cada isoforma detectada en la electroforesis capilar para las mutaciones del exón 27 ensayadas en el MGBR2 19-27.

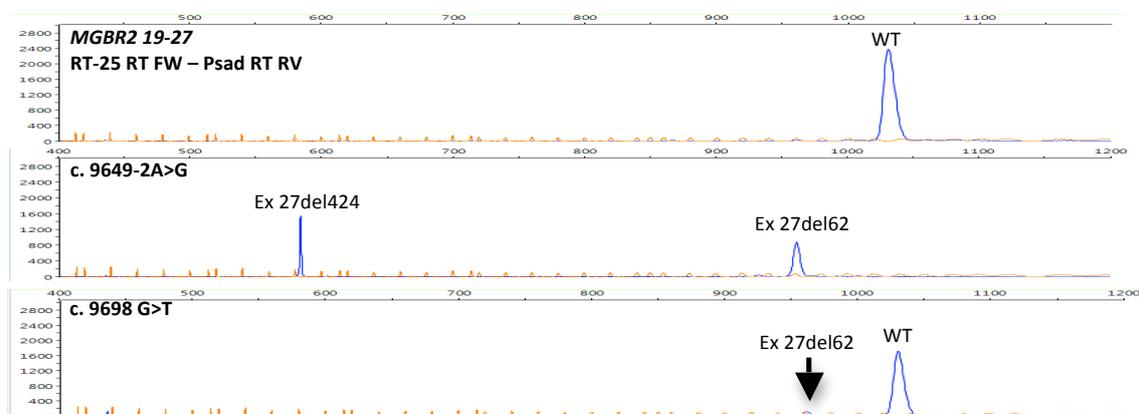
La mutación **c.9649-2A>G** es una variante intrónica registrada en la base de datos BIC. Según el estudio informático, la mutación elimina el sitio aceptor del exón 27 (Tabla 3). El ensayo funcional del MGBR2 19-27 con la mutación c.9649-2A>G mostró la alteración del proceso de *splicing* (Figuras 71, 73 y 75). La cuantificación de las isoformas detectadas mediante electroforesis capilar (Figura 72), indicó que el 72% de los transcritos no incluían los primeros 424 nt del exón 27, debido a que la maquinaria de *splicing* reconocía un sitio aceptor críptico a 424 nt *downstream* del natural. Según las predicciones informáticas, la probabilidad relativa de que este sitio fuera reconocido era de 0,75. El resto de los transcritos no incluían los primeros 62 nt

del exón 27 debido al reconocimiento de un sitio aceptor críptico en esa posición, que no había sido detectado por las utilidades bioinformáticas.



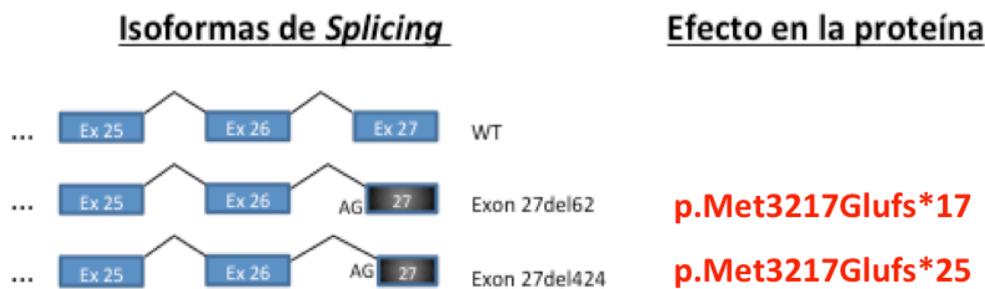
**Figura 73.** Secuencias de los productos RT-PCR procedentes de la mutación c.9649-2A>G. En el cromatograma se puede comprobar la presencia de transcritos que carecen tanto de los primeros 62 como de los 424 nt del exón 27.

La mutación **c.9698G>T** es una variante registrada en la base de datos LOVD y catalogada como *missense*. Según el estudio informático, la mutación crearía dos sitios silenciadores (Tabla 3). En el ensayo funcional con el MGBR2 19-27 con la mutación c.9698G>T se comprobó la existencia de una alteración en el proceso de *splicing* (Figuras 71 y 74). La cuantificación de las isoformas detectadas mediante electroforesis capilar (Figura 72) indicó que, aunque la mayoría de los transcritos son similares al WT, un 4% carecía de los primeros 424 nt del exón 27, debido al reconocimiento de un sitio aceptor críptico. Según las predicciones informáticas, la probabilidad relativa de que este sitio sea reconocido era de 0,75. Además se detectó un porcentaje de transcritos inferior al 1% que no presentaban los primeros 62 nt del exón 27 por el uso de un sitio aceptor críptico en esa posición, que, sin embargo, no había sido detectado por las utilidades bioinformáticas.



**Figura 74.** Electroforesis capilar de los ensayos funcionales de *splicing* de las mutaciones del exón 27 de *BRCA2* (MGBR2 19-27) en células HeLa.

Los ensayos funcionales de *splicing* con minigenes híbridos de mutaciones del exón 27 mostraron la presencia de dos diferentes tipos de transcritos aberrantes (Figura 75). La traducción de todos ellos sería interrumpida por la presencia de un codón de parada prematuro.



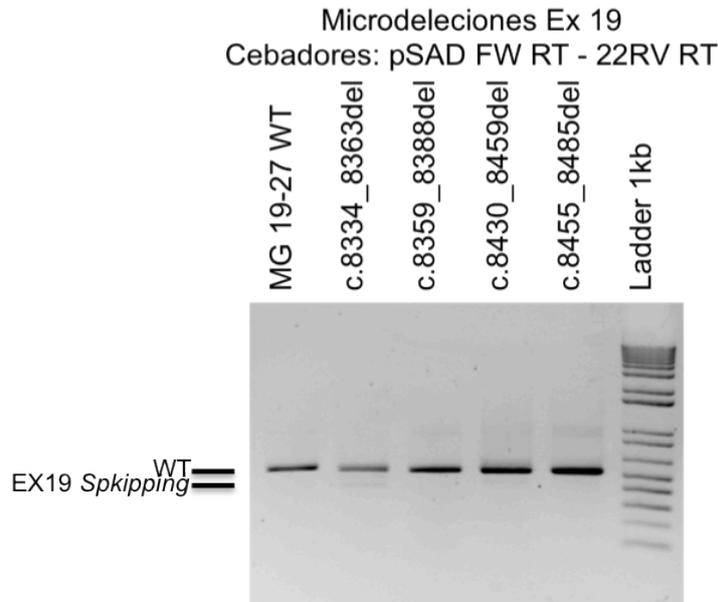
**Figura 75. Isoformas de *splicing* detectadas en el exón 27 del minigen MGBR2 19-27 y efecto predecible en la proteína de acuerdo a la nomenclatura HGVS.** Diagramas de representación del proceso de *splicing* que da lugar a las diferentes isoformas por mutaciones en el exón 27 y el efecto de las mismas en la traducción a proteína.

### **VI.3.b. Mapeo de elementos reguladores de *splicing* mediante microdeleciones solapantes exónicas en el MGBR2 19-27**

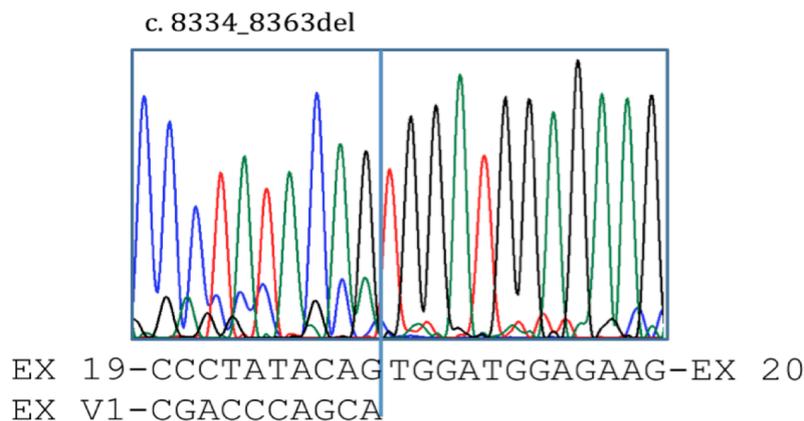
El estudio informático y los ensayos funcionales con minigenes híbridos que portan mutaciones en varios de los exones incluidos en el MGBR2 19-27 dejan patente que determinados sitios de *splicing* naturales (como el sitio aceptor del exón 23) tienen una baja probabilidad de ser identificados por la maquinaria de *splicing*. En estos casos, parece demostrado que la existencia de secuencias reguladoras de *splicing* (como los *enhancers*) facilitan el reconocimiento de los mismos. Con el objetivo concreto de determinar la presencia de SRE se realizaron 34 microdeleciones de 30 nt sobre MGBR2 19-27 de forma independiente, dos en cada extremo de los exones solapantes en 5 nt. Siempre se respetaron los dos primeros y los tres últimos nucleótidos de cada exón, puesto que se trata de nucleótidos muy conservados en los sitios naturales de *splicing*. La excepción es el exón 27, al que sólo se le realizaron dos microdeleciones en el extremo 5'. De todas ellas, sólo tres en tres exones diferentes tuvieron un impacto significativo en el *splicing*.

La microdelección de 30 nt **c.8334\_8363del** es la más próxima al sitio aceptor del exón 19. Según las predicciones bioinformáticas con HSF, la microdelección podría

eliminar una secuencia que fuese reconocida por el factor de *splicing* SC35. En los resultados del ensayo funcional con el MGBR2 19-27 con la delección c.8334\_8363del se pudo comprobar (Figuras 76 y 77) que parte de los transcritos no incluían el exón 19, por lo que el resultado encajaría con el predicho por el análisis informático.

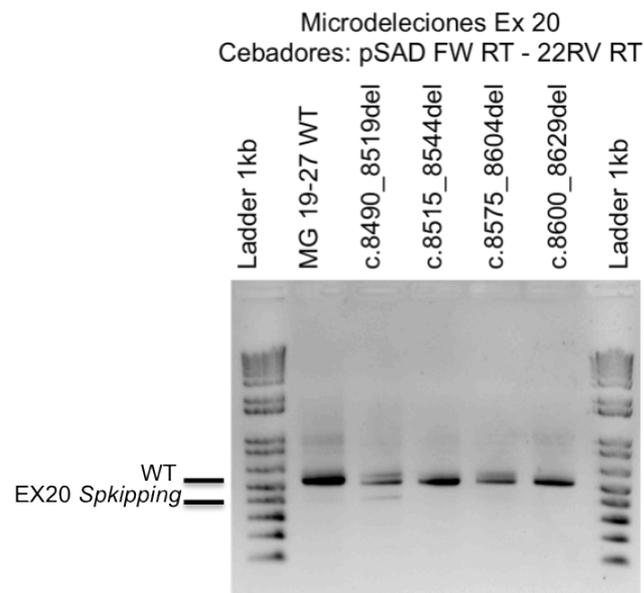


**Figura 76. Resultados de la RT-PCR del minigen MGBR2 19-20 *wild type* y las microdelecciones realizadas en el exón 19.** Resultado de electroforesis en gel de agarosa (1,5%) del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y los derivados del mismo con microdelecciones en el exón 19. La microdelección c.8334\_8363del alteró el *splicing* causando que parte de sus transcritos perdieran el exón 19.

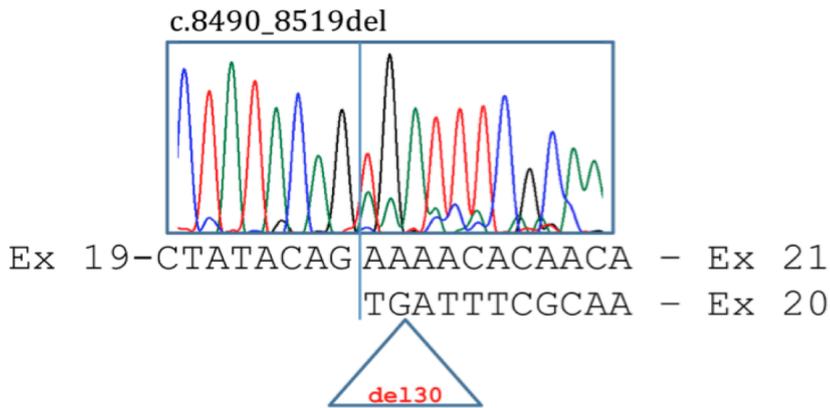


**Figura 77. Secuencia de los productos RT-PCR generados por la delección c.8334\_8363del en el exón 19 del MGBR2 19-27.** Resultado de la secuenciación de los productos RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 al que se le ha realizado la microdelección c.8334\_8363del. La secuencia muestra la carencia del exón 19 en parte de los transcritos maduros, mientras que el resto son similares al WT.

La microdelección **c.8490\_8519del** es la más próxima al sitio aceptor del exón 20. Según las predicciones bioinformáticas con HSF la microdelección podría eliminar diferentes secuencias que podrían ser reconocidas por los factores de *splicing* SRp40, SF2/ASF y SC35. En los resultados del ensayo funcional del MGBR2 19-27 con la deleción c.8490\_8519del se comprobó (Figuras 78 y 79) que parte de los transcritos carecían de exón 20. Según las predicciones informáticas el sitio aceptor del exón 20 sería débil (baja similitud con la secuencia consenso) con lo cual es probable que el reconocimiento del sitio aceptor del exón 20 sea promovido por una secuencia de unión a una proteína del tipo SR potenciadora de *splicing*.

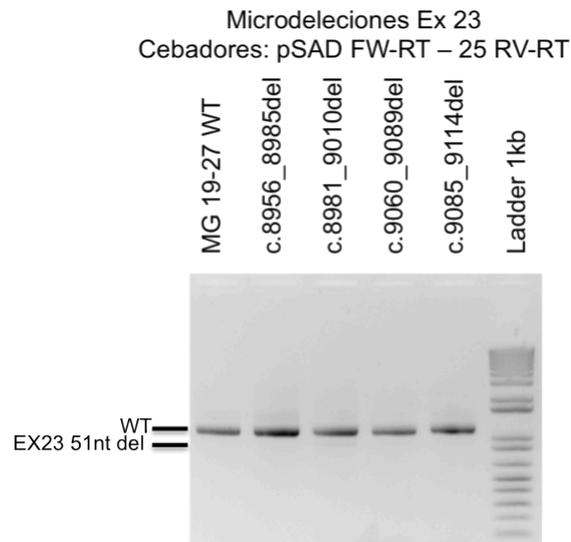


**Figura 78. Resultados de la RT-PCR del minigen MGBR2 19-27 *wild type* y las microdeleciones realizadas en el exón 20 del mismo.** Resultado de electroforesis en gel de agarosa (1,5%) del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y los derivados con diferentes microdeleciones en el exón 19. La microdelección c.8490\_8519del alteró el *splicing* causando que parte de sus transcritos carecieran del exón 20.

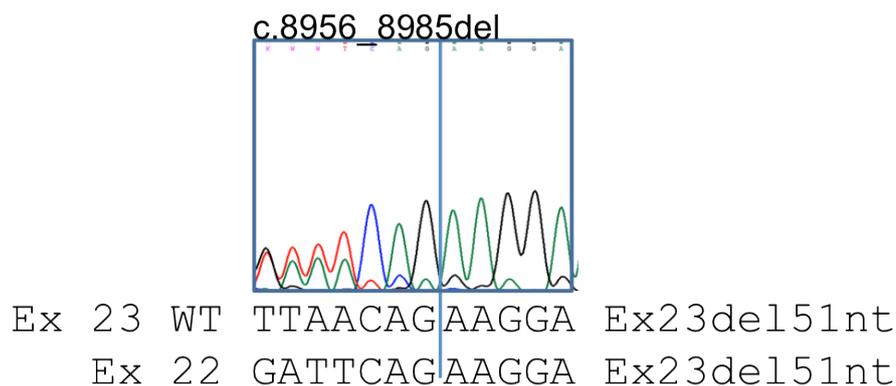


**Figura 79. Secuencias de los productos RT-PCR procedentes de la microdelección c.8490\_8519del en el exón 20 del MGBR2 19-27.** Resultado de la secuenciación de los productos RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 al que se le ha realizado la microdelección c.8490\_8519del. La secuencia muestra la carencia del exón 20 en parte los transcritos maduros, mientras que el resto son similares al WT.

La microdelección **c.8956\_8985del** es la más próxima al sitio aceptor del exón 23. Según las predicciones bioinformáticas con HSF la microdelección podría eliminar diferentes secuencias de unión de los factores de *splicing* SRp55 y SC35 (Figura 20). El ensayo funcional del MGBR2 19-27 con la delección c.8956\_8985del se detectó por diferentes métodos (Figuras 80 y 81). Parte de los transcritos habían perdido los primeros 51 nt del exón 23 porque era reconocido un sitio críptico más fuerte que el natural 51 nt *downstream*. Según las predicciones informáticas, la probabilidad relativa de que el sitio aceptor del exón 23 fuese reconocido era baja (0,51) y en función de los resultados de la microdelección c.8490\_8519del, parecía probable la existencia de una secuencia de unión a una proteína del tipo SR potenciadora de *splicing* que facilitara el reconocimiento del sitio aceptor del exón 23.



**Figura 80. Resultados de la RT-PCR del minigen MGBR2 19-20 *wild type* y las microdelecciones realizadas en el exón 23.** Electroforesis en gel de agarosa (1,5%) del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con el MGBR2 19-27 y los derivados del mismo a los que se le han realizado diferentes microdelecciones en el exón 23. La microdelección c.8956\_8985del alteró el *splicing* causando que parte de sus transcritos pierdan los primeros 51 nt del exón 23.

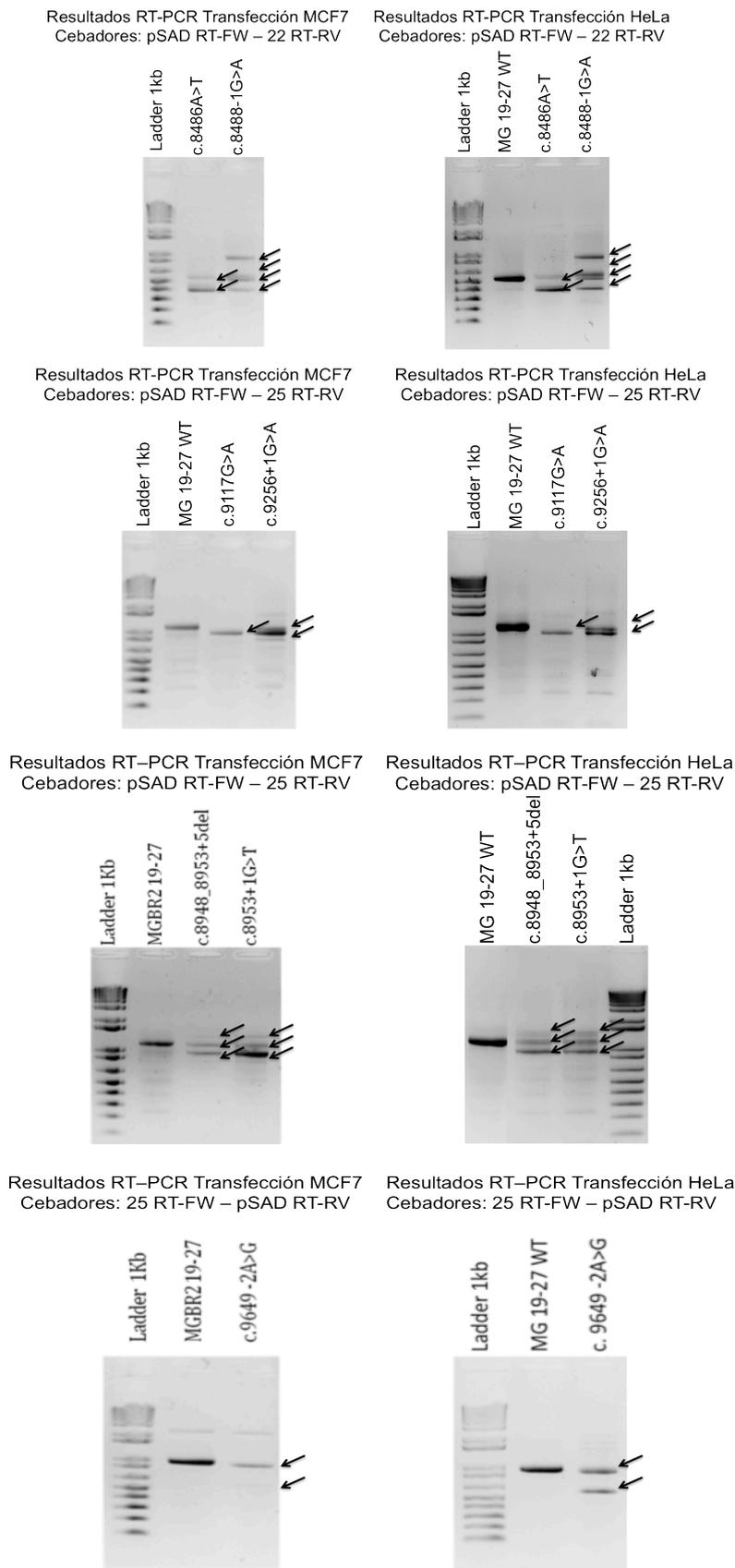


**Figura 81. Secuencias de los productos RT-PCR procedentes de la microdelección c.8956\_8985del en el exón 23 de MGBR2 19-27.** La secuencia muestra *skipping* del exón 23 en parte los transcritos maduros, mientras que el resto son similares al WT.

### VI.3.c. Comparación entre células HeLa y MCF-7

Los minigenes derivados del MGBR2 19-27 con las variantes c.8486A>C, c.8488-1G>A, c.9117G>A, c.9256G>T, c.8948\_8953+5del, c.8953+1G>T y c.9649-2A>G fueron transfectados tanto en células HeLa como en MCF-7. Las células MCF-7 proceden de epitelio mamario, tejido diana de la enfermedad en estudio, por lo que cabe pensar que sería el tipo celular adecuado para realizar los ensayos funcionales de *splicing* con minigenes híbridos, de cara a descartar los eventos de *splicing* alternativo específicos de tejido. Sin embargo, estas células requieren la adaptación de los protocolos específicos de cultivo y transfección porque su cinética de crecimiento es diferente al

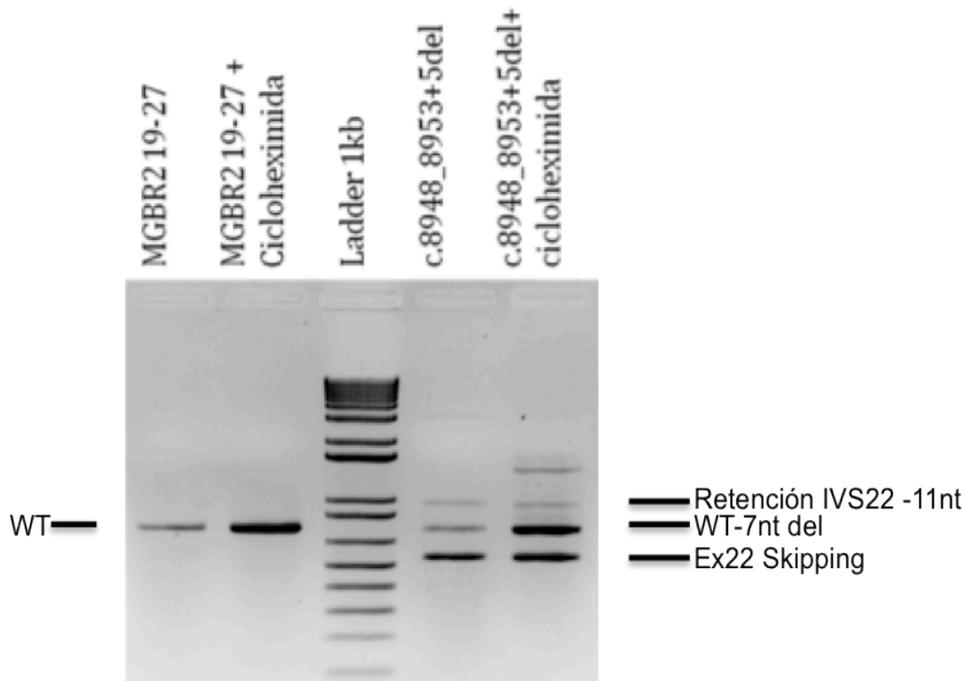
de las células HeLa. Para establecer una comparación entre los resultados procedentes de ambas líneas celulares se han comparado los resultados obtenidos de la transfección en células HeLa y MCF-7 de siete variantes con impacto en el *splicing*. Los resultados obtenidos mediante electroforesis en gel de agarosa de las RT-PCR de cada minigen ensayado muestran que no existen diferencias significativas entre las isoformas detectadas.



**Figura 82. RT-PCR de mutaciones en el MGBR2 19-27 transfectadas en células MCF-7.** Electroforesis en gel de agarosa (1,5%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células MCF-7 transfectadas con los minigenes derivados del MGBR2 19-27 con las mutaciones: c.8486A>C, c.8488-1G>A, c.9117G>A, c.9256G>T, c.8948\_8953+5del, c.8953+1G>T y c.9649-2A>G.

#### VI.3.d. Inhibición del mecanismo de *nonsense mediated mRNA decay*

El *nonsense mediated mRNA decay* (NMD) es un mecanismo celular que se encarga de degradar los transcritos que contienen codones prematuros de parada y que, por lo tanto, pueden dar lugar a proteínas truncadas. La eficiencia de este proceso depende del tipo celular y se ha comprobado que en células HeLa es muy alta (Linde *et al.*, 2007). Los resultados obtenidos hasta al momento, mostraban las isoformas detectadas procedentes de los minigenes una vez que ha actuado el NMD. Para observar el efecto del NMD sobre los transcritos, se realizaron transfecciones en células HeLa con el MGBR2 19-27 y un derivado del mismo con la mutación c.8948\_8953+5del del sitio donador del exón 22 a las que se les añadió cicloheximida, un inhibidor del NMD.



**Figura 83.** Efectos del inhibidor del NMD cicloheximida sobre los transcritos del minigen MGBR2 19-27 *wild type* y con la variante c.8948\_8953+5del. Electroforesis en gel de agarosa (1,5%) que muestra el resultado del ensayo RT-PCR de ARN procedente de células HeLa transfectadas con y sin cicloheximida con el MGBR2 19-27 y un derivado que contiene la mutación c.8948\_8953+5del.

Los resultados obtenidos mediante electroforesis en gel de agarosa de los amplificadores mediante RT-PCR del ARN producido por cada minigen ensayado con y sin cicloheximida (Figura 83) muestra que la inhibición del NMD está actuando en aquellos transcritos que producen un codón de parada prematuro. Por ejemplo, se pudo apreciar un incremento en la cantidad de transcritos en la isoforma *wild type* para el

MGBR2 19-27 y *wild type* del 7nt para la mutación c.8948\_8953+5del una vez inhibido el NMD. También se detectó una isoforma minoritaria no detectada con anterioridad en el minigen que contiene la mutación c.8948\_8953+5del. Probablemente se trate de una retención de intrón más, aparte del 22. La razón por la que se producen estos resultados es porque los exones incluidos en el MGBR2 19-27 no están en fase con el exón 1 del vector de acuerdo a como estarían en el contexto genómico *in vivo*. Para reproducir un contexto similar sería necesario que cada isoforma fuese ensayada teniendo en cuenta el efecto diferencial de actuación del NMD en función del tipo celular y por triplicado en minigenes que contengan +1 y +2 nt en el exón V1 del vector para mantener la fase de lectura.



# DISCUSIÓN



## I. Correlación entre *splicing* y enfermedad en las variantes detectadas en pacientes del programa de prevención del CMOH

El *splicing* tiene un papel crucial en la generación de la complejidad biológica, puesto que el procesamiento del pre-ARNm es una etapa esencial de la expresión génica. Como se ha mostrado anteriormente, el *splicing* requiere una compleja red de interacciones que permiten una regulación precisa, de modo que cualquier alteración de los factores reguladores implicados (secuencias, proteínas, estructura secundaria, etc) puede causar o modular la severidad de una determinada enfermedad (Buratti *et al.*, 2006; Cartegni *et al.*, 2002; Wang y Cooper, 2007).

La mutación intrónica **c.8488-1G>A** detectada en una paciente procedente del programa de Consejo Genético en CMOH de Castilla y León no se encuentra descrita en el BIC y por lo tanto nada se sabía ni de su posible alteración de la proteína ni de su posible efecto patogénico. Al parecer, la portadora había padecido cáncer de mama a los 66 y 74 años respectivamente así como de ovario a los 76. La historia familiar de la paciente aportaba datos relativamente escasos, sólo se sabía que una hermana fallecida había padecido cáncer de mama a los 51 años. Por la posición de la variante encontrada, se pudo deducir que la mutación c.8488-1G>A afectaba a una de las señales básicas para el *splicing*, el sitio aceptor del exón 20. Este tipo de alteraciones ha sido ampliamente reportada en la literatura, como por ejemplo:

- La variante c.5153-1G>A que altera el sitio aceptor del exón 19 del gen *BRCA1* e induce una deleción de 1 nt en el transcrito maduro. Se trata de una mutación fundadora en Castilla y León y de relevancia en consejo genético de cáncer de mama hereditario (Infante *et al.*, 2010).
- La mutación c.439-2A>C en el sitio aceptor del exón 5 del gen *BSCL2* (que codifica la proteína seipina) provoca la aparición de dos tipos de transcritos aberrantes: uno con la pérdida completa del exón 5 y otro con una pérdida parcial, ya que un sitio aceptor crítico es reconocido dentro del exón 5. El resultado fenotípico es la aparición de una lipodistrofia generalizada (Jeninga *et al.*, 2012).
- La eliminación del sitio aceptor del exón 3 (c.590-2A>G) del gen *FAM20A*, responsable de la amelogénesis imperfecta hipoplásica, provoca que este exón no sea incluido en los transcritos maduros (Cho *et al.*, 2011).

El análisis bioinformático con la aplicación web NNSplice no detectó el sitio aceptor del exón 20 de *BRCA2*. Este hecho se atribuye a que las secuencias consenso para los sitios de *splicing* son degeneradas, por lo que la probabilidad relativa de que el sitio aceptor del exón 20 sea reconocido o lo que es lo mismo, su similitud con la secuencia consenso de los sitios aceptores, es muy baja. Sin embargo, esto no parece un obstáculo para que funcione como tal, debido a que el proceso de reconocimiento de los exones sigue un modelo combinatorial, en el que una suma de factores (fortaleza de sitio, presencia de potenciadores de *splicing* o *enhancers*, contexto genómico, estructura secundaria, cinética de la transcripción, etc) facilitaría la detección precisa de los sitios donador y aceptor.

El resultado del ensayo funcional de RT-PCR en linfocitos corroboró la sospecha previa de que la mutación c.8488-1G>A afectaba al *splicing*. Se pudo comprobar la existencia de tres tipos de transcritos: uno canónico procedente del alelo *wild type* correspondiente y dos aberrantes generados por la mutación. El primero incluye una retención del intrón 19 y el segundo (detectado mediante secuenciación) había perdido los primeros 12 nt del exón 20, debido a que un sitio críptico aceptor 12 nt *downstream* del natural es reconocido por la maquinaria de *splicing*.

La mutación c.8488-1G>A también fue analizada mediante ensayos funcionales de *splicing* con los MGBR2 19-20 en pSPL3 y MGBR2 19-27 en pSAD que portaban dicha variante. Los resultados muestran la presencia de las dos isoformas aberrantes detectadas en el ensayo de RT-PCR de linfocitos, pero también la presencia de un tercer transcrito mayoritario en MGBR2 19-20 con delección del exón 20.

La diferencia entre los resultados de los ensayos de linfocitos y de minigenes podría explicarse por la falta de contexto genómico en los minigenes. Es posible que exones anteriores o posteriores al 19 estén involucrados en el reconocimiento y eficiencia del *splicing* de los exones 19 y 20 (Buratti *et al.*, 2006). De hecho, el MGBR2 19-27 prácticamente restaura el patrón de *splicing* de linfocitos (Figura 38-Resultados), apoyando la hipótesis de la necesidad los exones flanqueantes para un reconocimiento preciso de cada exón. Por tanto, sería recomendable la construcción de minigenes con el mayor número de exones de cara a reproducir los patrones de *splicing* fisiológicos y patológicos.

Por otra parte, la alteración del *splicing* causado por la microdelección de 30 nt c.8490\_8519del en el extremo 5' del exón 20 confirma la existencia de posibles elementos potenciadores de *splicing* que pueden facilitar el reconocimiento, por parte del espliceosoma, del sitio aceptor del exón 20. Sin embargo, el resultado no es del todo similar al que se produce con la alteración del sitio aceptor. En concreto, no se ha detectado la presencia de isoformas que hayan incluido el intrón 19. Los eventos de retención de intrón son un fenómeno poco frecuente en el *splicing* alternativo fisiológico porque generalmente están asociados a la introducción de codones de parada prematuros de la traducción (Kim *et al.*, 2008). Más aún, la frecuencia de este evento se duplica en células cancerosas, sugiriendo la inactivación de genes represores esenciales implicados en la progresión tumoral. Lo habitual son los eventos de pérdida de exón y selección alternativa de sitios de *splicing*. Se ha podido comprobar que estos eventos son los que contribuyen en mayor proporción a la complejidad del proteoma humano (Ast, 2004).

En cuanto a la arquitectura genómica para el reconocimiento de intrones/exones, el tamaño del intrón 19 parece saltarse la norma general de los genomas de eucariotas superiores, según la cual los exones cortos suelen estar rodeados de intrones largos y el *splicing* se produciría a través del modelo de definición de exón (Sakharkar *et al.*, 2005). En este caso, al tratarse de un intrón pequeño (398 nt), podría ocurrir lo contrario y que el *splicing* se produzca a través del modelo de reconocimiento de intrón, aunque se ha reportado que éste no funciona cuando el intrón supera los 250 nt (Fox-Walsh, 2008). En estos casos, también se ha descrito que las diferencias entre el contenido de GC del intrón respecto a los exones flanqueantes determinan también el reconocimiento del exón (Amit *et al.*, 2012). Sin embargo, no se hallaron diferencias significativas respecto al porcentaje de GC entre los exones 19 y 20 y el intrón 19 (datos no mostrados).

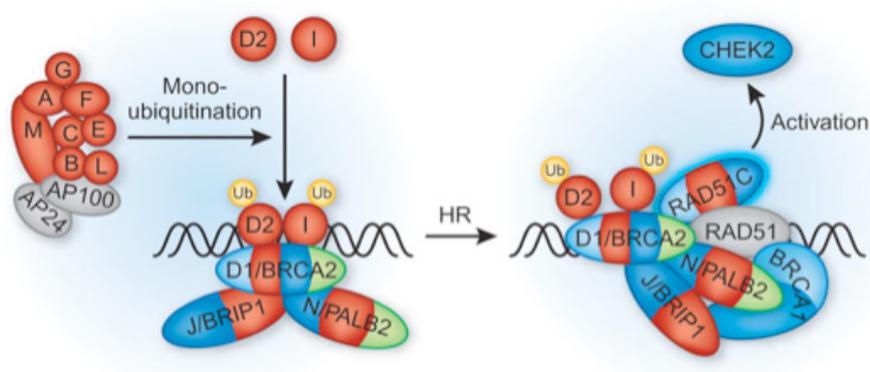
En el intrón 19 se identificó la presencia de una secuencia Alu *Sq* en orientación contraria. Los análisis bioinformáticos revelaron que esta secuencia Alu contenía múltiples putativos sitios de unión a proteínas SR que podrían estar implicadas en su reconocimiento. Además, gracias a la comparación evolutiva entre las secuencias del intrón 19 y de diferentes especies se pudo constatar la existencia de dos secuencias muy conservadas, flanqueantes a la inserción Alu. Con el objetivo de estudiar la

implicación de dichas secuencias en el reconocimiento del exón 20, se realizaron ensayos funcionales con minigenes procedentes del MGBR2 19-20 a los que se les delecionaron tanto las secuencias conservadas como la secuencia Alu de forma individual, pero no se observó ningún efecto sobre el *splicing*. Como se ha podido comprobar en estudios previos (Sanz *et al.*, 2010) los elementos potenciadores de *splicing* o *enhancers* suelen actuar de forma cooperativa. Es por ello por lo que se decidió realizar diferentes combinaciones de las tres delecciones en diferentes minigenes. El ensayo funcional con dos de los minigenes derivados del MGBR2 19-20 sobre los que se realizó la combinación de delecciones c.8487+31\_8487+42del-c.8487+54\_8487+355del y c.8487+31\_8487+42del-c.8487+54\_8487+355del-c.8488-49\_8488-44del afectaron al *splicing* produciendo un porcentaje pequeño de transcritos que perdieron el exón 19 (3%-7%). El resultado sugiere que el intrón 19 contendría secuencias que podrían ser necesarias para el reconocimiento del exón 19 y no del exón 20.

Los transcritos aberrantes derivados de la variante c.8488-1G>A con *skipping* del exón 20 o retención del intrón 19 provocan la aparición de un codón de parada prematuro en la traducción a proteína (p.Trp2830Lysfs\*13 y p.Trp2830Valfs\*1) que se vería afectados por la degradación del mecanismo de NMD. La excepción la forman los transcritos (<20%) que han perdido de los primeros 12 nt del exón 20 que sí mantiene la fase de lectura. La proteína resultante procedente de estos transcritos carece de 4 aminoácidos (p.Trp2830\_Lys2833del; W-M-E-K) pertenecientes al dominio de unión a ADN de BRCA2. Más aún, el Triptófano 2830 y el Glutámico 2832 están estrictamente conservados en mamíferos (IARC BRCA2 alignment, [http://agvgd.iarc.fr/BRCA2\\_Align.htm](http://agvgd.iarc.fr/BRCA2_Align.htm)) sugiriendo que juegan un papel funcional relevante.

Todos estos datos en conjunto sugieren que la variante c.8488-1G>A es una mutación patogénica o deletérea. Asimismo, en un trabajo previo se describió la presencia de la variante c.8488-1G>A en los dos alelos del gen *BRCA2* de un paciente con Anemia de Fanconi tipo D1 (Howlett *et al.*, 2002), apoyando una contribución importante al riesgo de padecer CMOH. En este trabajo, detectaron la presencia de transcritos con la delección de los primeros 12 nt del exón. La presencia del fenotipo de Anemia de Fanconi es debida a la inactivación de las dos copias de *BRCA2*. Como se puede apreciar en la Figura 1, la mayoría de proteínas codificadas por los genes

implicados en la Anemia de Fanconi forma un complejo para la monoubiquitinación de las proteínas FANCD2 y FANCI que indica focos de reparación en respuesta al daño en el ADN. A su vez, estas proteínas necesitan interactuar y formar un complejo multiproteico con proteínas codificadas por genes implicados en la reparación de roturas de ADN de doble cadena que, como hemos visto en la introducción, están relacionados tanto con la Anemia de Fanconi con la predisposición a padecer CMOH (genes *BRIP1*, *PALB2*, *RAD51C*). De esta forma, los síndromes de Anemia de Fanconi y CMOH están estrechamente relacionados (Figura 1; Levy-Lahad, 2010).



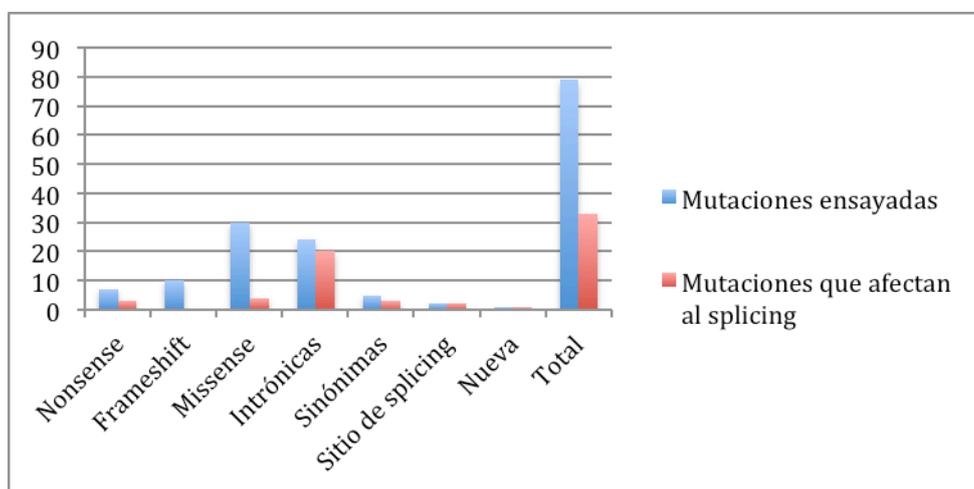
**Figura 1. Proteínas implicadas en Anemia de Fanconi y su relación con el CMOH.** Las proteínas relacionadas con la Anemia de Fanconi se indican con la letra del grupo de complementación (por ejemplo, A representa FANCA, B representa FANCB, C representa FANCG, D1 representa FANCD1, D2 representa FANCD2, E representa FANCD3, F representa FANCD4, G representa FANCG, I representa FANCI, M representa FANCM y FAAP24 (AP24 denota el complejo central FA)). Las proteínas son de color para indicar que son codificadas por un gen en el que las mutaciones se han asociado a la Anemia de Fanconi (rojo), los cánceres sólidos infantiles (verde), el cáncer de mama (azul oscuro). Algunas proteínas están marcadas con múltiples colores que representan asociaciones múltiples. No todas las interacciones de la proteína se muestran y la proximidad no implica una interacción directa.

La mutación **c.9026\_9030del** es una de las variantes más prevalentes en la población española. En Castilla y León se había detectado en el 7,6% de las familias positivas con mutaciones en el gen *BRCA2*. De acuerdo al software HSF, esta variante eliminaba 2 “*Exon Identity Elements*” y creaba 2 “*Intron Identity Elements*” de importancia funcional desconocida. Además, estudios previos de nuestro laboratorio indicaban que las mutaciones *frameshift* (deleciones o inserciones de nucleótidos) también podían afectar al proceso de *splicing*, como por ejemplo la c.470\_474del (Sanz *et al.*, 2010), lo cual propició la realización de ensayos funcionales de *splicing*. Los resultados tanto de RT-PCR de linfocitos como con minigenes no detectaron la presencia de transcritos aberrantes. Aunque la variante c.9026\_9030del no afecte al *splicing*, se trata de una mutación *frameshift* patogénica, ya que implica la aparición de

un codón de parada prematuro en la traducción (p.Tyr3009Serfs\*9). Por lo tanto, podemos concluir que es una variante de alto riesgo que confiere predisposición a CMOH.

## II. Cambio en el efecto de una mutación por alteración del *splicing*

En el presente trabajo se han analizado en relación a su efecto sobre el *splicing* mutaciones clasificadas en todos los tipos existentes: *frameshift*, *nonsense*, *missense*, intrónicas, sinónimas y de sitios de *splicing*. En la Figura 2, se muestra que salvo por las mutaciones seleccionadas clasificadas como *frameshift*, se ha podido constatar que más de una mutación de cada tipo ha resultado afectar al proceso de *splicing* y deberían ser reclasificadas.



**Figura 2. Diagrama de representación del número de mutaciones seleccionadas totales de cada tipo respecto a las que han resultado positivas por alterar el *splicing*.** Las Barras azules representan el número total de mutaciones seleccionadas para realizar ensayos funcionales de *splicing* y que están descritas en las bases de datos internacionales. Las rojas, el número de mutaciones positivas de las seleccionadas de cada tipo.

### II.1. Mutaciones frameshift y nonsense

Las variantes *nonsense* y *frameshift* son clasificadas directamente como mutaciones causantes de enfermedad sin necesidad de estudios adicionales por su efecto predecible en la proteína (truncamiento). En este trabajo se han analizado siete mutaciones *nonsense* registradas en las base de datos BIC y catalogadas de importancia clínica mediante ensayos funcionales con minigenes híbridos. Sorprendentemente, tres de ellas han alterado el proceso de *splicing* de forma parcial: la mutación c.8969G>A (p.Trp2990Ter) y c.9076C>T (p.Gln3026Ter) en el exón 23 y la c.9256G>T (p.Gly3086Ter) en el exón 24. Teniendo en cuenta que la única isoforma

observada que no produce un codón prematuro de parada en los ensayos funcionales de las variantes c.9076C>T y c.9256G>T es la que carece de los exones 23 y 24 y que puede estar relacionada con reacciones anómalas entre los sitios de *splicing* de los exones constitutivos del vector en MGBR2 23-24 por su escaso contexto genómico, no debería tenerse en consideración. Una prueba de que este suceso podría estar ocurriendo es que no se ha encontrado esta isoforma en los ensayos de las variantes c.9117G>A y c.9256+1G>A reproducidas en el MGBR2 19-27. Aunque las variantes c.9076C>T y c.9256G>T alteren el proceso de *splicing* el resto de los transcritos portan codones de parada prematuros, por lo que ambas pueden ser clasificadas como patogénicas. Sin embargo, la mutación c.8969G>A produce un 10% de transcritos en los que no han sido incorporados los primeros 51 nt del exón 23, en este caso la isoforma queda libre de la mutación y además está en fase de lectura (Val2985\_Thr3001del). Podríamos suponer para este caso un cambio en la penetrancia de la mutación. El efecto que puede tener esta isoforma respecto a la funcionalidad de la proteína puede ser diverso dependiendo de la importancia de los amino ácidos delecionados y debe ser confirmado experimentalmente mediante ensayos funcionales de proteína.

La conclusión más importante es que mutaciones *a priori* truncantes de proteína pueden ser reclasificadas como mutaciones de *splicing*, lo cual sugiere que la patogenicidad de cada variante debe ser interpretada cuidadosamente de forma individual. Asimismo, podría tener una repercusión en la penetrancia de una determinada mutación, sobre todo si provoca eventos de *splicing* que mantienen la pauta de lectura original. En conclusión, las mutaciones *nonsense* y *frameshift* con efecto parcial sobre el *splicing* van a tener un doble mecanismo etiopatogénico: truncamiento de proteína y alteración del *splicing*.

## II.2. Mutaciones *missense*

Se ha ensayado mediante análisis funcional con minigenes híbridos, un total de 30 mutaciones *missense* registradas en bases de datos internacionales. Cuatro de ellas alteraban el *splicing*: c.8378G>A, c.8486A>T (100%), c.9006A>T (impacto moderado) y c.9698G>T. La proporción de transcritos anómalos a causa de las mutaciones c.8378G>A y c.9698G>T es inferior al 15%, mientras que en la variante c.9006A>T

supera el 20%. Por último, la totalidad de los transcritos el efecto de la mutación c.8486A>T el *splicing* era completo. Solamente las isoformas aberrantes procedentes del cambio c.9698G>T provocan la aparición de un codón de parada prematuro (p.Met3217Glufs\*17) que darían lugar a una proteína truncada Sin embargo el mecanismo de NMD supuestamente no las degradaría porque la alteración se produce en el último exón del gen *BRCA2* donde se localiza el codón de parada canónico. El resto de transcritos aberrantes procedentes de mutaciones *missense* mantiene la fase de lectura con deleciones de 156 nt (Exón 19 *skippnig* - p.Ile2778\_Gln2829del - 52 aa) y 51 nt (Exón 23del51 nt – p.Val2985\_Thr3001del - 17 aa). Los aminoácidos codificados por el exón 19 están altamente conservados y forman parte del dominio OB2 de *BRCA2*, cuya función es la de interactuar con cadenas sencillas de ADN durante la reparación de roturas de doble hebra, por lo que podrían resultar esenciales para la función de la proteína. A pesar de la conservación evolutiva de los 17 aa perdidos del exón 23 y de que éstos se asocian con un función similar, dichos aminoácidos no pertenecen a un dominio concreto. El hecho de que estas mutaciones afecten al *splicing* debería replantear su clasificación como alelos de riesgo con importancia clínica.

La variante c.8486A>T debe ser reclasificada como mutación patogénica de *splicing* puesto que cumple con los indicadores biológicos de patogenicidad: alta proporción de transcritos afectados y pérdida de un dominio esencial de *BRCA2* (*DNA Binding Domain*). Sin embargo, aquellas variantes con impacto parcial podrían estar asociadas a un riesgo moderado o bajo de padecer CMOH. También, se han descrito variantes *missense*, tales como c.5123C>A (p.A1708E) de *BRCA1*, con un mecanismo patogénico doble: alteración de la funcionalidad de la proteína (Vallon-Christerson *et al* 2001) e impacto parcial en el *splicing* (Millevoi *et al* 2009; Sanz *et al* 2010).

### II.3. Mutaciones sinónimas

Las variantes sinónimas, que no implican un cambio de amino ácido en la proteína, típicamente son clasificadas como variantes neutras o polimorfismos. Sin embargo, numerosos autores han reportado efectos deletéreos de dichas variantes sobre el *splicing* mediante la alteración de elementos reguladores clave (Sanz *et al.*, 2010; Vidal *et al.*, 2009; Sauna y Kimchi-Sarfaty, 2011).

Cinco mutaciones **sinónimas** fueron ensayadas mediante minigenes híbridos de las que tres afectaron al proceso de *splicing*: c.8487G>A, c.8754G>A, c.9117G>A. En los tres casos las mutaciones se sitúan en el último nucleótido de su exón correspondiente y lindando con los nucleótidos altamente conservados GT del sitio donador del exón 19, 21 y 24 respectivamente. En todos los casos, los porcentajes de transcritos aberrantes superan el 80%. Las mutaciones c.8754G>A y c.9117G>A producen en su mayoría transcritos con codones de parada prematuros (p.Glu2919Valfs\*4, p.Val2985Glyfs\*3). La variante c.8487G>A provoca la aparición de isoformas que carecen del exón 19 y, aunque mantiene la fase de lectura, 52 aa no son incorporados en la traducción a proteína (p.Ile2778\_Gln2829del). Estos pueden resultar esenciales en la función de *BRCA2* ya que muchos aminoácidos están conservados evolutivamente y pertenecen al dominio OB2. La variante c.8487G>A es la única de las tres sinónimas positivas que está catalogada como de significado clínico desconocido. Los resultados probaron la existencia de un efecto funcional, y por lo tanto, podría ser catalogada como mutación patogénica.

#### II.4. Mutaciones de sitios de *splicing*

Las dos mutaciones catalogadas como de sitios de *splicing* (no estrictamente intrónicas) seleccionadas de la base de datos BIC y analizadas mediante ensayos funcionales con minigenes híbridos fueron c.9248\_9256+7del y c.8948\_8955+5del. En ambos casos se producen diversos transcritos con PTCs (p.Val3040Aspfs\*18, p.Val3072Asp\*18 y p.Gly2919Valfs\*8, p.Asp2983leufs\*3 y p.Asp2983valfs\*34 respectivamente) en los que siempre aparece un codón prematuro de parada. Por lo tanto, dichos transcritos podrán ser degradados por el NMD y aquellos que puedan “escapar” a este proceso producirán proteínas aberrantes escasamente funcionales. Las dos variantes han sido clasificadas como de importancia clínica.

#### II.5. Mutaciones intrónicas

Veinticuatro mutaciones intrónicas (sin contar con la nueva detectada c.8488-1G>A) registradas en la base de datos BIC, fueron seleccionadas para realizar ensayos funcionales de *splicing* con minigenes híbridos en base a los resultados bioinformáticos. Únicamente cuatro mutaciones ensayadas no afectaron al *splicing*:

c.9257-18C>A, c.9501+4A>G, c.9501+9A>C en el exón 25 y c.9649-5A>G en el exón 27. Clasificadas como intrónicas, las mutaciones en las posiciones conservadas de los extremos (-1, -2, +1 y +2) de los intrones correspondientes con el sitio aceptor y donador (nucleótidos AG-GT) son clasificadas directamente como de importancia clínica. Todas las variantes estudiadas en estas posiciones c.8487+1G>A, c.8488-2A>G, c.9117+1G>T, c.9117+1G>A, c.9118-2A>G, c.9256+1G>A, c.8755-1G>A, c.8953+1G>T, c.9257-1G>C, c.9502-2A>C, c. 9649-2A>G provocaban una alteración del *splicing* en la totalidad de los transcritos. Éstas fueron clasificadas como patogénicas o deletéreas. No obstante la penetrancia y el riesgo asociados a cada una de ellas podrían ser diferentes en función de el porcentaje y tipo de isoformas aberrantes producidas.

Por ejemplo, en las variantes c.8487+1G>A y c.9257-1G>C, aunque todos sus transcritos son aberrantes mantienen la pauta de lectura y darían lugar a proteínas que carecen de 52 aa (p.Ile2778\_Gln2829del) y 9 aa (p.[Gly3086Asp;Leu3087\_Asp3095del]) respectivamente. Como se ha mencionado anteriormente, la delección el exón 19 que codifica 52 aa que forman parte del dominio OB2 puede tener repercusión funcional en BRCA2. Por otro lado, 5 (VYLS D) de los 9 aa delecionados están altamente conservados en la evolución y formarían parte del dominio OB3 de unión a ADN monocatenario, esencial para la función de BRCA2. En la mutación c.9502-2A>C aunque todos sus transcritos son aberrantes, el 30% que carecen del exón 26 también conservan la pauta de lectura, pudiendo dar lugar a proteínas con delección de 49 aa (p.Asn3168\_Leu3216del). Estos transcritos podrían producir una proteína BRCA2 con sus características funcionales alteradas. En este caso, a priori, dichos aminoácidos no están conservados evolutivamente ni se relacionan con dominios funcionales. Sería procedente la realización de experimentos funcionales de proteína adicionales, puesto que pueden tener un papel importante en la estructura de BRCA2.

Otras dos mutaciones c.9501+3A>T y c.9502-12T>G próximas a los sitios de *splicing* producían transcritos aberrantes en menos de un 15% del total. En el transcrito de la variante c.9501+3A>T no mantiene la pauta de lectura y se produce la aparición de un codón de parada prematuro (p.Gly3085Glu\*3), mientras que en el transcrito inducido por c.9502-12T>G (p.Asn3168\_Leu3216del) guarda la ORF. La importancia clínica de estas variantes es desconocida pero podrían contribuir a la susceptibilidad a CMOH como alelo de baja-media predisposición. De interés especial

resulta la variante c.9501+3A>T puesto que es la que más registros presenta, hasta 15, de las mutaciones ensayadas que alteraron el *splicing* en las bases de datos internacionales.

Seis variantes intrónicas ensayadas y registradas en la base de datos BIC en posiciones lindantes con los nucleótidos conservados de los sitios aceptores y donadores (-3, +4 y +5) alteraron el *splicing* y la gran mayoría de sus transcritos fueron aberrantes.

Las mutaciones c.8954-3C>G, c.8754+4A>G, c.8754+5G>A y c.8754+5G>T produjeron transcritos aberrantes con codones de parada prematuros (p.Val2985Glyfs\*4 y p.Glu2919Valfs\*4). Curiosamente, la base de datos BIC las califica como de efecto clínico desconocido. Sin embargo, la variante c.8754G>A produce el mismo efecto en la alteración del *splicing* que las mutaciones c.8754+4A>G, c.8754+5G>A y c.8754+5G>T, por lo que su importancia clínica parece constatada. Por extensión, los transcritos de la variante c.8954-3C>G también podrán ser degradados en parte por el mecanismo del MND y los transcritos que no sean procesados de esta forma producirán proteínas aberrantes. La elevada proporción de transcritos aberrantes junto con la introducción de PTCs indican que estas variantes están implicadas en la etiopatogenia del CMOH. Las variantes c.8487+3A>G y c.8487+3A>C provocaron transcritos que habían perdido el exón 19 (p.Ile2778\_Gln2829del). De este modo se mantiene la fase de lectura y sus efectos en la proteína son similares a la variante patogénica c.8487+1G>A, por lo que estas variantes deberían ser reclasificadas como mutaciones de *splicing*.

Las variantes intrónicas c.8633-24\_8634del y c.8954-1\_8955delinAA registradas en la base de datos BIC, que consisten en deleciones de la región intrón-exón afectaron de forma drástica al correcto *splicing*. No obstante mientras que la c.8633-24\_8634del produce transcritos que incorporan codones de parada prematuros (p.Glu2878Glyfs\*4 y p.Gly2919Leufs\*3) la c.8954-1\_8955delinAA en el 80% de los transcritos mantiene la pauta de lectura, no serán traducidos los 7 primeros aminoácidos del exón 23, el resto de transcritos produce PTC (p.Val2985Glyfs\*3). Ambas mutaciones pueden ser clasificadas como de importancia clínica.

Es importante recalcar que no todas las variantes en posiciones cercanas a los sitios conservados aceptor y donador alteran el *splicing*. Por ejemplo, la variante

c.9501+4A>G no causa ningún efecto mientras que el efecto de la c.8754+4A>G es total aunque ambas afectan a un nucleótido altamente conservado (+4, A: 68% vs G: 12%). El contexto genómico y el grado de conservación evolutiva de las secuencias flanqueantes al sitio aceptor y donador estarían involucrados en estos efectos. Sin embargo, no puede establecerse una regla general y se hace necesario siempre proceder a evaluar la posible alteración del *splicing* mediante ensayos funcionales, más aún si cabe teniendo en cuenta la variedad de eventos que pueden ocurrir (*exon skipping*, uso de sitios alternativos, etc)

### III. Reproducibilidad de los resultados obtenidos en minigenes

Un punto muy importante de cara a la validación de los minigenes construidos es su capacidad de reproducir los patrones de *splicing* fisiológicos o inducidos por una determinada variante de ADN que se han obtenido en tejidos de pacientes. En este sentido, cabe subrayar que la mayoría de los estudios de *splicing* realizados en pacientes se hace sobre ARN de linfocitos, siendo muy poco frecuente el ensayo sobre el tejido diana de la enfermedad.

Un primer factor que avala la reproducibilidad del estudio es que cada minigen genera mayoritariamente un ARNm del tamaño y estructura exónica esperados (figuras 29, 30 y 37 de resultados), salvo en alguna construcción intermedia, como MGiBR2 17-18, que produce transcritos anómalos que posteriormente son corregidos mediante la inserción de exones vecinos (mantenimiento del contexto genómico; Figura 30).

En segundo lugar, es capaz incluso de reproducir los patrones de *splicing* alternativo fisiológicos como es el caso del *skipping* de los exones 18 y 17-18 descritos previamente (Walker *et al.*, 2010), en el contexto del minigen MGBR2 16-20.

En tercer lugar (y lo que es más importante desde el punto de vista de diagnóstico), en la mayoría de los casos los minigenes pueden replicar los efectos de variantes de ADN sobre el *splicing*. En la literatura se han encontrado datos correspondientes a once variantes ensayadas en los minigenes MGBR2 19-27, MGBR2 23-24, MGBR 19-20 de las cuales se hace un breve resumen a continuación.

La variante c.8487+1G>A registrada en un paciente con fuerte historial en CMOH ha sido estudiada mediante RT-PCR de linfocitos por Chen *et al.* (2006). Los resultados publicados muestran que aunque se trate de una mutación en heterocigosis curiosamente, existe un mayor porcentaje de transcritos aberrantes que *wild type*. El estudio de la variante en el minigen MGBR2 19-20 elimina el efecto de heterocigosis observándose un tipo único transcrito con *skipping* del exón 19, resultado que se corresponde con el descrito previamente mediante RT-PCR de linfocitos.

Los ensayos funcionales de *splicing* con el minigen de la mutación c.8754+4A>G donde se observa la inclusión de los primeros 46 nt del intrón 21 en los transcritos maduros, coinciden con los expuestos en el estudio en linfocitos publicado por Bonatti *et al.* (2006). Otro estudio reciente (Thomassen *et al.*, 2012) analizó la mutación c.8754+3G>C tanto en ARN de linfocitos como en ensayos con un minigen construido en pSPL3 que únicamente constaba del exón 21. Aunque esta variante no ha sido ensayada en este estudio, otras variantes del sitio donador, como c.8754G>A, y los tres posibles cambios en la posición c.8754+5G>A/T/C, mostraban el mismo efecto en el *splicing* que las variantes c.8754+4A>G y c.8754+3G>C. En todos los casos el sitio donador del exón 21 es eliminado y es reconocido un sitio críptico 46 nt *downstream* del natural que afecta al 100% de los transcritos.

Los ensayos funcionales de *splicing* en linfocitos de la variante c.8953+1G>T (Whiley *et al.*, 2011) son significativamente diferentes a los obtenidos en el contexto del minigen MGBR2 19-27 con dicha variante. Los resultados en linfocitos muestran dos tipos de transcritos: los que pierden el exón 22 y los que tienen una deleción de los últimos 31 nt del exón 22. En minigenes se observan tres tipos de transcritos: *skipping* del exón 22 (la única coincidencia) y dos transcritos más, uno con una inserción de 4 nt del intrón 22 debido al uso de un sitio donador críptico y otro con retención completa del intrón 22. Los resultados en minigenes son similares a los obtenidos en el ensayo de la mutación c.8948\_8953+5del que afecta también al sitio donador del exón 22. La diferencia entre los resultados obtenidos entre ARN de linfocitos y minigenes no podría ser atribuida a una falta de contexto genómico, pues el MGBR2 19-27 incorpora

suficientes exones para limitar este hecho. Una posible explicación sería que en el ensayo previamente publicado, el ARN de linfocitos se obtuvo tras un tratamiento con un inhibidor del NMD (cicloheximida) no efectuado en el caso del MGBR2 19-27 con la variante c.8953+1G>T. Otra posible explicación sería el tipo celular en el que se ha realizado el ensayo, ya que los ensayos funcionales de minigenes se realizaron en células HeLa y MCF7.

El resultado obtenido de la variante c.8954-3C>G del sitio aceptor del exón 23 y ensayada mediante el análisis funcional de *splicing* con el MGBR2 23-24 es similar al descrito en un estudio reciente (Menéndez *et al.*, 2012) para la variante c.8954-5A>G a través de RT-PCR de linfocitos. En ambos casos se genera un sitio aceptor alternativo muy cercano al natural para el exón 21 (2 y 4 nt *upstream*, respectivamente). Estos sitios alternativos nuevos tienen una probabilidad relativa (0,87 y 0,89 NNsplice) mayor de ser reconocidos por la maquinaria de *splicing* que el natural (0,51 NNsplice). El efecto en el *splicing* no es exactamente el mismo, ya que la variante c.8954-3C>G provoca transcritos que incorporan los últimos 2 nt del intrón 20, mientras que la mutación c.8954-5A>G incorpora 4 nt, al situarse el nuevo sitio aceptor 4 nt *upstream* del natural. El posible efecto en la proteína también es parecido, puesto que en ambos casos las isoformas producidas no mantienen la fase de lectura y contienen un codón prematuro de parada.

Interesantemente, la variante c.9117G>A situada en el último nucleótido del exón 23 estudiada en RT-PCR de linfocitos (Bonatti *et al.*, 2006) muestra el mismo resultado (100% de transcritos con *skipping* del exón 23) que el obtenido mediante el ensayo funcional de *splicing* con el MGBR2 19-27 que porta dicha mutación. Esta mutación también fue ensayada en el MGBR2 23-24. En este caso el resultado incluía un 25% de transcritos con pérdida de los exones 23 y 24. Quizá la falta de contexto genómico del minigen o el propio vector pSPL3 en el que fue construido hayan propiciado la aparición de dichos transcritos. En todo caso, este hallazgo ilustra las ventajas del nuevo vector pSAD.

La variante c.9118-2A>C del sitio aceptor del exón 24 fue analizada previamente mediante RT-PCR de linfocitos (Claes *et al.*, 2003). Los resultados obtenidos en nuestro estudio mediante ensayo funcional de *splicing* con el MGBR2 23-24 que porta dicha mutación son muy similares. En ambos trabajos la mayor parte de los transcritos carecen de los primeros 7 nt del exón 24 puesto que un sitio aceptor críptico interno es utilizado a causa de supresión del aceptor natural, y un 10% de transcritos que carece del exón 24. El análisis de un único alelo en el sistema de minigenes combinado con el análisis de fragmentos de la RT-PCR fluorescente en un secuenciador automático permiten alcanzar una mayor sensibilidad en la detección de transcritos minoritarios que en linfocitos o en cualquier otro tejido no podrían ser identificados. La consecuencia es muy similar ya que las dos isoformas detectadas contribuyen a la aparición de un codón prematuro de parada de la traducción.

En el mismo estudio (Claes *et al.*, 2003) se analizó también mediante RT-PCR de linfocitos otra variante incluida en este trabajo, c.9256+1G>A, que había sido ensayada tanto en el MGBR2 23-24 como en el MGBR2 19-27. Aunque prácticamente los resultados coincidían con los publicados en linfocitos y mostraban la pérdida del exón 24 en la gran mayoría de los transcritos maduros, en el estudio con minigenes se pudo identificar un nuevo transcrito aberrante que había perdido los últimos 46 nt del exón por uso de un sitio críptico donador que constituía el 25% de todas las isoformas de *splicing* detectadas. Nuevamente, este dato apoya firmemente la alta sensibilidad de la tecnología de los vectores reporteros de *splicing* para la detección de todos los transcritos generados por una determinada variante de ADN. Sendos transcritos aberrantes no mantienen la fase de lectura y por lo tanto aparecen codones de parada prematuro.

El estudio de la variante c.9257-1G>C en el sitio aceptor del exón 25, ensayada en linfocitos (Whiley *et al.*, 2011) mostró dos tipos de transcritos: uno con *skipping* del exón 25 y otro minoritario con una deleción “*in frame*” de los primeros 27 nt del exón 25 por uso de un sitio críptico aceptor en ensayos de RT-PCR de linfocitos con inhibición del NMD con cicloheximida. Sin embargo, en los resultados sin la inhibición del NMD no se detectó el *skipping* del exón 25. En este caso los resultados obtenidos

en los ensayos funcionales de *splicing* con el MGBR2 19-27 que porta dicha mutación fueron idénticos a los previamente reportados en RT-PCR de linfocitos sin inhibición del NMD. Al igual que ocurre en el caso de la variante c.8953+1G>T, el uso de la cicloheximida puede estar determinando, que en este caso los resultados no sean exactamente los mismos. Por el contrario, la variante c.9501+3A>T del sitio donador del exón 25 mostró resultados similares en linfocitos (Bonnet *et al.*, 2008) y en el minigen MGBR2 19-27, ambos consistían en la generación de dos transcritos: el canónico, y otro con la delección del exón 25, que en el caso del MGBR2 19-27, representaba un 13%.

Respecto a las mutaciones seleccionadas de las bases de datos por su posible alteración de elementos reguladores de *splicing* (*enhancers* y silenciadores) de acuerdo a las previsiones informáticas y ensayadas mediante minigenes híbridos, hemos encontrado coincidencias en nuestros resultados negativos con los publicados por otros grupos (Whiley *et al.*, 2010) para las variantes c.8972G>A (exón 23), c.9172A>G (exón 24) y c.9213G>T (exón 24).

Los resultados obtenidos en este trabajo mediante ensayos funcionales de *splicing* de minigenes han demostrado la alta fiabilidad del sistema al reproducir tanto el tamaño de ARNm esperado en cada minigen, como eventos fisiológicos de *splicing* alternativo. Asimismo, la reproducibilidad de los resultados procedentes de los ensayos con minigenes de este trabajo, con los publicados por otros grupos donde se han utilizado técnicas como ensayos RT-PCR de linfocitos, demuestra que la técnica de minigenes híbridos es altamente eficiente y robusta, y más aún en aquellos minigenes construidos en el nuevo vector pSAD. La estandarización de procedimientos de cuantificación de isoformas, de los tipos celulares de ensayo y el uso de inhibidores de NMD, permitiría una comparación más precisa y una mayor reproducibilidad de los resultados obtenidos en distintos tipos de ensayos funcionales.

#### IV. Técnicas de identificación de alteraciones de *splicing* provocadas por mutaciones puntuales en *cis*

La caracterización de transcritos aberrantes por causa de una mutación y su cuantificación son fundamentales para evaluar la patogenicidad de la misma.. Además, el tipo celular del que se obtienen los resultados y el posible efecto del mecanismo de NMD como proceso de degradación de transcritos con PTCs deben ser tenidos en cuenta de cara a mejorar las conclusiones sobre el grado de patogenicidad de una variante de ADN.

El método de elección para testar una variante de ADN con presunto impacto en el *splicing* debería ser el análisis directo en muestras de la paciente (Baralle *et al.*, 2009), pero esto no es siempre posible ya que la paciente podría no estar disponible. *A priori*, la técnica más adecuada sería la RT-PCR de células del tejido diana de la enfermedad (en este caso, epitelio mamario) de la paciente portadora de la variante para evitar los efectos derivados del *splicing* alternativo específico de tejido, pero esto constituye una tarea aún más complicada.

Las dificultades de esta metodología pueden ser solventadas utilizando un tipo de muestra diferente al tejido diana, pero que pertenezca igualmente a la persona portadora de la variante genética que se pretende analizar. Puesto que la extracción de sangre es un procedimiento de rutina en la práctica clínica, la extracción de ARN procedente de linfocitos es una alternativa adecuada para la realización de RT-PCR. Sin embargo, el *splicing* alternativo de tejido es un proceso que influye directamente en el tipo y la proporción de los transcritos detectados, pudiendo encontrarse diferencias significativas entre linfocitos y el tejido de interés, sobre todo si la mutación afecta a la expresión génica específica de tejido. Además, la inadecuada conservación de la sangre puede causar la detección de isoformas anómalas de *splicing* (Speevak *et al.*, 2003).

Una dificultad añadida en los ensayos de *splicing* en linfocitos es la cuantificación precisa de las isoformas inducidas por la variante. Por lo general la mutación se haya en uno de los alelos y cuando el efecto es parcial, es complicado cuantificar los transcritos canónicos procedentes del alelo mutante, a lo que hay que añadir el sesgo producido por la eliminación de transcritos con PTCs causada por el NMD que obligaría al uso de inhibidores de NMD en este proceso. Así, una ventaja radica en el hecho de

poder identificar y cuantificar todos los transcritos generados por una mutación independientemente del alelo *wild type*.

Los ensayos funcionales de *splicing* mediante minigenes híbridos solventan varios de los problemas que pueden aparecer con las técnicas de análisis directo de una muestra del paciente. Este tipo de aproximación no necesita ARN procedente de la persona portadora de la mutación a estudiar. Solamente se precisa de ADN (por ejemplo de un control anónimo), una molécula mucho más estable y fácil de obtener y conservar. De esta manera se evitan los artefactos procedentes de la mala conservación de las muestras de las que se extrae ARN. Técnicas de manipulación de la secuencia de ADN, utilizadas tras la construcción del minigen *wild type*, permiten recrear cualquier mutación tanto detectadas previamente en pacientes como generadas por primera vez en regiones que teóricamente pueden regular el procesamiento del pre-ARNm.

En cuanto a la cuestión del *splicing* alternativo específico de tejido, los ensayos de *splicing* con minigenes híbridos permiten analizar el efecto de una variante en diferentes tipos celulares y establecer comparaciones entre ellas para optimizar un ensayo concreto. Estudios previos demostraron que los resultados de ensayos funcionales con minigenes que recreaban mutaciones en los genes *BRCAs* utilizando tanto células HeLa como células de epitelio mamario no cancerosas MCF10A eran muy similares, presentando diferencias cuantitativas no significativas (Sanz *et al* 2010). En este estudio se ha realizado también una comparación entre los resultados obtenidos en la transfección de células HeLa y células epiteliales de mama cancerosas MCF-7 con los MGBR2 19-27 *wild type* y aquellos en los que se generaron las variantes c.8486A>C, c.8488-1G>A, c.9117G>A, c.9256G>T, c.8948\_8953+5del, c.8953+1G>T y c.9649-2A>G, no encontrando en ningún caso diferencias significativas. Según estos resultados, podemos afirmar que las células HeLa son una línea celular óptima para los ensayos funcionales híbridos por su alta eficiencia de transfección, estabilidad, su facilidad de manipulación y alto rendimiento de ARN producido a través del promotor de SV40 del minigen.

Los ensayos con minigenes híbridos permiten estudiar no sólo el efecto de una mutación concreta, sino mapear el conjunto de secuencias reguladoras esenciales para el reconocimiento correcto de un exón a través de cambios de nucleótidos específicos

o mediante la delección de regiones más amplias (microdeleciones solapantes) que podrían incluir varios elementos reguladores.

En los minigenes utilizados en este estudio se han podido detectar determinadas aberraciones en el proceso de *splicing* no identificadas en ensayos RT-PCR de linfocitos, ni descritos con anterioridad. Este es el caso de los transcritos que carecen de los exones 19-20 y 23-24 de forma conjunta, que aparecen de forma recurrente y minoritaria respectivamente en el MGBR2 19-20, MGBR2 23-24 y sus derivados que contienen las mutaciones ensayadas. La aparición de estos pequeños transcritos anómalos con sólo los exones constitutivos del vector es fruto de la falta de un contexto genómico amplio en el minigen debido a la dificultad de clonar varios exones en el mismo minigen, sobre todo en el caso de intrones de gran tamaño que impiden su inserción completa en el vector en una sola etapa de clonaje como por ejemplo, el intrón 24 de BRCA2 que tiene 14,5 Kb. Este problema se ha solucionado de forma específica para el conjunto de exones mencionado, mediante la clonación de una versión resumida de los intrones con 100-400 nucleótidos flanqueantes esenciales y, como ya se ha comentado, gracias al nuevo vector pSAD, que mejora en muchos aspectos al plásmido pSPL3 usado en la creación del MGBR2 19-20 y MGBR2 23-24. Entre ellas podemos destacar la reducción del tamaño del vector y el aumento de secuencias de restricción únicas dentro del sitio múltiple de clonación que incrementan su versatilidad y facilitan la inserción de fragmento de ADN más grandes. También, el diseño de protocolos derivados de la mutagénesis dirigida, como la PCR de inserción han contribuido a generar los minigenes con más exones del gen *BRCA2* conocidos hasta el momento. Hasta nueve exones con al menos 200 nt de intrón a cada lado han sido clonados en el vector pSAD dando lugar al MGBR2 19-27. Otra ventaja derivada de la construcción de grandes minigenes con el vector pSAD recae en la disminución de reacciones heterólogas entre los sitios de *splicing* del vector y los sitios de los exones clonados. Más aún, dado que el exón 27 es el último exón de BRCA2, en el MGBR2 19-27 se fusionó el exón V2 constitutivo del plásmido con el exón 27 del gen *BRCA2* (con su sitio aceptor natural), eliminando de esta forma el sitio aceptor de V2. Así, se redujo la posibilidad de que ocurrieran reacciones heterólogas entre los sitios de *splicing* del vector y los clonados. No sólo eso, se mantuvieron las secuencias de V2 donde hibrida el oligonucleótido RT-pSAD-RV, de modo que sólo se

amplificarán específicamente aquellos transcritos procedentes del minigen. La región genómica original de los exones 19 a 27 de BRCA2 tiene un tamaño de 28.717 pb que se ha resumido en un inserto del minigen MGBR2 19-27 con un tamaño de 5.620 pb (estructura en la Figura 37 - Resultados). Ambas regiones son equivalentes funcionalmente desde el punto de vista del *splicing*, ya que se produce un transcrito mayoritario con el exón V1 y los exones 19 a 27-V2. Para llegar a esta “versión resumida” se han acortado los intrones 20, 21, 24, 25 y 26. Cabe resaltar el intrón 24 con un tamaño genómico de 14,5 Kb se redujo a 418 pb (147 + 271). Teniendo en cuenta la integridad del ARNm producido por el minigen, estos resultados indican que la mayoría de las secuencias intrónicas no son necesarias para el correcto procesamiento del ARNm precursor.

Otro dato que avala la importancia de la arquitectura intrón-exón en el ámbito de los minigenes recae en los resultados obtenidos con las microdeleciones del MGBR2 16-20 donde la deleción c.7944\_7973del próxima al sitio donador del exón 17 y las deleciones c.7979\_8008del y c.8004\_8033del del exón 18 y cercanas al sitio aceptor del mismo han alterado el *splicing*. Estas regiones parecen claves para el reconocimiento de ambos exones 17 y 18 o del intrón 17. Dichos resultados contrastan con los obtenidos mediante estudios previos realizados con un minigen construido en pSPL3 donde sólo se incluía el exón 18 (Sanz *et al.*, 2010). Mutaciones ensayadas en este minigen y situadas en las regiones de las deleciones, como la c.7994A>G que afectaría, según los estudios informáticos, a posibles secuencias enhancer que facilitarían el reconocimiento del sitio aceptor del exón 18 y que en principio no causaron ningún efecto en el *splicing* (Sanz *et al.*, 2010), deberían ser reevaluadas en nuevos minigenes de mayor tamaño. Es plausible que alguna de las variantes sin efecto en el minigen del exón 18 den un resultado diferente en un contexto mayor como el minigen MGBR2 16-20. Teniendo en cuenta que el sitio donador del exón 17 es atípico (GC en vez de GT) parece lógico pensar que deben existir elementos reguladores que faciliten el reconocimiento de las fronteras entre el exón 17 y el intrón 17 y de éste último con el exón 18. La falta de este contexto en el minigen construido que contiene sólo el exón 18 podría estar influyendo en los resultados obtenidos y serían necesarios nuevos estudios funcionales de variantes previamente consideradas negativas en el nuevo minigen con los exones 16 a 20 de BRCA2.

La única variable que conserva el nuevo vector pSAD con respecto a su antecesor pSPL3 es que mantiene la misma secuencia promotora de SV40 y ésta es distinta a la del promotor natural del gen *BRCA2* por lo que el nivel de expresión y la velocidad de transcripción de los minigenes construidos es diferente a la natural. Según la teoría cinética de la transcripción (Luco *et al.*, 2010) comentada en la introducción, la velocidad de la transcripción está implicada en el reconocimiento y procesado del pre-ARNm.

El promotor vírico de los vectores pSAD y pSPL3 promueve una alta tasa de transcripción por lo que todos los transcritos aberrantes que contienen codones de parada prematuros difícilmente son eliminados por el mecanismo del NMD. En caso contrario sería imposible detectar este tipo de isoforma de *splicing* sin inhibir el NMD, hecho que no ocurre como se ha podido demostrar con nuestros resultados. La comparación de los resultados obtenidos de la transfección de células HeLa, de las que se ha descrito una alta eficiencia de este mecanismo (Linde *et al.*, 2007) con y sin inhibición del MND mediante cicloheximida demuestra que este mecanismo está presente en los ensayos funcionales de *splicing*. Sin embargo, la alta eficiencia del MND en células Hela no ha impedido detectar isoformas aberrantes que no están en fase, debido a la alta tasa transcripcional de los minigenes. La inhibición del NMD en comparación con el ensayo con minigenes sin cicloheximida de la variante c.8948\_8953+5del ha mostrado un aumento del porcentaje de la isoforma con deleción de 7 nt y se ha podido detectar una isoforma no identificada previamente en la que presumiblemente ocurre la retención de varias regiones intrónicas. El exón V1 del vector no estaba en fase con algunos de los transcritos generados. Para reproducir y contextualizar el efecto del MND en todos los transcritos producidos sería necesario generar los minigenes por triplicado, con tres diferentes vectores que incluyeran 0, 1 y 2 nt en el exón V1 para mantener el ORF de los exones clonados, pero la complejidad del ensayo y el gran esfuerzo técnico aumentaría considerablemente sin aportar datos significativos y diferenciados respecto al procedimiento habitual de inhibición del NMD.

El nuevo vector pSAD además ha corregido algunas de las carencias y problemas habituales en el manejo de pSPL3 a la hora de realizar ensayos funcionales con minigenes híbridos. La inclusión de regiones promotoras del gen en estudio en

sustitución de la que posee el vector pSAD incluso podría *a priori* mejorar la reproducibilidad del ensayo. Aunque los ensayos con minigenes híbridos requieren mayor complejidad respecto a la técnica de RT-PCR de linfocitos, la independencia del estudio de las muestras biológicas es una excelente ventaja en la práctica clínica. Nuevos sistemas que permitan la realización de ensayos funcionales a través de la manipulación de la secuencias de ADN en líneas celulares o la generación de modelos animales podrían constituir un avance metodológico importante que complementarían los datos que aportan los ensayos funcionales de *splicing* de diferentes variantes mediante minigenes híbridos.

#### V. Tipos de alteración en cis detectadas y efecto en el proceso de *splicing*

En este trabajo se analizaron 333 variantes de ADN detectadas en pacientes con CMOH y 9 artificiales (sin considerar las microdeleciones, ni la c.97806-14T>C procedente de una de colaboración) mediante estudios *in silico* con programas de predicción de alteración del *splicing*, de las que 88 fueron seleccionadas como putativas variantes con impacto en el *splicing*. Los resultados obtenidos mediante ensayos funcionales confirmaron que realmente 39 de ellas (44.3%) afectaban al proceso de *splicing*.

Aproximadamente el 80% (31/39) de las variantes que afectan al proceso de *splicing* se sitúan en la frontera exón-intrón (-3, -2, -1, +1, +2, +3, +4, +5 en intrones y último y penúltimo nucleótidos de exones) alterando los sitios donador y aceptor canónicos. En su mayoría, estas alteraciones afectaban a un porcentaje mayoritario o a la totalidad de los transcritos. Sin embargo, este hecho no se cumple para las regiones adyacentes a los sitios de *splicing*. Porcentajes menores del 30% de isoformas aberrantes se han cuantificado en las variantes c.8487+3A>T que afectan al sitio donador del exón 19, c.9256G>T en el exón 24, c.9501+3A>T en el exón 25. Por lo tanto, aunque los cambios nucleotídicos en estas posiciones son altamente sospechosos, no existe una regla general para predecir un efecto deletéreo total sobre el *splicing*. Aunque los programas de predicción informáticos de los sitios de *splicing* son bastante precisos, es necesario en todo caso realizar ensayos funcionales de *splicing* para caracterizar cualquier variante, porque es muy complejo predecir el patrón anómalo de *splicing* (exon *skipping*, uso de sitios alternativos, retención de

intrón) que inducirá cada variante. Todos los pronósticos informáticos deben ser corroborados funcionalmente.

Asimismo hay que mencionar aparte a las variantes c.8486A>T y c.8487G>A que disminuyen la probabilidad de reconocimiento del sitio donador del exón 19, pero también parecen afectar a *enhancers* de *splicing* según las predicciones informáticas. Su proximidad al sitio donador (penúltimo y último nucleótidos del exón, respectivamente) y otros resultados en posiciones similares, como los de la mutación c.8754G>A situada en el último nucleótido del exón 24, han propiciado que las hayamos clasificado como específicas de sitios de *splicing*. Ensayos de unión ARN-proteína serían esenciales para averiguar la causa del efecto observado.

Por su parte, dos variantes, la c.8755-1G>A en el sitio aceptor del exón 22 y la c.8954-3C>G del sitio aceptor del exón 23, han generado un nuevo sitio de aceptor de *splicing* muy cerca del natural beneficiándose tanto del punto de ramificación como del tracto de polipirimidina situados en el intrón 22. El efecto sobre el *splicing* fue casi total en la c.8755-1G>A, ya que el 80% de los transcritos no incluían el primer nucleótido del exón 22, mientras que en la mutación c.8954-3C>G fue total, produciendo transcritos aberrantes que afectaron a la fase de lectura (inserción de 2 nt). Un efecto similar fue descrito para la variante c.302-3C>G en el gen *BRCA1* (Sanz *et al.*, 2010). La microdelección de 30 nt c.8956\_8985del realizada en el exón 23, y las variantes c.8996G>A, c.8972G>A y c.9006A>T, inducían isoformas aberrantes, confirmando que el reconocimiento del sitio aceptor del exón 23 necesita elementos reguladores de *splicing* adicionales. Otro ejemplo de creación de sitios de *splicing* en la variante c.8948\_8953+3del en el sitio donador del exón 23. La mutación crea un nuevo sitio donador al eliminar los 6 último nucleótidos del exón 23 y los 5 primeros del intrón 23.

Todas las variantes que alteran elementos reguladores de *splicing* (7/39) y que resultaron positivas producen un efecto parcial minoritario en la alteración del *splicing*. Se pudo comprobar que tanto la creación de sitios silenciadores (ESS) (por ejemplo en las mutaciones c.8969G>A -crea hnRNP-A1 o B-, c.9076C>T y c.9698G>T) como la eliminación de secuencias *enhancers* de *splicing* (es el caso de la variante 9006A>T (SRp40) o la artificial doble c.[8609A>G;8611G>T] (SRp40)) tenían un efecto en el proceso de *splicing*.

Un efecto singular se pudo observar al ensayar la combinación en un mismo minigen de las variantes c.8972G>A y c.9006A>T que eliminaban las secuencias de unión de proteínas *enhancers* SC35 y SRp40 respectivamente. De forma individual la variante c.8972G>A no provocó ningún efecto observable pero la c.9006A>T provocaba un efecto parcial con un 24% de los transcritos aberrantes que no incorporaron los primeros 51 nt del exón 23. De forma conjunta, introduciendo ambas mutaciones en el mismo minigen, se detectó un aumento considerable en el porcentaje de transcritos aberrantes llegando al 40%. Este resultado puede implicar una complejidad añadida en el estudio del efecto de una mutación sobre el proceso de *splicing*, puesto que mutaciones consideradas simples polimorfismos pueden actuar de forma sinérgica en la modulación del *splicing*. Por otro lado, este resultado sugiere que los factores de *splicing* SRp55 y SRp40 actúan de forma cooperativa de cara al reconocimiento preciso del exón 23 (Acedo *et al.*, 2012).

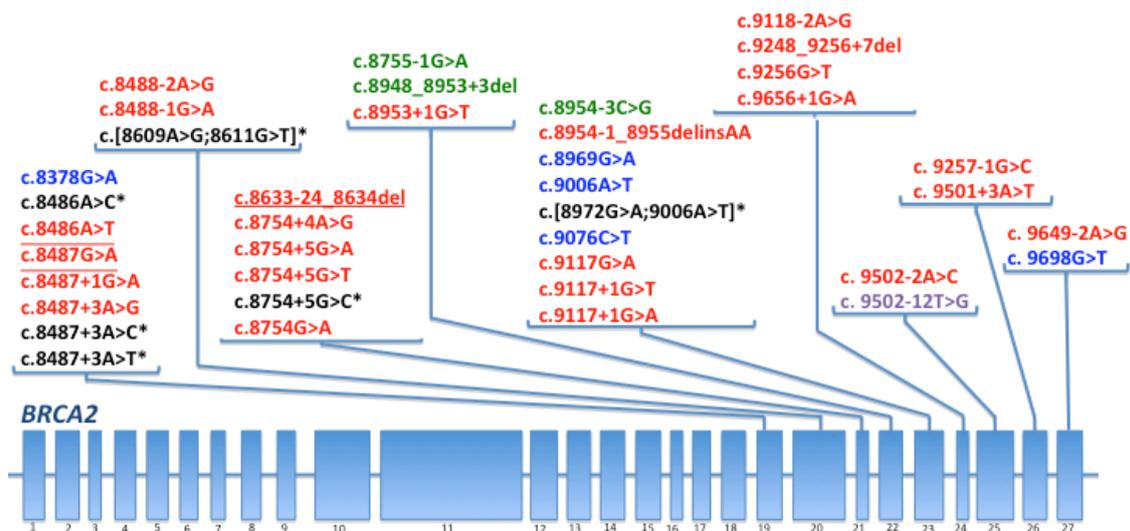
El mismo efecto ha sido observado en la combinación de deleciones realizadas en el intrón 19 del gen *BRCA2*, en las que se eliminó una secuencia Alu y dos secuencias altamente conservadas flanqueantes a la misma. No se observó efecto por separado, pero las combinaciones doble [c.8487+31\_8487+42del -c.8487+54\_8487+355del] y triple [c.8487+31\_8487+42del - c.8487+54\_8487+355del - c.8488-49\_8488-44del] provocaron el *skipping* del exón 19 en un 3% y un 7% respectivamente, sugiriendo que estas secuencias son fundamentales para la eliminación del intrón 19. La cuantificación de las isoformas detectadas reveló el doble de transcritos aberrantes con la triple deleción respecto a la doble, por lo que la alteración combinada de secuencias intrónicas puede también producir un efecto multiplicador en cuanto a la alteración del *splicing* similar al observado en la alteración de secuencias reguladoras exónicas del párrafo anterior y podría ser explicado por la cooperación de elementos reguladores para facilitar el reconocimiento de los sitios de *splicing*.

La eliminación simultánea de las secuencias conservadas y Alu reduce el tamaño del intrón 19 de 398pb a 98 pb. Esta disminución podría afectar al reconocimiento del intrón 19 puesto que se ha estimado que el tamaño mínimo que debe tener un intrón humano para ser reconocido por el espliceosoma es de 92±14pb (Yu *et al.*, 2002).

Más complejo es establecer la causa por la que la variante c.8378G>A del exón 19 altera el *splicing* (*skipping* del exón 19 en el 12% de los transcritos). Las

predicciones informáticas muestran dos efectos opuestos: eliminación de *enhancers* y creación de un silenciador (hnRNP-B). Las alteraciones de *splicing* (*skipping* del exón 19) inducidas por la microdelección c.8334\_8363del próxima al sitio aceptor del exón 19 y a la variante c.8378G>A sugieren que deben existir elementos reguladores de *splicing* promoviendo el reconocimiento del exón 19. Al encontrarse la variante c.8378G>A fuera de la secuencia de la microdelección c.8334\_8363del, estos datos sugieren la existencia de, al menos, dos elementos *enhancers* que podrían cooperar para facilitar el reconocimiento del sitio aceptor. Como se ha podido comprobar en este trabajo y en estudios previos (Sanz *et al.*, 2010), los factores de *splicing* actúan de forma sinérgica para facilitar un correcto procesado del pre-ARN. En el supuesto de que la variante c.8378G>A estuviera generando un sitio de reconocimiento de una proteína represora, podría interferir estéricamente en la unión de la proteína activadora del *splicing* que, como hemos detectado en con la microdelección c.8334\_8363del, se uniría a una secuencia cercana al sitio aceptor para facilitar su reconocimiento. En cualquier caso, es necesario realizar ensayos de interacción ARN-Proteína para poder aclarar la causa real de la anomalía de *splicing*.

Otra variante intrónica ensayada, relativamente alejada de los sitios de *splicing*, ha afectado al tracto de polipirimidina del sitio aceptor del exón 26. La variante c.9502-12T>G tiene un efecto parcial sobre el *splicing*. El tracto de polipirimidina junto con el punto de ramificación son secuencias necesarias para el reconocimiento del sitio aceptor por parte del espliceosoma. En este caso se sustituye una pirimidina (U, 53% en la secuencia consenso) por una purina (G, 10% en la secuencia consenso), de modo que influye levemente en la fortaleza del sitio aceptor (0,98 a 0,94 NNsplice). El efecto parcial (en el que el 8% de los transcritos no contienen el exón 26) es similar a otros ya documentados para otras variantes que afectan al tracto de polipirimidina (como por ejemplo para la mutación c.68-7T>A del sitio aceptor del exón 3 de *BRCA2*; Sanz *et al.*, 2010). No obstante, se ha mostrado que deleciones parciales del tracto de pirimidina, como la variante c.426-12\_426-8del del intrón 4 de *BRCA2*, también pueden causar alteraciones completas del patrón de *splicing* (*skipping* del exón 5; Zhang *et al.*, 2009; Sanz *et al.*, 2010).



**Figura 3. Variantes en *cis* ensayadas que afectan al *splicing*.** La figura representa la posición de las variantes del gen *BRCA2* ensayadas con impacto en el *splicing*. En color rojo se muestran las variantes que afectan a los sitios de *splicing*. En azul aquellas que eliminan *enhancers* de *splicing* o la crean silenciadores. En color morado las que afectan al tracto de polipirimidina y en verde las que crean un nuevo sitio de *splicing*. Las variantes identificadas en color negro son artificiales y las subrayadas podrían afectar el *splicing* de más formas además de alterar los sitios de *splicing*.

El análisis global de los resultados obtenidos muestra que, del total de mutaciones estudiadas, detectadas en pacientes con mutaciones en el gen *BRCA2* y registradas en las bases de datos internacionales, aproximadamente un cuarto (81/334, el 24%) cumplen los criterios de selección bioinformáticos para realizar ensayos funcionales de *splicing*. Más de un 40% (33/81) de las mutaciones seleccionadas han resultado positivas en cuanto a la alteración del *splicing*. Por lo tanto, se puede afirmar que los criterios de selección de variantes han sido óptimos. No obstante, aunque las predicciones de alteración del *splicing* de mutaciones que afectan a los sitios canónicos aceptor y donador son muy acertadas, no lo son tanto cuando se trata de mutaciones que crean o eliminan elementos reguladores de *splicing* (They *et al.*, 2011). Además, la predicción informática no indica ni el tipo de transcritos resultante (*exon skipping*, uso de sitios alternativos o retención de intrón) ni el porcentaje de los mismos, por lo que siempre se hará necesaria la validación funcional de las predicciones informáticas. Extrapolando los resultados obtenidos (33 mutaciones positivas de 334 analizadas) al total de variantes registradas en las bases de datos BIC para el gen *BRCA2*, actualmente 1.994, hipotéticamente 200 variantes afectarían al *splicing*, cuando en la realidad tan sólo 4 de este tipo están catalogadas específicamente como mutaciones de *splicing*. Además, si consideramos que todas las

variantes intrónicas de importancia clínica conocida han sido catalogadas como tal por afectar a los sitios de *splicing* canónicos (actualmente 43 mutaciones en la base de datos BIC) y comparamos ese número según los resultados obtenidos (22 mutaciones intrónicas alteraron los sitios de *splicing* de un total de 34 variantes positivas analizadas, un 66%), deberíamos suponer que unas 130 variantes intrónicas del total de 192 variantes intrónicas registradas en el BIC serían de importancia clínica y afectarían al *splicing*. Teniendo en cuenta estos porcentajes y extrapolando los mismos al conjunto de los genes *BRCAs*, podemos concretar que el 10% de las variantes registradas en los genes *BRCAs* afectan al proceso de *splicing*. Este porcentaje sería ligeramente superior al de variantes *nonsense* registradas (9,5%), por lo que la alteración del *splicing* podría ser considerado el segundo mecanismo más importante en la etiopatogenia del CMOH, por detrás del ocasionado por las mutaciones truncantes de proteína *frameshift* (25%).

## VI. Evolución y perspectivas del modelo de cáncer de mama y ovario hereditario

A partir de los años 90, el descubrimiento de la relación entre mutaciones detectadas en los genes *BRCA1* y *BRCA2* en ciertas familias con una alta frecuencia de casos con cáncer de mama transformó la visión de esta enfermedad y fue clasificada como un síndrome concreto, en el que es posible aplicar programas específicos de prevención en base a las características genéticas de los pacientes. No obstante, los datos actuales indican que, de media, el 16% de los casos de CMOH pueden ser explicados por mutaciones en estos genes (Thompson y Easton, 2004). Este porcentaje puede variar en función de las diferentes poblaciones, como por ejemplo, por la presencia de mutaciones fundadoras (Infante *et al.*, 2010). Es previsible que en el contexto globalizado de la sociedad del siglo XXI se homogenicen estos datos epidemiológicos. En la última década, el descubrimiento de las bases moleculares y la función de los genes *BRCAs* han posibilitado aumentar el conocimiento del espectro de predisposición a CMOH mediante el hallazgo de mutaciones en nuevos genes relacionados con la ruta de reparación de ADN, en la que participan de forma conjunta *BRCA1* y *BRCA2* (Apostolou y Fostira, 2013). Pero las mutaciones en los genes considerados de alta predisposición (*TP53*, *PTEN*, *CDH1*, *STK11*) son muy raras y en conjunto no representarían más del 4% de los casos (Thompson y Easton, 2004). Se

han descrito mutaciones raras en otros genes (*ATM*, *CHEK2*, *PALB2*, *BARD1*, *BRIP1*, *MRE11A*, *NBN*, *RAD50*, *RAD51C*, *XRCC2*, *RAD51D*) que confieren un riesgo moderado de padecer la enfermedad. Aún en el tiempo presente, se están identificando nuevas relaciones entre la enfermedad y mutaciones en genes que hasta la fecha no habían sido relacionados con el síndrome de CMOH como *Abraxas* (Solyom *et al.*, 2012) o *RBMX* (Adamson *et al.*, 2012), de los que se conocía su participación en la ruta de reparación de ADN en la que los genes *BRCA*s constituyen el núcleo central. En conclusión, cualquier gen que participa en esta ruta se convierte en un gen candidato de susceptibilidad a CMOH que debe ser investigado en las familias con test *BRCA* negativo.

Otra aproximación a la hora de dilucidar el espectro de predisposición del CMOH es la relativa a estudios de asociación alélica o GWAS en los que se relacionan variantes mucho más frecuentes en la población o SNPs con incrementos relativamente pequeños en el riesgo a padecer CMOH. Recientemente se han identificado nuevos 41 *loci* de susceptibilidad a cáncer de mama y en total se han descrito 89 relacionados con cáncer de mama y ovario (Michailidou *et al.*, 2013). En conjunto, todos los genes de susceptibilidad más todas estas variantes descritas en multitud de regiones genómicas podrían explicar hasta el 50% del espectro de predisposición. Se estima que más de 1000 *loci* podrían estar implicados en la predisposición a esta enfermedad (Michailidou *et al.*, 2013). También se ha descrito un alelo “protector”, el SNP (missense D302H), en una región codificante del gen *CASP8* que reduciría el riesgo de cáncer de mama (Rigas *et al.*, 2009).

La alta heterogeneidad genética del CMOH con hasta 18 genes de predisposición y 89 *loci* de susceptibilidad junto con las penetrancias variables de los diferentes alelos encontrados apoyan la existencia de un modelo poligénico según el cual la combinación de diferentes variantes genéticas de baja y moderada predisposición de distintos genes en un mismo paciente podrían incrementar de forma significativa el riesgo (Ponder *et al.*, 2005).

Uno de los problemas asociados a la alta heterogeneidad genética del CMOH en la práctica clínica radica en las limitaciones tanto económicas como técnicas a la hora de llevar a cabo estudios amplios que impliquen el rastreo mutacional de numerosos genes. Nuevas aproximaciones tecnológicas como la secuenciación masiva de ADN

están permitiendo el desarrollo de métodos coste-efectivos para detectar variantes en regiones muy amplias del genoma. Estas técnicas permiten analizar desde conjuntos de genes implicados en la ruta de reparación del ADN en la que ya se ha descrito su implicación con el riesgo a padecer CMOH hasta estudios de todo el exoma. Por ejemplo: secuenciando un conjunto de 507 genes relacionados con la ruta de reparación del ADN de doble hebra, se ha descrito la relación de mutaciones del gen *PPM1D* con la predisposición a CMOH (Ruark *et al.*, 2013) o mediante la secuenciación del exoma en pacientes con Anemia de Fanconi se han detectado mutaciones bialélicas en el gen *SLX4* (Shuste *et al.*, 2013). Así, también se detectaron mutaciones de *splicing* de *SLX4* en pacientes con fuerte historial de CMOH con un resultado previo negativo en el análisis de los genes *BRCA* (Bakker *et al.*, 2013).

## VII. Evaluación de la patogenicidad de variantes de ADN

Como hemos podido comprobar a lo largo de esta memoria de tesis, tan importante puede ser la detección de las variantes como la evaluación de su patogenicidad. Variantes detectadas en los genes de alta predisposición *BRCA1* y *BRCA2* pueden suponer un riesgo mayor o menor de desarrollar la enfermedad en función del efecto sobre la proteína. Datos procedentes de la historia familiar en combinación con estudios de cosegregación de la mutaciones con la enfermedad y diferentes modelos estadísticos ayudan a evaluar la patogenicidad de las variantes. Por ejemplo el método de análisis multifactorial integra datos epidemiológicos como: segregación familiar, co-ocurrencia en *trans* de otras mutaciones patogénicas, historia personal y familiar y características histopatológicas del tumor (Lindor *et al.*, 2012). No cabe duda de que las evidencias aportadas por los ensayos funcionales son la mejor prueba para evaluar el riesgo asociado a un cambio en el ADN.

Los resultados de nuestro estudio demuestran que la alteración proceso de *splicing* mediante variantes de ADN en *cis* de las secuencias de los genes *BRCA2* y *BRCA1* es el segundo mecanismo más frecuente de patogenicidad después de las mutaciones truncantes. Los ensayos funcionales de *splicing* y en concreto los realizados con minigenes híbridos son realmente útiles para evaluar la patogenicidad de determinadas variantes preseleccionadas. Todos los datos apuntan a que existe una correlación directa entre la penetrancia de una mutación (completa o parcial) y el tipo

de alteración de *splicing*. La importancia clínica de las mutaciones que alteran el proceso de *splicing* nos inducen a proponer un procedimiento rutinario de evaluación de cualquier variante de ADN detectada en una paciente dentro de un programa de prevención de CMOH u otro tipo de enfermedad hereditaria en las unidades de consejo genético. Antes de evaluar el efecto que tal cambio en el ADN pueda tener de forma directa en la proteína, debe realizarse un análisis bioinformático en relación a la posible alteración en el procesamiento del ARNm una vez realizada la transcripción, paso previo a la traducción a proteína en las células.

En caso de detectar *in silico* que la variante de ADN puede estar afectando a secuencias básicas o reguladoras esenciales para el correcto *splicing*, se debe proceder a realizar ensayos funcionales. La posibilidad de utilizar muestras del propio paciente es favorable a la realización de ensayos RT-PCR por ser rápidos y económicos. Sin embargo, la muestra biológica no siempre está disponible, convirtiendo los ensayos funcionales de *splicing* con minigenes en una herramienta valiosa para testar las variantes candidatas. Ambas pruebas deberían ser complementarias, sobre todo a la hora de identificar de forma precisa los transcritos resultantes y cuantificar los mismos. El proceso de identificación y cuantificación es esencial para evaluar la patogenicidad de una variante de ADN con impacto en el *splicing*. En muchas ocasiones este paso resulta imposible de realizar en muestras de paciente, debido a la interferencia del alelo *wild type* o la presencia de transcritos minoritarios difícilmente identificables, siendo necesaria la realización de ensayos funcionales con minigenes híbridos que contengan suficiente contexto genómico y estén contruidos en un vector reportero de *splicing* como pSAD, que mejora significativamente los resultados obtenidos respecto al vector pSPL3 u otros.

Un resultado bioinformático negativo requerirá una mayor profundidad en el análisis comparativo en relación a la conservación evolutiva de la secuencia afectada, así como de la posición respecto al exón-intrón particular, teniendo en cuenta estudios previos. Este es el caso de la variante que ha dado lugar a este trabajo, la c.8488-1G>A, en la que el estudio bioinformático era negativo porque la secuencia del sitio aceptor del exón 20 está muy poco conservada. Siempre y cuando se tengan indicios de que la variante puede alterar el *splicing* será necesario realizar ensayos funcionales. Si no fuese este el caso, la variante detectada deberá ser clasificada en función del efecto

directo que dicho cambio produce en la traducción a proteína: *nonsense*, *frameshift*, *missense*, sinónima o variante de efecto clínico desconocido (UVS).

Los indicadores biológicos de patogenicidad de una variante con impacto en el *splicing* son el alto porcentaje de transcritos anómalos y el efecto sobre la proteína (ya sea truncándola por un cambio en la pauta de lectura o bien delecionando dominios funcionales esenciales). La correcta caracterización y cuantificación de los transcritos es clave para determinar la penetrancia de una variante. Un porcentaje alto de transcritos que inducen un codón de parada prematuro o alteraciones (inserciones o deleciones) generadas por el proceso aberrante de *splicing* en regiones con una función esencial para la proteína deberían considerarse de alta penetrancia. Sin embargo, la penetrancia y el riesgo asociado a una variante de ADN con efecto parcial en el *splicing* sólo podrían ser evaluados mediante estudios epidemiológicos y el análisis estadístico multifactorial integrado, que incluye parámetros como segregación familiar, co-ocurrencia en *trans* de otras mutaciones patogénicas, características histopatológicas del tumor e historia personal y familiar (Lindor *et al.*, 2012). Con la conjunción de estos datos y la evidencia funcional, estas variantes podrían clasificarse según el IARC (Grupo de trabajo de variantes genéticas no clasificadas; Tabla 1), como C3 o incierta, en el sentido de que las evidencias recogidas indican que tienen alguna significancia clínica o C4, probablemente patogénica (Plon *et al.*, 2008).

**Tabla 1. Clasificación de mutaciones según el IARC.** Clasificación de grado de patogenicidad de una mutación según el IARC.

Clasificación IARC		
Clase	Descripción	Probabilidad de ser Patogénica
5	Definitivamente Patogénica	> 0,99
4	Probablemente Patogénica	0,95-0,99
3	Incierta	0,05-0,949
2	Probablemente no patogénica o de pequeña significancia clínica	0,001-0,049
1	No patogénica	< 0,001

## VIII. Aplicaciones y perspectivas de los estudios de *splicing*

### VIII.1. Terapias de *splicing*

El conocimiento acerca de las bases moleculares por las que se producen alteraciones en el proceso de *splicing* está promoviendo el desarrollo de diferentes

aproximaciones terapéuticas. Métodos que corrijan alteraciones tan frecuentes como la creación de sitios de *splicing* o la eliminación de los naturales podrían ofrecer terapias válidas para múltiples enfermedades hereditarias (Kole *et al* 2012). Por ejemplo, se ha determinado que la mutación más frecuente encontrada en pacientes con síndrome de envejecimiento prematuro o Progeria de Hutchinson-Gilford, activa un sitio aceptor dentro de uno de los exones del gen de la laminina A, provocando una delección de 50 aminoácidos. Una aproximación a terapias de este tipo propone el uso de oligonucleótidos complementarios al nuevo sitio aceptor, evitando que éste sea reconocido. Los resultados con modelos de ratones con Progeria han sido prometedores (Osorio *et al.*, 2011).

Modelos terapéuticos parecidos han sido ensayados en ratones transgénicos que expresan *SMN2* humano cuyo objetivo es la corrección futura del *splicing* del gen *SMN2* en pacientes con atrofia muscular espinal que tienen inactivado el gen *SMN1*. El exón 7 del gen *SMN2* sólo difiere en un nucleótido del correspondiente exón de *SMN1* (5 nt contando todo el gen). Sin embargo, esta variante impide el reconocimiento del exón 7 del gen *SMN2*, por la eliminación de un *enhancer* de *splicing* y/o la creación de un silenciador. El uso de oligonucleótidos complementarios a la región donde se encuentra la variante promueve la inclusión del exón 7 de *SMN2* en el ARN maduro de las motoneuronas de los ratones. Los oligos son administrados a través de una inyección intracerebroventricular a los ratones, que además desarrollan necrosis en la cola y en las orejas, simulando las necrosis que se observan en tejidos distales de algunos lactantes que desarrollan atrofia muscular espinal. Tras la inyección, se detectaron altos niveles de ARNm de *SMN2* con la inclusión del exón 7 en las motoneuronas de la médula espinal. Sorprendentemente, también se solventó la necrosis de las orejas y cola. Por lo tanto, se trata de una prometedora terapia en pacientes con atrofia muscular espinal (Hua *et al.*, 2010).

Por otro lado, también se han desarrollado terapias génicas basadas en provocar reacciones de *splicing* para que determinadas secuencias que incluyen mutaciones *frameshift* y *nonsense* no sean incorporadas en el transcrito maduro. Estas reacciones de *splicing* inducidas de forma artificial deben producir transcritos que mantengan la pauta de lectura. Es un tipo de terapia génica que se ha probado en modelos de ratones de Distrofia Muscular de Duchenne, donde se ligaron oligos antisentido en los

sitios de *splicing* del exón 23 del gen DMD que contenía una mutación *nonsense*. La falta del exón 23 en el ARNm maduro no altera el marco de lectura y en consecuencia, se ha comprobado una atenuación importante de fenotipo (Mann *et al.*, 2001; Lee y Dougherty, 2012).

La complejidad de estas técnicas recae en la estabilidad de los oligonucleótidos administrados de forma periférica y su inclusión en el núcleo celular. Modificaciones recientes en la estructura química de los oligonucleótidos utilizados han propiciado resultados prometedores mediante su administración periférica (Hua *et al.*, 2011), posibilitando el inicio de ensayos clínicos.

Otro tipo de alteraciones que implican cambios globales en los patrones de *splicing* que derivan en diversas alteraciones en la función celular, como la división. Por ejemplo, en las neoplasias de células hematopoyéticas como los síndromes mielodisplásicos y la leucemia crónica se establecen peores pronósticos y resistencia a quimioterapia cuando se detectan mutaciones en el factor de *splicing* SF3b1 (Quesada *et al.*, 2011). SF3b1 es una subunidad del complejo ribonucleoprotéico U2 cuya función es la de reconocer el punto de ramificación. Recientemente, se han propuesto posibilidades terapéuticas basadas en compuestos químicos derivados de la fermentación bacteriana que se unen a SF3b1 y poseen actividad antitumoral, inhibiendo la división celular de células cancerosas en cultivo y restringiendo la progresión tumoral en modelos animales (Bonnal *et al.*, 2012). Estudios actuales parecen indicar que estas sustancias afectan a la precisión con la que el espliceosoma reconoce los sitios de *splicing* (Corrionero *et al.*, 2011), de tal manera que no se inhibe el proceso de *splicing* si no que se induce la aparición de diferentes transcritos a través de *splicing* alternativo,

## VIII.2. Nuevas tecnologías de análisis global de patrones de *splicing*

La revolución en tecnologías de alto rendimiento desarrolladas en los últimos años está permitiendo estudios del transcriptoma completo, donde no sólo se puede evaluar la alteración en el proceso de *splicing* causada por mutaciones en determinados genes, sino también cambios globales en los patrones de *splicing* como los detectados en diversos tipos celulares tumorales. Estas metodologías analíticas

pueden revelar nuevos biomarcadores para el diagnóstico y tratamiento de enfermedades.

La tecnología de *microarrays* era, hasta hace muy poco tiempo, la única alternativa para realizar estudios clínicos con cantidades limitadas de ARN procedentes de muestras sanguíneas, biopsias o células. *Microarrays* como el *Human Exon Array 1.0 ST* de *Affymetrix* contienen más de 1,2 millones de sondas dirigidas a cada uno de los exones conocidos del genoma, proporcionando excelentes perfiles de transcripción. A la vez permite identificar eventos de *splicing* alternativo derivados de un proceso patológico. Nuevos diseños de *microarrays* como el *GeneChip® Human Transcriptome Array 2.0* han aumentado considerablemente el número de sondas hasta la asombrosa cifra de 7,2 millones, pudiendo identificar la mayoría de los eventos de *splicing* conocidos en el transcriptoma, ya que a las sondas propias de un *exon array* se unen sondas que reconocen las uniones exón-exón anotadas hasta el momento (*junction array*; Xu *et al.*, 2011). Aunque la tecnología de los *microarrays* es una magnífica herramienta, han surgido nuevas aproximaciones basadas en la secuenciación masiva (NGS) de ARN.

La secuenciación masiva de ARN requiere un mayor cantidad de muestra pero ofrece la posibilidad de detectar transcritos no identificados previamente o que están en muy baja proporción. Otra ventaja derivada de la secuenciación es la detección de variantes en la secuencia de nucleótidos. La disminución progresiva en el coste de la secuenciación, el multiplexado de las muestras y el aumento en la longitud de lectura son factores en constante evolución que favorecen su utilización. No obstante, estudios comparativos recientes muestran un alto nivel de concordancia entre los datos obtenidos con *microarrays* y secuenciación masiva de ARN (Raghavachari *et al.*, 2012).

Por lo tanto, el desarrollo de modelos que integren las grandes cantidades de datos procedentes de herramientas de alto rendimiento como *microarrays* y secuenciación masiva de ARN con ensayos funcionales de *splicing* contribuirá de forma significativa a un conocimiento más profundo de los diferentes mecanismos de regulación de *splicing* que generan diferentes transcritos bajo condiciones específicas.

En conclusión, los resultados obtenidos en este estudio aportan luz sobre el mecanismo de patogenicidad relacionado con la alteración del *splicing* de las variantes de ADN estudiadas. En este sentido, podemos afirmar que una importante fracción de variantes de ADN está asociada con *splicing* anómalo del gen *BRCA2* y éste podría incluso ser considerado como el principal mecanismo de inactivación génica en variantes de ADN no clasificadas.

El desarrollo de la medicina genómica personalizada ligado al uso de nuevas tecnologías de análisis de ADN como la ultrasecuenciación, está democratizando la realización de estudios genéticos dirigidos tanto para la prevención como el diagnóstico de enfermedades hereditarias. La ingente cantidad de información reportada en estos estudios necesita de herramientas funcionales que aporten información esencial para distinguir entre variantes neutras o patogénicas. Las nuevas herramientas desarrolladas en este trabajo para la realización de ensayos funcionales de *splicing*, nuevo vector pSAD y los minigenes construidos deberían ser incorporadas de rutina como instrumentos de apoyo en los estudios genéticos ya que permiten evaluar cualquier variante de ADN en relación con el proceso de *splicing*.

Por último y no menos importante, los estudios funcionales de *splicing* mediante minigenes híbridos permiten esclarecer los complejos mecanismos de regulación de la expresión génica en los eucariotas. De esta forma, contribuyen a establecer modelos que mejoran el conocimiento de las reglas para la definición de los exones.



# CONCLUSIONES



- I. La mutación c.8488-1G>A detectada en una paciente del Programa de Prevención del Cáncer Hereditario de la Junta de Castilla y León altera drásticamente el proceso de *splicing* tanto en RT-PCR de linfocitos como mediante el uso de minigenes híbridos. Por lo tanto, se puede clasificar como mutación patogénica.
  
- II. Se ha diseñado un nuevo plásmido reportero de *splicing*, pSAD, que mejora a los previamente disponibles en varios aspectos:
  - Reducción del tamaño del intrón del vector en 1.2 kb.
  - Mayor capacidad de selección por la introducción de un segundo marcador (LacZ).
  - Mayor versatilidad, con 27 sitios de restricción en el sitio de clonación múltiple.
  - Fortalecimiento del sitio aceptor del exón constitutivo 2 del vector.
  - Eliminación de sitios crípticos para evitar reacciones cruzadas de *splicing*.
  
- III. Los ensayos funcionales con minigenes híbridos en el vector pSAD presentan las siguientes ventajas:
  - Construcción de grandes minigenes que mantienen el contexto genómico y reproducen fielmente las reacciones de *splicing* fisiológicas, con la ventaja de no necesitar material biológico de los pacientes y permitir el análisis aislado del efecto de un único alelo.
  - Análisis simple y reproducible del impacto de las variantes de ADN sobre el *splicing*, que son introducidas mediante mutagénesis dirigida.
  - La electroforesis capilar de los transcritos es una técnica de alta resolución que permite la detección de transcritos minoritarios y una cuantificación precisa de todas las isoformas.
  
- IV. Los programas bioinformáticos de análisis de *splicing* son, por regla general, instrumentos muy útiles de predicción. Las predicciones relativas a la eliminación de sitios canónicos de *splicing* o la creación de nuevos sitios alternativos son mucho más fiables que las predicciones de otros elementos reguladores de *splicing* (*enhancers* y silenciadores). Los estudios funcionales de *splicing* contribuirán a mejorar la precisión de estas herramientas.

- V.** Un total de 33 mutaciones de ADN descritas en pacientes de CMOH en las bases de datos internacionales (BIC, LOVD, UMD) de las 80 analizadas (41%) alteraron el procesamiento de pre-ARNm del gen *BRCA2* mediante ensayos funcionales de *splicing* de los minigenes con los exones 19 a 27 de *BRCA2*.
- VI.** Cualquier tipo de mutación (*frameshift*, *nonsense*, *missense*, *sinónimas* e intrónicas) es capaz de alterar el *splicing*, pudiendo repercutir en su penetrancia y expresividad. La reclasificación de variantes de efecto fisiológico desconocido de *BRCA1* y *BRCA2* como mutaciones de *splicing* deletéreas repercute en el consejo genético, incrementando el número de familias CMOH que se pueden beneficiar de los protocolos de prevención indicados.
- VII.** Los elementos reguladores actúan cooperativamente para promover el reconocimiento de un exón, como por ejemplo las variantes c.8972G>A y c.9006A>T del exón 23 de *BRCA2* o las microdeleciones artificiales del intrón 19.
- VIII.** La incorporación de los estudios de *splicing* como pruebas complementarias en el diagnóstico genético de enfermedades hereditarias permitirá esclarecer el papel del creciente número de variantes de efecto clínico desconocido. El conocimiento preciso del mecanismo patogénico de una mutación permitirá el desarrollo de nuevas aproximaciones terapéuticas para enfermedades de origen genético. Asimismo, un análisis profundo de ambos genes podría generar cambios sustanciales en el modelo general de predisposición del CMOH.

# *BIBLIOGRAFÍA*



- Adamson, B., A. Smogorzewska, F. D. Sigoillot, R. W. King and S. J. Elledge (2012). "A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response." *Nat Cell Biol* **14**(3): 318-328.
- Alekseyenko, A. V., N. Kim and C. J. Lee (2007). "Global analysis of exon creation versus loss and the role of alternative splicing in 17 vertebrate genomes." *RNA* **13**(5): 661-670.
- Allo, M., V. Buggiano, J. P. Fededa, E. Petrillo, I. Schor, M. de la Mata, E. Agirre, M. Plass, E. Eyras, S. A. Elela, R. Klinck, B. Chabot and A. R. Kornblihtt (2009). "Control of alternative splicing through siRNA-mediated transcriptional gene silencing." *Nat Struct Mol Biol* **16**(7): 717-724.
- Amit, M., M. Donyo, D. Hollander, A. Goren, E. Kim, S. Gelfman, G. Lev-Maor, D. Burstein, S. Schwartz, B. Postolsky, T. Pupko and G. Ast (2012). "Differential GC content between exons and introns establishes distinct strategies of splice-site recognition." *Cell Rep* **1**(5): 543-556.
- Anczukow, O., M. Buisson, M. Leone, C. Coutanson, C. Lasset, A. Calender, O. M. Sinilnikova and S. Mazoyer (2012). "BRCA2 deep intronic mutation causing activation of a cryptic exon: opening toward a new preventive therapeutic strategy." *Clin Cancer Res* **18**(18): 4903-4909.
- Apostolou, P. and F. Fostira (2013). "Hereditary breast cancer: the era of new susceptibility genes." *Biomed Res Int* **2013**: 747318.
- Ast, G. (2004). "How did alternative splicing evolve?" *Nat Rev Genet* **5**(10): 773-782.
- Athanasiadis, A., A. Rich and S. Maas (2004). "Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome." *PLoS Biol* **2**(12): e391.
- Aubol, B. E., B. Nolen, J. Shaffer, G. Ghosh and J. A. Adams (2003). "Novel destabilization of nucleotide binding by the gamma phosphate of ATP in the yeast SR protein kinase Sky1p." *Biochemistry* **42**(44): 12813-12820.
  
- Bakker, J. L., S. E. van Mil, G. Crossan, N. Sabbaghian, K. De Leeneer, B. Poppe, M. Adank, H. Gille, H. Verheul, H. Meijers-Heijboer, J. P. de Winter, K. Claes, M. Tischkowitz and Q. Waisfisz (2013). "Analysis of the novel fanconi anemia gene SLX4/FANCP in familial breast cancer cases." *Hum Mutat* **34**(1): 70-73.
- Bakkour, N., Y. L. Lin, S. Maire, L. Ayadi, F. Mahuteau-Betzer, C. H. Nguyen, C. Mettling, P. Portales, D. Grierson, B. Chabot, P. Jeanteur, C. Branlant, P. Corbeau and J. Tazi (2007). "Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance." *PLoS Pathog* **3**(10): 1530-1539.
- Baralle, D., A. Lucassen and E. Buratti (2009). "Missed threads. The impact of pre-mRNA splicing defects on clinical practice." *EMBO Rep* **10**(8): 810-816.
- Barash, Y., J. A. Calarco, W. Gao, Q. Pan, X. Wang, O. Shai, B. J. Blencowe and B. J. Frey (2010). "Deciphering the splicing code." *Nature* **465**(7294): 53-59.
- Barreau, C., L. Paillard, A. Mereau and H. B. Osborne (2006). "Mammalian CELF/Bruno-like RNA-binding proteins: molecular characteristics and biological functions." *Biochimie* **88**(5): 515-525.
- Bass, B. L. (2002). "RNA editing by adenosine deaminases that act on RNA." *Annu Rev Biochem* **71**: 817-846.
- Batsche, E., M. Yaniv and C. Muchardt (2006). "The human SWI/SNF subunit Brm is a regulator of alternative splicing." *Nat Struct Mol Biol* **13**(1): 22-29.

- Berget, S. M. (1995). "Exon recognition in vertebrate splicing." *J Biol Chem* **270**(6): 2411-2414.
- Berget, S. M., C. Moore and P. A. Sharp (1977). "Spliced segments at the 5' terminus of adenovirus 2 late mRNA." *Proc Natl Acad Sci U S A* **74**(8): 3171-3175.
- Beroud, C., G. Collod-Beroud, C. Boileau, T. Soussi and C. Junien (2000). "UMD (Universal mutation database): a generic software to build and analyze locus-specific databases." *Hum Mutat* **15**(1): 86-94.
- Beyer, A. L. and Y. N. Osheim (1988). "Splice site selection, rate of splicing, and alternative splicing on nascent transcripts." *Genes Dev* **2**(6): 754-765.
- Bonatti, F., C. Pepe, M. Tancredi, G. Lombardi, P. Aretini, E. Sensi, E. Falaschi, G. Cipollini, G. Bevilacqua and M. A. Caligo (2006). "RNA-based analysis of BRCA1 and BRCA2 gene alterations." *Cancer Genet Cytogenet* **170**(2): 93-101.
- Bonnal, S., L. Vigevani and J. Valcarcel (2012). "The spliceosome as a target of novel antitumour drugs." *Nat Rev Drug Discov* **11**(11): 847-859.
- Bonnet, C., S. Krieger, M. Vezain, A. Rousselin, I. Tournier, A. Martins, P. Berthet, A. Chevrier, C. Dugast, V. Layet, A. Rossi, R. Lidereau, T. Frebourg, A. Hardouin and M. Tosi (2008). "Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene." *J Med Genet* **45**(7): 438-446.
- Brooks, S. C., E. R. Locke and H. D. Soule (1973). "Estrogen receptor in a human cell line (MCF-7) from breast carcinoma." *J Biol Chem* **248**(17): 6251-6253.
- Buratti, E. and F. E. Baralle (2004). "Influence of RNA secondary structure on the pre-mRNA splicing process." *Mol Cell Biol* **24**(24): 10505-10514.
- Buratti, E. and F. E. Baralle (2010). "The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation." *RNA Biol* **7**(4): 420-429.
- Buratti, E., M. Baralle and F. E. Baralle (2006). "Defective splicing, disease and therapy: searching for master checkpoints in exon definition." *Nucleic Acids Res* **34**(12): 3494-3510.
- Burn, T. C., T. D. Connors, K. W. Klinger and G. M. Landes (1995). "Increased exon-trapping efficiency through modifications to the pSPL3 splicing vector." *Gene* **161**(2): 183-187.
- Busch, A. and K. J. Hertel (2012). "Evolution of SR protein and hnRNP splicing regulatory factors." *Wiley Interdiscip Rev RNA* **3**(1): 1-12.
- Cartegni, L. and A. R. Krainer (2002). "Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1." *Nat Genet* **30**(4): 377-384.
- Cartegni, L., J. Wang, Z. Zhu, M. Q. Zhang and A. R. Krainer (2003). "ESEfinder: A web resource to identify exonic splicing enhancers." *Nucleic Acids Res* **31**(13): 3568-3571.
- Chalfant, C. E., K. Rathman, R. L. Pinkerman, R. E. Wood, L. M. Obeid, B. Ogretmen and Y. A. Hannun (2002). "De novo ceramide regulates the alternative splicing of caspase 9 and Bcl-x in A549 lung adenocarcinoma cells. Dependence on protein phosphatase-1." *J Biol Chem* **277**(15): 12587-12595.
- Chen, M. and J. L. Manley (2009). "Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches." *Nat Rev Mol Cell Biol* **10**(11): 741-754.

- Chen, X., T. T. Truong, J. Weaver, B. A. Bove, K. Cattie, B. A. Armstrong, M. B. Daly and A. K. Godwin (2006). "Intronic alterations in BRCA1 and BRCA2: effect on mRNA splicing fidelity and expression." *Hum Mutat* **27**(5): 427-435.
- Cho, S. H., F. Seymen, K. E. Lee, S. K. Lee, Y. S. Kweon, K. J. Kim, S. E. Jung, S. J. Song, M. Yildirim, M. Bayram, E. B. Tuna, K. Gencay and J. W. Kim (2012). "Novel FAM20A mutations in hypoplastic amelogenesis imperfecta." *Hum Mutat* **33**(1): 91-94.
- Choi, Y. D. and G. Dreyfuss (1984). "Isolation of the heterogeneous nuclear RNA-ribonucleoprotein complex (hnRNP): a unique supramolecular assembly." *Proc Natl Acad Sci U S A* **81**(23): 7471-7475.
- Chow, L. T., J. M. Roberts, J. B. Lewis and T. R. Broker (1977). "A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids." *Cell* **11**(4): 819-836.
- Claes, K., B. Poppe, E. Machackova, I. Coene, L. Foretova, A. De Paepe and L. Messiaen (2003). "Differentiating pathogenic mutations from polymorphic alterations in the splice sites of BRCA1 and BRCA2." *Genes Chromosomes Cancer* **37**(3): 314-320.
- Colapietro, P., C. Gervasini, F. Natacci, L. Rossi, P. Riva and L. Larizza (2003). "NF1 exon 7 skipping and sequence alterations in exonic splice enhancers (ESEs) in a neurofibromatosis 1 patient." *Hum Genet* **113**(6): 551-554.
- Corriero, A., B. Minana and J. Valcarcel (2011). "Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A." *Genes Dev* **25**(5): 445-459.
- Cramer, P., C. G. Pesce, F. E. Baralle and A. R. Kornblihtt (1997). "Functional association between promoter structure and transcript alternative splicing." *Proc Natl Acad Sci U S A* **94**(21): 11456-11460.
- De Conti, L., M. Baralle and E. Buratti (2012). "Exon and intron definition in pre-mRNA splicing." *Wiley Interdiscip Rev RNA*.
- De Klein, A., P. H. Riegman, E. K. Bijlsma, A. Helder, M. Muijtjens, M. A. den Bakker, C. J. Avezaat and E. C. Zwarthoff (1998). "A G-->A transition creates a branch point sequence and activation of a cryptic exon, resulting in the hereditary disorder neurofibromatosis 2." *Hum Mol Genet* **7**(3): 393-398.
- De la Mata, M., C. R. Alonso, S. Kadener, J. P. Fededa, M. Blaustein, F. Pelisch, P. Cramer, D. Bentley and A. R. Kornblihtt (2003). "A slow RNA polymerase II affects alternative splicing in vivo." *Mol Cell* **12**(2): 525-532.
- Desmet, F. O., D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres and C. Beroud (2009). "Human Splicing Finder: an online bioinformatics tool to predict splicing signals." *Nucleic Acids Res* **37**(9): e67.
- Disset, A., C. F. Bourgeois, N. Benmalek, M. Claustres, J. Stevenin and S. Tuffery-Giraud (2006). "An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements." *Hum Mol Genet* **15**(6): 999-1013.
- Dominski, Z. and R. Kole (1991). "Selection of splice sites in pre-mRNAs with short internal exons." *Mol Cell Biol* **11**(12): 6075-6083.
- Dowhan, D. H., E. P. Hong, D. Auboeuf, A. P. Dennis, M. M. Wilson, S. M. Berget and B. W. O'Malley (2005). "Steroid hormone receptor coactivation and alternative RNA

- splicing by U2AF65-related proteins CAPERalpha and CAPERbeta." *Mol Cell* **17**(3): 429-439.
- Dowling, D., S. Nasr-Esfahani, C. H. Tan, K. O'Brien, J. L. Howard, D. A. Jans, D. F. Purcell, C. M. Stoltzfus and S. Sonza (2008). "HIV-1 infection induces changes in expression of cellular splicing factors that regulate alternative viral splicing and virus production in macrophages." *Retrovirology* **5**: 18.
  - Edery, P., C. Marcaillou, M. Sahbatou, A. Labalme, J. Chastang, R. Touraine, E. Tubacher, F. Senni, M. B. Bober, S. Nampoothiri, P. S. Jouk, E. Steichen, S. Berland, A. Toutain, C. A. Wise, D. Sanlaville, F. Rousseau, F. Clerget-Darpoux and A. L. Leutenegger (2011). "Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA." *Science* **332**(6026): 240-243.
  - Elgar, G. and T. Vavouri (2008). "Tuning in to the signals: noncoding sequence conservation in vertebrate genomes." *Trends Genet* **24**(7): 344-352.
  - Erkkö, H., B. Xia, J. Nikkila, J. Schleutker, K. Syrjäkoski, A. Mannermaa, A. Kallioniemi, K. Pylkas, S. M. Karppinen, K. Rapakko, A. Miron, Q. Sheng, G. Li, H. Mattila, D. W. Bell, D. A. Haber, M. Grip, M. Reiman, A. Jukkola-Vuorinen, A. Mustonen, J. Kere, L. A. Aaltonen, V. M. Kosma, V. Kataja, Y. Soini, R. I. Drapkin, D. M. Livingston and R. Winqvist (2007). "A recurrent mutation in PALB2 in Finnish cancer families." *Nature* **446**(7133): 316-319.
  - Esteban-Cardenosa, E., M. Duran, M. Infante, E. Velasco and C. Miner (2004). "High-throughput mutation detection method to scan BRCA1 and BRCA2 based on heteroduplex analysis by capillary array electrophoresis." *Clin Chem* **50**(2): 313-320.
  - Fanale, D., V. Amodeo, L. R. Corsini, S. Rizzo, V. Bazan and A. Russo (2012). "Breast cancer genome-wide association studies: there is strength in numbers." *Oncogene* **31**(17): 2121-2128.
  - Faustino, N. A. and T. A. Cooper (2003). "Pre-mRNA splicing and human disease." *Genes Dev* **17**(4): 419-437.
  - Fokkema, I. F., J. T. den Dunnen and P. E. Taschner (2005). "LOVD: easy creation of a locus-specific sequence variation database using an "LSDB-in-a-box" approach." *Hum Mutat* **26**(2): 63-68.
  - Foulkes, W. D. (2008). "Inherited susceptibility to common cancers." *N Engl J Med* **359**(20): 2143-2153.
  - Fox-Walsh, K. L., Y. Dou, B. J. Lam, S. P. Hung, P. F. Baldi and K. J. Hertel (2005). "The architecture of pre-mRNAs affects mechanisms of splice-site pairing." *Proc Natl Acad Sci U S A* **102**(45): 16176-16181.
  - Gaildrat, P., S. Krieger, D. Di Giacomo, J. Abdat, F. Revillion, S. Caputo, D. Vaur, E. Jamard, E. Bohers, D. Ledemenev, J. P. Peyrat, C. Houdayer, E. Rouleau, R. Lidereau, T. Frebourg, A. Hardouin, M. Tosi and A. Martins (2012). "Multiple sequence variants of BRCA2 exon 7 alter splicing regulation." *J Med Genet* **49**(10): 609-617.
  - Gal-Mark, N., S. Schwartz, O. Ram, E. Eyraş and G. Ast (2009). "The pivotal roles of TIA proteins in 5' splice-site selection of alu exons and across evolution." *PLoS Genet* **5**(11): e1000717.

- Ghigna, C., S. Giordano, H. Shen, F. Benvenuto, F. Castiglioni, P. M. Comoglio, M. R. Green, S. Riva and G. Biamonti (2005). "Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene." *Mol Cell* **20**(6): 881-890.
- Gilbert, W. (1978). "Why genes in pieces?" *Nature* **271**(5645): 501.
- Graveley, B. R., A. N. Brooks, J. W. Carlson, M. O. Duff, J. M. Landolin, L. Yang, C. G. Artieri, M. J. van Baren, N. Boley, B. W. Booth, J. B. Brown, L. Cherbas, C. A. Davis, A. Dobin, R. Li, W. Lin, J. H. Malone, N. R. Mattiuzzo, D. Miller, D. Sturgill, B. B. Tuch, C. Zaleski, D. Zhang, M. Blanchette, S. Dudoit, B. Eads, R. E. Green, A. Hammonds, L. Jiang, P. Kapranov, L. Langton, N. Perrimon, J. E. Sandler, K. H. Wan, A. Willingham, Y. Zhang, Y. Zou, J. Andrews, P. J. Bickel, S. E. Brenner, M. R. Brent, P. Cherbas, T. R. Gingeras, R. A. Hoskins, T. C. Kaufman, B. Oliver and S. E. Celniker (2011). "The developmental transcriptome of *Drosophila melanogaster*." *Nature* **471**(7339): 473-479.
- Gudmundsdottir, K. and A. Ashworth (2006). "The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability." *Oncogene* **25**(43): 5864-5874.
- Gui, J. F., W. S. Lane and X. D. Fu (1994). "A serine kinase regulates intracellular localization of splicing factors in the cell cycle." *Nature* **369**(6482): 678-682.
- Gutierrez-Enriquez, S., J. Balmana, M. Baiget and O. Diez (2008). "Detection of the CHEK2 1100delC mutation by MLPA BRCA1/2 analysis: a worthwhile strategy for its clinical applicability in 1100delC low-frequency populations?" *Breast Cancer Res Treat* **107**(3): 455-457.
- Hakem, R., J. L. de la Pompa, C. Sirard, R. Mo, M. Woo, A. Hakem, A. Wakeham, J. Potter, A. Reitmair, F. Billia, E. Firpo, C. C. Hui, J. Roberts, J. Rossant and T. W. Mak (1996). "The tumor suppressor gene *Brca1* is required for embryonic cellular proliferation in the mouse." *Cell* **85**(7): 1009-1023.
- Hall, J. M., M. K. Lee, B. Newman, J. E. Morrow, L. A. Anderson, B. Huey and M. C. King (1990). "Linkage of early-onset familial breast cancer to chromosome 17q21." *Science* **250**(4988): 1684-1689.
- Hanahan, D. (1983). "Studies on transformation of *Escherichia coli* with plasmids." *J Mol Biol* **166**(4): 557-580.
- He, H., S. Liyanarachchi, K. Akagi, R. Nagy, J. Li, R. C. Dietrich, W. Li, N. Sebastian, B. Wen, B. Xin, J. Singh, P. Yan, H. Alder, E. Haan, D. Wiczorek, B. Albrecht, E. Puffenberger, H. Wang, J. A. Westman, R. A. Padgett, D. E. Symer and A. de la Chapelle (2011). "Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I." *Science* **332**(6026): 238-240.
- Hegele, A., A. Kamburov, A. Grossmann, C. Surlis, S. Wowro, M. Weimann, C. L. Will, V. Pena, R. Luhrmann and U. Stelzl (2012). "Dynamic protein-protein interaction wiring of the human spliceosome." *Mol Cell* **45**(4): 567-580.
- Hicks, M. J., B. J. Lam and K. J. Hertel (2005). "Analyzing mechanisms of alternative pre-mRNA splicing using in vitro splicing assays." *Methods* **37**(4): 306-313.
- Hiller, M., Z. Zhang, R. Backofen and S. Stamm (2007). "Pre-mRNA secondary structures influence exon recognition." *PLoS Genet* **3**(11): e204.
- House, A. E. and K. W. Lynch (2006). "An exonic splicing silencer represses spliceosome assembly after ATP-dependent exon recognition." *Nat Struct Mol Biol* **13**(10): 937-944.

- Howlett, N. G., T. Taniguchi, S. Olson, B. Cox, Q. Waisfisz, C. De Die-Smulders, N. Persky, M. Grompe, H. Joenje, G. Pals, H. Ikeda, E. A. Fox and A. D. D'Andrea (2002). "Biallelic inactivation of BRCA2 in Fanconi anemia." *Science* **297**(5581): 606-609.
- Hua, Y., K. Sahashi, G. Hung, F. Rigo, M. A. Passini, C. F. Bennett and A. R. Krainer (2010). "Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model." *Genes Dev* **24**(15): 1634-1644.
- Hua, Y., K. Sahashi, F. Rigo, G. Hung, G. Horev, C. F. Bennett and A. R. Krainer (2011). "Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model." *Nature* **478**(7367): 123-126.
- Huang, L. and M. F. Wilkinson (2012). "Regulation of nonsense-mediated mRNA decay." *Wiley Interdiscip Rev RNA* **3**(6): 807-828.
- Hubbard, T., D. Barker, E. Birney, G. Cameron, Y. Chen, L. Clark, T. Cox, J. Cuff, V. Curwen, T. Down, R. Durbin, E. Eyra, J. Gilbert, M. Hammond, L. Humniecki, A. Kasprzyk, H. Lehvaslaiho, P. Lijnzaad, C. Melsopp, E. Mongin, R. Pettett, M. Pocock, S. Potter, A. Rust, E. Schmidt, S. Searle, G. Slater, J. Smith, W. Spooner, A. Stabenau, J. Stalker, E. Stupka, A. Ureta-Vidal, I. Vastrik and M. Clamp (2002). "The Ensembl genome database project." *Nucleic Acids Res* **30**(1): 38-41.
- Ibrahim, E. C., T. D. Schaal, K. J. Hertel, R. Reed and T. Maniatis (2005). "Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers." *Proc Natl Acad Sci U S A* **102**(14): 5002-5007.
- Infante, M., M. Duran, A. Acedo, L. Perez-Cabornero, D. J. Sanz, M. Garcia-Gonzalez, E. Beristain, E. Esteban-Cardenosa, M. de la Hoya, A. Teule, A. Vega, M. I. Tejada, E. Lastra, C. Miner and E. A. Velasco (2010). "BRCA1 5272-1G>A and BRCA2 5374delTATG are founder mutations of high relevance for genetic counselling in breast/ovarian cancer families of Spanish origin." *Clin Genet* **77**(1): 60-69.
- Jeninga, E. H., M. de Vroede, N. Hamers, J. M. Breur, N. M. Verhoeven-Duif, R. Berger and E. Kalkhoven (2012). "A Patient with Congenital Generalized Lipodystrophy Due To a Novel Mutation in BSCL2: Indications for Secondary Mitochondrial Dysfunction." *JIMD Rep* **4**: 47-54.
- Jones, H. W., Jr., V. A. McKusick, P. S. Harper and K. D. Wu (1971). "George Otto Gey. (1899-1970). The HeLa cell and a reappraisal of its origin." *Obstet Gynecol* **38**(6): 945-949.
- Jurica, M. S. and M. J. Moore (2003). "Pre-mRNA splicing: awash in a sea of proteins." *Mol Cell* **12**(1): 5-14.
- Kaida, D., T. Schneider-Poetsch and M. Yoshida (2012). "Splicing in oncogenesis and tumor suppression." *Cancer Sci* **103**(9): 1611-1616.
- Karni, R., E. de Stanchina, S. W. Lowe, R. Sinha, D. Mu and A. R. Krainer (2007). "The gene encoding the splicing factor SF2/ASF is a proto-oncogene." *Nat Struct Mol Biol* **14**(3): 185-193.
- Kashima, T. and J. L. Manley (2003). "A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy." *Nat Genet* **34**(4): 460-463.
- Ke, S., X. H. Zhang and L. A. Chasin (2008). "Positive selection acting on splicing motifs reflects compensatory evolution." *Genome Res* **18**(4): 533-543.

- Keren, H., G. Lev-Maor and G. Ast (2010). "Alternative splicing and evolution: diversification, exon definition and function." Nat Rev Genet **11**(5): 345-355.
- Kim, E., A. Goren and G. Ast (2008). "Alternative splicing: current perspectives." Bioessays **30**(1): 38-47.
- Kim, Y. J. and H. S. Kim (2012). "Alternative splicing and its impact as a cancer diagnostic marker." Genomics Inform **10**(2): 74-80.
- Kole, R., A. R. Krainer and S. Altman (2012). "RNA therapeutics: beyond RNA interference and antisense oligonucleotides." Nat Rev Drug Discov **11**(2): 125-140.
- Konig, H., N. Matter, R. Bader, W. Thiele and F. Muller (2007). "Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation." Cell **131**(4): 718-729.
- Krawczak, M., N. S. Thomas, B. Hundrieser, M. Mort, M. Wittig, J. Hampe and D. N. Cooper (2007). "Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing." Hum Mutat **28**(2): 150-158.
- Kurahashi, H., K. Takami, T. Oue, T. Kusafuka, A. Okada, A. Tawa, S. Okada and I. Nishisho (1995). "Biallelic inactivation of the APC gene in hepatoblastoma." Cancer Res **55**(21): 5007-5011.
- Kuroyanagi, H. (2009). "Fox-1 family of RNA-binding proteins." Cell Mol Life Sci **66**(24): 3895-3907.
- Labrador, M. and V. G. Corces (2003). "Extensive exon reshuffling over evolutionary time coupled to trans-splicing in Drosophila." Genome Res **13**(10): 2220-2228.
- Lam, B. J. and K. J. Hertel (2002). "A general role for splicing enhancers in exon definition." RNA **8**(10): 1233-1241.
- Lareau, L. F., R. E. Green, R. S. Bhatnagar and S. E. Brenner (2004). "The evolving roles of alternative splicing." Curr Opin Struct Biol **14**(3): 273-282.
- Lee, H. L. and J. P. Dougherty (2012). "Pharmaceutical therapies to recode nonsense mutations in inherited diseases." Pharmacol Ther **136**(2): 227-266.
- Letunic, I., R. R. Copley and P. Bork (2002). "Common exon duplication in animals and its role in alternative splicing." Hum Mol Genet **11**(13): 1561-1567.
- Lev-Maor, G., O. Ram, E. Kim, N. Sela, A. Goren, E. Y. Levanon and G. Ast (2008). "Intronic Alus influence alternative splicing." PLoS Genet **4**(9): e1000204.
- Levesque, S., C. Morin, S. P. Guay, J. Villeneuve, P. Marquis, W. Y. Yik, S. Jiralerspong, L. Bouchard, S. Steinberg, J. G. Hacia, K. Dewar and N. E. Braverman (2012). "A founder mutation in the PEX6 gene is responsible for increased incidence of Zellweger syndrome in a French Canadian population." BMC Med Genet **13**: 72.
- Levitus, M., Q. Waisfisz, B. C. Godthelp, Y. de Vries, S. Hussain, W. W. Wiegant, E. Elghalbzouri-Maghrani, J. Steltenpool, M. A. Rooimans, G. Pals, F. Arwert, C. G. Mathew, M. Z. Zdzienicka, K. Hiom, J. P. De Winter and H. Joenje (2005). "The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J." Nat Genet **37**(9): 934-935.
- Levy-Lahad, E. (2010). "Fanconi anemia and breast cancer susceptibility meet again." Nat Genet **42**(5): 368-369.
- Lewis, J. D., E. Izaurralde, A. Jarmolowski, C. McGuigan and I. W. Mattaj (1996). "A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site." Genes Dev **10**(13): 1683-1698.

- Linde, L., S. Boelz, G. Neu-Yilik, A. E. Kulozik and B. Kerem (2007). "The efficiency of nonsense-mediated mRNA decay is an inherent character and varies among different cells." *Eur J Hum Genet* **15**(11): 1156-1162.
- Lindor, N. M., L. Guidugli, X. Wang, M. P. Vallee, A. N. Monteiro, S. Tavtigian, D. E. Goldgar and F. J. Couch (2012). "A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS)." *Hum Mutat* **33**(1): 8-21.
- Liu, M. and A. Grigoriev (2004). "Protein domains correlate strongly with exons in multiple eukaryotic genomes--evidence of exon shuffling?" *Trends Genet* **20**(9): 399-403.
- Lopez-Bigas, N., B. Audit, C. Ouzounis, G. Parra and R. Guigo (2005). "Are splicing mutations the most frequent cause of hereditary disease?" *FEBS Lett* **579**(9): 1900-1903.
- Lu, Z. X., P. Jiang and Y. Xing (2012). "Genetic variation of pre-mRNA alternative splicing in human populations." *Wiley Interdiscip Rev RNA* **3**(4): 581-592.
- Luco, R. F., Q. Pan, K. Tominaga, B. J. Blencowe, O. M. Pereira-Smith and T. Misteli (2010). "Regulation of alternative splicing by histone modifications." *Science* **327**(5968): 996-1000.
- Lux, M. P., P. A. Fasching and M. W. Beckmann (2006). "Hereditary breast and ovarian cancer: review and future perspectives." *J Mol Med (Berl)* **84**(1): 16-28.
- Mann, C. J., K. Honeyman, A. J. Cheng, T. Ly, F. Lloyd, S. Fletcher, J. E. Morgan, T. A. Partridge and S. D. Wilton (2001). "Antisense-induced exon skipping and synthesis of dystrophin in the mdx mouse." *Proc Natl Acad Sci U S A* **98**(1): 42-47.
- Martinez-Contreras, R., J. F. Fiset, F. U. Nasim, R. Madden, M. Cordeau and B. Chabot (2006). "Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing." *PLoS Biol* **4**(2): e21.
- McCracken, S., N. Fong, E. Rosonina, K. Yankulov, G. Brothers, D. Siderovski, A. Hessel, S. Foster, S. Shuman and D. L. Bentley (1997). "5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II." *Genes Dev* **11**(24): 3306-3318.
- Meijers-Heijboer, H., A. van den Ouweland, J. Klijn, M. Wasielewski, A. de Snoo, R. Oldenburg, A. Hollestelle, M. Houben, E. Crepin, M. van Veghel-Plandsoen, F. Elstrodt, C. van Duijn, C. Bartels, C. Meijers, M. Schutte, L. McGuffog, D. Thompson, D. Easton, N. Sodha, S. Seal, R. Barfoot, J. Mangion, J. Chang-Claude, D. Eccles, R. Eeles, D. G. Evans, R. Houlston, V. Murday, S. Narod, T. Peretz, J. Peto, C. Phelan, H. X. Zhang, C. Szabo, P. Devilee, D. Goldgar, P. A. Futreal, K. L. Nathanson, B. Weber, N. Rahman, M. R. Stratton and C. H.-B. C. Consortium (2002). "Low-penetrance susceptibility to breast cancer due to CHEK2(\*)1100delC in noncarriers of BRCA1 or BRCA2 mutations." *Nat Genet* **31**(1): 55-59.
- Menendez, M., J. Castellsague, M. Mirete, E. Pros, L. Feliubadalo, A. Osorio, M. Calaf, E. Tornero, J. del Valle, J. Fernandez-Rodriguez, F. Quiles, M. Salinas, A. Velasco, A. Teule, J. Brunet, I. Blanco, G. Capella and C. Lazaro (2012). "Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes." *Breast Cancer Res Treat* **132**(3): 979-992.
- J. Chang-Claude, S. E. Bojesen, M. K. Bolla, Q. Wang, E. Dicks, A. Lee, C. Turnbull, N. Rahman, Breast, C. Ovarian Cancer Susceptibility, O. Fletcher, J. Peto, L. Gibson, I.

- Dos Santos Silva, H. Nevanlinna, T. A. Muranen, K. Aittomaki, C. Blomqvist, K. Czene, A. Irwanto, J. Liu, Q. Waisfisz, H. Meijers-Heijboer, M. Adank, B. Hereditary, N. Ovarian Cancer Research Group, R. B. van der Luijt, R. Hein, N. Dahmen, L. Beckman, A. Meindl, R. K. Schmutzler, B. Muller-Myhsok, P. Lichtner, J. L. Hopper, M. C. Southey, E. Makalic, D. F. Schmidt, A. G. Uitterlinden, A. Hofman, D. J. Hunter, S. J. Chanock, D. Vincent, F. Bacot, D. C. Tessier, S. Canisius, L. F. Wessels, C. A. Haiman, M. Shah, R. Luben, J. Brown, C. Luccarini, N. Schoof, K. Humphreys, J. Li, B. G. Nordestgaard, S. F. Nielsen, H. Flyger, F. J. Couch, X. Wang, C. Vachon, K. N. Stevens, D. Lambrechts, M. Moisse, R. Paridaens, M. R. Christiaens, A. Rudolph, S. Nickels, D. Flesch-Janys, N. Johnson, Z. Aitken, K. Aaltonen, T. Heikkinen, A. Broeks, L. J. Veer, C. E. van der Schoot, P. Guenel, T. Truong, P. Laurent-Puig, F. Menegaux, F. Marme, A. Schneeweiss, C. Sohn, B. Burwinkel, M. P. Zamora, J. I. Perez, G. Pita, M. R. Alonso, A. Cox, I. W. Brock, S. S. Cross, M. W. Reed, E. J. Sawyer, I. Tomlinson, M. J. Kerin, N. Miller, B. E. Henderson, F. Schumacher, L. Le Marchand, I. L. Andrulis, J. A. Knight, G. Glendon, A. M. Mulligan, I. kConFab, G. Australian Ovarian Cancer Study, A. Lindblom, S. Margolin, M. J. Hooning, A. Hollestelle, A. M. van den Ouweland, A. Jager, Q. M. Bui, J. Stone, G. S. Dite, C. Apicella, H. Tsimiklis, G. G. Giles, G. Severi, L. Baglietto, P. A. Fasching, L. Haeberle, A. B. Ekici, M. W. Beckmann, H. Brenner, H. Muller, V. Arndt, C. Stegmaier, A. Swerdlow, A. Ashworth, N. Orr, M. Jones, J. Figueroa, J. Lissowska, L. Brinton, M. S. Goldberg, F. Labreche, M. Dumont, R. Winqvist, K. Pylkas, A. Jukkola-Vuorinen, M. Grip, H. Brauch, U. Hamann, T. Bruning, G. Network, P. Radice, P. Peterlongo, S. Manoukian, B. Bonanni, P. Devilee, R. A. Tollenaar, C. Seynaeve, C. J. van Asperen, A. Jakubowska, J. Lubinski, K. Jaworska, K. Durda, A. Mannermaa, V. Kataja, V. M. Kosma, J. M. Hartikainen, N. V. Bogdanova, N. N. Antonenkova, T. Dork, V. N. Kristensen, H. Anton-Culver, S. Slager, A. E. Toland, S. Edge, F. Fostira, D. Kang, K. Y. Yoo, D. Y. Noh, K. Matsuo, H. Ito, H. Iwata, A. Sueta, A. H. Wu, C. C. Tseng, D. Van Den Berg, D. O. Stram, X. O. Shu, W. Lu, Y. T. Gao, H. Cai, S. H. Teo, C. H. Yip, S. Y. Phuah, B. K. Cornes, M. Hartman, H. Miao, W. Y. Lim, J. H. Sng, K. Muir, A. Lophatananon, S. Stewart-Brown, P. Siriwanarangsang, C. Y. Shen, C. N. Hsiung, P. E. Wu, S. L. Ding, S. Sangrajrang, V. Gaborieau, P. Brennan, J. McKay, W. J. Blot, L. B. Signorello, Q. Cai, W. Zheng, S. Deming-Halverson, M. Shrubsole, J. Long, J. Simard, M. Garcia-Closas, P. D. Pharoah, G. Chenevix-Trench, A. M. Dunning, J. Benitez and D. F. Easton (2013). "Large-scale genotyping identifies 41 new loci associated with breast cancer risk." *Nat Genet* **45**(4): 353-361, 361e351-352.
- Miki, Y., J. Swensen, D. Shattuck-Eidens, P. A. Futreal, K. Harshman, S. Tavtigian, Q. Liu, C. Cochran, L. M. Bennett, W. Ding and et al. (1994). "A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1." *Science* **266**(5182): 66-71.
  - Millevoi, S., S. Bernat, D. Telly, F. Fouque, L. Gladiéff, G. Favre, S. Vagner and C. Toulas (2010). "The c.5242C>A BRCA1 missense variant induces exon skipping by increasing splicing repressors binding." *Breast Cancer Res Treat* **120**(2): 391-399.
  - Modrek, B. and C. Lee (2002). "A genomic view of alternative splicing." *Nat Genet* **30**(1): 13-19.
  - Moller-Krull, M., A. Zemmann, C. Roos, J. Brosius and J. Schmitz (2008). "Beyond DNA: RNA editing and steps toward Alu exonization in primates." *J Mol Biol* **382**(3): 601-609.

- Moore, M. J. and P. A. Sharp (1993). "Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing." *Nature* **365**(6444): 364-368.
- Mordes, D., X. Luo, A. Kar, D. Kuo, L. Xu, K. Fushimi, G. Yu, P. Sternberg, Jr. and J. Y. Wu (2006). "Pre-mRNA splicing and retinitis pigmentosa." *Mol Vis* **12**: 1259-1271.
- Nekrutenko, A. and W. H. Li (2001). "Transposable elements are found in a large number of human protein-coding genes." *Trends Genet* **17**(11): 619-621.
- Nilsen, T. W. and B. R. Graveley (2010). "Expansion of the eukaryotic proteome by alternative splicing." *Nature* **463**(7280): 457-463.
- Nowak, D. G., J. Woolard, E. M. Amin, O. Konopatskaya, M. A. Saleem, A. J. Churchill, M. R. Lodomery, S. J. Harper and D. O. Bates (2008). "Expression of pro- and anti-angiogenic isoforms of VEGF is differentially regulated by splicing and growth factors." *J Cell Sci* **121**(Pt 20): 3487-3495.
- Oliver, A. W., S. Swift, C. J. Lord, A. Ashworth and L. H. Pearl (2009). "Structural basis for recruitment of BRCA2 by PALB2." *EMBO Rep* **10**(9): 990-996.
- Osheim, Y. N., O. L. Miller, Jr. and A. L. Beyer (1985). "RNP particles at splice junction sequences on Drosophila chorion transcripts." *Cell* **43**(1): 143-151.
- Osher, D. J., K. De Leeneer, G. Michils, N. Hamel, E. Tomiak, B. Poppe, K. Leunen, E. Legius, A. Shuen, E. Smith, J. Arseneau, P. Tonin, G. Matthijs, K. Claes, M. D. Tischkowitz and W. D. Foulkes (2012). "Mutation analysis of RAD51D in non-BRCA1/2 ovarian and breast cancer families." *Br J Cancer* **106**(8): 1460-1463.
- Osorio, F. G., C. L. Navarro, J. Cadinanos, I. C. Lopez-Mejia, P. M. Quiros, C. Bartoli, J. Rivera, J. Tazi, G. Guzman, I. Varela, D. Depetris, F. de Carlos, J. Cobo, V. Andres, A. De Sandre-Giovannoli, J. M. Freije, N. Levy and C. Lopez-Otin (2011). "Splicing-directed therapy in a new mouse model of human accelerated aging." *Sci Transl Med* **3**(106): 106ra107.
- Padgett, R. A. (2012). "New connections between splicing and human disease." *Trends Genet* **28**(4): 147-154.
- Pajares, M. J., T. Ezponda, R. Catena, A. Calvo, R. Pio and L. M. Montuenga (2007). "Alternative splicing: an emerging topic in molecular and clinical oncology." *Lancet Oncol* **8**(4): 349-357.
- Pan, Q., O. Shai, L. J. Lee, B. J. Frey and B. J. Blencowe (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." *Nat Genet* **40**(12): 1413-1415.
- Pandya-Jones, A. (2011). "Pre-mRNA splicing during transcription in the mammalian system." *Wiley Interdiscip Rev RNA* **2**(5): 700-717.
- Park, D. J., F. Lesueur, T. Nguyen-Dumont, M. Pertesi, F. Odefrey, F. Hammet, S. L. Neuhausen, E. M. John, I. L. Andrulis, M. B. Terry, M. Daly, S. Buys, F. Le Calvez-Kelm, A. Lonie, B. J. Pope, H. Tsimiklis, C. Voegelé, F. M. Hilbers, N. Hoogerbrugge, A. Barroso, A. Osorio, R. Breast Cancer Family, C. Kathleen Cunningham Foundation Consortium for Research into Familial Breast, G. G. Giles, P. Devilee, J. Benitez, J. L. Hopper, S. V. Tavtigian, D. E. Goldgar and M. C. Southey (2012). "Rare mutations in XRCC2 increase the risk of breast cancer." *Am J Hum Genet* **90**(4): 734-739.

- Paronetto, M. P., T. Achsel, A. Massiello, C. E. Chalfant and C. Sette (2007). "The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x." J Cell Biol **176**(7): 929-939.
- Pedersen, J. S., G. Bejerano, A. Siepel, K. Rosenbloom, K. Lindblad-Toh, E. S. Lander, J. Kent, W. Miller and D. Haussler (2006). "Identification and classification of conserved RNA secondary structures in the human genome." PLoS Comput Biol **2**(4): e33.
- Perez-Valle, J. and J. Vilardell (2012). "Intronic features that determine the selection of the 3' splice site." Wiley Interdiscip Rev RNA **3**(5): 707-717.
- Pettigrew, C., N. Wayte, P. K. Lovelock, S. V. Tavtigian, G. Chenevix-Trench, A. B. Spurdle and M. A. Brown (2005). "Evolutionary conservation analysis increases the colocalization of predicted exonic splicing enhancers in the BRCA1 gene with missense sequence changes and in-frame deletions, but not polymorphisms." Breast Cancer Res **7**(6): R929-939.
- Plon, S. E., D. M. Eccles, D. Easton, W. D. Foulkes, M. Genuardi, M. S. Greenblatt, F. B. Hogervorst, N. Hoogerbrugge, A. B. Spurdle, S. V. Tavtigian and I. U. G. V. W. Group (2008). "Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results." Hum Mutat **29**(11): 1282-1291.
- Ponder, B. A., A. Antoniou, A. Dunning, D. F. Easton and P. D. Pharoah (2005). "Polygenic inherited predisposition to breast cancer." Cold Spring Harb Symp Quant Biol **70**: 35-41.
- Provencio Pulla, M. and F. Valcarcel Sanchez (2005). "[Locally advanced non-small cell lung cancer]." Rev Clin Esp **205**(8): 392-394.
- Quesada, V., L. Conde, N. Villamor, G. R. Ordonez, P. Jares, L. Bassaganyas, A. J. Ramsay, S. Bea, M. Pinyol, A. Martinez-Trillos, M. Lopez-Guerra, D. Colomer, A. Navarro, T. Baumann, M. Aymerich, M. Rozman, J. Delgado, E. Gine, J. M. Hernandez, M. Gonzalez-Diaz, D. A. Puente, G. Velasco, J. M. Freije, J. M. Tubio, R. Royo, J. L. Gelpi, M. Orozco, D. G. Pisano, J. Zamora, M. Vazquez, A. Valencia, H. Himmelbauer, M. Bayes, S. Heath, M. Gut, I. Gut, X. Estivill, A. Lopez-Guillermo, X. S. Puente, E. Campo and C. Lopez-Otin (2012). "Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia." Nat Genet **44**(1): 47-52.
- Raghavachari, N., J. Barb, Y. Yang, P. Liu, K. Woodhouse, D. Levy, C. J. O'Donnell, P. J. Munson and G. J. Kato (2012). "A systematic comparison and evaluation of high density exon arrays and RNA-seq technology used to unravel the peripheral blood transcriptome of sickle cell disease." BMC Med Genomics **5**: 28.
- Ramani, A. K., J. A. Calarco, Q. Pan, S. Mavandadi, Y. Wang, A. C. Nelson, L. J. Lee, Q. Morris, B. J. Blencowe, M. Zhen and A. G. Fraser (2011). "Genome-wide analysis of alternative splicing in *Caenorhabditis elegans*." Genome Res **21**(2): 342-348.
- Reed, R. and T. Maniatis (1986). "A role for exon sequences and splice-site proximity in splice-site selection." Cell **46**(5): 681-690.
- Reese, M. G., F. H. Eeckman, D. Kulp and D. Haussler (1997). "Improved splice site detection in Genie." J Comput Biol **4**(3): 311-323.

- Reid, S., D. Schindler, H. Hanenberg, K. Barker, S. Hanks, R. Kalb, K. Neveling, P. Kelly, S. Seal, M. Freund, M. Wurm, S. D. Batish, F. P. Lach, S. Yetgin, H. Neitzel, H. Ariffin, M. Tischkowitz, C. G. Mathew, A. D. Auerbach and N. Rahman (2007). "Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer." *Nat Genet* **39**(2): 162-164.
- Renwick, A., D. Thompson, S. Seal, P. Kelly, T. Chagtai, M. Ahmed, B. North, H. Jayatilake, R. Barfoot, K. Spanova, L. McGuffog, D. G. Evans, D. Eccles, C. Breast Cancer Susceptibility, D. F. Easton, M. R. Stratton and N. Rahman (2006). "ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles." *Nat Genet* **38**(8): 873-875.
- Robberson, B. L., G. J. Cote and S. M. Berget (1990). "Exon definition may facilitate splice site selection in RNAs with multiple exons." *Mol Cell Biol* **10**(1): 84-94.
- Rodriguez, A. O., M. Llacuachaqui, G. G. Pardo, R. Royer, G. Larson, J. N. Weitzel and S. A. Narod (2012). "BRCA1 and BRCA2 mutations among ovarian cancer patients from Colombia." *Gynecol Oncol* **124**(2): 236-243.
- Rosa-Rosa, J. M., G. Pita, M. Urioste, G. Llorca, J. Brunet, C. Lazaro, I. Blanco, T. Ramon y Cajal, O. Diez, M. de la Hoya, T. Caldes, M. I. Tejada, A. Gonzalez-Neira and J. Benitez (2009). "Genome-wide linkage scan reveals three putative breast-cancer-susceptibility loci." *Am J Hum Genet* **84**(2): 115-122.
- Roscigno, R. F. and M. A. Garcia-Blanco (1995). "SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome." *RNA* **1**(7): 692-706.
- Roy, R., J. Chun and S. N. Powell (2012). "BRCA1 and BRCA2: different roles in a common pathway of genome protection." *Nat Rev Cancer* **12**(1): 68-78.
- Roy, S. W. and W. Gilbert (2006). "The evolution of spliceosomal introns: patterns, puzzles and progress." *Nat Rev Genet* **7**(3): 211-221.
- Ruark, E., K. Snape, P. Humburg, C. Loveday, I. Bajrami, R. Brough, D. N. Rodrigues, A. Renwick, S. Seal, E. Ramsay, V. Duarte Sdel, M. A. Rivas, M. Warren-Perry, A. Zachariou, A. Champion-Flora, S. Hanks, A. Murray, N. A. Pour, J. Douglas, L. Gregory, A. Rimmer, N. M. Walker, T. P. Yang, J. W. Adlard, J. Barwell, J. Berg, A. F. Brady, C. Brewer, G. Brice, C. Chapman, J. Cook, R. Davidson, A. Donaldson, F. Douglas, D. Eccles, D. G. Evans, L. Greenhalgh, A. Henderson, L. Izatt, A. Kumar, F. Laloo, Z. Miedzybrodzka, P. J. Morrison, J. Paterson, M. Porteous, M. T. Rogers, S. Shanley, L. Walker, M. Gore, R. Houlston, M. A. Brown, M. J. Caufield, P. Deloukas, M. I. McCarthy, J. A. Todd, Breast, C. Ovarian Cancer Susceptibility, C. Wellcome Trust Case Control, C. Turnbull, J. S. Reis-Filho, A. Ashworth, A. C. Antoniou, C. J. Lord, P. Donnelly and N. Rahman (2013). "Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer." *Nature* **493**(7432): 406-410.
- Sakharkar, M. K., B. S. Perumal, K. R. Sakharkar and P. Kanguene (2005). "An analysis on gene architecture in human and mouse genomes." *In Silico Biol* **5**(4): 347-365.
- Sanz, D. J., A. Acedo, M. Infante, M. Duran, L. Perez-Cabornero, E. Esteban-Cardenosa, E. Lastra, F. Pagani, C. Miner and E. A. Velasco (2010). "A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients." *Clin Cancer Res* **16**(6): 1957-1967.
- Sauna, Z. E. and C. Kimchi-Sarfaty (2011). "Understanding the contribution of synonymous mutations to human disease." *Nat Rev Genet* **12**(10): 683-691.

- Schmucker, D., J. C. Clemens, H. Shu, C. A. Worby, J. Xiao, M. Muda, J. E. Dixon and S. L. Zipursky (2000). "Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity." Cell **101**(6): 671-684.
- Schor, I. E., N. Rascovan, F. Pelisch, M. Allo and A. R. Kornblihtt (2009). "Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing." Proc Natl Acad Sci U S A **106**(11): 4325-4330.
- Schuster, B., K. Knies, C. Stoepker, E. Velleuer, R. Friedl, B. Gottwald-Muhlhauser, J. P. de Winter and D. Schindler (2013). "Whole exome sequencing reveals uncommon mutations in the recently identified Fanconi anemia gene SLX4/FANCP." Hum Mutat **34**(1): 93-96.
- Seal, S., D. Thompson, A. Renwick, A. Elliott, P. Kelly, R. Barfoot, T. Chagtai, H. Jayatilake, M. Ahmed, K. Spanova, B. North, L. McGuffog, D. G. Evans, D. Eccles, C. Breast Cancer Susceptibility, D. F. Easton, M. R. Stratton and N. Rahman (2006). "Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles." Nat Genet **38**(11): 1239-1241.
- Shepard, P. J. and K. J. Hertel (2008). "Conserved RNA secondary structures promote alternative splicing." RNA **14**(8): 1463-1469.
- Shephard, N. D., R. Abo, S. H. Rigas, B. Frank, W. Y. Lin, I. W. Brock, A. Shippen, S. P. Balasubramanian, M. W. Reed, C. R. Bartram, A. Meindl, R. K. Schmutzler, C. Engel, B. Burwinkel, L. A. Cannon-Albright, K. Allen-Brady, N. J. Camp and A. Cox (2009). "A breast cancer risk haplotype in the caspase-8 gene." Cancer Res **69**(7): 2724-2728.
- Skotheim, R. I. and M. Nees (2007). "Alternative splicing in cancer: noise, functional, or systematic?" Int J Biochem Cell Biol **39**(7-8): 1432-1449.
- Smith, C. W. and J. Valcarcel (2000). "Alternative pre-mRNA splicing: the logic of combinatorial control." Trends Biochem Sci **25**(8): 381-388.
- Solyom, S., B. Aressy, K. Pylkas, J. Patterson-Fortin, J. M. Hartikainen, A. Kallioniemi, S. Kauppila, J. Nikkila, V. M. Kosma, A. Mannermaa, R. A. Greenberg and R. Winqvist (2012). "Breast cancer-associated Abraxas mutation disrupts nuclear localization and DNA damage response functions." Sci Transl Med **4**(122): 122ra123.
- Somyajit, K., S. Subramanya and G. Nagaraju (2010). "RAD51C: a novel cancer susceptibility gene is linked to Fanconi anemia and breast cancer." Carcinogenesis **31**(12): 2031-2038.
- Sorek, R., G. Ast and D. Graur (2002). "Alu-containing exons are alternatively spliced." Genome Res **12**(7): 1060-1067.
- Speevak, M. D., S. S. Young, H. Feilotter and P. Ainsworth (2003). "Alternatively spliced, truncated human BRCA2 isoforms contain a novel coding exon." Eur J Hum Genet **11**(12): 951-954.
- Steitz, J. A., G. Dreyfuss, A. R. Krainer, A. I. Lamond, A. G. Matera and R. A. Padgett (2008). "Where in the cell is the minor spliceosome?" Proc Natl Acad Sci U S A **105**(25): 8485-8486.
- Sterne-Weiler, T., J. Howard, M. Mort, D. N. Cooper and J. R. Sanford (2011). "Loss of exon identity is a common mechanism of human inherited disease." Genome Res **21**(10): 1563-1571.
- Sterner, D. A., T. Carlo and S. M. Berget (1996). "Architectural limits on split genes." Proc Natl Acad Sci U S A **93**(26): 15081-15085.

- Stoltzfus, C. M. and J. M. Madsen (2006). "Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing." Curr HIV Res **4**(1): 43-55.
- Stratton, M. R. and N. Rahman (2008). "The emerging landscape of breast cancer susceptibility." Nat Genet **40**(1): 17-22.
- Szabo, C., A. Masiello, J. F. Ryan and L. C. Brody (2000). "The breast cancer information core: database design, structure, and scope." Hum Mutat **16**(2): 123-131.
- Tardiff, D. F. and M. Rosbash (2006). "Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly." RNA **12**(6): 968-979.
- Taylor, M. D., N. Gokgoz, I. L. Andrulis, T. G. Mainprize, J. M. Drake and J. T. Rutka (2000). "Familial posterior fossa brain tumors of infancy secondary to germline mutation of the hSNF5 gene." Am J Hum Genet **66**(4): 1403-1406.
- Thanaraj, T. A., F. Clark and J. Muilu (2003). "Conservation of human alternative splice events in mouse." Nucleic Acids Res **31**(10): 2544-2552.
- Thery, J. C., S. Krieger, P. Gaildrat, F. Revillion, M. P. Buisine, A. Killian, C. Duponchel, A. Rousselin, D. Vaur, J. P. Peyrat, P. Berthet, T. Frebourg, A. Martins, A. Hardouin and M. Tosi (2011). "Contribution of bioinformatics predictions and functional splicing assays to the interpretation of unclassified variants of the BRCA genes." Eur J Hum Genet **19**(10): 1052-1058.
- Thomassen, M., A. Blanco, M. Montagna, T. V. Hansen, I. S. Pedersen, S. Gutierrez-Enriquez, M. Menendez, L. Fachal, M. Santamarina, A. Y. Steffensen, L. Jonson, S. Agata, P. Whiley, S. Tognazzo, E. Tornero, U. B. Jensen, J. Balmana, T. A. Kruse, D. E. Goldgar, C. Lazaro, O. Diez, A. B. Spurdle and A. Vega (2012). "Characterization of BRCA1 and BRCA2 splicing variants: a collaborative report by ENIGMA consortium members." Breast Cancer Res Treat **132**(3): 1009-1023.
- Thompson, D. and D. Easton (2004). "The genetic epidemiology of breast cancer genes." J Mammary Gland Biol Neoplasia **9**(3): 221-236.
- Thompson, J. D., T. J. Gibson and D. G. Higgins (2002). "Multiple sequence alignment using ClustalW and ClustalX." Curr Protoc Bioinformatics **Chapter 2**: Unit 2 3.
- Thorslund, T., M. J. McIlwraith, S. A. Compton, S. Lekomtsev, M. Petronczki, J. D. Griffith and S. C. West (2010). "The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA." Nat Struct Mol Biol **17**(10): 1263-1265.
- Tuerk, C. and L. Gold (1990). "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." Science **249**(4968): 505-510.
- Turunen, J. J., E. H. Niemela, B. Verma and M. J. Frilander (2012). "The significant other: splicing by the minor spliceosome." Wiley Interdiscip Rev RNA.
- Ule, J., A. Ule, J. Spencer, A. Williams, J. S. Hu, M. Cline, H. Wang, T. Clark, C. Fraser, M. Ruggiu, B. R. Zeeberg, D. Kane, J. N. Weinstein, J. Blume and R. B. Darnell (2005). "Nova regulates brain-specific splicing to shape the synapse." Nat Genet **37**(8): 844-852.

- Vallon-Christersson, J., C. Cayan, K. Haraldsson, N. Loman, J. T. Bergthorsson, K. Brondum-Nielsen, A. M. Gerdes, P. Moller, U. Kristoffersson, H. Olsson, A. Borg and A. N. Monteiro (2001). "Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families." *Hum Mol Genet* **10**(4): 353-360.
- Velasco, E., M. Infante, M. Duran, L. Perez-Cabornero, D. J. Sanz, E. Esteban-Cardenosa and C. Miner (2007). "Heteroduplex analysis by capillary array electrophoresis for rapid mutation detection in large multiexon genes." *Nat Protoc* **2**(1): 237-246.
- Vidal, C., A. Cachia and A. Xuereb-Anastasi (2009). "Effects of a synonymous variant in exon 9 of the CD44 gene on pre-mRNA splicing in a family with osteoporosis." *Bone* **45**(4): 736-742.
- Wachter, A. (2010). "Riboswitch-mediated control of gene expression in eukaryotes." *RNA Biol* **7**(1): 67-76.
- Walker, L. C., P. J. Whiley, F. J. Couch, D. J. Farrugia, S. Healey, D. M. Eccles, F. Lin, S. A. Butler, S. A. Goff, B. A. Thompson, S. R. Lakhani, L. M. Da Silva, I. kConFab, S. V. Tavtigian, D. E. Goldgar, M. A. Brown and A. B. Spurdle (2010). "Detection of splicing aberrations caused by BRCA1 and BRCA2 sequence variants encoding missense substitutions: implications for prediction of pathogenicity." *Hum Mutat* **31**(6): E1484-1505.
- Wang, B., S. Matsuoka, B. A. Ballif, D. Zhang, A. Smogorzewska, S. P. Gygi and S. J. Elledge (2007). "Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response." *Science* **316**(5828): 1194-1198.
- Wang, J., J. Zhang, K. Li, W. Zhao and Q. Cui (2012). "SpliceDisease database: linking RNA splicing and disease." *Nucleic Acids Res* **40**(Database issue): D1055-1059.
- Wang, M., H. Dotzlaw, S. A. Fuqua and L. C. Murphy (1997). "A point mutation in the human estrogen receptor gene is associated with the expression of an abnormal estrogen receptor mRNA containing a 69 novel nucleotide insertion." *Breast Cancer Res Treat* **44**(2): 145-151.
- Wang, Z. and C. B. Burge (2008). "Splicing regulation: from a parts list of regulatory elements to an integrated splicing code." *RNA* **14**(5): 802-813.
- Wang, Z., M. E. Rolish, G. Yeo, V. Tung, M. Mawson and C. B. Burge (2004). "Systematic identification and analysis of exonic splicing silencers." *Cell* **119**(6): 831-845.
- Wappenschmidt, B., A. A. Becker, J. Hauke, U. Weber, S. Engert, J. Kohler, K. Kast, N. Arnold, K. Rhiem, E. Hahnen, A. Meindl and R. K. Schmutzler (2012). "Analysis of 30 putative BRCA1 splicing mutations in hereditary breast and ovarian cancer families identifies exonic splice site mutations that escape in silico prediction." *PLoS One* **7**(12): e50800.
- Warzecha, C. C., T. K. Sato, B. Nabet, J. B. Hogenesch and R. P. Carstens (2009). "ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing." *Mol Cell* **33**(5): 591-601.
- Whiley, P. J., L. Guidugli, L. C. Walker, S. Healey, B. A. Thompson, S. R. Lakhani, L. M. Da Silva, I. kConFab, S. V. Tavtigian, D. E. Goldgar, M. A. Brown, F. J. Couch and A. B. Spurdle (2011). "Splicing and multifactorial analysis of intronic BRCA1 and BRCA2

- sequence variants identifies clinically significant splicing aberrations up to 12 nucleotides from the intron/exon boundary." *Hum Mutat* **32**(6): 678-687.
- Whiley, P. J., C. A. Pettigrew, B. L. Brewster, L. C. Walker, I. kConFab, A. B. Spurdle and M. A. Brown (2010). "Effect of BRCA2 sequence variants predicted to disrupt exonic splice enhancers on BRCA2 transcripts." *BMC Med Genet* **11**: 80.
  - Will, C. L. and R. Luhrmann (2011). "Spliceosome structure and function." *Cold Spring Harb Perspect Biol* **3**(7).
  - Winkler, W., A. Nahvi and R. R. Breaker (2002). "Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression." *Nature* **419**(6910): 952-956.
  - Wirth, B., L. Brichta and E. Hahnen (2006). "Spinal muscular atrophy: from gene to therapy." *Semin Pediatr Neurol* **13**(2): 121-131.
  - Wooster, R., G. Bignell, J. Lancaster, S. Swift, S. Seal, J. Mangion, N. Collins, S. Gregory, C. Gumbs and G. Micklem (1995). "Identification of the breast cancer susceptibility gene BRCA2." *Nature* **378**(6559): 789-792.
  - Wu, Q. and A. R. Krainer (1997). "Splicing of a divergent subclass of AT-AC introns requires the major spliceosomal snRNAs." *RNA* **3**(6): 586-601.
  
  - Xu, W., J. Seok, M. N. Mindrinos, A. C. Schweitzer, H. Jiang, J. Wilhelmy, T. A. Clark, K. Kapur, Y. Xing, M. Faham, J. D. Storey, L. L. Moldawer, R. V. Maier, R. G. Tompkins, W. H. Wong, R. W. Davis, W. Xiao, Inflammation and P. Host Response to Injury Large-Scale Collaborative Research (2011). "Human transcriptome array for high-throughput clinical studies." *Proc Natl Acad Sci U S A* **108**(9): 3707-3712.
  - Yu, J., Z. Yang, M. Kibukawa, M. Paddock, D. A. Passey and G. K. Wong (2002). "Minimal introns are not "junk"." *Genome Res* **12**(8): 1185-1189.
  - Yun, M. H. and K. Hiom (2009). "CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle." *Nature* **459**(7245): 460-463.
  
  - Zhang, J., C. C. Kuo and L. Chen (2011). "GC content around splice sites affects splicing through pre-mRNA secondary structures." *BMC Genomics* **12**: 90.
  - Zhang, L., R. Bacares, S. Boyar, C. Hudis, K. Nafa and K. Offit (2009). "cDNA analysis demonstrates that the BRCA2 intronic variant IVS4-12del5 is a deleterious mutation." *Mutat Res* **663**(1-2): 84-89.
  - Zhang, L., G. M. Vincent, M. Baralle, F. E. Baralle, B. D. Anson, D. W. Benson, B. Whiting, K. W. Timothy, J. Carlquist, C. T. January, M. T. Keating and I. Splawski (2004). "An intronic mutation causes long QT syndrome." *J Am Coll Cardiol* **44**(6): 1283-1291.
  - Zhang, X. H., K. A. Heller, I. Hefter, C. S. Leslie and L. A. Chasin (2003). "Sequence information for the splicing of human pre-mRNA identified by support vector machine classification." *Genome Res* **13**(12): 2637-2650.
  - Zuker, M. (1989). "Computer prediction of RNA structure." *Methods Enzymol* **180**: 262-288.

*Anexo I*

*Producción Científica*



**Patente:**

**Título:** P201231427 - Plásmido pSAD para ensayos funcionales de *splicing*.

**Autores:** Eladio A. Velasco, Alberto Acedo, Beatriz Díez.

**Artículos Publicados:**

1. **Título:** *Comprehensive splicing functional analysis of DNA variants of the BRCA2 gene by hybrid minigenes.*

**Autores:** **Alberto Acedo**, David J. Sanz, Mercedes Durán, Mar Infante, Lucía Pérez-Cabornero, Enrique Lastra, Cristina Miner and Eladio A. Velasco.

**Referencia:** Breast Cancer Res **14**(3): R87.

**Abstract:**

**INTRODUCTION:** The underlying pathogenic mechanism of a large fraction of DNA variants of disease-causing genes is the disruption of the splicing process. We aimed to investigate the effect on splicing of the BRCA2 variants c.8488-1G > A (exon 20) and c.9026\_9030del (exon 23), as well as 41 BRCA2 variants reported in the Breast Cancer Information Core (BIC) mutation database.

**METHODS:** DNA variants were analyzed with the splicing prediction programs NNSPLICE and Human Splicing Finder. Functional analyses of candidate variants were performed by lymphocyte RT-PCR and/or hybrid minigene assays. Forty-one BIC variants of exons 19, 20, 23 and 24 were bioinformatically selected and generated by PCR-mutagenesis of the wild type minigenes.

**RESULTS:** Lymphocyte RT-PCR of c.8488-1G>A showed intron 19 retention and a 12-nucleotide deletion in exon 20, whereas c.9026\_9030del did not show any splicing anomaly. Minigene analysis of c.8488-1G>A displayed the aforementioned aberrant isoforms but also exon 20 skipping. We further evaluated the splicing outcomes of 41 variants of four BRCA2 exons by minigene analysis. Eighteen variants presented splicing aberrations. Most variants (78.9%) disrupted the natural splice sites, whereas four altered putative enhancers/silencers and had a weak effect. Fluorescent RT-PCR of minigenes accurately detected 14 RNA isoforms generated by cryptic site usage, exon skipping and intron retention events. Fourteen variants showed total splicing disruptions and were predicted to truncate or eliminate essential domains of BRCA2.

**CONCLUSIONS:** A relevant proportion of BRCA2 variants are correlated with splicing disruptions, indicating that RNA analysis is a valuable tool to assess the pathogenicity of a particular DNA change. The minigene system is a straightforward and robust approach to detect variants with an impact on splicing and contributes to a better knowledge of this gene expression step.

2. **Título:** *A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patient.*

**Autores:** Sanz DJ, **Acedo A**, Infante M, Durán M Infante M, Pérez-Cabornero L, Esteban-Cardenosa E, Lastra E, Pagani F, Miner C,& Velasco EA.

**Referencia:** Clin Cancer Res **16**(6): 1957-1967.

**Abstract:**

**PURPOSE:** Most BRCA1/2 mutations are of unknown clinical relevance. An increasing amount of evidence indicates that there can be deleterious effects

through the disruption of the splicing process. We have investigated the effect of aberrant splicing of BRCA1/2 on hereditary breast/ovarian cancer (HBOC).

**EXPERIMENTAL DESIGN:** DNA variants were analyzed with splicing prediction programs to select putative splicing mutations. Splicing assays of 57 genetic variants were done by lymphocyte reverse transcription-PCR and/or hybrid minigenes in HeLa and nontumor breast epithelial cells.

**RESULTS:** Twenty-four BRCA1/2 variants of Spanish HBOC patients were bioinformatically preselected. Functional assays showed that 12 variants induced anomalous splicing patterns, 6 of which accounted for 58.5% of BRCA1 families. To further evaluate the defective splicing of BRCA1/2, we analyzed 31 Breast Cancer Information Core Database (BIC) and two artificial variants that were generated by mutagenesis. Sixteen variants induced different degrees of aberrant splicing. Altogether, anomalous splicing was caused by 28 BRCA1/2 variants of all types, indicating that any DNA change can disrupt pre-mRNA processing. We show that a wide range of regulatory elements can be involved, including the canonical and cryptic splice sites, the polypyrimidine tract, and splicing enhancers/silencers. Twenty mutations were predicted to truncate the BRCA proteins and/or to delete essential domains, thus supporting a role in HBOC.

**CONCLUSIONS:** An important fraction of DNA variants of BRCA1/2 presents splicing aberrations that may represent a relevant disease-causing mechanism in HBOC. The identification of splicing disruptions by functional assays is a valuable tool to discriminate between benign polymorphisms and pathogenic mutations.

**3. Título:** Capillary Electrophoresis analysis of Conventional Splicing Assays: IARC Analytical and Clinical Classification of 31 BRCA2 genetic variants.

**Autores:** Gorka Ruiz de Garibay, **Alberto Acedo**, Zaida García-Casado, Alicia Tosar, Atocha Romero, Pilar Garre, Gemma Llorca, Mads Thomassen, Orland Díez, Pedro Pérez-Segura, Eduardo Díaz-Rubio, Eladio A. Velasco, Trinidad Caldés, and Miguel de la Hoya.

**Referencia:** Human Mutation (Accepted).

**Abstract:** Rare sequence variants in “high-risk” disease genes often referred as Unclassified Variants (UVs) pose a serious challenge to genetic testing. However, UVs resulting in splicing alterations can be readily assessed by in vitro assays. Unfortunately, analytical and clinical interpretation of these assays is often challenging. Here we explore this issue by conducting splicing assays in 31 BRCA2 genetic variants. All variants were assessed by RT-PCR followed by capillary electrophoresis (EP) and direct sequencing. If assays did not produce clear-cut outputs (Class-2 or Class-5 according to analytical IARC guide-lines), we performed qPCR and/or Minigene assays. The latter were performed with a new splicing vector (pSAD) developed by authors of the present manuscript (Patent P201231427 CSIC). We have identified 3 clinically relevant Class-5 variants (c.682-2A>G, c.7617+1G>A, and c.8954-5A>G), and 27 analytical Class-2 variants (not inducing splicing alterations). In addition, we demonstrate that rs9534262 (c.7806-14T>C) is a BRCA2 splicing quantitative trait loci (sQTL).

**4. Título:** BRCA1 5272-1G>A and BRCA2 5374delTATG are founder mutations of high relevance for genetic counselling in breast/ovarian cancer families of Spanish origin.

**Autores:** Infante M, Durán M, **Acedo A**, Pérez-Cabornero L, Sanz DJ, García-González M, Beristain E, Esteban-Cardenosa E, Teulé A, Vega A, Tejada I, Lastra E, Miner C & Velasco EA.

**Referencia:** *Clin Genet* **77**(1): 60-69.

**Abstract:**

The distribution of BRCA1 and BRCA2 germ line mutations in breast/ovarian cancer families varies among different populations, which typically present a wide spectrum of unique mutations. Splicing mutation 5272-1G>A of BRCA1 and frameshift mutation 5374delTATG of BRCA2 are highly prevalent mutations in Castilla-Leon (Spain), accounting for 18.4% and 13.6% of BRCA1 and BRCA2 positive families, respectively. To test the presence of founder effects, 9 Spanish 5272-1G>A and 13 5374delTATG families were genotyped with polymorphic markers linked to BRCA1 or BRCA2. All the 5272-1G>A families shared a common haplotype in eight markers (1.1 Mb region) and the mutation age was estimated in 15 generations (approximately 380 years). A conserved haplotype associated to 5374delTATG was observed in four markers (0.82 Mb). The mutation occurred approximately 48 generations ago (approximately 1200 years). Each mutation likely arose from a common ancestor that could be traced to a small area of Castilla-Leon and expanded to other Spanish regions. They can have a significant impact on the clinical management of asymptomatic carriers as well as on the genetic screening strategy to be followed in populations with Spanish ancestries.

**5. Título:** *Two founder BRCA2 mutations predispose to breast cancer in young women.*

**Autores:** Infante M, Durán M, Lasa A, **Acedo A**, de la Hoya M, Esteban-Cardenosa E, Sanz DJ, Pérez-Cabornero L, Lastra E, Miner C & Velasco EA.

**Referencia:** *Breast Cancer Res Treat* **122**(2): 567-571.

**Abstract:**

The mutation spectrum of BRCA1 and BRCA2 presents a wide range of unique mutations in breast/ovarian cancer patients but recurrent mutations with founder effects have also been described. BRCA2 5344delAATA and 9538delAA are recurrent mutations in Castilla-Leon (Spain) representing 10.6% of BRCA2 positive families. By genotyping eleven chromosome 13 markers (4.3 Mb) we demonstrate that each mutation shows core haplotypes of 1.66 and 0.87 Mb, respectively, supporting a common ancestor in Castilla-Leon. Furthermore, both mutations are associated with earlier onset of breast cancer (5344delAATA: 37.4 years, P = 0.033; 9538delAA: 39.4 years, P = 0.008). The identification of founder effects improves the genetic screening strategy to be followed and facilitates the clinical management of asymptomatic carriers.

**6. Título:** *The highly prevalent BRCA2 mutation c.2808\_2811del (3036delACAA) is located in a mutational hotspot and has multiple origins.*

**Autores:** Mar Infante, Mercedes Durán, **Alberto Acedo**, Eva María Sánchez-Tapia, Beatriz Díez-Gómez, Alicia Barroso, María García-González, Lidia Feliubadaló, Adriana Lasa, Miguel de la Hoya, Eva Esteban-Cardenosa, Orland Díez, Cristina Martínez Bouzas, Javier Godino, Alexandre Teulé, Ana Osorio, Enrique Lastra, Rogelio González-Sarmiento, Cristina Miner and Eladio A. Velasco.

**Referencia:** *Carcinogenesis*. doi: 10.1093/carcin/bgt272

**Abstract:**

**BACKGROUND:** *BRCA2* 3036delACAA (c.2808\_2811del) is the most prevalent mutation in Spanish breast cancer patients and the second most reported in the mutation databases. We aimed to investigate a putative common genetic origin of this mutation in Spanish carrier families.

**METHODS:** Mutation carriers of 51 families were genotyped with seven microsatellite markers and four intragenic SNPs of *BRCA2*. Mutation age was calculated with the DMLE software. DNA secondary structure was predicted with the Quikfold program. Random mutagenesis experiments were performed over plasmid containing a 923-bp fragment of *BRCA2* surrounding the nucleotide 3036.

**RESULTS:** Three independent founder haplotypes associated with 3036delACAA were clearly distinguished that accounted for 23 families (West Castilla y León, WCL), 20 families (East Castilla y León, ECL) and 6 families (Mediterranean coast). However haplotypes could not be assigned in two families. Mutation age was estimated in a range of 45-68 generations for WCL families, and 45-71 generations for ECL families. Mutation 3036delACAA was located in a double-hairpin loop structure (*BRCA2* cDNA 3022-3053) that was proposed as mutational hotspots. To check this hypothesis, random mutagenesis was performed over these sequences. A total of 85 DNA variants were characterized, 53 of which (62.4%) were located in hairpin-loop structures. Interestingly, three mutations reported in the mutation databases (2908G>A, 3172delA and 3185insA) were replicated in the mutagenesis experiment. Moreover, five mutations were placed in the same hairpin-loop of 3036delACAA, and one of them (3036A>G) affected the same position.

**CONCLUSIONS:** Our results support that at least three different mutational events took place in the analyzed families to generate this mutation. DNA variants, including 3036delACAA, are preferentially concentrated in predicted hairpin-loops suggesting that these structures might represent mutational hotspots.

**7. Título:** *Characterization of new founder Alu-mediated rearrangements in MSH2 gene associated with a Lynch Syndrome phenotype.*

**Autores:** Lucía Pérez-Cabornero, Ester Borrás, Mar Infante, Eladio A. Velasco, **Alberto Acedo**, Enrique Lastra, Jorge Cuevas, Marta Pineda, Teresa Ramon and Cajal, Capellá Gabriel, Cristina Miner and Mercedes Durán Domínguez.

**Referencia:** *Cancer Prev Res (Phila)* **4**(10): 1546-1555.

**Abstract:**

It has been reported that large genomic deletions in the *MLH1* and *MSH2* genes are a frequent cause of Lynch syndrome in certain populations. Here, a cohort has been screened and two new founder rearrangements have been found in the *MSH2* gene. These mutations have been characterized by break point determination, haplotype analysis, and genotype-phenotype correlation. Mutations have been identified in the *MLH1*, *MSH2*, and *MSH6* genes in 303 subjects from 160 suspected Lynch syndrome unrelated families. All subjects were tested using heteroduplex analysis by capillary array electrophoresis. Multiplex ligation-dependent probe amplification was used to detect rearrangements in mutation-negative index patients and confirmed by reverse transcriptase PCR. The break point of the deletions was further characterized by the array comparative

genomic hybridization method. Immunohistochemical staining and microsatellite instability were studied in tumor samples. Hereditary nonpolyposis colorectal cancer-related phenotypes were evaluated. More than 16% (24 of 160) of the families had pathogenic mutations (8 MLH1, 15 MSH2, and 1 MSH6). Twelve of these families (50%) are carriers of a novel mutation. Seven of the 15 positive MSH2 families (47%) are carriers of a rearrangement. The exon 7 deletion and exon 4 to 8 deletion of MSH2 are new founder mutations. The segregation of a common haplotype, a similar phenotype, and anticipation effects were observed in these families. These findings will greatly simplify the diagnosis, counseling, and clinical care in suspected Lynch syndrome families and not just in specific geographic areas, so wide distribution may be explained by migration patterns.

**8. *Título:*** *Frequency of Rearrangements in Lynch Syndrome cases associated to MSH2. Characterization of a new deletion involving both EPCAM and the 5' part of MSH2.*

***Autores:*** Lucia Perez-Cabornero, Mar Infante, Eladio A. Velasco, Enrique Lastra, **Alberto Acedo**, Cristina Miner and Mercedes Durán Domínguez.

***Referencia:*** *Cancer Prev Res (Phila)* **4**(10): 1556-1562.

***Abstract:***

Lynch syndrome is caused by germline mutations in MSH2, MLH1, MSH6, and PMS2 mismatch repair genes and leads to a high risk of colorectal and endometrial cancer. It was recently shown that constitutional 3' end deletions of EPCAM could cause Lynch syndrome in tissues with MSH2 deficiency. We aim to establish the spectrum of mutations in MSH2-associated Lynch syndrome cases and their clinical implications. Probands from 159 families suspected of having Lynch syndrome were enrolled in the study. Immunohistochemistry and microsatellite instability (MSI) analyses were used on the probands of all families. Eighteen cases with MSH2 loss were identified: eight had point mutations in MSH2. In 10 Lynch syndrome families without MSH2 mutations, EPCAM-MSH2 genomic rearrangement screening was carried out with the use of multiplex ligation-dependent probe amplification and reverse transcriptase PCR. We report that large germline deletions, encompassing one or more exons of the MSH2 gene, cosegregate with the Lynch syndrome phenotype in 23% (8 of 35) of MSI families tested. A new combined deletion EPCAM-MSH2 was identified and characterized by break point analysis, encompassing from the 3' end region of EPCAM to the 5' initial sequences of the MSH2 (c.859-1860\_MSH2: 646-254del). EPCAM-MSH2 fusion transcript was isolated. The tumors of the carriers show high-level MSI and MSH2 protein loss. The clinical correlation provided evidence that the type of mutation and the extension of the deletions involving the MSH2 gene could have different implications in cancer predisposition. Thus, the identification of EPCAM-MSH2 rearrangements and their comprehensive characterization should be included in the routine mutation screening protocols for Lynch syndrome.

Premio mejor comunicación en el congreso de la Asociación Española de genética Humana, Madrid 12 de abril de 2013.

# New splicing vector pSAD: Splicing functional analysis of a hybrid "maxi-minigene" with exons 19 to 27 of BRCA2



Acedo A.<sup>1</sup>, B. Díez-Gómez<sup>1</sup>, A. Curiel<sup>1</sup>, C. Hernández-Moro<sup>1</sup>, M. Infante<sup>2</sup>, M. Durán<sup>2</sup> & EA. Velasco<sup>1</sup>

acedo@ibgm.uva.es / eavelsam@ibgm.uva.es



<sup>1</sup> Splicing & Cancer, Instituto de Biología y Genética Molecular (UVA-CSIC), Valladolid.  
<sup>2</sup> Cancer Prevention Program of Castilla y León, Instituto de Biología y Genética Molecular (UVA-CSIC), Valladolid.

**Introduction:** At least 50% of all mutations identified in the BRCA genes in breast/ovarian cancer patients are variants of unknown clinical significance (VUS). Several studies have shown the link between disease and splicing disruptions due to mutations, including predicted synonym, missense, nonsense or frameshift variants (Sanz et al., 2010). We aimed to investigate the role of BRCA2 splicing in breast/ovarian cancer. For this purpose, we designed a new splicing reported plasmid (pSAD, patent P201231427) and proceeded to construct a hybrid minigene with exons of the 3' end of BRCA2.

**Methods:** We followed a simple strategy consisting of bioinformatic analysis of DNA variants and splicing functional assays of hybrid minigenes. Splicing reported plasmids allow to perform functional analysis without the need of patient RNA. Because of the lack of efficient splicing vectors, we designed a new one, pSAD, which enabled us the creation of a "maxi-minigene" that include nine exons of the 3' end of BRCA2 gene (19 to 27, MGBR2 19-27, where exon 27 substituted the second (and last) vector exon. The wild type minigene produced a mRNA of the expected size (2174 nt). We validated the new MGBR2 19-27 with four splicing variants of exons 19,20,23 and 24, which we had already been tested in a previous report (Acedo et al., 2012). Other DNA variants reported at different sources and included in cloned BRCA2 exons were analyzed with splicing bioinformatic tools (NNSPLICE, ESEfinder, ESResearch and HSF). Thirty-six of 166 variants were bioinformatically selected and introduced by PCR-mutagenesis in the WT MGBR2 19-27. Mutant and WT Minigenes were transfected into HeLa cells. RT-PCR was performed with specific vector primers and splicing effect of all variants were visualized and quantified by gel and fluorescent capillary electrophoresis in an ABI3130 sequencer. Finally, RT-PCR products were sequenced to characterize the splicing outcome.

Table 1. Bioinformatic analysis of putative splicing variants of exons 19, 20, 21, 22, 23, 24, 25, 26 and 27 of BRCA2.

Mutation (HGVS nomenclature)	Source	Type of mutation	SUMMARY OF BIOINFORMATICS ANALYSIS
c.8486A>T	BIC	Missense	EXON 19 Donor site: 0.95 → 0.87 / [-] conserved SF2/ASF
c.8488-1G>A	Patient	Splice site	EXON 20 [-] Acceptor site
c.8633-24_8634del	BIC	Intronic	EXON 21 [-] Acceptor site
c.8647C>T	BIC	Missense	[+] 3E34, [+1] ESS
c.8723T>G	BIC	Missense	[+] SF2/ASF / [+1] ESS (hrNP-A1)
c.8732C>T	BIC	Missense	[+] SF2/ASF partially conserved / [+1] ESS (hrNP-A1)
c.8748insA	BIC	Frameshift	[+] SRP55 / [+1] ESS
c.8754A>G	BIC	Synonymous	[+] Donor site
c.8754-4A>G	BIC	Intronic	[+] Donor site
c.8754-1G>A	BIC	Intronic	Donor site: 0.98 → 0.43 / [+1] Cryptic donor site 7-nt downstream 0.73
c.8754-5G>T	BIC	Intronic	Donor site: 0.98 → 0.42 / [+1] Cryptic donor site 7-nt downstream 0.82
c.8754-5G>C	Artificial	Intronic	Donor site: 0.98 → 0.42 / [+1] Cryptic donor site 7-nt downstream 0.82
c.8755-10A>A	BIC	Intronic	EXON 22 [-] Acceptor site
c.8773C>T	BIC	Nonsense	[+] 3E5 / [+1] ESS (hrNP-PTB)
c.8789delA	BIC	Frameshift	[+] SRP55 partially conserved
c.8933T>G	BIC	Missense	[+] SRP55 / [+1] ESS
c.8948_8953del	BIC	Splice site	EXON 23 [-] Donor site
c.8953-1G>T	BIC	Intronic	[-] Donor site
c.9117G>A	BIC	Synonymous	EXON 23 [-] Donor site
c.9256-1G>A	BIC	Splice site	EXON 24 [-] Donor site
c.9257-18C>A	BIC	Intronic	EXON 25 [+] SRP40; [-] SRP40; [-] SRP55; [-] SF2/ASF; [-] 11E3E; [+1] 2TSS
c.9257-1G>C	BIC	Intronic	[-] Acceptor site; [-] SRP40; [-] SRP40; [-] SRP55; [+1] 4 ESE; [-] 2E3S
c.9275_9278del	BIC	Frameshift	Cryptic acceptor site 23-nt downstream 0.85 / [+1] 3E5E; [+1] 4E5E; [+1] 2E5S
c.9276T>G	BIC	Nonsense	Cryptic acceptor site 20-nt downstream 0.72 / [+1] SRP40; [+1] ESS
c.9301C>G	BIC	Missense	Cryptic acceptor site 27-nt downstream 0.60 / [+1] 3E23; [+1] 4E5E (hrNP-A1)
c.9454G>A	BIC	Missense	[+] SRP40; [-] SF2/ASF; [-] 3E3E; [-] 1 E5S
c.9464delC	BIC	Frameshift	[+] SRP40; [-] 1 E5S; [-] 1 E5S
c.9501G>A	BIC	Synonymous	Donor site: 0.99 → 0.60 / [-] SRP40; [+1] SRP40; [+1] SRP55; [+1] 4 ESE; [-] 2 E5S
c.9501-3A>T	BIC	Intronic	Donor site: 0.99 → 0.41
c.9501-4A>G	BIC	Intronic	Donor site: 0.99 → 0.78
c.9501-9A>C	Patient	Intronic	[+] E5E; [-] 2E5S; [-] 1 E5S
c.9502-12T>G	BIC	Intronic	EXON 26 Acceptor Site: 0.93 → 0.74
c.9502-2A>C	BIC	Intronic	[-] Acceptor Site
c.9505C>G	LOVD	Missense	[+] 2E5S (hrNP-A1)
c.9610A>G	BIC	Missense	[+] SRP40; [-] SF2/ASF
c.9611C>G	BIC	Missense	[+] SRP40
c.9613G>C	BIC	Missense	[+] 3E3S; [-] SRP40; [-] SRP55
c.9646C>T	BIC	Synonymous	[+] SF2/ASF; [-] 3E3S; [+1] hrNP-B
c.9649-5A>G	UMD	Intronic	EXON 27 c.9649-5A>G
c.9649-2A>G	BIC	Intronic	[-] Acceptor Site
c.9698delT	UMD	Missense	[+] 2E5S

HGVS nomenclature: Human Genome Variation Society, (<http://www.hgvs.org/mutnomen>); Source: BIC, The Breast Cancer Core Database; LOVD, Leiden Open (Source) Variation Database; UMD, Universal Mutation Database.

Bioinformatics analysis: [-], disruption; [+], creation; nt, nucleotide; ESS, Exonic Splicing Enhancer; ESS, Exonic Splicing Silencer; SF2/ASF, SC35, SRP40 and SRP55 are SR proteins that bind ESE/enhancer motifs; hrNP proteins bind ESS/silencer motifs.

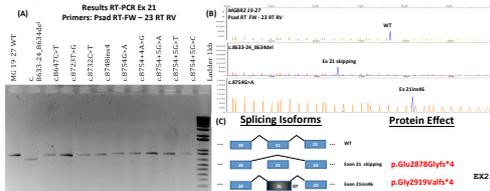


Figure 3. Example of results of variants analyzed in exon 21 of BRCA2 by comprehensive splicing functional analysis with MGBR2 19-27. (A) Agarose gel electrophoresis of RT-PCR products of wt and mutant MGBR2 19-27. (B) Capillary electrophoresis of fluorescent RT-PCR of MGBR2 19-27 wt and c.8633-24\_8634del and c.8754A>G. This figure exhibit three isoforms detected, wt and two aberrant: exon 21 skipping and 46 nt insertion in the position 3' of exon 21. (C) Predicted truncation effect on the BRCA2 protein by aberrant isoforms. (D) Sequence chromatograms of the aberrant splicing outcomes of the exon 21 in MGBR2 19-27.

**Results B.** Thirty-seven of 166 variants annotated from different sources were bioinformatically selected and their splicing effect were evaluated by splicing functional assays of hybrid minigenes. Forty per cent of selected variants presented splicing aberrations. Most variants (86.6%) disrupted the natural splice sites, whereas two altered putative silencers and the polypyrimidine tract and had weak effects. Fluorescent RT-PCR of minigenes accurately detected 14 different RNA isoforms generated by cryptic site usage, exon skipping and intron retention events. Eleven variants showed total splicing disruptions and were predicted to truncate or eliminate essential domains of BRCA2.

**Conclusions.** A relevant proportion of BRCA2 variants are correlated with splicing anomalies. The minigene system with pSAD vector is a straightforward and robust approach to detect variants with an impact on splicing and contributes to a better knowledge of this gene expression step. pSAD is a valuable tool to discriminate between benign and pathogenic DNA variants.

Figure 1. New Plasmid pSAD (PATENT P201231427). The most widespread splicing vector, pSPL3 (Invitrogen), was discontinued and show several disadvantages (e.g. size of more than 6 kb) so that we have developed a new splicing vector that we have called pSAD (for Splicing And Disease). Plasmid pSAD is based on pSPL3 and have the following advantages:

1. Reduction of intron size by 1.2 kb
2. Introduction of a second selection marker (pSPL3 has only resistance to Ampicillin).
3. New polylinker with more than 20 restriction sites (vs. 10 sites of pSPL3).
4. Strengthening of acceptor site of constitutive vector exon 2 by triple mutagenesis to avoid anomalous cross-reactions.
5. Elimination of old restriction sites from pSPL3 to increase versatility of the new multiple cloning site.

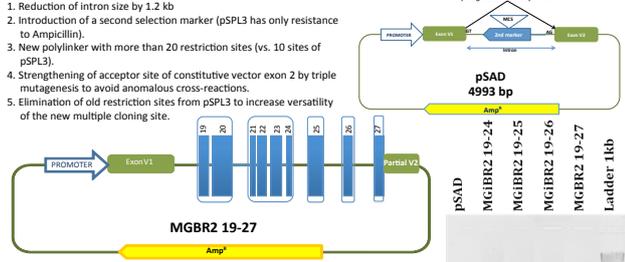


Figure 2. MGBR2 19-27. Nine exons (and part of their flanking introns) of the 3' end of BRCA2 gene were cloned into pSAD. In the final cloning process the V2 exon was substituted by last exon of BRCA2. The minigen resulting is named MGBR2 19-27. All intermediated minigenes and final MGBR2 19-20 were transfected into HeLa cells. Results of RT-PCR of them were visualized in agarose gel electrophoresis and sequenced. The wild type minigene produced a mRNA of the expected size (2,174 nucleotides).

**Results A.** We validated the new MGBR2\_19-27 minigene with four splicing variants (c.8486A>T, c.8488-1G>A, c.9117G>A and c.9256-1G>A) of exons 19, 20, 23 and 24 respectively, which we had already been tested in a previous report of two minigenes with exons 19-20 and 23-24 (Acedo et al., 2012). Improved results were obtained with pSAD without anomalous cross-reactions between splice sites of vector exons, due to the removal of cryptic splice sites and the strengthening of the acceptor site of V2 pSAD exon. Moreover, variant c.8488-1G>A (ivs19) reproduced the splicing outcome of lymphocyte RT-PCR of a carrier patient whereas a pSPL3 minigene 19-20 did not (Acedo et al., 2012). This may be due to an increase of the genomic-context within the minigene where 3' flanking exons were added to exon 20 as we first postulated.

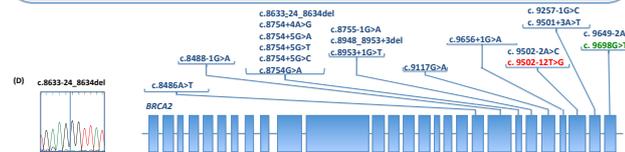


Figure 4. Variants with splicing aberrations analyzed in the MGBR2 19-27. Blue variants disrupted the natural splice sites, red variant altered polypyrimidine tract and the green variant altered exonic splicing silencers.

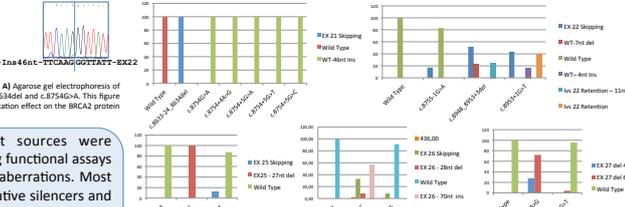


Figure 5. Quantification of RNA isoforms induced by variants of exons 21, 22, 25, 26, and 27 of minigene MGBR2 19-27. Semi-quantitative RT-PCR was performed in triplicate with an exonic FAM-primer. Mean peak areas of three independent experiments were used to calculate percentage of each RNA isoform.

**References:**

- Plasmid pSAD and minigene constructions are protected by the Spanish patent P201231427. "PLASMIDO pSAD PARA ENSAYOS FUNCIONALES DE SPlicing". Autores: E. Velasco, A. Acedo, B. Díez-Gómez. Centro Superior de Investigaciones Científicas (CSIC), Priority Patent Application filed.
- Sanz et al., 2010. A high proportion of DNA variants of BRCA1 and BRCA2 are associated with aberrant splicing in breast/ovarian cancer patients. Clinical Cancer Research, 16 (8) 1957-1967.
- Acedo et al., 2012. Comprehensive splicing functional analysis of DNA variants of the BRCA2 gene by hybrid minigenes. Breast Cancer Res 14: R87.
- Acción de I+D+i. This work was supported by grants P10/2010 of the Instituto de Salud Carlos III (Spanish Ministry of Science and Innovation), and C10K4AL32 (Consejería de Educación) and RD10/10 (Comunidad de Castilla y León) of the Regional Government of Castilla y León. Alberto Acedo is supported by the European Social Fund and Consejería de Educación de la Junta de Castilla y León under the P.O. Castilla y León 2007-2013.



La Comunicación

**Nuevo vector de splicing pSAD. Validación mediante ensayos funcionales de splicing en el "maxi-minigen" con los exones 19 a 27 de BRCA2**

**De Alberto Acedo Becares**

ha obtenido una

**BECA AEGH A LA MEJOR COMUNICACIÓN**

Madrid, 12 de Abril de 2013

Pablo Lapunzina  
Presidente del Comité Organizador  
Hospital Universitario La Paz



Asociación Española  
de Genética Humana  
www.aegh.org

Miguel Ángel Moreno Pelayo  
Presidente del Comité Organizador  
Hospital Universitario Ramón y Cajal

**Comunicación Oral. Congreso de la Sociedad Española de Genética.**

**Título:** Análisis funcional de splicing de variantes de ADN en los genes *BRCA1* y *BRCA2* mediante minigenes híbridos

**Autores:** Alberto Acedo, David J. Sanz, Mercedes Durán, Mar Infante, Lucía Pérez-Cabornero, Beatriz Díez, Cristina Miner y Eladio A. Velasco.

**Resumen:** Alrededor del 15% de los pacientes con cáncer de mama/ovario (BOC) portan variantes de ADN en *BRCA1/BRCA2* de efecto fisiológico desconocido lo cual dificulta el consejo genético. El análisis de los efectos deletéreos de las variantes de ADN en genes de enfermedades hereditarias se centra en la predicción del efecto sobre la función de la proteína, aunque la alteración del splicing está adquiriendo cada vez más relevancia. Mediante las aplicaciones informáticas NNSPLICE, ESEfinder, ESRsearch y HSF nos propusimos analizar mutaciones en los genes *BRCAs* que pudieran alterar elementos reguladores de splicing. En función de los resultados se llevaron a cabo ensayos funcionales de *splicing* mediante RT-PCR de linfocitos y/o experimentos con minigenes. Se seleccionaron 25 de 203 variantes de ADN detectadas en 1007 familias BOC de Castilla y León, 13 de las cuales afectaron al proceso de splicing. Además, seis de ellas fueron responsables del 50% de las familias *BRCA1* positivas. Dado este elevado porcentaje, decidimos investigar todas las variantes de ADN registradas en la base de datos internacional de mutaciones (BIC database) en exones de minigenes previamente construidos (*BRCA1* exones 5-6-7, 13 and 14 and *BRCA2* exones 3, 5-6-7, 18, 19-20 and 23-24). Un total de 74 mutaciones fueron seleccionadas con las herramientas bioinformáticas, generadas por mutagénesis dirigida sobre los minigenes *wild type* y transfectadas en células Hela. 34 variantes (45%) presentaban patrones anómalos de *splicing*.

En conjunto 47/99 de los cambios ensayados alteraban el proceso de *splicing*. De forma individual se vieron alterados: sitios de *splicing* (18), tracto de polipirimidina (2), creación de nuevos sitios de *splicing* (9), eliminación de *enhancer* (2), creación de silenciadores (4), ESE eliminación/ ESS creación (10) y sitios de *splicing*/ESE eliminación (2). Dichas variantes de *splicing* fueron de todo tipo: 4 inserciones /delecciones y 43 sustituciones de nucleótido simple (15 intrónicas, 17 *missense*, 7 *nonsense* y 4 variantes sinónimas).

En conclusión, una importante fracción de las variantes de *BRCA1/2* produce aberraciones en el proceso de *splicing*, representando un mecanismo patogénico relevante en HBOC. Los ensayos funcionales de *splicing* son un instrumento eficaz para discriminar entre polimorfismos benignos y mutaciones patogénicas.

